Target essential surrogate E. coli (TESEC) : a technology platform for tuberculosis drug discovery and the investigation into the mutational landscape of drug targets

Nadine Bongaerts

Université de Paris

École doctorale : **École doctorale Interdisciplinaire Européenne Frontières du Vivant**

Laboratoire : **Institut Cochin, INSERM U1001 Systems Engineering and Evolution Dynamics & Center for Research and Interdisciplinarity**

THÈSE DE DOCTORAT

présentée par : **Nadine BONGAERTS**

soutenue publiquement le : **6 décembre 2019**

Thèse de doctorat de : **Microbiologie**

Target Essential Surrogate *E. coli* (TESEC)

A technology platform for tuberculosis drug discovery and the investigation into the mutational landscape of drug targets

DIRIGÉE PAR

Ariel Lindner — *Directeur de recherche, INSERM, CRI*

RAPPORTEURS

Stéphane Lemaire — *Directeur de recherche, CNRS, Sorbonne Université*

Barbara Di Ventura — *Professeur, Université de Fribourg*

EXAMINATEURS

Anshu Bhardwaj — *Professeur Assistant, CSIR - Institute of Microbial Technology*

Gregory Bokinsky — *Professeur Assistant, Delft University of Technology*

Matthew Todd — *Professeur, University College London*

Résumé: La tuberculose (TB) causée par l'agent pathogène bactérien *Mycobacterium tuberculosis* (*Mtb*) reste la maladie infectieuse la plus meurtrière au monde. Il y a donc un besoin urgent de nouveaux antibiotiques dotés d'une efficacité améliorée et de nouveaux mécanismes d'action. Le but de ce projet de thèse était de mettre au point une méthode pour tester l'efficacité de librairies d'antibiotiques grâce à des bactéries *E. coli* modifiées génétiquement dofnt la croissance dépend de l'expression de l'action de ces antibiotiques sur des gènes cibles de *Mtb*. Cette approche pourrait compléter les méthodes existantes et pourrait également être utilisée comme plateforme communautaire pour la découverte d'antibiotiques. La construction d'une souche d'*E. coli* pour le criblage de médicaments nécessite la suppression d'un gène essentiel natif, suivie de la complémentation avec un gène équivalent de *Mtb* fonctionnel. L'expression de ce gène cible *Mtb* peut être induite par l'arabinose afin de générer une réponse d'expression log-linéaire. Nous avons montré que le contrôle précis de l'enzyme cible de *Mtb* permettait d'accroître la sensibilité du médicament aux inhibiteurs spécifiques de la cible lorsque l'expression était réduite, tandis que la surexpression de la cible a entraîné une augmentation de la résistance au médicament. Dans le cadre de cette thèse, nous avons cherché à valider cette approche en se concentrant sur l'alanine racémase (Alr), une cible classique pour *Mtb*. Pour ce faire, nous avons exposé notre *E. coli* exprimant Mtb Alr à une vaste collection de molécules naturelles en faisant varier les niveaux d'expression de *Mtb* Alr. Nous avons pu ainsi mesurer une différence de sensibilité à certains composés inhibiteurs de la cible en fonction des niveaux d'induction, ce qui nous a permis de découvrir un inhibiteur que les autres approches conventionnelles avaient jusqu'ici manqué : le « bénazépril », un médicament normalement utilisé contre l'hypertension. Pour confirmer ces résultats, nous avons d'abord démontrer la spécificité du bénazépril vis-à-vis de la Mtb Alr par des tests enzymatiques dans lesquels le médicament inhibait l'activité de l'alanine racémase. Puis, nous avons démontré que cette molécule était bien efficace directement sur des souches sauvages de *M. smegmatis* mc^2155 et sur *M. tuberculosis* H37Rv), avec des valeurs de CMI comprises entre 2,25 et 4,5 mM chez ces espèces. Nous avons tenté par la suite de démontrer le potentiel de cette approche pour aussi étudier et mieux comprendre les bases moléculaires de la résistance aux médicaments, notamment dûes à des mutations de leurs cibles. Pour continuer avec notre cible candidate, l'alanine racémase, des recherches antérieures ont suggérées que les mutations de l'alanine racémase sont souvent à l'origine de la résistance à la D-cyclosérine. Pour ce projet, nous avons construit plusieurs librairies de mutants *Mtb* Alr pour explorer de façon exhaustive le paysage des mutations pouvant conférer une résistance à la D-cyclosérine.

Nous avons non seulement validé plusieurs mutations liées à la résistance précédemment prédites par des études d'associations pangénomiques mais aussi prouvé que d'autres mutations précédemment proposées pourraient ne pas être impliquées dans la résistance ou n'affectant pas directement la liaison de la D-cycloserine à sa cible.

Ensemble, ces résultats valident le potentiel d'une nouvelle plate-forme technologique pouvant qui pourrait permettre de découvrir de nouvelles molécules pour lutter contre la tuberculose ainsi que les potentielles mutations qui pourraient conduire à une résistance potentielle. La plate-forme peut être étendue à des cibles de médicaments provenant de nombreuses autres bactéries pathogènes au-delà de *Mtb* et pourrait alimenter un pipeline de nouveaux médicaments candidats nécessaires pour lutter contre la montée de la résistance aux antibiotiques.

Mots clés : Biologie Synthetique, *Mycobacterium tuberculosis*, *Escherichia coli*, Antibiotiques

Abstract: Tuberculosis (TB) caused by the bacterial pathogen *Mycobacterium tuberculosis* (*Mtb*) remains the deadliest infectious disease in the world, killing 1.7 million people in 2017 [1]. There is an urgent need for novel TB antibiotics with improved efficacy and/or new mechanisms of action [1]. The aim of this PhD thesis project was to provide a safe, inexpensive and rapid alternative drug screening method based on engineered *E. coli* cells that depend on the expression of *Mtb* drug targets for growth. This approach would complement existing methods and could also be used as a community toolbox for collaborative antibiotic discovery.

The construction of a *E. coli* drug screening strain involved the deletion of a native essential gene, followed by the complementation with a functional equivalent *Mtb* gene. The *Mtb* target was expressed from an arabinose inducible genetic circuit [2] to generate a log-linear expression response. Fine-tuned control of the *Mtb* enzyme allowed for an increased drug sensitivity to target-specific inhibitors when expression was lowered, whereas target overexpression led to increased drug resistance.

While several *Mtb* drug targets were successfully engineered in *E. coli*, the drug screening efforts as part of this thesis focused primarily on the known TB drug target alanine racemase (Alr). Alanine racemase is inhibited by the second-line TB drug D-cycloserine that was useful as positive control during the development of the drug screening assays. *E. coli* expressing *Mtb* alanine racemase was exposed to compounds of the NIH natural product library and the Prestwick chemical library [3]. By comparing cell growth of *Mtb* Alr-expressing *E. coli* cells in presence of low and high inducer concentrations, a ∼100-fold difference in sensitivity could be achieved for target-specific inhibitors. This allowed for the discovery of inhibitors that conventional approaches could have missed.

As a result, the hypertension drug 'benazepril' was identified as hit compound and further explored for its potential antitubercular properties. Growth inhibition experiments on wild type *M. smegmatis* mc^2155 as well as *M. tuberculosis* H37Rv confirmed benazepril as an anti-mycobacterial compound with MIC values between 2.25-4.5 mM. Importantly, benazepril's specificity to *Mtb* alanine racemase was confirmed in enzyme assays where the drug inhibited alanine racemase activity.

This thesis also explored the use of synthetic *E. coli* strains with *Mtb* target expression to investigate the molecular basis of drug resistance. With cases of extensively drug-resistant TB on the rise, *Mtb* treatments will increasingly depend on the use of second-line antibiotics. Previous research [4, 5, 6] suggests that mutations of alanine racemase, frequently underlie D-cycloserine resistance. For this project, a library based on saturation mutagenesis of the *Mtb alr* gene and a rationally designed *Mtb* Alr mutant library were constructed to exhaustively explore the target mutational landscape surrounding D-cycloserine resistance.

We validated several resistance-linked mutations in *Mtb* Alr that previously predicted on the basis of genome wide association studies with clinical *Mtb* isolates and found evidence that other proposed mutations might not be involved in resistance or do not directly affect D-cycloserine-target binding. Eventually, knowledge of the resistance landscapes for more drug targets would support the monitoring of drug resistance in patients and the improved design of antibacterial drugs.

Together, these results established the foundation for a technology platform that has the potential to discover new TB inhibitors and resistance mutations. The platform can be extended with drug targets from numerous other pathogenic bacteria beyond *Mtb* and can fuel the pipeline with new drug candidates that are needed to fight the antibiotic crisis.

Key words: Synthetic Biology, *Mycobacterium tuberculosis*, *Escherichia coli*, Drug Discovery, Antibiotics

"It is not the strongest of the species that survives, not the most intelligent that survives. It is the one that is the most adaptable to change." — Charles Darwin

Acknowledgement

First of all, I would like to thank my thesis director Ariel Lindner. Thank you for your guidance and for allowing me to grow and learn. I also appreciated the freedom you gave me to explore my own scientific questions and interests. Your scientific ideas are inspirational, and they pushed me to the next level. Four years went fast, and it feels like yesterday that I arrived in Paris to meet with you and Jake. The interdisciplinary character of the CRI reflected me as a researcher, and it allowed me to be myself. I would also like to thank Jake Wintermute. Thank you for your supervision, advice, and our collaboration to establish the TESEC drug screening platform. Before joining the Ph.D. program, finding tuberculosis drugs with *E. coli* sounded like a crazy idea. But after spending four years to make it real, I believe that the future of TESEC is bright. In addition, I would like to thank you for your DJ-ing skills in the lab, the epic Synbio Mooc 1 hype intro, insights into American politics, and for introducing me to the finest French boxed wine brands.

I want to thank my Thesis Advisory Committee, Roland Brosch (Institut Pasteur), Jérôme Bonnet (CNRS), Rosalind Allen (University of Edinburgh) for their valuable guidance throughout my Ph.D. Your feedback helped me to stay on track, and after every TAC meeting, I felt even more motivated about my research.

Zainab Edoo and Michel Arthur (Centre de Recherche des Cordeliers), thank you for the valuable collaboration we had. I very much enjoyed our conversations, and your tuberculosis experiments were indispensable to proof the antimycobacterial effect of benazepril.

Martin Carballo Pacheco and Rosalind Allen (University of Edinburgh), thank you for the fruitful Skype calls and the nice collabora-

tion to combine my biological experiments with molecular dynamics simulations of Martin. I am curious about our future findings.

Then I would like to thank my master interns Sebastián Sosa Carrillo, Salam Abbara, Hanifa Bouziri, and Ayan Abukar. Sebastián, you were my very first intern, and the use of the *tolC* mutation we explored was essential in the end. Salam, you inspired and motivated me for my research by sharing your experiences as a medical doctor in infectious diseases. I was also impressed by how fast you developed lab skills. We had many productive and fun interactions, and I am incredibly thankful for the strong friendship we have as a result of your internship. Hanifa, thank you for all your hard work, for challenging me to explain science in French. Dear Ayan, after we met at iGEM and the GAP Summit, I convinced you to come to France for an internship, you stayed for a CRI master, and now you will follow my footsteps for your Ph.D. You impressed me with your quick thinking and scientific ideas, and I am sure you will do an excellent job for your Ph.D. I would also like to thank you for proof-reading parts of this Ph.D.

To my dear colleagues, Ivan, Chantal, Magia, Sébastien, Wei Lin, and Marie-Florence and Arnaud, thank you for the nice years we spent together in Cochin, for all your support, scientific feedback, kindness, and interesting conversations. Special thanks to Magia for the delicious Moroccan food. Arnaud, thank you for your valuable feedback, for always challenging my ideas and for putting me in touch with the tuberculosis lab in Cordeliers. I miss you in the office. Béatrice, without your help, I would have been completely lost in dealing with the French administration. Thanks to you, my integration in France was a lot easier. Dear Alvaro, thank you for always being there when I needed it. You are both a friend and mentor to me. I enjoyed the always stimulating conversations about the evolution of bacterial sex and all the big life questions, and I hope that our coincidental scientific collaboration will pay off soon. Hoatian, thank you for the nice office/lab time we spent and I apologize for eating all your chocolate..., I owe you some. Paulina, you did a fantastic job of getting the new lab up and running so quickly. Thanks for helping me to get the

weirdest drug compounds from companies around the world. Your cookies were also beneficial in maintaining my glucose levels during writing. Xiaohu and Yifan, thank you for helping me out with fixing my code when I was learning how to use Matlab. Anshu, you are an inspiration and a walking Wikipedia about mycobacteria. I very much appreciated our discussions. Radek, I am sure that our brainstorms will one day save the world. Rania, thank you for helping me to get started with LaTeX, coffee cups, and vitamin pills when I needed it. Dule, thank you for your support in the CRI, scientific feedback and the many enjoyable conversations. Gert-Jan, Remy, it's great to have some Dutch-speaking buddies in the lab. GJ, thanks for introducing and organizing the weekly CRI Thursday drinks. Ana, thank you for always sending me great articles and for the interesting philosophical conversations we had. I wish you all the best for your new position at Harvard. Aya, I appreciate that are always willing to help out and I am very glad to have you as a colleague. Hamid, thank you for your French lessons, for your support in the lab and for continuing the experiments while I was writing this thesis. Eugenia, thank you for supporting the final stages of my thesis.

I would also like to thank my colleagues at Hello Tomorrow for giving me a great start when I arrived in Paris for my Ph.D. Thank you for your friendship, collaboration, patience, flexibility and the fantastic work you did to grow Hello Tomorrow.
Dear Eva, my science communication partner in crime, I am thankful for your support as a friend and for the many years we worked together. I can always count on you and the many communication experiences we shared have also helped me in this Ph.D.

My dear Dutch-French family, thank you for your unconditional support and for always stimulating me to follow my curiosity and my passion. I would like to especially thank my parents for their help, for listening and for giving me the education I needed to reach my potential. Mama, I think your early biology lessons paid off. Papa, I guess your French genes brought me to Paris. Michiel, it's great to have a nerd brother who understands what I am doing. For my Ph.D., I stuck to Matlab, but I hope you can teach me Python later.

Catherine, Alain, Hugo, Alex, Marie, Christine & Denis, thank you for supporting me and making me feel very welcome in France. Being part of the family helped me to make this country my second home.
Last, but not least, I would like to thank my husband. Xavier, I would probably not have written this thesis if we had not met. The beginning of this thesis hallmarked the beginning of our life together in Paris. Thank you for always believing in me, your tremendous support and valuable advice. You are the best.

Contents

List of Figures

List of Tables

Abbreviations

aa	amino acid
ACE	angiotensin-converting-enzyme
AI	Antagonism Index
AMR	Antimicrobial resistance
APS	adenosine 5-phosphosulphate
BCG	bacille Calmette-Guerin
BSL	Biosafety Level
D-ala	D-alanine
DAP	*meso*-diaminopimelic acid
DCS	D-cycloserine
DMSO	Dimethyl sulphoxide
dNTP	Deoxynucleotide
DR-TB	drug-resistant tuberculosis
DST	Drug Susceptibility Testing
E. coli	*Escherichia coli*
epPCR	error-prone PCR
FRT	FLP recombinase recognition target
GlcNAc	N-acetylglucosamine
HCP	High-copy plasmid
HTS	High Throughput Screening
IC_{50}	Inhibitory Concentration 50%

IPTG	isopropyl-β-D-thiogalactopyranoside
L-ala	L-alanine
LCP	Low-copy plasmid
mAGP	mycolyl arabinogalactan–peptidoglycan
ManLAM	mannose-capped lipoarabinomannan
MD	molecular dynamics
MDR-TB	Multidrug-Resistant Tuberuclosis
meGFP	Mono Enhanced Green Fluorescent Protein
MIC	Minimal Inhibitory Concentration
MOA	Mechanism Of Action
MRSA	Systemic Methicillin Resistant *S. aureus*
Msmeg	Mycobacterium smegmatis
Mtb	Mycobacterium tuberculosis
MurNAc	N-acetylmuramic acid
NGS	Next Generation Sequencing
nt	nucleotide
NTA	nickel-charged nitrilotriacetic acid
OD	Optical Density
PAMPs	pathogen-associated molecular patterns
PAPS	3-phosphoadenosine 5-phosphosulfate
PBP	Penecillin Binding Proteins
PBS	Phosphate-Buffered Saline
PCR	Polymerase Chain Reaction
PG	Peptidoglycan
PRRs	pattern recognition receptors
REA	Response Envelope Analysis
RR-TB	Rifampicin resistant tuberculosis

SI	Synergy Index
SSMD	strictly standardized mean difference
T_m	Melting Temperature
TB	Tuberculosis
TESEC	Target Essential Surrogate *E. coli* strain
U.S.	United States
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
XDR-TB	Extensively drug-resistant tuberculosis
βME	β-mercaptoethanol

Introduction

0.1 The antibiotic crisis

The biological relationship between humans and bacteria is ambiguous. Our survival is barely possible without them [9], and yet, their presence has equally threatened our existence for thousands of years [10, 11]. With respect to human history, we live in a prosperous period in which we have acquired technologies to diagnose, treat, and prevent bacteria from endangering our health [12]. Among these, life-saving revolutionary steps were improved sanitation [13] combined with the medical use of sulfa drugs in the 1930s [14], the discovery of penicillin by Alexander Flemming in 1928 [15] and its large-scale production in the 1940s, closely followed by the Waksman platform [16] to find antibiotic-producing microbes. However, the life-saving ability of these innovations is strongly dependent on their responsible use. As early as 1945, Alexander Flemming warned about the risks of antibiotics during his Nobel lecture [17]: "The time may come when penicillin can be bought by anyone in the shops. Then there is the danger that the ignorant man may easily underdose himself and by exposing his microbes to non-lethal quantities of the drug make them resistant" (p. 93).

Today, the consequences of Flemming's early warnings are evident. The effectiveness of antibacterials is no longer guaranteed [18]. Decades

of misuse and overuse of antibiotics in agriculture [19] and in patients without bacterial infections [20, 21, 22] accelerated the selection and survival of bacteria that spontaneously acquired mutations or genetic elements that render drug-resistant phenotypes [23] (Figure 1). As a result, antimicrobial resistance (AMR) forms a global threat to our health and economies [24, 25].

Current efforts to drastically reduce AMR rates in the future are insufficient [25]. For example, the number of novel antibiotics that reach the market is declining [24, 25], and although the importance of handwashing is generally known, proper handwashing by medical personnel is still not standard practice [26]. The Review on Antimicrobial Resistance reported that without change of the current state of affairs, an alarming 10 million people could die from the consequences of antimicrobial resistance in 2050 [25].

0.2 Tuberculosis

Central to the issue of AMR is one of the deadliest diseases in the world: tuberculosis (TB). The disease has accumulated human victims throughout history and was an epidemic in Europe and North America in the 18th and 19th centuries [28]. The cause of the devastating disease was unknown, until the German physician and microbiologist Robert Koch, presented his paper 'Die Aetiologie der Tuberkulose' in 1882 [29]. Koch changed history by demonstrating that the causal agent of tuberculosis was the tubercle bacillus [29], also known as *Mycobacterium tuberculosis*. Although today, antibiotics are available to treat tuberculosis, the disease is still annually killing 1.5 million people [30]. And while tuberculosis is prevalent worldwide, 95% of

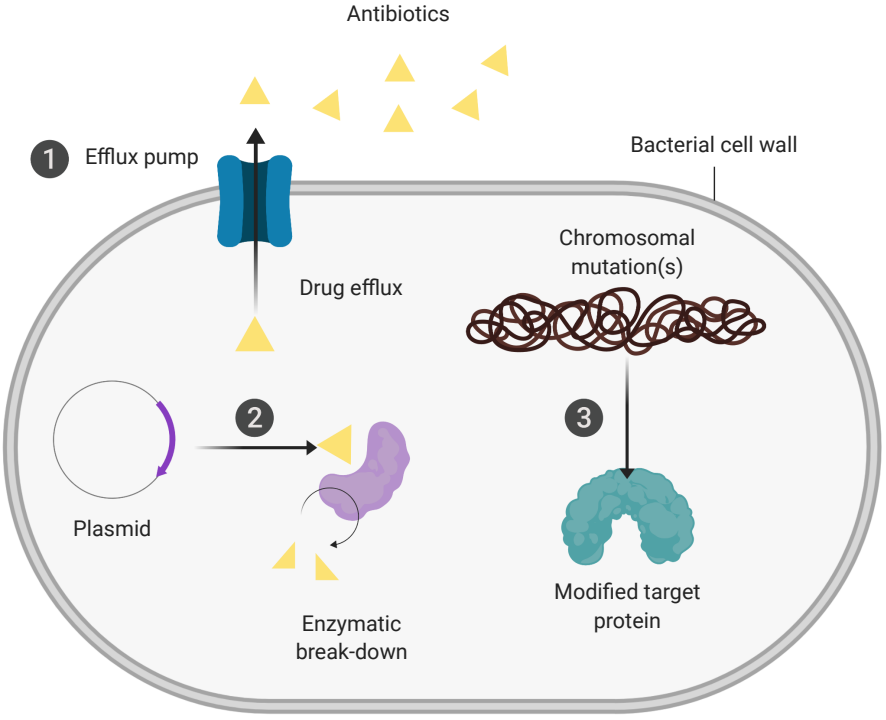

Figure 1: Antimicrobial resistance mechanisms
Bacteria have various ways to survive in the presence of antibiotics. 1. Drug molecules can be exported by the cell through the action of efflux pumps. 2. Plasmids carrying genes for antibiotic degrading enzymes can be passed between bacteria via horizontal gene transfer. 3. Mutations in the genes of drug target proteins can lead to structural modifications that prevent drug binding or inhibition. Figure based on figure 2 from Aslam *et al.* [27].

tuberculosis-caused deaths come from developing countries [30].

0.2.1 *Mycobacterium tuberculosis*

A tuberculosis infection starts with the inhalation of aerosol droplets containing *Mycobacterium tuberculosis* (*Mtb*) [41] (Figure 3). *Mtb* is a rod-shaped bacterium [42] and classified as acid-fast in reference to the acid-fast Ziehl-Neelsen stain that is required to detect the bacteria [43]. The mycobacterial cell wall displays exceptional features compared to gram-positive or gram-negative bacteria [32]. The core of the structure is composed of covalently connected layers of peptidogly-

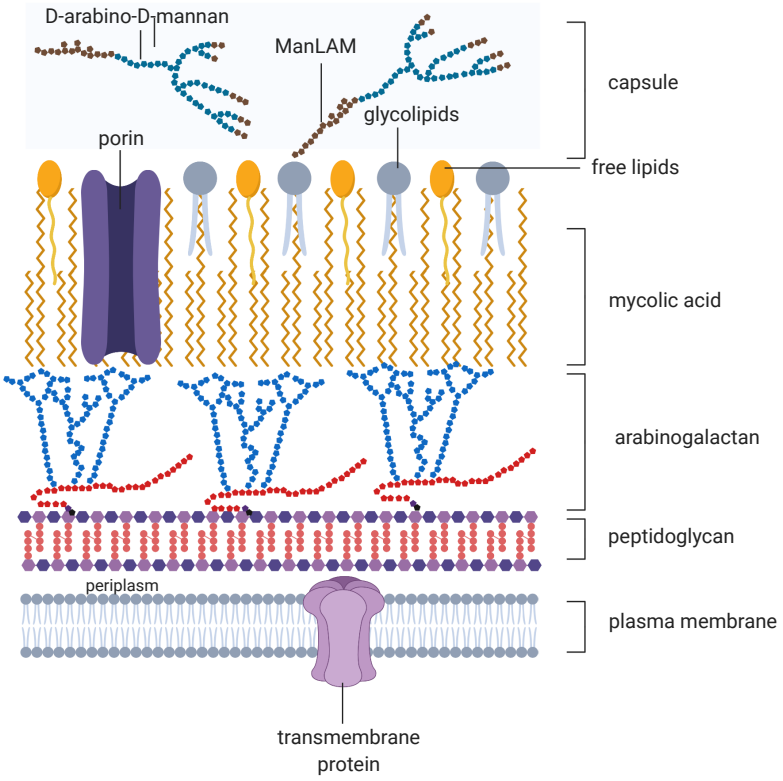

Figure 2: Schematic representation of the general composition of the mycobacterial cell envelope.

The *Mtb* cell envelope is a layered structure composed of (from top to bottom): 1. an outermembrane 'capsule', 2. the mycolyl arabinogalactan–peptidoglycan (mAGP) complex, 3. the plasma membrane connected to the cytosol of the cell. Image is not drawn to scale. Created by combining adaptations of figure 1 from [31], figure 1 from [32], and figure 1 from [33].

can, arabinoglycalactan and long-chain (C_{60}-C_{90}) mycolic acids [31], the so-called mycolyl arabinogalactan–peptidoglycan (mAGP) complex [44] (Figure 2). A variety of (glyco)lipids (e.g. trehalose mono- and dimycolates, sulfoglycolipids, phosphatidylinositol mannosides), and lipoglycans such as lipomannan, lipoarabinomannan, mannose-capped lipoarabinomannan (ManLAM) are embedded in the mycolic acid layer [31] (Figure 2). The mAGP complex is covered by a 'capsule' composed of proteins, polysaccharides, and a small fraction of lipids [31] that form the outer layer of the cell envelope.

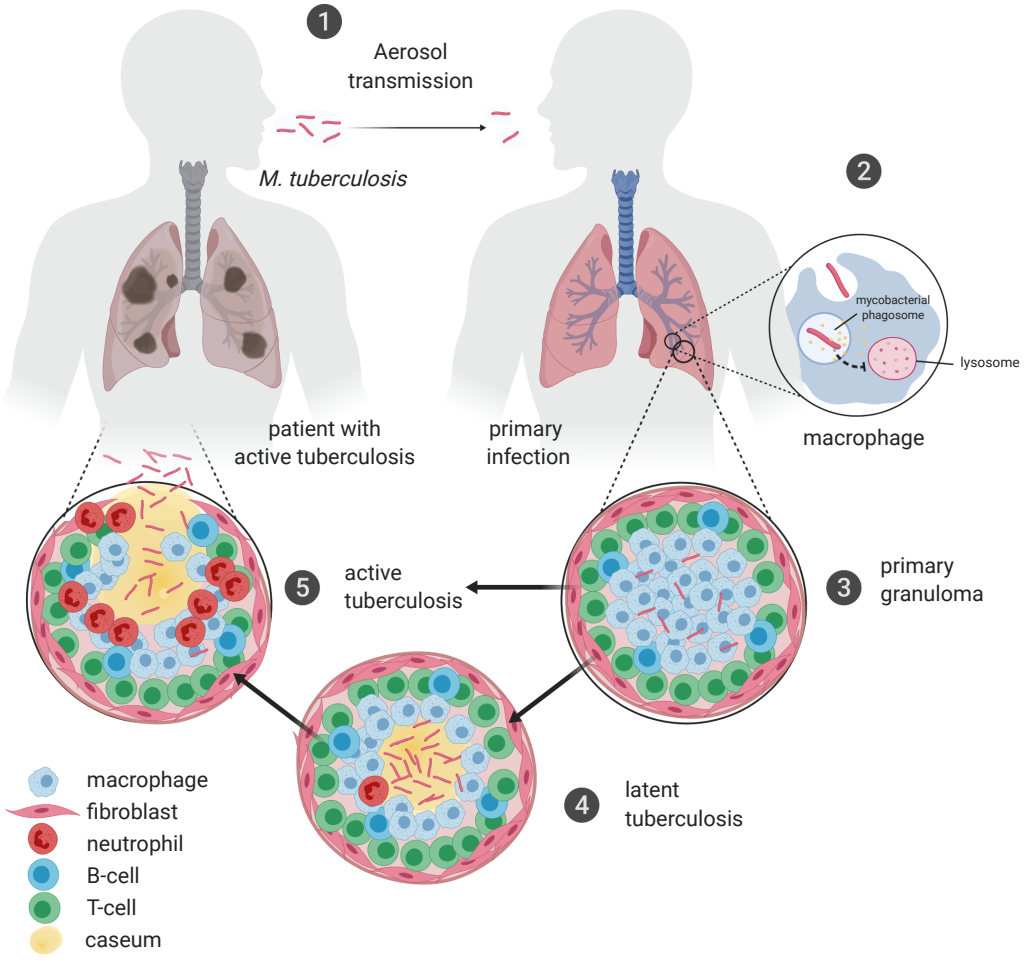

Figure 3: Tuberculosis infection.
1. Tuberculosis can be transmitted when a person with an active pulmonary tuberculosis sneezes or coughs. The resulting aerosol droplets contaminated with *Mtb* can travel to the lungs of a healthy individual where *Mtb* is phagocytosed by alveolar macrophages [34]. 2. *Mtb* survives within macrophages by preventing phagosome acidification and fushion with the phagosome [35, 36, 37, 38]. 3. A granuloma is formed in response to the tuberculosis infection. 4. In 90-95% of the individuals, mycobacterial growth is suppressed and the bacteria enter a latent state. However, 5-10% of the patients develop the active disease when the granuloma is unable to sustain the primary granuloma enabling bacteria and infected macrophages to spread throughout the lungs or other parts of the body [39, 40]. Created by combining adaptations of figure 1 from [40] and figure 1 from [36].

Beyond the importance for *Mtb* physiology [42], the architecture of the mycobacterial cell envelope is essential for its natural resistance towards a wide range of therapeutic agents [45, 46] and the modulation of the host immune response [47, 48, 49].

When *Mtb* bacilli arrive in the lungs, they encounter alveolar macrophages, neutrophils and dendritic cells [50]. These innate immune cells form the first line of defense by recognizing pathogen-associated molecular patterns (PAMPs) through interactions with pattern recognition receptors (PRRs) [51]. The complex interplay between pathogen-derived molecules and immune cell PRRs (e.g., Toll-like or mannose receptors) modulates the host immune response and determines the fate of mycobacteria [52, 51, 31]. Various *Mtb*-ligand-PRR interactions facilitate phagocytosis [38]. One way for *Mtb* to enter macrophages is through the binding of mannose caps on ManLAM (Figure 2) to the mannose receptor of the alveolar macrophages [53, 54, 51].

Within the macrophage, *Mtb*'s viability is dependent on its ability to survive in the phagosome. Although the intracellular space of macrophages is detrimental for most bacteria, *Mtb* has evolved several strategies to survive inside the vacuole [38]. Macrophages degrade bacteria through a multi-stage process in which the 'early phagesome' is acidified through the recruitment of V-ATPase that pumps cytosolic H^+ into the vesicle [55]. Matured phagosomes finally fuse with lysosomes that contain a variety of lytic enzymes and reactive oxygen species that degrade endocytosed bacteria [55]. However, *Mtb* is able to stabilize the phagosomal pH between 6.2-6.5 by preventing V-ATPase docking [35] as well as promoting V-ATPase degradation [37]. *Mtb* is also equipped to escape degradation by lytic enzymes. An important factor is *Mtb* protein kinase G (PknG) that blocks phagosome-lysosome fusion [38], as mycobacteria with inactive PknG are effectively eliminated [56, 38].

The proliferation of *Mtb* within macrophages triggers the release of cytokines that attract other macrophages, monocytes, neutrophils, T and B lymphocytes to the site of infection [40]. This cascade results in the formation of a structured cluster of immune cells termed granuloma [40]. Within the granuloma, *Mtb* pursues an attenuated lifestyle surrounded by a low pH, low oxygen, limited iron, and nutrients conditions [57]. Depending on the host immune response, granulomas can either contain or further spread the infection [58, 40]. When the immune response leads to uncontrolled tissue damage, the infection can create new granulomas and even spread to organs outside the lungs to create extrapulmonary tuberculosis [40].

0.2.2 Disease and treatment

The adaptive immune system plays a determining role in the progression of tuberculosis [50]. Although tuberculosis resides in approximately two billion people globally, only 10% of tuberculosis carriers develop disease symptoms when their immune system is unable to control the infection [59, 50]. When *Mtb* causes disease, bacteria actively replicate inside the patient to spread in the body and cause tissue damage [60]. In the majority of patients, the host immune system represses the tuberculosis infection resulting in latent tuberculosis that lacks disease symptoms and is not transmissible [60]. Unfortunately, patients with deficient immune systems, e.g., in people co-infected with HIV, have an increased risk of developing active tuberculosis. HIV-infected patients are twenty to thirty times more likely to progress to active TB disease [60].

At present, no effective TB vaccine is available. Although the bacille Calmette-Guérin (BCG) vaccine is used since the 1920s [61], its pro-

tection for adults is highly variable and only prevents children from developing the most serious forms of tuberculosis [1]. Therefore, tuberculosis therapies based on antibiotics are currently the best option to fight the disease. The standard treatment of active tuberculosis concerns a 6-month combination therapy that requires the administration of four types of antibiotics: rifampicin, isoniazid, pyrazinamide, and ethambutol. All four antibiotics in the first two months, and isoniazid and rifampicin in the 4 last months [62]. There are different treatment options for latent tuberculosis. In general, latent tuberculosis therapy consists of a 3 to 9-month administration of only rifampicin or isoniazid, or a combination of rifapentine and isoniazid, or isoniazid with rifampicin [63]. Current therapies are sub-optimal for several reasons. First of all, except for pyrazinamide, the aforementioned drugs are not effective against slow-growing or non-replicating cells [64]. Second, the cornerstone TB drug rifampicin is incompatible with HIV/AIDS medication because it induces the enzyme cytochrome P450. Cytochrome P450 causes antiretroviral drugs to be metabolized too quickly, thereby reducing its efficacy [65]. In addition, the long duration, costs, and adverse side effects related to the current therapies complicate adherence [66]. And finally, drug resistance against the most effective first-line tuberculosis drugs is internationally spread [1].

0.2.3 Drug-resistant tuberculosis

In 2017, an estimated 10 million people acquired tuberculosis, including more than half a million patients with drug-resistant tuberculosis (DR-TB) [1]. *Mtb* has both intrinsic and acquired drug resistance mechanisms [67]. Intrinsic resistance mechanisms are related to the mycobacterial cell wall with low permeability [45], the expression of a wide variety of drug efflux pumps [68], including pumps that are

only induced within the macrophage [69], and the natural expression of β-lactamase BlaC [70]. Moreover, *Mtb* displays an increased drug tolerance when it is starved of nutrients and exists in a non-replicating state [71].
Acquired drug resistance, on the other hand, occurs predominately through chromosomal point mutations in specific gene targets [72]. Interestingly, drug resistance acquired via horizontal gene transfer of mobile genetic elements has not been found in *Mtb* [73, 74].

The classification of drug resistance depends on the pathogen's susceptibility to the most important TB drugs. Rifampicin resistant tuberculosis (RR-TB) is not susceptible for the most effective first-line TB drug, multidrug-resistant tuberculosis (MDR-TB) is defined as mycobacteria without susceptibility for the two critical first-line TB drugs (isoniazid and rifampicin) and extensively drug resistance tuberculosis (XDR-TB) is recognized as MDR-TB with resistance to at least one drug in the fluoroquinolones class and one or more drugs in the second-line injectables class [1] (Figure 4).

Patients with DR-TB often receive ineffective treatments because DR-TB remains poorly diagnosed [1]. Since drug-resistant tuberculosis (DR-TB) is treatable with second-line antibiotics, rapid, inexpensive, and easy to perform drug-susceptibility assays are essential to determine an appropriate treatment for individual patients and to monitor the occurrence of drug-resistance [76, 77].
Generally, there are two types of drug susceptibility testing (DST): phenotypic or genetic. Phenotypic in vitro assays screen for mycobacterial growth inhibition in the presence of various drug concentrations and is still the most reliable approach [78]. At the same time, there

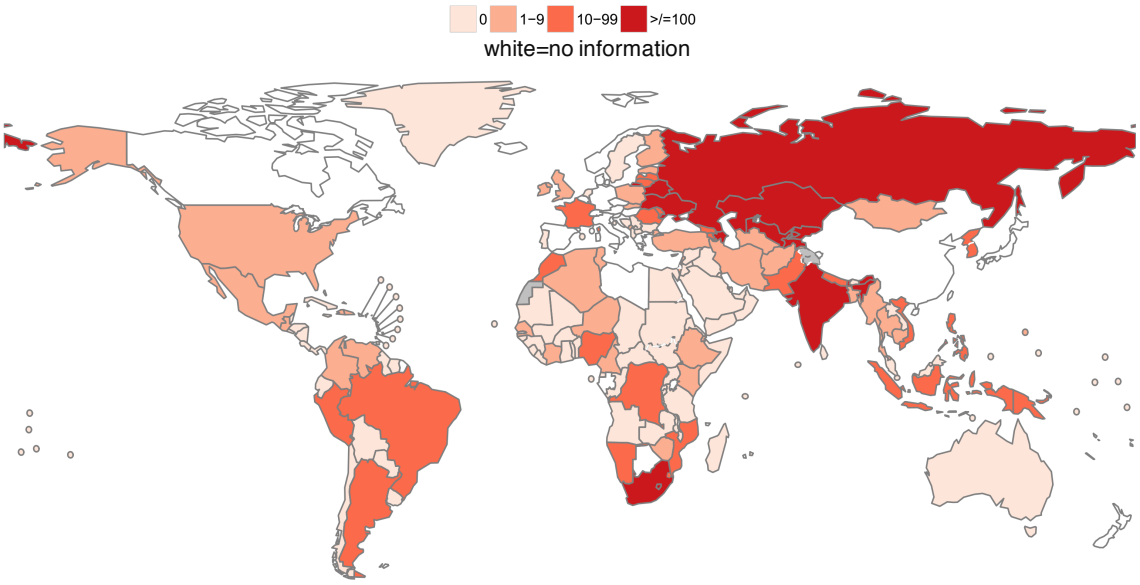

Figure 4: Number of patients with laboratory-confirmed XDR-TB started on treatment in 2017

Reprinted with permission from: Global MDR-TB situation: update 2018, WHO, slide 23, copyright WHO (2018) [75].

are also issues concerning the reproducibility of phenotypic methods as variations in protocols result in different outcomes [77]. As a consequence, there is no universally accepted standard method to determine the minimal inhibitory concentration (MIC) for *Mtb* [78]. In addition, these phenotypic assays demand resources, time, and biosafety precautions related to the cultivation of clinically isolated mycobacteria. In low-income countries with limited lab facilities, such assays are particularly difficult to conduct on a routine basis [79].

Alternatively, genetic tests that screen for known resistance genotypes allow for rapid and inexpensive detection of resistant strains in patients [79, 77]. Thanks to progress in Next Generation Sequencing and research into the molecular basis of resistance, numerous genetic tests (e.g., GeneXpert® MTB/RIF) [80, 81] have been developed to identify tuberculosis resistance mutations. While those tests have helped to improve the detection of resistance to first-line TB drugs,

the available number of genetic tests to verify the resistance to second-line TB drugs is still limited [82].

Thus, broadening the availability of genetic tests for additional antibiotics is important for the appropriate diagnosis and treatment of tuberculosis patients, especially for low-income countries.

0.2.4 Ending tuberculosis?

The economic benefits of investing in solutions to prevent or treat resistant tuberculosis outweigh the future costs related to treating patients with MDR- or XDR-TB [83]. The average costs per American patient increase from $47,000 for non-resistant TB infections to $301,000 for MDR-TB and $701,000 for patients with XDR-TB [84].

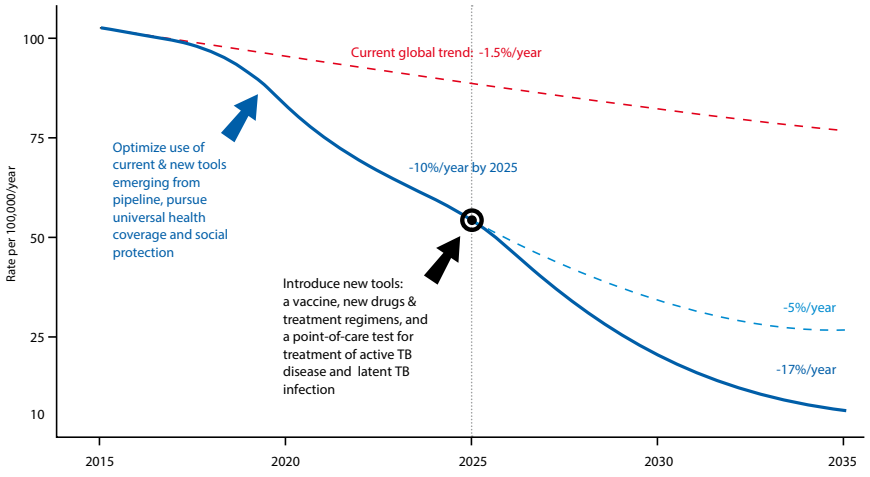

Figure 5: Projected acceleration in the decline of global tuberculosis incidence rates to target levels.
Reprinted with permission from The End TB Strategy report, WHO, Chapter: Strategy, vision, goal, milestones and targets, page 5, copyright WHO (2014) [85].

The World Health Organisation (WHO) outlined a strategy to reduce the number of deaths caused by tuberculosis 95% by 2035 compared to 2015 [86]. Key pillars are the development of vaccines, drugs, and improved treatment strategies for both active and latent tuber-

culosis. Such innovations could reduce the current global decline rate from 1.5%/year to 10%/year by 2025 and 17%/year by 2035 [86] (Figure 5). In line with this ambition, impact-driven private organizations and governments, recently pledged US$14B to the Global Fund to significantly reduce the burden of tuberculosis, but also AIDS and malaria by 2030 [87]. Such international efforts are crucial to change the tide, especially when, despite the health demands, commercial incentives are lacking [83]. Other initiatives such as the TB Alliance, the TuBerculosis Vaccine Initiative, and the Stop TB Partnership form multi-stakeholder programs involving government agencies, academia, pharmaceutical companies, and philanthropic institutions for the development of new tuberculosis drugs, vaccines or diagnostics.

As a result of these partnerships, three new tuberculosis drugs have been recently approved for the first time in more than 40 years: bedaquiline (Johnson & Johnson) [88], delamanid (Otsuka Pharmaceutical Co Ltd.) [89] and pretomanid (TB Alliance) [90]. Interestingly, pretomanid is the first tuberculosis drug that was developed by a non-profit organization and will be part of a six-month regimen with bedaquiline and high-dose linezolid (BPaL) to treat XDR-TB or non-responsive MDR pulmonary TB with a cure-rate of ∼90% [91]. This therapy is an important breakthrough because it improves the current two-year during XDR-TB therapy (without bedaquiline) that only cures ∼20% of XDR-TB patients [91].

However, Médecins Sans Frontières commented that the price of BPaL (∼$1000 per therapy) is currently two times too high to become widely accessible for those in need and that price reductions are crucial [92, 93, 94].

Moreover, bedaquiline resistance due to mutations in amongst others

its drug target ATP synthase have already been reported [95].
Therefore, drug development efforts are still needed to improve and lower the price of current tuberculosis therapies, as well as to ensure to the availability of effective tuberculosis drugs in the future.

To address this, the open source pharma approach represents an innovative alternative strategy to drastically lower the costs of pharmaceuticals [96]. In this model, an open community approach enables any researcher or institution to openly share experimental data, compounds, software, methods, and technologies to collectively form a pipeline for drug development [96].

This thesis aims to contribute to the open source pharma efforts by developing an alternative strategy to accelerate drug discovery against tuberculosis. More specifically, we want to establish a modular technology platform for hit compound identification based on non-pathogenic, easy-to-handle *E. coli* bacteria that replace the use of slow-growing, hazardous and hard-to-engineer *M. tuberculosis*.

Therefore, the central question that this thesis addresses is:

Can engineered E. coli strains with Mtb drug targets serve as an inexpensive and scalable platform for the development of tuberculosis antibiotics?

Moreover, insights into the resistance landscape of drug targets could aid the development of more effective drugs and the risk assessment of clinical *M. tuberculosis* isolates with unknown resistance phenotypes.

Hence, a secondary question that is also addressed is:

Could engineered E. coli strains provide new insights in drug resistance landscapes of Mtb drug targets by enabling us to predict resistance mutations?

0.3 Target-Essential Surrogate *E. coli* strains

TESEC stands for Target-Essential Surrogate *Escherichia coli* and describes engineered *E. coli* strains in which we replace an essential *E. coli* enzyme with an orthologous *Mtb* enzyme that catalyzes an identical biochemical reaction. As a result, TESEC strains are growth-dependent on the functional expression of the heterologous *Mtb* protein. This concept was developed by the iGEM Paris Bettencourt team [97] in 2013 and has been the starting point of the work described here. Our ambition is to use TESEC strains as a rapid, inexpensive, and scalable open science platform for the drug discovery of *Mtb* target-specific inhibitors. In theory, when a molecule inhibits the target activity of our TESEC strains, the same compound could be expected to inhibit the same biochemical reaction in *M. tuberculosis*.
TESEC-based drug discovery could offer several advantages over conventional approaches that involve whole-cell experiments with *M. tuberculosis* or target-based enzyme assays (Table 1). First of all, experiments with *M. tuberculosis* require a lab with biosafety level 2 or 3 [98, 99] and typically takes several weeks as the *Mtb* has a doubling time between 18 and 54 hours [100].

The genetic toolbox for *Mtb* is expanding [101] and supports the generation of *Mtb* deletion strains or strains with alternative protein

Assay Type:	**Whole cell-based (*Mtb*)**	**Enzyme-based**	**TESEC-based**
Assay Time	Weeks	Minutes	Hours
Biosafety level	2-3	1	1
Live-cell-context	Real pathogen	None	*E. coli*
Information on the mechanism of action	None	Yes	Partially
Technical difficulty	Difficult to genetically engineer	Enzyme isolation and assay development can be complex and time-consuming	Easy to genetically engineer
Relative costs	Moderate	Expensive	Low
Relative sensitivity	Fixed	High	High

Table 1: TESEC-based assays compared to conventional drug screening strategies such as whole cell exposure of compounds to *Mtb* strains or the use of biochemical assays.

.

expression that are useful for drug target identification and validation [102]. Yet, its slow-growth and thick mycobacterial cell wall combined with the limited number of antibiotic resistance markers [103, 104] make the genetic modification of *Mtb* [101] challenging compared to model bacteria such as *E. coli*.

Biochemical approaches that use isolated *Mtb* enzymes allow for the direct identification of target inhibitors. However, decades of efforts using this targeted approach failed to generate a new tuberculosis drug as they lacked whole-cell activity [102]. In addition, reagents and enzymes for biochemistry experiments are expensive compared to cultivation experiments with low-cost media. Therefore, the use of an *E. coli*-host for the early-stage TB drug discovery process offers profound benefits in time, costs, staff, and training. Although TESEC-based discovery does not guarantee effectiveness on whole-cell *Mtb* due to many organism-specific features, we do provide a live-cell context that may increase the biological effectiveness compared to a more artificial

biochemical environment.

0.4 Thesis structure

The **first chapter** describes how I characterized the general TESEC strategy for tuberculosis drug discovery, including a wide variety of TESEC strains I engineered with *Mtb* drug target expression and methods to control the growth dynamics.

In the **second chapter**, I demonstrate the establishment of a TESEC-based drug screening effort that allowed for the discovery of benazepril as *M. tuberculosis* inhibitor.

In the **third chapter**, I show another example of how engineered *E. coli* with *Mtb*-drug target expression can be used to validate and predict tuberculosis resistance genotypes.

Finally, I conclude this thesis with my reflections and perspectives on the impact of the work and provide recommendations for the future of the technology platform.

Chapter 1

Developing Target Essential Surrogate *E. coli* (TESEC) strains as Drug Discovery Tool

1.1 Abstract

The development of the TESEC platform for tuberculosis drug discovery included a series of design, decision making, and testing steps that are described in this chapter. First, we explored whether relevant *Mtb* drug target enzymes alanine racemase (Alr), 5-adenosinephosphosulphate reductase (CysH), anthranilate phosphoribosyltransferase (TrpD), anthranilate synthase component I (TrpE), aspartate-semialdehyde dehydrogenase (Asd), 4-hydroxy-tetrahydrodipicolinate synthase (DapA) and dihydrodipicolinate reductase (DapB) were able to rescue the growth of auxotropic *E. coli* cells in which the corresponding native gene product was absent. All implemented *Mtb* enzymes, with the exception of *Mtb* TrpE, could successfully replace the activity of the native *E. coli* equivalent. Moreover, we investigated the bacterial growth dynamics when the expression of *E. coli* or *Mtb* essential proteins were

varied through genetics and/or cultivation conditions.

To facilitate the optimization and future expansion of the TESEC toolbox, a golden gate library based on the previously published CIDAR MoClo parts library was constructed. This TESEC MoClo library contained all essential parts for TESEC strain construction: a new set of destination vectors with chloramphenicol resistance markers; additional essential pre- and suffixes to facilitate the construction of fusion enzymes; essential *E. coli* and *Mtb* genes; different promoters and an RBS library. Finally, we focussed on efforts using a TESEC strain with *Mtb* alanine racemase (Alr) expression, which led to the establishment of a reliable drug screening strategy that is further explored in Chapter 2.

1.2 Introduction

1.2.1 *Mtb* target requirements for use in TESEC strains

TESEC relies on homologous enzymes that are essential in both *E. coli* and *Mtb*. In 2013, the iGEM Bettencourt team [97] identified about 100 metabolic reactions in *E. coli* that could in theory be replaced by an essential *Mtb* homologue. Based on that list, pathways were prioritized for the development of TESEC strains. We took the following considerations into account:

- Targeted protein domains should be absent in the human proteome.
- The *Mtb* target should be essential for host survival and is preferably conserved among natural isolates
- *Mtb* targets that are up-regulated in the latent state, since in-

hibitors for essential targets in non-growing *Mtb* are necessary for the End TB strategy of the World Health Organisation [86]

.

Targets that met those requirements are explained in more detail in the following section of this chapter.

1.2.2 Primary *Mtb* drug targets

Peptidoglycan biosynthesis Peptidoglycan (PG) is a crucial cell wall component for nearly all bacterial species. Its network structure is important in the maintenance of cell morphology and for the prevention of osmotic lysis [105]. PG can be described as a biopolymer made of N-acetylmuramic acid (MurNAc) and N-acetylglucosamine (GlcNAc) units. MurNac monomers carry short oligopeptides composed of L-alanine, D-glutamate, *meso*-diaminopimelic acid (DAP) and D-alanine (D-ala) that form bridges between the $(\text{MurNAc-GlcNAc})_{\text{n}}$ strands (Figure 1.1) [106, 107]. In both *E. coli* and *Mtb*, PG forms the inner layer of the cell wall and features 3→4-linked D-ala-DAP or 3→3-linked DAP-DAP peptide bridges for support. However, mycobacterial PG also covers several unique features compared to other bacteria [108, 109].

One of them is a high degree of 3→3 cross-linking between DAP molecules [109]. Those DAP-DAP bridges were found to play an important role as virulence factor [110] and are more common in the stationary phase [111]. Other PG features specific to mycobacteria are the presence of N-glycosylated muramic acid in addition to N-acetylated residues [112] as well as the high degree of peptide cross-links (∼80%) compared to *E. coli* PG (∼30-50%) [113].

Many commonly used antibiotics target enzymes involved in peptidoglycan synthesis as instability of PG leads to bacterial cell death

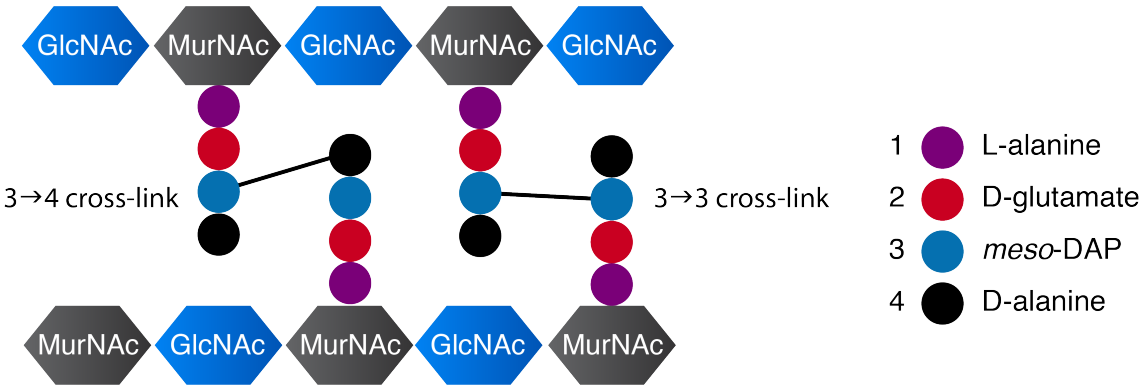

Figure 1.1: Schematic representation of peptide bridges in peptidoglycan of *Mtb* and *E. coli*.

GlcNAc = N-acetylglucosamine, MurNAc - N-acetylmuramic acid

[114].

While numerous essential biosynthesis enzymes involved in PG construction are suitable targets for antibiotic development, this thesis has focused on the synthesis of two indispensable components of PG: D-alanine and *meso*-DAP.

D-alanine is generated through alanine racemase (Alr) racemization activity of L-alanine into D-alanine. *Mtb* alanine racemase is a known *Mtb* target that is inhibited by the second-line drug D-cycloserine [115]. The use of a TESEC strain expressing *Mtb*-Alr in combination with D-cycloserine as an Alr-specific control drug was instrumental for the development and validation of TESEC-based applications (Chapter 2 & 3).

Moreover, *meso*-diaminopimelic acid biosynthesis was another target of interest, since *Mtb* and *E. coli* enzymes involved in DAP biosynthesis have been thoroughly characterized and may provide novel opportunities for antibacterial development [116, 117, 118, 119, 120]. Knock out studies in both *M. smegmatis* and *Mtb* identified that a deficient DAP pathway hampers mycobacterial growth ([121, 122, 123]. For example, the *Mtb* gene *asd* encodes for aspartate-semialdehyde dehydrogenase

(Asd) and is a branch point enzyme in the biosynthesis of isoleucine, threonine, methionine, and DAP [124]. A recent study showed the dependence of Asd expression for the growth, survival, and pathogenicity of *Mtb* in vitro [122]. Likewise, *E. coli* deletion mutants for *asd*, *dapA*, *dapB*, *dapD*, *dapE* are known to have an obligate requirement for DAP in the culture medium [125].

Tryptophan biosynthesis The pharmacological disruption of tryptophan biosynthetic enzymes is an attractive strategy to treat a tuberculosis infection. The tryptophan pathway does not exist in mammals, but it is essential for the survival of *Mtb* [126]. The necessity of tryptophan metabolism in *Mtb* becomes even more apparent in the context of the host immune response. IFN-γ mediated activation of macrophages triggers the expression of the tryptophan degradase indoleamine 2,3-dioxygenase to deplete cytosolic levels of tryptophan. This is an appropriate strategy to kill or inhibit intracellular pathogens that depend on the host to supply tryptophan [127].

However, in non-auxotrophic pathogens like *Mtb*, the reduction of extracellular tryptophan triggers the pathogen to upregulate tryptophan biosynthesis genes that allow survival of the immune response [128]. Therefore, inhibition of essential enzymes for tryptophan biosynthesis could complement the natural immune response and support the clearance of *Mtb*. Anthranilate synthase component I TrpE and type III phosphoribosyltransferase TrpD have been identified as potential drug targets within this pathway [126, 128]. These enzymes respectively catalyze the first and second reaction in L-tryptophan biosynthesis. Previous research showed that a *Mtb* trpD knockout mutant

lost its virulence in immunodeficient mice [129]. Zhang *et al.* reported that *Mtb* deletion strains lacking tryptophan biosynthesis genes *trpE* or *trpD* were unable to survive without extracellular tryptophan during cultivation [128]. In addition, they showed that the *trpE* deletion strain was significantly more susceptible to the defense mechanisms of CD4 T and macrophages [128].

Sulfur metabolism Sulfur has been identified as an essential element for *Mtb* survival in the host macrophage. Genes involved in *Mtb* sulfur metabolism have been reported to be upregulated in response to reactive oxidative stress, nutrient depletion, and reduced bacterial growth [130, 131, 132]. In both *E. coli* and *Mtb*, the incorporation of sulfur-containing amino acids such as cysteine and methionine relies on the import of sulfate and reduction to sulfite. The conversion of adenosine 5-phosphosulphate (APS) to sulfite (SO_3^{2-}) by CysH represents the first step in the sulfur assimilation pathway in *Mtb* [133]. Similarly, in *E. coli* APS is first converted into 3-phosphoadenosine 5-phosphosulfate (PAPS), whereafter CysH converts PAPS into sulfite. CysH appears to be critical to *Mtb* during defense against the immune system in mice [134].

1.2.3 Design features TESEC strains

Engineering control over cellular protein levels is central to metabolic engineering and synthetic biology in general [135]. For our TESEC strains, we strived for a system that would allow us to linearily and uniformely express essential proteins through exposure to varying inducer concentrations.

More specifically, we aimed at developing strains and conditions that would allow for the following features:

1. A TESEC background strain should not grow in the absence of essential target gene expression
2. TESEC's growth dynamics should depend on inducer levels in order to find minimal target level requirements
3. The cellular efflux of drug molecules should be minimized to increase the chance of finding target-specific inhibitors
4. Strain characterization and drug discovery assays should be simple and not rely on expensive reagents

1.2.4 Controlling the expression of *Mtb* essential genes for the discovery of hit compounds

We envisioned two possible TESEC-based drug screening strategies:

1. A differential screening based on **structural differences between homologous proteins** of *E. coli* and *Mtb*
2. A differential screening based on **differences in intracellular target protein concentrations**

Both strategies require controlled expression of the essential genes and have advantages and disadvantages.

Differential screening based on structural differences between homologous proteins of *E. coli* and *Mtb*

Sub-optimal expression levels of the target protein could negatively impact bacterial fitness [136] and potentially lead to false positives in the drug screen. For the first strategy, we wanted to find the minimal induction level at which TESEC growth was similar to the wild type growth. As such, the cell would endure minimal metabolic burden related to heterologous target under- or overexpression, while being sensitive to target-specific molecules.

For each drug target, we aimed at developing two TESEC strains. One strain with inducible expression of its native *E. coli* protein and another strain with the inducible expression of the *Mtb* homolog. As a result of *E. coli* and *Mtb* enzymes being evolutionarily distant (Table 1.1), differential effects between two strains can reveal the specificity of *Mtb*-specific inhibitors.
Narrow-spectrum antibiotics have various advantages over broad-spectrum antibiotics [137] as they reduce the development of antibiotic resistance in other microbes, and minimize the negative impact of the host microbiome. This is particularly relevant for tuberculosis treatments in which patients experience long-term exposure to antibiotics [1, 138].

On the other hand, although the protein sequences of our focus targets display a ∼30% homology (Table 1.1), important domains for protein activity are frequently conserved amongst essential enzymes

[124]. For example, several key residues in Asd are highly conserved across bacterial species [139]. Furthermore, a study [140] that compared several essential metabolic enzymes from *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* and *Enterobacter aurogenes*. *Mycobacterium tuberculosis*, *Mycobacterium smegmatis* and *Escherichia coli* revealed that drug pocket regions were highly conserved amongst species.

Pathway	*E. coli* K-12	*M. tuberculosis* H37Rv	Protein sequence Identity
meso-diaminopimelic acid biosynthesis	Asd	Asd	27%
meso-diaminopimelic acid biosynthesis	DapA	DapA	32%
meso-diaminopimelic acid biosynthesis	DapB	DapB	28%
D-alanine	DadX	Alr	29%
Sulfur metabolism	CysH	CysH	26%
Tryptophan metabolism	TrpD	TrpD	20%
Tryptophan metabolism	TrpE	TrpE	34%

Table 1.1: Protein sequence identities (%) of focus *E. coli* and *Mtb* essential enzyme targets.

For a TESEC-based drug discovery strategy that relies on structural differences between homologs, immediate exclusion of molecules that act on the *E. coli* essential enzymes could therefore also result in missed opportunities.

Although molecules may be found that specifically act on protein characteristics that are specific to *Mtb*, the chances of finding those are lower (especially in a lab-setting with a limited compound screening capacity).

And although it is likely that *E. coli* and *Mtb* enzymes have distinct affinities for target specific inhibitors, a differential screening that depends on those differences might be challenged by a lack of signal-to-noise. For example, several compounds that inhibit *Mtb* alanine racemase, were found to have similar inhibition constants for *E. coli* alanine racemase [141].

Differential screening based on differences in intracellular target protein concentrations.

An alternative TESEC-based drug screening relies on differences in target protein concentrations, instead of biochemical properties of proteins. Different target protein levels can make cells more or less vulnerable to target specific drugs [144, 143] (Figure 1.2c & d). The maximal differential signal for target-specific drugs can be obtained when we use conditions that allow for TESEC strains to grow with relatively low or high target expression levels. Because the required or tolerated activity is different for each essential enzyme, this strategy also required a system to fine-tune the target expression in each TESEC strain (Figure 1.2).

L-arabinose inducible system

We utilized a modified version of the L-arabinose inducible expression system with improved log-linear expression [2] of our target enzymes (details are explained in Chapter 2, Figure 2.1).

The L-arabinose inducible system of *Escherichia coli* (Figure 1.3) [145] is well characterized and inherently provided some of our desired features [146]. The activity of the system can be regulated through

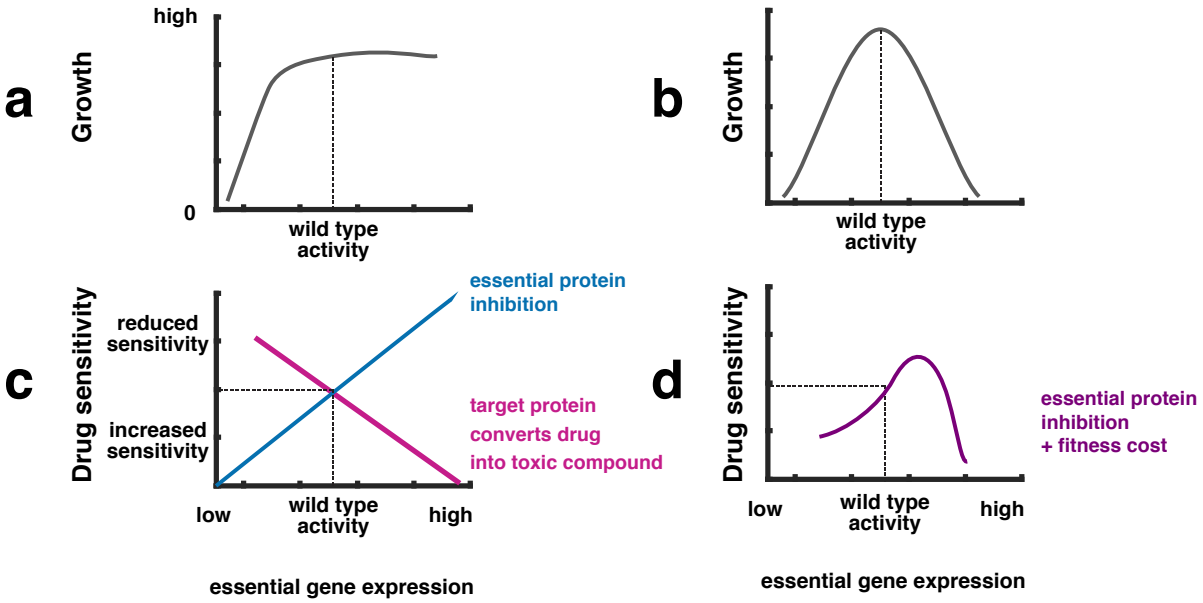

Figure 1.2: The expression levels of essential proteins affect growth and drug sensitivity of bacteria.
Top. Growth effects of varying the levels of essential proteins. a & b. For any given essential protein, the complete absence of expression results in the lack of bacterial growth. By increasing the expression levels (e.g. through the addition of an inducer molecule to the growth medium), growth is restored. The fitness effect of increasing expression levels depends on how much protein activity is required, and how much is tolerated. **Panel a**. Overexpression of a hypothetical essential protein has minimal impact on growth. **Panel b** Overexpression of a hypothetical essential protein negatively impacts the fitness of the strain **Bottom. Effect of the expression levels of essential proteins on their sensitivity to a specific drug.** Different mechanisms of action can result from drugs and their specific target proteins. For some drug-target pairs, increases in the essential protein levels correlate with the increase in drug resistance **(c - blue line)**. Another possibility is that the target-specific drug does not inhibit the essential protein function, but is converted into a toxic compound that is lethal for the cell in a dose-dependent manner **(c - pink line)**. Finally, in cases where overexpression is detrimental, drug resistance can follow a non-monotonic behavior where resistance decreases at those fitness-compromised ranges **(d - purple line)**. Modified from [142, 143].

the use of varying L-arabinose concentrations. In addition to this, relatively tight repression can be achieved since its transcription factor AraC has the dual role of acting as repressor as well as activator of gene expression (Figure 1.3) [147, 146]). Importantly, L-arabinose is an inexpensive molecule [148] and supports the development of a low-cost platform that is financially accessible.

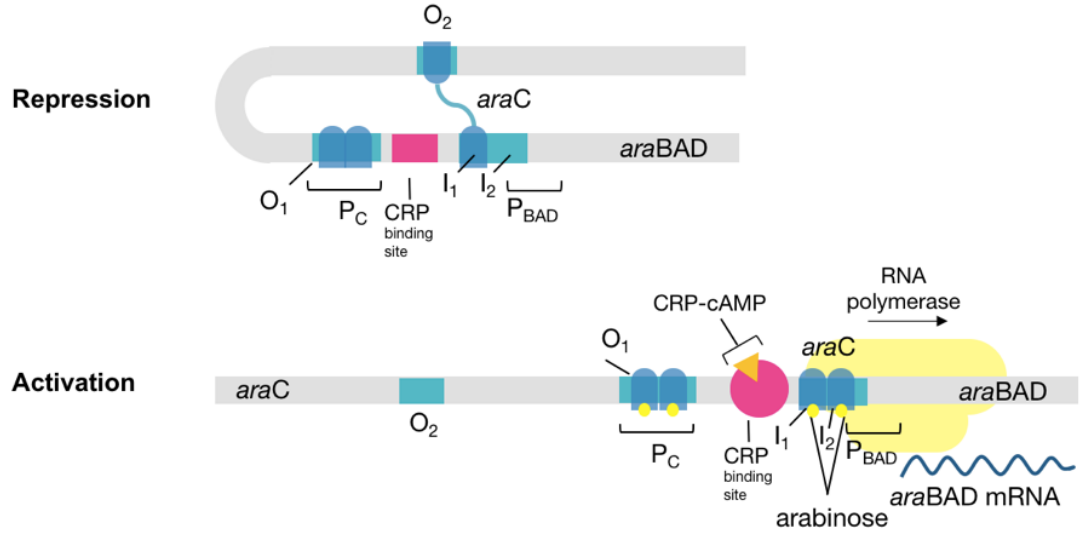

Figure 1.3: Repression and activation of the L-arabinose operon in *Escherichia coli*. AraC is a positive and negative regulator protein. In the absence of L-arabinose, AraC binds to O_2 and I_1 sites to form a loop that contributes to the repression of the *araBAD* genes. In the presence of L-arabinose, AraC undergoes a conformational change and migrates from O_2 to the I_2 site to release the DNA loop. The system requires another activator CRP-cAMP to promote the transcription of *araBAD*. When bound to cyclic AMP (cAMP), the catabolite repressor protein (CRP) binds to a CRP binding site next to I_{12}. Therefore, in media with glucose, in which cAMP levels are low, expression is significantly repressed. (Image is a modified version from Bio-Siva.50webs.org [149]).

Constitutive expression of L-arabinose transporters

To potentially improve the L-arabinose dose-response of our engineered cells, we engineered TESEC strains with an alternative regulation of the L-arabinose transport. The L-arabinose uptake is facilitated by three transport systems AraFGH, AraE and AraJ. AraFGH is a high affinity transporter (∼10 uM) [150, 151] and AraE is a low affinity high capacity arabinose/proton symporter (140 to 320 μM) [151] that serves as the dominant uptake system in the presence of intermediate to high arabinose levels [152]. Although the function of AraJ remains elusive [153], based on structural analogies, AraJ is expected to be a transporter for arabinose polymers [154] or an exporter of arabinose [155, 156]. The intracellular arabinose concentration positively regulates the transcription of *araFGH* and *araE* genes via the p_E, p_{FGH}

and p_J promoters, and as a result the number of arabinose transporters dictates the internal arabinose concentration and vice versa. Previous studies [153, 157, 158, 159] have shown that this feedback loop causes a non-linear activation of p_{BAD}-AraC in response to the extracellular arabinose concentration. Decoupling of this feedback loop by introducing independent AraE expression from a constitutive promoter improved the homogenous expression from the p_{BAD} promoter [157, 158]. Therefore, we aimed to explore the performance of TESEC strains in which the native p_E promoter was replaced by a constitutive promoter.

Reducing the bacterial efflux of drug compounds

The export of antimicrobial compounds from all classes is mediated by the expression of bacterial efflux pumps. This forms a major challenge in antibacterial resistance [160]. *E. coli*'s genome encodes for efflux systems that naturally protect the cell from the potentially detrimental effects of external agents. Its efflux machinery is composed of an inner and outer membrane part that together form a tripartite complex [161]. The outer membrane protein TolC [162] is the essential component of the tripartite efflux pumps that assembles in various combinations with inner membrane proteins [161, 163]. As such, families of efflux pumps have been classified that excrete different types of compounds [163]: AcrAB-TolC (Resistance-Nodulation-Cell Division family) for antibiotics, HlyBD-TolC (Type I secretion system) for hemolysin, EmrAB-TolC (Major Facilitator Superfamily) for antibiotics and bile, and MacAB-TolC (MacB Superfamily) for antibiotics and toxins. Without TolC, all mentioned efflux systems are defective and can not offer bacterial protection [161]. Because mycobacterial drug efflux is significantly divergent from *E. coli* [164, 69, 165] we decided to remove TolC expression in our *E. coli* background strain to

allow a broader range of compounds to access our *Mtb* targets of interest. The importance of this deletion for our drug screening application will be covered in Chapter 2.

1.2.5 Final TESEC-based drug screening strategy

Finding the right conditions and genetic backgrounds to achieve the intended strain behavior (Section 1.2.3) was not evident for all TESEC strains.

Apart from alanine racemase expression, all TESEC strains that expressed their native *E. coli* enzyme grew normally in the absence of L-arabinose. Expression from the feed-back controlled L-arabinsose inducible system [2] in various contexts did not sufficiently lower basal expression levels in order to find the minimal induction condition for growth.
Although glucose acts as a catabolic repressor of the expression by lowering the levels of cyclic AMP [166], glucose could not stop inducer-independent growth. Leaky expression provided sufficient enzyme for *E. coli* growth and made it difficult to find screening conditions that generated *E. coli*-*Mtb*-TESEC pairs with equal sensitivities to drugs.

This led us to focus on the second strategy that takes use of differences in target level expression. Experiments in which we exposed a TESEC strain with *Mtb*-Alr expression to series of L-arabinose and D-cycloserine concentrations showed that the strain's drug sensitivity range was dose-dependent on the inducer concentration. Even when growth profiles at different induction levels resembled each other in the absence of D-cycloserine, the sensitivity of the cell to D-cycloserine was significantly increased or reduced (Chapter 2 Figure 2.4). A dif-

ferential drug screening involving a hypersensitive TESEC strain with low *Mtb* alanine expression or a hyperresistant TESEC strain in the presence of a relatively high L-arabinose concentration, could, therefore, allow us to find target-specific inhibitors. Previous studies that adopted similar chemical-genetic strategies [167, 102, 168] showed that these approaches are powerful to find relevant hit compounds or investigate mechanisms of action for known inhibitors. In addition, clues about the specificity of the drug to the *Mtb*-target could be determined by a counter screen using wild type *E. coli*. With this set-up, I screened 60 compounds of the NIH/NCI natural product set III [169]. As expected for this small screen, no novel hits were discovered. However, this result provided confidence that the screening method was robust and did not lead to significant false positive rates. Importantly, maximal sensitivity of the TESEC with low *Mtb*-Alr expression was necessary to confirm to positive control D-cycloserine as a 'hit' at the standard library compound concentration of 0.1 mM.

Using the same strategy, we later developed a drug screening effort with 1,280 compounds of the Prestwick Chemical Library [3]. The results of this drug screen that led to the discovery of TB hit compound benazepril are described in Chapter 2.

In addition to TESEC-*Mtb*-Alr, five other characterized TESEC strains with expression of *Mtb*-Asd, *Mtb*-DapA, *Mtb*-DapB, *Mtb*-CysH and *Mtb*-TrpD are ready to be screened based on the low and high target expression strategy.

Finally, because further genetic optimization of expression is presumably necessary for other targets of interests, I built a golden gate

library based on the CIDAR MoClo [7] format and adapted it to the needs of the TESEC platform. This golden gate library standardizes and simplifies the construction of expression vectors to aid the optimization of strains and facilitate collaborations to expand TESEC as a shared technology platform.

1.3 Results

1.3.1 Characterization of auxotrophic deletion strains

The first step in the development of a TESEC strain is the deletion of the essential gene (or set of genes) in *E. coli* that we ultimately want to target with drugs. The auxotrophic phenotypes of deletion strains were verified by cultivation in medium supplemented with varying concentrations of an essential nutrient that rescued the bacterial growth (Appendix Figure A.1). All tested *E. coli* deletion strains were growth-dependent on the external addition of the molecule they were unable to produce. The growth-dependence was verified in M9 medium when the essential amino acid was present in LB medium.

1.3.2 Growth-rescue through the dose-dependent expression of *Mtb*-target enzymes

To complement the gene deletion of our auxotrophic background strains, plasmids carrying the essential enzymes responsible for the biosynthesis of *meso*-diaminopimelic acid/lysine, cysteine, tryptophan and D-alanine were constructed.

We equipped deletion strains with a genetic circuit to generate *E. coli* strains with growth-dependency of essential target gene expression. This system, based on Daniel and co-workers [2], is a modified

version of the native arabinose inducible operon (Figure 1.3) and was chosen to allow a logarithmically linear expression over a wide range of L-arabinose concentrations. It includes a low copy (∼5 copies/cell) plasmid to generate AraC under control of the p_{BAD} promoter and a high copy plasmid (15-20 copies/cell) to express the essential genes under control of the p_{BAD} promoter (Figure 2.1).

1.3.3 Characterization TESEC strains in fructose and glucose medium

TESEC strains carrying essential enzymes responsible for the biosynthesis of *meso*-diaminopimelic (Asd, DapA and DapB), D-alanine (Alr) and cysteine (CysH) were characterized in M9 fructose and glucose medium with a variety of L-arabinose concentrations (Figure 1.4). The use of fructose as a carbon source generally resulted in increased growth reduction for elevating inducer concentrations when TESEC strains expressed *E. coli* enzymes (Figure 1.4.A). For the same condition, TESEC strain TS008 with Alr_{Mtb} expression displayed a similar growth profile as TS009 with $\text{DadX}_{E.\ coli}$ production. For all other homologous enzyme pairs, the growth response of *E. coli*- versus *Mtb*-TESEC strains in the presence of L-arabinose concentrations varied.

In general, relatively high L-arabinose concentrations were necessary to restore growth of *Mtb*-TESEC strains in glucose medium. On the contrary, it was not possible to arrest the growth of most *E.coli*-TESEC strains in absence of the inducer (except TS009 with *E. coli*-DadX expression). Even in medium supplemented with glucose that represses p_{BAD} expression [145] (Figure 1.4.B). This suggested that very low levels of *E. coli*'s essential proteins (leaky expression) were sufficient to establish growth. Compared to *E. coli*-TESEC strains, *Mtb*-TESEC strains required exposure to higher arabinose levels to rescue bacterial growth. Possibly because of a lower catalytic activity

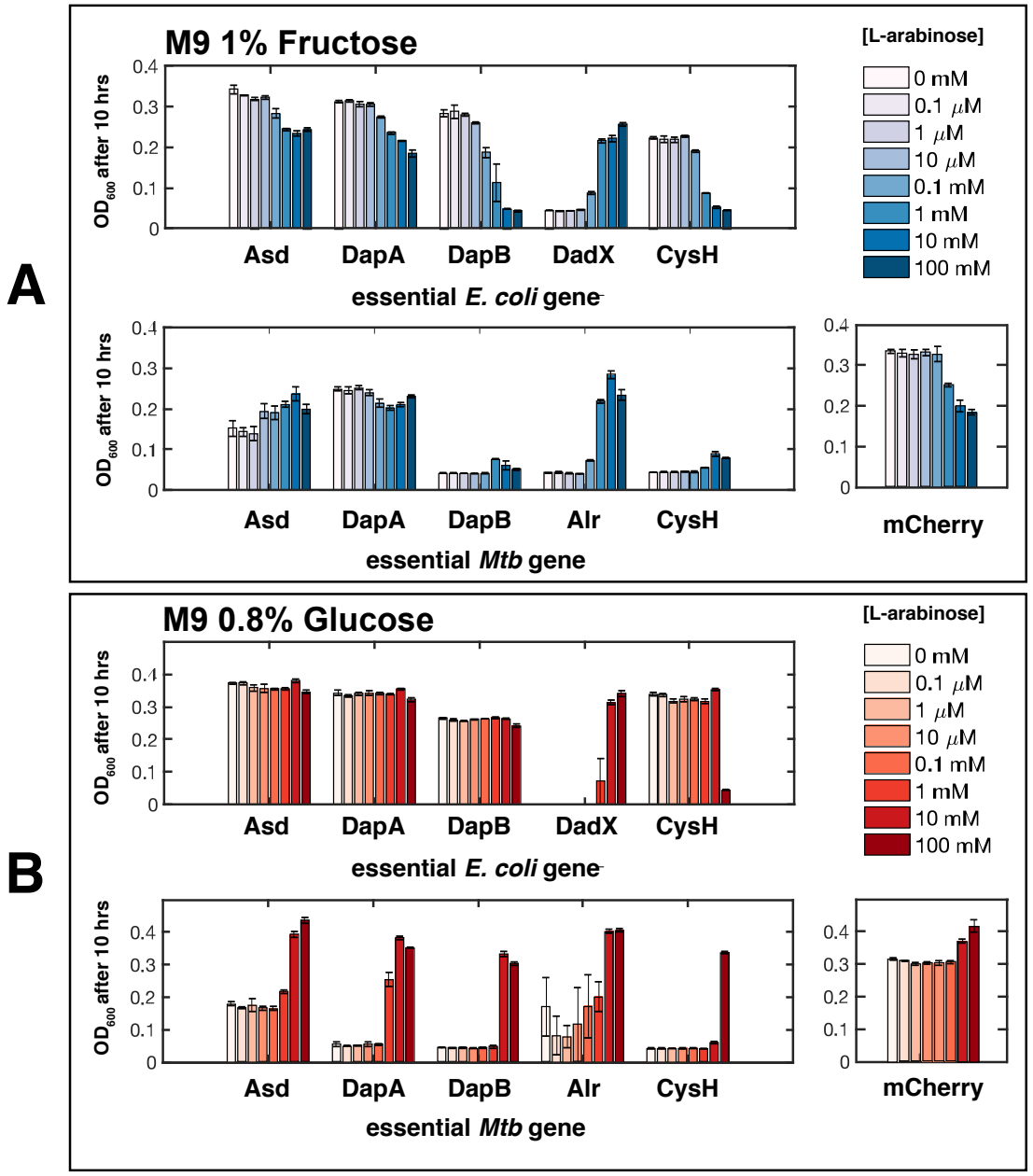

Figure 1.4: Testing the impact of various induction and media conditions on TESEC strains expression essential *E. coli* or *Mtb* proteins.
The bar graphs represent final OD_{600}'s of TESEC strains TS002 (*Mtb*-asd), TS003 (*E. coli*-asd), TS004 (*E. coli*-dapA), TS005 (*E. coli*-dapA), TS006 (*Mtb*-dapB), TS007 (*E. coli*-dapB), TS008 (*Mtb*-Alr), TS009 (*E. coli*-Alr), TS010 (*Mtb*-CysH), TS011 (*E. coli*-CysH) after 10 hours of growth in **M9 1% fructose (panel A)** and **M9 0.8% glucose (panel B)** in varying L-arabinose concentrations. The indicated essential gene was deleted from the *E. coli* chromosome and transformed with the two-plasmid system including: 1. a 'low copy' number plasmid generating the AraC under control of the p_{BAD} promoter and 2. a 'high copy' plasmid carrying the essential gene from *E. coli* or *Mtb* homolog under control of the p_{BAD} promoter. In the non-auxotrophic control strain, mCherry replaced the essential gene expressed from the high copy plasmid. Error bars represent data from three technical replicates.

of *Mtb* essential enzymes.

1.3.4 Constitutive AraE or AraC expression

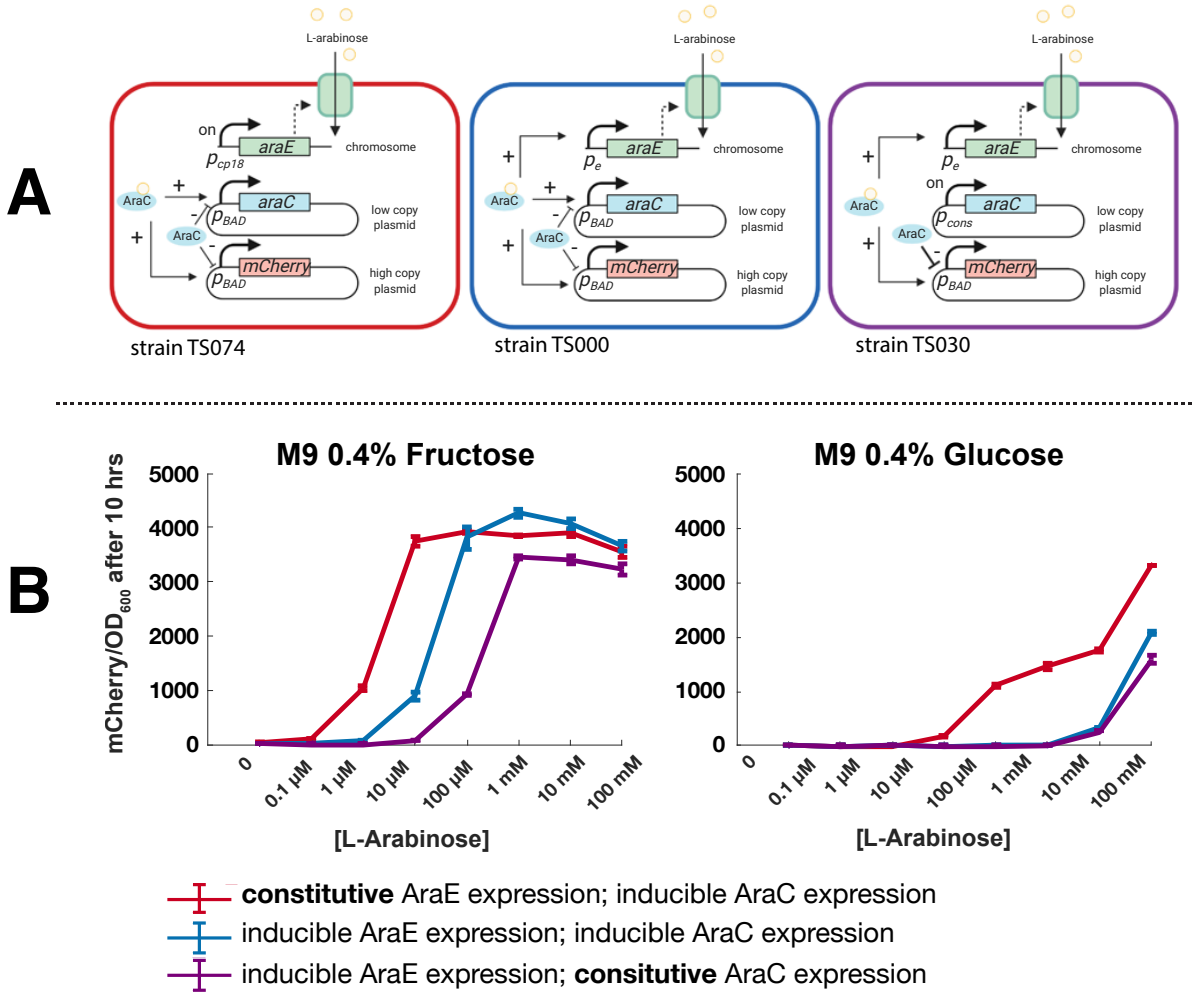

Figure 1.5: Effect of constitutive AraE or constitutive AraC expression on mCherry expression.
A. Schematic representations of strains TS074, TS000 and TS030 and their L-arabinose inducible expression systems. The color of the illustrated strain matches the results color in the graphs of panel B. **B.** mCherry expression in strain TS000 with the wild type AraE expression and inducible AraC expression (blue) compared to strain TS074 with constitutive AraE expression from the cp_{18} promoter (red) or strain TS030 with constitutive AraC expression (purple). All strains expressed mCherry under the control of p_{BAD}. Cultivation took place in M9 0.4% fructose or 0.4% glucose medium and 0-100 mM L-arabinose concentrations. Depicted are the fluorescence measurements divided by the OD_{600} after 10 hours. Error bars represent data from three technical replicates.

TESEC strains were constructed with differential AraE or AraC

expression. Variants carried either the wild type p_E promoter or the constitutive promoter cp_{18} from Keasling *et al.* 2001 [157] upstream the chromosomal *araE* gene (Figure 1.5, top panel, red strain).
In addition, the effect of constitutive AraC expression was tested by including plasmid pD.048 with AraC expression from BioBrick promoter J23106 [170] instead of V91 with inducible AraC expression from p_{BAD} (Figure 1.5, top panel, purple strain).

To verify the inducer-expression response of strains with altered L-arabinose systems, non-auxotrophic strains with the aforementioned backgrounds expressed *mcherry* under control of p_{BAD}. Moreover, strains were cultured in M9 medium with either 0.4% fructose or 0.4% glucose as their carbon source.
As expected, the various strain backgrounds and culture conditions impacted the mCherry expression dynamics (Figure 1.5). Compared to strains with native AraE expression, constitutive expression of AraE resulted in a shift of the dynamic range towards the lower inducer concentrations in fructose medium (Figure 1.5.B - left graph, red line). Constant AraE expression in glucose medium resulted in a broadening of the mCherry expression range (Figure 1.5.B - right graph, red line). The opposite effect was observed when AraC was constitutively expressed. The expression dynamics shifted to the higher end of L-arabinose in fructose containing medium, and a small reduction in expression levels was detected in glucose medium (Figure 1.5.B, purple lines).

Using this information, I tested whether more control over essential gene expression could be achieved when we used a background strain with araE expression from cp_{18} combined with cultivation in

glucose medium. It was not further investigated whether the continuous expression of AraE affected the homogeneity of expression as was described previously [158]. In addition, TESEC strains with constitutive expression of AraC were constructed and tested in glucose conditions to investigate whether the expression of essential *E. coli* proteins could be lowered.

TESEC strains with constitutive AraE expression

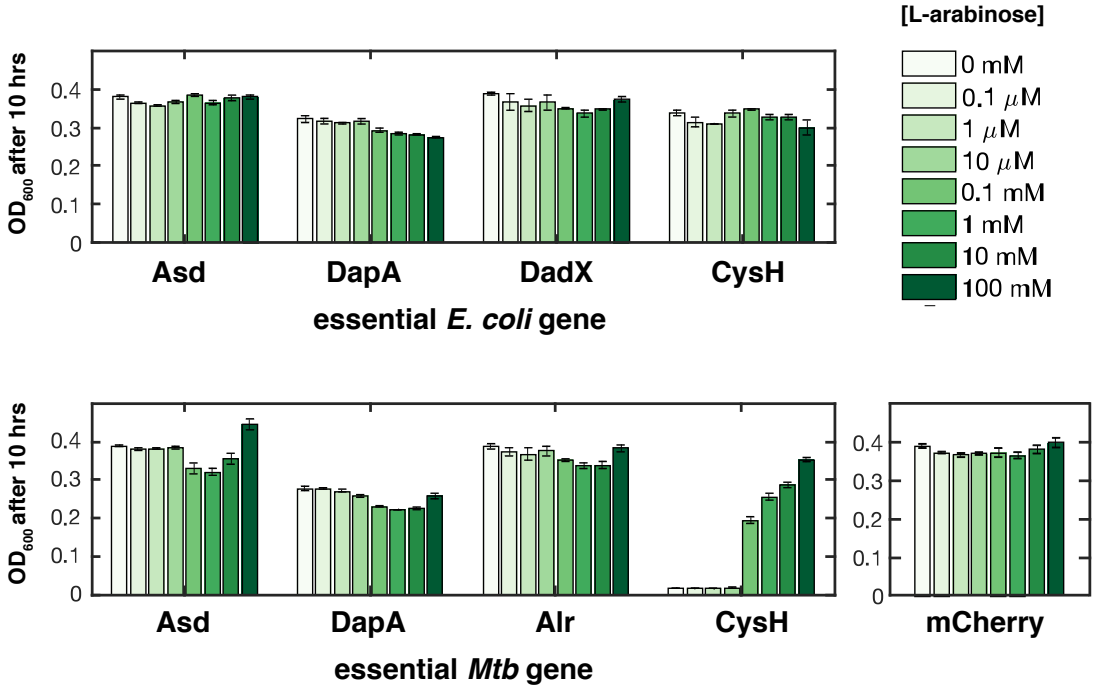

Figure 1.6: Effect of constitutive L-arabinose transporter AraE expression. TESEC strains TS082 ($Asd_{E.\ coli}$), TS085 ($DapA_{E.\ coli}$), TS093 ($Alr_{E.\ coli}$), TS095 ($CysH_{E.\ coli}$), (Asd_{Mtb}), TS084 ($DapA_{Mtb}$), TS092 (Alr_{Mtb}), TS094 ($CysH_{Mtb}$), TS083 control strain TS074 (mCherry) with initial OD_{600} 0.05 were exposed to 0-100 mM L-arabinose concentrations in M9 %0.4 glucose medium. After 10 hours of cultivation at 37°C, OD_{600} measurement were taken. Results represent the mean and standard deviations of three technical replicates. Fluorescence measurements of internal controls TS000 and TS074 with mCherry production confirmed the same expression dynamics as depicted in figure 1.5 (data not shown here).

When TESEC strains with constitutive expression of AraE were assayed, extra control over bacterial growth was observed for TESEC

TS094 with $CysH_{Mtb}$ (Figure 1.6). In the new strain background (cp_{18}-*araE*), growth could be modulated over a range of 10 μM – 100 mM L-arabinose concentrations instead of a single L-arabinose concentration at 100 mM (p_E-*araE*) (Figure 1.4.B - TS010 with $CysH_{Mtb}$ expression). Unfortunately, continuous expression of AraE led to undesired L-arabinose-independent growth for all other TESEC strains (beyond T094) with either *Mtb*- or *E. coli*-target expression. Even TESEC strains that displayed L-arabinose-dependent growth when AraE was expressed from the wild type promoter cp_{18} (Figure 1.4). For example, TESEC-Asd_{Mtb} with p_E-*araE* in figure 1.4.B compared to TESEC-Asd_{Mtb} with p_{cp18}-*araE* in figure 1.6.

Overall, I observed that a growth shift to low levels of L-arabinose (and even in absence) for TESEC strains with constitutive AraE expression corresponded with the earlier mCherry results (Figure 1.5). Fluorescence data also showed that mCherry expression in strain TS074 with constitutive AraE expression was increased in lower L-arabinose levels compared to strain TS000 with native AraE expression (Figure 1.5).

TESEC strains with constitutive AraC expression

Next, I tested the impact of constitutive AraC expression. Since AraC has a dual role in repression and activation [171] of p_{BAD} promoters, it was hypothesized that constitutive AraC expression could tighten the repression in absence of L-arabinose. This could be useful to gain increased control over the growth dynamics of TESEC strains that displayed growth in absence of the inducer (mainly *E. coli*-TESEC strains).

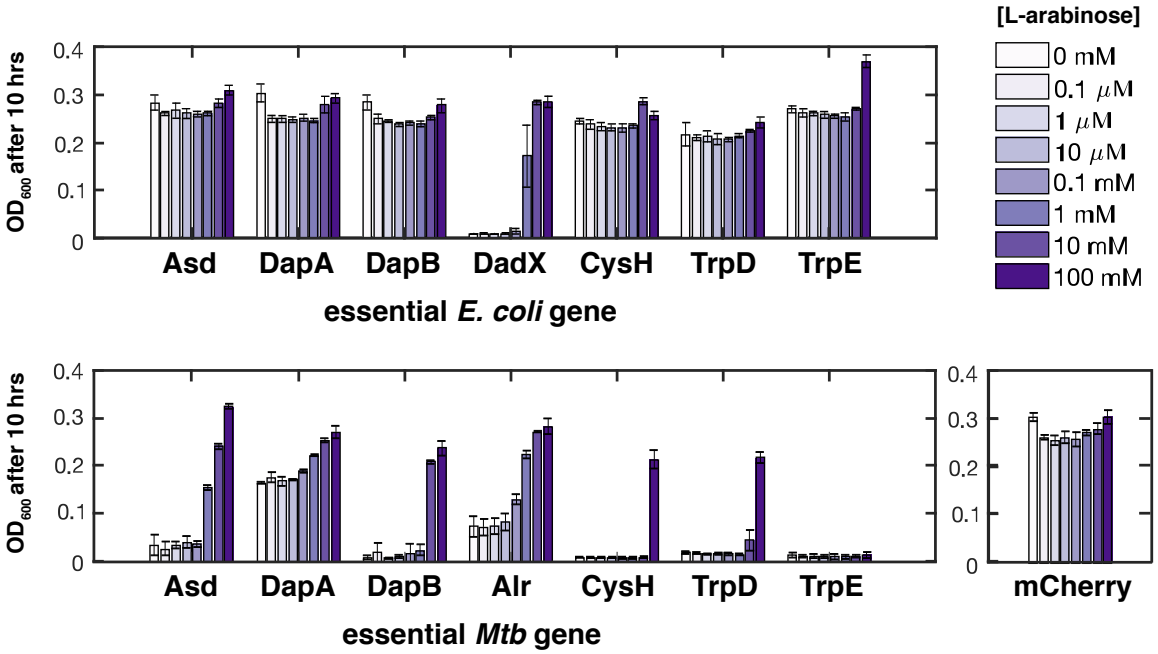

Figure 1.7: Effect of constitutive AraC expression
TESEC strains TS032 ($Asd_{E.\ coli}$), TS035 ($DapA_{E.\ coli}$), TS038 ($DapB_{E.\ coli}$), TS047 ($DadX_{E.\ coli}$), TS050 ($CysH_{E.\ coli}$), TS041 ($TrpD_{E.\ coli}$), TS044 ($TrpE_{E.\ coli}$), TS031 (Asd_{Mtb}), TS034 ($DapA_{Mtb}$), TS037 ($DapB_{Mtb}$), TS046 (Alr_{Mtb}), TS049 ($CysH_{Mtb}$), TS040 ($TrpD_{Mtb}$), TS043 ($TrpE_{Mtb}$), control strain TS030 (mCherry) with initial OD_{600} 0.05 were cultivated in M9 %0.4 glucose medium with varying L-arabinose concentrations (0-100 mM). Data represents the OD_{600} measurements of three technical replicates.

In parallel, we constructed additional TESEC strains with $TrpD_{E.\ coli}$ or $TrpD_{Mtb}$, and $TrpE_{E.\ coli}$ or $TrpE_{Mtb}$ target expression.

As before, strains were cultivated in M9 glucose medium to determine their inducer-growth dynamics. Constitutive AraC expression in glucose medium did not significantly reduce the basal expression of TESEC strains with $Asd_{E.\ coli}$, $DapA_{E.\ coli}$, $DapB_{E.\ coli}$, $TrpD_{E.\ coli}$ and $TrpE_{E.\ coli}$ to generate sub-optimal growth. However, the alternative AraC expression improved the control over the growth dynamics of TESEC strain TS034 with *Mtb* DapA expression. Although TS034 grew in absence of L-arabinose, its growth was significantly reduced compared to growth in the highest inducer concentration (Figure 1.7

- *Mtb* gene, DapA). This strain could be useful in a low versus high expression drug screen, because sub-optimal growth indicates that the strain is expressing sub-optimal levels of the essential protein and thus more sensitive for target-specific drugs compared to a wild-type *E. coli* strain.

While, $TrpD_{Mtb}$ was able to rescue the growth of the K004 deletion strain ($\Delta trpD$), growth rescue for the $\Delta trpE$ strain (K005) was not detected in this experiment, and neither in replicate experiments (data not shown here).

Selected drug screening conditions for each TESEC strain

In various TESEC strain characterization assays, I showed how the bacterial growth dynamics could be influenced using genetics or varying carbon sources in the growth medium. Table 1.2 summarizes the results from figure 1.4, 1.6 and 1.7. For each applied condition, the range of L-arabinose concentrations is specified that allows each TESEC strain to grow.

For drug screening based on structural differences between *E. coli* and *Mtb* enzymes, we aimed to find conditions that resulted in normal *E. coli* growth with a minimal induction level for each TESEC strain. Using these strains and conditions, we intended to find compounds that differently inhibit the growth of a TESEC strain with *Mtb* enzyme expression compared to a strain that expresses the same enzyme from *E. coli*.

Finding minimal induction conditions for each TESEC strain was complicated since most *E. coli*-TESEC strains grew in the absence of

L-arabinose (even in the presence of glucose) (Figure 1.4, 1.6 and 1.7). Despite testing different systems, basal expression sufficiently rescued bacterial growth.

The only target that could be used for a differential drug screen based on structural target differences is Alr_{Mtb}/$DadX_{E.\ coli}$. Both *E. coli*-TESEC strain with $DadX_{E.\ coli}$ expression and *Mtb*-TESEC strain with Alr_{Mtb} expression displayed inducer-dependent growth dynamics similar conditions (Figure 1.4 & 1.7). However, to develop a versatile drug screening assay that works for all *Mtb* targets of interest, we decided to focus on the second drug screening strategy based on low and high target expression.

For drug screening experiments based on low versus high target expression levels, conditions that allow for mild growth reduction in low inducer concentrations as well as growth in the higher end of the inducer range are preferred. This maximizes the change in drug sensitivities for different induction levels, and thus the signal that can be obtained from the inhibition of a target-specific compound. Regardless, strains with similar growth at all inducer concentrations could be used as well. Even when growth profiles look similar at various induction levels, the drug sensitivity is still increased or decreased because of variations in the cellular target levels (Figure 1.2). In this case, the challenge is that the drug sensitivity of the TESEC strain is more difficult to predict in absence of a control drug. Therefore, I highlighted in green, conditions that allow one to find induction levels that render TESEC strains with somewhat reduced growth compared to a wild type strain.

		Genetic or cultivation condition			
		M9 1% Fructose p_{BAD}-$araC$ p_{E}-$araE$	M9 0.8% glucose p_{BAD}-$araC$ p_{E}-$araE$	M9 0.4% Glucose p_{BAD}-$araC$ p_{cp18}-$araE$	M9 0.4% Glucose p_{Cons}-$araC$ p_{E}-$araE$
***E. coli*-TESEC**	Asd	0-100 mM	0-100 mM	0-100 mM	0-100 mM
	DapA	0-100 mM	0-100 mM	0-100 mM	0-100 mM
	DapB	0-100 mM	0-100 mM	not tested	0-100 mM
	DadX	0.1 mM - 100 mM	1-100 mM	0-100 mM	1-100 mM
	CysH	0-1 mM	0-10 mM	0-100 mM	0-100 mM
	TrpD	not tested	not tested	0-100 mM	0-100 mM
	TrpE	not tested	not tested	0-100 mM	0-100 mM
***Mtb*-TESEC**	Asd	0-100 mM	0-100 mM	0-100 mM	0-100 mM
	DapA	0-100 mM	1-100 mM	0-100 mM	0-100 mM
	DapB	no growth	10-100 mM	not tested	10-100 mM
	Alr	0.1 mM - 100 mM	0 - 100 mM	0-100 mM	0-100 mM
	CysH	no growth	100 mM	0.1 mM - 100 mM	100 mM
	TrpD	not tested	not tested	not tested	100 mM
	TrpE	not tested	not tested	not tested	no growth

Table 1.2: TESEC strains and their growth range in various genetic and cultivation conditions.
The table provides an overview of the results from figures 1.4, 1.6 and 1.7. Each concentration range refers to the L-arabinose levels that resulted in *E. coli* growth. Green high lights represent a relevant condition for drug screening purposes.

1.3.5 Primary Drug screen with 60 compounds of the NIH/NCI Natural Product Set III

The *Mtb*-TESEC strain with Alr expression was used to validate the application of the drug screening strategy based on low and high target expression. We chose alanine racemase as a first target for the development of the assay, because its activity is blocked by the known inhibitor D-cycloserine [172], which could serve as a positive control. An important advantage of our drug screen compared to more conventional assays is the increased drug sensitivity we achieved through the simple adjustment of inducer concentrations.

Accordingly, when the known *Mtb*-alanine racemase inhibitor D-cycloserine was added to TS008 that grew in 0-100 mM range of L-arabinose in M9 0.4% fructose, we observed a ∼30-fold increase in the sensitivity for the strain with the low $\mathrm{Alr_{Mtb}}$ expression compared to the strain with high target enzyme expression (Figure 1.8). Since

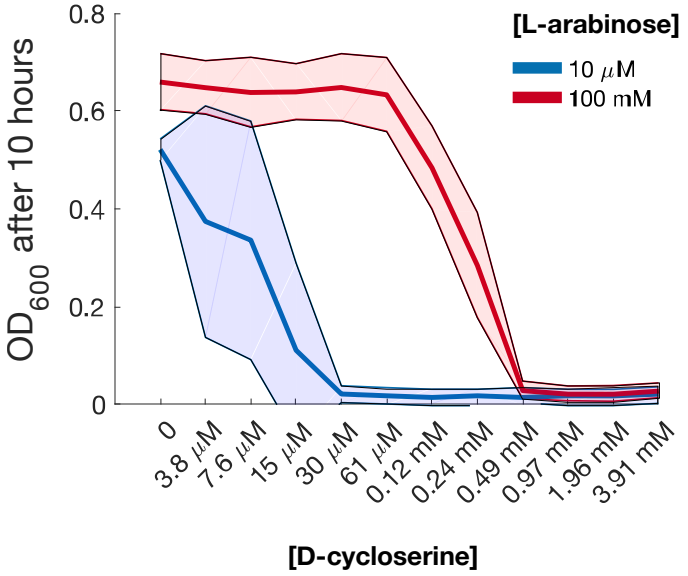

Figure 1.8: Increased drug sensitivity for TS008 with low (in 10 μM L-arabinose) and high (in 100 mM L-arabinose) Alr$_{Mtb}$ expression.
Shaded error bars represent the standard deviations of six independent experiments.

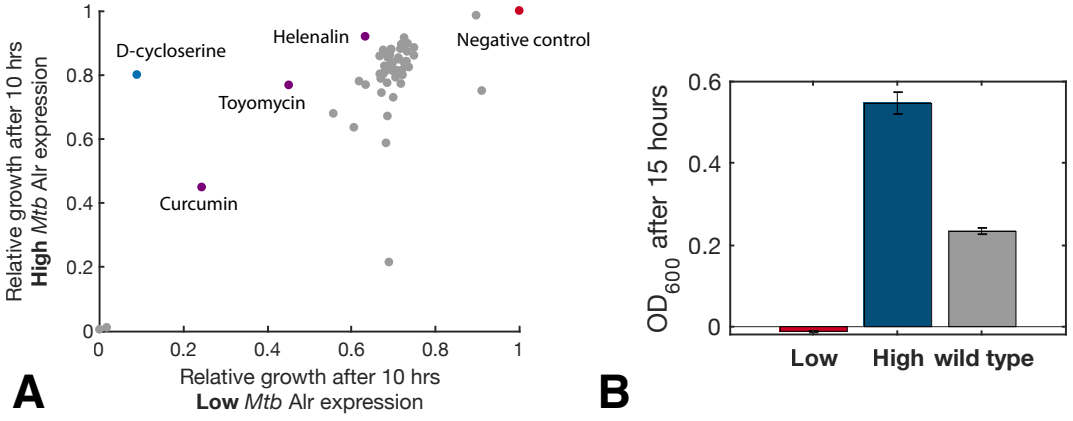

Figure 1.9: Preliminary Library screen with 60 compounds from the NIH Natural Product Set III.
90 μL cultures of TS008 (initial OD$_{600}$ = 0.05) with low Alr$_{Mtb}$ in the presence of 10 μM and high Alr$_{Mtb}$ expression with 100 mM L-arabinose were exposed to 0.1 mM compound concentrations. **A. Differential effect of compounds D-cycloserine (positive control), curcumin, toyomycin and helenalin.** OD$_{600}$ measurements were used to compare the growth of TS008 in the two induction conditions. OD-values were normalised for the growth of the strain in absence of compound. **B. Differential D-cycloserine (0.1 mM) inhibition of TESEC strain TS008 with low or high *Mtb* Alr expression and strain D20 with wild type expression of DadX and Alr.**

drug library screens typically involve the screening of compounds at a standard drug concentration of ∼0.1 mM, increased drug sensitivity could be an advantage.

Using the target-specific drug sensitivity range of TS008, I conducted a preliminary screen with 60 compounds the NHI Natural Products Set III library [169]. This library consists of a diverse set of substances from marine organisms, bacteria, fungi and plants. TS008 was exposed the compounds in the presence of 0.1 μM L-arabinose, referred to as 'low *Mtb* Alr expression', or 100 mM L-arabinose, 'high *Mtb* Alr expression'. For comparison, the non-auxotropic TESEC parent strain D20 that expressed its native alanine racemases was also included in this screen. Differences in OD_{600} after 10 hours of cultivation between the high and low induced conditions were determined. For each compound, differences in Z-scores were determined as well. Z-scores reflect the number of standard deviations from the mean growth in all 60 compounds. These values allowed me to score compounds on the basis of their ability to reduce the growth of the low versus the highly induced TS008 strain. Positive control D-cycloserine displayed the expected differential 'hit signature' (Figure 1.9.A & B). The growth of the low induced strain was completely inhibited by 0.1 mM D-cycloserine (more than the *E. coli* control strain), whereas the high induced strain presented relatively normal growth (Figure 1.9.B). I used the $\Delta OD_{\text{high-low induction}}$ and $\Delta Z_{\text{high-low induction}}$ scores of D-cycloserine, respectively ∼0.5 and ∼4, as threshold to identify hit compounds from the library. As expected from a screen with only 60 compounds, none of the library compounds yielded $\Delta Z_{\text{high-low induction}}$-scores >4 or $\Delta OD_{\text{high-low induction}}$ values >0.5 to qualify as significant 'hits'. 90% of the Z-scores for both the low and high induced strain

were between -1 and 1. Known antibiotics, chlortetracycline [173] hydrochloride and streptonigrin [174] disabled the growth of all strains. Three molecules with the highest $\Delta OD_{\text{high-low induction}}$ scores from this screen were curcumin, toyocamycin and helenanin (Figure 1.9.A). Helenin is known to have antibacterial properties against *E. coli* [175], but its precise mechanism of action is not reported. Toyomycin is capable of binding in reverse to a GC-specific DNA ligand in the minor groove which inhibits transcription [176]. Consequently, this could have made TS008 with low Alr_{Mtb} expression more vulnerable compared to TS008 with high Alr_{Mtb} expression. Curcumin is known for its antibacterial properties [140, 177] and prevents Z-ring assembly of *B. subtilis* 168 [177]). Overall, this screen showed the importance of having a sensitive system and that despite different cellular Alr_{Mtb} levels, TS008 responded similarly to the majority of compounds. This validated that potential fitness differences due to alternative Alr_{Mtb} quantities did not make the cell extra vulnerable to a random compound and that the screening strategy was reliable.

1.4 Discussion

Research with *Mycobacterium tuberculosis* is challenging due to the pathogenic nature [178] and slow-growth [100] of the bacteria. This work represents an endeavor to develop alternative methods for the early-stage discovery of tuberculosis inhibitors based on engineered *E. coli*.

Despite the common use of *E. coli* strains for the production and isolation of *Mtb* enzymes [172, 179], to our best knowledge, *E. coli* has not been explored as a drug screening chassis for tuberculosis.

Towards this goal, I constructed TESEC strains with the expression of seven different *Mtb* essential enzymes and explored various ways to vary their expression levels and indirectly influence *E. coli* growth. In general, *E. coli* was easier to control when *Mtb* essential enzymes were expressed. Except *Mtb* TrpE (Figure 1.7), conditions were found that allowed all other mycobacterial enzymes to rescue *E. coli* growth over a range of inducer concentrations. The inability for TrpE to rescue *E. coli* growth could be explained as follows. In *E. coli*, anthranilate is synthesized from chorismate by the protein complex anthranilate synthase composed of both TrpD and TrpE [180]. Consequently, the *Mtb* TrpE may not able to effectively assemble with $\text{TrpD}_{E.\ coli}$ to function as TrpDE.

Obtaining sufficient control over *E. coli* targets was challenged by basal expression levels that led to full bacterial growth (Figure 1.4 1.6 & 1.6). A possible explanation is that essential proteins in *E. coli* are typically expressed at a low copy numbers per cell (∼10/cell) [181]. However, because we want to achieve maximal sensitivity to target-specific drugs, we required a system that allowed us to lower the expression level incrementally to only a few copies per cell. This could have been achieved through for example the addition of protein degradation tags [182], alternative RBS sequences [183], expression systems that use tetracycline induction [184] or more modern approaches such as the tCRISPRi system from Li *et al.* [185]. In attempts to solve this issue, I built an RBS library of various strengths (Appendix section A.0) and constructed a variant of the p_{BAD} promoter in which the O_2 half-site by the I_1 half-site for which AraC has a stronger binding affinity that may increase the promoter repression in absence of inducer

[186, 171] (Figure 1.3).

I halted efforts to further optimize *E. coli*-TESEC strains when we prioritized the drug screening method based on differences in expression levels instead of variations in biochemical characteristics between *E. coli* and *Mtb* enzymes. Nevertheless, the genetic tools I built may still be relevant for future strain optimizations.

TS008 with increased and reduced sensitivities to Alr-specific inhibitor D-cycloserine (Figure 1.8) was useful for drug screening. A preliminary differential drug screen based on different target level concentrations with with 60 compounds from the NIH library showed that strain generated a clear hit signature for the D-cycloserine (Figure 1.9) and grew normally (independent on the Alr expression level) in the presence of most other compounds.

This strain was further improved by deleting *tolC* encoding for the essential outer membrane component TolC of various drug efflux pumps. Because this gene deletion was essential for the discovery of tuberculosis inhibitor benezapril (Chapter 2), I constructed additional *Mtb*-TESEC strains (with Asd, DapA, DapB, CysH or TrpD expression) that include the same $\Delta tolC$ genotype (and $\Delta entC$ as explained in Chapter 2).

Together, this work laid the foundation of the current TESEC discovery platform. The strains, plasmids, methods and strategies I developed, tested or advanced are scalable and can be used to find *Mtb* target inhibitors.

1.5 Contributions

For this work I have been supervised by Dr. Ariel Lindner and Dr. Edwin Wintermute. The master internships of Sebastián Sosa Carrillo and Ayan Abukar I supervised contributed to the construction of TESEC strains and TESEC MoClo library constructs.

1.6 Methods

1.6.1 Strains & plasmids

A complete list of deletion strains, TESEC strains and plasmids (including strains and plasmids that not mentioned in the results section) are listed in Appendix A.1, A.2 and A.3.

1.6.2 Chemicals and reagents

Unless mentioned otherwise, chemicals including antibiotics, amino acids, sugars and salts were purchased from Sigma.

1.6.3 Storage of bacterial stocks

All bacterial strains were stored at -80°C in v/v 15% glycerol.

1.6.4 Plasmid construction

All *Mtb* and *E. coli* genes were synthesized (IDT) after codon optimisation for expression in *E. coli* K-12 and the removal of unwanted type IIS restriction sites. Genes were flanked by BsaI restriction sites and inserted downstream the p_{BAD} promoter of pTS000 (the construction of this plasmid is described in the method section of Chapter 2.7.3).

1.6.5 TESEC MoClo Library construction

An extension of the CIDAR MoClo golden gate library [7] was developed for the rapid and standardized assembly of TESEC expression vectors with various promoters, RBS sequences and genes. When possible, CIDAR MoClo pre- suffixes were adopted. However, to facilate the construction of fusion target enzymes with N- or C- terminal reporter genes (such as GFP or mCherry), additional assembly sequences

were added (Table 1.3). A complete overview of all current TESEC MoClo parts and plasmids can be found in Appendix Table A.0.

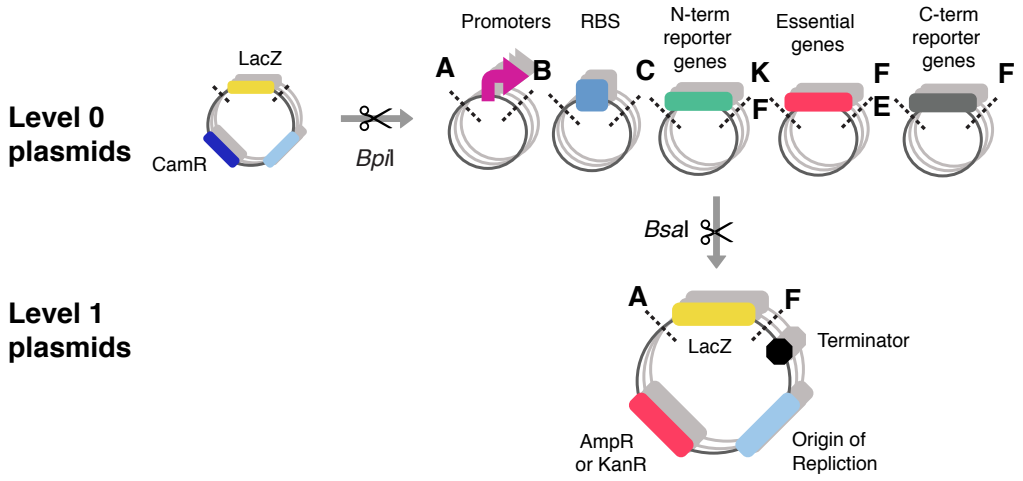

Figure 1.10: Schematic overview over the MoClo TESEC parts library

Pre-/suffix	TESEC MoClo	Pre-/suffix	CIDAR MoClo
A	GGAG	A	GGAG
B	TACT	B	TACT
C	AATG	C	AATG
D	TACC		
E	GGTG		
F	AGGT	D	AGGT
G	GCTT	E	GCTT
H	CGCT	F	CGCT
I	TGCC	G	TGCC
J	ACTA	H	ACTA
K	TTCT		

Table 1.3: TESEC MoClo pre- and suffixes compared to CIDAR MoClo [7]

A description of the construction of destination vectors and parts can be found in Chapter 2.7.3.

1.6.6 Gene deletions

E. coli gene deletions were introduced using reported methods from [125, 187]. For gene deletions that were present in strains of the Keio collection [125], P1 phage transduction was used to transfer the deletion of interest into the desired background strain. For required genotypes that were not available in the Keio collection, gene deletions in *E. coli* K-12 MG1655 were introduced based on the same protocol of Datsenko *et al.* [187]. Subsequently, P1 phage transduction was used to copy the gene deletion into other background strains. The replacement of the p_E promoter with the cp_{18} promoter in front of the chromosomal *araE* gene was done using P1 phage lysate obtained from strain BW2727 [157].

Preparation of deletion cassettes

Each gene deletion cassette was amplified from pKD13 using primers with 50-nt homologies up and downstream of the target site in *E. coli*'s genome (Table 1.4). Homology regions were based on the genome sequence with NCBI accession number NC_000913. PCR reactions were done as described below.

Phusion PCR reaction mix:

- 20 μL 5X HF Buffer (NEB)
- 1 μL Thermo Scientific® Phusion High-Fidelity DNA Polymerase
- 1 μL Deoxynucleotide (dNTP) Solution Mix (NEB)
- 10 μM Forward Primer (Table 1.4)
- 10 μM Reverse primer (Table 1.4)

Primer ID	**Primer description**	***E. coli* target gene**	**Template DNA**	**Sequence**	**Length (nt)**	**Tm (°C)**
NB.31	FW.pKD13-asd	***asd***	pKD13	GACCGGCACATTTATACAGCACA CATCTTTGCAGGAAAAAAACGCT TATGATTCCGGGGATCCGTCGA CC	70	63.9
NB.32	RV.pKD13-asd	***asd***	pKD13	CATCGCGCCCCAGATTTAATGAA TAAAGATTACGCCAGTTGACGA AGCATTGTAGGCTGGAGCTGCT TCG	70	62.2
NB.33	FW.pKD13-dapA	***dapA***	pKD13	CATACCAAACGTACCATTGAGAC ACTTGTTTGCACAGAGGATGGC CCATGATTCCGGGGATCCGTCG ACC	70	63.9
NB.34	RV.pKD13-dapA	***dapA***	pKD13	CAGAGTAAGCCATCAAATCTCCC TAAACTTTACAGCAAACCGGCAT GCTTTGTAGGCTGGAGCTGCTT CG	70	62.2
NB.35	FW.pKD13-dapB	***dapB***	pKD13	TGGTTACTCTGAAAACGGTCTAT GCAAATTAACAAAAGAGAATAG CTATGATTCCGGGGATCCGTCG ACC	70	63.9
NB.36	RV.pKD13-dapB	***dapB***	pKD13	ATTTTTGCACCATAACAAATATT TTGTGGTTACAAATTATTGAGAT CAAGTGTAGGCTGGAGCTGCTT CG	70	62.2
NB.102	FW.pKD13-folA	***folA****	pKD13	TATCGGGAAATCTCAATGATCA GTCTGATTGCGGCGTTAGCGGT AGATCGCATTCCGGGGATCCGT CGACC	71	59.6
NB.103	RV.pKD13-folA	***folA****	pKD13	AAGACGCGACCGGCGTCGCATC CGGCGCTAGCCGTAAATTCTATA CAAAATGTAGGCTGGAGCTGCT TCG	70	62.2
NB.104	FW.pKD13-folB	***folB***	pKD13	ACATTTCGCCGGAAGTGGTATCC GGTTAGCCAAAAAGCAGGCAGG ACGTGATGATTCCGGGGATCCG TCGACC	73	61.6
NB.105	RV.pKD13-folB	***folB***	pKD13	AACTACTGCGTATGACCAGGTTA TAACCGTTTGGTTTAACAGCTGT AAAATGTAGGCTGGAGCTGCTT CG	70	59.6

Table 1.4: Primers for the amplification of gene deletion cassettes.
***** Attempts to construct a *folA* deletion strain were unsuccessful to date. Later it was found that a *folA* deletion is lethal unless *thyA* is removed first [188].

- 100 ng pKD13
- dH_2O until 100 μL total reaction volume

Phusion PCR thermocycles:

1. 98°C for 30 seconds
2. 98°C for 10 seconds
3. 55°C for 15 seconds
4. 72°C for 30 seconds
5. 30 cycles back to step 2
6. 72°C for 7 minutes
7. Hold at 4 °C

PCR was followed by the digestion of pKD13 through the addition of 1 μL of DpnI (NEB) to the PCR reaction mixture and incubation at 37°C for 3 hours. DNA samples with deletion cassettes were purified and dialyzed against dH_2O over a Millipore MF Membrane Filter 0.025 μm (30 minutes at room temperature).

Transformation of deletion cassettes

An overnight culture of NB003 carrying pKD46 was diluted 1:100 in 5 mL LB with 100 μg/mL ampicillin and grown at 30°C; 150 rpm until the early exponential phase (OD_{600} 0.1-0.2) was reached. Next, 1 mM of L-arabinose was added to the medium to induce the expression of λ red from pKD46 during cultivation (30°C; 150 rpm) until $OD_{0.6}$. NB003 was harvested by centrifugation at 4000 rpm and made electrocompetent as described before [189].

To transfer the deletion cassette to *E. coli*, 200 μL ice-cold competent NB0003 was electroporated with 200 ng of salt-free deletion cassette using the MicroPulser Electroporator (Bio-Rad) and electroporation cuvets (ThermoFisher Scientific). Transformants were selected

on LB agar medium containing 50 μg/mL kanamycin and essential amino acid that successful clones should not auto-produce. Correct replacement of the target gene by the kanamycin resistance gene was confirmed by DreamTaq PCR and by verification of the auxotrophic phenotype.

Removal of Kan^r markers

Kan^r markers were removed from strains that resulted from P1 transductions or recombination with deletion cassettes by transforming them with pCP20 as described [187]. Colonies that grew after an overnight at 30°C on LB with 100 μg/mL ampicillin were picked for an overnight cultivation at 43°C in LB without antibiotics and with one of the essential amino acids. Successful recombination of the FRT flanking to remove the kanamycin resistance gene was verified by PCR and by testing if the bacteria lost the ability to grow in LB medium with 50 μg/mL kanamycin.

1.6.7 Verification of cloning and PCR

The composition of plasmids, or, the presence of genomic deletions/insertions were routinely verified using DreamTaq Green PCR Master Mix (2X) (ThermoFischer Scientific) following the supplier's instructions. PCR samples were loaded on 1% agarose, 0.5x TBE gels with SYBR Safe (ThermoFischer Scientific) for visualization. The DNA sequences of plasmids were continuously verified by Sanger sequencing (Eurofins Genomics) and subsequently analyzed with bioinformatics software tool Geneious.

1.6.8 Isolation of plasmids, PCR fragments and genomic DNA

PCR fragments were purified using Wizard® SV Gel and PCR Clean-Up System (Promega), plasmids were isolated from bacterial cultures with Wizard® Plus SV Minipreps DNA Purification System (Promega) and genomic DNA was isolated using Wizard® Genomic DNA Purification Kit (Promega). All procedures were performed according to the Promega's protocols.

1.6.9 Characterization of deletion strains

Amino acid requirements of auxotrophic phenotypes of TESEC background strains with essential gene deletions were experimentally verified by cultivation in dilution series of amino acids. Singly colonies of K001, K002, K003, K004, K005, D18 or D19 were grown overnight in 5 mL of:

Medium for overnight cultivations:

K001 (Δ*asd*), K002 (Δ*dapA*), K003 (Δ*dapB*) LB with 20 μg/mL 20 *meso*-DAP

D18 (Δ*alr* Δ*dadX*) LB with 100 μg/mL D-alanine

D19 (Δ*cysH*) M9 0.8% glucose medium including 1 g/L Drop-out Mix Complete, Adenine Rich, w/o Yeast Nitrogen Base (Powder) (USBiological Life Sciences)

K004 (Δ*trpD*), K005 (Δ*trpE*) M9 0.4% glucose medium including 1 g/L Drop-out Mix Complete, Adenine Rich, w/o Yeast Nitrogen Base (Powder) (USBiological Life Sciences)

Overnight cultures were washed twice in medium lacking essential amino acids. 96-well plates were prepared with 180 μL cultures of deletion strains (initial OD_{600} 0.005) and dilution series of their responding essential amino acid (as described below). 20 μL of mineral oil was added to each well to minimize liquid evaporation.

Experimental medium for each deletion strain:

K001 (Δ*asd*), K002 (Δ*dapA*), K003 (Δ*dapB*) LB with 0-120 μM *meso*-DAP dilution series

D18 (Δ*alr* Δ*dadX*) LB with 0-1000 μM D-alanine dilution series

D19 (Δ*cysH*) M9 0.8% glucose medium with 1 g/L Drop-out Mix Synthetic Minus Histidine, Methionine, Cysteine, Adenine Rich w/o Yeast Nitrogen Base (DO, Dropout, Powder) (USBiological Life Sciences)

K004 (Δ*trpD*), K005 (Δ*trpE*) M9 0.4% glucose medium with 1 g/L Drop-out Mix Synthetic Minus Tryptophan w/o Yeast Nitrogen Base (DO, Dropout, Powder) (USBiological Life Sciences) supplemented with 0-10 μM L-tryptophan.

D20 (control strain) D20 is the parent strain of all deletion strain and served as control in all growth rescue experiments. This strain was grown in the same medium as the tested deletion strain.

Time interval measurements of the OD_{600} were taken using Infinite 200Pro plate reader (TECAN).

1.6.10 Growth conditions of TESEC strains

All TESEC strains (annotated as TS*XXX*) were always incubated at 37°C. Liquid overnight cultures were grown with agitation at 150 rpm (Multitron Standard, INFORS HT). Shaking conditions during experiments varied and are detailed in the sections below.

Solid media For each experiment, glycerol stocks were streaked on LB agar plates with 100 μg/mL ampicillin, 50 μg/mL kanamcyin. In addition:

- 20 μg/mL *meso*-DAP was added for TESEC strains with Δasd, $\Delta dapA$ or $\Delta dapB$ gene deletions.
- 100 μg/mL D-alanine was added for TESEC strains with Δalr $\Delta dadX$ gene deletions.

Liquid media Growth conditions of TESEC strains varied. For overnight cultivation, single colonies were inoculated in M9 medium without L-arabinose and with the supplement that the cell required in absence of essential target activity.

Carbon sources were alternated in experiments to analyze the impact on growth.

The cultivation medium generally contained:

Base M9 medium:

- 1x M9 Minimal Salts (Difco^{TM})
- 2 mM $MgSO_4$
- 0.1 mM $CaCl_2$
- 0.03% (w/v) vitamin B1

- x% of fructose or glucose (as mentioned in the results)
- 50 μg/mL kanamycin
- 100 μg/mL ampicillin

Depending on the TESEC strain, alternative amino acid mixes and supplements were used:

TESEC strains with Asd, DapA or DapB expression

Overnight cultivation - Base M9 medium with:

- 1 g/L Drop-out Mix Synthetic Minus Histidine, Methionine, Cysteine, Adenine Rich w/o Yeast Nitrogen Base (DO, Dropout, Powder) (USBiological Life Sciences)
- 20 μg/mL *meso*-DAP
- 40 μg/mL L-methionine

Growth assay - Base M9 medium with:

- 1 g/L Drop-out Mix Complete, Adenine Rich, w/o Yeast Nitrogen Base (Powder) (USBiological Life Sciences).
- 0-100 mM L-arabinose

TESEC strains with Alr expression

Overnight cultivation - Base M9 medium with:

- 1 g/L Drop-out Mix Complete, Adenine Rich, w/o Yeast Nitrogen Base (Powder) (USBiological Life Sciences)
- 100 μg/mL D-alanine

Growth assay - Base M9 medium with:

- 1 g/L Drop-out Mix Complete, Adenine Rich, w/o Yeast Nitrogen Base (Powder) (USBiological Life Sciences)
- 0-100 mM L-arabinose

TESEC strains with TrpD expression

Overnight cultivation - Base M9 medium with:

- 1 g/L Drop-out Mix Complete, Adenine Rich, w/o Yeast Nitrogen Base (Powder) (USBiological Life Sciences)

Growth assay - Base M9 medium with:

- 1 g/L Drop-out Mix Synthetic Minus Tryptophan w/o Yeast Nitrogen Base (DO, Dropout, Powder) (USBiological Life Sciences)
- 0-100 mM L-arabinose

TESEC strains with CysH expression

Overnight cultivation - Base M9 medium with:

- 1 g/L Drop-out Mix Complete, Adenine Rich, w/o Yeast Nitrogen Base (Powder) (USBiological Life Sciences)

Growth assay - Base M9 medium with:

- 1 g/L Drop-out Mix Synthetic Minus Histidine, Methionine, Cysteine, Adenine Rich w/o Yeast Nitrogen Base (DO, Dropout, Powder) (USBiological Life Sciences) for induction experiments.
- 0-100 mM L-arabinose

1.6.11 Preparation of L-arabinose dilution series

L-arabinose dilution series in experimental medium were routinely prepared in 96 Nunc^{TM} DeepWell^{TM} plates (ThermoFischer Scientific) using the TECAN Freedom Evo liquid handling robot.

1.6.12 Characterization of TESEC strains

Overnight cultures of TESEC strains were washed two times and diluted to the same initial assay OD_{600}. Triplicate 96-well plates were prepared with 90 μL bacterial cultures that were grown in M9 medium with dilution series of 0-100 mM L-arabinose. Microtiter plates were incubated without agitation at 37°C. The TECAN Freedom Evo robot routinely transferred the microtiter plates from the incubator to the Infinite 200Pro plate reader (TECAN) to take OD_{600} measurements. The background OD_{600} was subtracted from all OD_{600} measurements. The OD_{600} after 10 hours was used to compare the bacterial growth between the different TESEC strains. At this point in time, the control strain TS000 reached stationary phase, while a clear OD_{600} difference could be observed for strains with reduced growth.

1.6.13 Data analysis

All data analyses were conducted and visualized in MATLAB (Version R2017a, Mathworks).

1.6.14 Primary Drug screen with 60 compounds of the NIH/NCI Natural Product Set III

Overnight cultures of D20 (control strain with native alanine racemase expression) and TS008 (TESEC with Alr_{Mtb} expression) in M9 0.4% fructose medium were washed and diluted. Triplicate 96-well plates contained 90 μL bacterial cultures (with an initial OD_{600} of 0.05)

in M9 0.4% fructose and compounds of the NIH Natural Products Set III (Plate ID 13120880060) [169] at final compound concentrations of 0.1 mM in 1% DMSO. Positive control cultures contained 0.1 mM of D-cycloserine and negative control samples were prepared with solely 1% DMSO. The library compounds were tested on:

- TS008 induced with 0.1 μM L-arabinose for low Alr_{Mtb} expression
- TS008 induced with 100 mM L-arabinose for high Alr_{Mtb} expression
- D20 without L-arabinose

96-well plates were incubated at 37°C and time-course OD_{600} measurements (TECAN, Spark® multimode reader) were automatically taken by TECAN Freedom Evo liquid handling robot. A time-point at 10 hours was chosen to compare the effects of different drug molecules. At this time, cells without drug exposure reached the early stationary phase and it therefore provided a combined measure of the lag, exponential and bacterial yield. Because TS008 with low induction reached a lower final OD_{600} than TS008 with high induction, values were normalised by growth of reach strain in absence of the drug. Next, growth differences of TS008 with low and high Alr_{Mtb} expression were determined by subtracting the mean relative growth of TS008 in 100 mM L-arabinose by the mean relative growth of TS008 in 0.1 μM L-arabinose.

Chapter 2

Anti-mycobacterial activity of ACE-inhibitor benazepril identified with a rationally engineered tuberculosis surrogate *E. coli* strain

Nadine Bongaerts, Zainab Edoo, Ariel B Lindner, Edwin H Wintermute

manuscript in preparation

2.1 Abstract

Tuberculosis (TB) remains one of the deadliest diseases in the world. The threat of resistant *Mycobacterium tuberculosis* (*Mtb*) requires a continuous supply of new hit compounds to the global drug pipeline.

Because whole-cell screening for growth inhibitors of *Mtb* is complicated by the pathogen's slow growth and stringent biocontainment requirements, we developed a synthetic biology strategy for assaying *Mtb* drug targets in engineered *E. coli*. We delete an essential *E. coli* metabolic enzyme and replace it with an *Mtb*-derived functional ana-

log.

In this way, compounds that disrupt the activity of the *Mtb* target will inhibit the growth of the *E. coli* host. Our approach allows screening efforts to focus on specific drug targets while retaining key features of the live-cell context. As a proof-of-concept, we constructed an *E. coli* expression model for *Mtb* alanine racemase. We used feedback-controlled expression circuits to precisely set *Mtb* enzyme levels and achieve a 2-log fold drug sensitivity range for the positive control drug D-cycloserine. A drug screen of the Prestwick Chemical Library with 1,280 approved, off-patent drugs revealed benazepril as an alanine racemase-specific inhibitor, a result validated with biochemical assays and in live *Mtb*.

This is the first report of direct anti-tubercular activity of an Angiotensin-Converting Enzyme (ACE) inhibitor, and suggests a mechanism for clinical data associating the use of ACE inhibitors with reduced risk for active *Mtb* infection. More generally, the framework we describe could be applied for high throughput target-based screening of inhibitors for numerous enzymes from bacterial pathogens.

2.2 Introduction

Existing treatments for *Mycobacterium tuberculosis* infection are challenged by emerging antibiotic resistance [190]. New tuberculosis drugs are a global health priority [1], as are new technologies to support early stage drug discovery.

Of the recent generation of novel anti-*Mtb* compounds, most were

discovered through whole-cell screening [191]. In this technique, compound libraries are assayed for their ability to inhibit *Mtb* growth in batch culture [192]. The whole-cell approach has the advantage of testing directly for the desired antibacterial effect, but is insensitive to some potentially therapeutic activities. For example, *Mtb* may have vulnerabilities specific to host-pathogen interactions [193] or to the non-growing latent phenotype [194], that are not targetable in exponentially growing cells. A wider range of drug mechanisms can be explored with target-based screening. This approach uses custom biochemical assays, usually on purified cellular components, to test the disruption of specific molecular interactions with a proposed role in pathogenesis. As biological models of *Mtb* infection improve, many attractive targets have been proposed [195, 196, 197]. However, hit compounds from target-based screens have generally failed to advance through the drug development pipeline [198, 199, 200]. A common complaint is that this technique does not prioritize membrane permeation, which is essential for any targeted antibacterial to function in situ [201].

Targeted whole-cell assays make use of genetically modified *Mtb* to combine features of the target-based and whole-cell approaches. Increasing expression of a drug's target, for example, by introducing extra copies of the relevant gene, sometimes produces a resistant phenotype [202, 203]. Similarly, genetic knock-downs may cause drug sensitivity [204, 167]. Differential growth assays can then implicate the modified gene in the drug's mechanism of action. This technique has been used to confirm the target of known anti-tuberculars [205], and has been successful in identifying inhibitors against selected Mycobacterial pathways [206, 207].

While recent advances [208, 207] have let to the *Mtb* genetic modification toolbox, *E. coli* remains a preferred model organism because of its fast growth rate, the ease of genetic modification of and basic laboratory requirements. Here we explore the use of an engineered *E. coli* host to assay the *Mtb*-derived drug target in order to develop a scalable method.

By deleting an essential *E. coli* metabolic enzyme and expressing an *Mtb* analog, we create a Target Essential Surrogate *E. coli* (TESEC) strain whose viability depends on the heterologous mycobacterial gene.

Small molecule inhibitors that act on the expression, maturation, or function of the mycobacterial gene will, therefore, suppress *E. coli* growth. False positives and nuisance compounds that act non-specifically can be identified by counter-screening for growth inhibition of cells with native *E. coli* enzyme expression. Precise control of heterologous protein production is key to generate hypomorphic phenotypes (with reduced protein activity) and hypermorphic phenotypes (with increased target activity) that vary in their sensitivity to target-specific inhibitors.

To this end, we adopt an inducible, feedback-controlled gene expression circuit that provides a log-linear response to inducer over a wide dynamic range [2]. We show that this strategy substantially improves the sensitivity and precision of our assay. The final drug sensitivity was also influenced by dismantling the universal outer membrane protein TolC of *E. coli*'s tripartite efflux pumps involved in actively secreting proteins, toxins, and a variety of antimicrobial agents

[209, 210, 211]. Because drug efflux mechanisms generally differ between *Mycobacterium tuberculosis* and *E. coli* [165], we deleted *tolC* in our host to prevent that potential *Mtb*-inhibitors that are exported by *E. coli*.

As a proof-of-concept, we constructed a TESEC model for *Mtb* alanine racemase, a well-known and attractive drug target involved in the production of peptidoglycan pre-cursor D-alanine [212]. Although *Mtb* and *E. coli* are evolutionary distant relatives [213], their peptidoglycan structures resemble each other [116] (Section 1.2.2). Several established inhibitors of alanine racemase have been discovered [214], including the second-line TB-drug D-cycloserine, which is therapeutically less preferred due to off-target affinities and systemic toxicity [141]. The absence of human alanine racemase homologs [141] and the evolved *Mtb* resistance to D-cycloserine [215, 6] further motivates efforts to target this enzyme by other means.

A high-throughput screen with the Prestwick Chemical Library [3] of approved, off-patent drugs revealed that benazepril was able to suppress the growth of *E. coli* isomorphs with hypomorphic *Mtb* alanine racemase but not hypermorphic *Mtb* alanine racemase expression nor the wild type *E. coli* control. Subsequent validation experiments confirmed that benazepril was an inhibitor of *Mycobacterium smegmatis* mc^2155 and *Mycobacterium tuberculosis* H37Rv in whole-cell growth assays. Benazepril was also shown to inhibit *Mtb* alanine racemase activity in vitro.

This work is a new horizon for target-based whole-cell screening and shows that engineered *E. coli* host can allow for the finding of relevant tuberculosis hits in drug libraries that have been screened exhaustively

on tuberculosis.

Moreover, we designed our engineered *E. coli* host and plasmid-based gene expression systems for maximum versatility and re-use. Many other enzymes, from *Mtb* and other pathogens, could potentially be screened within this modular framework.

2.3 Results

2.3.1 Development Target Essential Surrogate *E. coli* strain with controlled *Mtb* alanine racemase expression for use in drug discovery

A target essential surrogate *E. coli* strain was constructed with growth-dependency for *Mtb* alanine racemase expression (Figure 2.1). The background strain lacks its native alanine racemase genes *dadX* and *alr* and is therefore unable to grow in the absence of D-alanine (Appendix Figure A.1).

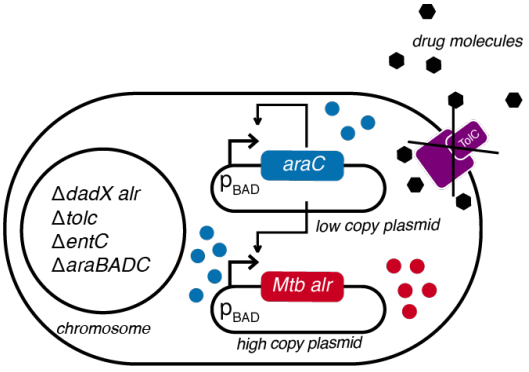

Figure 2.1: Schematic representation of Alr$_{Mtb}$ Target Essential *E. coli* strain TS144. TESEC strain TS144 carried chromosomal deletions in *dadX, alr, tolC, entC, and araBADC*. A two-plasmid system was used for log-linear *Mtb* alanine racemase expression based on [2]. This circuit consist of a low-copy plasmid (LCP) and a high-copy plasmid (HCP). The LCP includes a positive feedback loop to generate the transcription factor AraC that activates transcription on the HCP to express *Mtb* Alr. Deletion in *tolC* renders deficient TolC-related efflux pumps and prevents the export of a variety of drug molecules. A deletion in *entC* prevents the accumulation of sidophores in the periplasm when TolC is absent [216].

In addition, *tolC* was deleted from the chromosome of the drug

screening strain to prevent the potential export of library compounds through TolC-related multidrug efflux pumps [162]. Previous research demonstrated significant decreases in the MIC (minimal inhibitory concentration) for a $\Delta tolC$ strain compared to a $tolC^+$ strain exposed to diverse antibiotics [217, 218, 211].

Although TolC is not essential for *E. coli* survival, Vega *et al.* showed that a $\Delta tolC$ *E. coli* can have abnormal cell shapes and a reduced growth rate when cultured in minimal, iron lacking medium [216]. In strains lacking TolC, enterobactin cannot be transported to extracellular environment and is consequently accumulated in the *E. coli* periplasm [216]. Vega *et al.* restored the wild type phenotype by supplementing minimal medium with extra iron or by disrupting the iron-specific sidophore enterobactin production via a deletion in *entC* [216].

We verified whether the $\Delta tolC$ deletion in our background strain displayed an abnormal phenotype in the media that was intended for drug screening. When $\Delta tolC$ deletion strain K028 was cultivated in M9 fructose medium, an extended lag phase (Figure 2.2) and an elongated cell morphology in the stationary phase (Figure 2.3) was observed. To prevent unnecessary membrane stress due to enterobactin buildup, we included the same deletion in *entC*.
Although the additional deletion in *entC* reduced the prolonged lag phase of the $\Delta tolC$ strain, it was not sufficient to fully restore wild type growth 2.2). Regardless, microscope images revealed that the normal cell morphology was restored for the $\Delta tolC$ $\Delta entC$ strain K031 (Figure 2.3).

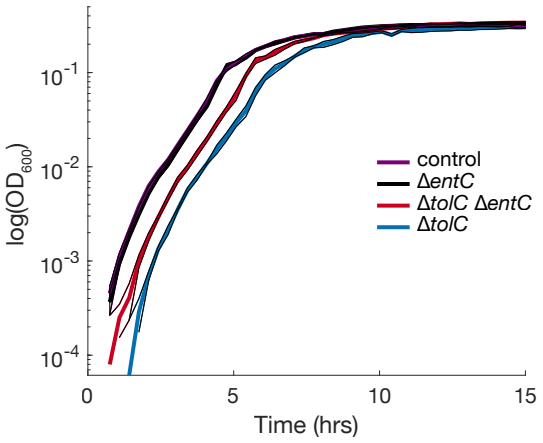

Figure 2.2: Increased lag phase for Δ*tolC* deletion strain.
K028 (Δ*tolC*), K031 (Δ*tolC* Δ*entC*), K123 (Δ*entC*), control strain D20 (parent strain with TolC and EntC expression) were grown in M9 medium for 15 hours at 37°C. Results are the mean values and standard deviations from three independent experiments.

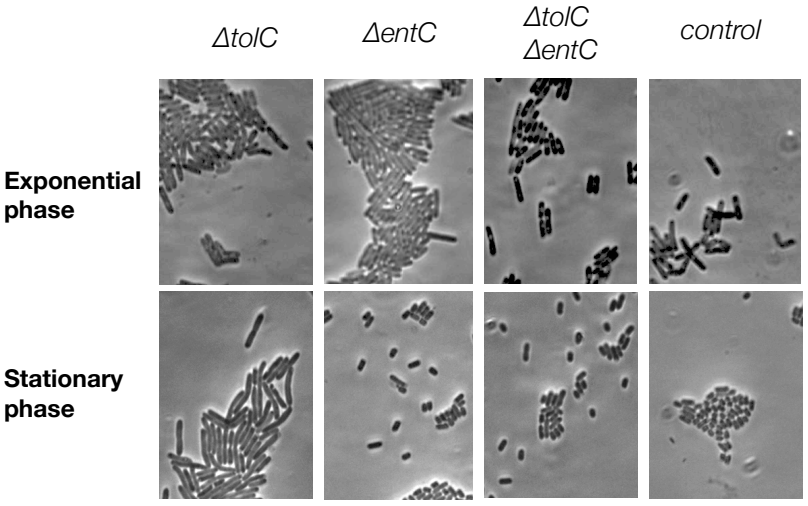

Figure 2.3: Elongated cell morphologies for stationary phase Δ*tolC* cells in minimal medium.
Phase-contrast microscopy images (63x magnification) of strains K028 (Δ*tolC*), K123 (Δ*entC*), K031(Δ*tolC* Δ*entC*) and the control strain D20 (parent strain with TolC and EntC expression) grown in M9 0.4% fructose medium at 37°C, 150 rpm. **A. Exponential phase cells.** Overnight cell cultures were diluted 1:100 and grown until OD_{600} = 0.3-0.5. **B. Stationary phase cells.** Bacterial cells after 20 hours of cultivation. A second deletion in *entC* restored normal morphology to a stationary phase Δ*tolC* mutant grown in M9 medium.

Finally, an arabinose-inducible, two-plasmid expression circuit based on Daniel *et al.* [2] was introduced to allow for differential expression of the target enzyme *Mtb* alanine racemase.

The system consists of a low-copy (∼5) plasmid expressing AraC and a high-copy (∼15-20) plasmid expressing *Mtb* alanine racemase (Figure 2.1). Although induction over 0-100 mM L-arabinose concentrations resulted in comparable bacterial growth (in absence of inhibitor) (Figure 2.4.A), the drug sensitivity for Alr inhibitor D-cycloserine, expressed as half-maximal inhibitory concentration (IC_{50}), varried over a ∼2-log range (Figure 2.4.B). This indicated that the genetic circuit allowed us to alter the cellular alanine racemase concentrations while maintaining *E. coli* growth.

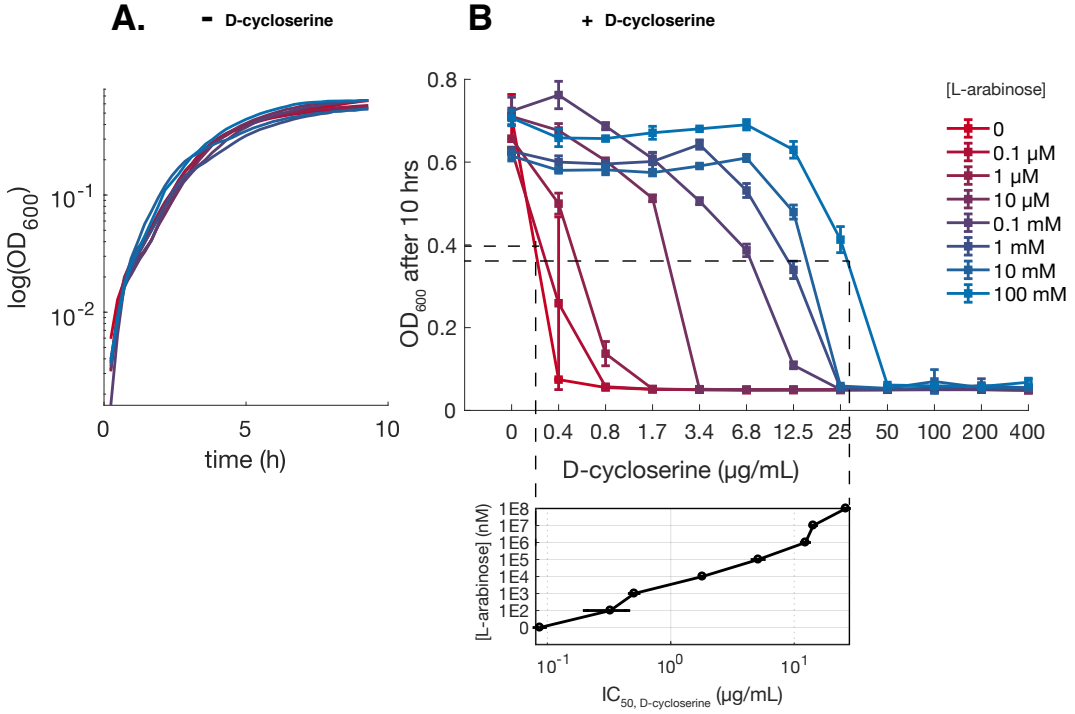

Figure 2.4: TESEC-Alr$_{Mtb}$ growth in absence (A) and presence (B) of D-cycloserine. TESEC strain TS144 with Alr$_{Mtb}$ expression was grown in M9 fructose medium with L-arabinose (0-100 mM) and D-cycloserine series (0-400 μg/mL). **A. Growth curves in absence of D-cycloserine. B. Induction of TS144 renders a 2-log drug sensitivity range for D-cycloserine.** OD_{600} measurements at 10 hours of incubation were used to plot growth inhibition curves. Log-linear variations in inducer levels resulted in log-changes of the D-cycloserine sensitivity of the strain, allowing for an IC_{50} range between 0.09±6.5E-3 μg/mL to 26.21±0.97 μg/mL D-cycloserine. Error bars represent the standard deviations of three technical replicates.

2.3.2 Screening of the Prestwick Chemical Library

We screened TESEC strain TS144 with controlled *Mtb* alanine racemase expression against the Prestwick Chemical library with 1,280 FDA approved drug molecules [3]. To find Alr_{Mtb}-specific inhibitors, TS144 was induced with 0.1 μM and 10 mM L-arabinose for respectively low and high Alr_{Mtb} expression. In addition, we included negative control strain TS163 with native alanine racemase expression ($DadX_{Ec}$ and Alr_{Ec}) and a high copy plasmid with *gfp* instead of alr_{Mtb}. The final screening concentration of each chemical compound was 0.1 mM and each screening condition was tested in triplicate.

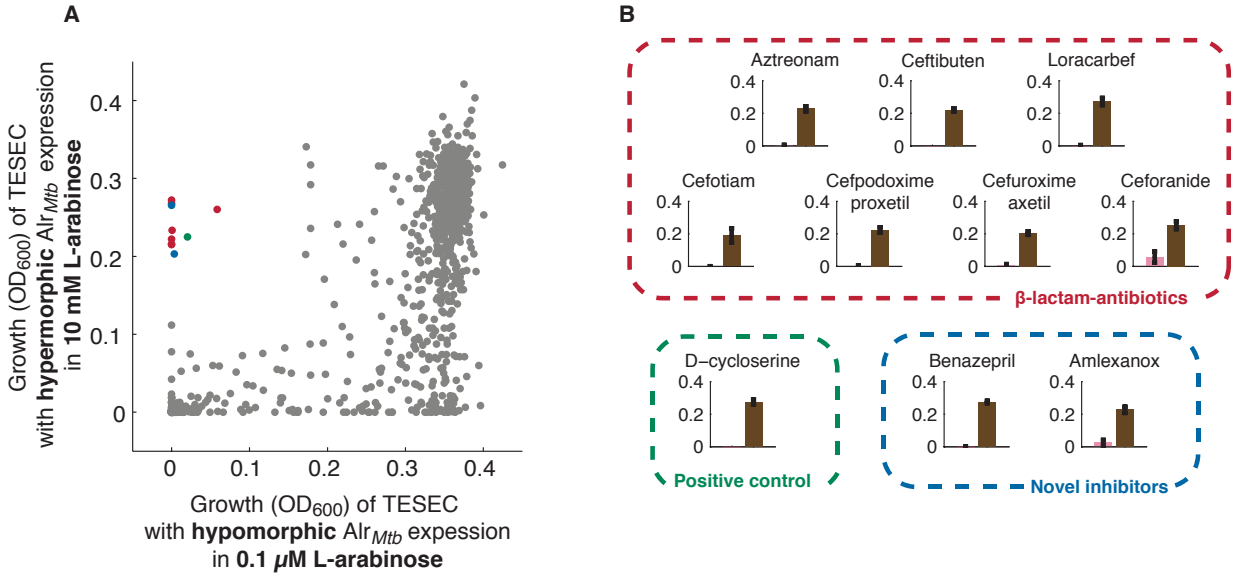

Figure 2.5: Prestwick Library Screening Results.
100 μL bacterial cultures of TESEC TS144 in M9 0.4% fructose medium were exposed to Prestwick chemical library compounds [3] at the standard compound concentration of 0.1 mM. OD_{600} measurements of triplicate 96-well microtiter plates were taken after 10 hours of incubation at 37°C. **A.** Hit identification through OD_{600} comparisons of TS144 isomorphs with hypermorphic and hypomorphic Alr_{Mtb} expression. **B.** Hit compound signatures. Brown and pink bars represent the OD values after 10 hours in respectively 'high' induction at 10 mM L-arabinose and 'low' induction at 0.1 μM L-arabinose.

For ten compounds, growth was significantly impacted under low induction (OD_{600}<0.1) compared to high induction (OD_{600}>0.2) (Figure

2.5). As expected, the positive control D-cycloserine was one of the ten 'hits'. Next, we used the strictly standardized difference metric (SSMD) [219] to determine the significance of the growth inhibition. Thresh hold values below -5 indicate that the effect is 'extremely strong' [219]. All ten compounds had SSMD-scores below -8 for the low induction condition. Hence, the growth reduction was significant. The growth of the wild type control was insignificantly reduced compared to the TESEC hypomorph.

Interestingly, the majority of hit compounds (7/10) were classified as cephalosporins or monobactams. These are classes of β-lactams that target penicillin-binding proteins involved in, amongst others, the cross-linking and binding of d-amino acids in peptidoglycan [212]. While resistance to newer generations of β-lactams such as cefuroxime axetil, loracarbef, ceforanide (second-generation cephalosporins) and ceftibuten, cefotiam, cefpodoxime proxetil (third-generation cephalosporins), is typically associated with the presence of extended-spectrum beta-lactamase enzymes [220], our screening results hinted at the existence of a β-lactam resistance mechanism related to alanine racemase over-expression.

Importantly, two hits had no previously reported antibacterial properties: benazepril HCl and amlexanox.

2.3.3 Validation of primary hits

To validate the findings from the drug screen, I assessed the relation between Alr_{Mtb} expression and the primary hit compounds by growing TS144 in a two-dimensional gradient of ten-fold L-arabinose and two-

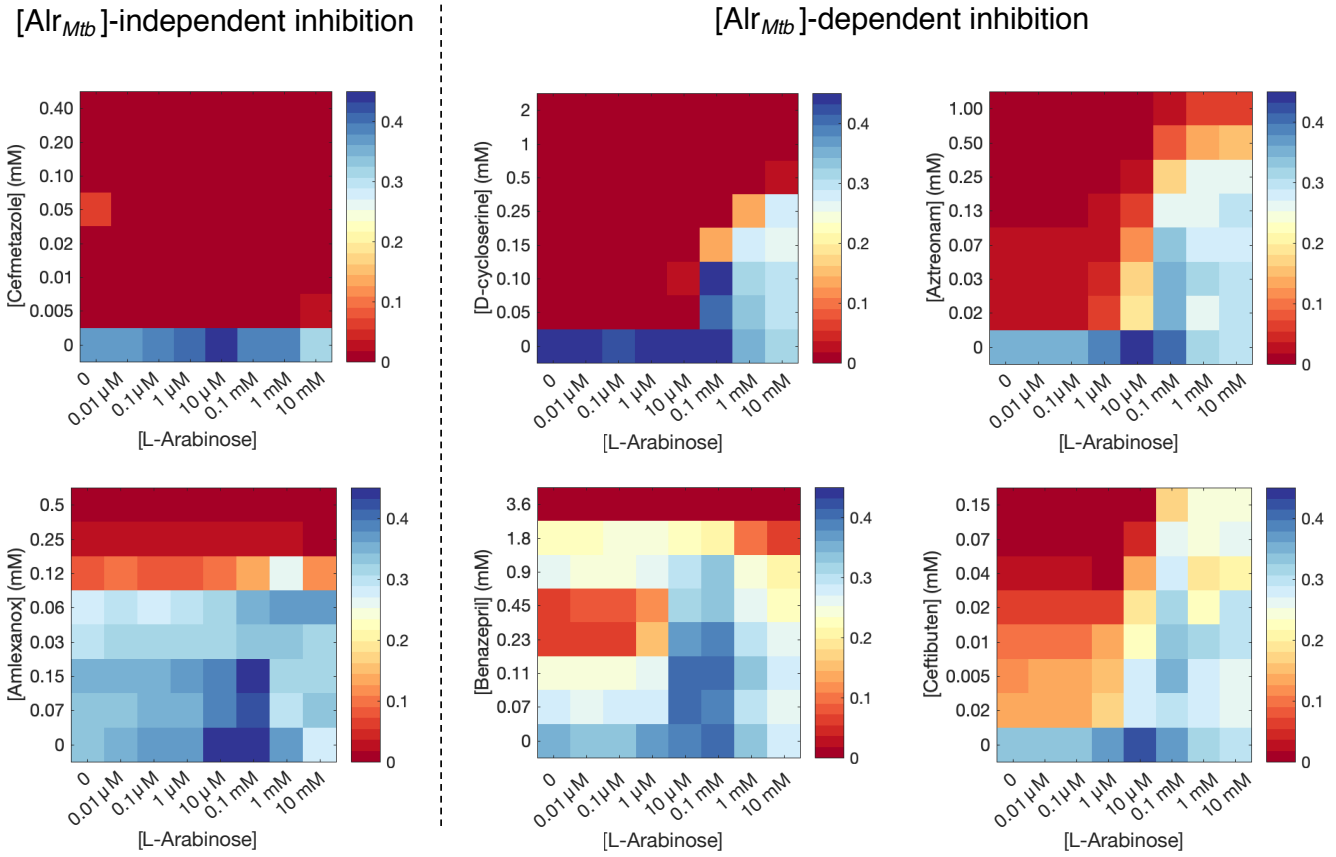

Figure 2.6: Chemical-genetic profiles of hit compounds with [Alr$_{Mtb}$]-independent (left) and -dependent (right) growth inhibition.

TS144 was grown in 384-well plates with 50 μL 0.4% Fructose medium with 0 – 10 mM L-arabinose concentrations and a range of test compound concentrations. OD_{600} measurements were taken after 10 hours of growth at 37°C, 600 rpm. Results represent the mean values of three technical replicates.

fold drug concentrations to generate more detailed chemical-genetic profiles (Figure 2.6). Growth inhibition by negative control cefmetazole (a Prestwick library compound that was not identified as hit), and hit compound amlexanox was independent of Alr$_{Mtb}$ expression levels. Increased L-arabinose concentrations could not increase resistance for those compounds. In contrast, drug resistance with elevated Alr$_{Mtb}$ levels was confirmed for D-cycloserine (positive control), and for β-lactam hit compounds aztreonam and ceftibuten. Interestingly, increased expression levels of *Mtb* alanine racemase also provided protection to the

angiotensin-converting enzyme (ACE) inhibitor benazepril.

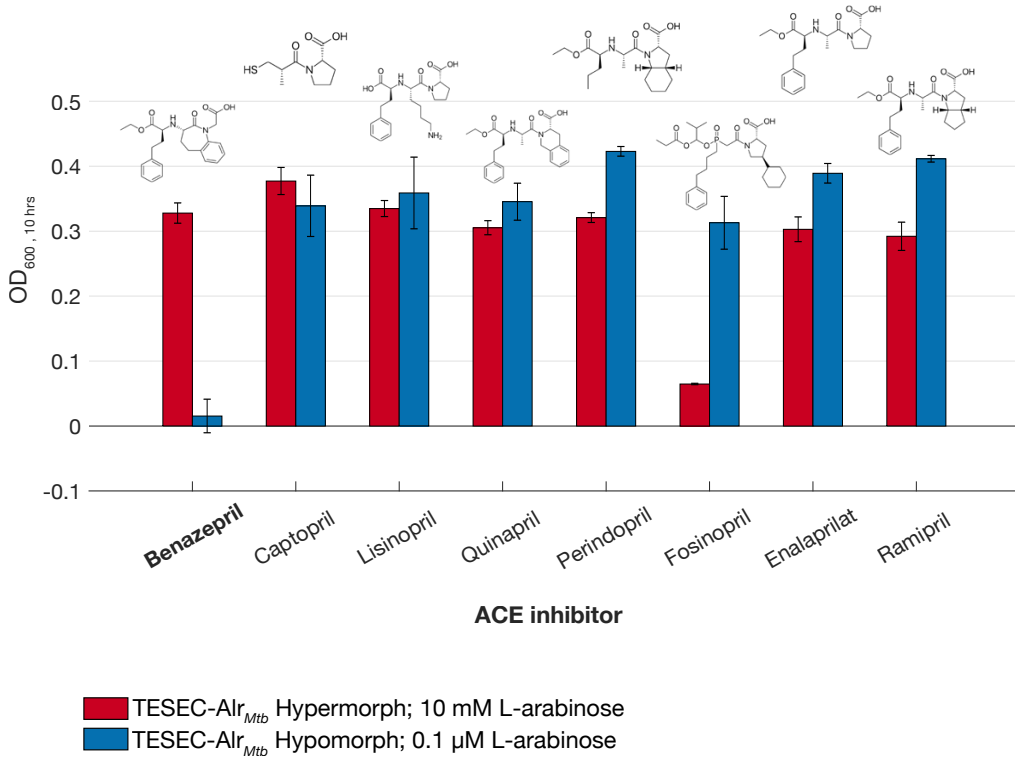

Figure 2.7: Prestwick Library screening results for ACE-inhibitors.

Since β-lactams have been exhaustively tested on *M. tuberculosis* due to their effectiveness in combination treatments with clavulanic acid and other tuberculosis drugs [221, 222], we decided to focus our attention on the unexpected finding that ACE-inhibitor benazepril seemed to display Alr_{Mtb}-level dependent antibacterial properties. Interestingly, other ACE-inhibitors in the Prestwick Library, with similar chemical structures, were not identified as primary hits (Figure 2.7).

2.3.4 Growth rescue of TESEC-Alr_{Mtb} by D-alanine supplementation

Next, we tested whether the supplementation of D-alanine to the TS144 culture medium could rescue bacterial growth in the presence of MIC concentrations of either benazepril or D-cycloserine (control)

(Figure 2.8). Simple addition of the d-amino acid compound should allow for growth rescue if D-alanine is the only essential compound that the TESEC strain is unable to sufficiently produce in the presence of benazepril. Similar to higher induction of Alr_{Mtb} expression, the supplementation of D-alanine allowed TS144 to grow with benazepril (Figure 2.8). This pointed again at the possibility that D-alanine production was inhibited by benazepril.

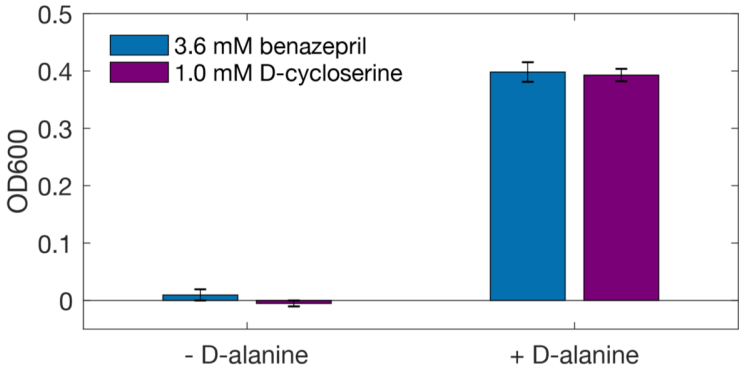

Figure 2.8: D-alanine rescues TS144 growth inhibition by benazepril.
Strain TS144 with Alr_{Mtb} expression was grown in fructose M9 medium with 0.1 mM L-arabinose, 3.6 mM benazepril or 1 mM D-cycloserine (positive control), and with or without 100 μg/mL D-alanine. D-alanine supplementation to the cultivation medium allowed for growth rescue in the presence of either drugs. The bar graph represents the OD_{600} of three technical replicates that was determined after 10 hours of growth at 37°C, 600 rpm. Results are representative for two independent assays with different medium conditions.

2.3.5 Synergistic effect of DCS with benazepril

To obtain additional information about benazepril's possible mechanism of action, I conducted an experiment to study the combined inhibitory effect of benazepril and D-cycloserine on TS144. Drug synergy happens when the effect of one drug is enhanced in combination with another drug. This can occur when two drugs act subsequently in the same pathway [223]. D-cycloserine inhibits D-alanine racemase, as well as D-ala-D-ala ligase (Ddl). Ddl is responsible for the catalysis of two D-alanine molecules into the dipeptide D-alanyl-D-alanine. Prosser *et al.* found that D-cycloserine binding is stronger for D-ala-D-

ala ligase [224] compared to alanine racemase. Therefore, synergistic inhibition between D-cycloserine and benazepril could indicate that benazepril inhibits alanine racemase, or, other enzymes involved in peptidoglycan synthesis.
To investigate this, I cultivated TESEC strain TS144 at a single L-arabinose concentration (75 μM) and exposed the bacteria to a checkerboard of D-cycloserine and benazepril series (Figure 2.9.A). TS164 with native *E. coli* alanine racemase expression and an empty 'high copy' vector served as control (Figure 2.9.B). D-cycloserine inhibition of the control TS164 was slightly stronger in high benazepril concentrations, but benazepril only had no effect on growth. Oppositely, a strong combinatorial inhibition effect was observed for TS144.

To evaluate whether the combined inhibitory effect of benazepril and D-cycloserine on TS144 was synergistic, I used a response envelope analysis (REA). This method combines the Bliss Independence [226] model and the generalized Loewe Additivity model [227, 228] to specify whether the drug combination is synergistic or antagonistic without the need for prior knowledge of the inhibition mechanisms [225]. The corresponding REA bioinformatics tool allowed me to calculate a Synergy Index (SI) and Antagonism Index (AI) of respectively 0.056 and 0.00042. Because SI was higher than AI, benazepril and D-cycloserine acted synergistally on TS144. Hence, this experiment pointed again in the direction that benazepril affected alanine racemase activity, or, at least that peptidoglycan production was affected.

2.3.6 Possible drug efflux of benazepril via TolC-related efflux

I also verified whether background deletions *tolC* and *entC* were important for the discovery of benazepril. To examine this, I exposed

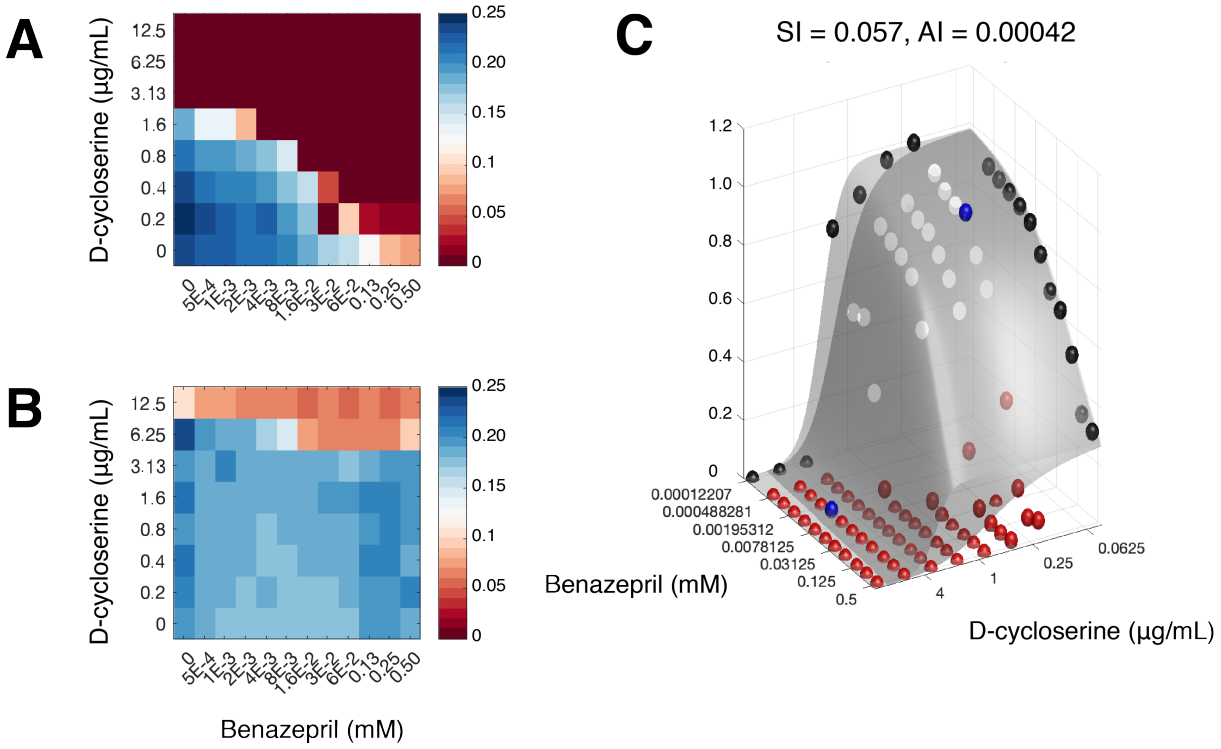

Figure 2.9: Benazepril-D-cycloserine drug synergy detected for TESEC-Alr$_{Mtb}$ strain TS144.
A. TS144 and control strain TS164 with native alanine racemase expression and an empty plasmid **B.** were grown in M9 fructose medium supplemented with 0.75 μM L-arabinose. Colors represent the mean OD_{600} values of three technical replicates after 10 hours of incubation at 37°C. **C. Synergistic effect of D-cycloserine-benazepril determined by Response Envelope Analysis (REA)** [225]. REA was performed with data in panel A. Drug synergy between D-cycloserine and benazepril was confirmed as Synergy Index (SI) > Antagonism Index (AI). The response envelope (grey) results from individual drug inhibition responses. Each point defines a synergistic, antagostic or additive effect represented by the respective colors red, blue or white. The black dots are the drug-response curves of each individual drug. SI and AI, were respectively SI = 0.056 and AI = 0.00042.

$\Delta tolC\Delta entC$ and $tolC^{+}entC^{+}$ TESEC strains (respectively TS008 and TS144) with identical Alr$_{Mtb}$ expression to 0.1 mM concentrations of benazepril, its derivative benazeprilat, D-cycloserine (positive control) and DMSO (negative control).

In the human body, benazepril is converted into the active angiotensin-converting enzyme (ACE) inhibitor benazeprilat through the enzymatic hydrolysis of its ethyl ester [229]. After several hours of be-

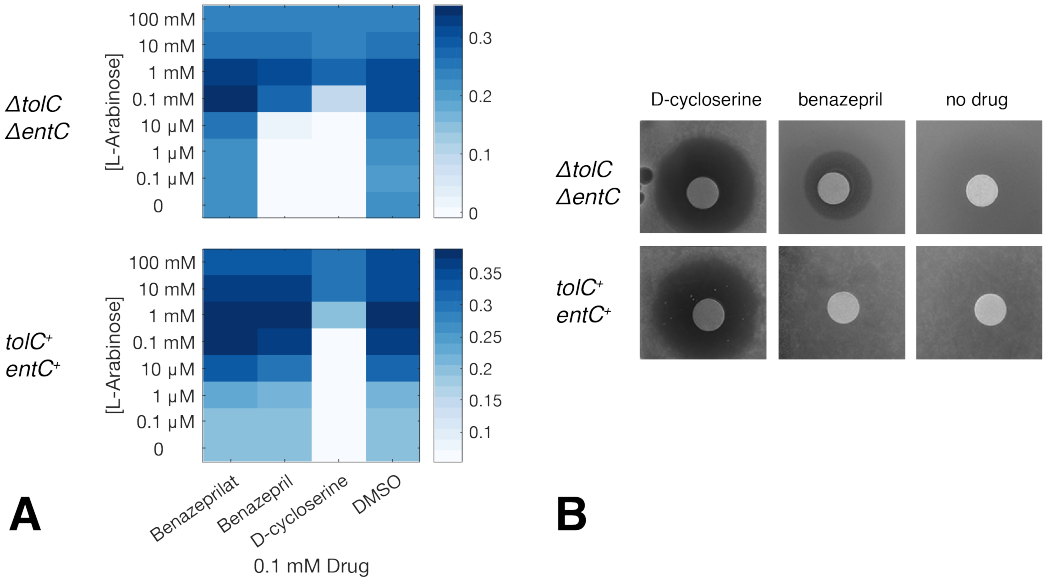

Figure 2.10: Benazepril does not affect TESEC-Alr$_{Mtb}$ with TolC & EntC expression **A.** Standard compound concentrations of 0.1 mM in 0.2% DMSO were exposed to strains TS144 (top) an TS008 (bottom). Colors represent the OD_{600} values after 10 hours of growth at 37°C. **B.** Disk diffusion assay to verify the benazepril susceptibility of strains TS144 (top row) and TS008 (bottom row). Bacteria were grown in M9 fructose soft agar medium with 0.1 mM L-arabinose. Disks contained 10 μL of 500 mM benazepril in DMSO, 10 μL of 245 mM of D-cycloserine (positive control) or 10 μL DMSO ('no drug', negative control). Photos were taken after 24 hours of incubation at 37°C.

nazepril administration, the patients' blood contains a higher plasma concentration of benazeprilat than the prodrug [229]. Given this, I included benazeprilat in the experiment to verify whether its activity on the TESEC strains.

As expected, D-cycloserine had a comparable inhibitory effect on both TS008 and TS144 (Figure 2.10.A). However, benazepril had no effect on TS008 that lacked chromosomal deletions in *tolC* and *entC*. A result that was also confirmed with a disk diffusion assay (Figure 2.10.A&B). This suggested that TolC-mediated efflux could be responsible for the lack of benazepril activity on TS008. An additional validation experiment with a $\Delta tolC$ $entC^+$ strain and $\Delta entC$ $tolC^+$ strain should still be done to confirm with certainty that only the lack of

TolC was responsible for this observation. Interestingy, benazeprilat had no effect on either TESEC strain (Figure 2.10.A).

2.3.7 Alr_{Mtb} inhibition by benazepril.

Benazepril's activity was further assayed on Alr_{Mtb} from cell lysate to determine the specificity for alanine racemase activity. I adopted a biochemical assay in which a coupled reaction is detected by fluorescence measurements. First, D-alanine is converted into L-alanine by *Mtb* alanine racemase, followed by the catalysis of L-alanine and NAD^+ by *B. subtilis* L-alanine dehydrogenase into NADH and pyruvate [230].
Cell lysate of TESEC strain without native alanine racemase expression and with an expression vector that expressed *mCherry* instead of Alr_{Mtb} was used as control lysate.

Reactions were performed in 0-20 mM substrate concentrations of D-alanine and 0-9 mM concentrations of benazepril. The initial velocities of the reactions were determined using linear regression of fluorescent measurements in time. Subsequently, those values were used to construct Michaelis-Menten, Lineweaver-Burk and a Quotient velocity plots [231].

No alanine racemase activity was detected for cell lysate without alanine racemase, as increasing substrate concentrations of D-alanine did not change the initial velocity of the reaction (data not shown). On the other hand, alanine racemase activity was reduced for increasing concentration of benazepril (Figure 2.12).

Linear regression analysis from Lineweaver-Burk plots (Figure 2.12.B)

was used to determine kinetic parameters K_m and maximum relative velocities in the absence an presence of benazepril (Table 2.11). Moreover, I applied a graphical method described by Yoshino and Murakami [231] to obtain additional information about the mode of alanine racemase inhibition by benazepril.

Benazepril (mM)	Relative max. velocity (v_{max}/V_{max})	Km (mM)
0	1	0.82
2.25	0.53	0.63
4.5	0.26	0.57
9	0.12	0.53

Figure 2.11: *Mtb* alanine racemase activity in the presence of 0-9 mM benazepril
Values were estimated by determining the slope and intercept of the Lineweaver-Burk plot (Figure 2.12.B)

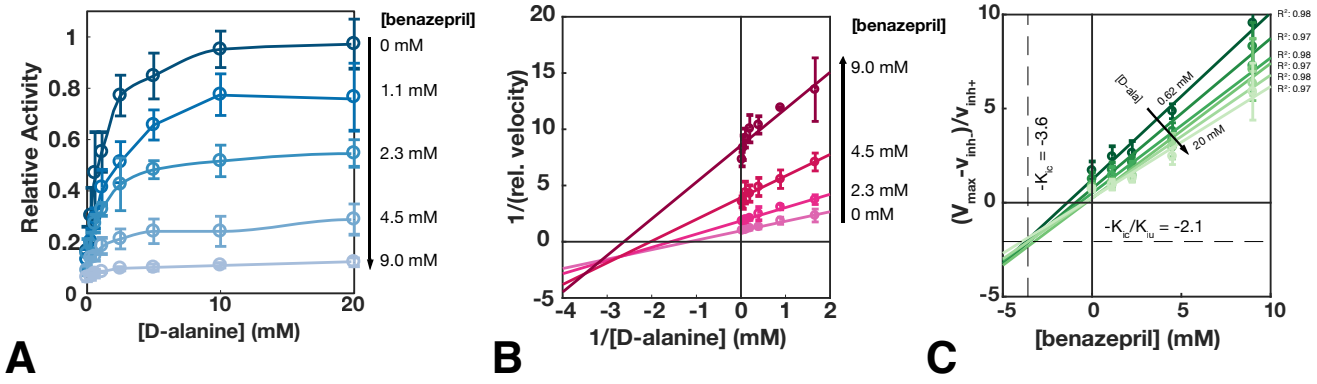

Figure 2.12: Benazepril inhibition of Alr_{Mtb}
96-well plates containing cell lysate with *Mtb* Alr in 110 mM $NaHCO_3$ buffer (pH 10.5) containing 10 mM β-NAD, 2.5E-2 units of *B. subtilis* L-alanine dehydrogenase in a 2D gradient of D-alanine (0 - 40 mM) and benazepril (0 – 9 mM) series were incubated at 30°C. The production of NADH was determined by absorbance measurements at 340 nm. **A. Michaelis-Menten plot.** Blue lines represent measured racemase activity in cell lysate expressing *Mtb* Alr. **C. Lineweaver-Burk plot. B. Quotient velocity plot** suggested mixed uncompetitive inhibition by benazepril. K_{ic} = 3.6 mM and K_{iu} = 1.7 mM. Data points and error bars represent the mean values and standard deviations of 4 independent experiments.

The values presented here should be interpreted cautiously, as as-

says with pure alanine racemase should be performed to accurately determine the kinetic parameters. These experiments have not yet been performed due to severe delays in the delivery of L-alanine dehydrogenase.

Still, based on the current data, it was determined that both V_{max} and K_m were reduced for increasing concentrations of the benazepril (Table 2.11).

In addition, inhibition constants for competitive inhibition (K_{ic}) and uncompetitive inhibition (K_{iu}) were estimated (Figure 2.12.C) as respectively 3.6 mM and 1.7 mM. Since $K_{ic} > K_{iu}$, this could be an indication that benazepril might act as a mixed uncompetitive inhibiter of *Mtb* alanine racemase.

2.3.8 Whole-cell activity of benazepril on *M. smegmatis* and *M. tuberculosis*

To further investigate benazepril's potential as antimycobacterial agent, we examined benazepril's activity on *Mycobacterium smegmatis* mc^2155 and *Mycobacterium tuberculosis* H37Rv. In parallel, I used chemical-genetic profiling of TESEC-strain TS123 with *M. smegmatis* mc^2155 alanine racemase expression (Figure 2.13.A) to verify $Alr_{M.\ smegmatis}$ level-dependent inhibition by benazepril. I identified a comparable inhibition profile for TS123 as for strain TS144 with Alr_{Mtb} expression (Figure 2.6). However, unlike Alr_{Mtb} expression in TS144, $Alr_{M.\ smegmatis}$ expression in high L-arabinose concentrations (10 and 100 mM) was detrimental for growth. This illustrates the importance of using a system that allows for fine-tuned expression.

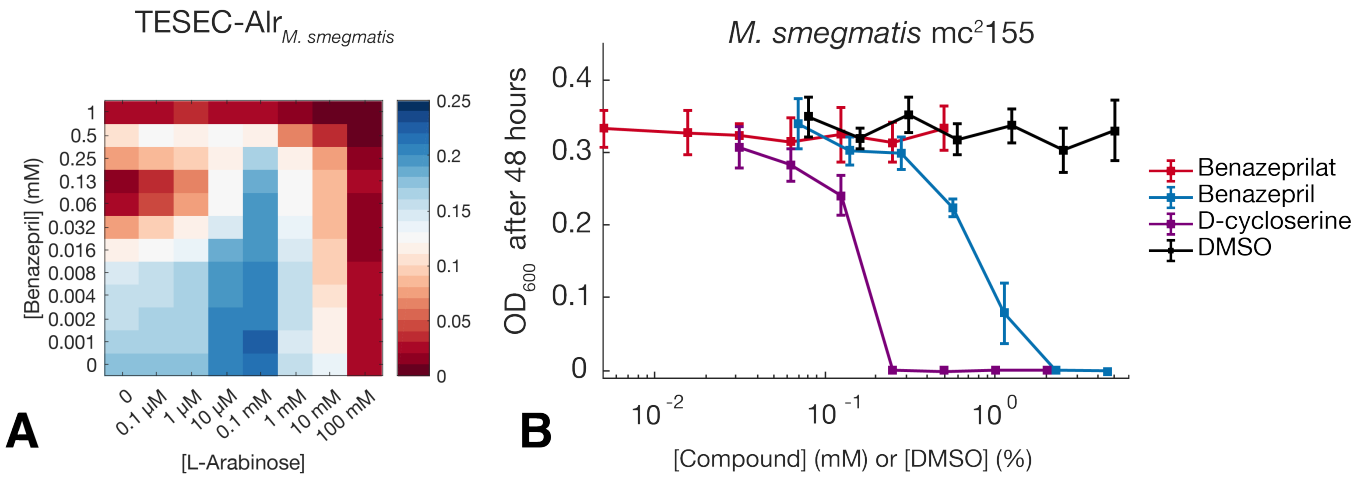

Figure 2.13: Benazepril inhibits both TESEC-Alr$_{M.\ smegmatis}$ and wild type *M. smegmatis*.
A. Alr$_{M.\ smegmatis}$-dose dependent benazepril inhibition of TESEC strain TS123. 90 μL TS123 cultures were grown in M9 0.4% fructose medium with 0-100 mM L-arabinose and 0-1 mM benazepril concentrations. Colors represent the mean OD_{600} values of three technical replicates after 10 hours of incubation at 37°C. **B. *Mycobacterium smegmatis* mc^{2}155 growth inhibition by benazepril.** 100 μL *M. smegmatis* cultures in 96-well plates were incubated for 48 hours (37°C; 750 rpm) in 7H9 medium with dilutions series of benazepril (0-9 mM), benazeprilat (0-0.5 mM), DMSO (0-5%) as negative control and D-cycloserine (0-2 mM) as positive control. Data points are the means and standard deviations of 3 independent experiments.

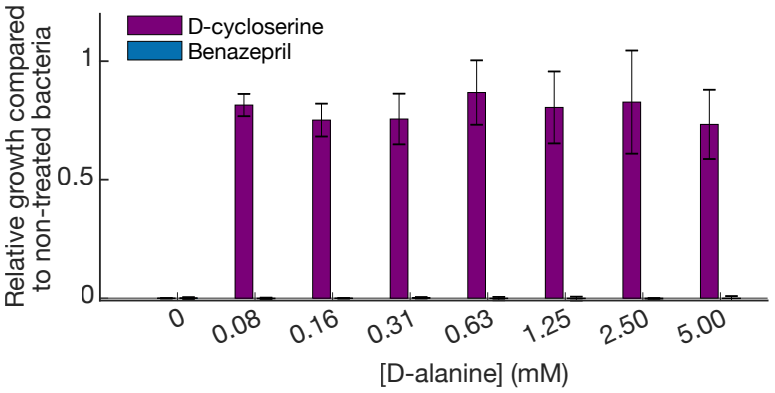

Figure 2.14: *M. smegmatis* growth inhibition by benazepril is not rescued by D-alanine supplementation.
Serial dilutions of 0–5 mM D-alanine were added to *M. smegmatis* cultures growing in either 0.5 mM D-cycloserine, 4.5 mM benazepril or 1% DMSO. Bar graphs represent the relative bacterial growth in the presence of D-cycloserine or benazepril and varying D-alanine concentrations compared to bacterial growth in the presence of 1% DMSO. After 48 hours of incubation at 37°C, OD_{600} measurements were taken to verify the bacterial yield. Bar graphs represent the mean values and standard deviations of 3 independent experiments.

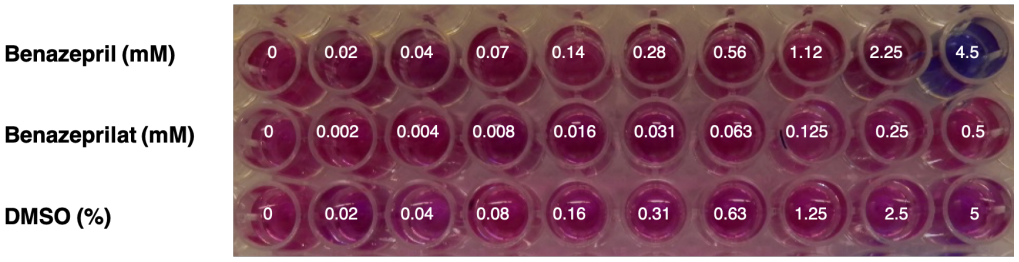

Figure 2.15: *M. tuberculosis* H37Rv growth inhibition by benazepril.
96-well plate image of *M. tuberculosis* H37Rv cultures in the presence of benazepril, benazeprilat and DMSO taken after 8-10 days of incubation at 37°C. Resazurin (0.0025%, w/v) was added one day before the image was taken. Bacterial growth was inhibited when resazurin color changed from blue to pink.

Compound	Dilution range	MIC range	MAD
Benazepril	0 – 4.5 mM	2.25 - 4.5 mM	2.25 mM
Benazeprilat	0 – 0.5 mM	>0.5 mM	-
DMSO	0 – 5%	>5%	-
Isoniazid	0 – 4 µg/mL	0.031 – 0.062 µg/mL	0.0146 µg/mL
D-cycloserine	0 – 2 mM	0.125 mM	0 mM

Table 2.1: Susceptibility assay of *Mycobacterium tuberculosis* H37Rv.
Dilution range, MIC range, absolute deviation (MAD) for each tested compound. Results are based on five biological replicates.

Compound	Required [D-alanine] to rescue *Mtb* H37Rv growth
2.25 mM Benazepril	no benazepril inhibition
4.5 mM Benazepril	no growth in the presence of 0-5 mM D-alanine
0.125 mM D-cycloserine	≥ 0.02 mM
0.25 mM D-cycloserine	≥ 0.04 mM

Table 2.2: *M. tuberculosis* H37Rv growth at MIC concentrations of benazepril and D-cycloserine and a range of 0–5 mM D-alanine concentrations.
Results represent the median values of 5 biological replicates.

Mycobacterium smegmatis mc^{2}155 and *Mycobacterium tuberculosis* H37Rv were grown in 7H9 medium with serial dilutions of benazepril, its derivative benazeprilat and the positive control D-cycloserine (Figure 2.13, 2.15, Table 2.1). The benazepril MIC for *M. smegmatis*

mc^2155 was determined as >2.25 mM based on optical density measurements (Figure 2.13) and viability screening on agar plates (Figure B.1). The MIC of benezapril is therefore approximately 10 times higher than D-cycloserine under the same growth conditions. *M. tuberculosis* had a comparable benazepril MIC range of 2.25–4.5 mM based on visible growth and cell viability using the fluorometric resazurin reduction method (Table 2.1 & Figure 2.15), while whole-cell exposure to benazeprilat had no detectable inhibitory effect on either *M. smegmatis* nor *M. tuberculosis* (Figures 2.13, 2.15 & Table 2.1).

As expected, supplementation of D-alanine rescued *M. smegmatis* and *M. tuberculosis* growth in the presence of D-cycloserine. However, unlike TS144 (Figure 2.8), D-alanine addition had no effect no effect on *M. smegmatis* nor *M. tuberculosis* cultures grown in 4.5 mM benazepril (Figure 2.14 & Table 2.1). This revealed that benazepril could have additional targets beyond alanine racemase.

2.4 Discussion

This study demonstrates the opportunity of using rationally engineered *Mtb*-target-dependent *E. coli* strains as inexpensive, scalable, and easy-to-adopt alternatives for the discovery tool of tuberculosis inhibitors. Our strain design used the relationship between target expression level and target-specific drug resistance as a means to identify hit compounds [143, 232]. However, it should be mentioned that target overexpression or depletion could also increase, decrease, or have no effect on bacterial growth in the presence of drug [143]. This complexity depends partly on the details of the drug's mechanism and partly on the broader phenotypic consequences of perturbing gene expression.

2.4.1 TESEC-Alr$_{Mtb}$ allowed for a broad drug sensitivity range.

For this proof of concept, we went beyond naïve *Mtb*-Alr target overexpression and used a genetic circuit [233] for a log-linear inducer-expression response. Fine-tuned control over heterologous gene expression was important to generate a broad range of foreign target expression levels that allowed for host-cell growth. Importantly, we obtained a linear relation between the log-concentrations of L-arabinose and the MIC_{50} values of the positive control drug D-cycloserine (Figure 2.4). Next, we used TS144 hypomorph with reduced Alr$_{Mtb}$ expression and TS144 hypermorph with high expression of Alr$_{Mtb}$ with the aim to vary the sensitivity of our TESEC strain to alanine racemase specific inhibitors. This strategy allowed for the hit identification of known alanine racemase inhibitor D-cycloserine, a variety of β-lactams and benazepril as the only compound with no previously reported antibacterial activity (Figure 2.5).

2.4.2 Alanine racemase overexpression renders β-lactam resistance

Chemical-genetic profiling of TS144 in a 2D-gradient of L-arabinose and hit compounds confirmed the importance of alanine racemase activity on the bacterial susceptibility to β-lactams ceftibuten and aztreonam as well as ACE inhibitor benazepril (Figure 2.6). This study is not the first report that finds an association between the intracellular D-alanine pool and the bacterial susceptibility to beta-lactams. Previous studies with systemic methicillin-resistant *S. aureus* (MRSA), and notably *M. tuberculosis* [234, 235, 236] discovered drug synergies between D-cycloserine and other compounds targeting peptidoglycan

biosynthesis. When the internal D-alanine concentration is low, the incorporation of incomplete tripeptide monomers into peptidoglycan could destabilize the bacterial membrane and create increased vulnerability to β-lactams that target penicillin-binding proteins involved in cross-linking peptides [212, 234].
Vice versa, TS144 was less susceptible to β-lactam activity when alanine racemase was overexpressed. To our knowledge, this is the first report that provides evidence of a resistance mechanism in *E. coli* where elevated alanine racemase activity can increase the resistance to β-lactams.

2.4.3 Understanding benazepril's mechanism of action

ACE-inhibitor benazepril was further investigated for its potential to inhibit Alr_{Mtb}. A hint for Alr_{Mtb} specificity was obtained by analyzing benazepril's effect in the presence of D-cycloserine. Drug-drug combinations can improve, neutralize, or add to each other's effectiveness leading to respectively drug synergy, antagonism, or additivity [237]. In general, improved efficacy for combinations of drugs can lead to useful insights to design more effective therapeutic strategies, and may also provide information about the mechanism of action [238]. Interestingly, drug synergy was found when TS144 was exposed to checkerboard titration of benazepril and D-cycloserine (Figure 2.9). While various explanations for drug synergism exists [239, 238], synergy can occur when two drugs target different sites of the same protein, or when different targets in the same pathway are affected [238]. Both scenario's could be a possible for combination between D-cycloserine and benazepril. Because D-cycloserine also binds to D-ala-D-ala ligase [224], benazepril could act on a different site of alanine racemase

or the second D-cycloserine target D-ala-D-ala ligase. A second indication that the D-alanine pathway was affected by benazepril was found when supplementation of the medium with D-alanine allowed for TS144 growth rescue (Figure 2.8).

More conclusive evidence for Alr_{Mtb} specificity by benazepril was generated by in vitro assays with cell lysate of TESEC strain TS008 with *Mtb* alanine racemase expression (Figure 2.12). Dose-dependent inhibition was observed for increasing benazepril concentrations, and analysis of the quotient velocity plot provided indications of mixed alanine racemase inhibition by benazepril. Yet, to be confident about biochemistry parameters and the mode of inhibition, additional experiments with purified Alr_{Mtb} should be performed. I succesfully isolated Alr_{Mtb} (Appendix Figure B.2) and the concluding experiments will be performed in the near-future.

Finally, we confirmed benazepril as an antimycobacterial compound for *Mycobacterium smegmatis* and importantly of *Mycobacterium tuberculosis* (Figure 2.13 & 2.15). Interestingly, while both mycobacteria were susceptible for benazepril, TESEC strain TS008 with TolC and EntC expression was not inhibited by the drug (Figure 2.10). Hence, the deletion background of TS144 was crucial to detect the antibacterial activity of benazepril and it indicates the possibility that TolC-related membrane pumps can export benazepril from *E. coli*'s cell (Figure 2.10).

Similarly, Schaller *et al.* found that *E. coli* $\Delta tolC$ was found to be more susceptible for the first-line TB drug pyrazinoic acid (POA) then wild type *E. coli*. Weak acids such as POA are efficiently removed from

E. coli through TolC-mediated efflux [164], while acid removal is deficient in *Mycobacterium tuberculosis* [164, 240]. A TolC-defficient *E. coli* could therefore be important to mimic *Mtb*'s poor efflux response to acidic compounds and explain the natural mycobacterial susceptibility to the weak acid benazepril (pKa 3.73).

Despite previous screening efforts [241] where the Prestwick Library was exposed to mycobacteria, benazepril has not previously reported as antimycobacterial. Although this is the first study that demonstrates the antimycobacterial properties of benazepril, this is not the first study that describes a link between an ACE inhibitor and tuberculosis. A Taiwanese population study [242] found a correlation between the chronic use of ACE inhibitors and the prevalence of active tuberculosis, but could not provide a mechanistic explanation why the use of ACE inhibitors provided protection against the development of tuberculosis symptoms [242].

In humans, benazepril's ester group is cleaved in the liver to generate active ACE inhibitor benazeprilat. However, in contrast to the ACE inhibiting effect, only exposure benazepril and not benazeprilat resulted in *M. smegmatis* and *M. tuberculosis* growth inhibition. Possibly, the low log P value of benazeprilat (log P 0.15) does not allow benazeprilat to enter the cell as most TB drugs have a log P value above 1 [243]. Oppositely, benazepril's log P value of 3.3 is comparable with that of recent TB leads with high lipophilicities (log P 4 on average) that target the tuberculosis membrane [244]. In addition, preliminary biochemical assays in which I exposed benazeprilat to cell lysate with Alr_{Mtb} did not signal that benazeprilat could reduce alanine racemase activity (data not shown).

I also found that benazepril's antimycobacterial action differs between TESEC and the tested mycobacteria. For example, drug-drug combination experiments with benazepril and D-cycloserine revealed an additive inhibition effect on *M. smegmatis* (Appendix Figure B.4) instead of a synergistic effect that was observed for TS144 (Figure 2.9). Since not all mycobacterial D-cycloserine targets are known [245] and additivity can result from drugs with the same target as well as unrelated targets [238] it is not possible to draw firm mechanistic conclusions from these experiments.

Nevertheless, the fact that D-alanine could not prevent mycobacterial growth inhibition by benazepril (2.14 & Table 2.2) strongly indicates the existence of secondary benazepril targets in *M. smegmatis* and *M. tuberculosis*.

2.4.4 Alanine transferase as possible drug target of benazepril

Because studies both report the essentially as well as nonessentiality of alanine racemase in *M. tuberculosis*, it could be beneficial if benazepril also acts on other targets than alanine racemase.
While Awasthy *et al.* [246] showed that *M. tuberculosis* with defective alanine racemase expression was not viable in macrophages and mice, work from Houlouska *et al.* and Mortuza *et al.* [247, 248] suggests that there is an alternative enzyme to generate D-alanine in mycobacteria. In that substitute pathway, an alanine transferase is responsible for the reversible catalysis of pyruvate + D-glutamate $\rightleftarrows$ α-ketoglutarate + D-alanine. Mortuza *et al.* showed that an the gene product of MSMEG_5795 that was previously annotated as encoding

for 4-amino-4-deoxychorismate lyase acts as alanine tranferase. They showed that overexpression of MSMEG_5795 could rescue growth of a *M. smegmatis* (Δalr) deletion strain [248].

Interestingly, in our study, the benazepril inhibition of an alanine transferase (Sigma) from an unknown organism was detected (Figure B.3). The degree similarity between the enzyme from Sigma and the proposed alanine transferase from *M. tuberculosis* and *M. smegmatis* could not be verified, because Sigma did not disclose the origin of their alanine transferase.
However, based on this preliminary result, it would be interesting to test benazepril on isolated mycobacterial alanine transferase as there might be a possibility that benazepril has inhibitory activity on MSMEG_5795 or the *M. tuberculosis* homologue Rv0812. Moreover, inhibition of MSMEG_5795 or Rv0812 could be important to make D-cycloserine inhibition more effective. It is known that Alr overexpression provides D-cycloserine resistance [249] as overproduction of D-alanine outcompetes binding of D-cycloserine's second target D-alanine-D-alanine ligase (Ddl) [224]. In theory, the same could be achieved through upregulation of D-alanine transferase. Therefore, alanine transferase inhibitors could be clinically relevant to prevent DCS resistance. Although we did not further investigate benazepril's mechanism beyond its inhibitory effect on Alr_{Mtb}, follow-up studies could be interesting to discover potentially unknown tuberculosis vulnerabilities. Finally, because millimolar range concentrations were necessary to inhibit *M. tuberculosis* growth, future work related to hit to lead optimization of benazepril would be useful to improve the drug's antimicrobial efficacy.

2.5 Conclusion

Tuberculosis is a global health problem [178] that calls for innovative solutions for drug development. *Mycobacterium tuberculosis* is a pathogenic, slow-growing organism that is difficult to genetically manipulate and risky to work with [99]. *E. coli* K-12 is a supreme model organism. It is safe, well-studied, fast-growing, easy to engineer, and preferred for experimentation thanks to its versatility [250]. The TESEC-based discovery of antimycobacterial benazepril demonstrates that, despite important organism-specific characteristics such as membrane architecture, rationally engineered *E. coli* strains can serve as a platform for tuberculosis drug discovery. We believe that this proof-of-concept has the potential to become widely adopted by the research community.

2.6 Contributions

AL and EW supervised the study. NB and EW constructed strains and plasmids, designed and performed the drug screens. EW wrote code, analyzed the drug screen data, and wrote the abstract and introduction. NB designed, performed and analyzed hit validation experiments with *E. coli*, *M. smegmatis* and enzyme assays. NB wrote the manuscript. ZE performed and analyzed the experiments with *M. tuberculosis*. I would like to thank Hamid Mebrouki for the disk diffusion assay, and Michel Arthur (Centre de Recherche des Cordeliers) for the valuable conversations and for offering the *M. smegmatis* mc^2155 strain.

2.7 Materials and Methods

2.7.1 Strains, plasmids & primers

Strains, plasmids and primers of this study are listed in Appendix B.1, B.2 & B.3.

2.7.2 Gene deletions in *E. coli*

Phage P1 transductions used to delete *E. coli* genes according to a method described by [187]. Kan^r markers were subsequently removed by transforming plasmid pCP20 with flipase expression.

2.7.3 Construction of expression plasmids

Plasmids were constructed using Golden Gate cloning and standard restriction/ligation cloning. PCR amplification for cloning purposes was performed using Phusion® high fidelity polymerase (NEB) following the manufacturer's instructions. PCR products were purified using the Wizard® SV Gel and PCR Clean-Up System (Promega). The destination vector pTS000 was derived from P37.pRD131. First, an internal BsaI restriction site in Amp^r gene was removed using site-directed mutagenesis based on a published protocol [251] using primers NB.27 and NB.28. A DNA fragment carrying the *lacZ* gene for *E. coli* β-galactosidase under the control of the lac promoter flanked by KpnI and XmaI restriction sites was synthesized (IDT) and cloned into P37.pRD131 backbone to replace the original mCherry gene using KpnI and XmaI restriction enzymes (NEB) and T4 DNA ligase (HC) (Promega). Genes alr_{Mtb} and alr_{Msmeg} were codon optimized (IDT Codon Optimization Tool) for expression in *E. coli* K-12 and synthesized *de novo* as gBlock (IDT). Next, alr_{Mtb} was amplified with primers NB.43 and NB.44 (Appendix Table B.3) to add flanking BsaI

restriction sites. Golden gate cloning was used to replace the original LacZ expression cassette with alr_{Mtb} or alr_{Msmeg} in backbone plasmid pTS000 to generate respectively pTS008 and pTS068.

A second destination vector DVA_AD was derived from pTS000 inspired by the CIDAR MoClo [7] assembly standard to facilitate modular cloning of multiple parts and future target genes of interest. Based on this standard, level 0 destination plasmids annotated as DVC_XX (Destination Vector with Cam^r resistance_Prefix Suffix) were constructed by combining the pSB1C3 backbone of biobrick Bba_J04450 and the lacZ expression cassette amplified from CIDAR MoClo vector DVA_CD. CIDAR MoClo suffixes and prefixes were adopted and expanded when needed. Linearized vector and insert fragments amplified with MoClo primers (Appendix Table B.3) were digested with BbsI (NEB) and assembled following the golden gate reaction protocol below. Primers for the construction of level 0 plasmids (L0_XXX) are listed in Appendix Table B.3. Plasmids pTS059, pTS067, and pTS072 were assembled using level 0 plasmids listed in Appendix Table B.2.

10 μL Golden gate reaction cocktails contained:

- 40 fmol vector
- 40 fmol insert(s)
- 0.2 μL T4 ligase (HC) (Promega)
- 1.0 μL T4 Buffer (Promega)
- 0.4 μL BbsI or BsaI-HF (NEB)

Golden gate thermocycles:

1. 37°C for 2 min

2. 16°C for 5 min

3. 50 cycles back to step 2

4. 50°C for 10 min

5. 80°C for 5 min

6. Hold at 4°C

2.7.4 Bacterial transformations

5-10 μL (golden gate) ligation reactions were transformed to 30 μL of chemically competent Dh5-α cells. Transformants were selected on LB agar plates with either 50 μg/mL chloramphenicol, 100 μg/mL ampicillin or 50 μg/mL kanamycin. White colonies from transformations with golden gate reaction mix were selected on solid LB plates with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and 20 mg/μL 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside (X-gal) in addition to an appropriate antibiotic. Experimental strains were constructed by transforming 100 ng plasmid(s) to 30 μL chemically competent deletion strains and plated on selective LB agar plates. Colonies were grown overnight at 37°C.

2.7.5 Plasmid isolation and sequence analysis

Plasmids were routinely isolated using Wizard® Plus SV Minipreps DNA Purification (Promega). Correct plasmid construction was verified by sanger sequencing (Eurofins Genomics). Sequences were evaluated using Geneious Bioinformatics software.

2.7.6 PCR verification of chromosomal deletions

Routine PCR verification of chromosomal deletions was done using verification primers listed in Appendix Table B.3 and Thermo Scientific DreamTaq Green PCR Master Mix (2x) following the supplier manual.

2.7.7 Culture conditions

E. coli strains were streaked on Luria-Bertani (LB) 1.5% agar plates supplemented with 50 μg/mL kanamycin (Sigma), 100 μg/mL ampicillin (Sigma), 100 μg/mL D-alanine (Sigma) when appropriate. Single colonies were grown in Luria-Bertani (LB) medium or M9 medium with 1x M9 Minimal Salts (Difco^{TM}), 1 g/L Drop-out Mix Complete, Adenine Rich, w/o Yeast Nitrogen Base (Powder) (USBiological Life Sciences), 2 mM $MgSO_4$ and 0.1 mM $CaCl_2$ and 0.4% (w/v) D-fructose (Sigma) medium 0.03% (w/v) vitamin B1 (Sigma) with appropriate antibiotics and D-alanine when necessary. All cultivations were performed at 37°C.

M. smegmatis mc^2155 was cultured at 37°C in Middlebrook 7H9 base broth (Sigma) supplemented with 10% Middlebrook ADC growth supplement (v/v) (Sigma), 0.2% (v/v) glycerol and 0.1% (v/v) Tween® 80 (Sigma). Solid medium was prepared from Middlebrook 7H9 medium or LB medium with 1.5% agar.

M. tuberculosis H37Rv was cultured at 37°C in Middlebrook 7H9 broth (BD Difco) containing 0.2% (v/v) glycerol (VWR) and 10% (v/v) OADC supplement (BD Difco).

2.7.8 Characterization of *tolC entC* deletion strains

Overnight cultures of K028, K031, K123, D20 were diluted until OD_{600} 0.005 in M9 0.4% fructose medium and transferred to 96-well plates. Bacterial growth was determined by measuring OD_{600} of 100 μL microwell cultures over a time course of 15 hours in the Spark® multimode reader from TECAN. In parallel, potential differences in cell shapes between the strains during the stationary phase (after 20 hours of cultivation) and exponential growth (OD_{600} = 0.3-0.5) were analyzed using phase-contrast microscopy (Zeiss).

2.7.9 Prestwick Chemical Library® Drug Screens

The Prestwick Chemical Library® compounds in DMSO (Sigma) were distributed in 96-well black Greiner Bio-One plates using the TECAN Freedom Evo liquid handling robot. Single colonies of TS144 or TS163 were inoculated in M9 medium supplemented with 50 μg/mL kanamycin, 100 μg/mL ampicillin and 100 μg/mL D-alanine for overnight cultivation (37°C; 150 rpm). The next day, 40 mL of overnight culture was centrifuged for 10 minutes at 4000 rpm, and the cell pellet was washed twice by resuspending in M9 medium without D-alanine. Finally, the washed cell culture was diluted until OD_{600} 0.025 in the final assay medium (M9 medium with 50 μg/mL kanamycin and/or 100 μg/mL ampicillin and a defined arabinose concentration). The Prestwick Chemical Library® was screened against Alr_{Mtb} expressing strain TS144 or the control strain TS163 with native alanine racemase expression. TS144 was screened in the presence of 0.1 μM or 10 mM L-arabinose concentrations. The control strain TS163 was screened without L-arabinose. For each drug screen, 150 μL of washed bacterial suspensions were distributed with the Multidrop dispenser (Thermo

Fischer Scientific) to 96-well plates containing 1.5 μL Prestwick Library compounds for a final DMSO concentration of 1% and compound concentration of 0.1 mM. The optical density was measured at 600 nm (TECAN, Spark® multimode reader) after 0 and 10 hours of incubation at 37°C (non-shaking conditions).

2.7.10 Statistical analysis of drug screening results

Hit compounds were selected by comparing growth measurements of drug-treated TS144 under high (10 mM) L-arabinose induction for hypermorphic Alr_{Mtb} expression to those of the same strain under low (0.1 μM) arabinose to generate a hypermorphic strain. The assay was performed in three biological replicates and background-subtracted to compensate for absorbance by the drug and media. A scatter plot of median absorbance values revealed a subset of 10 drugs allowing significant growth ($OD_{600} > 0.2$) under high induction but reduced growth ($OD_{600} < 0.1$) under low induction. The statistical significance of these hit compounds was assessed by calculating the strictly standardized mean difference (SSMD) metric of OD under drug treatment to DMSO-only controls. For each of the 10 hit compounds, SSMD values less than -8 indicated significant reduction in OD values for drug-treated cells under low arabinose induction. Under high L-arabinose induction, SSMD values greater than -5 indicated weak or insignificant growth inhibition.

2.7.11 Chemical-genetic profiling of hit compounds

Hit compounds that resulted from the initial drug screen were re-evaluated for their potential mechanism of action by growing TS144 in the presence of a two-dimensional gradient of two-fold 'hit' or 'control' compound dilution series (Table 2.3) and ten-fold L-arabinose dilution

series (0 – 10 mM). 50 μL bacterial cultures were grown in Nunc^{TM} 384-Well Optical Bottom Plates (Thermo Fischer Scientific). OD_{600} measurements were taken after 0 and 10 hours of incubation at 37°C, with shaking at 600 rpm.

Compound (Supplier; stock concentration; solvent)	**Compound concentration range (mM)**	**[Solvent] (%, v/v)**
Aztreonam (TCI Chemicals; 215 mM; DMSO)	0 - 1.02	0.5
Amlexanox (TCI Chemicals; 50 mM; DMSO)	0 - 0.51	1.0
Benazepril HCl (TCI Chemicals; 4.5 mM; DSMO)	0 - 0.36	0.1
D-cycloserine (Sigma-Aldrich; 979 mM; H_2O)	0 - 1.00	0.2
Ceftibuten hyrate (TCI Chemicals;15 mM; DMSO)	0 - 0.15	1.0
Cefmetazole sodium salt (Sigma-Aldrich; 100 mM; H_2O)	0 - 0.40	0.4
Fosinopril (Sigma-Aldrich; 43 mM; H2O)	0 - 0.43	1.0
Cefmetazole sodium salt (Sigma-Aldrich; 10 mM; H_2O)	0 - 0.04	0.4

Table 2.3: Chemical compounds, stock concentrations and dilution series.

2.7.12 Response envelope analysis to quantify the effect of drug-drug combinations

A published bioinformatics approach called Response Envelope Analysis (REA) [225] was used to determine synergism, additivity, or antagonism of drug-drug pairs on *E. coli* or *M. smegmatis* growth.The overall effect of the drug combination was calculated using the REA toolbox in MATLAB (Version R2017a, Mathworks). REA combines the Bliss Independence model [226] and a generalized Loewe Additivity model [227, 228] to make 3D landscape plots based on the local inhibitory effects of drug-drug checkerboard combinations. The overall effect of the local inhibitory effects are scored as Synergy Index (SI) and Antagonism Index (AI). As such, SI and AI scores provided insights to whether the drug combination was synergistic (for SI>AI),

antagonistic (for SI<AI) or additive (for SI=AI).

2.7.13 Susceptibility assays of *M. smegmatis* mc^{2}155

MICs were determined by a microplate two-fold dilution method. Single *M. smegmatis* mc^{2}155 colonies were inoculated in 5 mL 7H9 medium for ∼45 hours of cultivation at 37°C and 150 rpm agitation speed. Flat black 96-well microtiter plates (Greiner Bio-One) with 100 μL microwell cultures were prepared with diluted pre-cultures (OD_{600} 0.01) and two-fold drug dilution in 7H9 medium. Each well contained a final DMSO concentration of 1% and 1x10^5 CFUs. Final concentration ranges of benazepril, benazeprilat, D-cycloserine, and DMSO (control) were respectively: 0 – 4.5 mM; 0 – 0.5 mM; 0 – 2 mM; 0 – 5%. Plates were covered with an aluminum cover seal (Croning) during incubation at 37°C, 750 rpm in the Titramax 1000 shaker (Heidolph) and OD_{600} measurements were taken at 0, 24 and 48 hours of incubation with the Spark® multimode reader from TECAN. To verify potential bactericidal activity, ten-fold-dilution series were made from 48 h-cultures in wells with 0, 0.56, 1.13, 2.25 and 4.5 mM benazepril and 0, 0.22 0.56, 0.13, 0.25, 0.5 mM D-cycloserine and 1 μL aliquots were spot plated onto 7H9 1.5% agar plates. After 48 hours of incubation of 37°C images of the plates were taken. The MIC was determined as the lowest concentration of inhibitor that prevented growth.

2.7.14 Susceptibility assays of *Mycobacterium tuberculosis* H37Rv

The antibacterial activity of hit compound benazepril (Sigma) and control compounds isoniazid (Sigma), DMSO (VWR) and D-cycloserine (Sigma) was determined by the microdilution method on 5 biological replicates. To investigate potential growth rescue by D-alanine, a fixed concentration of D-cycloserine (0.125 or 0.25 mM) or benazepril (2.25

or 4.5 mM) (corresponding to their respective MIC value and one additional concentration) was added to each well containing two-fold serial dilutions of D-alanine (0 – 5 mM). Individual wells of a 96-well plate containing the growth medium and two-fold serial dilutions of the tested compounds were inoculated with 1x10^5 CFUs in a total volume of 200 μL per well. The plates were incubated at 37°C for 7-9 days and resazurin (0.0025%, w/v) was added to each well before a further incubation of 24 h. The MIC was defined as the lowest drug concentration that prevents the resazurin color change from blue to pink.

2.7.15 Preparation of *E. coli* cell extracts

E. coli $\Delta alr \Delta dadX$ deletion strains TS008 and TS028 with inducible expression of respectively Alr_{Mtb} or mCherry were grown in 200 mL LB medium with 100 μg/mL ampicillin, 50 μg/mL kanamycin and 100 μg/mL D-alanine at 150 rpm and 37°C. Alr_{Mtb} and mCherry expression was induced in early exponential phase through the addition of L-arabinose to a final concentration of 10 mM. After 6.5 hours of induction, 50 mL of the cell cultures were harvested by centrifugation at 3,434 x g for 30 min at 4°C. Next, the collected bacterial cells were chemically disrupted in 2 mL B-PER® Bacterial Protein Extraction Reagent (Thermo Scientific) with 50 μg/mL Lysozyme from chicken egg white (Sigma Aldrich) and protease inhibitor (cOmplete^{TM} ULTRA Tablets, EDTA-free; Roche). The lysate was cleared by centrifugation at 20,000 x g for 15 minutes and dialyzed in a in a Slide-A-Lyzer Dialysis Cassette (20,000 MWCO) (Pierce, Rockford, IL) for 48 hours at 4°C against 200 mL of 50 mM potassium phosphate, pH 7.5. Protein concentrations in the dialyzed samples were estimated by measuring absorbance at 280 nm on a NanoDrop (NanoDrop 2000,

Thermo Scientific). The protein concentrations were ∼15-20 mg/mL and the samples were stored at –20°C.

2.7.16 Reagents for alanine racemase purification

B-PER™ Complete Bacterial Protein Extraction Reagent, 3-mL HisPur Ni-NTA spin columns, Phosphate-Buffered Saline (PBS) (10X, 200 mM sodium phosphate, 3 M sodium chloride, pH 7.4), Imidazole (2M, pH 7.4), Pierce™ Rapid Gold BCA Protein Assay Kit were purchased from Thermo Fischer Scientific. Trizma® base (Sigma Aldrich) and Sodium Chloride (Merck) were used to prepare alanine racemase buffer. Merck™ Amicon™ Ultra Centrifugal Filter Units (10,000 MWCO) were used to concentrate the isolated protein sample.

2.7.17 Protein purification of *Mtb* alanine racemase

Mycobacterium tuberculosis H37Rv alanine racemase was purified from strain TS167 that lacked native alanine racemase activity. An overnight culture of TS167 in LB with 100 μg/mL ampicillin and 50 μg/mL kanamycin, 100 μg/mL D-alanine was washed twice and diluted (1:100) in two Erlenmeyer flasks with 250 mL of LB supplemented with 100 mM L-arabinose, 100 μg/mL ampicillin and 50 μg/mL kanamycin. The bacterial cell cultures were grown for 20 hours (37°C, 150 rpm) and harvested by centrifugation at 3,434 x g for 30 min at 4°C. Next, 2 grams of cell pellet was chemically disrupted in 10 mL B-PER® Complete Bacterial Protein Extraction Reagent and 1/2 protease inhibitor tablet (cOmplete™ ULTRA Tablet, EDTA-free; Roche). The lysate was cleared by centrifugation at 20,000 x g for 15 minutes and transferred to 3-mL HisPur Ni-NTA spin columns. HisPur Ni-NTA columns were eluted with 1x PBS buffer with increasing concentrations of imidazole (5 – 500 mM). Elution fractions were ver-

ified on an 4-15% Mini-PROTEIN® TGX Protein gel (Bio-Rad) and fractions with a single *Mtb* alanine racemase band (200 mM, 250 mM and 500 mM imidazole) were combined and dialyzed in a in a Slide-A-Lyzer Dialysis Cassette (10,000 MWCO) (Pierce, Rockford, IL) for 72 hours at 4°C against 20 mM Tris, 100 mM NaCl pH 8.0. Finally, the protein sample was concentrated using an Amicon^{TM} Ultra Centrifugal Filter Unit and centrifugation at 3,434 x g for 30 min at 4°C.

2.7.18 Quantification of isolated *Mtb* alanine racemase

Protein concentrations were determined using the Pierce^{TM} Rapid Gold BCA Protein Assay and measured following the manufacturer's protocol. 20 μL standards with pre-determined BSA concentrations in Tris buffer (20 mM Tris, 100 mM NaCl, pH 8) and sample with isolated Alr_{Mtb}. were mixed with 200 μL working reagent. After incubation at room temperature for 5 minutes, absorbance measurements at 460 nm were taken. Final *Mtb* alanine racemase concentration was determined based on linear regression analysis of the standard curve in Microsoft Excel.

2.7.19 *Mtb* alanine racemase inhibition assay

The coupled D-alanine racemase assay was executed in flat black 96-well microtiter plates (Greiner Bio-One). 50 μL reaction cocktails consisted of 0.02 mg/mL cell lysate from TS008 (with Alr_{Mtb} expression) or TS028 (with mCherry expression) in potassium phosphate (50 mM, pH 7.5), 0.05 units/mL *Bacillus subtillus* L-alanine dehydrogenase, varying D-alanine concentrations (0-20 mM) and varying benazepril concentrations (0-9 mM, 0.2% DMSO) in $NaHCO_3$ buffer (110 mM, pH 10.5). Fluorescence associated with NADH production was measured in an TECAN Spark reader with excitation/emission

wavelengths 340/460 nm. Reaction without D-alanine were used as control samples and for background measurements. Initial reaction rates were determined in MATLAB (Version R2017a, Mathworks) by linear regression of the fluorescence measurements in time. The relative activity was determined as the initial velocity in a specific substrate/inhibitor concentration divided by the maximal initial velocity in the presence of 20 mM D-alanine without benazepril. The experiment was conducted four times.

Chapter 3

Synthetic *E. coli* mutant libraries provide new insights into the D-cycloserine resistance landscape of *M. tuberculosis* alanine racemase

3.1 Abstract

Insights into the resistance landscapes of tuberculosis drug targets are valuable for the development of rapid AMR diagnostics and for the prioritization of hit compounds. With cases of drug-resistant *Mycobacterium tuberculosis* (*Mtb*) on the rise, *Mtb* treatments will increasingly depend on the use of second-line antibiotics such as D-cycloserine (DCS). As for all antibiotics, increased DCS usage could lead to increased frequencies of DCS resistant *Mtb*. Previous research suggest that mutations in alanine racemase (Alr), frequently underlie D-cycloserine resistance.

However, a surprisingly limited number of the reported mutations in Alr have been studied in detail to prove their contribution to DCS

resistance. Therefore, this study aims at validating and identifying DCS resistant *Mtb alr* genotypes using recombinant *E. coli*. More specifically, the goal is to scan the Alr_{Mtb} mutational landscape of single amino acid changes and determine which mutations are directly linked to DCS resistance phenotypes.

For that purpose, I made two synthetic Alr_{Mtb} mutant libraries. One based classical error-prone PCR and a second based on rational DNA design to cover all possible single non-synonymous Alr_{Mtb} mutations. The libraries were heterologously expressed in *E. coli* strains that rely on Alr_{Mtb} mutant expression for bacterial growth.

Mutant library screens and in-depth characterization of *E. coli* strains with Alr_{Mtb} mutant expression confirmed the previously reported mutations A308G, Y388D, Y338C, L113R, P311S, and D344N as potentially direct contributors to DCS resistance phenotypes and identified 130 additional non-reported non-synonymous Alr_{Mtb} mutations that could lead to DCS resistance phenotypes.

A significant advantage of this approach is that it facilitates the study of pathogenic resistance in a fast-growing non-pathogenic host. While it is still uncertain which of the predicted drug resistance mutations in this study are translatable to clinical *Mtb* isolates, we believe that our foresight could provide advantageous warning signs when patients carry *Mtb* with mutations that were validated in this study. Moreover, the systematic screening of mutant libraries can be used to prioritize new drug candidates that effectively inhibit D-cycloserine resistant alanine racemases or for which target-based resistance is modest compared to other inhibitors.

3.2 Introduction

3.2.1 Detection and treatment of MDR-TB

Monitoring and studying multi-drug-resistant tuberculosis (MDR-TB) is essential to improve patient outcomes [72]. When resistance is detected for first-line TB drugs, appropriate second-line antibiotics must be used as therapeutic replacements. Those second-line antibiotics are categorized in groups A, B, and C, referring to the respective order of clinical preference, based on their effectiveness and adverse side effects [178].

The latest guidelines on treating MDR-TB by the World Health Organisation (WHO) [178] recommend a combination of three group A antibiotics (levofloxacin or moxifloxacin, bedaquiline, and linezolid) with one group B antibiotic (clofazimine, cycloserine or terizidone).

To increase the chance that any of those treatments result in a positive outcome, it is necessary to verify *Mtb* 's susceptibility to the scheduled antibiotics. In recent years, genetic biomarkers related to first- and second-line antibiotic resistance have been well-characterized [72] and have led to the development of rapid genetic-based diagnostics [252]. Ideally, such tests would be available for all TB drugs. However, the group B antibiotic D-cycloserine (DCS) is an example of a second-line bacteriostatic agent for which the genetic basis of resistance is not fully understood and for which genetic tests are unavailable today [6, 253].

Despite the severe psychiatric and central nervous system toxicity side effects of DCS [254, 255], the drug may also have advantages. A

recent study [253] showed that the drug lacks cross-resistance with other tuberculosis drugs and that the spontaneous emergence of D-cycloserine resistance mutations is slow compared to other TB drugs.

Concomitantly, *M. tuberculosis* strains with DCS resistance have been identified in patients [4, 6], such that appropriate drug susceptibility testing (DST) is important to prevent unnecessary exposure to the toxic drug and aid medical professionals to prescribe the right treatment. In the absence of a molecular diagnostic, phenotypic screening is usually still an option. However, in the specific case of DCS, phenotypic DST is technically not possible since in vitro tests with *Mtb* strains consistently generate results that do not match with the clinical outcomes [256, 257]. This lack of testing options further motivates the study of the genetic basis of DCS resistance in order to support the future development of innovative DST solutions.

3.2.2 DCS mechanism of action and resistance

Research into DCS's mechanism of action suggests that the drug primarily operates as a competitive inhibitor of two major enzymes involved in bacterial peptidoglycan synthesis: alanine racemase (Alr), and D-alanine D-alanine ligase (Ddl) (Figure 3.1) [258, 6]). Although DCS has been on the market for decades [259], the scientific community is still debating details of its mechanism of action and how *Mtb* becomes resistant to the drug.

Studies have pointed at Ddl as the main target during DCS inhibition since DCS inhibition is stronger for Ddl than for Alr [224, 247]. On the other hand, the most effective cause of DCS resistance appeared to be related to the modification and overexpression of Alr [6, 249, 260, 253]. This can be explained by the competitive binding

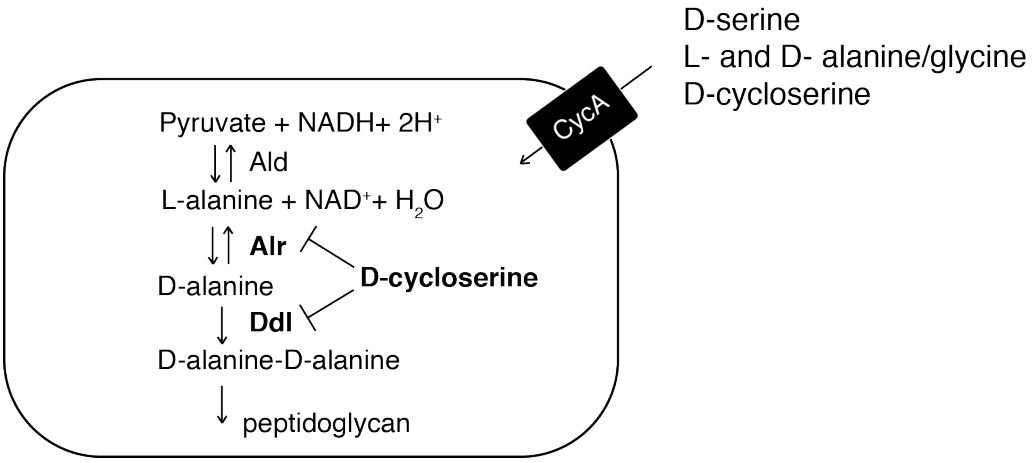

Figure 3.1: Mechanism of action by D-cycloserine.
L-alanine is produced via L-alanine hydrogenase (Ald) activity and converted into D-alanine by alanine racemase (Alr). CycA-mediated uptake allows D-cycloserine to reach the inside the of the cell where the drug inhibits the D-alanine and D-alanine-D-alanine production through competive inhibition of Alr and D-alanine-D-alanine ligase (Ddl). Modified from Desjardins *et al.* (2016), figure 3 [6].

between D-alanine and D-cycloserine to Ddl. When alanine racemase activity is maintained in the presence of DCS, D-alanine can outcompete DCS-binding to Ddl [224] thereby facilitating *Mtb*'s survival.

Other DCS resistance mechanisms, such as Ddl overexpression [260], mutations in the CycA transporter [261, 262] or the loss of function of L-alanine dehydrogenase (Ald) [6] have been reported as well. Although mutations Ald were confirmed as contributors to DCS resistance of *Mtb* clinical isolates, conclusive evidence that the other DCS resistance mechanisms play a role in *M. tuberculosis* remains to be elucidated [6].

3.2.3 Studying *Mtb* DCS resistance in engineered *E. coli*

DCS resistance is studied using mycobacterial model organisms [260, 261], enzyme assays and bioinformatics approaches [215] as well as by making associations between whole-genome sequencing and phenotypic screening of drug-resistant *Mtb* [6, 4, 263]. Few studies involving

drug-resistant *Mtb* strains [253, 6] validate mutations that are likely to be involved in DCS resistance. Moreover, to our best knowledge, the re-engineering of suspected DCS resistance mutations into wild type *Mtb* strains has not been done so far.

This study attempts to establish a high-throughput approach to exhaustively predict all possible single amino acid mutations that lead to Alr_{Mtb}-dependent D-cycloserine resistance.

We hypothesized that we could confirm the role of the Alr_{Mtb} specific mutations in drug-resistant TB when the same non-synonymous mutations would also render increased DCS resistance in our *E. coli*-host. When the same mutation generates DCS resistance in two organisms that are evolutionary distant, it would be more likely that the amino acid change results in a structural modification that affects drug binding.

On the contrary, it is possible that specific *alr* mutations in DCS resistant *Mtb* do not decrease the DCS susceptibility of our engineered host. Such instances could be explained in different ways. First of all, it is possible that a mutation or mutations in another protein than Alr is merely responsible for the DCS resistance phenotype of *Mtb*. Drug-resistant clinical *Mtb* isolates frequently carry mutations in proteins that are not directly targeted by DCS [6]. Moreover, it is possible that certain DCS resistance mutations are co-dependent on host-specific factors related to e.g. epistasis, genetic background, codon usage, protein folding, interactions with other proteins, or variations in protein expression level.

We used different approaches to investigate Alr-dependent D-cycloserine resistance:

1. The introduction of previously reported Alr_{Mtb} mutants from DR-TB clinical isolates into our chassis

2. Screens for D-cycloserine resistance mutants using libraries of synthetic *E. coli* strains with alr_{Mtb} mutant genotypes

3. Molecular dynamics simulations to obtain complementary information about how experimentally identified mutations impact DCS-Alr_{Mtb} binding

Similar to the approach described in Chapter section 2.3.1, we used TESEC strains in which the *Mtb* functional homolog replaces the activity of *E. coli*'s native alanine racemase enzymes.
Based on this strain design, I reconstructed six previously reported *Mtb alr* mutations in *E. coli*. My work confirms that three out of six are directly responsible for the reported DCS resistance.

Novel insights into the Alr_{Mtb} DCS resistance landscape could eventually accelerate the development of rapid genetic-based tests. In addition, the method described in this chapter complements the drug screening strategy described in Chapter 2. TESEC strains with drug-resistant *Mtb* targets could be used to screen for hits compounds for drug-resistant tuberculosis. Similarly, target-based drug resistance of hit compounds could be assessed in advance by exposing the drugs to mutant libraries. This information could be beneficial for the prioritization or optimization of hits molecules.

3.3 Results

3.3.1 Six reported *Mtb* alanine racemase mutants rescue *E. coli* growth

We first focused on expressing six reported Alr_{Mtb} mutants [6] in an auxotrophic *E. coli* strain ($\Delta alr \Delta dadX$) to asses the feasibility of *E. coli*-based discovery of *Mtb* drug resistance genotypes. The selected Alr mutants were previously identified in *Mtb* clinical isolates with increased resistance for DCS (Table 3.1). In some of the reported drug-resistant *Mtb* strains (e.g. TKK_03_0026, TKK_04_0075 & TKK-01-0050), Alr mutations co-existed with mutations in other enzymes involved in D-alanine-D-alanine biosynthesis (Table 3.1).

			Mutations in enzymes:				
Strain name	Spoligo	Drug Susceptibility	Alr	Ald	Ddl	CycA	DCS MIC (µg/mL) in Löwenstein-Jensen media
TKK_03_0026	Unknown	susceptible	A308G	WT	WT	A18T	20
TKK_02_0004	LAM4	XDR	K157E	WT	WT	WT	60
TKK_02_0055	Beijing	XDR	L113R	WT	WT	WT	60
TKK_04_0090	LAM4	mono	L113R	WT	WT	WT	40
TKK_04_0105	LAM4	XDR	M343T	WT	WT	WT	>60
TKK_02_0007	LAM4	XDR	R397G	WT	WT	WT	25
TKK_04_0075	T1	Pre-XDR	Y338D	1 bp deletion at 132	WT	WT	40
TKK-01-0050	T1	MDR	Y338D	1 bp deletion at 132	WT	WT	40

Table 3.1: DCS resistant *Mtb* clinical isolates with non-synonymous mutations in enzymes of the D-alanine-D-alanine biosynthesis pathway. Published data from Desjardins *et al.*, 2016 [6].
The DCS MIC's of *Mtb* clinical isolates were determined in Löwenstein-Jensen medium. The D-cycloserine MIC of wild type *M. tuberculosis* in the reference study was determined as 15 μg/mL.

Plasmids carrying Alr_{Mtb} with non-synonymous mutations A308G, K157E, L113R, M343T, R397G or Y388D in upstream a p_{BAD} promoter were inducibly expressed in TESEC strains through the addi-

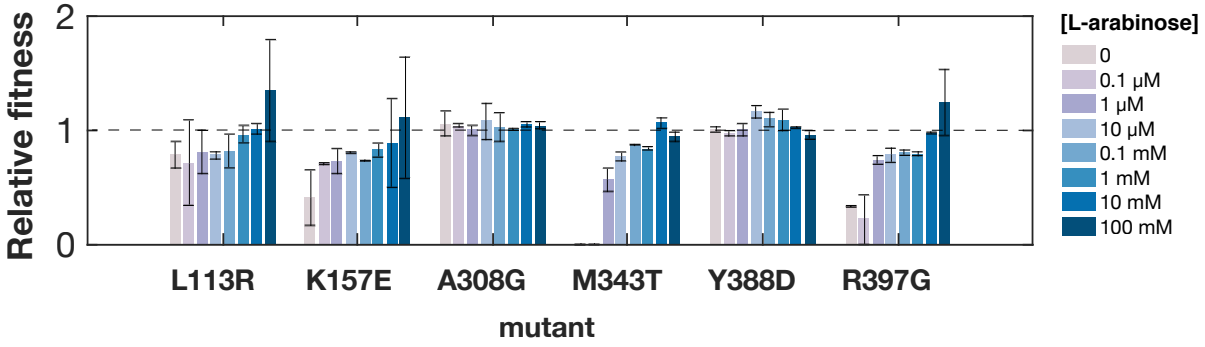

Figure 3.2: Relative fitness of *E. coli* expressing Alr_{Mtb} mutants.
E. coli strains MUT.0001 (Alr_{Mtb} A380G), MUT.0002 (Alr_{Mtb} M343T), MUT.0003 (Alr_{Mtb} Y388D), MUT.0005 (Alr_{Mtb} K157E), MUT.0006 (Alr_{Mtb} L113R), MUT.0007 (Alr_{Mtb} R397G) and TS008 (control with wild type Alr_{Mtb} expression) were grown at 37°C in 90 μL cultures with M9 0.4% fructose medium and 0-100 mM L-arabinose concentrations. Growth profiles were obtained by measuring the OD_{600} of each well of the 96-well microtiter plates over 20 hours. The relative fitness was determined as the maximal growth rate of the mutant strain divided by the maximal growth rate of the control strain. Bar graphs represent the mean values and standard deviations of three independent experiments.

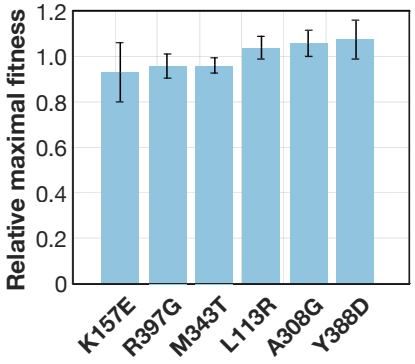

Figure 3.3: Relative maximal fitness of *E. coli* expressing Alr_{Mtb} mutants K157E, R397G, M343T, L113R, A308G, Y388D.
The maximum achievable growth rate for each mutant was calculated with data from figure 3.2. Growth rates of each TESEC-Alr_{Mtb} mutant strain and the control strain (TS008) were determined in the presence of L-arabinose series to find the maximal growth rate at a specific induction level. The relative maximal fitness was defined as the maximal growth rate of the TESEC-Alr_{Mtb} mutant strain divided by the maximal growth rate of the control strain TS008. Bar graphs represent the mean values and standard deviations of three independent experiments.

tion L-arabinose. Next, I asssayed the mutant strains for their fitness in the absence of DCS (Figure 3.2 & 3.3) and their ability to resist DCS (Figure 3.4) compared to the control TESEC strain TS008 with wild type Alr_{Mtb} expression.

All Alr_{Mtb} mutants with single amino acid substitutions were able

to functionally rescue *E. coli* growth (Figure 3.2). However, for some mutants, the relative bacterial fitness was highly dependent on inducer level that was present in the culture medium. For example, whereas leaky expression in absence of L-arabinose provided sufficient Alr_{Mtb} activity for growth of the control strain TS008, and for most mutant strains, the TESEC strain with Alr_{Mtb} M343T expression required a minimum inducer concentration of 1 μM L-arabinose to grow. Interestingly, all seven (including the control) strains had similar growth rates in the presence of 10 mM L-arabinose.

I observed that *E. coli* strains MUT.0003 (Alr_{Mtb} Y388D) and MUT.0001 (Alr_{Mtb} A308G) have comparable growth rates to the control strain over the full spectrum of tested L-arabinose concentrations. This indicates that those mutations lead to minimal fitness changes. I also analysed whether variations in expression levels allow any of the six mutants to achieve the same maximal fitness as the control strain (Figure 3.3). My results indicate that expression level variations could indeed compensate for the fitness loss of any of the tested non-synonymous mutations in Alr_{Mtb}. Moreover, I found that differences in the maximal achievable fitness were insignificant.

3.3.2 Three out of six reported *Mtb* alanine racemase mutants rendered D-cycloserine resistant *E. coli* strains

To verify DCS resistance, I cultivated the six TESEC strains with Alr_{Mtb} mutant expression to combinations of L-arabinose and D-cycloserine series (Figure 3.4). When TESEC-Alr_{Mtb} mutants displayed reduced susceptibility to DCS compared to the control TESEC with wild type Alr_{Mtb} expression, they were coined as having a phenotype with increased DCS.

Because the DCS susceptibility depended on the Alr_{Mtb} expression level, seperate DCS IC_{50} values were determined for varying induction levels. IC_{50} values that were determined in the presence of relatively low L-arabinose concentrations (<0.1 mM) displayed high variability (not all data shown). It was therefore difficult to reliably determine whether the DCS susceptibility was increased or reduced. More conclusive data resulted from conditions involving the higher end of the applied L-arabinose concentrations (0.1-100 mM) (Figure 3.4).

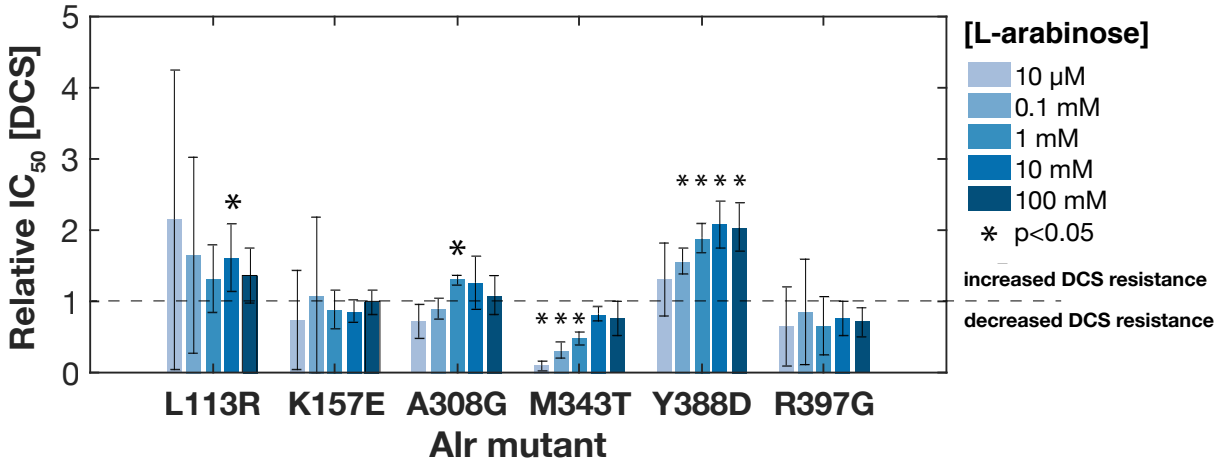

Figure 3.4: Relative IC_{50} values for D-cycloserine inhibition of *E. coli* expressing Alr_{Mtb} mutants.
The relative IC_{50} values were determined by the OD_{600} after 10 hours of cultivation in 96-well plates of each mutant strain divided by the OD_{600} at the same time point of the control strain (TS008) that expressed the wild type Alr_{Mtb}. All strains were cultivated at 37°C in M9 0.4% fructose medium with L-arabinose and D-cycloserine series. IC_{50} values were determined by fitting a hill-function on the dose-response curves. The horizontal dotted black line at y=1 indicates the normalized control value. Values above 1, therefore represent an increased resistance for D-cycloserine compared to the control strain, whereas values below 1 represent an increased sensitivity to D-cycloserine. Bar graphs represent the mean values, and standard deviations of 3 biological replicates. P-values were calculated with a two-sample *t*-test.

Conclusively, Alr_{Mtb} mutants A308G, Y388D and L113R confer significant DCS resistance (Figure 3.4). Notably, the resistance phenotype of TESEC-Alr_{Mtb} mutants A308G and L113R was significantly more dependent on the level of induction compared to TESEC-Alr_{Mtb} mutant Y388D. This suggests that compensatory mechanisms are nec-

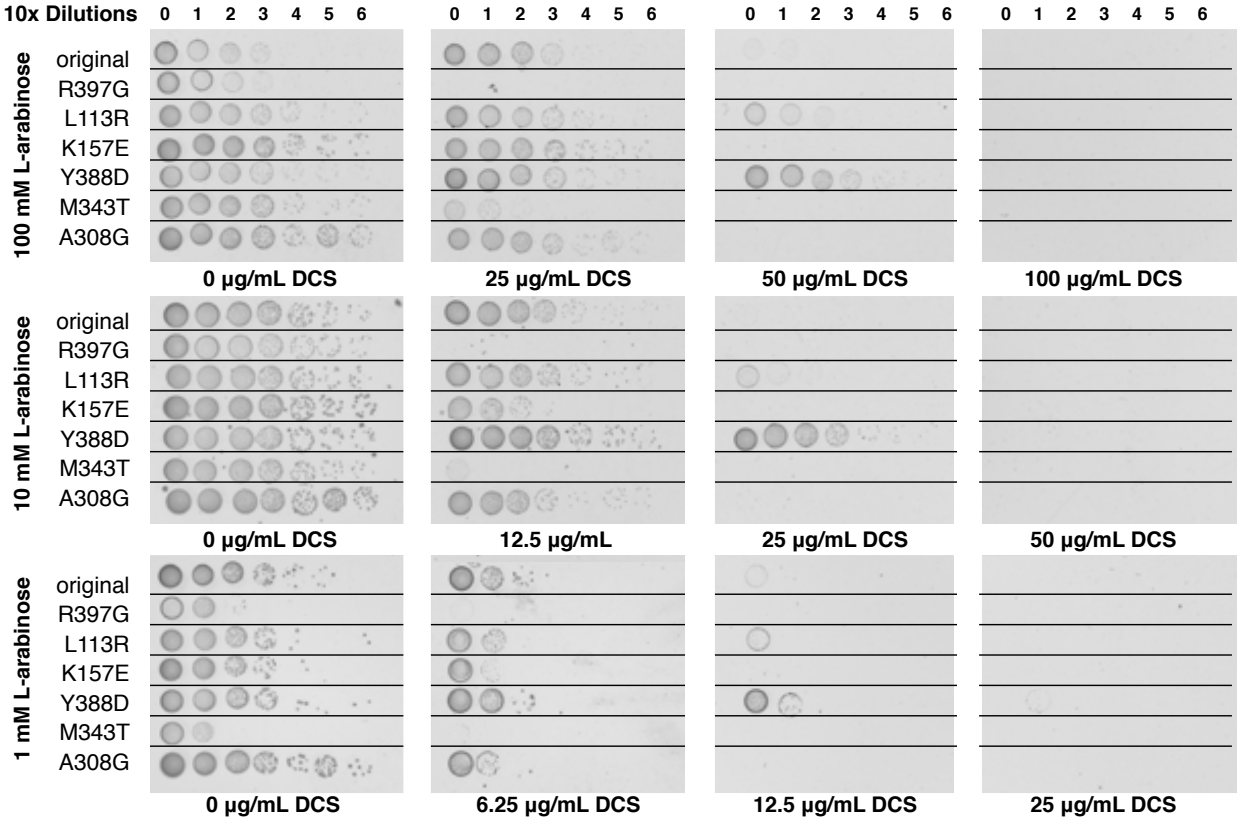

Figure 3.5: Increased D-cycloserine resistance on solid M9 medium for TESEC strains with Alr_{Mtb} L113R and Alr_{Mtb} Y388D expression.
Overnight cultures of *E. coli* strains MUT.0001 (Alr_{Mtb} A380G), MUT.0002 (Alr_{Mtb} M343T), MUT.0003 (Alr_{Mtb}Y388D), MUT.0005 (Alr_{Mtb} K157E), MUT.0006 (Alr_{Mtb} L113R), MUT.0007 (Alr_{Mtb} R397G) and TS008 (control with wild type Alr_{Mtb} expression) were washed and diluted until OD_{600} 0.05. Ten-fold dilution series of the cell suspensions were prepared and 10 μL volumes were transferred onto M9 0.4% fructose agar plates with 1, 10 or 100 mM L-arabinose and a series of DCS concentrations. DCS concentration range was varied per inducer concentration since increased Alr_{Mtb} expression generally led to increased DCS resistance. Photos were taken after 19 hours of incubation at 37°C. Colors are inverted for clarity.

essary to benefit from A308G or L113R mediated resistance, when the mutants are insufficiently expressed.

I also verified the impact of the six non-synonymous mutations on solid M9 medium in order to find plating conditions that discriminate mutants with reduced or increased DCS resistance phenotypes. Such conditions were required to select DCS-resistant mutants from synthetic *E. coli* Alr_{Mtb} mutant libraries.

In line with the observations in liquid medium (Figure 3.4), TESEC-Alr_{Mtb} mutants grew equally well on a M9 agar plate that was supplemented with 10 mM L-arabinose (without DCS). The TESEC with Alr_{Mtb} Y388D expression exhibited superior survival on medium with the highest DCS concentrations. As expected from earlier results (Figure 3.4), Alr_{Mtb} L113R expression provided increased DCS resistance compared to the control strain. However, in this assay, no extra DCS resistance for Alr_{Mtb} A308G expression was detected.

Together, these results provided confidence that it was possible to detect the contribution of Alr_{Mtb} mutations to DCS susceptibility phenotypes. Moreover, information about the importance of the expression level to compensate for fitness loss was obtained as well.

3.3.3 Construction and characterization of two synthetic *E. coli* libraries with Alr_{Mtb} mutant expression

To identify unknown Alr_{Mtb} genotypes that could lead to DCS resistance phenotypes, two synthetic mutant libraries were constructed. The first library was generated using error-prone PCR (epPCR) amplification of the alr_{Mtb} gene. The transformation of the epPCR plasmid library to Dh5-α cells resulted in 8.5x10^5 colonies that carried alr_{Mtb} genes with an unknown number and distribution of mutations. Because a sub-library with all possible single nucleotide substitutions (x=1) was desired for this study, the chance that those variants were present was estimated with the online bioinformatics tool PEDEL [8]. PEDEL calculates the completeness of each sub-library with x number of mutations based on the number of transformants (L), the average number of mutations per *alr* sequence (m) and the number of PCR cycles. For L = 8.5x10^5, sequence length N = 1227 bps, m = 1.43

x mutations[a]	Estimated portion of the library with x mutations (%)[b]	Number of individuals with exactly x mutations[c]	Number of possible sequences with exactly x mutations	Number of different variants[d]	Completeness of the library (%)
0	$26.7x10^{0}$	$2.3x10^{5}$	1	$1.0x10^{0}$	100
1	$32.5x10^{0}$	$2.7x10^{5}$	$3.7x10^{3}$	$3.7x10^{3}$	100
2	$22.4x10^{0}$	$1.9x10^{5}$	$6.8x10^{6}$	$1.9x10^{5}$	2.8
3	$11.4x10^{0}$	$9.6x10^{4}$	$8.3x10^{9}$	$9.6x10^{4}$	$1.2x10^{-3}$
4	$4.7x10^{0}$	$4.0x10^{4}$	$7.6x10^{12}$	$4.0x10^{4}$	$5.2x10^{-7}$
5	$1.7x10^{0}$	$1.4x10^{4}$	$5.6x10^{15}$	$1.4x10^{4}$	$2.5x10^{-10}$
6	$5.2x10^{-3}$	$4.4x10^{3}$	$3.4x10^{18}$	$4.4x10^{3}$	$1.3x10^{-13}$
7	$1.5x10^{-3}$	$1.2x10^{3}$	$1.8x10^{21}$	$1.2x10^{3}$	$6.9x10^{-17}$
8	$3.8x10^{-4}$	$3.2x10^{2}$	$8.2x10^{23}$	$3.2x10^{2}$	$3.9x10^{-20}$
9	$9.1x10^{-5}$	$7.7x10^{1}$	$3.3x10^{26}$	$7.7x10^{1}$	$2.3x10^{-23}$
10	$2.1x10^{-5}$	$1.7x10^{1}$	$1.2x10^{29}$	$1.7x10^{1}$	$1.4x10^{-26}$
11	$4.5x10^{-6}$	$3.7x10^{0}$	$4.0x10^{31}$	$3.7x10^{0}$	$9.2x10^{-30}$
12	$8.8x10^{-7}$	$0.7x10^{0}$	$1.2\ x10^{34}$	$0.7x10^{0}$	$6.1x10^{-33}$
13	$1.7x10^{-7}$	$0.1x10^{0}$	$3.4x10^{36}$	$0.1x10^{0}$	$4.2x10^{-36}$

Table 3.2: epPCR Library composition predicted by PEDEL [8].
[a] The number of mutations per *Mtb alr* sequence.
[b] The estimated percentage of variants with x mutations given a mean number of point mutations of 1.43
[c] The predicted number of sequences in epPCR library with exactly x mutations.
[d] The estimated number of different sequences in the sub-library comprising sequences with exactly x mutations.

(based on Sanger sequencing of random 127 colonies) and the number of PCR cycles that was used = 30, it was predicted that the sub-library with x=1, was indeed present (Figure 3.6 & Table 3.2). The PEDEL analysis further suggested that approximately 26.7% of the library contained the original *Mtb alr* sequence, that 32.5% contained a single mutation and that 40.9% contained more than one mutation (Table 3.2).

I Sanger sequenced 127 colonies from the epPCR library that grew

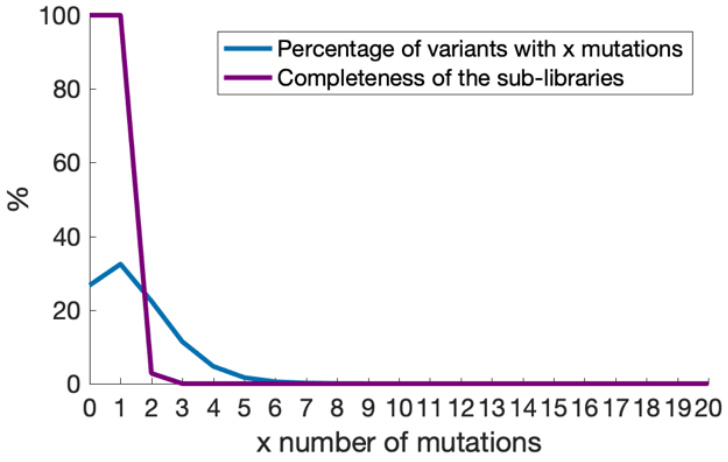

Figure 3.6: Distribution and completeness of sub-libraries within the epPCR Alr$_{Mtb}$ mutant library.
Graphical representation of data in Table 3.2. PEDEL predicted that the sub-library with all possible single nucleotide variants (x=1) was represented as 32.5% of the individuals in the constructed epPCR Alr$_{Mtb}$ mutant library.

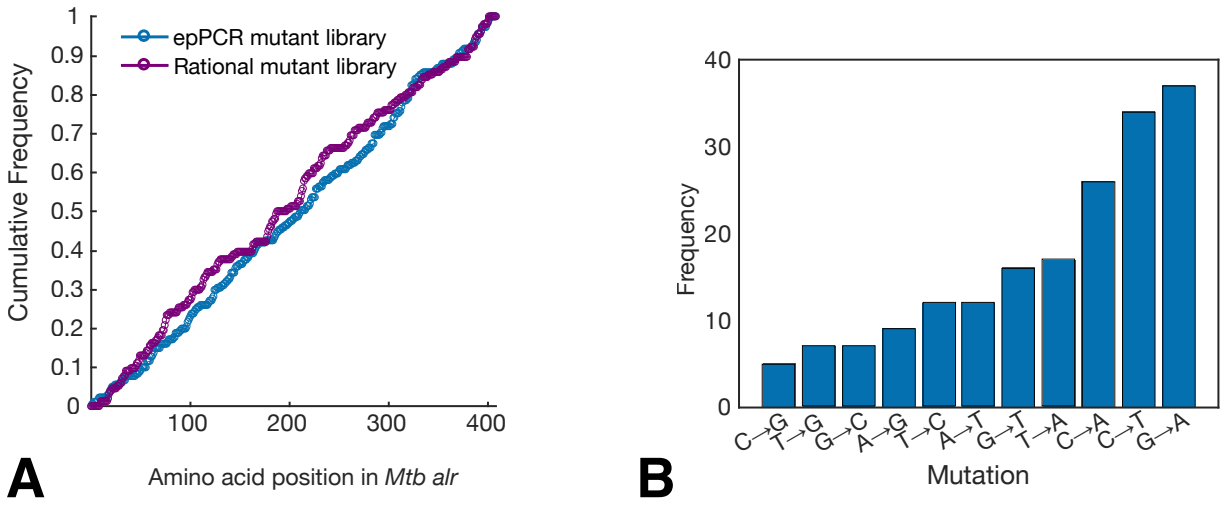

Figure 3.7: Characterization of the distribution of mutations within the epPCR and rational saturation Alr$_{Mtb}$ mutant libraries.
A random selection colonies of the epPCR library and the Rational Saturation library that grew on LB plates with D-alanine were Sanger sequenced for the identification of mutations within the *Mtb alr* gene. **A. Uniform distribution of mutations within the *Mtb alr* gene fragments in the epPCR and Rational Saturation library.** The amino acid positions 182 mutations that were found in the epPCR library and 154 mutations from the Rational saturation library were uniformly distributed over the *Mtb alr* gene. **B. Mutational bias of the epPCR mutant library.** Sanger sequencing of 182 mutations revealed unequal frequencies for different DNA nucleotide substitutions.

on non-selective plates with D-alanine to obtain additional information about the distribution of the mutations within the *Mtb alr* sequence

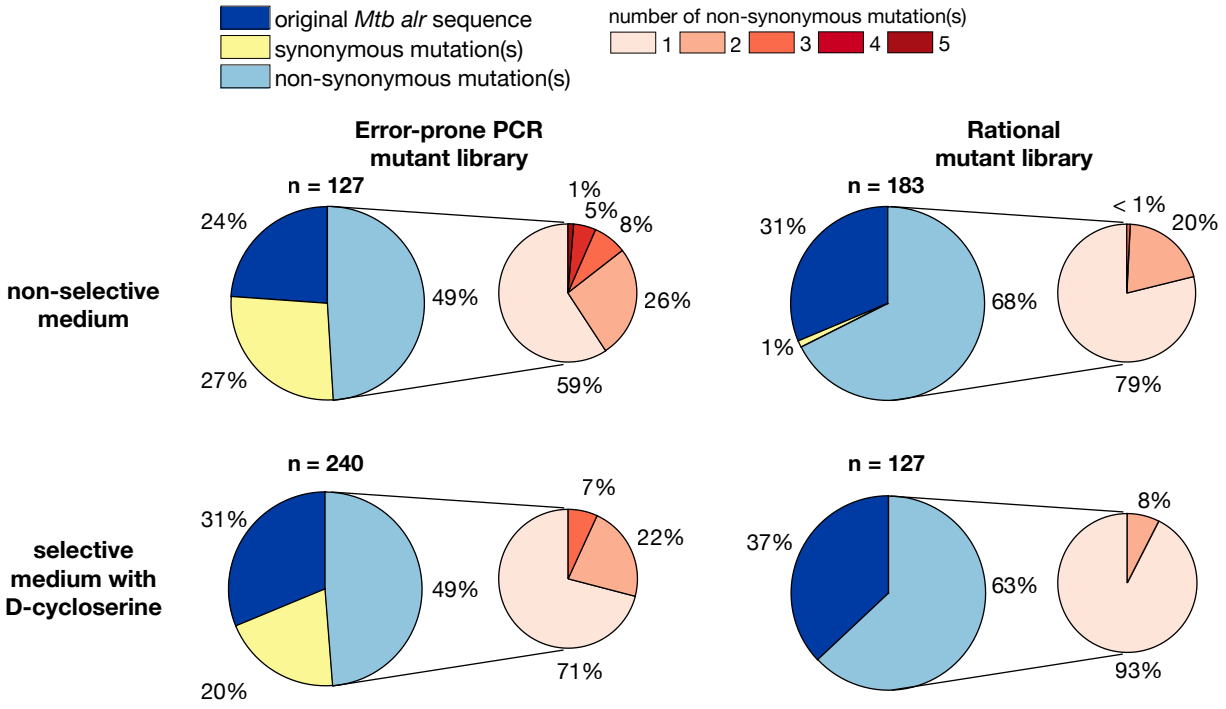

Figure 3.8: Distribution of mutations in the epPCR and rational saturation Alr_{Mtb} mutant libraries before and after selection on DCS containing medium.
Colors represent the percentage of colonies with the original *Mtb alr* sequence (dark blue), colonies with one or more synonymous mutations (medium blue) and at least one or more non-synonymous mutations (light blue). The distribution of non-synonymous mutations per colony are represented in the orange shaded pie chart. **Mutations in non-selective medium.** Colonies the epPCR library and the rational saturation library were randomly picked from LB agar plates with 100 μg/mL D-alanine for plasmid extraction. The alr_{Mtb} genes of 127 epPCR library colonies and 183 rational saturation library colonies were verified by Sanger sequencing. **Mutations in selective medium with DCS.** 240 colonies from the epPCR library (that either grew on LB agar plates supplemented with 100 mM L-arabinose and >45 μg/mL DCS or M9 0.4% fructose medium with 10 mM L-arabinose and 50 μg/mL DCS were analysed) and 127 colonies of the rational saturation library (from M9 0.4% fructose agar plates with 10 mM L-arabinose and 50 μg/mL DCS) were picked. All colonies were miniprepped whereafter Sanger sequencing was used to identify mutations in alr_{Mtb}.

and the potential mutational bias related to the method of error-prone PCR. Coherent with results from the PEDEL analysis, 24% of the *alr* sequences had no mutations (Figure 3.8). Importantly, 49% of the sequenced epPCR colonies had at least one non-synonymous mutation and no more than five non-synonymous mutations.

Sanger sequencing data also provided information about the dis-

tribution of the mutations within the *Mtb alr* sequence and the potential mutational bias related to the method of epPCR. While the verified mutations were homogeneously distributed over the *alr* gene (Figure 3.7.A), the types of nucleotide substitutions varied in frequencies (Figure 3.7.B). According to the supplier's manual, Mutazyme II DNA polymerase has similar substitution frequencies for A→N, T→N (50.7%) as for G→N, C→N (43.8%), but varying preferences for pyrimidine or purine changes [264]:

- 28.5% A→T or T→A
- 25.5% G→A or C→T
- 17.5% A→G or T→C
- 14.1% G→T or C→A
- 4.7% A→C or T→G
- 4.1% G→C or C→G

This explains the relative high mutation frequencies for e.g. G→A and the low occurrence of C→G in the epPCR library. Notably, out of 182 mutations, not a single A→C mutation was identified. As a result, it was possible that not all single nucleotide variations were present and that certain substitutions were more common than others.

Because the exact composition of an epPCR library is difficult to control, I constructed an additional Alr_{Mtb} mutant library based on *de novo* DNA synthesis. This library was rationally designed to cover every single codon mutation of *Mtb* alanine racemase gene to overcome the limitation of mutational bias and to guarantee the presence of all non-synonymous mutations of interest. As for the epPCR mutant

library, Sanger sequencing analysis confirmed that the spread of mutations was equally distributed over the *Mtb alr* gene (Figure 3.7.A). However, sequencing data also showed that a significant percentage (31%) of the rational saturation library contained the original *Mtb alr* gene. Compared to the epPCR library, a more significant 68% of the library contained at least one non-synonymous mutation, of which 79% percent carried a single non-synonymous mutation. I estimated the completeness of the rational saturation library by calculating the degeneracy of the library *D*. The degeneracy is defined the number of unique individuals within the library [265] *D* that can be calculated as follows:

$$D = D_{\max}(1 - e^{-T/D_{\max}})$$

T is the number of transformants and $D_{\max}$ is the maximal number of possible diverse individuals. Under the assumption that the maximal number of variants in the rational saturation library should theoretically be 7753, and given 9.7x10^5 transformants, D should encompass the maximum number of possible variants. However, it should be mentioned that this calculation does not provide any information concerning the diversity and frequencies of each variant in the library.

Based on the estimations above, we assumed that the two libraries bore sufficient variation to cover most if not the entire spectrum of possible non-synonymous mutations in Alr_{Mtb}.

3.3.4 *E. coli*-based screening for DCS resistant Alr_{Mtb} mutants & establishing screening conditions

To identify Alr_{Mtb} mutants that generate *E. coli* cells with increased DCS resistance, exponential phase cell cultures of synthetic Alr_{Mtb} mu-

tant libraries were plated on selective agar medium with L-arabinose and DCS. We performed two resistance screenings: one with the ep-PCR library on LB medium with 100 mM and a series of 0-100 μg/mL DCS concentrations, and a second with both mutant libraries on M9 0.4%fructose medium supplemented with 10 mM L-arabinose and 0-50 μg/mL DCS concentrations.

Strain TS008 with wild type Alr_{Mtb} expression was included as negative control in every resistance screen.

Before plating on selective media, 50 mL bacterial cultures were grown until early exponential phase ($\text{OD}_{600} \approx$ 0.2-0.3) in either LB or M9 0.4% fructose medium supplemented with D-alanine. Glycerol stocks with the mutant libraries were diluted in 50 mL for an initial OD_{600} of 0.05 (for $>10^8$ cells, $\sim 10^2$x the mutant library size).

Because mutations in the D-cycloserine/D-alanine transporter CycA can result in acquired *E. coli* resistance for D-cycloserine [266, 267], I aimed to minimize the selection for spontaneous *cycA* mutants by pre-culturing the bacteria in the presence of D-alanine and in absence of L-arabinose. In this manner, mutants required a functional CycA transporter to benefit from D-alanine in the medium for growth.

In addition, cultures were not grown for more than three generations to minimize competition between mutant strains as a result of to potential fitness differences (due to leaky Alr expression, Figure 3.2) in medium without D-cycloserine.

After plating on selective medium, the MIC was defined as the DCS plate concentration at which a sharp decline in colony numbers was ob-

served. It was expected that the number of surviving colonies on plates with DCS concentrations above the MIC would be greater for mutant library colonies than the control strain. However, several screening attempts on LB medium resulted in comparable colony counts for both the control strain (TS008) and the mutant libraries.

I expected that the use of LB medium was not optimal when I found reports that the D-alanine/D-cycloserine transporter CycA is repressed in rich medium [268, 267]. This could therefore offer natural *E. coli* resistance to the drug. In line with the literature, I confirmed that D-alanine import is less repressed in M9 0.4% fructose medium compared to LB for concentrations <50 μg/mL, but not for D-alanine concentrations >50 μg/mL (Appendix Figure C.1). While it remained unclear whether reduced CycA expression could explain comparable colony counts for the control strain, we picked colonies that survived on plates with $\geq$ 45 μg/mL DCS to check what Alr_{Mtb} mutations were present.

To establish improved screening conditions, I used the finding that six TESEC strains with different Alr_{Mtb} mutations grew comparably well on M9 0.4% fructose agar plates supplemented with 10 mM L-arabinose (Figure 3.5). Importantly, DCS-resistant *E. coli* strains with Alr_{Mtb} Y388D and L113R were able to survive on plates with 50 μg/mL DCS, whereas the control strain TS008 could not (Figure 3.5). Hence, I performed the screen at this induction level to maximize fitness compensation for mutants of interest and eliminate growth of the control strain on >MIC concentrations of D-cycloserine.

As expected (based on earlier results (Figure 3.5), plates with M9

medium containing ≤25 μg/mL DCS and 10 mM L-arabinose were covered with bacteria. A sharp decline in colony numbers was observed for plates with 50 μg/mL DCS. However, colony counts for the TS008 control, epPCR mutant library and rational mutant library were still comparable and respectively: 156, 184, 130.

Additional control experiments should still be performed to improve our understanding of why this occurs. The anwser may lie in the diverse ways *E. coli* could become resistant to DCS [266, 267]. Despite this, we decided to pursue the analysis of colonies that survived on DCS concentrations that were detrimental for most bacteria, and we planned followup experiments to discriminate between DCS resistance contributions of Alr_{Mtb} mutants versus possible background mutations.

3.3.5 Alr_{Mtb} mutations in *E. coli* strains with increased DCS resistance

All colonies that grew on >MIC concentrations of DCS were isolated and miniprepped to identify the potential presence of mutations in the alr_{Mtb} region. Sanger sequencing data of insufficient quality was discarded from the data set. A total of 96 colonies of the control strain TS008, 240 colonies from the epPCR mutant library, and 127 colonies from the rational saturation library were analyzed for the presence of mutations in the p_{BAD} promoter, RBS and/or alr_{Mtb} sequence. The type (synonymous or non-synonymous) and number of mutations per colony was determined (Figure 3.8). Colonies (from non-selective conditions) that did not require functional Alr_{Mtb} expression for their survival occasionally carried truncated alr_{Mtb} sequences or genes with stop codons. Such modifications were absent in colonies isolated from selective conditions. In general, no mutations in the promoter region nor indels, insertions or stop codons in the coding region of alr_{Mtb} were

found in colonies with increased DCS resistance. All 96 DCS resistant colonies of the TS008 control were absent of mutations in the coding region of alr_{Mtb}. We did not investigate (e.g. by whole genome sequencing) whether or what *E. coli* background mutations could instead have been responsible for the DCS resistant phenotypes.

Sequencing data further disclosed that about one-third of the selected DCS resistant library colonies carried wild type alr_{Mtb} genes and that the majority had at least one non-synonymous mutation (Figure 3.8). When we compared the sequencing data from colonies that were selected in the presence of DCS with the data from colonies that were picked from non-selective medium, fractions of cells expressing more than one non-synonymous mutation was drastically curtailed. This trend could be explained by the destabilizing effect of a typical mutation. Tawfik *et al.* found that the probability that a single mutation is detrimental for the protein function is ∼36% on average [269]. As a result, the accumulation of mutations exponentially decreases the protein fitness (W) following a relation that can be expressed as $W \approx e^{-0.36n}$ [269], in which 'n' is the number of mutations. Consequently, on average, 1, 2, 3, 4 or 5 protein mutations reduce the relative protein fitness to respectively 70%, 49%, 34%, 24% and 16%. Hence, Alr_{Mtb} mutants with more than one non-synonymous mutation are often not viable.

3.3.6 Selection bias of mutations with high initial frequencies

Next, I analysed the abundance of each Alr_{Mtb} mutation. Because the resulting Alr_{Mtb} mutation frequencies after selection could have been influenced by the initial composition of the mutant library, we

used Next-Generation Sequencing (NGS) of the epPCR library to determine the initial mutational frequencies (Appendix C.2) (NGS data to determine the initial mutation frequencies within the rational saturation library is pending).

I observed that 60 out of 194 unique (synonymous and non-synonymous) mutations in epPCR library colonies from selective plates had initial frequencies ≥0.01%. I also compared which mutations were both present on non-selective and selective media (Appendix Table C.2). For verified colonies of the epPCR library, 30 Alr_{Mtb} mutations were identified on non-selective as well as selective conditions. Thirteen of those thirty mutations were identified by NGS as having initial frequency ≥0.1% in the epPCR library. In general, 20% of the Alr_{Mtb} mutations from randomly verified epPCR library colonies that grew on non-selective plates were present with an initial frequency ≥0.1%.

For the verified colonies of the rational saturation library, only three mutations were both isolated from colonies on non-selective and selective conditions (Appendix Table C.2).

These comparisons indicated that the initial library composition did influence the selection of mutations and signalled the possibility that background mutations in the *E. coli* genome may have provided certain strains with the ability to survive.

At the same time, the majority of mutations were not detected before selection on >MIC DCS concentrations. Moreover, non-synonymous mutations that were found in both non-selective and selective conditions did not appear more than twice in DCS resistant colonies (Fig-

ure 3.9). Importantly, mutations Y388D, Y388C and D344N with the highest frequencies were not found in non-selective conditions (Figure 3.9 & Appendix Table C.2 % Table C.2). Finally, while NGS data revealed 13 Alr_{Mtb} mutations with an initial frequency of ≥1% in the library (Appendix Table C.5), solely two of them were recovered once from DCS resistant epPCR library colonies. These findings suggested that not all isolated DCS resistant colonies were the result of background mutations and/or the initial mutational frequencies and that certain mutations in Alr_{Mtb} were indeed selected for their assistance in the DCS resistance phenotype.

3.3.7 Non-synonymous mutations with the highest frequencies (≥4) after selection

Non-synonymous mutations with the highest frequencies Y388D, Y388C and D344N (Figure 3.9) were previously found in *Mtb* clinical isolates (Figure 3.10 & Appendix C.1). Strikingly, the mutation with the highest selection frequency was Alr_{Mtb} Y388D (Figure 3.9) that we characterized as significantly resistant to DCS in earlier experiments (Figure 3.4 & 3.5). Non-synonymous L266M was only found in combination with the previously reported D344N mutation. This suggested that their frequent selection could have been co-dependent on the presence of D344N.

3.3.8 Similarities between Alr_{Mtb} mutations of DCS resistant *E. coli* strains and *Mtb* clinical isolates

In total, five previously reported non-synonymous mutations (Y338D, Y338C, D344N, L113R and P311S) were identified in DCS resistant *E. coli* mutants (Figure 3.10 - right Venn Diagram). Like Y388D, L113R was validated earlier (Figure 2.4). Unlike high frequency mu-

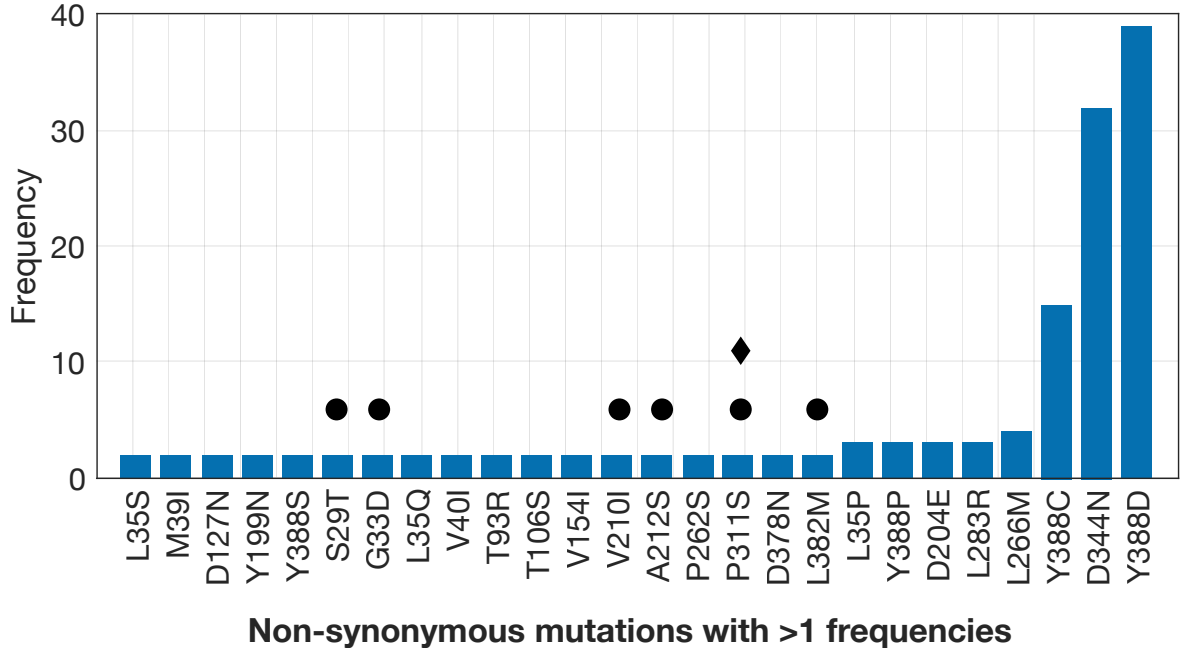

Figure 3.9: Non-synonymous Alr_{Mtb} mutations that were selected with frequencies >1.

377 colonies from the epPCR and rational saturation libraries that survived >MIC DCS concentrations were Sanger sequencing and analysed for the presence of mutations in alr_{Mtb}. The bar plot provides a ranking of the mutation frequencies. Bullet points (•) indicate that the mutation had an initial frequency $\geq 0.01\%$ within the epPCR library. Diamonds (♦) indicate that the mutation was also found in colonies that were isolated from non-selective plates.

tations Y338D, Y388C and D344N, L113R was only selected once and P311S was selected twice. Although all five non-synonymous mutations resulted from the different libraries and conditions, they were all recovered from epPCR colonies that grew on LB medium with 100 mM L-arabinose and high DCS concentrations $\geq 90\mu$g/mL.

DCS resistant *E. coli* mutants also expressed Alr_{Mtb} with alternative amino acids at positions that were also modified in *Mtb* clinical isolates (Figure 3.10) - left Venn Diagram). 23% percent of the reported positions with non-synonymous mutations were also found in DCS resistant *E. coli* mutants.

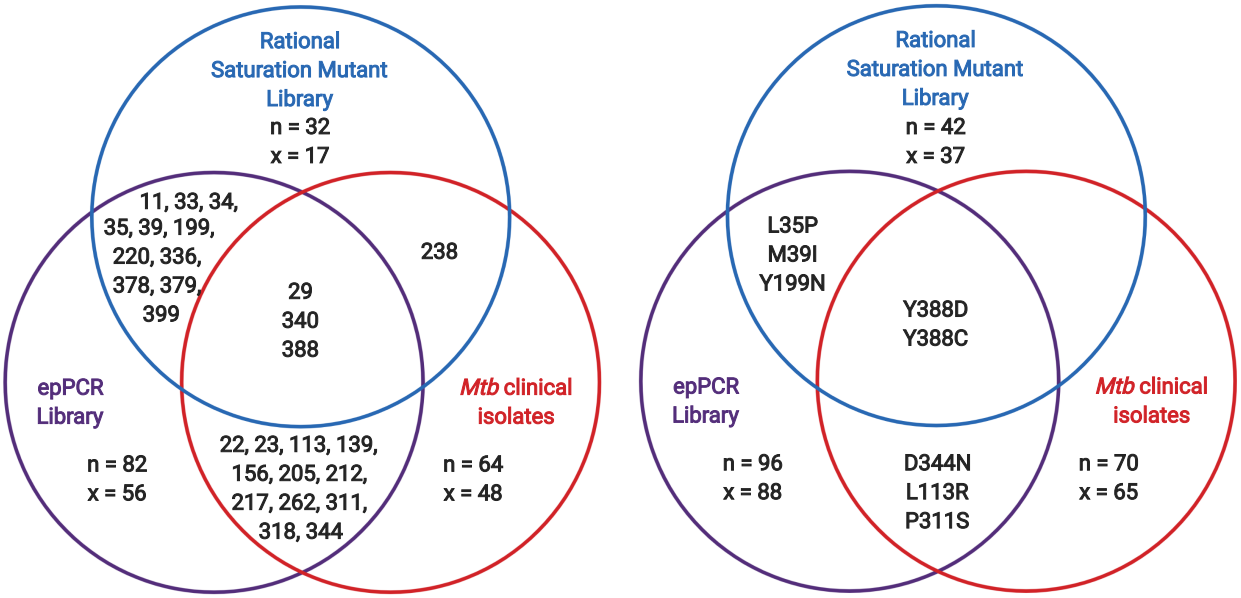

Figure 3.10: Common Alr_{Mtb} amino acid positions of non-synonymous mutations (left Venn diagram) and common non-synonymous mutations (right Venn diagram) found in DCS resistant isolates of *E. coli* Alr_{Mtb} mutant libraries and reported in clinical *Mtb* isolates.
Left Venn Diagram. The numbers in the Venn Diagram represent the location of amino acids within Alr_{Mtb} with non-synonymous mutation. E.g. at least one individual from the rational saturation mutation library and the a reported non-synonymous mutation Alr_{Mtb} were found at the 238^{th} amino acid. However, the amino acid change was different. 'n' represents the total number of unique positions of non-synonymous mutations. 'x' represents the number of non-synonymous mutation positions that were only found in either the epPCR or rational mutant library, or in *Mtb* clinical isolates. **Right Venn Diagram.** The overlapping circles in the Venn diagram show which non-synonymous mutations were shared between DCS resistant members of the rational saturation library (light blue), the epPCR library (dark blue) and the non-synonymous mutations that were previously reported in *Mtb* clinical isolates (Appendix C.1). 'n' represents the total number of unique non-synonymous mutations. 'x' indicates the number of non-synonymous mutations that were only found in a specific mutant library or only in *Mtb* clinical isolates.

3.3.9 Mutation 'hot spots' and diversity

I found that certain positions within the Alr_{Mtb} enzyme were more frequently mutated than others (Figure 3.11). Moreover, some amino acids were substituted by more than one non-synonymous mutation. Mutations at locations 11, 24, 29, 33, 35, 185, 210, 311, 379 and 388 displayed the highest diversity of amino acids substitutions (Table 3.3). Interestingly, except for P311S, Y388D and Y388C, none of the other

non-synonymous mutations in (Table 3.3) have been reported in *Mtb* clinical isolates to date. While two non-synonymous mutations were found at position 311, including the previously reported Alr_{Mtb} P311S [4]. The other previously reported non-synonymous mutation at this amino acid position, P311L [270, 5], was not identified in this study.

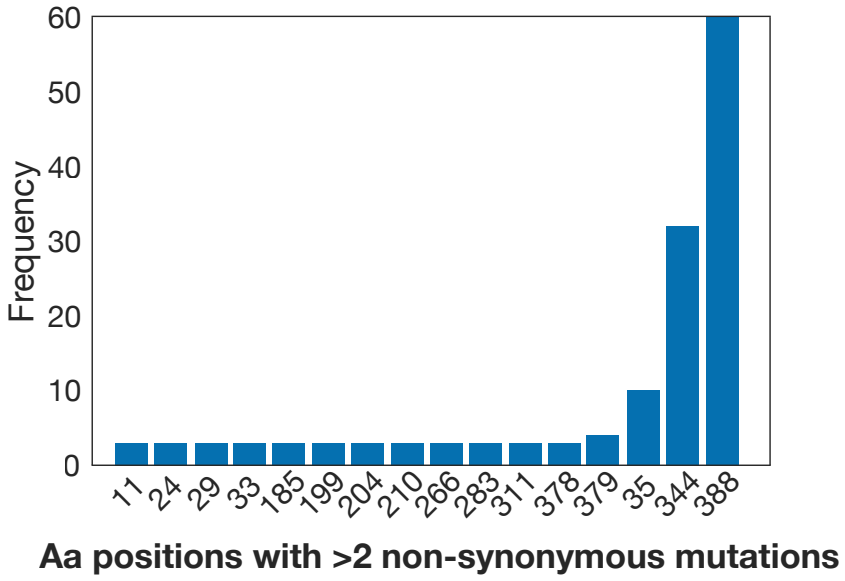

Figure 3.11: Mutation 'hot spots' in Alr_{Mtb}. Amino acid positions at which more than two non-synonymous mutation were identified.

3.3.10 Preliminary results of Alr_{Mtb} mutant validations

Plasmids carrying candidate DCS resistance mutations Alr_{Mtb} Y388C (previously reported) and P311R together with D139E and V257L (mutants P311L and P311S were previously reported) [4, 270, 5]), were retransformed into the original background strain to verify whether the DCS resistance phenotype could be reproduced. An arbitrary selection of plasmids from non-selected conditions with mutated Alr_{Mtb} genotypes (Q30P, L250M, T14R) was also retransformed and served as controls with 'low probability' to confer resistance. MUT.0003 with Alr_{Mtb} Y388 expression served as positive control. After, the

Original aa	**P11**	**A24**	**S29**	**G33D**	**L35**
Non- synonymous mutations	P11S (1x) P11Q (1x) P11H (1x)	A24S (1x) A24V (1x) A24E (1x)	S29T (2x) S29H (1x)	G33D (2x) G33E (1x)	L35P (3x) L35S (2x) L35Q (2x) L35D (1x) L35R (1x) L35F (1x)
Original aa	**E185**	**V210**	**P311**	**W379**	**Y388**
Non- synonymous mutations	E185S (1x) E185I (1x) E185D (1x)	V210I (2x) V210A (1x)	P311S (2x) P311R (1x)	W379Y (1x) W379L (1x) W379C (1x)	Y388D (39x) Y388C (10x) Y388P (3x) Y388S (2x) Y388N (1x)

Table 3.3: Positions in Alr_{Mtb} that were replaced by ≥2 alternative amino acids. Non-synonymous mutations that were found at positions 11, 24, 29, 33, 35, 185, 210, 311, 379 and 388 within Alr_{Mtb}. More details about the colonies with these mutations can be found in Appendix Table C.4.

strains were exposed to a range of L-arabinose and DCS concentrations and their relative DCS resistance was determined compared to control strain TS008 with wild type Alr_{Mtb}.

As expected, the all 'low probability' control mutants did not display signs of increased DCS resistance (Figure 3.12). The average fold differences in DCS IC_{50} values (Figure 3.12) over the four L-arabinose concentrations between the candidate mutants MUT.0103 (Alr_{Mtb} P311R, D139E, V257L), MUT.0090 (Alr_{Mtb} Y388C), MUT.0003 (Alr_{Mtb} Y388D) and control were respectively 1.7, 2.0, 3.2. Hence, using an additional validation round, we were able to confirm that mutations Alr_{Mtb} Y388C and Alr_{Mtb} P311R, D139E, V257L were contributors to *E. coli*'s DCS resistance. Because P311S and P311L were identified in *Mtb* clinical isolates [4, 270, 5], we will further investigate whether the presence of P311R alone is sufficient for the resistance phenotype.

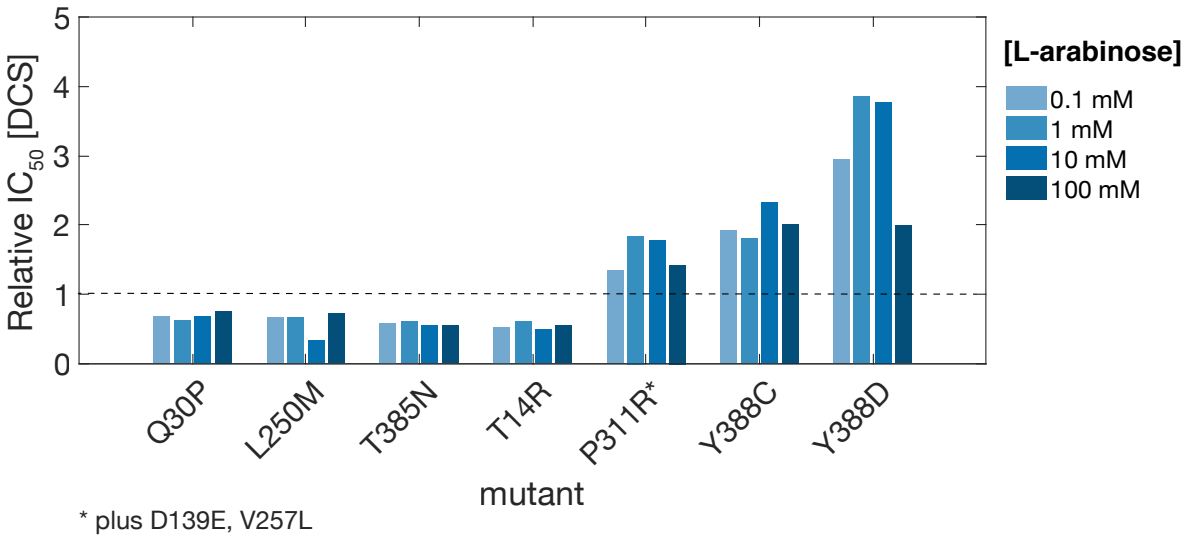

Figure 3.12: Re-evaluation of candidate DCS resistance mutations.
90 μL cultures of *E. coli* strains MUT.0048 (Alr_{Mtb} Q30P), MUT.0049 (Alr_{Mtb} L250M), MUT.0050 (Alr_{Mtb} T385N), MUT.0051 (Alr_{Mtb} T14R), MUT.0080 (Alr_{Mtb} S22T), MUT.0103 (Alr_{Mtb} (P311R, D139E, V257L), MUT.0090 (Alr_{Mtb} Y388C), MUT.0003 (Alr_{Mtb} Y388D) and TS008 (control with wild type Alr_{Mtb} expression) were grown in M9 medium supplemented with 0.1, 1, 10, or 100 mM L-arabinose concentrations and 12 DCS concentrations ranging from 0 to 400 μg/mL. OD_{600} measurements at 10 hours of cultivation at 37°C were used to generate growth-inhibition profiles for each strain at each inducer level. Four-parameter Hill equations were fitted on the inhibition profiles to determine the DCS concentration at which growth was inhibited for 50% (IC_{50}). IC_{50}-values for each strain expressing a Alr_{Mtb} mutant were normalized for the IC_{50} of the control strain TS008. The relative IC_{50} values therefore represent the fold increase or reduction in DCS resistance compared to the control strain. Bar plots are the mean values of two technical replicates.

3.3.11 Molecular dynamics simulations revealed the resistance mechanism of the Y388D mutation

In collaboration with the University of Edinburgh, atomistic simulations of *Mtb* alanine racemase and its interaction with D-cyloserine and L-alanine were performed by Dr. Martin Carballo Pacheco to investigate the influence of Alr_{Mtb} mutations (L113R, K157E, A308G, M343T, R379G, Y388D, Y388C).

We considered that DCS is located inside the binding pocket when the distance is under 1.4 nm and that DCS is in the vicinity of its active conformation when the distance between DCS amine group and the PLP aldimine carbon (PLP is bound to alanine racemase) is less

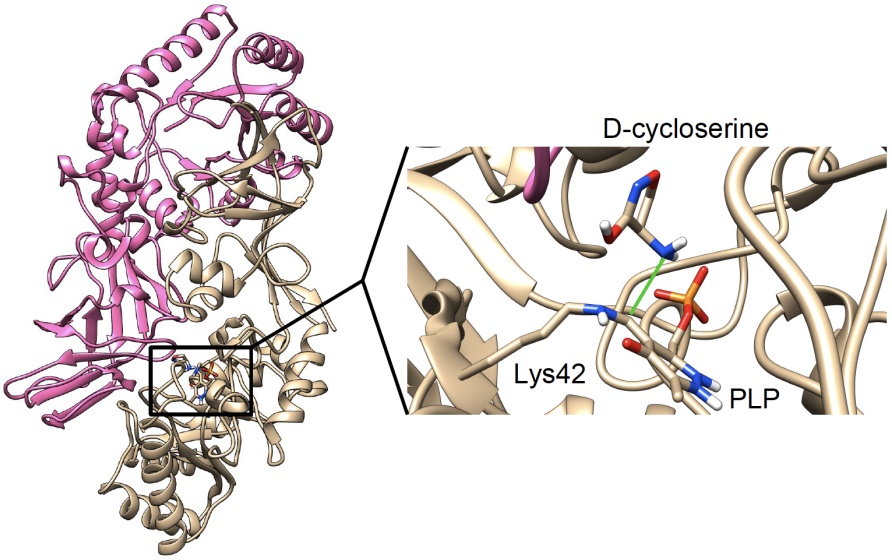

Figure 3.13: Structure of Mycobacterium tuberculosis alanine racemase (chain A in pink and chain B in light brown).
On the right, the binding pocket of alanine racemase with a bound DCS is plotted. PLP is covalently bound to Lys42 and DCS is located so as to minimize the distance between its amine group and the PLP aldimine carbon (see green line) for the nucleophilic attack which is needed for the inhibition reaction between DCS and alanine racemase.

that 0.7 nm. These assumptions were based on the hypothesis that resistant mutants will force DCS to be far from its active conformation, while sensitive mutants will allow DCS to stay close to the active conformation. As observed in (Figure 3.14), the simulations tend to support this hypothesis. Mutants with a significant increase in DCS IC_{50} as observed in *E. coli* 3.4, in general do not allow DCS to stay in its active conformation (in particular Y388D and L113R). On the other hand, sensitive mutants, with decreased DCS IC_{50}, allow D-cycloserine to stay in the active conformation (e.g., M343T or the wild type).

Calculations to determine the location of DCS inside the binding pocket during the entirety of the simulations provided new insights into the possible resistance mechanisms caused by the different mutations. For mutants which are DCS sensitive, such as the wild type and Alr_{Mtb} M343T, it appears that D-cycloserine stays in proximity to PLP during the entirety of the simulation. However, for mutants

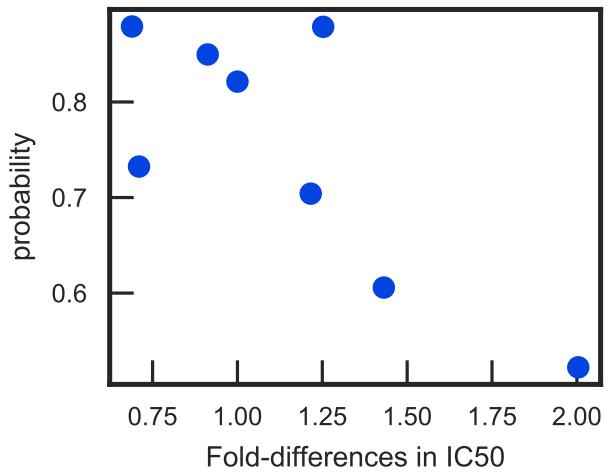

Figure 3.14: Probability that D-cycloserine stays close to its active conformation when inside the binding pocket for different mutants as a function of the experimentally measured relative IC_{50}.
The probability of DCS staying in the active position is defined as the simulated time the distance between the amine group in D-cycloserine and the PLP aldimine carbon is under 0.7 nm, divided by the simulated time the same distance is under 1.4 nm. All simulations are up to 1 microsecond or when DCS leaves the binding pocket. The fold-difference in DCS IC_{50} was determined using *E. coli* strains with Alr_{Mtb} mutant expression. Resistant mutants appear to force D-cycloserine outside of the active conformations while sensitive mutants allow DCS to stay close to the its active conformation.

such as Y388D or L113R, which were validated as resistant in *E. coli*, there appear to be several alternative configurations. One that has the D-cycloserine close to PLP and others of which do not. It appears that the stabilization of these non-catalytic binding positions could be the cause of the resistance. In addition, when we investigated the various binding positions for the two Y388 mutants in more detail, we observed that DCS can be bound mostly in two different positions, one of which is the correct catalytic position while the other is stabilized by the interaction between DCS and the mutant residues C or D. This might provide an explanation for the significant resistance that was observed experimentally.

3.4 Discussion

3.4.1 Establishing an *E. coli*-based approach to analyse the potential contribution of alr_{Mtb} genotypes in DCS resistance

This work demonstrates that Alr_{Mtb} mutants can be tested in engineered *E. coli* host-cells to verify their role in bacterial DCS resistance phenotypes. We explored two ways in which TESEC strains with mutant *Mtb* enzymes could be applicable.

I first tested whether Alr mutants M343T, L113R, Y388D, R397G and K157E, which are associated with DCS resistance in *Mtb* [6], could confer resistance in our *E. coli* host strain.

Using our tunable expression system, loss of bacterial fitness due to the introduced protein modification was compensated by precise controlling of the enzyme expression level (Figure 3.3). This way, it was possible to measure the potential resistance contribution of the mutated protein at conditions that did not generally reduce the bacterial fitness.

From the six previously reported Alr_{Mtb} mutants, Alr_{Mtb} L113R and Alr_{Mtb} Y388D provided the most significant increase in DCS resistance in both liquid medium (Figure 3.2) and on plates (Figure 3.5). On the other hand, a reduced susceptibility for the TESEC with Alr_{Mtb} A308G expression was only observed in liquid M9 medium supplemented with 1 mM L-arabinose. These results imply that some mutations confer DCS resistance depending on the conditions.

Noteworthy, Alr_{Mtb} mutants M343T, R397G and K157E did not

provide any additional DCS resistance to our *E. coli* strains. While Alr_{Mtb} M343T in *Mtb* clinical isolates has been associated with a DCS MIC of >60 μg/mL DCS (>4 fold the wild type *Mtb* MIC) [6, 215], increased DCS resistance was not detected for a TESEC strain expressing the same Alr_{Mtb} mutation (Figure 3.4 & 3.5). Moreover, in vitro experiments from Nakatani *et al.* [215] with isolated Alr_{Mtb} mutants reported a 50-fold increase in enzymatic DCS IC_{50} for Alr_{Mtb} Y388D and an immeasurably high DCS IC_{50} for Alr_{Mtb} M343T. While these in vitro results support that the suspected Y388D and M343T mutations are contributing to *Mtb* DCS resistance, the increase in the reported IC_{50} values were ∼10-fold off from the reported MIC differences between the mutant *Mtb* strains and the wild type *Mtb* strains [215]. This difference shows that drawing direct conclusions between enzyme assay data and MIC's of *Mtb* clinical isolates is not evident.

In our approach, single mutations were able to increase the DCS resistance with below 10-fold differences compared to the control strain. This fold difference is in the same order of magnitude as reported for DCS resistant *Mtb* strains (Table 3.1). Interestingly, Nakatani *et al.* also described that a 10-fold higher enzyme concentration was necessary for Alr_{Mtb} M343T to obtain sufficient racemase activity. This is in line with the observation in our study that more inducer was required to rescue bacterial growth of strain MUT.0002 with Alr_{Mtb} M343T expression (Figure 3.2).

Both results thus point at the fitness impact that this mutation may have on the protein activity and suggests that the presence of compensatory mechanisms may be necessary in case M344T leads to *Mtb* DCS resistance. Moreover, while the Y388D mutation was re-

ported in strains that carried a 1 base pair deletion in Ald (Table 3.1), Nakatani *et al.* reported that this mutation was present in *Mtb* strains with both high and low susceptibility for DCS [215]. Again, this in line with our finding that the presence of Y388D alone is sufficient to generate the DCS resistance phenotype.

It is well-understood that the impact of mutations is dependent on a variety of factors such as the genetic context (epistasis), the environment, or the composition and size of the population [271]. Because antibiotic resistance emerges from complex interactions it is possible that certain Alr_{Mtb} mutations are important for DCS resistance in clinical *Mtb* strains, but will not lead to DCS resistance in our model strain. Likewise, the impact of drug-resistance mutations also varies between *Mtb* strains[272]. This co-dependence between resistance mutations and the bacterial genetic background is not well understood [272]. More research is needed to understand how drug resistance varies between strains in order to improve future *Mtb* therapies.

At the same time, when the same mutation provides DCS survival advantages in both *Mtb* as well as in its distant evolutionary relative *E. coli*, one could hypothesize that the impact of this mutation on the drug resistance phenotype is less context-dependent.

3.4.2 Benefits of screening both the epPCR an rational Alr_{Mtb} mutant libraries

Synthetic Alr_{Mtb} mutant expressing *E. coli* libraries were constructed to find potential DCS resistance mutations that are currently unknown. To maximize the chance that all possible single amino acid variants of the protein were present, we used both epPCR and DNA synthesis

to generate modified alr_{Mtb} gene fragments. Based on Sanger sequencing of epPCR library colonies that grew in non-selective conditions and predictions of mutational distributions using bioinformatics, we estimated a 100% completeness for the epPCR sub-library with all possible alr_{Mtb} single nucleotide variants. However, to ensure that amino acid changes were included that require a full codon change, a rational saturation Alr_{Mtb} library was constructed. This allowed for the screening of potential DCS resistance mutations that are more difficult to evolve.

Both libraries complemented each other. The advantage of mutations that were found in the epPCR library was that they frequently displayed unique mutational signatures that provided information about how 'accidental' the presence of that mutation was. E.g. when the same mutation occurs in combination with various other mutations, it could signal that the finding of this Alr_{Mtb} mutant is not the result from an initial high frequency in the population, but is rather selected for it contribution to the survival in presence of DCS. On the other hand, not a single isolate of the epPCR library carried a full codon change. This is why the rational saturation library offered complementary advantages. For example, certain non-synonymous mutations at mutational 'hot spots' in Alr_{Mtb} that required a full codon change were solely isolated from the rational saturation library (Appendix Table C.4). One of them was the Y388P mutation at the often cited amino acid position 388.

3.4.3 Identification of previously described Alr_{Mtb} mutations in our *E. coli* screens

Only five (D334N, L113R, Y338D, Y338C, P311S) out of 70 non-synonymous Alr_{Mtb} mutations that were previously found in *Mtb* clinical isolates (with and without phenotypically confirmed DCS resistance) were also detected in DCS resistant *E. coli* cells from the recombinant mutant libraries of this study. We re-validated three Alr_{Mtb} mutants (L113R, Y338D, Y338C) from those five as being responsible for increased DCS resistance in our *E. coli* host. It should still be verified if the DCS resistance phenotype can also be reproduced when the plasmid carrying alr_{Mtb} D344N or alr_{Mtb} P311S is re-transformed to the original background strain.

3.4.4 Challenges related to screening for *Mtb* drug resistance mutations in *E. coli*

Despite succesfully identifying five non-synonymous mutations and 16 positions of non-synonymous mutations that were previously associated with DCS resistance in *Mtb*, 48 reported positions of non-synonymous mutations in Alr_{Mtb} were not found. A variety of reasons could explain this result. First of all, it is possible that the reported mutations in *Mtb* clinical isolates have no role in DCS resistance as the dataset also included strains with wild type DCS susceptibility or strains for which DCS resistance was not evaluated. A second possibility is that the Alr_{Mtb} mutation is important for *Mtb* DCS resistance, but only in the presence of another compensatory mutation(s) or specific genetic background. Thirdly, certain DCS resistant Alr_{Mtb} mutants may not be viable in *E. coli* due to protein misfolding, the presence of an additional lethal mutation [273] or the lack of *Mtb*-specific chaperones [274]. Another reason is that some resistance mutations

may require a certain expression level that was not provided during the resistance screen (as seen for A308G and L113R in liquid M9 medium - Figure 3.4). To include resistance mutations that depend on the latter scenario, additional resistance screens at alternative L-arabinose concentrations are currently being performed. Moreover, it is possible that the initial library frequency of DCS resistance mutations was too low or even absent. While screening two different mutant libraries should minimize this chance, it remains probable that certain mutations were therefore not identified. Finally, screening for Alr_{Mtb}-specific resistance mutations is challenging in a host that is able to evolve alternative ways to resist DCS [267, 266]. Because spontaneous mutations in *E. coli*'s chromosome could have provided cellular resistance to DCS, we included another round of phenotypic screening for which plasmids with suspected Alr_{Mtb} mutations were retransformed to the orginal background *E. coli*. Although understanding DCS resistance in *E. coli* was not the focus of this study, the genomes of isolated DCS resistant colonies with wild type Alr_{Mtb} expression may still provide a valuable asset to examine whether mutations can be detected in bacterial pathways that have not been described earlier [267, 266].

3.4.5 Alr_{Mtb} positions Y388 and D344

Our current data supports previous conclusions that Y388 and D344 are important positions for Alr_{Mtb} resistance mutations [253, 270] (Figure 3.9). Interestingly, we find evidence that there might be more flexibility at the 388 position than what has been reported so far as *E. coli* strains were isolated with expression of unreported mutations Alr_{Mtb} Y338P, Y388S and Y388N. Because of the diversity at the 388 amino acid position, the influence of all 20 amino acid changes at this

position is currently tested by constructing TESEC strains expressing all non-synonymous variants of Alr_{Mtb} Y388.

Surprisingly, while the mutational frequency of position 344 was second highest, no other variant than D344N was found. This may indicate that less variety is accepted at this position. As for Y388, we will determine the impact of all alternative amino acids at this position in followup experiments.

3.4.6 Molecular dynamics provided new insights into the impact of Y388D

Molecular dynamics simulations allowed us to observe a trend between the experimentally measured IC_{50}'s from *E. coli* strains (Figure 3.4 & 3.12) and the probability that DCS stays close to its catalytically active position (Figure 3.14). However, it was not possible to find a common mechanistic reason for this among all the mutants. In the case of Y388D, it appears to be clear that D388 interacts directly with DCS and removes the drug from its catalytically active position. More work is needed to explain the impact of each Alr_{Mtb} mutation. For the near-future, we hope that the established methods will provide mechanistic insights into why unreported Alr_{Mtb} mutations that were discovered during our *E. coli*-based resistance screening may cause DCS resistance.

3.4.7 Synonymous mutations

This work has primarily focused on the validation and identification of non-synonymous mutations, but further explorations into the contribution of synonymous mutations may provide complementary insights. While synonymous have been falsely coined as 'neutral' or 'silent' in the past [275], research into the effects of synonymous mutations has

led to the understanding that they can significantly impact translation rates by altering protein folding, mRNA stability or introduction of rare codons [276]. Some of the encountered synonymous mutations in study may therefore play an important role in the observed DCS resistance by stabilizing the Alr_{Mtb} mutant protein through the facilitation of the folding process or increase in the protein production. For example, synonymous mutation L271L was found five times. Once alone and three times in combination with D344N. Moreover, synonymous mutations in Alr_{Mtb} are also reported for *Mtb* clinical isolates (Appendix C.1), but their possible role in DCS resistance has not been investigated.

3.5 Conclusion

This is the first study that investigates and predicts DCS resistance alr_{Mtb} genotypes in a non-mycobacterial organism. Because genetic engineering is significantly faster in *E. coli* and since this study shows that at least a subset of Alr_{Mtb}-dependent DCS resistance can be directly translated to *E. coli*, the data generated in this study may hold valuable predictions about how DCS resistance could evolve in *Mtb*. Additional work is still required to gain more confidence about resistance mutations that were identified during the resistance screenings with *E. coli* expressing Alr_{Mtb} library detected. Nevertheless, we are confident that with the current data and strategies we will be able to confirm more currently unknown alr_{Mtb} DCS resistance genotypes in the near future.

3.6 Contributions

For this project, I have been supervised by Dr. Ariel Lindner (CRI, INSERM). Under the supervision of Dr. Rosalind Allen (University of Edinburgh), Dr. Martin Carballo (University of Edinburgh) has supported this work with molecular dynamics simulations. For the experimental part, I mentored 3 second-year MSc. students: Salam Abara (construction and testing of the first mutant strains, epPCR Alr_{Mtb} mutant library), Hanifa Loupido (verification the DCS resistance phenotypes and construction of numerous mutant strains), Ayan Abukar (testing of DCS resistance screenings methods). Xiahou Song (CRI, INSERM) has done the computational analyses to determine growth parameters of *E. coli* strains with Alr_{Mtb} expression. Dr. Jake Wintermute (CRI, INSERM) has provided advice, feedback and protocols to characterize TESEC strains.

3.7 Materials and Methods

3.7.1 Strains, Plasmids & Primers

Lists with strains and plasmids that were used in this study can be found in Appendix C.6 & C.7.

Primer	DNA Sequence (5'→3')	Experimental Purpose	T_m (°C)	nt length
NB.128-F	GGGTTATTGTCTCATGAGCGG A	PCR verification of plasmu	60.2	22
NB.129-R	CTCGCCGCAGCCGAAC	PCR verification of pTS-plasmids	61.3	16
NB.43-F	ATATGGTCTCATACCATGAAA CGCTTTTGGGAAAATGT	Amplification of alr_{Mtb} with BsaI-sites for pTS008 construction	58.3	38
NB.44-R	ATATGGTCTCACGGGTTAGCG GTTTTCAGCTTCGC	Amplification of alr_{Mtb} with BsaI-sites for pTS008 construction	59.5	35

Table 3.4: Forward (-F) and reverse (-R) primers used in this study.
All primers were manufactured by IDT.

3.7.2 Construction of *E. coli* strains expressing Alr_{Mtb} mutants associated with DCS resistance

The original DNA sequence of *Mtb*'s *alr* gene was codon optimized and synthesized by IDT. In addition, six variants of the codon-optimized *Mtb alr* gene were synthesized that carried one of the following non-synonymous mutations: A308G, M343T, Y388D, K157E, L113R or R397G. Those non-synonymous mutations were previously associated with D-cycloserine resistance [6]. Gene fragements were flanked by type IIS BsaI restriction sites and embedded upstream the $_{BAD}$ promoter of the pTS000 plasmid using golden gate assembly. The resulting plasmids were transformed to Dh5-α cells and selected using blue white screening on LB agar plates with 0.2 mg/mL X-gal, 10 mM IPTG, 100 μg/mL ampicillin. Correct plasmid construction was

verified by Sanger sequencing (Eurofins) before transformation to the final destination strain TS072. Transformants were selected on LB agar plates with 100 μg/mL ampicillin, 50 μg/mL ampicillin and 100 μg/mL D-alanine.

3.7.3 Construction of the epPCR Alr_{Mtb} mutant library

error-prone PCR GeneMorph® II Random Mutagenesis Kit (Aligent) was used according to the supplier's instructions to generate point mutations in the *Mtb* alanine racemase gene. Primers NB.43 and NB.44 were used in the epPCR reaction to amplify the *Mtb alr* gene and add BsaI flanking sites and a quantity of 750 ng pTS008 was used as template DNA to achieve a low mutation rate of ∼1-3 nucleotide changes per *Mtb alr* sequence.

epPCR thermocycles:

1. 95°C for 2 min
2. 95°C for 30 sec
3. 55°C for 30 sec
4. 72°C for 1 min 25 sec
5. 30x to step 2
6. 72°C for 10 min
7. 4°C on hold

Golden gate assembly of mutant library plasmids PCR fragments were digested with DpnI (NEB) and purified using Wizard® SV Gel and PCR Clean-Up System (Promega). Next, golden gate assembly [277]

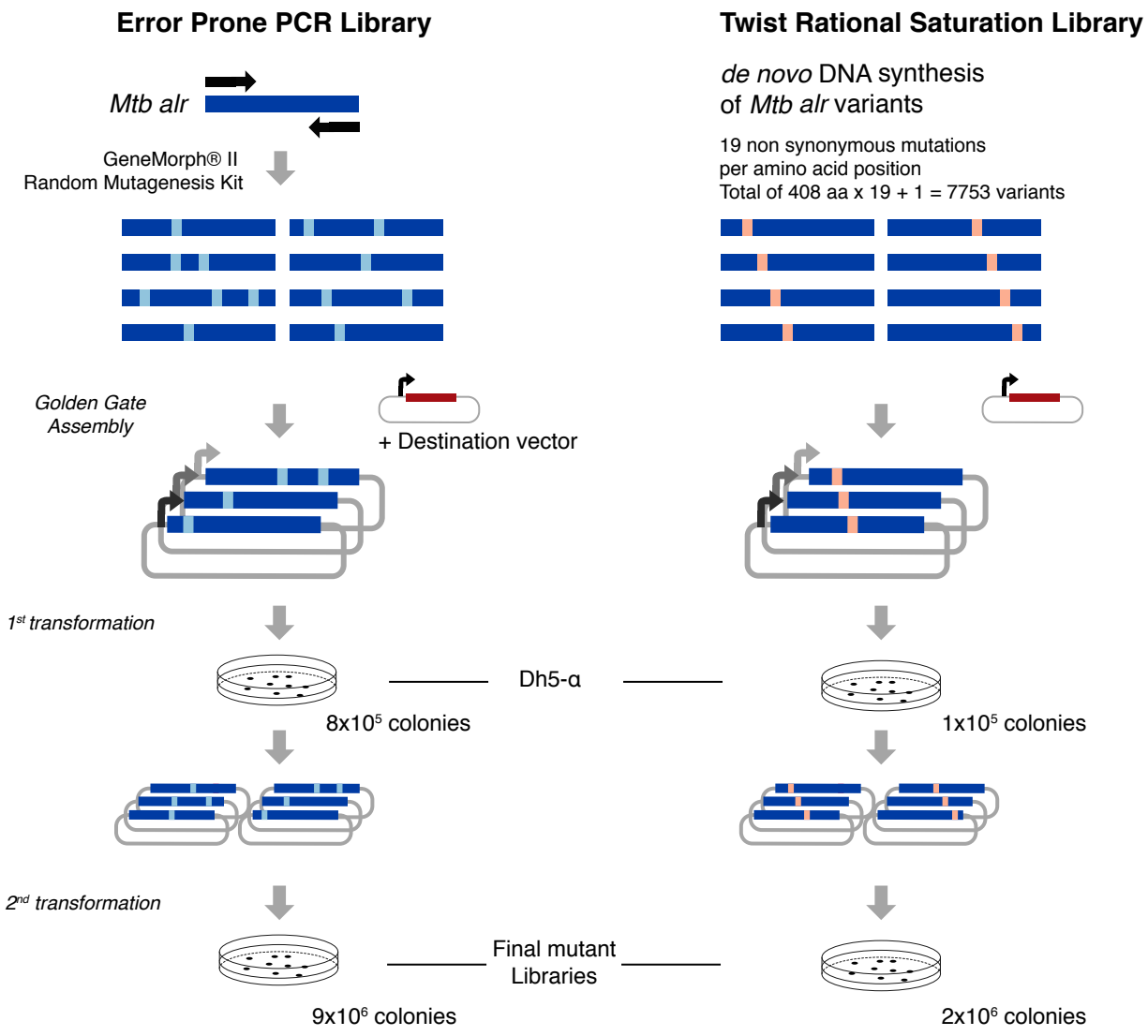

Figure 3.15: Schematic overview of the construction of synthetic libraries expressing *Mtb* alanine racemase mutants.
Error Prone PCR Library (left pane) was constructed using the Gene morph® II Random Mutagenesis Kit to generate random mutations within the *Mtb alr* gene. The Rational Saturation mutant Library (right pane) was rationally designed and synthesized by Twist Bioscience. This library was composed of all possible single non-synonymous amino acid variants of the Alr_{Mtb} enzyme. For both library constructions, the *Mtb alr* gene was flanked with standard Type IIS BsaI restriction sites that allowed for golden gate assembly into a standard backbone plasmid pTS000. The golden gate reaction mix was first transformed to Dh5-α cells from which the plasmid libraries were extracted and re-transformed into the final background strain TS072.

was used to insert the epPCR fragments upstream the *textsubscript-BAD* promoter of the pTS000 destination vector.

10 μL Golden gate reaction cocktails contained:

- 80 fmol vector
- 40 fmol insert(s)
- 0.2 μL T4 ligase (HC) (Promega)
- 1.0 μL T4 ligase Buffer (Promega)
- 0.4 μL BbsI or BsaI-HF (NEB)

Golden Gate thermocycles:

1. 37°C for 2 min
2. 16°C for 5 min
3. 50x back to step 2
4. 50°C for 10 min
5. 80°C for 5 min
6. 4°C on hold

Bacterial transformations Golden ligations were dialysed for 30 minutes against dH_2O, whereafter 15 electrotransformations were perfomed with 4 μL of dialysed golden gate reaction mix and 200 μL electrocompetent Dh5-α cells. Transformed cells were recovered in LB medium for 2 hours at 30°C and plated on selective LB plates with 0.2 mg/mL X-gal, 10 mM IPTG and 100 μg/mL ampicillin. Ten-fold dilution series of the transformation mix were plated in parallel to estimate to total number of colonies. After incubation at 37°C for 24 hours, and estimated $8x10^5$ surviving colonies were collected from the LB plates and combined a conical tube. Less than 0.3% of the transformants were blue, implying that the golden gate reaction was efficient. The

cell suspension was miniprepped using Wizard® Plus SV Minipreps DNA Purification (Promega) to obtain the plasmids carrying the epPCR library. Additionally, library cell suspensions were aliquoted as 15% glycerol stocks that were stored at -80°C.
Finally, the miniprepped epPCR library was retransformed to the final background strain TS072 and plated on LB agar plates with 100 μg/mL ampicillin, 50 μg/mL kanamycin and 100 μg/mL D-alanine. The final epPCR Alr_{Mtb} mutant library was prepared by combining all $\sim$9x10^6 transformants.

3.7.4 Construction of the rational saturation Library Alr_{Mtb} mutant Library

A library of alr_{Mtb} with all possible single non-synonymous amino acid variations was synthesized by Twist Bioscience. The quality of six out of the 408 amino acid positions were below the companies quality threshold (Supplementary material C.3), indicating that these variants could have been absent or underrepresented in the final library. All synthesized DNA fragments were combined in a single tube and used for a round of golden gate cloning into the backbone vector pTS000 as described before (Section 3.7.3). Electrotransformations of the golden gate reactions to Dh5-α cells resulted in an estimated 9.7x10^5 colonies that were collected and stored as aliquots in glycerol. In parallel, the plasmids of the Dh5-α transformants were extracted and transformed to electrocompetent TS072 resulting in 1.9x10^6 colonies. The final Rational Saturation library was aliquoted and stored as glycerol stocks.

3.7.5 Validation of DCS resistant phenotypes of *E. coli* strains expressing six Alr_{Mtb} mutants (A308G, M343T, Y388D, K157E, L113R or R397G).

Single *E. coli* colonies that grew on Luria-Bertani (LB) 1.5% agar plates supplemented with 50 μg/mL kanamycin (Sigma), 100 μg/mL ampicillin (Sigma) and 100 μg/mL D-alanine (Sigma) were inoculated in M9 medium (1x M9 salts, 1 g/L Drop-out Mix Complete, Adenine Rich, w/o Yeast Nitrogen Base (Powder) (USBiological Life Sciences), 2 mM $MgSO_4$ and 0.1 mM $CaCl_2$ and 0.4% (w/v) D-fructose (Sigma) medium 0.03% (w/v) vitamin B1 (Sigma) with 100 μg/mL ampicillin, 50 μg/mL kanamycin and 100 μg/mL D-alanine for overnight growth at 37°C, 150 rpm agitation speed. The next day, cell cultures were washed twice in M9 0.4% fructose medium without D-alanine and diluted for a final initial OD_{600} of 0.05. 90 μL cell cultures were grown 96-well microtiter plates in serial dilutions of 10-fold L-arabinose concentrations (0-100 mM L-arabinose) and two-fold D-cycloserine concentrations (0-400 μg/mL). The TECAN Freedom Evo robot and the Spark® multimode reader were used to monitor bacterial growth by taking OD_{600} measurements for multiple 96-well plates during 20 hours. The 96-well plates were incubated without agitation at 37°C. Inhibition concentrations at which the bacterial growth was reduced by 50% (IC_{50}) were determined by fitting a four-parameter hill-function on OD_{600}-[DCS] curves for 0-100 mM L-arabinose concentrations. The relative resistance was determined by dividing the IC_{50} values of each mutant strain by the IC_{50} value of the TS008 control strain.

3.7.6 Calculation of growth rates

An in-house python script was used to determine the growth rate based on OD_{600} measurements. The background (measured from the

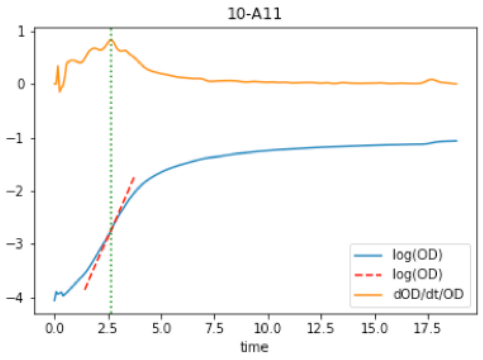

Figure 3.16: Growth curve analysis

medium only) was subtracted from each measurement, and the OD_{600} values were then smoothed by using a moving average filter (size = 5). The relationship of OD and growth rate over time is defined by . The growth rate (μ) is regarded as the slope coefficient by applying a local linear regression (window's size = 5); from this we calculated of the bacterial population. As the μ can be calculated from any two points that we place on the plotted growth curve (Figure 3.16):

$$\mu = lnOD_2 - lnOD_1/t_2 - t_1$$

3.7.7 Screening for Alt_{Mtb} mutants related to DCS resistance.

Glycerol stocks of TS008 (control strain), the epPCR Alr_{Mtb} mutant library and the rational saturation Alr_{Mtb} mutant library were thawed and diluted until OD_{600} 0.05 in 50 mL with M9 0.4% fructose or LB medium, 100 μg/mL D-alanine, 100 μg/mL ampicillin, 50 μg/mL kanamycin. 50 mL cell cultures were grown in Erlenmeyer flasks at 37°C with agitation (750 rpm) until they reached the exponential phase (OD_{600} 0.2-0.3). Next, the cells suspensions were centrifuged for 20 minutes at 4000 rpm and washed with M9 0.4% fructose or LB medium to remove the residual D-alanine. After two wash steps, the cell cultures were concentrated until OD_{600} 1 of which 250 μL was plated on M9 0.4% fructose agar plates with 0, 12.5, 25, or 50

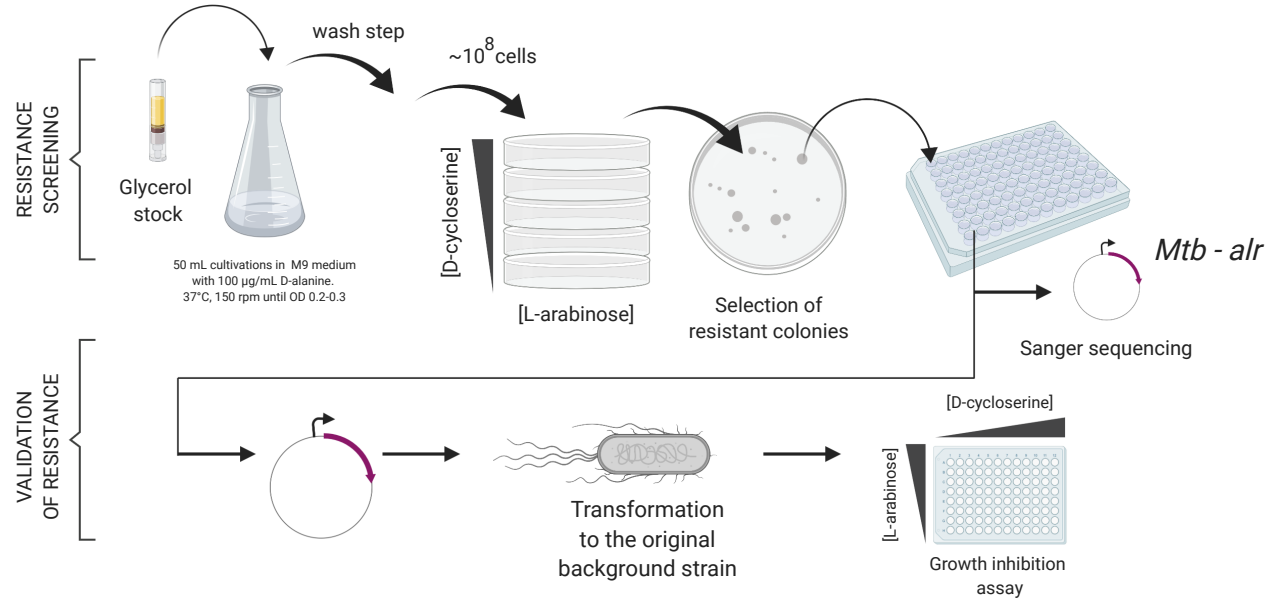

Figure 3.17: Screens and validation assays to identify DCS resistance mutations within *Alr_{Mtb}*.

μg/mL DCS, 100 μg/mL ampicillin, 50 μg/mL kanamycin and 10 mM L-arabinose. For the resistant screen in LB medium, cells suspensions were plated on LB agar medium with 100 mM L-arabinose, 100 μg/mL ampicillin, 50 μg/mL kanamycin, and a range of 0-100 μg/mL DCS concentrations. After overnight incubation at 37°C, all colonies were isolated that grew on the plates with M9 0.4% fructose medium and 50 μg/mL DCS. DCS resistant colonies on LB medium were picked from plates that grew in the presence of >45 μg/mL DCS. The selection of resistant colonies were subsequently inoculated in 100 μL non-selective medium with 100 μg/mL D-alanine, 100 μg/mL ampicilin and 50 μg/mL kanamycin in 96-well plates for overnight incubation at 37°C. Cell cultures were transferred to 96-well plates with LB agar, 100 μg/mL ampicilin and 100 μg/mL D-alanine and sent to Eurofins for downstream plasmid extraction and Sanger sequencing with primers NB.128 and NB.129 to verify the alr_{Mtb} gene and its promoter region. Sequencing results were imported into Geneious and compared with the original reference sequence for the identification of mutations.

3.7.8 Analysis of Sanger sequencing data

Sanger sequencing data was evaluated using Geneious Bioinformatics software. Mutations were manually annotated and stored in Microsoft Excel. MATLAB (Version R2017a, Mathworks) was used to analyse and compare sets of annotated data.

3.7.9 Next-generation sequencing.

The initial composition of the constructed mutant libraries was verified by next-generation sequencing (NGS). Plasmids were extracted from the alr_{Mtb} mutant library suspensions and the alr_{Mtb} region was amplified by Q5 PCR (NEB) as described below.

NEBNext Ultra II Q5® Master Mix:

- 25 μL NEBNext Ultra II Q5 Master Mix (M0544)
- 100 ng of library plasmid DNA
- 5 μL 10 μM forward primer NB.128
- 5 μL 10 μM reverse primer NB.129
- dH_2O to 50 μL

Q5 PCR Thermocycles:

1. 98°C - 30 sec
2. 98°C - 10 sec
3. 55°C - 30 sec
4. 65°C - 45 sec

5. 10 cycles to step 2
6. 65°C - 5 min
7. 4°C - hold

The amplified alr_{Mtb} fragments were analyzed by Illumina Nextseq500 (75 paired-end reads) in the Génomique platform of the Cochin Institute (Paris, France). FASTQ files of the paired-end reads were analyzed with *breseq* [278] software to generate a frequency table that ranks and high lights mutations that made up >0.06% of the library (Supplementary Materials Section C.2).

3.7.10 Validation of candidate Alr_{Mtb} mutations

Plasmids carrying suspected DCS resistance mutations in Alr_{Mtb} were retransformed to strain TS072. Next, the strains were assayed to determine their DCS IC_{50}-values as described before in section 3.7.5.

Molecular dynamics simulations

3.7.11 Protein structures

The simulations were started from the 1.9 Å crystal structure of *Mycobacterium tuberculosis* alanine racemase (PDB code: 1XFC) [279]. Alanine racemase is a dimer with two active sites, which are located in the interface between the two monomers. The structure in 1XFC contains residues 11-173 and 179-381 of chain A and 11-263 and 279-381 of chain B. Most of these residues are closed to the active site of the PLP bound to chain A, therefore, to minimize the effects of the missing residues only the interaction between D-cycloserine and alanine and the active site of chain B were studied. All termini and missing residues were capped with acetyl and N-methyl groups so as to not introduce artifcial charges inside the structure of the protein. The structure of *Mtb* alanine racemase bound to D-cycloserine is unknown. Hence, to study the interaction of *Mtb* alanine racemase with D-cycloserine, we superimpose it on the structure of the *Geobacillus stearothermophilus* alanine racemase bound to D-pyridoxil-N,O-cycloserylamide-5-monophosphate, an inhibitor derived from D-cycloserine (PDB code: 1EPV) [260]. For the starting conformation of the simulations of *Mtb* alanine racemase in combination with L-alanine, we superimposed the structure of *Mtb* alanine racemase to the structure of the *Bacillus stearothermophilus* alanine racemase bound to N-(5'-phospopyridoxyl-L-alanine), which is a reaction intermediate analog for L-alanine (PDB code: 1L6F) [280]. Mutants were created using UCSF Chimera [281].

3.7.12 Force field parametrization

The molecular dynamics (MD) simulations were performed with the popular Amber99SB*ILDN force field. Force field parameters for the protein were obtained from the standard databases. However, new parameters were determined for D-cycloserine, lysine-bound PLP, and the free L-alanine. The three molecules were parametrized with the protonation states needed for the established reaction mechanisms for both L-alanine [280] and D-cycloserine [282] reactions (Figure 3.18).

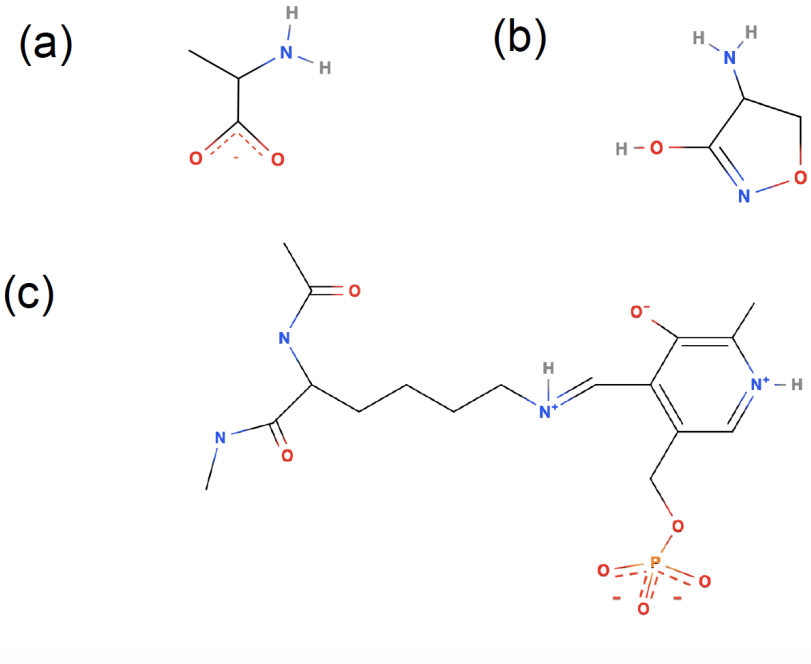

Figure 3.18: Structure and protonation state simulated for (a) L-alanine, (b) D-cycloserine and (c) lysine-bound PLP.

This means that L-alanine was simulated with a negatively charged acid group and a neutral amino group, which is needed for the nucleophilic attack of the PLP aldimine carbon. Also, the D-cycloserine was simulated with a ketone group instead of an alcohol group, as discussed by Fenn *et al.* [260]. Finally, PLP was bound to lysine as needed for the start of the reaction of both L-alanine [282] and D-cycloserine [280]. Lysine-bound PLP was parametrized capped with

an acetyl group and a N-methyl group, to take into account that lysine-bound PLP is located inside a protein backbone and not in isolation. The force field parameters were obtained using the general Amber force field (GAFF) [283]. Charges were obtained using the standard protocol for Amber99 force fields [284]. First, missing hydrogens were added using the Avogadro software [285]. All structures were first optimized using Gaussian09 [286] at the B3LYP/6-31G* level and then refined at the MP2/6-31G* level. The electrostatic potential was then calculated at the HF/6-31G* level. Charges were then determined using the Restrained ElectroStatic Potential (RESP) method [287] using the Antechamber package [283] Parameters were converted to Gromacs format using acpype [288].

3.7.13 Molecular dynamics simulations

All systems were simulated using the Amber99SB*ILDN [284, 289, 290, 291, 292] force field in combination with the Tip4p-Ew water model [293]. Alanine racemase was always simulated as a dimer. The proteins were first introduced in a rhombic dodecahedric box with a size large enough so that the protein was at least 1.2 nm away from the border. The system was then solvated and ions were added to neutralize the system and to simulate an ion concentration of 150 mM. Each system contains around 170,000 atoms. The sytems were first minimized using the steepest descent algorithm. They were then equilibrated using a 1-ns NVT and a 1-ns NPT simulation. Production runs were then performed at 310 K using the stochastic velocity rescale thermostat [294] and at 1 bar using the Parrinello-Rahman barostat [295, 296]. Electrostatic interactions were calculated using the particle mesh Ewald algorithm [297] with a cutoff of 1 nm and a Fourier spacing of 0.16 nm. Van der Waals interactions were calculated with a cutoff

of 1 nm. Protein bonds were constrained using the Lincs algorithm [298, 299] and water internal motions were constrained using the Settle algorithm [300]. All simulation were performed using Gromacs 2016.3 [301]. Protein visualization was performed using UCSF Chimera [281] and VMD.

Perspectives

The work in this thesis represents the first steps of the development of a technology platform based on engineered *E. coli* strains for tuberculosis drug discovery and the investigation into genetic causes of drug resistance.
With the help of colleagues throughout my PhD and thanks to collaborations I established, I developed the fundamental building blocks (e.g. genetic parts, strains, and methods) to enable the TESEC-based identification of target inhibitors and resistance mutations.

Based on my experience, I outline several opportunities as well as possible improvements to maximize the potential of TESEC.

Additional drug targets

In Chapter 1, I described how I implemented seven high priority *Mtb* drug targets in *E. coli*. Six out seven tested essential *Mtb* genes could functionally replace homologous native *E. coli* genes. This flexibility indicates the potential for an expanded TESEC strain collection with bacteria that express other target genes of interest from *M. tuberculosis*, or from other pathogenic bacteria. This could be particularly relevant for priority pathogens with problematic antibiotic resistance, and for which new medicines are critical [302].

The basics to construct new TESEC strains based on the design principles in this work is relatively straight forward. However, finding optimal growth conditions, genetic backgrounds, and expression levels that allow for a normal growing chassis with maximum sensitivity to target-specific drugs can be challenging. In general, functional heterologous protein production depends the transcription, translation, protein folding or aggregation rates [303]. In addition, homolog proteins from different organisms have different catalytic properties [115]. These biological factors impact *E. coli* physiology and determine the expression range that is acceptable for bacterial growth. Given this, and based on my experience, the use of an expression system that allows for increased control over protein expression is crucial. However, case-by-case TESEC strain optimization is still necessary to fine-tune conditions that render useful stratins for a robust and efficient drug screening assay.

Increasing drug influx

The absence of TolC-related drug efflux pumps in TESEC-Alr$_{Mtb}$ was crucial to identify benazepril as mycobacterial inhibitor (Chapter 2). Therefore, I equipped all ready-to-screen TESEC strains with the same *tolC* (and *entC*) deletion. However, in addition to reduced drug efflux, TESEC strains with an increased drug influx could allow even more compounds to reach the target protein of interest. Krishnamoorthy *et al.* reported that the outer membrane siderophore transporter mutant FhuAΔC/Δ4L allows for the increased passage of a variety of antibiotics [217]. The implementation of this mutation might be a useful TESEC feature to explore. On the other hand, the naturally reduced membrane permeability of the *E. coli* might also be useful to

mimic the difficult to penetrate the mycobacterial cell envelope.

Alternative drug screening strategies

The TESEC drug screening strategy in chapter 2 was based on low and high expression of a single *Mtb* drug target to influence the drug sensitivity. For future work, the same strategy could be used to simultaneously screen a collection of TESEC strains that express different drug targets. When differential growth is observed between high and low induction of a collection of TESEC strains, plasmids could be extracted and sequenced by NGS to determine which target was affected. This strategy is similar to the impressive recent work with engineered *M. tuberculosis* by Johnson *et al.* [207]. However, this high-throughput approach with *E. coli* instead of *Mtb* would significantly reduce the workload, costs, time and harzardous risks.

The TESEC platform can also be used to predict in a single screen whether other pathogenic bacteria could be susceptible to, for example, *Mtb* target-specific compounds. In this case, a TESEC strain library should be constructed that expresses the same homolog protein from a variety of pathogenic organisms. Next, this library could be exposed to the compound to determine the narrow- or broad-spectrum selectivity.

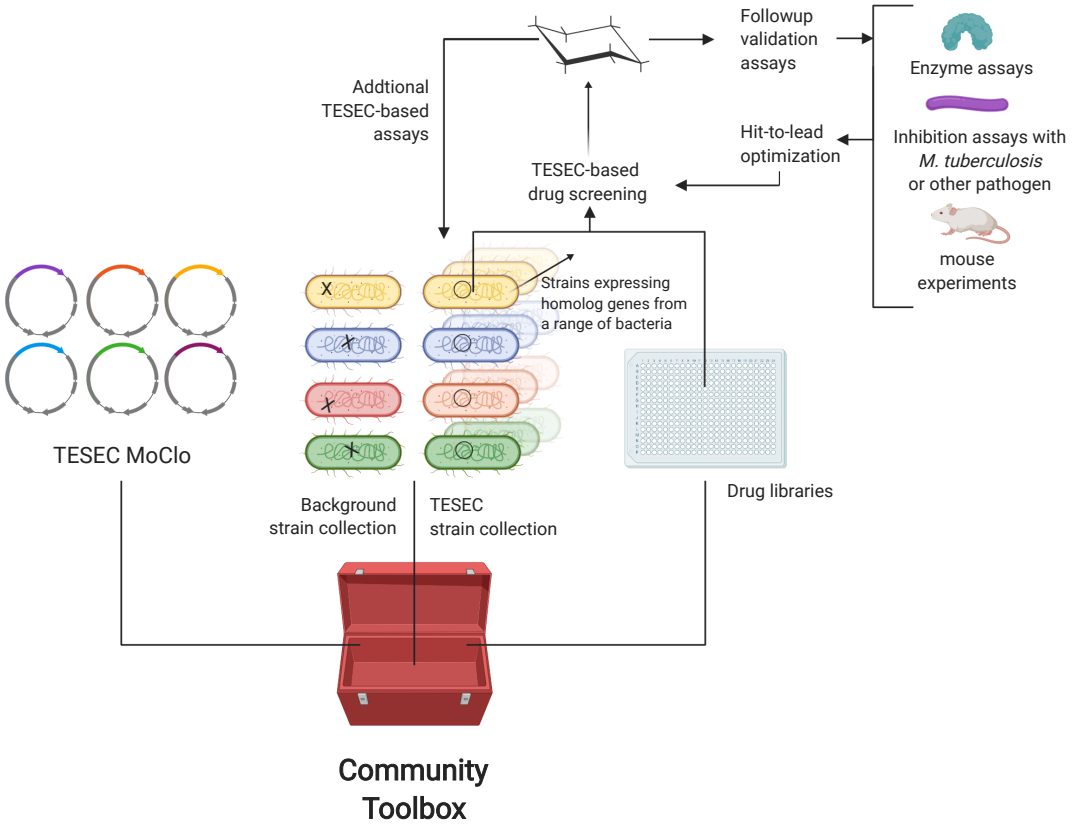

Figure 3.19: TESEC-based drug discovery.
The TESEC community toolbox consists of a TESEC MoClo library, a variety of background strains lacking native essential *E. coli* genes as well as genetic modifications to improve the drug sensitivity. TESEC-based drug screening can be conducted with single TESEC strains or a collection of strains. Hit compound identification can be done based on differential screening for growth between low and high expression conditions or differences in targets. The development of antituberculosis drugs is generally challenging because drug molecules have to overcome several barriers related to the granuloma, passage through the membrane of the infected host-cell, as well as the difficult-to-penetrate envelope of *Mtb* itself [71]. Also, molecules should escape from potential drug efflux or drug degradation activities in the pathogen and effectively inhibit an essential *Mtb* protein. Although TESEC-based discovery does provide a bacterial cell context, it does not mimic all barriers. Followup assays to test the effectiveness on *M. tuberculosis* in vitro and in vivo are therefore important. However, instead of testing thousands of molecules on pathogenic bacteria, selective hits from a TESEC-screen can be assayed at various compound concentrations. Molecules resulting from hit-to-lead optimization to improve the selectivity, specificity, and sensitivity can be directly tested on the pathogen or routed for another round of TESEC-based testing.

Biosafety & biosecurity advantages

The pathogenic and virulent nature of *Mycobacterium tuberculosis* poses direct risks to researchers who work with this organism as well

as their environment. Laboratory-acquired infections of tuberculosis occur, even when biosafety measures are considered [99]. This is an important risk factor in tuberculosis research, especially when gain-of-function experiments are conducted that could make the pathogen e.g., more virulent or resistant to medication. The study described in Chapterr 3 to validate and predict *Mtb* D-cycloserine resistance genotypes is an example of a gain-of-function experiment. Because I used the non-pathogenic *E. coli* K-12 MG1655 strain (with a F-minus genotype) and nonconjugative ColE1 derived plasmids [304] biosafety and biosecurity risks related to gain-of-function studies are significantly reduced. Similar experiments with *Mycobacterium tuberculosis* could yield tuberculosis with increased drug resistance and would require an assessment to determine whether the benefits of the study outweigh the potential risks [305].

Automation

Laboratory automation technologies are valuable assets for the high-throughput characterization of TESEC strains or drug screening experiments. Unfortunately, hurdles I faced during the thesis were mainly related to the continuous failure of the in-house TECAN EVO ware robot. This situation severely challenged the many large-scale multivariable strain characterization experiments I envisioned for this thesis and delayed many experiments by weeks and even moths. Moreover, a significant amount of my data was lost due to measurements that were randomly interrupted in the robot. When the robot platform functioned as desired, I was able to make rapid progress. However, it was frequently difficult to obtain replicate results when the robot failed to finish the experiment. Despite the significant efforts I invested in solving technical problems in collaboration with the company, impor-

tant technical issues were not resolved before the end of my thesis.

This year, we started to use the Labcyte Echo® Acoustic Liquid Handling machine that uses acoustics to dispense reagents into microtiter plates. This machine is handy to prepare drug library plates, dilution series, combinatorial DNA libraries, and significantly aids in-house drug screening efforts. Moreover, thanks to the possibility of dispensing nanoliter volumes, compound expenses can be significantly reduced. Still, subsequent assays to obtain rich data sets of complete bacterial growth curves (for either high-throughput drug screens or strain characterization) requires a functional plate-handling robot. With the current lab facilities, numerous strains can be built, tested and optimized in-house, but collaborations with third parties that have advanced automation for drug screening is important to maximize the potential finding of inhibitors. In a first attempt to scale TESEC, I initiated conversations with the NIH to explore the potential for a later compound screen with the validated drug screening strain TS144 with Alr_{Mtb} expression.

Enzyme purification

The strength of the TESEC system is that the initial drug screening does not require the isolation of enzymes. However, followup work to validate the mechanism of action of drugs do require further testing of enzymes. Although I was able to successfully purify Alr_{Mtb}, obtaining the right quantity and quality of the enzyme products can be challenging without specialized equipment. Therefore, I believe that this project would benefit from strong collaborations with a laboratory that has facilities in place to purify enzymes.

Access to compounds

Although TESEC-based screening requires inexpensive growth medium and antibiotics, it was not always easy to obtain specific chemical compounds. Some were not available for sale, or low compound quantities were highly-priced. Collaborations with chemists and pharmaceutical companies could minimize this barrier.

Open source pharma

The collection of TESEC strains and TESEC MoClo plasmids I constructed are ready to be converted into a 'TESEC Starters Kit' to be used as an open source pharma plug-and-play platform for the discovery of antibacterials against *Mtb*. Students, researchers, and organizations could contribute in various ways. Students can easily assemble TESEC MoClo plasmids, generate chromosomal deletions, or conduct small-scale drug screening and strain characterization tests. Professional entities such as pharmaceutical companies can contribute by providing compounds or screening validated TESEC strains. Strong collaborations with laboratories that have complementary expertise in biochemistry, medicinal chemistry, or tuberculosis research are essential for followup validations of hit compounds and hit-to-lead optimization efforts. Finally, an online-infrastructure to share, protocols, instruction videos, DNA sequences, TESEC strain characterization, and drug screening data would be needed to support an international open source pharma community around TESEC.

Personal reflection

I strongly believe in the importance of bridging science with society. This is why, throughout my science education, I have spent a significant amount of my spare time on educating the public, or training other PhD students in science communication to help them engage with numerous stakeholders and transform their research projects into stories that can be understood by non-scientific communities.

It is natural to me to seek interactions with other stakeholders and I see it as a personal responsibility to share my knowledge with others. When I got the opportunity, I tried to raise awareness about the challenges related to antimicrobial resistance or developments in synthetic biology.

My involvement in Hello Tomorrow and my recent acceptance as a lecturer for Singularity University allow me to engage with various international stakeholders and spark the debate about new technological developments.

I believe that, on the long term, such efforts will help science to reach its potential. Scientists can come up with elegant solutions for complex problems, but if society does not accept them or does not want to invest in them, this science is lost.

Portfolio

Scientific conferences

- **Selected as one of the 100 participants of the Gap Summit 2016**
  Cambridge University, United Kingdom, 4-6 April 2016

- **Poster: Target Essential Surrogate *E. coli* strains**
  SB 7.0 Conference, 13-16 June 2017, Singapore

- **Panel on synthetic biology and policy making**
  iGEM Jamboree, Boston, USA, 14 November 2017.

- **Breakout session: Science Communication**
  EUSynBioS Symposium 2018: Engineering Biology for a Better Future, Toulouse, France, 22 October 2018.

- **Poster and Short talk: *Mtb* resistance mutations based on screening synthetic mutant libraries**
  Combating resistance: Microbes an vectors - 2018 Institut Pasteur International Network Symposium, 15-16 November 2018

- **Presentation: Fighting Tuberculosis**
  GSK Innovation Summit, Brussels, 21 February 2019.

- **Panelist: Open R&D: Across the Stages**
  Open Source Pharma meeting 2019, CRI, Paris, 22 March 2019

- **Panelist: 'Le futur des Biotechs et des Medtechs'**
  BIOTechno forum Paris 2019, 14 June 2019

Public Outreach

- **TEDx presentation: Building better bridges between science and society for a better future**
  TEDx Saclay, 12 November 2015

- **Presentation: Synthetic Biology**
  Fraunhofer-Symposium, München, Germany, 22 February 2016

- **Presentation: The war against superbugs**
  Campus Party Conference, Utrecht, The Netherlands, 28 May 2016

- **Presentation: Synthetic Biology to fight against Antibiotic Resistant Bacteria.**
  European Health Parliament, EU, Brussels, 5 December 2017.

- **Scientific advice for popular science book: The Future (De Toekomst - Dutch)**
  Published by De Speld, September 2019.

- **Book contribution - Nature: Collaborations in Design: Cooper Hewitt Design Triennial**
  Published by Cooper Hewitt, Smithsonian Design Museum, May 2019.

Scientific Publications

- Solis-Escalante D, Kuijpers NG, **Bongaerts N**, *et al.* amdSYM, a new dominant recyclable marker cassette for Saccharomyces cerevisiae. *FEMS Yeast Res.* 2013;13(1):126–139.

- Brinkman EK, Schipper K, **Bongaerts N**, Voges MJ, Abate A, Wahl SA. A toolkit to enable hydrocarbon conversion in aqueous environments. *J Vis Exp.* 2012;(68):e4182. Published 2012 Oct 2. doi:10.3791/4182

Teaching

- **Supervision of four second-year master students for 3-month internships**
  Center for Research and Interdisciplinarity, 2015-2018

- **Teaching Assistant Synthetic Biology Course**
  Center for Research and Interdisciplinarity, 2015-2017

- **Video lessons for the SynBio1 Mooc**
  Center for Research and Interdisciplinarity, 2015-2017

- **iGEM Bettencourt Supervisor**
  iGEM 2016, March-November 2016

- **Lecturer for the international organization Singularity University**
  Non-expert level lectures to inform industry and government about developments in biotechnology, Singularity University, 2018-present.

Entrepreneurship

- **Co-founder of Science Matters**
  Agency for science communication, 2011-present

- **Board member of Hello Tomorrow**
  International non-profit organisation for science entrepreneurship, 2015-present

Bibliography

[1] World Health Organization. GLOBAL TUBERCULOSIS REPORT 2018. 1–277, 2018.

[2] Daniel R, Rubens JR, Sarpeshkar R, and Lu TK. Synthetic analog computation in living cells. *Nature*, 497(7451):619–623, 2013.

[3] PRESTWICK CHEMICAL LIBRARY. Http://www.prestwickchemical.com/libraries-screening-lib-pcl.html, Accessed: Sept. 21, 2019.

[4] Chernyaeva EN, Shulgina MV, Rotkevich MS, Dobrynin PV, Simonov SA *et al.* Genome-wide *Mycobacterium tuberculosis* variation (GMTV) database: a new tool for integrating sequence variations and epidemiology. *BMC genomics*, 15(1):308, 2014.

[5] Coll F, Phelan J, Hill-Cawthorne GA, Nair MB, Mallard K *et al.* Genome-wide analysis of multi- and extensively drug-resistant *Mycobacterium tuberculosis*. *Nature genetics*, 50(5):764–764, 2018.

[6] Desjardins CA, Cohen KA, Munsamy V, Abeel T, Maharaj K *et al.* Genomic and functional analyses of *Mycobacterium tuberculosis* strains implicate ald in D-cycloserine resistance. *Nature genetics*, 48(5):544–551, 2016.

[7] Iverson SV, Haddock TL, Beal J, and Densmore DM. CIDAR MoClo: Improved MoClo Assembly Standard and New *E. coli* Part Library Enable Rapid Combinatorial Design for Synthetic and Traditional Biology. *ACS Synthetic Biology*, 5(1):99–103, 2016.

[8] Patrick WM, Firth AE, and Blackburn JM. User-friendly algorithms for estimating completeness and diversity in randomized protein-encoding libraries. *Protein engineering*, 16(6):451–457, 2003.

[9] Gilbert JA and Neufeld JD. Life in a World without Microbes. *PLOS Biology*, 12(12):e1002020, 2014.

[10] Lotfy WM. Plague in Egypt: Disease biology, history and contemporary analysis: A minireview. *Journal of advanced research*, 6(4):549–554, 2015.

[11] Dobson AP and Carper ER. Infectious Diseases and Human Population History on JSTOR. *BioScience*, 46(Disease Ecology):115–126, 1996.

[12] Schlipköter U and Flahault A. Communicable Diseases: Achievements and Challenges for Public Health. *Public Health Reviews*, 32(1):90–119, 2010.

[13] Aiello AE, Larson EL, and Sedlak R. Hidden heroes of the health revolution Sanitation and personal hygiene. *American Journal of Infection Control*, 36(10):S128–S151, 2008.

[14] Jayachandran S, Lleras-Muney A, and Smith KV. Modern medicine and the twentieth century decline in mortality: Evidence on the impact of sulfa drugs. *American Economic Journal: Applied Economics*, 2(2):118–46, 2010.

[15] Fleming A. *On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B. influenzae. 1929.*, volume 79. World Health Organization, Laboratories of the Inoculation Department, St Mary's Hospital, London., 2001.

[16] Clardy J, Fischbach MA, and Currie CR. The natural history of antibiotics. *Current Biology*, 219:R437–R441, 2019.

[17] Fleming A. Penicillin. Nobel Lecture, 11 December 1945. Https://www.nobelprize.org/uploads/2018/06/fleming-lecture.pdf, Accessed: Oct. 10, 2019.

[18] World Health Organization. No Time to Wait: Securing the future from drug-resistant infections, 2019.

[19] Landers TF, Cohen B, Wittum TE, and Larson EL. A Review of Antibiotic Use in Food Animals: Perspective, Policy, and Potential:. *Public Health Reports*, 127(1):4–22, 2012.

[20] Li Y. China's misuse of antibiotics should be curbed. *BMJ*, 348, 2014.

[21] Soyka LF, Robinson DS, Lachant N, and Monaco J. The Misuse of Antibiotics for Treatment of Upper Respiratory Tract Infections in Children. *Pediatrics*, 55(4):552–556, 1975.

[22] Pechère JC. Patients' interviews and misuse of antibiotics. *Clinical Infectious Diseases*, 33(Supplement 3):S170–S173, 2001.

[23] Durão P, Balbontín R, and Gordo I. Evolutionary Mechanisms Shaping the Maintenance of Antibiotic Resistance. *Trends in microbiology*, 26(8):677–691, 2018.

[24] Cooper MA and Shlaes D. Fix the antibiotics pipeline. *Nature*, 472(7341):32–32, 2011.

[25] O'Neill J. Tackling Drug-Resistant Infections Globally, 2016. Https://amr-review.org/sites/default/files/160518_Final%20paper_with%20cover.pdf, Accessed: Aug. 12, 2019.

[26] Erasmus V, Daha TJ, Brug H, Richardus JH, Behrendt MD *et al.* Systematic Review of Studies on Compliance with Hand Hygiene Guidelines in Hospital Care. *Infection control and hospital epidemiology*, 31(3):283–294, 2010.

[27] Aslam B, Wang W, Arshad MI, Khurshid M, Muzammil S *et al.* Antibiotic resistance: a rundown of a global crisis. *Infection and drug resistance*, 11:1645–1658, 2018.

[28] Daniel TM. The history of tuberculosis. *Respiratory Medicine*, 100(11):1862–1870, 2006.

[29] Sakula A. Robert Koch: centenary of the discovery of the tubercle bacillus, 1882. *Thorax*, 37(4):246–251, 1982.

[30] World Health Organization. Tuberculosis, 2019. Https://www.who.int/news-room/fact-sheets/detail/tuberculosis, Accessed: Oct. 28, 2019.

[31] Jackson M. The Mycobacterial Cell Envelope—Lipids. *Cold Spring Harbor perspectives in medicine*, 4(10):a021105–a021105, 2014.

[32] Brown L, Wolf JM, Prados-Rosales R, and Casadevall A. Through the wall: extracellular vesicles in Gram-positive bacteria, mycobacteria and fungi. *Nature reviews. Microbiology*, 13(10):620–630, 2015.

[33] Medjahed H, Gaillard JL, and Reyrat JM. Mycobacterium abscessus: a new player in the mycobacterial field. *Trends in microbiology*, 18(3):117–123, 2010.

[34] Knechel NA. Tuberculosis: pathophysiology, clinical features, and diagnosis. *Critical Care Nurse*, 29(2):34–43, 2009.

[35] Wong D, Bach H, Sun J, Hmama Z, and Av-Gay Y. *Mycobacterium tuberculosis* protein tyrosine phosphatase (PtpA) excludes host vacuolar-H+–ATPase to inhibit phagosome acidification. *Proceedings of the National Academy of Sciences*, 108(48):19371–19376, 2011.

[36] Warner DF and Mizrahi V. The survival kit of *Mycobacterium tuberculosis*. *Nature Medicine*, 13(3):282–284, 2007.

[37] Queval CJ, Song OR, Carralot JP, Saliou JM, Bongiovanni A *et al.* *Mycobacterium tuberculosis* Controls Phagosomal Acidification by Targeting CISH-Mediated Signaling. *Cell Reports*, 20(13):3188–3198, 2017.

[38] Pieters J. *Mycobacterium tuberculosis* and the Macrophage: Maintaining a Balance. *Cell Host & Microbe*, 3(6):399–407, 2008.

[39] Gideon HP and Flynn JL. Latent tuberculosis: what the host "sees"? *Immunologic Research*, 2011.

[40] Balcells ME, Yokobori N, Hong BY, Corbett J, and Cervantes J. The lung microbiome, vitamin D, and the tuberculous granuloma: A balance triangle. *Microbial Pathogenesis*, 131:158–163, 2019.

[41] Russell DG. *Mycobacterium tuberculosis* : here today, and here tomorrow. *Nature reviews. Molecular cell biology*, 2(8):569–578, 2001.

[42] Cook GM, Berney M, Gebhard S, Heinemann M, Cox RA *et al.* Physiology of Mycobacteria. *Advances in Microbial Physiology*, 55:81–319, 2009.

[43] Vilchèze C and Kremer L. Acid-Fast Positive and Acid-Fast Negative *Mycobacterium tuberculosis*: The Koch Paradox. *Microbiology spectrum*, 5(2), 2017.

[44] Brennan PJ. Structure, function, and biogenesis of the cell wall of *Mycobacterium tuberculosis*, journal = Tuberculosis, year = 2003, volume = 83, number = 1-3, pages = 91–97, month = feb.

[45] Jarlier V and Nikaido H. Mycobacterial cell wall: structure and role in natural resistance to antibiotics. *FEMS Microbiology Letters*, 123(1):11–18, 1994.

[46] Nikaido H. Preventing drug access to targets: cell surface permeability barriers and active efflux in bacteria. *Seminars in Cell & Developmental Biology*, 12(3):215–223, 2001.

[47] Kleinnijenhuis J, Oosting M, Joosten LAB, Netea MG, and Van Crevel R. Innate Immune Recognition of *Mycobacterium tuberculosis*. *Clinical and Developmental Immunology*, 2011, 2011.

[48] Ishikawa E, Mori D, and Yamasaki S. Recognition of Mycobacterial Lipids by Immune Receptors. *Trends in Immunology*, 38(1):66–76, 2017.

[49] Faridgohar M and Nikoueinejad H. New findings of Toll-like receptors involved in *Mycobacterium tuberculosis* infection. *Pathogens and Global Health*, 111(5):256–264, 2017.

[50] O'Garra A, Redford PS, McNab FW, Bloom CI, Wilkinson RJ *et al.* The immune response in tuberculosis. *Annual Review of Immunology*, 31(1):475–527, 2013.

[51] Torrelles JB and Schlesinger LS. Diversity in *Mycobacterium tuberculosis* mannosylated cell wall determinants impacts adaptation to the host. *Tuberculosis*, 90(2):84–93, 2010.

[52] Jo EK, Yang CS, Choi CH, and Harding CV. Intracellular signalling cascades regulating innate immune responses to Mycobacteria: branching out from Toll-like receptors. *Cellular microbiology*, 9(5):1087–1098, 2007.

[53] Schlesinger LS. Macrophage phagocytosis of virulent but not attenuated strains of *Mycobacterium tuberculosis* is mediated by mannose receptors in addition to complement receptors. *The Journal of Immunology*, 150(7):2920–2930, 1993.

[54] Schlesinger LS, Hull SR, and Kaufman TM. Binding of the terminal mannosyl units of lipoarabinomannan from a virulent strain of *Mycobacterium tuberculosis* to human macrophages. *The Journal of Immunology*, 152(8):4070–4079, 1994.

[55] Uribe-Querol E and Rosales C. Control of Phagocytosis by Microbial Pathogens. *Frontiers in Immunology*, 8:191, 2017.

[56] Nguyen L and Pieters J. The Trojan horse: survival tactics of pathogenic mycobacteria in macrophages. *Trends in Cell Biology*, 15(5):269–276, 2005.

[57] Ehlers S and Schaible UE. The Granuloma in Tuberculosis: Dynamics of a Host–Pathogen Collusion. *Frontiers in Immunology*, 3, 2013.

[58] Kiran D, Podell BK, Chambers M, and Basaraba RJ. Host-directed therapy targeting the *Mycobacterium tuberculosis* granuloma: a review. *Seminars in Immunopathology*, 38(2):167–183, 2015.

[59] Flynn JL, Chan J, and Lin PL. Macrophages and control of granulomatous inflammation in tuberculosis. *Mucosal Immunology*, 4(3):271–278, 2011.

[60] Jilani TN, Avula A, Gondal AZ, and Siddiqui AH. Active Tuberculosis. In *StatPearls [Internet]*, 101–114. StatPearls Publishing, Treasure Island (FL), 2019.

[61] Luca S and Mihaescu T. History of BCG Vaccine. *Mædica*, 8(1):53–58, 2013.

[62] Lechartier B, Rybniker J, Zumla A, and Cole ST. Tuberculosis drug discovery in the post-post-genomic era. *EMBO molecular medicine*, 6(2):158–168, 2014.

[63] World Health Organization. Latent tuberculosis infection: updated and consolidated guidelines for programmatic management, 2018.

[64] Mandal S, Njikan S, Kumar A, Early JV, and Parish T. The relevance of persisters in tuberculosis drug discovery. *Microbiology*, 165(5):492–499, 2019.

[65] Mallolas J, Sarasa M, Nomdedeu M, Soriano A, López-Púa Y *et al.* Pharmacokinetic interaction between rifampicin and ritonavir-boosted atazanavir in hiv-infected patients*. *HIV Medicine*, 8(2):131–134, 2007.

[66] Munro SA, Lewin SA, Smith HJ, Engel ME, Fretheim A *et al.* Patient Adherence to Tuberculosis Treatment: A Systematic Review of Qualitative Research. *PLOS Medicine*, 4(7):e238, 2007.

[67] Nguyen L. Antibiotic resistance mechanisms in M. tuberculosis: an update. *Archives of toxicology*, 90(7):1585–1604, 2016.

[68] Balganesh M, Dinesh N, Sharma S, Kuruppath S, Nair AV *et al.* Efflux Pumps of *Mycobacterium tuberculosis* Play a Significant Role in Antituberculosis Activity of Potential Drug Candidates. *Antimicrobial Agents and Chemotherapy*, 56(5):2643–2651, 2012.

[69] Szumowski JD, Adams KN, Edelstein PH, and Ramakrishnan L. Antimicrobial Efflux Pumps and *Mycobacterium tuberculosis* Drug Tolerance: Evolutionary Considerations. In *Pathogenesis of Mycobacterium tuberculosis and its Interaction with the Host Organism*, 81–108. Springer, Berlin, Heidelberg, Berlin, Heidelberg, 2012.

[70] Hugonnet JE, Tremblay LW, Boshoff HI, Barry CE, and Blanchard JS. Meropenem-Clavulanate Is Effective Against Extensively Drug-Resistant *Mycobacterium tuberculosis*. *Science*, 323(5918):1215–1218, 2009.

[71] Sarathy J, Dartois V, Dick T, and Gengenbacher M. Reduced Drug Uptake in Phenotypically Resistant Nutrient-Starved Nonreplicating *Mycobacterium tuberculosis*. *Antimicrobial Agents and Chemotherapy*, 57(4):1648–1653, 2013.

[72] Dookie N, Rambaran S, Padayatchi N, Mahomed S, and Naidoo K. Evolution of drug resistance in *Mycobacterium tuberculosis*: a review on the molecular determinants of resistance and im-

plications for personalized care. *The Journal of antimicrobial chemotherapy*, 73(5):1138–1151, 2018.

[73] Gillespie SH. Evolution of Drug Resistance in *Mycobacterium tuberculosis*: Clinical and Molecular Perspective. *Antimicrobial Agents and Chemotherapy*, 46(2):267–274, 2002.

[74] Eldholm V and Balloux F. Antimicrobial Resistance in *Mycobacterium tuberculosis*: The Odd One Out. *Trends in microbiology*, 24(8):637–648, 2016.

[75] World Health Organization. Global MDR-TB situation: update 2018, 2018. Https://www.who.int/tb/areas-of-work/drug-resistant-tb/MDR_TB_2018.pptx, Accessed: Oct. 29, 2019.

[76] World Health Organization. Antimicrobial Resistance. Technical report, 2014.

[77] Kim SJ. Drug-susceptibility testing in tuberculosis: methods and reliability of results. *The European respiratory journal*, 25(3):564–569, 2005.

[78] Schön T, Miotto P, Köser CU, Viveiros M, Böttger E *et al. Mycobacterium tuberculosis* drug-resistance testing: challenges, recent developments and perspectives. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*, 23(3):154–160, 2017.

[79] World Health Organization. Noncommercial Culture and Drug-Susceptibility Testing Methods for Screening Patients at Risk for Multidrug-Resistant Tuberculosis: Policy Statement, 2011.

[80] UNITAID. Tuberculosis Medicines Technology and Market Landscape. 1–79, 2014.

[81] World Health Organization. The use of next-generation sequencing technologies for the detection of mutations associated with drug resistance in *Mycobacterium tuberculosis* complex: technical guide, 2018. Https://apps.who.int/iris/bitstream/handle/10665/274443/WHO-CDS-TB-2018.19-eng.pdf, Accessed: Oct. 27, 2019.

[82] Theron G, Peter J, Richardson M, Barnard M, Donegan S *et al.* The diagnostic accuracy of the GenoType(®) MTBDRsl assay for the detection of resistance to second-line antituberculosis drugs. *The Cochrane database of systematic reviews*, 7(10):CD010705, 2014.

[83] World Bank Group. Drug-resistant infections: a threat to our economic future, 2017.

[84] CDC. The costly burden of drug-resistant TB in the U.S., 2018. Https://www.cdc.gov/nchhstp/newsroom/docs/factsheets/costly-burden-dr-tb-508.pdf, Accessed: Oct. 23, 2019.

[85] World Health Organization. The End TB Strategy. 1–26, 2014.

[86] World Health Organization. Implementing the end TB strategy: the essentials. Geneva: The World Health Organization; 2015, 2019.

[87] Global Fund. Global Fund Donors Pledge US$14 Billion in Fight to End Epidemics, 2019.

[88] Mahajan R. Bedaquiline: First FDA-approved tuberculosis drug in 40 years. *International Journal of Applied and Basic Medical Research*, 3(1):1–2, 2013.

[89] Sotgiu G, Pontali E, Centis R, D'Ambrosio L, and Migliori GB. Delamanid (OPC-67683) for treatment of multi-drug-resistant

tuberculosis. *Expert Review of Anti-infective Therapy*, 13(3):305–315, 2014.

[90] Keam SJ. Pretomanid: First Approval. *Drugs*, 405(6789):1–7, 2019.

[91] TB Alliance. Pretomanid and BPaL Regimen for Treatment of Highly-Resistant Tuberculosis, 2019. Https://www.fda.gov/media/128001/download, Accessed: Oct. 29, 2019.

[92] Médicins Sans Frontières. MSF Demands Johnson & Johnson Reduce Price of Lifesaving TB Drug, 2019. Https://www.msf.org/johnson-johnson-must-halve-price-lifesaving-tb-drug-bedaquiline, Accessed: Oct. 28, 2019.

[93] Médicins Sans Frontières. Promising new tuberculosis drug pretomanid approved, but will it be affordable?, 2019. Https://www.msf.org/promising-new-tuberculosis-drug-approved-will-it-be-affordable, Accessed: Oct. 28, 2019.

[94] Médicins Sans Frontières. Price announced for new lifesaving TB drug pretomanid still too high, 2019. Https://www.msfaccess.org/price-announced-new-lifesaving-tb-drug-pretomanid-still-too-high, Accessed: Oct. 29, 2019.

[95] Nguyen T, Anthony RM, Bañuls AL, Van Anh Nguyen T, Vu DH *et al.* Bedaquiline resistance: its emergence, mechanism, and prevention. *Clinical Infectious Diseases*, 66(10):1625–1630, 2017.

[96] Balasegaram M, Kolb P, McKew J, Menon J, Olliaro P *et al.* An open source pharma roadmap. *PLOS Medicine*, 14(4):e1002276, 2017.

[97] iGEM Paris Bettencourt 2013. Http://2013.igem.org/Team:Paris_Bettencourt, Accessed: Sept. 14, 2019.

[98] European Centre for Disease Prevention and Control. Handbook on tuberculosis laboratory diagnostic methods in the European Union, 2018. Https://www.ecdc.europa.eu/sites/portal/files/documents/TB-handbook-updated-2018.pdf, Accessed: Oct. 22, 2019.

[99] Weinstein RA and Singh K. Laboratory-Acquired Infections. *Clinical Infectious Diseases*, 49(1):142–147, 2009.

[100] Gill WP, Harik NS, Whiddon MR, Liao RP, Mittler JE *et al.* A replication clock for *Mycobacterium tuberculosis*. *Nature Medicine*, 15(2):211–214, 2009.

[101] Borgers K, Vandewalle K, Festjens N, and Callewaert N. A guide to Mycobacterium mutagenesis. *The FEBS journal*, 286(19):3757–3774, 2019.

[102] Yuan T and Sampson NS. Hit Generation in TB Drug Discovery: From Genome to Granuloma. *Chemical Reviews*, 118(4):1887–1916, 2018.

[103] Movahedzadeh F and Bitter W. Ins and outs of mycobacterial plasmids. *Methods in molecular biology (Clifton, N.J.)*, 465(11):217–228, 2009.

[104] Goude R, Roberts DM, and Parish T. Electroporation of mycobacteria. *Methods in molecular biology (Clifton, N.J.)*, 1285(12):117–130, 2015.

[105] Royet J and Dziarski R. Peptidoglycan recognition proteins: pleiotropic sensors and effectors of antimicrobial defences. *Nature reviews. Microbiology*, 5(4):264–277, 2007.

[106] Rogers HJ. Bacterial growth and the cell envelope. *Bacteriological reviews*, 34(2):194–214, 1970.

[107] Vollmer W, Blanot D, and De Pedro MA. Peptidoglycan structure and architecture. *FEMS Microbiology Reviews*, 32(2):149–167, 2008.

[108] Hett EC and Rubin EJ. Bacterial growth and cell division: a mycobacterial perspective. *Microbiology and molecular biology reviews : MMBR*, 72(1):126–56– table of contents, 2008.

[109] Payne KM and Hatfull GF. Mycobacteriophage endolysins: diverse and modular enzymes with multiple catalytic activities. *PloS one*, 7(3):e34052, 2012.

[110] Gupta R, Lavollay M, Mainardi JL, Arthur M, Bishai WR *et al.* The *Mycobacterium tuberculosis* protein LdtMt2 is a nonclassical transpeptidase required for virulence and resistance to amoxicillin. *Nature Medicine*, 16(4):466–469, 2010.

[111] Lavollay M, Arthur M, Fourgeaud M, Dubost L, Marie A *et al.* The peptidoglycan of stationary-phase *Mycobacterium tuberculosis* predominantly contains cross-links generated by L,D-transpeptidation. *Journal of Bacteriology*, 190(12):4360–4366, 2008.

[112] Raymond JB, Mahapatra S, Crick DC, and Pavelka MS. Identification of the namH gene, encoding the hydroxylase responsible for the N-glycolylation of the mycobacterial peptidoglycan. *The Journal of biological chemistry*, 280(1):326–333, 2005.

[113] Böth D, Schneider G, and Schnell R. Peptidoglycan remodeling in *Mycobacterium tuberculosis*: comparison of structures and catalytic activities of RipA and RipB. *Journal of Molecular Biology*, 413(1):247–260, 2011.

[114] Nikolaidis I, Favini-Stabile S, and Dessen A. Resistance to antibiotics targeted to the bacterial cell wall. *Protein science : a publication of the Protein Society*, 23(3):243–259, 2014.

[115] Strych U, Penland RL, Jimenez M, Krause KL, and Benedik MJ. Characterization of the alanine racemases from two mycobacteria. *FEMS Microbiology Letters*, 196(2):93–98, 2001.

[116] Usha V, Lloyd AJ, Lovering AL, and Besra GS. Structure and function of *Mycobacterium tuberculosis* meso-diaminopimelic acid (DAP) biosynthetic enzymes. *FEMS Microbiology Letters*, 330(1):10–16, 2012.

[117] Bouvier J, Richaud C, Higgins W, Bögler O, and Stragier P. Cloning, characterization, and expression of the dapE gene of *Escherichia coli*. *Journal of Bacteriology*, 174(16):5265–5271, 1992.

[118] Scapin G, Reddy SG, Zheng R, and Blanchard JS. Three-dimensional structure of *Escherichia coli* dihydrodipicolinate reductase in complex with NADH and the inhibitor 2,6-pyridinedicarboxylate. *Biochemistry*, 36(49):15081–15088, 1997.

[119] Cohen GN. Aspartate-semialdehyde dehydrogenase from *Escherichia coli*. In *Glutamate, Glutamine, Glutathione, and Related Compounds*, 600–602. Elsevier, 1985.

[120] Dogovski C, C S, R S, Downton M, Hor L *et al.* Enzymology of Bacterial Lysine Biosynthesis. In *Biochemistry*. InTech, 2012.

[121] Pavelka MS and Jacobs WR. Biosynthesis of diaminopimelate, the precursor of lysine and a component of peptidoglycan, is an essential function of *Mycobacterium smegmatis*. *Journal of Bacteriology*, 178(22):6496–6507, 1996.

[122] Meng J, Yang Y, Xiao C, Guan Y, Hao X *et al.* Identification and Validation of Aspartic Acid Semialdehyde Dehydrogenase as a New Anti-*Mycobacterium tuberculosis* Target. *International journal of molecular sciences*, 16(10):23572–23586, 2015.

[123] Sassetti CM, Boyd DH, and Rubin EJ. Genes required for mycobacterial growth defined by high density mutagenesis. *Molecular Microbiology*, 48(1):77–84, 2003.

[124] Vyas R, Tewari R, Weiss MS, Karthikeyan S, and IUCr. Structures of ternary complexes of aspartate-semialdehyde dehydrogenase (Rv3708c) from *Mycobacterium tuberculosis* H37Rv. *Acta Crystallographica Section D: Biological Crystallography*, 68(6):671–679, 2012.

[125] Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y *et al.* Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology*, 2(1):2006.0008, 2006.

[126] Lin X, Xu S, Yang Y, Wu J, Wang H *et al.* Purification and characterization of anthranilate synthase component I (TrpE) from *Mycobacterium tuberculosis* H37Rv. *Protein expression and purification*, 64(1):8–15, 2009.

[127] Murray HW, Szuro-Sudol A, Wellner D, Oca MJ, Granger AM *et al.* Role of tryptophan degradation in respiratory burst-independent antimicrobial activity of gamma interferon-

stimulated human macrophages. *Infection and immunity*, 57(3):845–849, 1989.

[128] Zhang YJ, Reddy MC, Ioerger TR, Rothchild AC, Dartois V *et al.* Tryptophan biosynthesis protects mycobacteria from CD4 T-cell-mediated killing. *Cell*, 155(6):1296–1308, 2013.

[129] Lee CE, Goodfellow C, Javid-Majd F, Baker EN, and Shaun Lott J. The crystal structure of TrpD, a metabolic enzyme essential for lung colonization by *Mycobacterium tuberculosis*, in complex with its substrate phosphoribosylpyrophosphate. *Journal of Molecular Biology*, 355(4):784–797, 2006.

[130] Raman K, Yeturu K, and Chandra N. targetTB: a target identification pipeline for *Mycobacterium tuberculosis* through an interactome, reactome and genome-scale structural analysis. *BMC systems biology*, 2:109, 2008.

[131] Hampshire T, Soneji S, Bacon J, James BW, Hinds J *et al.* Stationary phase gene expression of *Mycobacterium tuberculosis* following a progressive nutrient depletion: a model for persistent organisms? *Tuberculosis*, 84(3-4):228–238, 2004.

[132] Paritala H and Carroll KS. New targets and inhibitors of mycobacterial sulfur metabolism. *Infectious Disorders Drug Targets*, 13(2):85–115, 2013.

[133] Williams SJ, Senaratne RH, Mougous JD, Riley LW, and Bertozzi CR. 5'-adenosinephosphosulfate lies at a metabolic branch point in mycobacteria. *The Journal of biological chemistry*, 277(36):32606–32615, 2002.

[134] Senaratne RH, De Silva AD, Williams SJ, Mougous JD, Reader JR *et al.* 5'-Adenosinephosphosulphate reductase (CysH) pro-

tects *Mycobacterium tuberculosis* against free radicals during chronic infection phase in mice. *Molecular Microbiology*, 59(6):1744–1753, 2006.

[135] Chen YY, Galloway KE, and Smolke CD. Synthetic biology: advancing biological frontiers by building synthetic systems. *Genome biology*, 13(2):240–10, 2012.

[136] Dekel E and Alon U. Optimality and evolutionary tuning of the expression level of a protein. *Nature*, 436(7050):588–592, 2005.

[137] Melander RJ, Zurawski DV, and Melander C. Narrow-spectrum antibacterial agents. *MedChemComm*, 9(1):12–21, 2018.

[138] Hong BY, Maulén NP, Adami AJ, Granados H, Balcells ME *et al.* Microbiome Changes during Tuberculosis and Antituberculous Therapy. *Clinical microbiology reviews*, 29(4):915–926, 2016.

[139] Shafiani S, Sharma P, Vohra RM, and Tewari R. Cloning and characterization of aspartate-beta-semialdehyde dehydrogenase from *Mycobacterium tuberculosis* H37 Rv. *Journal of applied microbiology*, 98(4):832–838, 2005.

[140] Naz S, Ngo T, Farooq U, and Abagyan R. Analysis of drug binding pockets and repurposing opportunities for twelve essential enzymes of ESKAPE pathogens. *PeerJ*, 5, 2017.

[141] Azam MA and Jayaram U. Inhibitors of alanine racemase enzyme: a review. *Journal of enzyme inhibition and medicinal chemistry*, 31(4):517–526, 2016.

[142] Keren L, Hausser J, Lotan-Pompan M, Vainberg Slutskin I, Alisar H *et al.* Massively Parallel Interrogation of the Effects of Gene Expression Levels on Fitness. *Cell*, 166(5):1282–1294.e18, 2016.

[143] Palmer AC and Kishony R. Opposing effects of target overexpression reveal drug mechanisms. *Nature communications*, 5:4296, 2014.

[144] Schnappinger D. Genetic Approaches to Facilitate Antibacterial Drug Development. *Cold Spring Harbor perspectives in medicine*, 5(7):a021139, 2015.

[145] Guzman LM, Belin D, Carson MJ, and Beckwith J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *Journal of Bacteriology*, 177(14):4121–4130, 1995.

[146] Marschall L, Sagmeister P, and Herwig C. Tunable recombinant protein expression in *E. coli*: enabler for continuous processing? *Applied Microbiology and Biotechnology*, 100(13):5719–5728, 2016.

[147] Englesberg E, Irr J, Power J, and Lee N. Positive control of enzyme synthesis by gene C in the L-arabinose system. *Journal of Bacteriology*, 90(4):946–957, 1965.

[148] L-Arabinose. Https://www.abcam.com/l-arabinose-ubiquitous-naturally-occuring-saccharide-ab146251.html, Accessed: Sept. 21, 2019.

[149] Bio-siva. Http://biosiva.50webs.org/transcription.htm, Accessed: Nov. 6, 2017.

[150] Scripture JB, Voelker C, Miller S, O'Donnell RT, Polgar L *et al.* High-affinity L-arabinose transport operon. Nucleotide sequence and analysis of gene products. *Journal of Molecular Biology*, 197(1):37–46, 1987.

[151] Chen J and Gottesman S. Spot 42 Small RNA Regulates Arabinose-Inducible araBADPromoter Activity by Repressing Synthesis of the High-Affinity Low-Capacity Arabinose Transporter. *Journal of Bacteriology*, 199(3):e00691–16, 2017.

[152] Siegele DA and Hu JC. Gene expression from plasmids containing the araBAD promoter at subsaturating inducer concentrations represents mixed populations. *Proceedings of the National Academy of Sciences of the United States of America*, 94(15):8168–8172, 1997.

[153] Fritz G, Megerle JA, Westermayer SA, Brick D, Heermann R *et al.* Single cell kinetics of phenotypic switching in the arabinose utilization system of E. coli. *PloS one*, 9(2):e89532, 2014.

[154] Reeder T and Schleif R. Mapping, sequence, and apparent lack of function of araJ, a gene of the *Escherichia coli* arabinose regulon. *Journal of Bacteriology*, 173(24):7765–7771, 1991.

[155] Carolé S, Pichoff S, and Bouch JP. *Escherichia coli* gene ydeA encodes a major facilitator pump which exports L-arabinose and isopropyl-beta-D-thiogalactopyranoside. *Journal of Bacteriology*, 181(16):5123–5125, 1999.

[156] Bost S, Silva F, and Belin D. Transcriptional activation of ydeA, which encodes a member of the major facilitator superfamily, interferes with arabinose accumulation and induction of the *Escherichia coli* arabinose PBAD promoter. *Journal of Bacteriology*, 181(7):2185–2191, 1999.

[157] Keasling JD, Wanner BL, Skaug T, Datsenko KA, and Khlebnikov A. Homogeneous expression of the PBAD promoter in *Escherichia coli* by constitutive expression of the low-affinity high-

capacity AraE transporter. *Microbiology*, 147(12):3241–3247, 2001.

[158] Khlebnikov A, Skaug T, and Keasling JD. Modulation of gene expression from the arabinose-inducible araBAD promoter. *Journal of industrial microbiology & biotechnology*, 29(1):34–37, 2002.

[159] Khlebnikov A, Risa O, Skaug T, Carrier TA, and Keasling JD. Regulatable arabinose-inducible gene expression system with consistent control in all cells of a culture. *Journal of Bacteriology*, 182(24):7029–7034, 2000.

[160] Webber MA and Piddock LJV. The importance of efflux pumps in bacterial antibiotic resistance. *Journal of Antimicrobial Chemotherapy*, 51(1):9–11, 2003.

[161] Hinchliffe P, Symmons MF, Hughes C, and Koronakis V. Structure and operation of bacterial tripartite pumps. *Annual review of microbiology*, 67(1):221–242, 2013.

[162] Koronakis V, Sharff A, Koronakis E, Luisi B, and Hughes C. Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature*, 405(6789):914–919, 2000.

[163] Greene NP, Kaplan E, Crow A, and Koronakis V. Antibiotic Resistance Mediated by the MacB ABC Transporter Family: A Structural and Functional Perspective. *Frontiers in microbiology*, 9:1923, 2018.

[164] Schaller A, Guo M, Gisanrin O, and Zhang Y. *Escherichia coli* genes involved in resistance to pyrazinoic acid, the active compo-

nent of the tuberculosis drug pyrazinamide. *FEMS Microbiology Letters*, 211(2):265–270, 2002.

[165] Niederweis M, Danilchanka O, Huff J, Hoffmann C, and Engelhardt H. Mycobacterial outer membranes: in search of proteins. *Trends in microbiology*, 18(3):109–116, 2010.

[166] Miyada CG, Stoltzfus L, and Wilcox G. Regulation of the araC gene of *Escherichia coli*: catabolite repression, autoregulation, and effect on araBAD expression. *Proceedings of the National Academy of Sciences*, 81(13):4120–4124, 1984.

[167] Park SW, Casalena DE, Wilson DJ, Dai R, Nag PP *et al.* Target-based identification of whole-cell active inhibitors of biotin biosynthesis in *Mycobacterium tuberculosis*. *Chemistry & Biology*, 22(1):76–86, 2015.

[168] Cacace E, Kritikos G, and Typas A. Chemical genetics in drug discovery. *Current Opinion in Systems Biology*, 4:35–42, 2017.

[169] Natural Products Branch (NPB) | Developmental Therapeutics Program (DTP). Https://dtp.cancer.gov/organization/npb/default.htm, Accessed: Sept. 21, 2019.

[170] Part:BBa J23106 - parts.igem.org. Http://parts.igem.org/Part:BBa_J23106, Accessed: Aug. 5, 2019.

[171] Schleif R. AraC protein, regulation of the l-arabinose operon in *Escherichia coli*, and the light switch mechanism of AraC action. *FEMS Microbiology Reviews*, 34(5):779–796, 2010.

[172] Strych U, Penland RL, Jimenez M, Krause KL, and Benedik MJ. Characterization of the alanine racemases from two mycobacteria. *FEMS Microbiology Letters*, 196(2):93–98, 2001.

[173] Ian Chopra MR. Tetracycline Antibiotics: Mode of Action, Applications, Molecular Biology, and Epidemiology of Bacterial Resistance. *Microbiology and Molecular Biology Reviews*, 65(2):232–260, 2001.

[174] Daly JW, Ferreira D, Gould SJ, Haslam E, Robins DJ *et al.* Streptonigrin. In *Progress in the Chemistry of Organic Natural Products*, 77–114. Springer, Vienna, 1982.

[175] Coss E, Kealey C, Brady D, and Walsh P. A laboratory investigation of the antimicrobial activity of a selection of western phytomedicinal tinctures. *European Journal of Integrative Medicine*, 19:80–83, 2018.

[176] Hsu CW, Chuang SM, Wu WL, and Hou MH. The Crucial Role of Divalent Metal Ions in the DNA-Acting Efficacy and Inhibition of the Transcription of Dimeric Chromomycin A3. *PloS one*, 7(9):e43792, 2012.

[177] Rai D, Singh JK, Roy N, and Panda D. Curcumin inhibits ftsz assembly: an attractive mechanism for its antibacterial activity. *Biochemical Journal*, 410(1):147–155, 2008.

[178] World Health Organaization. WHO consolidated guidelines on drug-resistant tuberculosis treatment, 2019.

[179] Dey S, Lane JM, Lee RE, Rubin EJ, and Sacchettini JC. Structural characterization of the *Mycobacterium tuberculosis* biotin biosynthesis enzymes 7,8-diaminopelargonic acid synthase and dethiobiotin synthetase . *Biochemistry*, 49(31):6746–6760, 2010.

[180] Li SL, Hanlon J, and Yanofsky C. Separation of anthranilate synthetase components I and II of *Escherichia coli*, Salmonella typhimurium, and Serratia marcescens and determination of their amino-terminal sequences by automatic Edman degradation. *Biochemistry*, 13(8):1736–1744, 1974.

[181] Taniguchi Y, Choi PJ, Li GW, Chen H, Babu M *et al.* Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. *Science*, 329(5991):533–538, 2010.

[182] McGinness KE, Baker TA, and T SR. Engineering Controllable Protein Degradation. *Molecular Cell*, 22:701–707, 2006.

[183] Salis HM, Mirsky EA, and Voigt CA. Automated design of synthetic ribosome binding sites to control protein expression. *Nature Biotechnology*, 27(10):946–950, 2009.

[184] Bertram R and Hillen W. The application of Tet repressor in prokaryotic gene regulation and expression. *Microbial Biotechnology*, 1(1):2–16, 2008.

[185] Li XT, Jun Y, Erickstad MJ, Brown SD, Parks A *et al.* tCRISPRi: tunable and reversible, one-step control of gene expression. *Scientific reports*, 6(1):39076, 2016.

[186] Seabold RR and Schleif RF. Apo-AraC actively seeks to loop. *Journal of Molecular Biology*, 278(3):529–538, 1998.

[187] Datsenko KA and Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America*, 97(12):6640–6645, 2000.

[188] Ahrweiler PM and Frieden C. Construction of a fol mutant strain of *Escherichia coli* for use in dihydrofolate reductase mutagenesis experiments. *Journal of Bacteriology*, 170(7):3301–3304, 1988.

[189] Sambrook J, Fritsch EF, and Maniatis T. Molecular Cloning: A Laboratory Manual. 2 nd Edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 1989.

[190] Dean AS, Cox H, and Zignol M. Epidemiology of Drug-Resistant Tuberculosis. *Advances in experimental medicine and biology*, 1019(Suppl 3):209–220, 2017.

[191] Mdluli K, Kaneko T, and Upton A. The tuberculosis drug discovery and development pipeline and emerging drug targets. *Cold Spring Harbor perspectives in medicine*, 5(6):a021154–a021154, 2015.

[192] Franzblau SG, DeGroote MA, Cho SH, Andries K, Nuermberger E *et al.* Comprehensive analysis of methods used for the evaluation of compounds against *Mycobacterium tuberculosis*. *Tuberculosis (Edinburgh, Scotland)*, 92(6):453–488, 2012.

[193] Stanley SA and Cox JS. Host–Pathogen Interactions During *Mycobacterium tuberculosis* infections. In *Pathogenesis of Mycobacterium tuberculosis and its Interaction with the Host Organism*, 211–241. Springer, Berlin, Heidelberg, Berlin, Heidelberg, 2013.

[194] Dutta NK and Karakousis PC. Latent tuberculosis infection: myths, models, and molecular mechanisms. *Microbiology and molecular biology reviews : MMBR*, 78(3):343–371, 2014.

[195] Alladi SM. Advances in Computational Studies of Potential Drug Targets in *Mycobacterium tuberculosis*. *Current topics in medicinal chemistry*, 18(13):1062–1074, 2018.

[196] Campaniço A, Moreira R, and Lopes F. Drug discovery in tuberculosis. New drug targets and antimycobacterial agents. *European journal of medicinal chemistry*, 150:525–545, 2018.

[197] Chiarelli LR, Mori G, Esposito M, Orena BS, and Pasca MR. New and Old Hot Drug Targets in Tuberculosis. *Current Medicinal Chemistry*, 23(33):3813–3846, 2016.

[198] Sams-Dodd F. Target-based drug discovery: is something wrong? *Drug discovery today*, 10(2):139–147, 2005.

[199] Payne DJ, Gwynn MN, Holmes DJ, and Pompliano DL. Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nature reviews. Drug discovery*, 6(1):29–40, 2007.

[200] Silver LL. Challenges of antibacterial discovery. *Clinical microbiology reviews*, 24(1):71–109, 2011.

[201] Machado D, Girardini M, Viveiros M, and Pieroni M. Challenging the Drug-Likeness Dogma for New Drug Discovery in Tuberculosis. *Frontiers in microbiology*, 9:1367, 2018.

[202] Banerjee A, Dubnau E, Quemard A, Balasubramanian V, Um KS *et al.* inhA, a gene encoding a target for isoniazid and ethionamide in *Mycobacterium tuberculosis*. *Science*, 263(5144):227–230, 1994.

[203] Goude R, Amin AG, Chatterjee D, and Parish T. The arabinosyltransferase EmbC is inhibited by ethambutol in *Mycobacterium tuberculosis*. *Antimicrobial Agents and Chemotherapy*, 53(10):4138–4146, 2009.

[204] Bonnett SA, Ollinger J, Chandrasekera S, Florio S, O'Malley T *et al.* A Target-Based Whole Cell Screen Approach To Identify

Potential Inhibitors of *Mycobacterium tuberculosis* Signal Peptidase. *ACS infectious diseases*, 2(12):893–902, 2016.

[205] Melief E, Kokoczka R, Files M, Bailey MA, Alling T *et al.* Construction of an overexpression library for *Mycobacterium tuberculosis*. *Biology methods & protocols*, 3(1):bpy009, 2018.

[206] Abrahams GL, Kumar A, Savvi S, Hung AW, Wen S *et al.* Pathway-selective sensitization of *Mycobacterium tuberculosis* for target-based whole-cell screening. *Chemistry & Biology*, 19(7):844–854, 2012.

[207] Johnson EO, LaVerriere E, Office E, Stanley M, Meyer E *et al.* Large-scale chemical-genetics yields new M. tuberculosis inhibitor classes. *Nature*, 1, 2019.

[208] Chhotaray C, Tan Y, Mugweru J, Islam MM, Adnan Hameed HM *et al.* Advances in the development of molecular genetic tools for *Mycobacterium tuberculosis*. *Journal of genetics and genomics = Yi chuan xue bao*, 45(6):281–297, 2018.

[209] Morona R, Manning PA, and Reeves P. Identification and characterization of the TolC protein, an outer membrane protein from *Escherichia coli*. *Journal of Bacteriology*, 153(2):693–699, 1983.

[210] Lewis K. Translocases: a bacterial tunnel for drugs and proteins. *Current biology : CB*, 10(18):R678–81, 2000.

[211] Nichols RJ, Sen S, Choo YJ, Beltrao P, Zietek M *et al.* Phenotypic Landscape of a Bacterial Cell. *Cell*, 144(1):143–156, 2011.

[212] Typas A, Banzhaf M, Gross CA, and Vollmer W. From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nature reviews. Microbiology*, 10(2):123–136, 2011.

[213] Fu LM and Fu-Liu CS. Is *Mycobacterium tuberculosis* a closer relative to Gram-positive or Gram-negative bacterial pathogens? *Tuberculosis (Edinburgh, Scotland)*, 82(2-3):85–90, 2002.

[214] Anthony KG, Strych U, Yeung KR, Shoen CS, Perez O *et al.* New classes of alanine racemase inhibitors identified by high-throughput screening show antimicrobial activity against *Mycobacterium tuberculosis*. *PloS one*, 6(5):e20374, 2011.

[215] Nakatani Y, Opel-Reading HK, Merker M, Machado D, Andres S *et al.* Role of Alanine Racemase Mutations in *Mycobacterium tuberculosis* d-Cycloserine Resistance. *Antimicrobial Agents and Chemotherapy*, 61(12):e01575–17, 2017.

[216] Vega DE and Young KD. Accumulation of periplasmic enterobactin impairs the growth and morphology of *Escherichia coli*tolC mutants. *Molecular Microbiology*, 91(3):508–521, 2014.

[217] Krishnamoorthy G, Wolloscheck D, Weeks JW, Croft C, Rybenkov VV *et al.* Breaking the Permeability Barrier of *Escherichia coli* by Controlled Hyperporination of the Outer Membrane. *Antimicrobial Agents and Chemotherapy*, 60(12):7372–7381, 2016.

[218] Sulavik MC, Houseweart C, Cramer C, Jiwani N, Murgolo N *et al.* Antibiotic susceptibility profiles of *Escherichia coli* strains lacking multidrug efflux pump genes. *Antimicrobial Agents and Chemotherapy*, 45(4):1126–1136, 2001.

[219] Zhang XD. Illustration of ssmd, z score, ssmd*, z* score, and t statistic for hit selection in rnai high-throughput screens. *Journal of Biomolecular Screening*, 16(7):775–785, 2011.

[220] Paterson DL and Bonomo RA. Extended-spectrum beta-lactamases: a clinical update. *Clinical microbiology reviews*, 18(4):657–686, 2005.

[221] Payen MC, Muylle I, Vandenberg O, Mathys V, Delforge M *et al.* Meropenem-clavulanate for drug-resistant tuberculosis: a follow-up of relapse-free cases. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease*, 22(1):34–39, 2018.

[222] Ramón-García S, González Del Río R, Villarejo AS, Sweet GD, Cunningham F *et al.* Repurposing clinically approved cephalosporins for tuberculosis therapy. *Scientific reports*, 6(1):34293, 2016.

[223] Zimmermann GR, Lehár J, and Keith CT. Multi-target therapeutics: when the whole is greater than the sum of the parts. *Drug discovery today*, 12(1-2):34–42, 2007.

[224] Prosser GA and de Carvalho LPS. Metabolomics Reveal d-Alanine: d-Alanine Ligase As the Target of d-Cycloserine in *Mycobacterium tuberculosis*. *ACS Medicinal Chemistry Letters*, 4(12):1233–1237, 2013.

[225] Du D, Chang CH, Wang Y, Tong P, Chan WK *et al.* Response envelope analysis for quantitative evaluation of drug combinations. *Bioinformatics (Oxford, England)*, 30:679, 2019.

[226] Berenbaum MC. The expected effect of a combination of agents: the general solution. *Journal of theoretical biology*, 114(3):413–431, 1985.

[227] LOEWE S. Effect of combinations : mathematical basis of problem. *Arch. Exp. Pathol. Pharmakol.*, 114(5-6):313–326, 1926.

[228] Foucquier J and Guedj M. Analysis of drug combinations: current methodological landscape. *Pharmacology Research & Perspectives*, 3(3):e00149, 2015.

[229] Whalen JJ. Definition of the effective dose of the converting-enzyme inhibitor benazepril. *American heart journal*, 117(3):728–734, 1989.

[230] Cupp C. Enzymatic Assay of D-Alanine Racemase (EC 5.1.1.1), 2004. Https://www.sigmaaldrich.com/technical-documents/protocols/biology/enzymatic-assay-of-d-alanine-racemase.html, Accessed: Jan. 11, 2018.

[231] Yoshino M and Murakami K. A graphical method for determining inhibition constants. *Journal of enzyme inhibition and medicinal chemistry*, 24(6):1288–1290, 2009.

[232] Palmer AC, Chait R, and Kishony R. Nonoptimal Gene Expression Creates Latent Potential for Antibiotic Resistance. *Molecular biology and evolution*, 35(11):2669–2684, 2018.

[233] Daniel R, Rubens JR, Sarpeshkar R, and Lu TK. Synthetic analog computation in living cells. *Nature*, 497(7451):619–623, 2013.

[234] Gallagher LA, Shears RK, Fingleton C, Alvarez L, Waters EM *et al.* Impaired alanine transport or exposure to D-cycloserine increases the susceptibility of MRSA to beta-lactam antibiotics. *biorxiv.org*, 2019.

[235] Sieradzki K and Tomasz A. Suppression of beta-lactam antibiotic resistance in a methicillin-resistant Staphylococcus aureus through synergic action of early cell wall inhibitors and

some other antibiotics. *Journal of Antimicrobial Chemotherapy*, $39(\text{suppl}_1) : 47 - -51, 1997$.

[236] Prosser GA, Rodenburg A, Khoury H, de Chiara C, Howell S *et al.* Glutamate Racemase Is the Primary Target of β-Chloro-d-Alanine in *Mycobacterium tuberculosis*. *Antimicrobial Agents and Chemotherapy*, 60(10):6091–6099, 2016.

[237] Roell KR, Reif DM, and Motsinger-Reif AA. An Introduction to Terminology and Methodology of Chemical Synergy-Perspectives from Across Disciplines. *Frontiers in pharmacology*, 8:158, 2017.

[238] Jia J, Zhu F, Ma X, Cao Z, Cao ZW *et al.* Mechanisms of drug combinations: interaction and network perspectives. *Nature reviews. Drug discovery*, 8(2):111–128, 2009.

[239] Cokol M, Chua HN, Tasan M, Mutlu B, Weinstein ZB *et al.* Systematic exploration of synergistic drug pairs. *Molecular Systems Biology*, 7(1):544, 2011.

[240] Zhang Y, Scorpio A, Nikaido H, and Sun Z. Role of acid pH and deficient efflux of pyrazinoic acid in unique susceptibility of *Mycobacterium tuberculosis* to pyrazinamide. *Journal of Bacteriology*, 181(7):2044–2049, 1999.

[241] Kanvatirth P, Jeeves RE, Bacon J, Besra GS, and Alderwick LJ. Utilisation of the Prestwick Chemical Library to identify drugs that inhibit the growth of mycobacteria. *PloS one*, 14(3):e0213713, 2019.

[242] Wu JY, Lee MTG, Lee SH, Lee SH, Tsai YW *et al.* Angiotensin-Converting Enzyme Inhibitors and Active Tuberculosis: A Population-Based Study. *Medicine*, 95(19):e3579, 2016.

[243] Chen H, Nyantakyi SA, Li M, Gopal P, Aziz DB *et al.* The Mycobacterial Membrane: A Novel Target Space for Anti-tubercular Drugs. *Frontiers in microbiology*, 9:1627, 2018.

[244] Chen H, Nyantakyi SA, Li M, Gopal P, Aziz DB *et al.* The Mycobacterial Membrane: A Novel Target Space for Anti-tubercular Drugs. *Frontiers in microbiology*, 9:1285, 2018.

[245] Lohrasbi V, Talebi M, Bialvaei AZ, Fattorini L, Drancourt M *et al.* Trends in the discovery of new drugs for *Mycobacterium tuberculosis* therapy with a glance at resistance. *Tuberculosis (Edinburgh, Scotland)*, 109:17–27, 2018.

[246] Awasthy D, Bharath S, Subbulakshmi V, and Sharma U. Alanine racemase mutants of *Mycobacterium tuberculosis* require D-alanine for growth and are defective for survival in macrophages and mice. *Microbiology (Reading, England)*, 158(Pt 2):319–327, 2012.

[247] Halouska S, Fenton RJ, Zinniel DK, Marshall DD, Barletta RG *et al.* Metabolomics analysis identifies d-Alanine-d-Alanine ligase as the primary lethal target of d-Cycloserine in mycobacteria. *Journal of proteome research*, 13(2):1065–1076, 2014.

[248] Mortuza R, Aung HL, Taiaroa G, Opel-Reading HK, Kleffmann T *et al.* Overexpression of a newly identified d-amino acid transaminase in *Mycobacterium smegmatis* complements glutamate racemase deletion. *Molecular Microbiology*, 107(2):198–213, 2018.

[249] Cáceres NE, Harris NB, Wellehan JF, Feng Z, Kapur V *et al.* Overexpression of the D-alanine racemase gene confers resistance to D-cycloserine in *Mycobacterium smegmatis*. *Journal of Bacteriology*, 179(16):5046–5055, 1997.

[250] Cronan JE. *Escherichia coli as an Experimental Organism*. American Cancer Society, 2014.

[251] Edelheit O, Hanukoglu A, and Hanukoglu I. Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies. *BMC biotechnology*, 9(1):61, 2009.

[252] UNITAID. TUBERCULOSIS - Diagnostics Technology and Market Landscape. 1–97, 2014.

[253] Evangelopoulos D, Prosser GA, Rodgers A, Dagg BM, Khatri B *et al.* Comparative fitness analysis of D-cycloserine resistant mutants reveals both fitness-neutral and high-fitness cost genotypes. *Nature communications*, 10(1):4177–11, 2019.

[254] Bankier RG. Psychisis Associated with Cycloserine. *Canadian Medical Association journal*, 93(1):35–37, 1965.

[255] Curry International Tuberculosis Center. Drug-Resistant Tuberculosis: A Survival Guide for Clinicians, 3rd edition. 1–324, 2016.

[256] Kam KM, Sloutsky A, Yip CW, Bulled N, Seung KJ *et al.* Determination of critical concentrations of second-line anti-tuberculosis drugs with clinical and microbiological relevance. *The International Journal of Tuberculosis and Lung Disease*, 14(3):282–288.

[257] Reller LB, Weinstein MP, and Woods GL. Susceptibility Testing for Mycobacteria. *Clinical Infectious Diseases*, 31(5):1209–1215, 2000.

[258] Lambert MP and Neuhaus FC. Mechanism of D-cycloserine action: alanine racemase from *Escherichia coli* W. *Journal of Bacteriology*, 110(3):978–987, 1972.

[259] EPSTEIN IG, NAIR KG, and BOYD LJ. Cycloserine, a new antibiotic, in the treatment of human pulmonary tuberculosis: a preliminary report. *Antibiotic medicine & clinical therapy (New York, NY)*, 1(2):80–93, 1955.

[260] Feng Z and Barletta RG. Roles of *Mycobacterium smegmatis* D-alanine:D-alanine ligase and D-alanine racemase in the mechanisms of action of and resistance to the peptidoglycan inhibitor D-cycloserine. *Antimicrobial Agents and Chemotherapy*, 47(1):283–291, 2003.

[261] Köser CU, Bryant JM, Becq J, Török ME, Ellington MJ *et al.* Whole-genome sequencing for rapid susceptibility testing of M. tuberculosis. *The New England journal of medicine*, 369(3):290–292, 2013.

[262] Chen JM, Uplekar S, Gordon SV, and Cole ST. A point mutation in cycA partially contributes to the D-cycloserine resistance trait of Mycobacterium bovis BCG vaccine strains. *PloS one*, 7(8):e43467, 2012.

[263] Coll F, Phelan J, Hill-Cawthorne GA, Nair MB, Mallard K *et al.* Genome-wide analysis of multi- and extensively drug-resistant *Mycobacterium tuberculosis*. *Nature genetics*, 50(2):307–316, 2018.

[264] Aligent Technologies, Inc. GeneMorph II Random Mutagenesis Kit, 2015. Https://www.agilent.com/cs/library/usermanuals/public/200550.pdf, Accessed: Oct. 10, 2019.

[265] Bosley AD and Ostermeier M. Mathematical expressions useful in the construction, description and evaluation of protein libraries. *Biomolecular engineering*, 22(1-3):57–61, 2005.

[266] Fehér T, Cseh B, Umenhoffer K, Karcagi I, and Pósfai G. Characterization of cycA mutants of *Escherichia coli*. An assay for measuring in vivo mutation rates. *Mutation research*, 595(1-2):184–190, 2006.

[267] Baisa G, Stabo NJ, and Welch RA. Characterization of *Escherichia coli* D-cycloserine transport and resistant mutants. *Journal of Bacteriology*, 195(7):1389–1399, 2013.

[268] Pulvermacher SC, Stauffer LT, and Stauffer GV. The Small RNA GcvB Regulates sstT mRNA Expression in *Escherichia coli*. *Journal of Bacteriology*, 191(1):238–248, 2009.

[269] Tokuriki N and Tawfik DS. Stability effects of mutations and protein evolvability. *Current opinion in structural biology*, 19(5):596–604, 2009.

[270] Portelli S, Phelan JE, Ascher DB, Clark TG, and Furnham N. Understanding molecular consequences of putative drug resistant mutations in *Mycobacterium tuberculosis*. *Scientific reports*, 8(1):15356, 2018.

[271] Loewe L and Hill WG. The population genetics of mutations: good, bad and indifferent. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 2010.

[272] Borrell S and Gagneux S. Infectiousness, reproductive fitness and evolution of drug-resistant *Mycobacterium tuberculosis*. *The International Journal of Tuberculosis and Lung Disease*, 13(12):1456–1466, 2009.

[273] Tokuriki N and Tawfik DS. Stability effects of mutations and protein evolvability. *Current opinion in structural biology*, 19(5):596–604, 2009.

[274] Canova MJ, Kremer L, and Molle V. The *Mycobacterium tuberculosis* GroEL1 chaperone is a substrate of Ser/Thr protein kinases. *Journal of Bacteriology*, 191(8):2876–2883, 2009.

[275] King JL and Jukes TH. Non-Darwinian evolution. *Science*, 164(3881):788–798, 1969.

[276] Zwart MP, Schenk MF, Hwang S, Koopmanschap B, de Lange N *et al.* Unraveling the causes of adaptive benefits of synonymous mutations in TEM-1 β-lactamase. *Heredity*, 121(5):406–421, 2018.

[277] Engler, Carola, Kandzia, Romy, and Marillonnet, Sylvestre. A One Pot, One Step, Precision Cloning Method with High Throughput Capability. *PloS one*, 3(11):e3647, 2008.

[278] Barrick JE, Colburn G, Deatherage DE, Traverse CC, Strand MD *et al.* Identifying structural variation in haploid microbial genomes from short-read resequencing data using breseq. *BMC genomics*, 15(1):1–17, 2014.

[279] LeMagueres P, Im H, Ebalunode J, Strych U, Benedik MJ *et al.* The 1.9 A crystal structure of alanine racemase from *Mycobacterium tuberculosis* contains a conserved entryway into the active site, 2005.

[280] Watanabe A, Yoshimura T, Mikami B, Hayashi H, Kagamiyama H *et al.* Reaction mechanism of alanine racemase from bacillus stearothermophilus : X-ray crystallographic studies of the enzyme bound withn-(5-phosphopyridoxyl)alanine. *Journal of Biological Chemistry*.

[281] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM *et al.* Ucsf chimera—a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 2004.

[282] Fenn TD, Stamper GF, Morollo AA, and Ringe D. A Side Reaction of Alanine Racemase: Transamination of Cycloserine †,‡. *Biochemistry*, 42(19):5775–5783, 2003.

[283] Wang J, Wolf RM, Caldwell JW, Kollman PA, and Case DA. Development and testing of a general amber force field. *Journal of Computational Chemistry*, 2004.

[284] Wang J, Cieplak P, and Kollman PA. How well does a restrained electrostatic potential (resp) model perform in calculating conformational energies of organic and biological molecules? *Journal of Computational Chemistry*, 21(12):1049–1074, 2000.

[285] Hanwell MD, Curtis DE, Lonie DC, Vandermeersch T, Zurek E *et al.* Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *Journal of Cheminformatics*, 4(1):1–17, 2012.

[286] Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA *et al.* Gaussian09 Revision E.01. Gaussian Inc. Wallingford CT 2009.

[287] Bayly CI, Cieplak P, Cornell W, and Kollman PA. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the resp model. *The Journal of Physical Chemistry*, 97(40):10269–10280, 1993.

[288] Sousa da Silva AW and Vranken WF. ACPYPE - AnteChamber PYthon Parser interfacE. *BMC Research Notes*, 5(1):367, 2012.

[289] Hornak V, Abel R, Okur A, Strockbine B, Roitberg A *et al.* Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics*, 65(3):712–725, 2006.

[290] Larsen KL, Piana S, Palmo K, Maragakis P, Klepeis JL *et al.* Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Structure, Function, and Bioinformatics*, 78(8):1950–1958, 2010.

[291] Best RB and Hummer G. Optimized molecular dynamics force fields applied to the helixcoil transition of polypeptides. *The Journal of Physical Chemistry B*, 113(26):9004–9015, 2009.

[292] Piana S, Lindorff-Larsen K, and Shaw DE. How Robust Are Protein Folding Simulations with Respect to Force Field Parameterization? *Biophysical journal*, 100(9):L47–L49, 2011.

[293] Horn HW, Swope WC, Pitera JW, Madura JD, Dick TJ *et al.* Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *The Journal of Chemical Physics*, 120(20):9665–9678, 2004.

[294] Bussi G, Donadio D, and Parrinello M. Canonical sampling through velocity rescaling. *The Journal of Chemical Physics*, 126(1):014101, 2007.

[295] Parrinello M and Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics*, 52(12):7182–7190, 1998.

[296] Nosé S and Klein ML. Constant pressure molecular dynamics for molecular systems. *Molecular Physics*, 50(5):1055–1076, 2006.

[297] Darden T, York D, and Pedersen L. Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *The Journal of Chemical Physics*, 98(12):10089–10092, 1998.

[298] Hess B, Bekker H, Berendsen HJC, and Fraaije JGEM. LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, 18(12):1463–1472, 1997.

[299] Hess B. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *Journal of Chemical Theory and Computation*, 4(1):116–122, 2007.

[300] Miyamoto S and Kollman PA. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of Computational Chemistry*, 13(8):952–962, 1992.

[301] Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC *et al.* Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1-2:19 – 25, 2015.

[302] World Health Organization. Global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics. Https://www.who.int/medicines/publications/WHO-PPL-Short_Summary_25Feb-ET_NM_WHO.pdf, Accessed: Oct. 25, 2019.

[303] Morel N and Massoulié J. Comparative Expression of Homologous Proteins. A NOVEL MODE OF TRANSCRIPTIONAL REGULATION BY THE CODING SEQUENCE FOLDING COMPATIBILITY OF CHIMERAS. *The Journal of biological chemistry*, 275(10):7304–7312, 2000.

[304] Dougan G, Crosa JH, and Falkow S. Mobilization of the *Escherichia coli* Plasmid ColE1 (Colicin E1) and ColE1 Vectors Used in Recombinant DNA Experiments. *The Journal of Infectious Diseases*, 137, 1978.

[305] Selgelid MJ. Gain-of-Function Research: Ethical Analysis. *Science and Engineering Ethics*, 22(4):923–964, 2016.

Appendix A

Chapter I

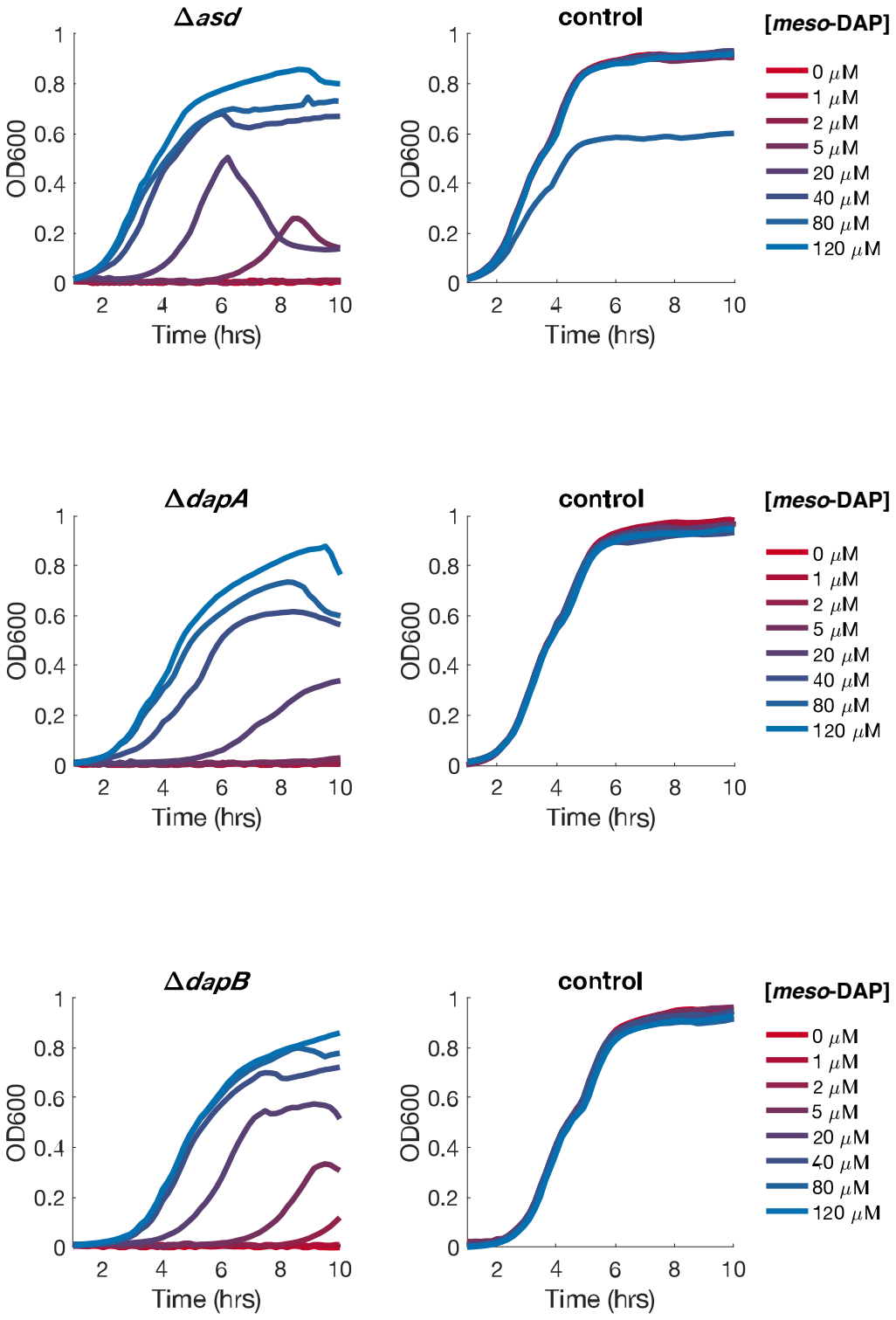

Figure A.1: Growth rescue of auxotrophic *E. coli* deletion strains through the external supplementation of amino acids.
meso-DAP supplementation in LB medium restored the growth of deletion strains K001 (Δ*asd*), K002 (Δ*dapA*) and K003 (Δ*dapB*). Strain D20 (parent strain) served as the control strain in all experiments.

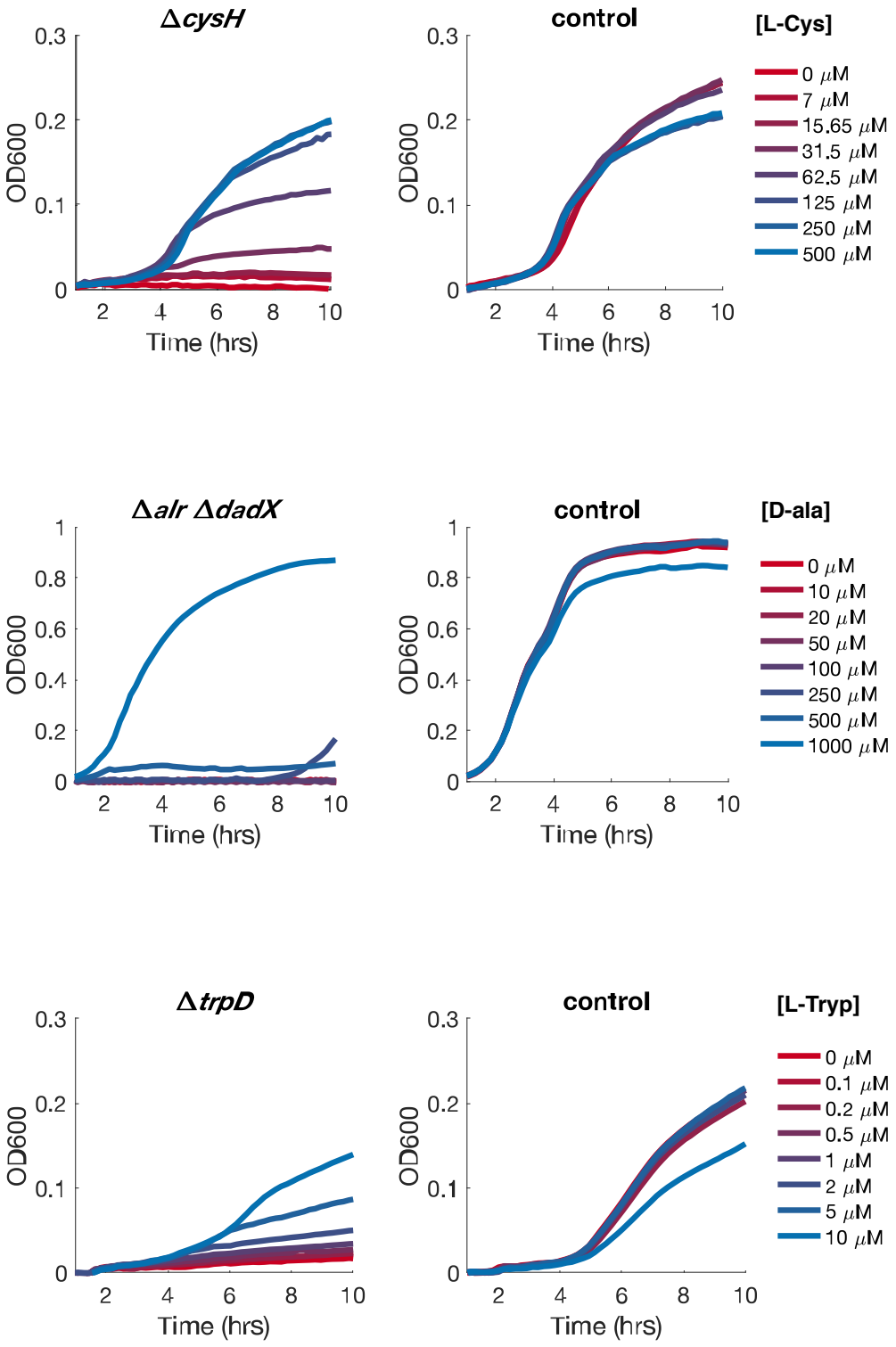

Figure A.2: Growth rescue of auxotrophic *E. coli* deletion strains through the external supplementation of amino acids.

The supplementation of D-alanine in LB medium restored the growth of D18 ($\Delta alr \Delta dadX$) and the addition of L-cysteine to M9 medium with 0.8% glucose restored the growth of a $\Delta cysH$. L-tryptophan was added to M9 medium with 0.4% glucose to restore the growth of $\Delta trpD$ and $\Delta trpE$ (next page).

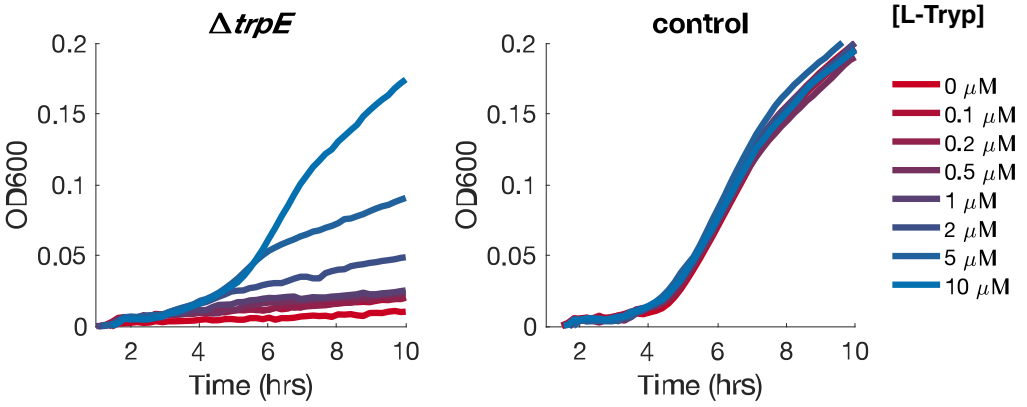

Figure A.3: Growth rescue of auxotrophic *E. coli* deletion strains through the external supplementation of amino acids.

	Strain ID	Host strain	Plasmid ID	Part description	Pre-/suffix	Antibiotic Resistance Marker
				Level 0 Destination Vectors		
	MC.001	Dh5-α	DVC_AB	LacZ	AB	Chloramphenicol
	MC.002	Dh5-α	DVC_AC	LacZ	AC	Chloramphenicol
	MC.003	Dh5-α	DVC_BC	LacZ	BD	Chloramphenicol
	MC.004	Dh5-α	DVC_CD	LacZ	CD	Chloramphenicol
	MC.005	Dh5-α	DVC_CE	LacZ	CE	Chloramphenicol
	MC.006	Dh5-α	DVC_CF	LacZ	CF	Chloramphenicol
	MC.007	Dh5-α	DVC_DG	LacZ	DG	Chloramphenicol
	MC.008	Dh5-α	DVC_DE	LacZ	DE	Chloramphenicol
	MC.009	Dh5-α	DVC_FG	LacZ	FG	Chloramphenicol
	MC.010	Dh5-α	DVC_FH	LacZ	FH	Chloramphenicol
	MC.011	Dh5-α	DVC_EF	LacZ	EF	Chloramphenicol
	MC.046	Dh5-α	DVC_BF	LacZ	BF	Chloramphenicol
	MC.053	Dh5-α	DVC_KF	LacZ	KF	Chloramphenicol
	MC.054	Dh5-α	DVC_CK	LacZ	CK	Chloramphenicol
				Level 0 Plasmids with parts		
pBAD - RBS variants	MC.017	Dh5-α	L0_006	p_{BAD}-RBS1	AC	Chloramphenicol
	MC.070	Dh5-α	L0_048	*araC* - p_c - p_{BAD} - RBS1 (wild type)	AC	Chloramphenicol
	MC.071	Dh5-α	L0_049	*araC* - p_c - p_{BAD} - RBS1 (O2>I1)	AC	Chloramphenicol
	MC.073	Dh5-α	L0_050	p_{BAD}-RBS1 (O2>I1)	AC	Chloramphenicol
RBS Library *	MC.088	Dh5-α	L0_065	p_{BAD}-RBS Salis (T.I.R 1)_AC	AC	Chloramphenicol
	MC.089	Dh5-α	L0_066	p_{BAD}-RBS Salis (T.I.R 2)_AC	AC	Chloramphenicol
	MC.090	Dh5-α	L0_067	p_{BAD}-RBS Salis (T.I.R 3)_AC	AC	Chloramphenicol
	MC.091	Dh5-α	L0_068	p_{BAD}-RBS Salis (T.I.R 10)_AC	AC	Chloramphenicol
	MC.092	Dh5-α	L0_069	p_{BAD}-RBS Salis (T.I.R 11)_AC	AC	Chloramphenicol
	MC.093	Dh5-α	L0_070	p_{BAD}-RBS Salis (T.I.R 12)_AC	AC	Chloramphenicol
	MC.094	Dh5-α	L0_071	p_{BAD}-RBS Salis (T.I.R 50)_AC	AC	Chloramphenicol
	MC.095	Dh5-α	L0_072	p_{BAD}-RBS Salis (T.I.R 100)_AC	AC	Chloramphenicol
	MC.096	Dh5-α	L0_073	p_{BAD}-RBS Salis (T.I.R 110)_AC	AC	Chloramphenicol
	MC.097	Dh5-α	L0_074	p_{BAD}-RBS Salis (T.I.R 120)_AC	AC	Chloramphenicol
	MC.098	Dh5-α	L0_075	p_{BAD}-RBS Salis (T.I.R 1000-1)_AC	AC	Chloramphenicol
	MC.099	Dh5-α	L0_076	p_{BAD}-RBS Salis (T.I.R 1000-2)_AC	AC	Chloramphenicol
	MC.100	Dh5-α	L0_077	p_{BAD}-RBS Salis (T.I.R 1100)_AC	AC	Chloramphenicol
Transcroption Factors & Reporters	MC.020	Dh5-α	L0_009	linker-*mCherry*	EF	Chloramphenicol
	MC.029	Dh5-α	L0_018	*meGFF* (no stop) - (GGGS)3	CK	Chloramphenicol
	MC.064	Dh5-α	L0_042	(GGGS)3 - meGFF (no start)	EF	Chloramphenicol
	MC.022	Dh5-α	L0_011	*luxR*	CF	Chloramphenicol
	MC.023	Dh5-α	L0_012	*luxR*-linker-*GFP*	CF	Chloramphenicol
	MC.062	Dh5-α	L0_040	*araC*	CF	Chloramphenicol
	MC.021	Dh5-α	L0_010	*araC* -linker-*GFP*	CF	Chloramphenicol
	MC.063	Dh5-α	L0_041	*araC*-linker-mCherry	CF	Chloramphenicol
	MC.066	Dh5-α	L0_044	*mE-GFP*	CF	Chloramphenicol
	MC.069	Dh5-α	L0_047	*mCherry*	CF	Chloramphenicol
	MC.030	Dh5-α	L0_019	*E.coli-asd* (no start)	KF	Chloramphenicol

Table A.0: TESEC MoClo parts and plasmids.

Essential *E. coli* genes	MC.031	Dh5-α	L0_020	*E.coli-cysH* (no start)	KF	Chloramphenicol
	MC.032	Dh5-α	L0_021	*E. coli-dadX* (no start)	KF	Chloramphenicol
	MC.033	Dh5-α	L0_022	*E.coli-dapA* (no start)	KF	Chloramphenicol
	MC.034	Dh5-α	L0_023	*E.coli-dapB* (no start)	KF	Chloramphenicol
	MC.035	Dh5-α	L0_024	*E.coli-trpD* (no start)	KF	Chloramphenicol
	MC.036	Dh5-α	L0_025	*E.coli-trpE* (no start)	KF	Chloramphenicol
	MC.074	Dh5-α	L0_051	*E.coli-asd*	CF	Chloramphenicol
	MC.068	Dh5-α	L0_046	*E. coli - asd* - LAA tag	CF	Chloramphenicol
	MC.075	Dh5-α	L0_052	*E.coli-cysH*	CF	Chloramphenicol
	MC.067	Dh5-α	L0_045	*E. coli - cysH* - LAA tag	CF	Chloramphenicol
	MC.076	Dh5-α	L0_053	*E. coli-dadX*	CF	Chloramphenicol
	MC.077	Dh5-α	L0_054	*E.coli-dapA*	CF	Chloramphenicol
	MC.078	Dh5-α	L0_055	*E.coli-dapB*	CF	Chloramphenicol
	MC.079	Dh5-α	L0_056	*E.coli-trpD*	CF	Chloramphenicol
	MC.080	Dh5-α	L0_057	*E.coli-trpE*	CF	Chloramphenicol
Essential *Mtb* genes	MC.055	Dh5-α	L0_033	*Mtb-alr* (no start)	KF	Chloramphenicol
	MC.056	Dh5-α	L0_034	*Mtb-asd* (no start)	KF	Chloramphenicol
	MC.057	Dh5-α	L0_035	*Mtb-cysH* (BpiI mutated) (no start)	KF	Chloramphenicol
	MC.058	Dh5-α	L0_036	*Mtb-dapA* (no start)	KF	Chloramphenicol
	MC.059	Dh5-α	L0_037	*Mtb-dapB* (no start)	KF	Chloramphenicol
	MC.060	Dh5-α	L0_038	*Mtb-trpD* (no start)	KF	Chloramphenicol
	MC.061	Dh5-α	L0_039	*Mtb-trpE* (no start)	KF	Chloramphenicol
	MC.081	Dh5-α	L0_058	*Mtb-alr*	CF	Chloramphenicol
	MC.082	Dh5-α	L0_059	*Mtb-asd*	CF	Chloramphenicol
	MC.083	Dh5-α	L0_060	*Mtb-cysH* (internal BpiI site removed)	CF	Chloramphenicol
	MC.084	Dh5-α	L0_061	*Mtb-dapA*	CF	Chloramphenicol
	MC.085	Dh5-α	L0_062	*Mtb-dapB*	CF	Chloramphenicol
	MC.086	Dh5-α	L0_063	*Mtb-trpD*	CF	Chloramphenicol
	MC.087	Dh5-α	L0_064	*Mtb-trpE*	CF	Chloramphenicol
Level 1 Destination Vectors **						
	MC.072	Dh5-α	DVA_p15A	LacZ	AF	Ampicillin
	MC.107	Dh5-α	DVA_colE1	LacZ	AF	Ampicillin

* T.I.R. = Transcription Initiation Rate calculated with the Salis RBS
** Destination vectors carry as terminator sequence downstream the F-suffix

Table A.0: TESEC MoClo parts and plasmids.

Strain ID	Parent strain	Relevant characteristics	Reference or constructor
BW25113		F-, Δ(*araD-araB*)567, Δ*lacZ*4787(::rrnB-3), λ-, rph-1, Δ(*rhaD-rhaB*)568, hsdR514	Datsenko and Wanner, 2000
D17	BW25113	*ΔcysH::FRT ΔaraC::Kan^r*	EW - This study
D18	BW25113	*ΔdadX::FRT Δalr::FRT ΔaraC::FRT*	EW - This study
D19	BW25113	*ΔcysH::FRT ΔaraC::FRT*	EW - This study
D20	BW25113	*ΔaraC::FRT*	EW - This study
K001	D20	*ΔaraC::FRT Δasd::FRT*	NB - This study
K002	D20	*ΔaraC::FRT ΔdapA::FRT*	NB - This study
K003	D20	*ΔaraC::FRT ΔdapB::FRT*	NB - This study
K004	D20	*ΔaraC::FRT ΔtrpD::FRT*	Sebastián Sosa Carrillo - This study
K005	D20	*ΔaraC::FRT ΔtrpE::FRT*	Sebastián Sosa Carrillo - This study
K006	K002	*ΔdapA::FRT Δasd::FRT*	NB - This study
K007	K002	*ΔdapA::FRT ΔdapB::FRT*	NB - This study
K008	BW25113	*ΔtolC::Kan^r*	NB - This study
K009	K001	*ΔaraC::FRT Δasd::FRT ΔtolC::Kan^r*	NB - This study
K010	BW25113	*ΔtrpE::Kan^r*	NB - This study
K011	BW25113	*ΔtrpD::Kan^r*	NB - This study
K012	D20	*ΔaraC::FRT ΔtrpE::Kan^r*	Sebastián Sosa Carrillo - This study
K013	D20	*ΔaraC::FRT ΔtrpD::Kan^r*	Sebastián Sosa Carrillo - This study
K014	D20	*ΔaraC::Kan^r ΔtrpD::Kan^r*	Sebastián Sosa Carrillo - This study
K015	D20	*ΔaraC::Kan^r ΔtrpE::FRT*	Sebastián Sosa Carrillo - This study
K016	D20	*ΔaraC::FRT ΔtrpE::FRT*	Sebastián Sosa Carrillo - This study
K017	D20	*ΔaraC::FRT ΔtrpE::FRT*	Sebastián Sosa Carrillo - This study
K018	JW0063	*ΔaraC::Kan^r*	Datsenko & Wanner, 2000
K019	K002	*ΔaraC::FRT ΔdapA::FRT ΔtolC::Kan^r*	NB - This study
K020	K004	*ΔaraC::FRT ΔtrpD::FRT ΔtolC::Kan^r*	NB - This study
K021	K005	*ΔaraC::FRT ΔtrpE::FRT ΔtolC::Kan^r*	NB - This study
K022	D18	*ΔaraC::FRT Δalr::FRT ΔdadX::FRT ΔtolC::Kan^r*	NB - This study
K023	D19	*ΔaraC::FRT ΔcysH::FRT ΔtolC::Kan^r*	NB - This study
K024	D20	*ΔaraC::FRT ΔtolC::Kan^r*	NB - This study
K025	K019	*ΔaraC::FRT ΔdapA::FRT ΔtolC::FRT*	Sebastián Sosa Carrillo - This study
K026	K021	*ΔaraC::FRT ΔtrpE::FRT ΔtolC::FRT*	Sebastián Sosa Carrillo - This study
K027	K009	*ΔaraC::FRT Δasd::FRT ΔtolC::FRT*	Sebastián Sosa Carrillo - This study
K028	K024	*ΔaraC::FRT ΔtolC::FRT*	Sebastián Sosa Carrillo - This study
K029	D20	*ΔaraC::FRT ΔentC::Kan^r*	Sebastián Sosa Carrillo - This study
K030	K028	*ΔaraC::FRT ΔtolC::FRT ΔentC::Kan^r*	Sebastián Sosa Carrillo - This study
K031	K030	*ΔaraC::FRT ΔtolC::FRT ΔentC::FRT*	Sebastián Sosa Carrillo - This study
K031	K030	*ΔaraC::FRT ΔtolC::FRT ΔentC::FRT*	Sebastián Sosa Carrillo - This study
K032	Keio Collection	*ΔtolC::Kan^r*	NB - This study
K033	Keio Collection	*ΔfolB::Kan^r*	Sebastián Sosa Carrillo - This study
K034	Keio Collection	*ΔfolP::Kan^r*	Sebastián Sosa Carrillo - This study
K035	Keio Collection	*ΔthyA::Kan^r*	Sebastián Sosa Carrillo - This study
K036	Keio Collection	*ΔentC::Kan^r*	NB - This study
K037	K031	*ΔaraC::FRT ΔtolC::FRT ΔentC::FRT ΔtrpE::Kan^r*	Sebastián Sosa Carrillo - This study
K038	K031	*ΔaraC::FRT ΔtolC::FRT ΔentC::FRT ΔtrpD::Kan^r*	Sebastián Sosa Carrillo - This study
K039	K031	*ΔaraC::FRT ΔtolC::FRT ΔentC::FRT ΔcysH::Kan^r*	Sebastián Sosa Carrillo - This study
K040	K031	*ΔaraC::FRT ΔtolC::FRT ΔentC::FRT ΔaraC::Kan^r*	Sebastián Sosa Carrillo - This study
K041	K031	*ΔaraC::FRT ΔtolC::FRT ΔentC::FRT Δalr::Kan^r*	Sebastián Sosa Carrillo - This study
K042	K031	*ΔaraC::FRT ΔtolC::FRT ΔentC::FRT ΔdadX::Kan^r*	Sebastián Sosa Carrillo - This study
K043	D20	*ΔaraC::FRT ΔfolB::Kan^r*	Sebastián Sosa Carrillo - This study
K044	K031	*ΔaraC::FRT ΔtolC::FRT ΔentC::FRT ΔfolB::Kan^r*	Sebastián Sosa Carrillo - This study
K045	K041	*ΔaraC::FRT ΔtolC::FRT ΔentC::FRT Δalr::FRT*	Sebastián Sosa Carrillo - This study
K046	K029	*ΔaraC::FRT ΔentC::FRT*	Sebastián Sosa Carrillo - This study
K047	K042	*ΔaraC::FRT ΔtolC::FRT ΔentC::FRT ΔdadX::FRT*	Sebastián Sosa Carrillo - This study
K048	K038	*ΔaraC::FRT ΔtolC::FRT ΔentC::FRT ΔtrpD::FRT*	Sebastián Sosa Carrillo - This study

Table A.1: List of deletion strains.

K049	K039	*ΔaraC::FRT ΔtolC::FRT ΔentC::FRT ΔcysH::FRT*	Sebastián Sosa Carrillo - This study
K050	K031	*ΔaraC::FRT ΔtolC::FRT ΔentC::FRT Δasd::Kan^r*	NB - This study
K051	K031	*ΔaraC::FRT ΔtolC::FRT ΔentC::FRT ΔdapA::Kan^r*	NB - This study
K052	Keio Collection	*ΔaraC::FRT ΔsdiA::Kan^r*	NB - This study
K053	K045	*ΔaraC::FRT ΔtolC::FRT ΔentC::FRT Δalr::FRT ΔdadX::Kan^r*	NB - This study
K054	JW1901-5	*ΔsdiA::Kan^r*	Datsenko & Wanner, 2000
K056	K031	*ΔaraC::FRT ΔtolC::FRT ΔentC::FRT ΔsdiA::Kan^r*	NB - This study
K057	NB.057	*ΔpE-araE::cp8-RBS-araE-Cam^r*	NB - This study
K058	NB.003	*ΔpE-araE::cp8-RBS-araE-Cam^r*	NB - This study
K063	D20	*ΔaraC::FRT ΔaraHGF::Kan^r*	NB - This study
K064	D20	*ΔaraC::FRT ΔaraEp-531::Kan^r cp8 - araE535*	NB - This study
K065	D20	*ΔaraC::FRT ΔaraEp-531::Kan^r cp18 - araE533*	NB - This study
K066	D20	*ΔaraC::FRT ΔaraEp-531::Kan^r cp13 - araE533*	NB - This study
K067	Keio strain	*ΔaraJ::Kan^r*	NB - This study
K068	K064	*ΔaraC::FRT ΔaraEp-531::FRT cp8 - araE535*	NB - This study
K069	K065	*ΔaraC::FRT ΔaraEp-531::FRT cp18 - araE533*	NB - This study
K070	K066	*ΔaraC::FRT ΔaraEp-531::FRT cp13 - araE533*	NB - This study
K071	K001	*Δasd::FRT ΔaraC771::FRT ΔaraEp-531::Kan^r cp8 - araE535*	NB - This study
K072	K001	*ΔaraC::FRT ΔaraEp-531::KanR cp18 - araE533*	NB - This study
K073	K001	*ΔaraC::FRT Δasd::FRT ΔaraEp-531::Kan^r φPcp18 - araE533*	NB - This study
K074	K002	*ΔaraC::FRT ΔdapA::FRT ΔaraEp-531::Kan^r cp8 - araE535*	NB - This study
K075	K002	*ΔdapA::FRT ΔaraC::FRT ΔaraEp-531::Kan^r cp18 - araE533*	NB - This study
K076	K002	*ΔdapA::FRT ΔaraC::FRT ΔaraEp-531::Kan^r cp13 - araE533*	NB - This study
K077	K003	*ΔdapB::FRT ΔaraC::FRT ΔaraEp-531::Kan^r cp8-araE535*	NB - This study
K078	K003	*ΔdapB::FRT ΔaraC::FRT ΔaraEp-531::Kan^r cp18 - araE533*	NB - This study
K079	K003	*ΔdapB::FRT ΔaraC771::FRT ΔaraEp-531::Kan^r cp13 - araE533*	NB - This study
K080	K004	*ΔtrpD::FRT ΔaraC771::FRT ΔaraEp-531::Kan^r cp8 - araE535*	NB - This study
K081	K004	*ΔtrpD::FRT ΔaraC::FRT ΔaraEp-531::Kan^r cp18 - araE533*	NB - This study
K082	K004	*ΔtrpD::FRT ΔaraC::FRT ΔaraEp-531::Kan^r cp13 - araE533*	NB - This study
K083	K005	*ΔtrpE::FRT ΔaraC771::FRT ΔaraEp-531:: Kan^r cp8 - araE535*	NB - This study
K084	K005	*ΔtrpE::FRT ΔaraC::FRT ΔaraEp-531:: Kan^r cp18 - araE533*	NB - This study
K085	K005	*ΔtrpE::FRT ΔaraC::FRT ΔaraEp-531::Kan^r cp13 - araE533*	NB - This study
K086	D18	*Δalr::FRT ΔdadX::FRT ΔaraC::FRT ΔaraEp-531::Kan^r cp8 - araE535*	NB - This study
K087	D18	*Δalr::FRT ΔdadX::FRT ΔaraC::FRT ΔaraEp-531::Kan^r cp18 - araE533*	NB - This study
K088	D18	*Δalr::FRT ΔdadX::FRT ΔaraC::FRT ΔaraEp-531::Kan^r cp13 - araE533*	NB - This study
K089	D19	*ΔcysH::FRT ΔaraC::FRT ΔaraEp-531::Kan^r cp8 - araE535*	NB - This study
K090	D19	*ΔcysH::FRT ΔaraC::FRT ΔaraEp-531::Kan^r cp18 - araE533*	NB - This study
K091	D19	*ΔcysH::FRT ΔaraC::FRT ΔaraEp-531::Kan^r cp13 - araE533*	NB - This study
K092	K068	*ΔaraC771::FRT ΔaraFGH::Kan^r ΔaraEp-531::FRT cp8 - araE535*	NB - This study
K093	K069	*ΔaraC::FRT ΔaraFGH::Kan^r ΔaraEp-531::FRT cp18 - araE533*	NB - This study
K094	K070	*ΔaraC::FRT ΔaraFGH:: ΔaraEp-531::FRT cp13 - araE533*	NB - This study

Table A.1: List of deletion strains.

K095	K071	*ΔaraC::FRT ΔasdA::FRT ΔaraEp-531::FRT - cp8 - araE535*	NB - This study
K096	K072	*ΔaraC::FRT ΔasdA::FRT ΔaraEp-531::FRT - cp18 - araE533*	NB - This study
K097	K073	*ΔaraC::FRT ΔasdA::FRT ΔaraEp-531::FRT - cp13 - araE533*	NB - This study
K098	K074	*ΔaraC::FRT ΔdapA::FRT ΔaraEp-531::FRT - cp8 - araE535*	NB - This study
K099	K075	*ΔaraC::FRT ΔdapA::FRT ΔaraEp-531::FRT - cp18 - araE533*	NB - This study
K100	K076	*ΔaraC::FRT ΔdapA::FRT ΔaraEp-531::FRT - cp13 - araE533*	NB - This study
K101	K077	*ΔaraC::FRT ΔdapB::FRT ΔaraEp-531::FRT - cp8 - araE535*	NB - This study
K102	K078	*ΔaraC::FRT ΔdapB::FRT ΔaraEp-531::FRT - cp18 - araE533*	NB - This study
K103	K079	*ΔaraC::FRT ΔdapB::FRT ΔaraEp-531::FRT - cp13 - araE533*	NB - This study
K104	K080	*ΔaraC::FRT ΔtrpD::FRT ΔaraEp-531::FRT - cp8 - araE535*	NB - This study
K105	K081	*ΔaraC::FRT ΔtrpD::FRT ΔaraEp-531::FRT - cp18 - araE533*	NB - This study
K106	K082	*ΔaraC::FRT ΔtrpD::FRT ΔaraEp-531::FRT - cp13 - araE533*	NB - This study
K107	K083	*ΔaraC::FRT ΔtrpE::FRT ΔaraEp-531::FRT - cp8 - araE535*	NB - This study
K108	K084	*ΔaraC::FRT ΔtrpE::FRT ΔaraEp-531::FRT - cp18 - araE533*	NB - This study
K109	K085	*ΔtrpE::FRT ΔaraC::FRT ΔaraEp-531:: Kan^r cp13 - araE533*	NB - This study
K110	K086	*ΔdadX::FRT Δalr::FRT ΔaraC::FRT ΔaraEp-531::Kan^r cp8 - araE535*	NB - This study
K111	K087	*Δalr::FRT ΔdadX::FRT ΔaraC::FRT ΔaraEp-531::Kan^r cp18 - araE533*	NB - This study
K112	K088	*Δalr::FRT ΔdadX::FRT ΔaraC::FRT ΔaraEp-531::Kan^r cp13 - araE533*	NB - This study
K113	K089	*ΔcysH::FRT ΔaraC::FRT ΔaraEp-531::cp8 - araE535*	NB - This study
K114	K090	*ΔcysH::FRT ΔaraC::FRT ΔaraEp-531::cp18 - araE533*	NB - This study
K115	K091	*ΔcysH::FRT ΔaraC::FRT ΔaraEp-531::cp13 - araE533*	NB - This study
K116	K092	*ΔaraC::FRT ΔaraEp-531::FRT cp8 - araE535 ΔaraFGH::FRT*	NB - This study
K117	K093	*ΔaraC::FRT ΔaraEp-531::FRT cp18 - araE533 ΔaraFGH::FRT*	NB - This study
K118	K094	*ΔaraC::FRT ΔaraEp-531::FRT cp13 - araE533 ΔaraFGH::FRT*	NB - This study
K119	K050	*ΔaraC::FRT ΔtolC::FRT ΔentC::FRT ΔasdA::FRT*	NB - This study
K120	K051	*ΔaraC::FRT ΔtolC::FRT ΔentC::FRT ΔdapA::FRT*	NB - This study
K121	K053	*ΔaraC::FRT ΔtolC::FRT ΔentC::FRT Δalr::FRT ΔdadX::FRT*	NB - This study
K122	K031	*ΔaraC::FRT ΔtolC::FRT ΔentC::FRT ΔdapB::Kan^r*	NB - This study
K123	K029	*ΔaraC::FRT ΔentC::FRT*	NB - This study
K124	K122	*ΔaraC::FRT ΔtolC::FRT ΔentC::FRT ΔdapB::FRT*	NB - This study

Table A.1: List of deletion strains.

Strain ID	Background strain	Description of genotype	Plasmid(s)	Short plasmid description	ATB resistance marker	Constructor
TS000	D20	*ΔaraC::FRT*	P37.pRD131b V91	*pBAD-mCherry* *pBAD-araC*	AmpR KanR	NB-This study
TS001	D20	*ΔaraC::FRT*	P40.pRD215a P39.pRD152	*pLux-mCherry* *pLux-araC-gfp*	AmpR KanR	NB-This study
TS002	K001	*ΔaraC::FRT* *Δasd::FRT*	pTS002 V91	*pBAD-Mtb-asd* *pBAD-araC*	AmpR KanR	NB-This study
TS003	K001	*ΔaraC::FRT* *Δasd::FRT*	pTS003 V91	*pBAD-E. coli-asd* *pBAD-araC*	AmpR KanR	NB-This study
TS004	K002	*ΔaraC::FRT* *ΔdapA::FRT*	pTS004 V91	*pBAD-Mtb-dapA* *pBAD-araC*	AmpR KanR	NB-This study
TS005	K002	*ΔaraC::FRT* *ΔdapA::FRT*	pTS005 V91	*pBAD-E. coli-dapA* *pBAD-araC*	AmpR KanR	NB-This study
TS006	K003	*ΔaraC::FRT* *ΔdapB::FRT*	pTS006 V91	*pBAD-Mtb-dapB* *pBAD-araC*	AmpR KanR	NB-This study
TS007	K003	*ΔaraC::FRT* *ΔdapB::FRT*	pTS007 V91	*pBAD-E. coli-dapB* *pBAD-araC*	AmpR KanR	NB-This study
TS008	D18	*ΔaraC::FRT* *Δalr::FRT* *ΔdadX::FRT*	pTS008 V91	*pBAD-Mtb-alr* *pBAD-araC*	AmpR KanR	NB-This study
TS009	D18	*ΔaraC::FRT* *Δalr::FRT* *ΔdadX::FRT*	pTS009 V91	*pBAD-E. coli-dadX* *pBAD-araC*	AmpR KanR	NB-This study
TS010	D19	*ΔaraC::FRT* *ΔcysH::FRT*	pTS010 V91	*pBAD-Mtb-cysH* *pBAD-araC*	AmpR KanR	NB-This study
TS011	D19	*ΔaraC::FRT* *ΔcysH::FRT*	pTS011 V91	*pBAD-E. coli-cysH* *pBAD-araC*	AmpR KanR	NB-This study
TS012	D18	*ΔaraC::FRT* *Δalr::FRT* *ΔdadX::FRT*	pTS008 pD.048	*pBAD-Mtb-alr* *pCons-araC*	AmpR KanR	NB-This study
TS013	D18	*ΔaraC::FRT* *Δalr::FRT* *ΔdadX::FRT*	pTS009 pD.048	*pBAD-E.coli-dadX* *pCons-araC*	AmpR KanR	NB-This study
TS014	D19	*ΔaraC::FRT* *ΔcysH::FRT*	pTS010 pD.048	*pBAD-Mtb-cysH* *pCons-araC*	AmpR KanR	NB-This study
TS015	D19	*ΔaraC::FRT* *ΔcysH::FRT*	pTS011 pD.048	*pBAD-E.coli-cysH* *pCons-araC*	AmpR KanR	NB-This study
TS016	K001	*ΔaraC::FRT* *Δasd::FRT*	pTS022 P39.pRD152	*pLux-Mtb-asd* *pLux-LuxR-gfp*	AmpR KanR	NB-This study
TS017	K001	*ΔaraC::FRT* *Δasd::FRT*	pTS023 P39.pRD152	*pLux-E.coli-asd* *pLux-LuxR-gfp*	AmpR KanR	NB-This study
TS018	D20	*ΔaraC::FRT*	V91 P37.pRD131b	*pBAD-araC* *pBAD-mCherry*	KanR AmpR	NB-This study
TS019	K001	*ΔaraC::FRT* *Δasd::FRT*	V91 P37.pRD131b	*pBAD-araC* *pBAD-mCherry*	AmpR KanR	NB-This study
TS020	K002	*ΔaraC::FRT* *ΔdapA::FRT*	V91 P37.pRD131b	*pBAD-araC* *pBAD-mCherry*	KanR AmpR	NB-This study
TS021	K003	*ΔaraC::FRT* *ΔdapB::FRT*	V91 P37.pRD131b	*pBAD-araC* *pBAD-mCherry*	KanR AmpR	NB-This study
TS022	K004	*ΔaraC::FRT* *ΔtrpD::FRT*	V91 pTS017	*pBAD-araC* *pBAD-Mtb-trpD*	KanR AmpR	NB-This study
TS023	K004	*ΔaraC::FRT* *ΔtrpD::FRT*	V91 pTS016	*pBAD-araC* *pBAD-E.coli-trpD*	KanR AmpR	NB-This study
TS024	K004	*ΔaraC::FRT* *ΔtrpD::FRT*	V91 P37.pRD131b	*pBAD-araC* *pBAD-mCherry*	KanR AmpR	NB-This study
TS025	K005	*ΔaraC::FRT* *ΔtrpE::FRT*	V91 pTS015	*pBAD-araC* *pBAD-Mtb-trpE*	KanR AmpR	NB-This study
TS026	K005	*ΔaraC::FRT* *ΔtrpE::FRT*	V91 pTS014	*pBAD-araC* *pBAD-E.coli-trpE*	KanR AmpR	NB-This study
TS027	K005	*ΔaraC::FRT* *ΔtrpE::FRT*	V91 P37.pRD131b	*pBAD-araC* *pBAD-mCherry*	KanR AmpR	NB-This study
TS028	D18	*ΔaraC::FRT* *Δalr::FRT* *ΔdadX::FRT*	V91 P37.pRD131b	*pBAD-araC* *pBAD-mCherry*	KanR AmpR	NB-This study
TS029	D19	*ΔaraC::FRT*	V91	*pBAD-araC*	KanR	NB-This study

Table A.2: List of TESEC strains.

		ΔcysH::FRT	P37.pRD131b	*pBAD-mCherry*	AmpR	
TS030	D20	*ΔaraC::FRT*	pD.048 P37.pRD131b	*pCons-araC* *pBAD-mCherry*	KanR AmpR	NB-This study
TS031	K001	*ΔaraC::FRT* *Δasd::FRT*	pD.048 PTS002	*pCons-araC* *pBAD-Mtb-asd*	KanR AmpR	NB-This study
TS032	K001	*ΔaraC::FRT* *Δasd::FRT*	pD.048 PTS003	*pCons-araC* *pBAD-E.coli-asd*	KanR AmpR	NB-This study
TS033	K001	*ΔaraC::FRT* *Δasd::FRT*	pD.048 P37.pRD131b	*pCons-araC* *pBAD-mCherry*	KanR AmpR	NB-This study
TS034	K002	*ΔaraC::FRT* *ΔdapA::FRT*	pD.048 PTS004	*pCons-araC* *pBAD-Mtb-dapA*	KanR AmpR	NB-This study
TS035	K002	*ΔaraC::FRT* *ΔdapA::FRT*	pD.048 PTS005	*pCons-araC* *pBAD-E.coli-dapA*	KanR AmpR	NB-This study
TS036	K002	*ΔaraC::FRT* *ΔdapA::FRT*	pD.048 P37.pRD131b	*pCons-araC* *pBAD-mCherry*	KanR AmpR	NB-This study
TS037	K003	*ΔaraC::FRT* *ΔdapB::FRT*	pD.048 PTS006	*pCons-araC* *pBAD-Mtb-dapB*	KanR AmpR	NB-This study
TS038	K003	*ΔaraC::FRT* *ΔdapB::FRT*	pD.048 PTS007	*pCons-araC* *pBAD-E.coli-dapB*	KanR AmpR	NB-This study
TS039	K003	*ΔaraC::FRT* *ΔdapB::FRT*	pD.048 P37.pRD131b	*pCons-araC* *pBAD-mCherry*	KanR AmpR	NB-This study
TS040	K004	*ΔaraC::FRT* *ΔtrpD::FRT*	pD.048 pTS017	*pCons-araC* *pBAD-Mtb-trpD*	KanR AmpR	NB-This study
TS041	K004	*ΔaraC::FRT* *ΔtrpD::FRT*	pD.048 pTS016	*pCons-araC* *pBAD-E.coli-trpD*	KanR AmpR	NB-This study
TS042	K004	*ΔaraC::FRT* *ΔtrpD::FRT*	pD.048 P37.pRD131b	*pCons-araC* *pBAD-mCherry*	KanR AmpR	NB-This study
TS043	K005	*ΔaraC::FRT* *ΔtrpE::FRT*	pD.048 pTS015	*pCons-araC* *pBAD-Mtb-trpE*	KanR AmpR	NB-This study
TS044	K005	*ΔaraC::FRT* *ΔtrpE::FRT*	pD.048 pTS014	*pCons-araC* *pBAD-E.coli-trpE*	KanR AmpR	NB-This study
TS045	K005	*ΔaraC::FRT* *ΔtrpE::FRT*	pD.048 P37.pRD131b	*pCons-araC* *pBAD-mCherry*	KanR AmpR	NB-This study
TS046	D18	*ΔaraC::FRT* *Δalr::FRT* *ΔdadX::FRT*	pD.048 pTS008	*pCons-araC* *pBAD-Mtb-alr*	KanR AmpR	NB-This study
TS047	D18	*ΔaraC::FRT* *Δalr::FRT* *ΔdadX::FRT*	pD.048 pTS009	*pCons-araC* *pBAD-E.coli-dadX*	KanR AmpR	NB-This study
TS048	D18	*ΔaraC::FRT* *Δalr::FRT* *ΔdadX::FRT*	pD.048 P37.pRD131b	*pCons-araC* *pBAD-mCherry*	KanR AmpR	NB-This study
TS049	D19	*ΔaraC::FRT* *ΔcysH::FRT*	pD.048 pTS010	*pCons-araC* *pBAD-Mtb-cysH*	KanR AmpR	NB-This study
TS050	D19	*ΔaraC::FRT* *ΔcysH::FRT*	pD.048 pTS011	*pCons-araC* *pBAD-E.coli-cysH*	AmpR KanR	NB-This study
TS051	D19	*ΔaraC::FRT* *ΔcysH::FRT*	pD.048 P37.pRD131b	*pCons-araC* *pBAD-mCherry*	AmpR KanR	NB-This study
TS052	D20	*ΔaraC::FRT*	P37.pRD131b	*pBAD-mCherry*	AmpR	NB-This study
TS053	K001	*ΔaraC::FRT* *Δasd::FRT*	pTS002	*pBAD-Mtb-asd*	AmpR	NB-This study
TS054	K001	*ΔaraC::FRT* *Δasd::FRT*	pTS003	*pBAD-E.coli-asd*	AmpR	NB-This study
TS055	K002	*ΔaraC::FRT* *ΔdapA::FRT*	pTS004	*pBAD-Mtb-dapA*	AmpR	NB-This study
TS056	K002	*ΔaraC::FRT* *ΔdapA::FRT*	pTS005	*pBAD-E.coli-dapA*	AmpR	NB-This study
TS057	K003	*ΔaraC::FRT* *ΔdapB::FRT*	pTS006	*pBAD-Mtb-dapB*	AmpR	NB-This study
TS058	K003	*ΔaraC::FRT* *ΔdapB::FRT*	pTS007	*pBAD-E.coli-dapB*	AmpR	NB-This study
TS059	K004	*ΔaraC::FRT* *ΔtrpD::FRT*	pTS017	*pBAD-Mtb-trpD*	AmpR	NB-This study
TS060	K004	*ΔaraC::FRT* *ΔtrpD::FRT*	pTS016	*pBAD-E.coli-trpD*	AmpR	NB-This study
		ΔaraC::FRT		*pBAD-Mtb-trpE*		

Table A.2: List of TESEC strains.

TS061	K005	*ΔtrpE::FRT*	pTS015		AmpR	NB-This study
TS062	K005	*ΔaraC::FRT* *ΔtrpE::FRT*	pTS014	*pBAD-E.coli-trpE*	AmpR	NB-This study
TS063	D18	*ΔaraC::FRT* *Δalr::FRT* *ΔdadX::FRT*	pTS008	*pBAD-Mtb-alr*	AmpR	NB-This study
TS064	D18	*ΔaraC::FRT* *Δalr::FRT* *ΔdadX::FRT*	pTS009	*pBAD-E.coli-dadX*	AmpR	NB-This study
TS065	D19	*ΔaraC::FRT* *ΔcysH::FRT*	pTS010	*pBAD-Mtb-cysH*	AmpR	NB-This study
TS066	D19	*ΔaraC::FRT* *ΔcysH::FRT*	pTS011	*pBAD-E.coli-cysH*	AmpR	NB-This study
TS071	D20	*ΔaraC::FRT*	pTS000 V91	*pBAD-lacZ* *pBAD-araC*	AmpR KanR	Salam Abbara - This study
TS072	D18	*ΔaraC::FRT* *Δalr::FRT* *ΔdadX::FRT*	V91	*pBAD-araC*	KanR	Salam Abbara - This study
TS073	K068	*ΔaraC::FRT* *ΔaraEp-531::cp8-araE535*	V91 P37.pRD131b	*pBAD-araC* *pBAD-mCherry*	AmpR KanR	NB-This study
TS074	K069	*ΔaraC::FRT* *ΔaraEp-531::cp18-araE533*	V91 P37.pRD131b	*pBAD-araC* *pBAD-mCherry*	AmpR KanR	NB-This study
TS075	K070	*ΔaraC::FRT* *ΔaraEp-531::cp13-araE533*	V91 P37.pRD131b	*pBAD-araC* *pBAD-mCherry*	AmpR KanR	NB-This study
TS076	K068	*ΔaraC::FRT* *ΔaraEp-531::cp8-araE535*	P36.pRD123 P37.pRD131b	*pBAD-araC-gfp* *pBAD-mCherry*	AmpR KanR	NB-This study
TS077	K069	*ΔaraC::FRT* *ΔaraEp-531::cp18-araE533*	P36.pRD123 P37.pRD131b	*pBAD-araC-gfp* *pBAD-mCherry*	AmpR KanR	NB-This study
TS078	K070	*ΔaraC::FRT* *ΔaraEp-531::cp13-araE533*	P36.pRD123 P37.pRD131b	*pBAD-araC-gfp* *pBAD-mCherry*	AmpR KanR	NB-This study
TS079	K068	*ΔaraC::FRT* *ΔaraEp-531::cp8-araE535*	pD.048 P37.pRD131b	*pCons-araC* *pBAD-mCherry*	KanR AmpR	NB-This study
TS080	K069	*ΔaraC::FRT* *ΔaraEp-531::cp18-araE533*	pD.048 P37.pRD131b	*pCons-araC* *pBAD-mCherry*	KanR AmpR	NB-This study
TS081	K070	*ΔaraC::FRT* *ΔaraEp-531::cp13-araE533*	pD.048 P37.pRD131b	*pCons-araC* *pBAD-mCherry*	KanR AmpR	NB-This study
TS082	K096	*Δasd::FRT* *ΔaraC::FRT* *ΔaraEp-531::FRT-cp18-araE533*	pTS002 V91	*pBAD-Mtb-asd* *pBAD-araC*	AmpR KanR	NB-This study
TS083	K096	*Δasd::FRT* *ΔaraC::FRT* *ΔaraEp-531::FRT-cp18-araE533*	pTS003 V91	*pBAD-E.coli-asd* *pBAD-araC*	AmpR KanR	NB-This study
TS084	K099	*ΔdapA::FRT* *ΔaraC::FRT* *ΔaraEp-531::FRT-cp18-araE533*	pTS004 V91	*pBAD-Mtb-dapA* *pBAD-araC*	AmpR KanR	NB-This study
TS085	K099	*ΔdapA::FRT* *ΔaraC::FRT* *ΔaraEp-531::FRT-cp18-araE533*	pTS005 V91	*pBAD-E.coli-dapA* *pBAD-araC*	AmpR KanR	NB-This study
TS088	K105	*ΔtrpD::FRT* *ΔaraC::FRT* *ΔaraEp-531::FRT-cp18-araE 533*	V91 pTS017	*pBAD-araC* *pBAD-Mtb-trpD*	AmpR KanR	NB-This study
TS089	K105	*ΔtrpD::FRT* *ΔaraC771::FRT* *ΔaraEp-531::FRT-cp18-araE 533*	V91 pTS016	*pBAD-araC* *pBAD-E.coli-trpD*	AmpR KanR	NB-This study
TS090	K108	*ΔtrpE::FRTΔaraC771::FRTΔaraEp-531::FRT-cp18-araE 533*	V91 pTS015	*pBAD-araC* *pBAD-Mtb-trpE*	AmpR KanR	NB-This study
TS091	K108	*ΔtrpE::FRT* *ΔaraC771::FRT* *ΔaraEp-531::FRT-cp18-araE 533*	V91 pTS014	*pBAD-araC* *pBAD-E.coli-trpE*	AmpR KanR	NB-This study
TS092	K111	*Δalr::FRT* *ΔdadX::FRT* *ΔaraC::FRT* *ΔaraEp-531::KanR cp18-araE 533*	pTS008 V91	*pBAD-Mtb-alr* *pBAD-araC*	AmpR KanR	NB-This study
TS093	K111	*Δalr::FRT ΔdadX::FRT* *ΔaraC771::FRT* *ΔaraEp-531::KanR cp18-araE 533*	pTS009 V91	*pBAD-E.coli-dadX* *pBAD-araC*	AmpR KanR	NB-This study

Table A.2: List of TESEC strains.

TS094	K114	*ΔcysH::FRT* *ΔaraC771::FRT* *ΔaraEp-531::KanR cp18-araE 533*	pTS010 V91	*pBAD-Mtb-cysH* *pBAD-araC*	AmpR KanR	NB-This study
TS095	K114	*ΔcysH::FRT* *ΔaraC::FRT* *ΔaraEp-531::KanR cp18-araE 533*	pTS011 V91	*pBAD-E.coli-cysH* *pBAD-araC*	AmpR KanR	NB-This study
TS100	K102	*ΔdapB::FRT* *ΔaraC::FRT* *ΔaraEp-531::FRT-cp18-araE 533*	pTS006	*pBAD-Mtb-dapB*	AmpR	NB-This study
TS101	K102	*ΔdapB::FRT* *ΔaraC::FRT* *ΔaraEp-531::FRT-cp18-araE 533*	pTS007	*pBAD-E.coli-dapB*	AmpR	NB-This study
TS110	TS072	*ΔaraC::FRT* *Δalr::FRT* *ΔdadX::FRT*	V91 pTS009	*pBAD-araC* *pBAD-E. coli -dadX*	KanR AmpR	NB-This study
TS111	K099	*ΔdapA::FRT* *ΔaraC::FRT* *ΔaraEp-531::FRT-cp18-araE 533*	V91 P37.pRD131b	*pBAD-araC* *pBAD-mCherry*	AmpR KanR	NB-This study
TS112	TS166	*ΔaraC::FRT* *ΔtolC::FRT* *ΔentC::FRT* *Δalr::FRT* *ΔdadX::FRT*	V91 pTS009	*pBAD-araC* *pBAD-E. coli -dadX*	KanR AmpR	NB-This study
TS113	K105	*ΔtrpD::FRT* *ΔaraC::FRT* *ΔaraEp-531::FRT-cp18-araE 533*	V91 P37.pRD131b	*pBAD-araC* *pBAD-mCherry*	AmpR KanR	NB-This study
TS114	K108	*ΔtrpE::FRT* *ΔaraC::FRT* *ΔaraEp-531::FRT-cp18-araE 533*	V91 P37.pRD131b	*pBAD-araC* *pBAD-mCherry*	AmpR KanR	NB-This study
TS115	K111	*Δalr::FRT ΔdadX::FRT* *ΔaraC::FRT* *ΔaraEp-531::KanR cp18-araE 533*	V91 P37.pRD131b	*pBAD-araC* *pBAD-mCherry*	AmpR KanR	NB-This study
TS116	K114	*ΔcysH::FRT* *ΔaraC::FRT* *ΔaraEp-531::KanR cp18-araE 533*	V91 P37.pRD131b	*pBAD-araC* *pBAD-mCherry*	AmpR KanR	NB-This study
TS117	K117	*ΔaraC::FRT* *ΔaraEp-531::FRT cp18-araE 533ΔaraFGH::FRT*	P37.pRD131b	*pBAD-mCherry*	AmpR	NB-This study
TS118	K117	*ΔaraC::FRT* *ΔaraEp-531::FRT cp18-araE 533 ΔaraFGH::FRT*	V91 P37.pRD131b	*pBAD-araC* *pBAD-mCherry*	AmpR KanR	NB-This study
TS119	K117	*ΔaraC::FRT* *ΔaraEp-531::FRT cp18-araE 533 ΔaraFGH::FRT*	pD.048 P37.pRD131b	*pCons-araC* *pBAD-mCherry*	AmpR KanR	NB-This study
TS120	K117	*ΔaraC::FRT* *ΔaraEp-531::FRT cp18-araE 533 ΔaraFGH::FRT*	V91 pTS000	*pBAD-araC* *pBAD-lacZ*	AmpR KanR	NB-This study
TS121	TS166	*ΔaraC::FRT* *ΔtolC::FRT* *ΔentC::FRT* *Δalr::FRT* *ΔdadX::FRT*	V91 pTS059	*pBAD-araC* *pBAD-meGFP-Mtb -alr*	AmpR KanR	NB-This study
TS122	K117	*ΔaraC::FRT* *ΔaraEp-531::FRT cp18-araE 533* *ΔaraFGH::FRT*	pTS000	*pBAD-lacZ*	AmpR	NB-This study
TS123	TS166	*ΔaraC::FRT* *ΔtolC::FRT* *ΔentC::FRT* *Δalr::FRT* *ΔdadX::FRT*	V91 pTS068	*pBAD-araC* *pBAD-alr*$_{M.\ smegmatis}$	KanR AmpR	NB-This study
TS124	TS166	*ΔaraC::FRT* *ΔtolC::FRT* *ΔentC::FRT* *Δalr::FRT* *ΔdadX::FRT*	V91 pMUT-alr-006	*pBAD-araC* *pBAD-Mtb-alr-Y388D*	KanR AmpR	NB-This study
TS125	K105	Δ*araC*::FRT Δ*trpD*::FRT Δ*araEp*-531::FRT - cp18 - *araE*533	pTS017	*pBAD-Mtb-trpD*	AmpR	NB-This study
TS128	TS072	*ΔaraC::FRT* *Δalr::FRT* *ΔdadX::FRT*	V91 pTS072	*pBAD-araC* *pBAD-Mtb-short-alr*	KanR AmpR	NB-This study

Table A.2: List of TESEC strains.

TS127	K105	*ΔaraC::FRT* *ΔtrpD::FRT* *ΔaraEp-531::FRT - cp18 - araE533*	pTS016	*pBAD-E. coli-trpD*		NB-This study
TS130	K114	*ΔcysH::FRT* *ΔaraC::FRT* *ΔaraEp-531::cp18 - araE533*	pTS010	*pBAD-Mtb-cysH*		NB-This study
TS131	K114	*ΔcysH::FRT* *ΔaraC::FRT* *ΔaraEp-531::cp18 - araE533*	pD.048 pTS011	*pCons-araC* *pBAD-E. coli-cysH*	KanR AmpR	NB-This study
TS132	K114	*ΔcysH::FRT* *ΔaraC::FRT* *ΔaraEp-531::cp18 - araE533*	pTS011	*pBAD-E. coli-cysH*	AmpR	NB-This study
TS133	D20	*ΔaraC::FRT*	pD.048 pTS000	*pCons-araC* *pBAD-lacZ*	KanR AmpR	NB-This study
TS135	D18	*ΔdadX::FRT* *Δalr::FRT* *ΔaraC::FRT*	pD.048	*pCons-araC*	KanR	NB-This study
TS136	M29	*ΔaraC::KanR*	pTS042	*araC-pBAD-RBS1 (wt)-mE-GFP in DVA_colE1_AF*	AmpR	NB-This study
TS137	M29	*ΔaraC::KanR*	pTS043	*araC-pBAD-RBS1 (wt)-mE-GFP in DVA_p15A_AF*	AmpR	NB-This study
TS139	M29	*ΔaraC::KanR*	pTS045	*araC-pBAD-RBS1 (wt)-mCherry in DVA_colE1_AF*	AmpR	NB-This study
TS140	M29	*ΔaraC::KanR*	pTS046	*araC-pBAD-RBS1 (wt)-mCherry in DVA_p15A_AF*	AmpR	NB-This study
TS142	M29	*ΔaraC::KanR*	pTS048	*araC-pBAD-RBS1 (O2>I1)-mE-GFP in DVA_colE1_AF*	AmpR	NB-This study
TS143	M29	*ΔaraC::KanR*	pTS049	*araC-pBAD-RBS1 (O2>I1)-mE-GFP in DVA_p15A_AF*	AmpR	NB-This study
TS144	K121	*ΔaraC::FRT* *ΔtolC::FRT* *ΔentC::FRT* *Δalr::FRT* *ΔdadX::FRT*	pTS008V91	*pBAD-Mtb-alr* *pBAD-araC*	AmpR KanR	NB-This study
TS145	M29	*ΔaraC::KanR*	pTS051	*araC-pBAD-RBS1 (O2>I1)-mCherry in DVA_colE1_AF*	AmpR	NB-This study
TS146	M29	*ΔaraC::KanR*	pTS052	*araC-pBAD-RBS1 (O2>I1)-mCherry in DVA_p15A_AF*	AmpR	NB-This study
TS147	K119	*ΔaraC::FRT* *ΔtolC::FRT* *ΔentC::FRT* *Δasd::FRT*	pTS060 V91	*pBAD-mEGFP-Mtb-asd* *pBAD-araC*	AmpR KanR	NB-This study
TS148	K001	*ΔaraC::FRT* *Δasd::FRT*	pTS060 V91	*pBAD-mEGFP-Mtb-asd* *pBAD-araC*	AmpR KanR	NB-This study
TS149	K119	*ΔaraC::FRT* *ΔtolC::FRT* *ΔentC::FRT* *Δasd::FRT*	pTS002 V91	*pBAD-Mtb-asd* *pBAD-araC*	AmpR KanR	NB-This study
TS150	K120	*ΔaraC::FRT* *ΔtolC::FRT* *ΔentC::FRT* *ΔdapA::FRT*	pTS062 V91	*pBAD-mEGFP-Mtb-dapA* *pBAD-araC*	AmpR KanR	NB-This study
TS151	K002	*ΔaraC::FRT* *ΔdapA::FRT*	pTS062 V91	*pBAD-mEGFP-Mtb-dapA* *pBAD-araC*	AmpR KanR	NB-This study
TS152	K120	*ΔaraC::FRT* *ΔtolC::FRT* *ΔentC::FRT* *ΔdapA::FRT*	pTS004 V91	*pBAD-Mtb-dapA* *pBAD-araC*	AmpR KanR	NB-This study
TS153	K003	*ΔaraC::FRT* *ΔdapB::FRT*	pTS063 V91	*pBAD-mEGFP-Mtb-dapB* *pBAD-araC*	AmpR KanR	NB-This study
TS154	K048	*ΔaraC::FRT* *ΔtolC::FRT*	pTS064 V91	*pBAD-mEGFP-Mtb-trpD* *pBAD-araC*	AmpR KanR	NB-This study

Table A.2: List of TESEC strains.

		ΔentC::FRT *ΔtrpD::FRT*				
TS155	K004	*ΔaraC::FRT* *ΔtrpD::FRT*	pTS064 V91	*pBAD-mEGFP-Mtb-trpD* *pBAD-araC*	AmpR KanR	NB-This study
TS156	K048	*ΔaraC::FRT* *ΔtolC::FRT* *ΔentC::FRT* *ΔtrpD::FRT*	pTS017 V91	*pBAD-Mtb-trpD* *pBAD-araC*	AmpR KanR	NB-This study
TS157	K005	*ΔaraC::FRT* *ΔtolC::FRT* *ΔentC::FRT* *ΔtrpE::FRT*	pTS065 V91	*pBAD-Mtb-trpE* *pBAD-araC*	AmpR KanR	NB-This study
TS158	TS072	*ΔaraC::FRT* *Δalr::FRT* *ΔdadX::FRT*	pTS059 V91	*pBAD-meGFP-Mtb-alr* *pBAD-araC*	AmpR KanR	NB-This study
TS159	K049	*ΔaraC::FRT* *ΔtolC::FRT* *ΔentC::FRT* *ΔcysH::FRT*	pTS010 V91	*pBAD-Mtb-cysH* *pBAD-araC*	AmpR KanR	NB-This study
TS160	D20	*ΔaraC::FRT*	pTS066 V91	*pBAD-meGFP* *pBAD-araC*	AmpR KanR	NB-This study
TS161	K028	*ΔaraC::FRT* *ΔtolC::FRT* *ΔentC::KanR*	pT067	*araC-pBAD-RBS1 (wt)-mEGFP in DVA_colE1_AF*	AmpR	NB-This study
TS162	M29	*ΔaraC::KanR*	pTS067	*araC-pBAD-RBS1 (wt)-mEGFP in DVA_colE1_AF*	AmpR	NB-This study
TS163	K031	*ΔaraC::FRT* *ΔtolC::FRT* *ΔentC::FRT*	pTS066 V91	*pBAD-meGFP* *pBAD-araC*	AmpR KanR	NB-This study
TS164	K031	*ΔaraC::FRT* *ΔtolC::FRT* *ΔentC::FRT*	pTS000 V91	*pBAD-lacZ* *pBAD-araC*	AmpR KanR	NB-This study
TS166	K121	*ΔaraC::FRT* *ΔtolC::FRT* *ΔentC::FRT* *Δalr::FRT* *ΔdadX::FRT*	V91	*pBAD-araC*	KanR	NB-This study
TS167	TS072	*ΔaraC::FRT* *Δalr::FRT* *ΔdadX::FRT*	V91 pTS072	*pBAD-araC* *pBAD-His6-Mtb-short-alr*	AmpR KanR	NB-This study
TS168	K124	*ΔaraC::FRT* *ΔtolC::FRT* *ΔentC::FRT* *ΔdapB::FRT*	V91 pTS006	*pBAD-araC* *pBAD-dapB*	AmpR KanR	NB-This study

Table A.2: List of TESEC strains.

Plasmid ID	Description	Antibiotic Resistence Marker	Reference
pKD13		AmpR	Datsenko & Wanner, 2000
pCP20		AmpR	Datsenko & Wanner, 2000
V91	*pBAD-araC*	KanR	EW
pD.048	*pCons-araC*	KanR	EW
P40.pRD215	*pLux-mCherry*	AmpR	Daniel *et al.* 2013
P37.pRD131	*pBAD-mCherry*	AmpR	Daniel *et al.* 2013
pTS000	derived, *lacZ*::*pBAD* upstream *lacZ*	AmpR	NB
pTS001	derived, *lacZ*::*pLux* upstream *lacZ*	AmpR	NB
pTS002	pTS000 derived, *lacZ*::*Mtb-asd*	AmpR	NB
pTS003	pTS000 derived, *lacZ*::*E. coli-asd*	AmpR	NB
pTS004	pTS000 derived, *lacZ*::*Mtb-dapA*	AmpR	NB
pTS005	pTS000 derived, *lacZ*::*E. coli-dapA*	AmpR	NB
pTS006	pTS000 derived, *lacZ*::*Mtb-dapB*	AmpR	NB
pTS007	pTS000 derived, *lacZ*::*E. coli-dapB*	AmpR	NB
pTS008	pTS000 derived, *lacZ*::*Mtb-alr*	AmpR	NB
pTS009	pTS000 derived, *lacZ*::*E. coli-dadX*	AmpR	NB
pTS010	pTS000 derived, *lacZ*::*Mtb-cysH*	AmpR	NB
pTS011	pTS000 derived, *lacZ*::*E.coli-cysH*	AmpR	NB
pTS012	pTS000 derived, *lacZ*::*Mtb-dfrA*	AmpR	NB
pTS013	pTS000 derived, *lacZ*::*E. coli-folA*	AmpR	NB
pTS014	pTS000 derived, *lacZ*::*E. coli-trpE*	AmpR	NB
pTS015	pTS000 derived, *lacZ*::*Mtb-trpE*	AmpR	NB
pTS016	pTS000 derived, *lacZ*::*E. coli-trpD*	AmpR	NB
pTS017	pTS000 derived, *lacZ*::*Mtb-trpD*	AmpR	NB
pTS018	pTS000 derived, *lacZ*::*E. coli-folB*	AmpR	NB
pTS019	pTS000 derived, *lacZ*::*Mtb-folB*	AmpR	NB
pTS020	pTS000 derived, *lacZ*::*E. coli-folP*	AmpR	NB
pTS021	pTS000 derived, *lacZ*::*Mtb-folP1*	AmpR	NB
pTS022	pTS001 derived, *lacZ*::*Mtb-asd*	AmpR	NB
pTS023	pTS001 derived, *lacZ*::*E. coli-asd*	AmpR	NB
pTS024	pTS001 derived, *lacZ*::*Mtb-alr*	AmpR	NB
pTS025	pTS001 derived, *lacZ*::*E. coli-dadX*	AmpR	NB
pTS026	pTS001 derived, *lacZ*::*Mtb-cysH*	AmpR	NB
pTS027	pTS001 derived, *lacZ*::*E. coli-cysH*	AmpR	NB
pTS028	pTS001 derived, *lacZ*::*Mtb-folB*	AmpR	NB
pTS029	pTS001 derived, *lacZ*::*E. coli-folB*	AmpR	NB
pTS030	pTS001 derived, *lacZ*::*Mtb-folA*	AmpR	NB
pTS031	pTS001 derived, *lacZ*::*E. coli-folA*	AmpR	NB
pTS032	pTS001 derived, *lacZ*::*Mtb-trpE*	AmpR	NB
pTS033	pTS001 derived, *lacZ*::*E. coli-trpE*	AmpR	NB
pTS034	pTS001 derived, *lacZ*::*Mtb-trpD*	AmpR	NB

Table A.3: List of plasmids.

pTS035	pTS001 derived, *lacZ*::*E. coli-trpD*	AmpR	NB
pTS036	pTS001 derived, *lacZ*::*Mtb-folP1*	AmpR	NB
pTS037	pTS001 derived, *lacZ*::*E. coli-folP1*	AmpR	NB
pTS038	pTS001 derived, *lacZ*::*Mtb-dapA*	AmpR	NB
pTS040	pTS001 derived, *lacZ*::*Mtb-dapB*	AmpR	NB
pTS041	pTS001 derived, *lacZ*::*E. coli-dapB*	AmpR	NB

References

Daniel, R., Rubens, J. R., Sarpeshkar, R. & Lu, T. K. Synthetic analog computation in living cells. Nature 497, 619–623 (2013).

Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc. Natl. Acad. Sci. U.S.A. 97, 6640–6645 (2000).

Table A.3: List of plasmids.

Appendix B

Chapter II

Original Strain name	**Relevant characteristic(s)**	**Description / Experimental Purpose**	**Gene, P1 lysate from the Keio collection** (Datsenko & Wanner, 2000) **x recipient strain**	**Source or reference**
Escherichia coli strains				
E. coli DH5α	*endA1 hsdR17 supE44 recA1 relA1 (lacZYA-argF)*	Plasmid amplification		NEB
BW25113	*Δ(araD-araB)567 ΔlacZ4787*(::rrnB-3) *λ, rph-1 Δ(rhaD-rhaB)568 hsdR514*	Parent strain		(Datsenko & Wanner, 2000)
JW0063	BW25113 Δ$araC_{Ec}$: :Kan^r			(Datsenko & Wanner, 2000)
D20	JW0063 Δ$araC_{Ec}$: :frt			This study
K028	BW25113 Δ$araC_{Ec}$: :frt, Δ$tolC_{Ec}$: :frt		$tolC_{Ec}$, JW5503 x D20	This study
K123	BW25113 Δ$araC_{Ec}$: :frt, Δ$entC_{Ec}$: :frt		$entC_{Ec}$, JW0585 x D20	This study
K031	BW25113 Δ$araC_{Ec}$: :frt, Δ$tolC_{Ec}$: :frt, Δ$entC_{Ec}$: :frt		$entC_{Ec}$, JW0585 x K028	This study
D18	BW25113 Δ$araC_{Ec}$: :frt, Δ$dadX_{Ec}$: :frt, Δalr_{Ec} : :frt		alr_{Ec}, JW4013 x D15	This study
K045	BW25113 Δ$araC_{Ec}$: :frt, Δ$tolC_{Ec}$: :frt, Δ$entC_{Ec}$: :frt Δalr_{Ec} : :frt		alr_{Ec}, JW4013 x K031	This study
K121	BW25113 Δ$araC_{Ec}$: :frt, Δ$tolC_{Ec}$: :frt, Δ$entC_{Ec}$: :frt Δalr_{Ec} : :frt Δ$dadX_{Ec}$: :frt		$dadX_{Ec}$, JW1179 x K045	This study
TS008	D18 + pTS008 & V91	TESEC strain with Alr_{Mtb} expression without gene deletions in *tolC entC* / Verification of drug efflux		This study
TS144	K122 + pTS008 & V91	Drug screening TESEC strain with Alr_{Mtb} expression / Prestwick Library Screen		This study
TS163	K031 + pTS066 & V91	Control strain with native alanine racemase expression / Prestwick Library Screen & Characterization of 2-plasmid system		This study
TS164	K031 + pTS000 & V91			

Table B.1: List of strains.

		meGFP-Alr$_{Mtb}$ expression from the 2-plasmid system		
TS123	K121 + pTS068 + V91	TESEC strain expressing Alr$_{Msmeg}$		
TS167	K122 + pTS072 & V91	TESEC strain expressing His$_6$-Alr$_{Mtb}$ / His-tag purification of Alr$_{Mtb}$		This study
Mycobacterial strains				
Mycobacterium smegmatis mc^{2}155		Hit validation of benazepril		American Type Culture Collection (ATCC)
Mycobacterium tuberculosis H37Rv		Hit validation of benazepril		Martin Pavelka, University of Rochester Medical Center

Table B.1: List of strains.

Plasmid	Relevant characteristic(s)	Antibiotic resistance marker	Source / Reference
P37.pRD131	pBAD/mCherry	Amp^r	Daniel, Rubens, Sarpeshkar, & Lu (2013)
V67	pBAD/meGFP	Cam^r	This work
V91	pBAD/*araC*	Kan^r	This work
Bba_J04450		Cam^r	iGEM
pCP20		Cam^r, Amp^r	Datsenko, K. A. & Wanner (2000).
pTS000	P37.pRD131 derivative internal BasI site removed, lacZ :: mCherry	Amp^r	This work
pTS008	pTS000 derivatived, lacZ :: alr_{Mtb}	Amp^r	This work
pTS068	pTS000 derivatived, lacZ :: alr_{Msmeg}	Amp^r	This work

MoClo plasmids **

Destination vectors	Description			Antibiotic resistance marker *	Reference
DVC_AC	Level 0 destination vector			Cam^r	This study
DVC_CI	Level 0 destination vector			Cam^r	This study
DVC_ID	Level 0 destination vector			Cam^r	This study
DVC_CD	Level 0 destination vector			Cam^r	This study
DVA_AD	Level 1 destination vector			Amp^r	This study
Level 0 & 1 vectors	**Part(s) description**	**Part source(s)**	**Destination vector**	**Antibiotic resistance marker**	
L0_006	pBAD RBS	P37.pRD131	DVC_AC	Cam^r	This study
L0_018	meGFP(without stop codon) $(GGGS)_3$	V67	DVC_CI	Cam^r	This study
L0_033	$alr_{Mtb\text{-}short}$ (without start codon)	pTS008	DVC_ID	Cam^r	This study
L0_044	meGFP	V67	DVC_CD	Cam^r	This study
L0_079	His_6-$alr_{Mtb\text{-}short}$	pTS008	DVC_CD	Cam^r	This study
pTS059	pBAD/meGFP/$(GGGS)_3$/ *alr|subscript{Mtb}*	L0_006 + L0_018 + L0_033	DVA_AD	Amp^r	This study
pTS066	pBAD/meGFP	L0_006 + L0_018	DVA_AD	Amp^r	This study
pTS072	pBAD/His_6-$alr_{Mtb\text{-}short}$	L0_006 + L0_79	DVA_AD	Amp^r	This study

* Amp^r - ampicillin resistance; Kan^r - kanamycin resistance; Cam^r – chloramphenicol resistance
** CIDAR MoClo prefix and suffixes (5'→3'): A = GGAG, C = AATG, D = AGGT and extra pre-, suffix 'I' = TTCT (this study) for fusion proteins.

Table B.2: List of plasmids.

Primer	DNA Sequence (5'→3')	Experimental Purpose	T_m (°C)	nt length
NB.142-F	AACCAGGCTTGAGTATAGCCT	PCR verification of chromosomal $araC_{Ec}$ deletion	58.5	21
NB.143-R	CTACTCCGTCAAGCCGTCAAT	PCR verification of chromosomal $araC_{Ec}$ deletion	60.1	21
NB.157-F	TCTGCTAGAATCCGCAATAATTTTA	PCR verification of chromosomal $tolC_{Ec}$ deletion	57.0	25
NB.158-R	GAAGAATGCGGCAGATAACC	PCR verification of chromosomal $entC_{Ec}$ deletion	57.0	20
NB.128-F	GGGTTATTGTCTCATGAGCGGA	PCR verification of pTS-plasmids	60.2	22
NB.129-R	CTCGCCGCAGCCGAAC	PCR verification of pTS-plasmids	61.3	16
NB.27-F	TGCTGCAATGATACCGCGGGACCCACGCTCACCG	Mutagenesis of BsaI site within AmpR of P37.pRD131	78.1	34
NB.28-R	CGGTGAGCGTGGGTCCCGCGGTATCATTGCAGCA	Mutagenesis of BsaI site within AmpR of P37.pRD131	78.1	34
NB.43-F	ATAT**GGTCTC**ATACCATGAAACGCTTTTGGGAAAATGT	Amplification of alr_{Mtb} with **BsaI**-sites for pTS008 construction	58.3	38
NB.44-R	ATAT**GGTCTC**ACGGGTTAGCGGTTTTCAGCTTCGC	Amplification of alr_{Mtb} with **BsaI**-sites for pTS008 construction	59.5	35
NB.616-F	ATAGGTGAAGACATAATG**CATCACCATCATCACCAC**ACACCGATCTCACAGACTCC	Amplification of His6-$alr_{Mtb\text{-short}}$ for L0_079 construction	58.8	55
NB.617-R	GATTAGAAGACATACCTTTAGCGGTTTTCAGCTTCGC	Amplification of His6-$alr_{Mtb\text{-short}}$ for L0_079 construction	59.5	36
NB.311 (RV-**A**)	CTTAAGGAAGACAT**CTCC**TGAGACCTCTAGAAGCGGCCGCGAATTC	Amplification of DVC_AX backbones	63.1	46
NB.233 (FW-**C**)	ATGGTAGAAGACGT**AATG**AGAGACCTACTAGTAGCGGCCGCTGCA	Amplification of DVC_XC backbones	64.3	45
NB.306 (FW-**A**)	GACCTAC**GGAG**ATGTCTTCTGCACCATAT	Amplification of DVC_AX LacZ expression cassette	50.3	29
NB.307 (RV-**C**)	ACGC**TACC**ATTACGTCTTCCCCGC	Amplification of DVC_XC LacZ expression cassette	55.0	24
NB.313 (RV-**C**)	CTTAAGGAAGACAT**CATT**TGAGACCTCTAGAAGCGGCCGCGAATTC	Amplification of DVC_CX backbone	63.1	46
NB.397 (FW-**I**)	ATGGTAGAAGACGT**TTCT**AGAGACCTACTAGTAGCGGCCGCTGC	Amplification of DVC_XI backbone	64.3	44
NB.250 (FW-**C**)	GACCTAC**AATG**ATGTCTTCTGCACCATAT	Amplification of DVC_CX LacZ expression cassettes	50.3	29
NB.396 (RV-**I**)	ACGCTAC**AGAA**ACGTCTTCCCCGCGC	Amplification of DVC_XI LacZ expression cassettes	59.3	26
NB.399 (RV-**I**)	CTTAAGGAAGACAT**AGAA**TGAGACCTCTAGAAGCGGCC	Amplification of DVC_IX backbones	63.1	38
NB.236 (FW-**D**)	ATGGTAGAAGACGT**AGGT**AGAGACCTACTAGTAGCGGCCGCTGCA	Amplification of DVC_XD backbones	64.3	45
NB.398 (FW-**I**)	GACCTAC**TTCT**ATGTCTTCTGCACCATATGCG	Amplification of DVC_IX LacZ expression cassette	58.8	32
NB.255 (RV-**D**)	ACGCTAC**ACCT**ACGTCTTCCCCGC	Amplification of DVC_XD LacZ expression cassette	50.4	24

Table B.3: List of primers.

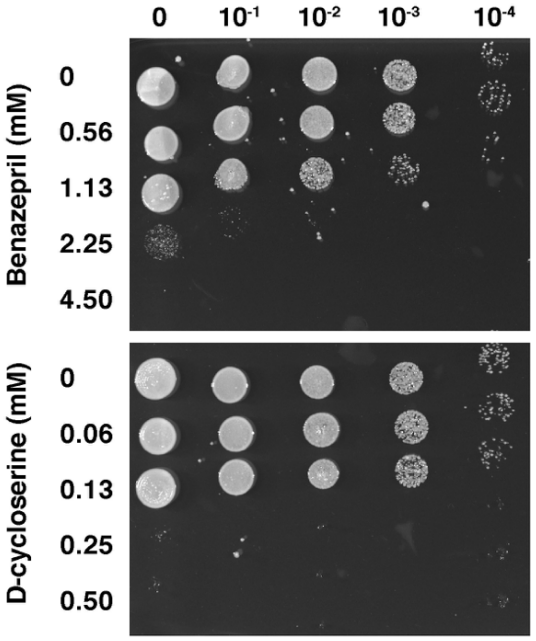

Figure B.1: Antibacterial activity of benazepril and D-cycloserine on *M. smegmatis* mc^{2}155.
M. smegmatis liquid cultures were exposed to D-cycloserine (0-0.5 mM) and benazepril (0-4.5 mM) series for 48 hours. After two days of incubation, serial dilutions were prepared of the bacterial cultures and spotted as 1 μL aliquots onto 7H9 agar plates. A photo was taken after 48 hours of incubation at 37°C. The depicted image is representative for five independent experiments.

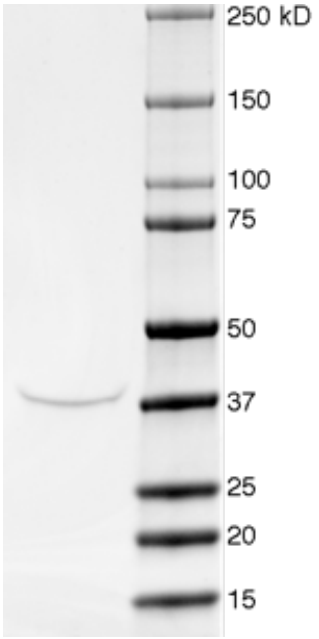

Figure B.2: Isolated His$_6$-Alr$_{Mtb}$-short (~41 kDa monomer).
15 μL of protein sample (first lane) on a 4-15% Mini-PROTEIN® TGX Protein gel stained with Coomassie blue and 5 μL Precision Plus Protein^{TM} Dual Color Standards (Bio-Rad) (second lane).

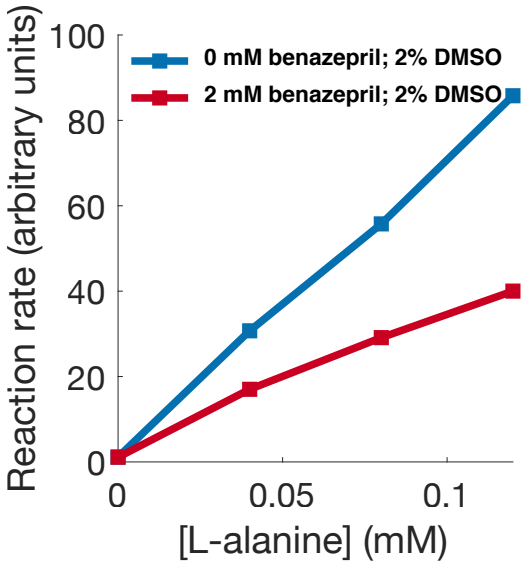

Figure B.3: Detection of alanine transferase inhibition by benazepril
Inhibition of alanine transferase in Alanine Assay Kit (MAK001) from Sigma. The assay was performed with 0-0.2 mM L-alanine concentrations, and with 0 or 2 mM benazepril. The assay was executed according to the supplier's instructions. The reaction rate was determined based on linear regression of fluorescence time course measurements (λ_{ex} = 535/ λ_{em} = 587 nm).

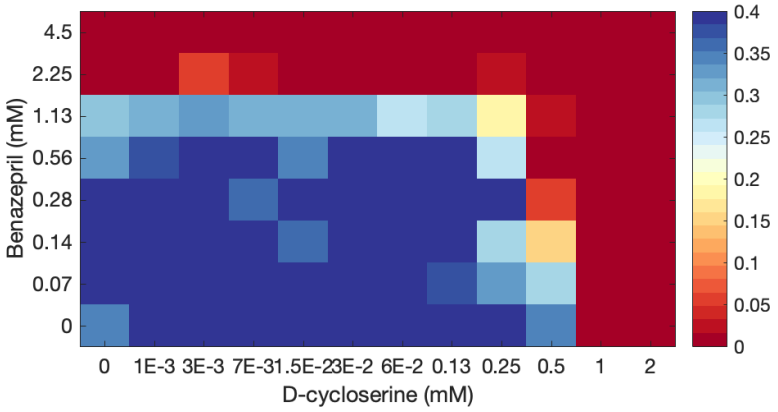

Figure B.4: Additive effect of benazepril and D-cycloserine on *M. smegmatis* mc^{2}155
100 μL M. smegmatis cultures in 7H9 medium were cultivated in 96-well plates with checkerboard titrations of benazepril and D-cycloserine drug combinations. Colors represent the mean OD_{600} values of three technical replicates taken after 48 hours of incubation at 37°C, 750 rpm. Based on a Response Envelope Analysis, SI = 0.039 and AI = 0.027.

Appendix C

Chapter III

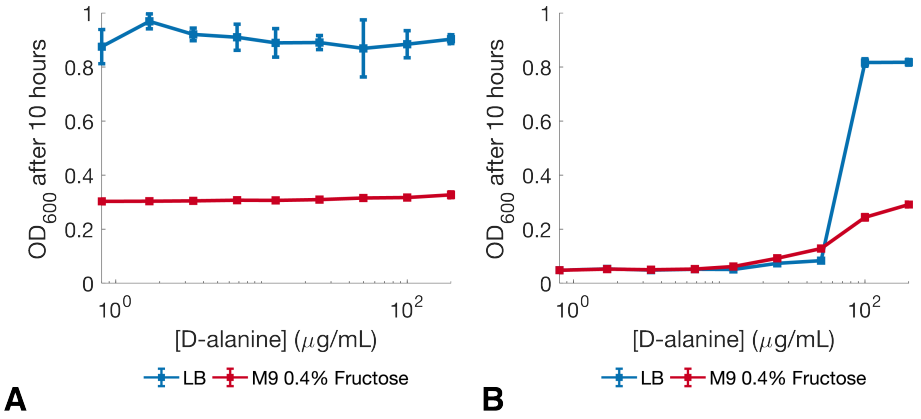

Figure C.1: CycA repression in LB medium
Non-auxotropic strain D20 **(A)** and auxotrophic strain D18 ($\Delta alr \Delta dadX$) **(B)** were grown at 37°C in the presence of 0-200 μg/mL D-alanine series in either LB or M9 0.4% fructose medium. Error bars represent the standard deviations of three technical replicates.

Source	original aa	aa position within Alr_{Mtb}	new aa	
Coll et al. (2018)	F	4	Y	non-synonymous
GMTV Database	E	6	D	non-synonymous
GMTV Database	N	12	H	non-synonymous
Coll et al. (2018)	N	12	H	non-synonymous
ReSeqTB	R	18	R	synonymous
Coll et al. (2018)	G	19	R	non-synonymous
Coll et al. (2018)	G	19	S	non-synonymous
GMTV Database	S	22	L	non-synonymous
Coll et al. (2018)	S	22	L	non-synonymous
Desjardins et al. (2016)	S	22	L	non-synonymous
Coll et al. (2018)	L	23	M	non-synonymous
Coll et al. (2018)	T	26	I	non-synonymous
Coll et al. (2018)	S	29	F	non-synonymous
Portelli & Phelan et al. (2018)	A	38	V	non-synonymous
Coll et al. (2018)	A	38	V	non-synonymous
GMTV Database	N	48	N	synonymous
GMTV Database	A	56	A	synonymous
Portelli & Phelan et al. (2018)	G	71	S	non-synonymous
GMTV Database	G	71	S	non-synonymous
Coll et al. (2018)	G	71	S	non-synonymous
Portelli & Phelan et al. (2018)	H	72	Y	non-synonymous
Coll et al. (2018)	H	72	Y	non-synonymous
Portelli & Phelan et al. (2018)	T	75	A	non-synonymous
Portelli & Phelan et al. (2018)	T	75	M	non-synonymous
Coll et al. (2018)	T	75	A	non-synonymous
Coll et al. (2018)	T	75	M	non-synonymous
GMTV Database	L	82	R	non-synonymous
GMTV Database	A	107	T	non-synonymous
GMTV Database	L	113	R	non-synonymous
Desjardins et al. (2016)	L	113	R	non-synonymous
ReSeqTB	L	113	R	non-synonymous
Portelli & Phelan et al. (2018)	P	122	S	non-synonymous

Table C.1: Published *alr* mutations of *M. tuberculosis* clinical isolates.

Coll et al. (2018)	P	122	S	non-synonymous
Portelli & Phelan et al. (2018)	L	133	R	non-synonymous
Portelli & Phelan et al. (2018)	D	139	H	non-synonymous
GMTV Database	D	139	H	non-synonymous
Coll et al. (2018)	D	139	H	non-synonymous
Portelli & Phelan et al. (2018)	D	139	G	non-synonymous
Coll et al. (2018)	D	139	G	non-synonymous
Portelli & Phelan et al. (2018)	E	140	D	non-synonymous
Coll et al. (2018)	E	140	D	non-synonymous
Portelli & Phelan et al. (2018)	T	155	A	non-synonymous
Coll et al. (2018)	T	155	A	non-synonymous
Portelli & Phelan et al. (2018)	V	156	A	non-synonymous
Coll et al. (2018)	V	156	A	non-synonymous
GMTV Database	V	156	V	synonymous
Portelli & Phelan et al. (2018)	K	157	E	non-synonymous
Coll et al. (2018)	K	157	E	non-synonymous
Desjardins et al. (2016)	K	157	E	non-synonymous
Portelli & Phelan et al. (2018)	T	160	A	non-synonymous
Coll et al. (2018)	T	160	A	non-synonymous
GMTV Database	N	163	N	synonymous
GMTV Database	R	164	H	non-synonymous
GMTV Database	F	172	F	synonymous
GMTV Database	A	174	A	synonymous
GMTV Database	A	184	A	synonymous
Portelli & Phelan et al. (2018)	D	186	G	non-synonymous
Coll et al. (2018)	D	186	G	synonymous
Portelli & Phelan et al. (2018)	A	187	S	non-synonymous
Coll et al. (2018)	A	187	S	non-synonymous
GMTV Database	A	200	D	non-synonymous
Portelli & Phelan et al. (2018)	D	205	G	non-synonymous

Table C.1: Published *alr* mutations of *M. tuberculosis* clinical isolates.

Coll et al. (2018)	D	205	G	non-synonymous
Portelli & Phelan et al. (2018)	A	212	D	non-synonymous
Coll et al. (2018)	A	212	D	non-synonymous
Portelli & Phelan et al. (2018)	F	215	V	non-synonymous
GMTV Database	F	215	V	non-synonymous
Coll et al. (2018)	F	215	V	non-synonymous
Portelli & Phelan et al. (2018)	A	217	V	non-synonymous
Coll et al. (2018)	A	217	V	non-synonymous
GMTV Database	L	219	L	synonymous
ReSeqTB	L	219	L	synonymous
Portelli & Phelan et al. (2018)	E	224	G	non-synonymous
GMTV Database	E	224	G	non-synonymous
GMTV Database	L	234	L	synonymous
ReSeqTB	L	234	L	synonymous
Portelli & Phelan et al. (2018)	S	235	W	non-synonymous
Coll et al. (2018)	S	235	W	non-synonymous
Portelli & Phelan et al. (2018)	S	238	L	non-synonymous
GMTV Database	S	238	L	non-synonymous
Coll et al. (2018)	S	238	L	non-synonymous
Portelli & Phelan et al. (2018)	T	247	M	non-synonymous
Coll et al. (2018)	T	247	M	non-synonymous
GMTV Database	P	252	A	non-synonymous
GMTV Database	A	256	S	non-synonymous
Portelli & Phelan et al. (2018)	L	260	V	non-synonymous
Coll et al. (2018)	L	260	V	non-synonymous
Portelli & Phelan et al. (2018)	S	261	N	non-synonymous
Coll et al. (2018)	S	261	N	non-synonymous
Portelli & Phelan et al. (2018)	P	262	Q	non-synonymous
Coll et al. (2018)	P	262	Q	non-synonymous
Portelli & Phelan et al. (2018)	D	268	G	non-synonymous
GMTV Database	D	268	G	non-synonymous
Coll et al. (2018)	D	268	G	non-synonymous
Portelli & Phelan et al. (2018)	G	270	E	non-synonymous
Coll et al. (2018)	G	270	E	non-synonymous
Portelli & Phelan et al. (2018)	M	275	I	non-synonymous
Coll et al. (2018)	M	275	I	non-synonymous
Portelli & Phelan et al. (2018)	V	284	L	non-synonymous
Coll et al. (2018)	V	284	L	non-synonymous
Portelli & Phelan et al. (2018)	I	287	V	non-synonymous

Table C.1: Published *alr* mutations of *M. tuberculosis* clinical isolates.

Coll et al. (2018)	I	287	V	non-synonymous
Portelli & Phelan et al. (2018)	A	308	G	non-synonymous
GMTV Database	A	308	G	non-synonymous
Coll et al. (2018)	A	308	G	non-synonymous
Desjardins et al. (2016)	A	308	G	non-synonymous
Portelli & Phelan et al. (2018)	P	311	L	non-synonymous
GMTV Database	P	311	S	non-synonymous
Coll et al. (2018)	P	311	L	non-synonymous
Portelli & Phelan et al. (2018)	D	316	E	non-synonymous
Coll et al. (2018)	D	316	E	non-synonymous
Portelli & Phelan et al. (2018)	V	318	M	non-synonymous
Coll et al. (2018)	V	318	M	non-synonymous
GMTV Database	S	321	L	non-synonymous
ReSeqTB	S	321	L	non-synonymous
Portelli & Phelan et al. (2018)	R	325	P	non-synonymous
Coll et al. (2018)	R	325	P	non-synonymous
GMTV Database	R	334	Q	non-synonymous
Portelli & Phelan et al. (2018)	R	340	L	non-synonymous
GMTV Database	R	340	L	non-synonymous
Coll et al. (2018)	R	340	L	non-synonymous
Portelli & Phelan et al. (2018)	M	343	T	non-synonymous
GMTV Database	M	343	T	non-synonymous
Desjardins et al. (2016)	M	343	T	non-synonymous
Köse et al. (2013)	M	343	T	non-synonymous
Chen et al. (2017)	D	344	N	non-synonymous
Portelli & Phelan et al. (2018)	L	350	P	non-synonymous
Coll et al. (2018)	L	350	P	non-synonymous
GMTV Database	P	352	A	non-synonymous
Portelli & Phelan et al. (2018)	A	363	T	non-synonymous
Coll et al. (2018)	A	363	T	non-synonymous
Portelli & Phelan et al. (2018)	I	364	V	non-synonymous
Coll et al. (2018)	I	364	V	non-synonymous
Portelli & Phelan et al. (2018)	E	373	G	non-synonymous
Coll et al. (2018)	E	373	G	non-synonymous
GMTV Database	I	386	I	synonymous
ReSeqTB	I	386	I	synonymous
Portelli & Phelan et al. (2018)	Y	388	C	non-synonymous
Coll et al. (2018)	Y	388	C	non-synonymous

Table C.1: Published *alr* mutations of *M. tuberculosis* clinical isolates.

Portelli & Phelan et al. (2018)	Y	388	D	non-synonymous
GMTV Database	Y	388	D	non-synonymous
Coll et al. (2018)	Y	388	D	non-synonymous
Desjardins et al. (2016)	Y	388	D	non-synonymous
ReSeqTB	Y	388	D	non-synonymous
Portelli & Phelan et al. (2018)	R	397	G	non-synonymous
Coll et al. (2018)	R	397	G	non-synonymous
Desjardins et al. (2016)	R	397	G	non-synonymous
Portelli & Phelan et al. (2018)	R	397	L	non-synonymous
Coll et al. (2018)	R	397	L	non-synonymous
Portelli & Phelan et al. (2018)	T	401	I	non-synonymous
Coll et al. (2018)	T	401	I	non-synonymous
Coll et al. (2018)	E	406	V	non-synonymous

References

Coll, F., Phelan, J., Hill-Cawthorne, G. A., Nair, M. B., Mallard, K., Ali, S., et al. (2018). Genome-wide analysis of multi- and extensively drug-resistant Mycobacterium tuberculosis. Nature Genetics, 50(5), 764–764. http://doi.org/10.1038/s41588-018-0074-3

Desjardins, C. A., Cohen, K. A., Munsamy, V., Abeel, T., Maharaj, K., Walker, B. J., et al. (2016). Genomic and functional analyses of Mycobacterium tuberculosis strains implicate ald in D-cycloserine resistance. Nature Genetics, 48(5), 544–551. http://doi.org/10.1038/ng.3548

GMTV Database, viewed August 2019, https://mtb.dobzhanskycenter.org/cgi-bin/beta/main.py#custom/world

Portelli, S., Phelan, J. E., Ascher, D. B., Clark, T. G., & Furnham, N. (2018). Understanding molecular consequences of putative drug resistant mutations in Mycobacterium tuberculosis. Scientific Reports, 8(1), 15356. http://doi.org/10.1038/s41598-018-33370-6

ReSeqTB Database, viewed August 2019, https://platform.reseqtb.org

Table C.1: Published *alr* mutations of *M. tuberculosis* clinical isolates.

High Frequencies mutations in Error Prone PCR *Mtb alr* mutant library

position in gen	mutation	frequency	annotation	gene
1153	A→G	5,10%	T385A (ACT→GCT) ‡	*alr* →
1142	A→G	4,40%	D381G (GAC→GGC) ‡	*alr* →
1159	C→A	3,20%	H387N (CAC→AAC) ‡	*alr* →
1145	T→G	2,90%	L382R (CTG→CGG) ‡	*alr* →
1191	T→G	2,10%	R397R (CGT→CGG) ‡	*alr* →
1179	A→C	2,00%	S393S (TCA→TCC) ‡	*alr* →
1130	A→C	1,60%	Q377P (CAA→CCA)	*alr* →
1173	A→T	1,40%	V391V (GTA→GTT) ‡	*alr* →
1053	T→G	1,30%	G351G (GGT→GGG)	*alr* →
1135	T→A	1,30%	W379R (TGG→AGG) ‡	*alr* →
1162	T→A	1,20%	Y388N (TAC→AAC) ‡	*alr* →
1195	A→T	1,10%	T399S (ACC→TCC) ‡	*alr* →
1155	T→G	0,96%	T385T (ACT→ACG) ‡	*alr* →
1177	T→A	0,72%	S393T (TCA→ACA) ‡	*alr* →
1043	T→G	0,69%	V348G (GTC→GGC) ‡	*alr* →
1182	G→C	0,67%	P394P (CCG→CCC) ‡	*alr* →
514	T→A	0,64%	F172I (TTC→ATC) ‡	*alr* →
1003	T→A	0,59%	C335S (TGC→AGC)	*alr* →
1133	A→C	0,54%	D378A (GAC→GCC) ‡	*alr* →
967	G→T	0,50%	G323C (GGC→TGC)	*alr* →
1157	T→A	0,48%	I386N (ATT→AAT) ‡	*alr* →
213	T→G	0,39%	G71G (GGT→GGG) ‡	*alr* →
1017	T→G	0,35%	G339G (GGT→GGG) ‡	*alr* →
761	G→A	0,33%	G254D (GGT→GAT)	*alr* →
1207	C→A	0,33%	R403S (CGC→AGC) ‡	*alr* →
1015	G→T	0,30%	G339C (GGT→TGT) ‡	*alr* →
440	G→A	0,30%	R147H (CGT→CAT) ‡	*alr* →
1156	A→T	0,29%	I386F (ATT→TTT) ‡	*alr* →
628	G→A	0,28%	V210I (GTA→ATA)	*alr* →
191	T→G	0,28%	V64G (GTC→GGC) ‡	*alr* →
1164	C→A	0,27%	Y388* (TAC→TAA) ‡	*alr* →
1047	T→G	0,24%	D349E (GAT→GAG) ‡	*alr* →
350	G→A	0,23%	G117D (GGT→GAT) ‡	*alr* →
719	C→T	0,23%	T240M (ACG→ATG) ‡	*alr* →
24	G→T	0,23%	V8V (GTG→GTT)	*alr* →
234	T→G	0,22%	A78A (GCT→GCG) ‡	*alr* →
746	A→T	0,22%	D249V (GAT→GTT) ‡	*alr* →
424	C→A	0,22%	L142M (CTG→ATG) ‡	*alr* →
32	C→A	0,22%	P11H (CCT→CAT)	*alr* →
373	C→A	0,20%	L125M (CTG→ATG)	*alr* →

Table C.2: Mutation frequencies in the epPCR *Mtb* Alr mutant library

675	G→T	0,19%	Q225H (CAG→CAT) ‡	*alr* →
1046	A→T	0,18%	D349V (GAT→GTT) ‡	*alr* →
655	C→A	0,18%	L219M (CTG→ATG)	*alr* →
124	C→A	0,18%	L42I (CTC→ATC)	*alr* →
774	C→T	0,18%	Y258Y (TAC→TAT)	*alr* →
1192	A→T	0,17%	I398F (ATT→TTT) ‡	*alr* →
339	G→A	0,17%	L113L (CTG→CTA) ‡	*alr* →
906	C→A	0,17%	P302P (CCC→CCA) ‡	*alr* →
1184	G→C	0,17%	R395P (CGC→CCC)	*alr* →
560	C→A	0,16%	A187D (GCT→GAT)	*alr* →
1143	C→A	0,16%	D381E (GAC→GAA) ‡	*alr* →
387	G→T	0,16%	Q129H (CAG→CAT) ‡	*alr* →
1009	G→T	0,15%	G337C (GGT→TGT) ‡	*alr* →
429	C→T	0,15%	H143H (CAC→CAT)	*alr* →
1034	A→T	0,15%	Q345L (CAG→CTG) ‡	*alr* →
1199	G→C	0,15%	R400P (CGT→CCT) ‡	*alr* →
1196	C→A	0,15%	T399N (ACC→AAC) ‡	*alr* →
210	C→T	0,15%	Y70Y (TAC→TAT)	*alr* →
331	G→T	0,14%	A111S (GCG→TCG) ‡	*alr* →
25	G→T	0,14%	G9C (GGT→TGT)	*alr* →
700	C→A	0,14%	L234M (CTG→ATG) ‡	*alr* →
758	C→A	0,14%	P253H (CCT→CAT) ‡	*alr* →
1132	G→C	0,13%	D378H (GAC→CAC) ‡	*alr* →
555	A→T	0,13%	E185D (GAA→GAT)	*alr* →
981	A→T	0,13%	E327D (GAA→GAT) ‡	*alr* →
874	G→T	0,13%	G292C (GGC→TGC) ‡	*alr* →
577	C→A	0,13%	L193M (CTG→ATG) ‡	*alr* →
983	T→A	0,13%	V328D (GTT→GAT) ‡	*alr* →
1137	G→T	0,13%	W379C (TGG→TGT) ‡	*alr* →
1128	C→A	0,12%	A376A (GCC→GCA)	*alr* →
207	C→A	0,12%	G69G (GGC→GGA)	*alr* →
254	G→A	0,12%	G85D (GGT→GAT) ‡	*alr* →
1193	T→A	0,12%	I398N (ATT→AAT) ‡	*alr* →
133	A→T	0,12%	I45F (ATT→TTT) ‡	*alr* →
198	G→T	0,12%	K66N (AAG→AAT) ‡	*alr* →
568	C→A	0,12%	L190M (CTG→ATG) ‡	*alr* →
67	C→A	0,12%	L23I (CTC→ATC)	*alr* →
811	C→A	0,12%	L271M (CTG→ATG) ‡	*alr* →
520	G→T	0,11%	A174S (GCG→TCG) ‡	*alr* →
502	G→T	0,11%	G168C (GGT→TGT) ‡	*alr* →
1011	T→G	0,11%	G337G (GGT→GGG) ‡	*alr* →
370	C→A	0,11%	L124M (CTG→ATG)	*alr* →
812	T→G	0,11%	L271R (CTG→CGG) ‡	*alr* →

Table C.2: Mutation frequencies in the epPCR *Mtb* Alr mutant library

63	C→A	0,11%	T21T (ACC→ACA)	*alr* →
634	G→T	0,10%	A212S (GCG→TCG)	*alr* →
1031	A→T	0,10%	D344V (GAT→GTT) ‡	*alr* →
122	A→C	0,10%	D41A (GAC→GCC) ‡	*alr* →
690	A→T	0,10%	E230D (GAA→GAT) ‡	*alr* →
587	A→T	0,10%	H196L (CAT→CTT)	*alr* →
1161	C→A	0,10%	H387Q (CAC→CAA) ‡	*alr* →
975	T→G	0,10%	R325R (CGT→CGG)	*alr* →
1020	T→A	0,10%	R340R (CGT→CGA) ‡	*alr* →
226	C→T	0,10%	R76C (CGC→TGC) ‡	*alr* →
382	G→C	0,10%	V128L (GTT→CTT) ‡	*alr* →
953	T→G	0,10%	V318G (GTT→GGT)	*alr* →
123	C→G	0,10%	D41E (GAC→GAG) ‡	*alr* →
742	T→A	0,10%	F248I (TTC→ATC) ‡	*alr* →
243	C→A	0,10%	A81A (GCC→GCA)	*alr* →
814	G→A	0,10%	V272M (GTG→ATG) ‡	*alr* →
887	G→T	0,10%	G296V (GGG→GTG)	*alr* →
154	C→A	0,10%	L52I (CTC→ATC)	*alr* →
508	G→T	0,09%	A170S (GCA→TCA) ‡	*alr* →
513	A→T	0,09%	Q171H (CAA→CAT)	*alr* →
945	C→A	0,09%	A315A (GCC→GCA)	*alr* →
803	A→T	0,09%	D268V (GAT→GTT)	*alr* →
736	C→A	0,09%	L246M (CTG→ATG)	*alr* →
204	C→T	0,09%	D68D (GAC→GAT) ‡	*alr* →
779	T→G	0,09%	L260R (CTT→CGT) ‡	*alr* →
543	G→T	0,09%	Q181H (CAG→CAT) ‡	*alr* →
901	G→T	0,09%	A301S (GCA→TCA)	*alr* →
744	C→A	0,09%	F248L (TTC→TTA) ‡	*alr* →
898	A→T	0,09%	I300F (ATC→TTC) ‡	*alr* →
181	T→A	0,09%	L61M (TTG→ATG) ‡	*alr* →
420	A→T	0,09%	E140D (GAA→GAT) ‡	*alr* →
83	T→A	0,09%	I28N (ATC→AAC) ‡	*alr* →
684	T→G	0,09%	R228R (CGT→CGG) ‡	*alr* →
457	A→C	0,09%	T153P (ACC→CCC) ‡	*alr* →
785	C→A	0,09%	P262Q (CCG→CAG) ‡	*alr* →
202	G→T	0,09%	D68Y (GAC→TAC) ‡	*alr* →
327	A→T	0,09%	V109V (GTA→GTT)	*alr* →
1167	A→T	0,09%	E389D (GAA→GAT) ‡	*alr* →
934	A→T	0,09%	I312F (ATT→TTT) ‡	*alr* →
337	C→A	0,09%	L113M (CTG→ATG) ‡	*alr* →

Table C.2: Mutation frequencies in the epPCR *Mtb* Alr mutant library

1225	T→A	0,08%	*409K (TAA→AAA) ‡	*alr* →
614	A→T	0,08%	D205V (GAT→GTT)	*alr* →
786	G→T	0,08%	P262P (CCG→CCT) ‡	*alr* →
856	A→T	0,08%	S286C (AGT→TGT)	*alr* →
463	A→C	0,08%	T155P (ACG→CCG) ‡	*alr* →
1040	T→A	0,08%	M347K (ATG→AAG) ‡	*alr* →
854	A→T	0,08%	K285I (AAA→ATA) ‡	*alr* →
882	T→A	0,08%	S294S (TCT→TCA) ‡	*alr* →
152	T→G	0,08%	V51G (GTC→GGC) ‡	*alr* →
939	C→A	0,08%	G313G (GGC→GGA) ‡	*alr* →
526	T→A	0,08%	L176I (TTA→ATA) ‡	*alr* →
590	T→A	0,08%	M197K (ATG→AAG) ‡	*alr* →
767	C→A	0,08%	A256D (GCT→GAT) ‡	*alr* →
730	C→A	0,08%	P244T (CCA→ACA)	*alr* →
734	A→C	0,08%	D245A (GAC→GCC) ‡	*alr* →
552	T→G	0,08%	A184A (GCT→GCG) ‡	*alr* →
881	C→A	0,08%	S294Y (TCT→TAT) ‡	*alr* →
1154	C→G	0,08%	T385S (ACT→AGT) ‡	*alr* →
238	A→C	0,08%	T80P (ACG→CCG) ‡	*alr* →
521	C→A	0,08%	A174E (GCG→GAG) ‡	*alr* →
711	T→A	0,08%	S237R (AGT→AGA) ‡	*alr* →
860	T→A	0,07%	I287N (ATT→AAT) ‡	*alr* →
1223	G→C	0,07%	R408P (CGC→CCC)	*alr* →
741	A→T	0,07%	T247T (ACA→ACT) ‡	*alr* →
893	C→G	0,07%	T298R (ACA→AGA) ‡	*alr* →
867	G→T	0,07%	A289A (GCG→GCT) ‡	*alr* →
471	A→T	0,07%	K157N (AAA→AAT) ‡	*alr* →
347	C→A	0,07%	P116Q (CCG→CAG) ‡	*alr* →
652	T→A	0,07%	F218I (TTC→ATC)	*alr* →
855	A→T	0,07%	K285N (AAA→AAT) ‡	*alr* →
348	G→T	0,07%	P116P (CCG→CCT) ‡	*alr* →
179	A→T	0,07%	Q60L (CAG→CTG) ‡	*alr* →
1124	C→G	0,07%	T375S (ACC→AGC) ‡	*alr* →
1186	G→C	0,07%	G396R (GGG→CGG)	*alr* →
605	A→T	0,07%	K202M (AAG→ATG) ‡	*alr* →
1008	A→T	0,07%	P336P (CCA→CCT) ‡	*alr* →
542	A→T	0,07%	Q181L (CAG→CTG) ‡	*alr* →
358	T→A	0,07%	F120I (TTT→ATT) ‡	*alr* →
834	A→T	0,07%	K278N (AAA→AAT) ‡	*alr* →
917	A→T	0,07%	N306I (AAT→ATT) ‡	*alr* →
334	T→G	0,07%	W112G (TGG→GGG)	*alr* →
609	C→A	0,07%	P203P (CCC→CCA)	*alr* →
1172	T→G	0,07%	V391G (GTA→GGA) ‡	*alr* →

Table C.2: Mutation frequencies in the epPCR *Mtb* Alr mutant library

600	A→T	0,07%	A200A (GCA→GCT)	*alr* →
131	C→A	0,07%	A44D (GCT→GAT)	*alr* →
233	C→A	0,07%	A78D (GCT→GAT) ‡	*alr* →
523	A→T	0,07%	M175L (ATG→TTG) ‡	*alr* →
36	C→A	0,07%	N12K (AAC→AAA)	*alr* →
1019	G→T	0,07%	R340L (CGT→CTT) ‡	*alr* →
135	T→A	0,07%	I45I (ATT→ATA) ‡	*alr* →
740	C→G	0,07%	T247R (ACA→AGA) ‡	*alr* →
421	T→A	0,07%	L141M (TTG→ATG)	*alr* →
66	C→A	0,07%	S22S (TCC→TCA) ‡	*alr* →
378	C→A	0,07%	A126A (GCC→GCA)	*alr* →
108	C→A	0,07%	A36A (GCC→GCA)	*alr* →
84	C→A	0,07%	I28I (ATC→ATA) ‡	*alr* →
847	C→A	0,07%	L283I (CTC→ATC) ‡	*alr* →
252	C→A	0,06%	A84A (GCC→GCA)	*alr* →
182	T→A	0,06%	L61* (TTG→TAG) ‡	*alr* →
71	C→A	0,06%	A24E (GCG→GAG)	*alr* →
359	T→A	0,06%	F120Y (TTT→TAT) ‡	*alr* →
676	G→T	0,06%	G226C (GGC→TGC)	*alr* →
606	G→T	0,06%	K202N (AAG→AAT) ‡	*alr* →
444	C→A	0,06%	T148T (ACC→ACA)	*alr* →
113	C→A	0,06%	A38E (GCG→GAG)	*alr* →
686	T→A	0,06%	F229Y (TTT→TAT) ‡	*alr* →
988	A→T	0,06%	I330F (ATT→TTT) ‡	*alr* →
691	G→T	0,06%	V231L (GTG→TTG)	*alr* →
221	C→A	0,06%	A74D (GCT→GAT) ‡	*alr* →
481	G→T	0,06%	G161* (GGA→TGA) ‡	*alr* →
799	G→T	0,06%	G267C (GGT→TGT)	*alr* →
139	C→G	0,06%	H47D (CAT→GAT) ‡	*alr* →
412	T→A	0,06%	L138M (TTG→ATG) ‡	*alr* →
100	T→A	0,06%	L34I (TTA→ATA) ‡	*alr* →
453	C→A	0,06%	T151T (ACC→ACA) ‡	*alr* →
1044	C→A	0,06%	V348V (GTC→GTA) ‡	*alr* →
138	A→T	0,06%	E46D (GAA→GAT) ‡	*alr* →
620	T→A	0,06%	I207N (ATT→AAT) ‡	*alr* →
1176	C→A	0,06%	T392T (ACC→ACA) ‡	*alr* →
873	A→T	0,06%	E291D (GAA→GAT)	*alr* →
899	T→A	0,06%	I300N (ATC→AAC) ‡	*alr* →
748	C→A	0,06%	L250M (CTG→ATG) ‡	*alr* →
1144	C→A	0,06%	L382M (CTG→ATG) ‡	*alr* →
1174	A→T	0,06%	T392S (ACC→TCC) ‡	*alr* →
56	G→T	0,06%	G19V (GGC→GTC) ‡	*alr* →
935	T→A	0,06%	I312N (ATT→AAT) ‡	*alr* →

Mutant Library	Mutation			Initial frequency >0.01%	Type
epPCR					
epPCR	G	9	S		non-synonymous
epPCR	S	22	P		non-synonymous
epPCR	L	61	M	yes	non-synonymous
epPCR	A	99	A		synonymous
epPCR	L	100	L		synonymous
epPCR	F	120	F		synonymous
epPCR	L	124	L		synonymous
epPCR	L	125	M	yes	non-synonymous
epPCR	D	139	E	yes	non-synonymous
epPCR	L	142	M	yes	non-synonymous
epPCR	T	155	T	yes	synonymous
epPCR	T	160	T	yes	synonymous
epPCR	G	168	S		non-synonymous
epPCR	T	177	T	yes	synonymous
epPCR	V	210	I	yes	non-synonymous
epPCR	A	217	A	yes	synonymous
epPCR	A	222	A		synonymous
epPCR	V	227	V		synonymous
epPCR	L	234	L	yes	synonymous
epPCR	Y	258	Y	yes	synonymous
epPCR	G	259	G		synonymous
epPCR	V	263	V		synonymous
epPCR	L	271	L		synonymous
epPCR	S	286	N		non-synonymous
epPCR	P	302	P	yes	synonymous
epPCR	P	311	S		non-synonymous
epPCR	L	326	L	yes	synonymous
epPCR	I	330	I		synonymous
epPCR	G	353	G		synonymous
epPCR	S	393	P		non-synonymous
Rational					
Rational	M	183	V	unknown	non-synonymous
Rational	A	220	D	unknown	non-synonymous
Rational	R	340	S	unknown	non-synonymous

Table C.2: Mutations from library colonies that were picked from both non-selective and selective plates

Pass/Fail Table

Position	WT	A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y	* Expected	Average	Observed	>50%
1	M	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
2	K	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
3	R	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
4	F	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
5	W	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	19	5.26%	19	PASS
6	E	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
7	N	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
8	V	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	19	5.26%	19	PASS
9	G	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
10	K	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
11	P	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
12	N	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
13	D	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
14	T	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	19	5.26%	19	PASS
15	T	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	19	5.26%	19	PASS
16	D	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
17	G	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
18	R	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
19	G	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
20	T	FAIL	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	FAIL	PASS	PASS	PASS	WT	PASS	PASS	PASS	19	5.26%	17	PASS
21	T	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	19	5.26%	19	PASS
22	S	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
23	L	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
24	A	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
25	M	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
26	T	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	FAIL	PASS	PASS	19	5.26%	18	PASS
27	P	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
28	I	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
29	S	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
30	Q	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
31	T	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	19	5.26%	19	PASS
32	P	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS

Table C.3: Rational Saturation Library Quality Report from Twist Bioscience.

33	G	PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
34	L	PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
35	L	PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
36	A	WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
37	E	PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
38	A	WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
39	M	PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
40	V	PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS	19	5.26%	19	PASS
41	D	PASS FAIL WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS FAIL FAIL PASS PASS PASS PASS	19	5.26%	16	PASS
42	L	PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
43	G	PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
44	A	WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
45	I	PASS PASS PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
46	E	PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
47	H	PASS PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
48	N	PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
49	**V**	**PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS**	**19**	**5.26%**	**19**	**PASS**
50	**R**	**PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS**	**19**	**5.26%**	**19**	**PASS**
51	V	PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS	19	5.26%	19	PASS
52	L	PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS FAIL PASS PASS PASS PASS PASS PASS	19	5.26%	18	PASS
53	R	PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS	19	5.26%	19	PASS
54	E	PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
55	H	FAIL FAIL PASS PASS PASS PASS WT PASS PASS PASS PASS FAIL PASS PASS FAIL FAIL PASS PASS PASS PASS	19	5.26%	14	PASS
56	A	WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
57	G	PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
58	H	PASS PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS FAIL PASS PASS PASS PASS PASS PASS PASS	19	5.26%	18	PASS
59	A	WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
60	Q	PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS WT FAIL PASS PASS PASS PASS PASS	19	5.26%	18	PASS
61	L	PASS PASS PASS PASS PASS FAIL PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	18	PASS
62	M	PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
63	A	WT PASS PASS PASS PASS FAIL PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	18	PASS
64	V	PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS	19	5.26%	19	PASS
65	V	PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS	19	5.26%	19	PASS
66	K	PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS

Table C.3: Rational Saturation Library Quality Report from Twist Bioscience.

67	A	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
68	D	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
69	G	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
70	Y	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	19	5.26%	19	PASS
71	G	FAIL	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	18	PASS
72	H	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
73	G	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
74	A	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
75	T	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	19	5.26%	19	PASS
76	R	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
77	V	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	19	5.26%	19	PASS
78	A	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
79	Q	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
80	T	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	19	5.26%	19	PASS
81	A	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
82	L	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
83	G	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
84	A	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
85	G	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
86	A	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
87	A	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
88	E	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
89	L	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
90	G	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
91	V	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	19	5.26%	19	PASS
92	A	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
93	T	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	19	5.26%	19	PASS
94	V	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	19	5.26%	19	PASS
95	D	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
96	E	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
97	A	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
98	L	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
99	A	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS
100	L	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	WT	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS	19	5.26%	19	PASS

Table C.3: Rational Saturation Library Quality Report from Twist Bioscience.

101	R	PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS	19	5.26%	19	PASS
102	A	WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
103	D	PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
104	G	FAIL PASS PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	18	PASS
105	I	PASS PASS PASS PASS PASS FAIL PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	18	PASS
106	T	PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS PASS	19	5.26%	19	PASS
107	A	WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
108	P	PASS PASS FAIL PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS	19	5.26%	18	PASS
109	V	PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS	19	5.26%	19	PASS
110	L	PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
111	A	WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
112	W	PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS	19	5.26%	19	PASS
113	L	PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
114	H	PASS PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
115	P	PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
116	P	PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
117	G	PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
118	I	PASS PASS PASS PASS PASS PASS PASS WT PASS FAIL PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	18	PASS
119	D	PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
120	F	PASS PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
121	G	PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
122	P	PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
123	A	WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
124	L	PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
125	L	PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
126	A	WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
127	D	PASS PASS WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
128	V	PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS	19	5.26%	19	PASS
129	Q	PASS PASS PASS PASS PASS PASS PASS PASS FAIL PASS PASS PASS PASS WT PASS PASS PASS PASS PASS PASS	19	5.26%	18	PASS
130	V	PASS PASS PASS PASS PASS FAIL PASS PASS PASS PASS PASS PASS FAIL FAIL PASS PASS PASS WT PASS PASS	19	5.26%	16	PASS
131	A	WT PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS	19	5.26%	19	PASS
132	V	PASS PASS PASS PASS PASS PASS FAIL PASS PASS PASSV_COIPASS PASS FAIL PASS PASS PASS WT V_COIPASS	17	5.88%	15	PASS
133	S	PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS PASS PASS	19	5.26%	19	PASS
134	S	PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS PASS WT PASS PASS PASS PASS	19	5.26%	19	PASS

Table C.3: Rational Saturation Library Quality Report from Twist Bioscience.

Mutant Library	Colony ID	Medium	[DCS] (µg/mL)	original codon	codon start	codon end	new codon	originial aa	aa Position	new aa	Mutation
								Y	**388**	**C**	
				CCG	1054	1056	CCA	P	352	P	s
epPCR	m14.10	A	90	TAC	1162	1164	TGC	**Y**	**388**	**C**	ns
			95	CTG	1144	1146	CTC	L	382	L	s
epPCR	m15.04		95	TAC	1162	1164	TGC	**Y**	**388**	**C**	ns
epPCR	m15.12	A	95	TAC	1162	1164	TGC	**Y**	**388**	**C**	ns
			95	GGT	25	27	AGT	G	9	S	ns
			95	ACC	316	318	AGC	T	106	S	ns
epPCR	m15.13	A	95	TAC	1162	1164	TGC	**Y**	**388**	**C**	ns
			95	TCA	85	87	ACA	S	29	T	ns
epPCR	m15.17	A	95	TAC	1162	1164	TGC	**Y**	**388**	**C**	ns
			95	GCG	634	636	TCG	A	212	S	ns
			95	GCG	1072	1074	GTG	A	358	V	ns
epPCR	m15.20	A	95	TAC	1162	1164	TGC	**Y**	**388**	**C**	ns
epPCR	m6.03	A	50	TAC	1162	1164	TGC	**Y**	**388**	**C**	ns
			60	CTG	748	750	TTG	L	250	L	s
			60	ACA	892	894	AGA	T	298	R	ns
			65	TAC	1162	1164	TGC	**Y**	**388**	**C**	ns
epPCR	m8.01	A	60	TAC	1162	1164	TGC	**Y**	**388**	**C**	ns
			50	GGC	310	312	GGT	G	104	G	s
epPCR	col 079		50	TAC	1162	1164	TGC	**Y**	**388**	**C**	ns
Rational	col 128	B	50	TAC	1162	1164	TGC	**Y**	**388**	**C**	ns
Rational	col 127	B	50	TAC	1162	1164	TGC	**Y**	**388**	**C**	ns
Rational	col 055	B	50	TAC	1162	1164	TGC	**Y**	**388**	**C**	ns
Rational	col 117	B	50	TAC	1162	1164	TGC	**Y**	**388**	**C**	ns
epPCR	col 106	B	50	TAC	1162	1164	TGC	**Y**	**388**	**C**	ns
								Y	**388**	**D**	
			60	GAT	379	381	GAC	D	127	D	s
			60	GGC	949	951	GGT	G	317	G	s
epPCR	m8.06	A	60	TAC	1162	1164	GAC	**Y**	**388**	**D**	ns
			50	GTA	628	630	ATA	V	210	I	ns
epPCR	col 156	B	50	TAC	1162	1164	GAC	**Y**	**388**	**D**	ns
Rational	col 014	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
Rational	col 015	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
			50	GTG	1042	1044	GTT	V	348	V	s
Rational	col 016	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
Rational	col 023	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
Rational	col 035	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
			50	GTG	1042	1044	GTT	V	348	V	s
Rational	col 039	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
			50	GTG	1042	1044	GTT	V	348	V	s
Rational	col 040	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
			50	GTG	1042	1044	GTT	V	348	V	s
Rational	col 041	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
			50	GTG	1042	1044	GTT	V	348	V	s

Table C.4: Overview of library strains with non-synonymous mutations at positions amino acid position Y388, P11, L35, P311 within Alr_{Mtb}. A = LB medium + 100 mM L-arabinose, B = M9 0.4% Fructose medium + 10 mM L-arabinose. s = synonymous and ns = non-synonymous

Rational	col 044	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
Rational	col 048	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
			50	GTG	1042	1044	GTT	V	348	V	s
Rational	col 049	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
Rational	col 051	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
			50	GTG	1042	1044	GTT	V	348	V	s
Rational	col 052	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
Rational	col 053	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
Rational	col 054	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
Rational	col 056	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
Rational	col 059	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
Rational	col 062	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
			50	CCT	31	33	CCC	P	11	P	s
Rational	col 063	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
			50	GTG	1042	1044	GTT	V	348	V	s
Rational	col 069	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
			50	GTG	1042	1044	GTT	V	348	V	s
Rational	col 072	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
			50	GTG	1042	1044	GTT	V	348	V	s
Rational	col 075	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
			50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
			50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
			50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
Rational	col 083	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
Rational	col 084	B	50	GTG	1042	1044	GTT	V	348	V	s
Rational	col 084	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
Rational	col 089	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
Rational	col 091	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
Rational	col 101	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
			50	GTC	1042	1044	GTT	V	348	V	s
Rational	col 103	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
Rational	col 104	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
Rational	col 107	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
Rational	col 109	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
			50	GTC	1042	1044	GTT	V	348	V	s
Rational	col 121	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
Rational	col 123	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
			50	GTC	1042	1044	GTT	V	348	V	s
Rational	col 124	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
Rational	col 125	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
			50	CCT	31	33	CCC	P	11	P	s
Rational	col 063	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
			50	GTC	1042	1044	GTT	V	348	V	s
Rational	col 129	B	50	TAC	1162	1164	GAT	**Y**	**388**	**D**	ns
								Y	**388**	**N**	
Rational	col 116	B	50	TAC	1162	1164	AAC	**Y**	**388**	**N**	ns
								Y	**388**	**S**	
Rational	col 029	B	50	TAC	1162	1164	AGC	**Y**	**388**	**S**	ns

Table C.4: Overview of library strains with non-synonymous mutations at positions amino acid position Y388, P11, L35, P311 within Alr_{Mtb}. A = LB medium + 100 mM L-arabinose, B = M9 0.4% Fructose medium + 10 mM L-arabinose. s = synonymous and ns = non-synonymous

Rational	col 074	B	50	TAC	1162	1164	AGC	**Y**	**388**	**S**	ns
								Y	**388**	**P**	
			50	GAT	379	381	AAC	D	127	N	ns
Rational	col 050	B	50	TAC	1162	1164	CCG	**Y**	**388**	**P**	ns
Rational	col 097	B	50	TAC	1162	1164	CCG	**Y**	**388**	**P**	ns
			50	GAC	379	381	AAC	D	127	N	ns
Rational	col 082	B	50	TAC	1162	1164	CCG	**Y**	**388**	**P**	ns
								P	**11**	**H**	
			60	CCT	31	33	CAT	**P**	**11**	**H**	ns
			60	GGC	1057	1059	GGG	G	353	G	s
epPCR	m8.07	A	60	CCG	1102	1104	TCG	P	368	S	ns
								P	**11**	**S**	
epPCR	col 092	B	50	CCT	31	33	TCT	**P**	**11**	**S**	ns
								P	**11**	**Q**	
Rational	col 010	B	50	CCT	31	33	CAG	**P**	**11**	**Q**	ns
								L	**35**	**Q**	
			90	CTG	103	105	CAG	**L**	**35**	**Q**	ns
epPCR	m14.08	A	90	GAA	286	288	GAG	E	96	E	s
			95	CTG	103	105	CAG	**L**	**35**	**Q**	ns
epPCR	m15.14	A	95	ACC	433	435	ACT	T	145	T	s
								L	**35**	**P**	
epPCR	m15.06	A	95	CTG	103	105	CCG	L	35	P	ns
			60	GTG	22	24	ATG	V	8	M	ns
			60	CTG	103	105	CCG	**L**	**35**	**P**	ns
epPCR	m8.04	A	60	CTT	292	294	CTA	L	98	L	s
								L	**35**	**R**	
			50	CTG	103	105	CGG	**L**	**35**	**R**	ns
epPCR	col 071	B	50	GGG	1150	1152	GGT	G	384	G	s
								L	**35**	**S**	
Rational	col 025	B	50	CTG	103	105	AGC	**L**	**35**	**S**	ns
Rational	col 057	B	50	CTG	103	105	AGC	**L**	**35**	**S**	ns
								L	**35**	**F**	
			50	CTG	103	105	TTT	**L**	**35**	**F**	ns
Rational	col 073	B	50	TGG	1135	1137	TTG	W	379	L	ns
								L	**35**	**D**	
			50	CTG	103	105	GAT	**L**	**35**	**D**	ns
Rational	col 081	B	50	ACC	1195	1197	AAA	T	399	K	ns
								L	**35**	**P**	
Rational	col 112	B	50	CTG	103	105	CCG	**L**	**35**	**P**	ns
								P	**311**	**S**	* / **
			90	GCG	922	924	GCC	A	308	A	s
epPCR	m14.01	A	90	CCG	931	933	TCG	**P**	**311**	**S**	ns
epPCR	m15.16	A	95	CCG	931	933	TCG	**P**	**311**	**S**	ns
								P	**311**	**R**	**
			95	GAT	415	417	GAA	D	139	E	ns
			95	GTC	769	771	CTC	V	257	L	ns
epPCR	m15.09	A	95	CCG	931	933	CGG	**P**	**311**	**R**	ns

Table C.4: Overview of library strains with non-synonymous mutations at positions amino acid position Y388, P11, L35, P311 within Alr_{Mtb}. A = LB medium + 100 mM L-arabinose, B = M9 0.4% Fructose medium + 10 mM L-arabinose. s = synonymous and ns = non-synonymous

References:

*** P311S was also published by** Chernyaeva, E. N., Shulgina, M. V., Rotkevich, M. S., Dobrynin, P. V., Simonov, S. A., Shitikov, E. A., et al. (2014). Genome-wide Mycobacterium tuberculosis variation (GMTV) database: a new tool for integrating sequence variations and epidemiology. *BMC Genomics*, *15*(1), 308.

**** Non-synonymous mutation at the same aa position P311L was cited in:**

Coll, F., Phelan, J., Hill-Cawthorne, G. A., Nair, M. B., Mallard, K., Ali, S., et al. (2018). Genome-wide analysis of multi- and extensively drug-resistant Mycobacterium tuberculosis. Nature Genetics, 50(2), 307–316.
Portelli, S., Phelan, J. E., Ascher, D. B., Clark, T. G., & Furnham, N. (2018). Understanding molecular consequences of putative drug resistant mutations in Mycobacterium tuberculosis. Scientific Reports, 8(1), 15356.

Table C.4: Overview of library strains with non-synonymous mutations at positions amino acid position Y388, P11, L35, P311 within $Alr_{Mtb.}$

Mutant Library	Colony ID	Medium	[DCS] (µg/mL)	original codon	codon start	codon end	new codon	originial aa	aa Position	new aa	Mutation
								V	**94**	**V**	
epPCR	col 041	B	50	GTG	280	282	GTA	V	94	V	s
				ACG	529	531	GCG	T	177	A	ns
				GCT	649	651	GAT	A	217	D	ns
epPCR	col 062	B	50	GTG	280	282	GTT	V	94	V	s
epPCR	col 112	B	50	GTG	280	282	GTT	V	94	V	s
				CTG	1144	1146	ATG	L	382	M	ns
Rational	col 114	B	50	GTG	280	282	GTT	V	94	V	s
				TAT	595	597	AAC	Y	199	N	ns
								L	**100**	**L**	
epPCR	m3.05	A	35	CTG	298	300	CTC	L	100	L	s
epPCR	col 021	B	50	CTG	298	300	CTA	L	100	L	s
				CGT	682	684	AGT	R	228	S	ns
epPCR	col 126	B	50	CTG	298	300	TTG	L	100	L	s
				ACC	316	318	AGC	T	106	S	ns
				ACC	451	453	ACT	T	151	T	s
epPCR	col 134	B	50	GTG	22	24	GTT	V	8	V	s
				CTG	298	300	CTA	L	100	L	s
				ACC	1123	1125	ACA	T	375	T	s
								L	**271**	**L**	
epPCR	col 040	B	50	CTG	811	813	CTT	L	271	L	s
epPCR	col 050	B	50	CTG	811	813	TTG	L	271	L	s
				GAT	1030	1032	AAT	D	344	N	ns
epPCR	col 063	B	50	CTG	811	813	TTG	L	271	L	s
				GGT	1009	1011	GGC	G	337	G	s
				GAT	1030	1032	AAT	D	344	N	ns
epPCR	col 064	B	50	CTG	811	813	TTG	L	271	L	s
				GGT	1009	1011	GGC	G	337	G	s
				GAT	1030	1032	AAT	D	344	N	ns
epPCR	col 141	B	50	GCG	520	522	GCA	A	174	A	s
				ACG	529	531	ACA	T	177	T	s
				CTG	811	813	TTG	L	271	L	s
				GCC	838	840	CCC	A	280	P	ns
								A	**308**	**A**	
epPCR	m14.01	A	90	GCG	922	924	GCC	A	308	A	s
			90	CCG	931	933	TCG	P	311	S	ns
epPCR	col 044	B	50	CTG	370	372	CTA	L	124	L	s
			50	GAT	613	615	GAA	D	205	E	ns
			50	GCG	922	924	GCA	A	308	A	s
epPCR	col 073	B	50	GCG	922	924	GCA	A	308	A	s
epPCR	col 145	B	50	GCG	922	924	GCA	A	308	A	s
								Q	**354**	**Q**	
epPCR	col 030	B	50	GAT	1030	1032	AAT	D	344	N	ns
			50	CAG	1033	1035	CAA	Q	345	Q	s
epPCR	col 031	B	50	GAT	1030	1032	AAT	D	344	N	ns
			50	CAG	1033	1035	CAA	Q	345	Q	s

Table C.5: Overview of mutations in library strains that carried mutations V94V, L100L, L271L, A308A, Q345Q, L266M and V348V.
A = LB medium + 100 mM L-arabinose, B = M9 0.4% Fructose medium + 10 mM L-arabinose. s = synonymous and ns = non-synonymous

epPCR	col 042	B	50	GAT	1030	1032	AAT	D	344	N	ns
			50	CAG	1033	1035	CAA	Q	345	Q	s
epPCR	col 067	B	50	GAT	1030	1032	AAT	D	344	N	ns
			50	CAG	1033	1035	CAA	Q	345	Q	s
								L	**266**	**M**	
epPCR	m8.10	A	60	CTG	796	798	ATG	L	266	M	ns
			60	GAT	1030	1032	AAT	D	344	N	ns
			60	CGC	1207	1209	CGT	R	403	R	s
epPCR	col 043	B	50	CTG	796	798	ATG	L	266	M	ns
			50	GAT	1030	1032	AAT	D	344	N	ns
epPCR	col 054	B	50	CTG	796	798	ATG	L	266	M	ns
			50	GAT	1030	1032	AAT	D	344	N	ns
			50	CGC	1207	1209	CGT	R	403	R	s
epPCR	col 055	B	50	CTG	796	798	ATG	L	266	M	ns
			50	GAT	1030	1032	AAT	D	344	N	ns
			50	CGC	1207	1209	CGT	R	403	R	s
								V	**348**	**V**	
Rational	col 016	B	50	GTG	1042	1044	GTT	V	348	V	s
				TAC	1162	1164	GAT	Y	388	D	ns
Rational	col 039	B	50	GTG	1042	1044	GTT	V	348	V	s
				TAC	1162	1164	GAT	Y	388	D	ns
Rational	col 040	B	50	GTG	1042	1044	GTT	V	348	V	s
				TAC	1162	1164	GAT	Y	388	D	ns
Rational	col 041	B	50	GTG	1042	1044	GTT	V	348	V	s
				TAC	1162	1164	GAT	Y	388	D	ns
Rational	col 044	B	50	GTG	1042	1044	GTT	V	348	V	s
				TAC	1162	1164	GAT	Y	388	D	ns
Rational	col 049	B	50	GTG	1042	1044	GTT	V	348	V	s
				TAC	1162	1164	GAT	Y	388	D	ns
Rational	col 052	B	50	GTG	1042	1044	GTT	V	348	V	s
				TAC	1162	1164	GAT	Y	388	D	ns
Rational	col 069	B	50	GTG	1042	1044	GTT	V	348	V	s
				TAC	1162	1164	GAT	Y	388	D	ns
Rational	col 072	B	50	GTG	1042	1044	GTT	V	348	V	s
				TAC	1162	1164	GAT	Y	388	D	ns
Rational	col 075	B	50	GTG	1042	1044	GTT	V	348	V	s
				TAC	1162	1164	GAT	Y	388	D	ns
Rational	col 084	B	50	GTG	1042	1044	GTT	V	348	V	s
				TAC	1162	1164	GAT	Y	388	D	ns
Rational	col 103	B	50	GTC	1042	1044	GTT	V	348	V	s
				TAC	1162	1164	GAT	Y	388	D	ns
Rational	col 121	B	50	GTC	1042	1044	GTT	V	348	V	s
				TAC	1162	1164	GAT	Y	388	D	ns
Rational	col 124	B	50	GTC	1042	1044	GTT	V	348	V	s
				TAC	1162	1164	GAT	Y	388	D	ns
Rational	col 129	B	50	GTC	1042	1044	GTT	V	348	V	s
				TAC	1162	1164	GAT	Y	388	D	ns

Table C.5: Overview of mutations in library strains that carried mutations V94V, L100L, L271L, A308A, Q345Q, L266M and V348V.
A = LB medium + 100 mM L-arabinose, B = M9 0.4% Fructose medium + 10 mM L-arabinose. s = synonymous and ns = non-synonymous

Escherichia coli strains	Relevant characteristic(s)	Description / Experimental Purpose	Source or reference
E. coli DH5α	*endA1 hsdR17 supE44 recA1 relA1* (*lacZYA-argF*)	Plasmid amplification	NEB
BW25113	Δ(*araD-araB*)567 Δ*lacZ4787*(::*rrnB-3*) λ, rph-1 Δ(*rhaD-rhaB*)568 hsdR514	Parent strain	(Datsenko & Wanner, 2000)
JW0063	BW25113 $\Delta araC_{Ec}$: :Kan^r		(Datsenko & Wanner, 2000)
D20	JW0063 $\Delta araC_{Ec}$: :frt		See Chapter II
D18	BW25113 $\Delta araC_{Ec}$: :frt, $\Delta dadX_{Ec}$: :frt, Δalr_{Ec} : :frt		See Chapter II
TS008	D18 + V91 + pTS008	TESEC strain with wild type Alr_{Mtb} expression	See Chapter II
TS072	D18 + V91	Background strain for the construction of mutant libraries	This study
MUT.0001	D18 + pMUT-alr-001 + V91	Characterization of Alr_{Mtb} mutants associated with DCS resistance	This study
MUT.0002	D18 + pMUT-alr-004 + V91	Characterization of Alr_{Mtb} mutants associated with DCS resistance	This study
MUT.0003	D18 + pMUT-alr-006 + V91	Characterization of Alr_{Mtb} mutants associated with DCS resistance	This study
MUT.0005	D18 + pMUT-alr-002 + V91	Characterization of Alr_{Mtb} mutants associated with DCS resistance	This study
MUT.0006	D18 + pMUT-alr-003 + V91	Characterization of Alr_{Mtb} mutants associated with DCS resistance	This study
MUT.0007	D18 + pMUT-alr-005 + V91	Characterization of Alr_{Mtb} mutants associated with DCS resistance	This study
MUT.0048	D18 + pMUT-alr-008 + V91	Characterization of Alr_{Mtb} mutants associated with DCS resistance	This study
MUT.0049	D18 + pMUT-alr-009 + V91	Characterization of Alr_{Mtb} mutants associated with DCS resistance	This study
MUT.0050	D18 + pMUT-alr-010 + V91	Characterization of Alr_{Mtb} mutants associated with DCS resistance	This study
MUT.0051	D18 + pMUT-alr-011 + V91	Characterization of Alr_{Mtb} mutants associated with DCS resistance	This study

Table C.6: List of strains.

MUT.0080	D18 + pMUT-alr-040 + V91	Characterization of Alr_{Mtb} mutants associated with DCS resistance	This study
MUT.0090	D18 + pMUT-alr-049 + V91	Characterization of Alr_{Mtb} mutants associated with DCS resistance	This study
MUT.0103	D18 + pMUT-alr-062 + V91	Characterization of Alr_{Mtb} mutants associated with DCS resistance	This study

Mutant Libraries	**Relevant characteristic(s)**	**Experimental purpose**	
epPCR Alr_{Mtb} mutant library	D18 with plasmids: V91 + pTS000 derived plasmid library carrying epPCR gene fragments of alr_{Mtb}	Screening for DCS resistance mutations within alr_{Mtb}	This study
Ration Saturation Alr_{Mtb} mutant library	D18 with plasmids: V91 + pTS000 derived plasmid library carrying synthesized gene fragments of alr_{Mtb}	Screening for DCS resistance mutations within alr_{Mtb}	This study

Table C.6: List of strains.

Plasmid	Relevant characteristic(s)	Antibiotic resistance marker	Source / Reference
V91	pBAD/araC	Kan^r	See Chapter II
pTS000	P37.pRD131 derived plasmid	Amp^r	See Chapter II
pTS008	pTS000 derived, pBAD/lacZ :: alr_{Mtb}	Amp^r	See Chapter II
pMUT-alr-001	pTS000 derived , pBAD/lacZ :: alr_{Mtb} A308G	Amp^r	This study
pMUT-alr-002	pTS000 derived, pBAD/ lacZ :: alr_{Mtb} M343T	Amp^r	This study
pMUT-alr-003	pTS000 derived, pBAD/ lacZ :: alr_{Mtb} L113R	Amp^r	This study
pMUT-alr-004	pTS000 derived, pBAD/ lacZ :: alr_{Mtb} M334T	Amp^r	This study
pMUT-alr-005	pTS000 derived, pBAD/ lacZ :: alr_{Mtb} R397G	Amp^r	This study
pMUT-alr-006	pTS000 derived, pBAD/ lacZ :: alr_{Mtb} Y388D	Amp^r	This study
pMUT-alr-008	pTS000 derived, pBAD/ lacZ :: alr_{Mtb} Q30P	Amp^r	This study
pMUT-alr-009	pTS000 derived, pBAD/ lacZ :: alr_{Mtb} L250M	Amp^r	This study
pMUT-alr-010	pTS000 derived, pBAD/ lacZ :: alr_{Mtb} T385N	Amp^r	This study
pMUT-alr-011	pTS000 derived, pBAD/ lacZ :: alr_{Mtb} T14R	Amp^r	This study
pMUT-alr-040	pTS000 derived, pBAD/ lacZ :: alr_{Mtb} S22T	Amp^r	This study
pMUT-alr-049	pTS000 derived, pBAD/ lacZ :: alr_{Mtb} Y388C	Amp^r	This study
pMUT-alr-062	pTS000 derived, pBAD/ lacZ :: alr_{Mtb} D139E, V257L, P311R	Amp^r	This study

Table C.7: List of plasmids.