INVESTIGATION OF ANTIBIOTIC TARGETS IN THE DECAPRENYL-PHOSPHORYLARABINOSE BIOSYNTHESIS IN *M. TUBERCULOSIS*

by

SZILVIA TÓTH

a thesis submitted to

University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

School of Biosciences

College of Life and Environmental Sciences

University of Birmingham

July 2018

Declaration

The work presented in this thesis is original except where citing relevant references. Studies were conducted from 2014 to 2017 in the School of Biosciences, University of Birmingham, B15 2TT, UK and also in research and development facility of GlaxoSmithKline, laboratory for Diseases of the Developing World, Tres Cantos, Spain.

No part of this work has been submitted for a degree or a diploma to this or any other university.

Abstract

An estimated 1.67 million people died of tuberculosis (TB) in 2016 and it is a threat to human life on a global-scale. To shorten current treatments and battle drug resistant strains it is important to discover and develop new drugs against the causative agent, *Mycobacterium tuberculosis*.

Phenotypic screens have delivered potent hit and lead molecules in the past but the need to target new pathways in *M. tuberculosis* metabolism to circumvent pre-existing drug resistance mechanisms. The decaprenyl-phosphorylarabinose (DPA) biosynthetic pathway produces irreplaceable building blocks of the unique mycobacterial cell wall structure. Decaprenyl-phosphorylribose (DPR) Epimerase sub-units 1 and 2 (DprE1 and DprE2) are essential enzymes of this pathway. Whilst DprE1 has received much attention, however DprE2 has been overlooked as a potential drug target.

In this thesis the role of a small disordered region of DprE1 in ligand binding and physiological function has been investigated. This region appears to interact with the substrate and based on enzyme activity assays, physiological importance of specific residues was found to be in line with the level of conservation between species.

In addition, a high-throughput screen against a DprE2 overexpressing *M. bovis* BCG strain has been utilised to discover new hit compounds against this target, coupled with the development of an enzymatic assay to evaluate specific DprE2 inhibition. The screen resulted nitrofuran hit compounds that probably act as prodrugs, which was indicated by spontaneous resistant mutants resembling the mutations necessary to activate new TB pro-drug pretomanid.

Acknowledgement

I would like to take this opportunity to thank all the people who helped me in completing this work and overcome obstacles along the way.

Firstly, I am most grateful to my supervisors, Prof. Gurdyal Besra and Dr. Klaus Fütterer, who guided me in these past four years and are role models to look up to in science research. I am thankful for their global influence on my professional development as well as for every small constructive criticism that made this thesis better.

I am thankful for my colleagues that helped my laboratory work in two different locations. The members of the Besra, Alderwick and Bhatt laboratories in University of Birmingham have always been supportive and a pleasure to work with. In particular, I would like to thank the support of Sarah M. Batt, Natacha Veerapen, Patrick Moynihan, Katherine Abrahams, Usha Veeraraghavan and Cristian Varela postdoctoral members of the lab who were knowledgeable and patient to answer my questions at all times. My fellow PhD students at the time, Monika Jankute, Panchali Kanvatirth, Asma Javid have also been always kind and a pleasure to work with.

Regarding my Spanish research placement, I am honestly grateful to the people of GSK Tres Cantos, our coordinator Joel Lelievre, my mentors Esther Perez-Herran and Beatriz Rodriguez-Miquel and the TB and Malaria research teams. Also, to Open Lab fellows Shipra Grover, Kaj Kreutzfeldt, Cynthia Tallant and Hugo Fraga who were once more very helpful in laboratory work and troubleshooting.

Finally, I am happy that as one of four Marie Sklodowska-Curie fellows of the project CooperaTB I could work along Giacomo Chiodarelli, Bogdan Duma and Ramon Soto Garcia. This journey would have not been the same without your company.

Table of contents

List of figures

List of tables

List of abbreviations

%inh	percentage of inhibition
ADC	albumin - dextrose - catalase supplement
Aft	arabinofuranosyl-transferase
AG	arabinogalactan
BA	benzyl alcohol
B_{max}	overall binding capacity
Brij-35	ethoxylated lauryl alcohol (detergent)
BTP	bis-tris propane
CHAPS	(3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonat
ctrl	control
Da	Dalton (unit)
DCPIP	dichlorophenolindophenol
DEAE sepharose	diethylaminoethanol sepharose
DMSO	dimethyl sulfoxide
dNTP	deoxynucleoside triphosphate mixture
DPA	decaprenyl-β-D-arabinose
DPR	decaprenyl-β-D-phosphorylribose
DprE	decaprenyl-phosphorylribose epimerase
DPX	decaprenylphosphoryl-2-keto-β-D-erythropentofuranose
EDTA	ethylenediamine tetraacetic acid
ETH	ethionamide
FAD	flavin adenine dinucleotide (oxidised)
$FADH_2$	flavin adenine dinucleotide (reduced)
FF	fast flow (column)
FPR	farnesyl-phosphorylribose
GGPP	geranylgeranyl-pyrophosphate
GGPR	geranylgeranyl-phosphorylribose
GSK	GlaxoSmithKline

GtrA	Rv3789 protein
h	Hill-coefficent
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
IEX	ion exchange chromatography
IMAC	immobilised metal affinity chromatography
INH	isoniazid
IPTG	isopropyl-β-D-thiogalactoside
K_{av}	apparent distribution coefficient
K_d	dissociation constant
kDa	kilodalton
LAM	lipoarabinomannan
LB	Luria-Bertani
MBP	maltose binding protein
MIC	minimal inhibitory concentration
Middlebrook 7H10	mycobacterial solid medium (7H9 + agar)
Middlebrook 7H9	mycobacterial liquid medium
MilliQ water	ultrapure deionised water made by Millipore system
MM	molecule marker
Mt-DprE1	DprE1 of *M. tuberculosis* (Rv3790)
Mt-DprE2	DprE2 of *M. tuberculosis* (Rv3791)
Mt-GtrA	GtrA of *M. tuberculosis* (Rv3789)
MTB	*Mycobacterium tuberculosis*
MWCO	molecular weight cut-off
NaCl	sodium chloride
NAD^+	nicotinamide adenine dinucleotide (oxidised)
NADH	nicotinamide adenine dinucleotide (reduced)
$NADP^+$	nicotinamide adenine dinucleotide phosphate (oxidised)
NADPH	nicotinamide adenine dinucleotide phosphate (reduced)
Ni-NTA	Nickel-nitrilotriacetic acid
OADC	oleic acid - dextrose - catalase supplement
oe	overexpressor

PCR	polymerase chain reaction
PDB	Protein Data Bank (website)
PEG-400	poly(ethylene glycol), 400 monomer/polimer
PG	peptidoglycan
PZA	pyrazinamide
R square	coefficient of determination
REMA	resazurin microplate assay protocol
RFU	relative fluorescence unit
RIF	rifampicin
rpm	revolutions per minute
SD	standard deviation
SDS-PAGE	sodium dodecyl sulfate polyakrylamide gel electrophoresis
SEC	size exclusion chromatorgaphy
SPOTi	spot culture growth inhibition
SSE	sum of squared errors
SUMO	small ubiquitin-like modifier protein
TB	tuberculosis
TCEP	Tris(2-carboxyethyl)phosphine hydrochloride
Tm	melting temperature
TP	total protein
Tris-HCl	Tris(hydroxymethyl)aminomethane hydrochloride
TSA	thermal shift assay = differential scanning fluorimetry
Tween-80	polyoxyethylenesorbitan monooleate
v/v %	volume per volume percentage
V_{void}	void volume
w/v %	weight per volume percentage
wt	wild type
λ_{max}	fluorescence emission peak

1 Introduction

1.1 Tuberculosis disease

1.1.1 Current global status

Tuberculosis (TB) is an infectious disease caused by *Mycobacterium tuberculosis* (MTB) and is the leading cause of death by a single infectious agent in the world (WHO, 2017). The incidence of tuberculosis was estimated to be 10.4 million cases globally in 2016, most of which were from South-East Asia and Africa (WHO, 2017). Current treatment failure rates are estimated to be around 17%, which can increase by infection with drug-resistant strains of MTB. Tuberculosis infection in 2016 resulted in over 1.67 million deaths, with TB now considered within the Sustainable Development Goals developed by the United Nations and specifically addressed in the End TB Strategy by the World Health Organisation (WHO, 2017).

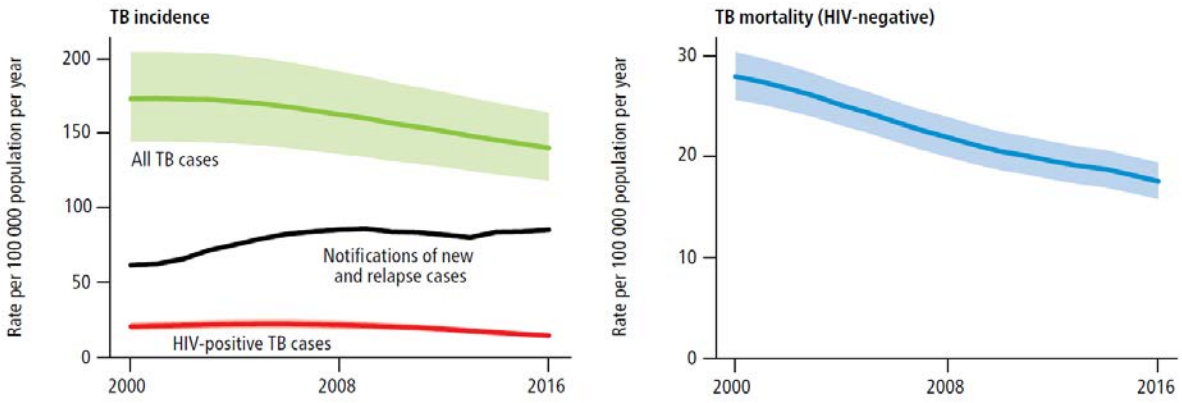


Figure 1.1. Global trends in TB incidence and mortality rates in recent years. (WHO, 2017)). Shaded areas represent uncertainty intervals.

Owing to global coordinated efforts, tuberculosis incidence is on the decline (see **Figure 1.1.**). Current initiatives are focusing on speeding up this trend, among other means, by improving prevention, diagnosis, increasing financial contribution and resources, setting up a monitoring framework, addressing HIV co-infection, and specifically preventing an epidemic of drug-resistant TB (WHO, 2017).

Drug-resistance of pathogenic bacteria is a worldwide and pressing matter. Inadequate drug treatment of tuberculosis has given rise to resistance against first-line drugs in tuberculosis therapy and led to multi-drug resistant (MDR) and extensively drug resistant (XDR) strains. MDR-TB refers to infections by strains that are isoniazid (INH) and rifampicin (RIF) resistant, while XDR-TB strains are resistant to both INH and RIF, and in addition resistant to a fluoroquinolone (like moxifloxacin), and at least on second-line injectable agent (Chan E.D. and Iseman M.D., 2008). In 2016, approximately 600 000 new cases of TB were RIF resistant and 490 000 cases were of MDR-TB. Approximately, 78% of global drug-resistant cases are estimated to never receive the required care and specific drug regimen (Chan E.D. and Iseman M.D., 2008).

Drug resistant tuberculosis consequently has a worse treatment outcome than drug sensitive strains. The current treatment against MDR and XDR-TB with second-line drugs has success rate of 52% for MDR-TB, and 28% for XDR-TB as determined by latest cohort data from 2013 by the World Health Organisation. A key issue is that the treatment is more time consuming.

Clinically, another complication within the last 25 years is the frequency of TB co-infection with the human immunodeficiency virus (HIV), which requires special patient management (WHO, 2017). In immunocompromised patients with HIV infection a latent TB infection is more likely to progress into an active TB infection (Getahun et al., 2015). Furthermore, drug

interactions and toxicity are even more severe due to simultaneously administered TB drugs and anti-retroviral therapy.

Another important comorbidity is caused by diabetes mellitus. Diabetes is listed among risk factors by the WHO in tuberculosis infections, along with unhealthy alcohol consumption, smoking and undernourishment (WHO, 2017). Diabetes is associated with increased risk of tuberculosis (Jeon and Murray, 2008). The interconnection of the two illnesses is not completely understood, but as a general expectation the increase in living standards and changes in dietary habits, the number of diabetics will rise in the future enhancing tuberculosis susceptibility on a global population level (Jeon et al., 2012).

1.1.2 Historical background of TB

Tuberculosis is an ancient disease. Based on tracking of genetic mutation markers the most recent common ancestor of the *Mycobacterium tuberculosis* complex (MTBC) emerged about 70,000 years ago in the upper paleolithic age in North-East Africa (Comas et al., 2013). The group of MTBC species share an exceptional genetic similarity at the nucleotide level and consist of eight obligate pathogens (Comas et al., 2013). MTB, *Mycobacterium africanum* and *Mycobacterium canettii* prefer a human host organism, while others like *M. bovis* prefer an animal host, however direct contact with these animals can transfer them to humans and develop symptoms (Kazda et al., 2009).

According to the most recent evidence MTBC species evolved parallel to the human host and followed early human migration over the world (Hershkovitz et al., 2015). Furthermore, unlike HIV or Ebola, these human mycobacterial pathogens are not of zoonotic origin,

instead the animal pathogens, such as *Mycobacterium bovis*, delineated at a later point from ancestral bacteria than MTB (Smith et al., 2009).

Many mycobacteria related to the MTBC are fast-growing saprophytes that are only opportunistic human pathogens, however, the slow-growing sub-cluster within the genus that includes MTBC, *Mycobacterium leprae* and *Mycobacterium ulcerans* are the etiological agents of leprosy and Buruli ulcer, respectively (Smith et al., 2009).

Beyond phylogenomic analysis and predictions there is archaeological evidence of MTB infections in ancient times. The oldest human remains with definitive tuberculosis infection by MTBC, possibly MTB is dated 6200-5500 BC (Hershkovitz et al., 2015). Signs of tuberculosis infections have been found in Egyptian mummies and written about in ancient Sanskrit and Chinese medical literature (Galagan, 2014). Hippocrates used the term phthisis meaning 'wasting away' in the 5th century BC (Galagan, 2014), which echoes with the later name of the disease, consumption, common symptom of weight loss during tuberculosis. In short, TB infections have been part of history for a long time.

During the industrial revolution, tuberculosis was followed by the expansion of global population movements and became a pandemic. By the end of the 18th century it was responsible for one quarter of all fatalities in the UK according to Parish registers (Davis, 2000). Diagnosis of tuberculosis was limited to pulmonary tuberculosis in most cases and often considered hereditary due to entire families being infected (Davis, 2000).

Robert Koch, a Microbiologist and later Nobel Laurate published in 1882 that MTB was the causative agent of tuberculosis (Koch, 1884). In this article, he wrote that every seventh person in Berlin died of tuberculosis at the time, and not enough attention was given to the disease.

„If the importance of a disease for mankind is measured from the number of fatalities which are due to it, then tuberculosis must be considered much more important than those most feared infectious diseases, plague, cholera, and the like.” From re-print of the 1882 article.

Based on the work of Jean-Antoine Villemin, who demonstrated transmissibility in rabbits and between human and rabbit, it became evident that tuberculosis is contagious (Villemin, 2015). These results led Koch to identify the microbe and to write a list of criteria applied for general pathogenic bacterial identification called Koch's postulates (Koch, 1884).

Koch also developed the tuberculin detection reagent, which is a skin test to detect people infected by tuberculosis (Koch, 1890)(WHO, 2017). Moreover, his visualisation technique was improved by Ziehl and Neelsen and became standard for acid fast bacilli staining (Neelsen, 1892). These tests are over 100 years old and yet are still the first-line diagnostic tools in use today (WHO, 2017).

Following the realisation that tuberculosis is a contagious disease, isolation of patients into sanatoria became common practice for those who could afford treatment (Davis, 2000). Aside from the beneficial impact of fresh air, rest and sufficient nutrition on the patient's health, isolation prevented further contact and chance of transmitting the disease. The tuberculosis epidemic was mitigated before Koch and antibiotics by improvements in sanitation and living standards in the 19th century (Fairchild and Oppenheimer, 1998). Addressing poverty and malnutrition are historically effective methods in lowering tuberculosis burden (Fairchild and Oppenheimer, 1998).

Diagnostics for tuberculosis improved at the turn of the century. The sputum smear test with Ziehl-Neelsen staining however practical, it is mostly reliable in advanced stages of TB infection. A positive sputum smear test indicates 70% mortality if patients are not treated (Tiemersma et al., 2011). Therefore, the development of medical radiography in the early

1900s was an important new tool in non-invasive tuberculosis diagnosis. X-rays, since their discovery in 1895 by Wilhelm Röntgen, have been considered for clinical use and by the middle of the century mass-screening of the population became routine test for tuberculosis (Davis, 2000).

The BCG (Bacillus Calmette-Guérin) vaccine was introduced in France in 1921 for the immunisation of children (Davis, 2000). BCG is an attenuated variant of *M. bovis* which was continuously passaged for 13 years on bile potato to decrease virulence (Calmette, 1931). Mass-vaccination with BCG is in place since the 1930s and in non-randomised human studies efficacy varies between 0-80% (Andersen and Doherty, 2005).

The microbiologist Paul Ehrlich at the end of the 19th century postulated the existence of compounds, a 'magic bullet', that would kill a microorganism without harming the human host (Murray et al., 2015). He synthesised arsphenamine (Salvarsan) in 1910, the first drug active against syphilis, later in 1929 penicillin was discovered by Alexander Fleming (Fleming, 1929). However, none of these compounds were potent against MTB.

Anti-mycobacterial drugs were discovered in the 1940s and the following three decades (see **Figure 1.2.** for chronology). Most of the early antitubercular drugs (e.g. streptomycin, STM) were discovered using animal models infected with the *M. tuberculosis* (Murray et al., 2015). Aside from the cost and ethical issue, animal studies were very efficient in discovering drugs like pyrazinamide (PZA), a potent drug *in vivo*, whose efficacy cannot be easily reproduced *in vitro* and today most likely it would evade discovery with current drug discovery strategies (Zhang and Mitchison, 2003).

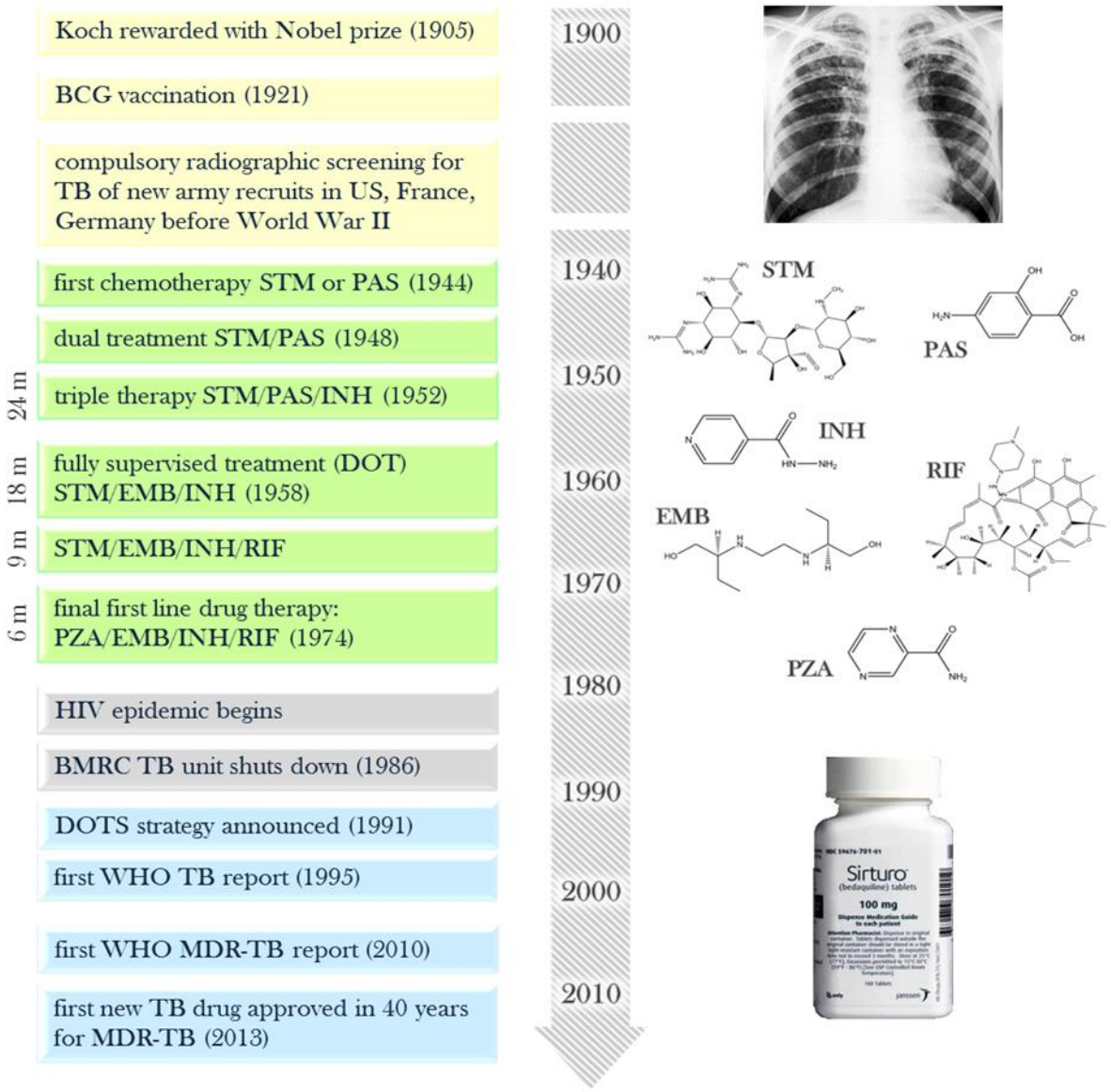


Figure 1.2. Important events in TB drug discovery in the 20th century. Yellow boxes represent pre-chemotherapy era, green boxes stand for first-line drug development, blue boxes display the re-emergence of TB. Next to therapy descriptions the length of treatment is indicated in months (m). DOT is directly observed therapy, DOTS directly observed therapy short-course, BMRC is British Medical Research Council. Images show a chest X-ray of a TB patient (with permission from Radiological Society of North America); molecule structures of the first-line drugs, and bedaquilin (Sirturo), the first Federal Drug Administration-approved TB drug in 40 years. Dates from (Murray et al., 2015) and (Iseman, 2002).

Historically, anti-tuberculosis drugs have been discovered by different approaches. STM and RIF discovery was based on the idea that soil microorganisms other than *Penicillium* moulds could also use antibiotic compounds in their natural environment to compete with other

bacteria and fungi (Murray et al., 2015). Meanwhile, *para*-aminosalicylic acid (PAS) was hypothesised to be a toxic alternative substrate for MTB, based on structural similarity to salicylic acid (Lehmann, 1946). In the case of PZA, it was observed that vitamin B3 inhibited MTB growth. Modification of nicotinamide (amide derivative of niacin, vitamin B3) lead to the discovery of PZA, then INH and ethionamide (ETH), structures illustrated in **Figure 1.3.** (Zhang and Mitchison, 2003).

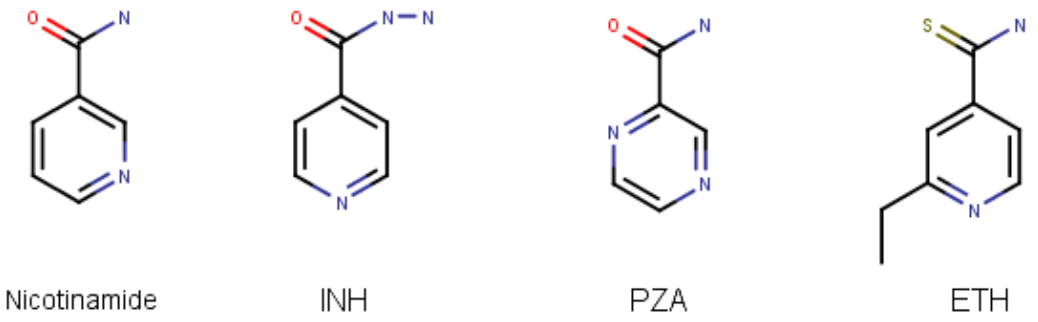


Figure 1.3. Structures of related first and second-line antituberculotic drugs. Nicotinamide is the amide of vitamin B3, INH is isoniazide, PZA is pyrazinamide, ETH is ethionamide.

1.1.3 Pathobiology of tuberculosis with implications to drug development

TB is an infectious disease with special pathological characteristics which affect drug development. The most notable features are dormancy, emerging resistance against existing drugs and the unique barriers of the microenvironment that limit drug access to the target. During drug development we need to assess compound activity against all of these distinctive populations, dormant, resistant, secluded populations in order to forecast precisely compound efficacy.

1.1.3.1 Latent TB infection

Tuberculosis infections remain latent in most cases (see **Figure 1.4.**) (WHO, 2017), which causes no symptoms and tends to be sub-clinical. According to the WHO, approximately one third of the global population has latent tuberculosis (WHO, 2017). It can reactivate with a probability of 5-15% later in life thus developing secondary tuberculosis (Getahun et al., 2015). The probability of reactivation depends mostly on elapsed time. Reactivation events which happen within the first five years following primary infection account for 85% of all reactivations, and 95% develop within the first ten years (Salgame et al., 2015). Besides reactivation, secondary tuberculosis may develop by reinfection as well, especially in high tuberculosis burden environments (Salgame et al., 2015).

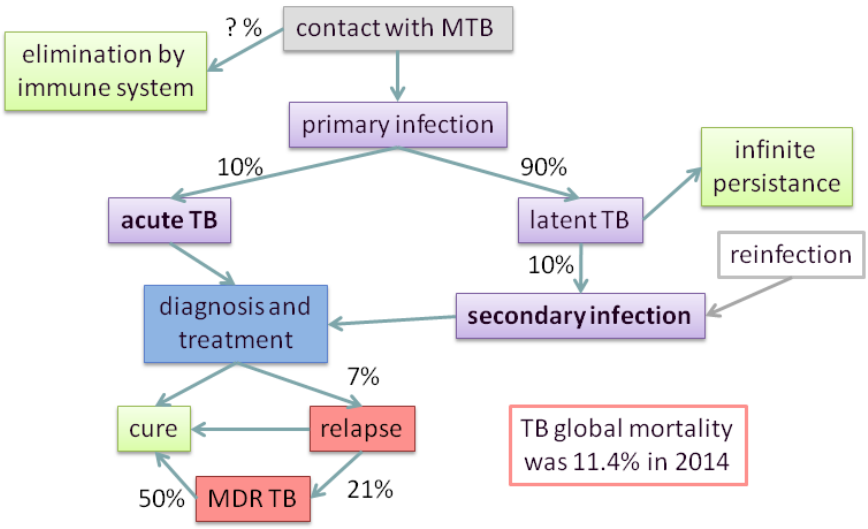


Figure 1.4. Overview of TB transmission and possible course of infection. Percentages show probabilities of that step based on data from 2015 (WHO, 2017).

Latent tuberculosis is identified by immunological tests using tuberculin skin test or interferon gamma release assays (WHO, 2017). However, these tests are only indicative of prior *M. tuberculosis* exposure. The tuberculin skin test shows false positive for individuals vaccinated with *M. bovis BCG* and false negative with immunocompromised patients

(Getahun et al., 2015). Therefore, the WHO guidelines do not limit latent tuberculosis treatment pending a positive tuberculin skin test, but instead recommend nine months of prophylactic INH treatment if latent tuberculosis risk is high (WHO, 2017).

M. tuberculosis in latent tuberculosis displays an altered phenotype termed dormant or non-replicating (Alnimr, 2015). Although, this phenotype is not exclusively observable in latent tuberculosis infections, there is a logical connection between delayed active disease and dormant bacilli. *M. tuberculosis* in the non-replicating persistence (NRP) state grows slower *in vitro*, is not culturable on solid media, stains differently and shows altered susceptibility to antimycobacterial drugs (Alnimr, 2015). Due to the metabolically altered NRP state bacilli have higher tolerance to cell-wall targeting drugs, like INH (Mitchison and Davies, 2012) and need "sterilising" drugs to get eliminated, such as PZA (Zhang and Mitchison, 2003). There are specialised *in vitro* and *in vivo* models particularly directed at testing drug efficacy against NRP *M. tuberculosis* (Alnimr, 2015).

1.1.3.2 Unique environmental niches

M. tuberculosis is an obligate human parasite with no scientifically proven environmental reservoir. Tuberculosis is an airborne infection, yet unlike other bacteria infecting the lungs, it targets specifically the lower lungs (Cambier et al., 2014). Avoiding large bacterial competition and often highly active immune responses in the upper respiratory tract, it is transmitted most efficiently by small droplets directed to the pulmonary alveoli (Cambier et al., 2014).

M. tuberculosis is a facultative intracellular pathogen targeting primarily (but no exclusively) alveolar macrophages (Ganbat et al., 2016). During its intracellular state *M. tuberculosis* endures various environmental conditions. It survives acidic conditions and lytic enzymes of the phagolysosome, nitric oxide and reactive oxygen species all aimed at

eliminating the bacterium (Ehrt and Rhee, 2013). Yet interestingly, *M. tuberculosis* seems to override endocytic pathways, prevents lysosome and phagosome fusion, and overcomes resource scarcity, such as iron or oxygen, and starts replicating to the level where the host cell undergoes necrosis (Baer et al., 2015)(Cambier et al., 2014).

Granuloma, an aggregation of macrophages and other cells, is the hallmark formation of tuberculosis infection (Cambier et al., 2014). In this structure macrophages are increasingly microbicidal and the intracellular environment is hostile towards *M. tuberculosis* (Cambier et al., 2014). The pathogen is present extracellularly outside the granuloma, intracellularly and inside the granuloma caseum – all of which result in phenotypically different sub-populations of *M. tuberculosis* (Baer et al., 2015) (see **Figure 1.5.** granuloma).

In primary tuberculosis the characteristic lesion is the caseating granuloma (Monin and Khader, 2014). It has a lipid-rich necrotic centre, surrounded by foamy and epithelioid macrophages and lymphocytes. This structure does not incorporate blood vessels in the centre, hence the centre of the granuloma is hypoxic and isolated from nutrient sources (see **Figure 1.5.**). Primary granulomas grow up to 2-3 cm in size and this is the 'tubercle' structure the disease has been named after.

Primary granulomas can undergo fibrosis and calcification and they can remain contained indefinitely, although in cases of infection with high bacillary load the granuloma can progress into a dysregulated inflammatory granuloma (Monin and Khader, 2014). These structures incorporate a large number of neutrophils and are the centres of unrestricted inflammation that impairs respiratory function (Monin and Khader, 2014), which is a typical symptom that leads patients to seek clinical help.

The stressful milieu of the granuloma leads to drug-tolerance in *M. tuberculosis* (Dartois, 2014; Baer et al., 2015). Replication slows down, generation time is prolonged from 24 hours

to more than 96 hours (Baer et al., 2015) and drugs targeting cell wall synthesis (INH, EMB) become ineffective. Especially in the caseum of the granuloma, persisting bacteria become phenotypically tolerant to most drugs, except towards RIF, PZA, bedaquiline and fluoroquinolones that inhibit metabolic processes still active in this phenotype (Baer et al., 2015). In general, accessing and effectively killing all these different subpopulations of *M. tuberculosis* is a complex problem of chemotherapy and is currently only possible with a multitude of drugs administered in combination.

To sum up**, Figure 1.5.** illustrates typical barriers that a drug compound needs to cross from the absorption or intravenous injection to reach the molecular drug target within the bacterium, all of which need a relevant model in laboratory testing. Animal models are the most relevant in studies at the granuloma level, however the pathology of human tuberculosis may not be mimicked accurately in some of these models. For example mice, which are the most widely used animal model, infected with MTB depending on mice strain, can develop specific types of granulomas, without hypoxia or necrotic caseum for example (Myllymäki et al., 2015). Therefore, correct understanding and application of mice strains immensely influence these experiments (Kramnik and Beamer, 2016). Surprisingly, important studies concerning granuloma formation and innate immune responses have been performed in zebrafish models with *Mycobacterium marinum*, because the granulomas that develop in various organs of the fish are very much alike the human granulomas (Parikka et al., 2012). For the level of bacterium and molecular target engagement whole cell assays are routinely applied. This topic is elaborated upon in Chapter 1.5.

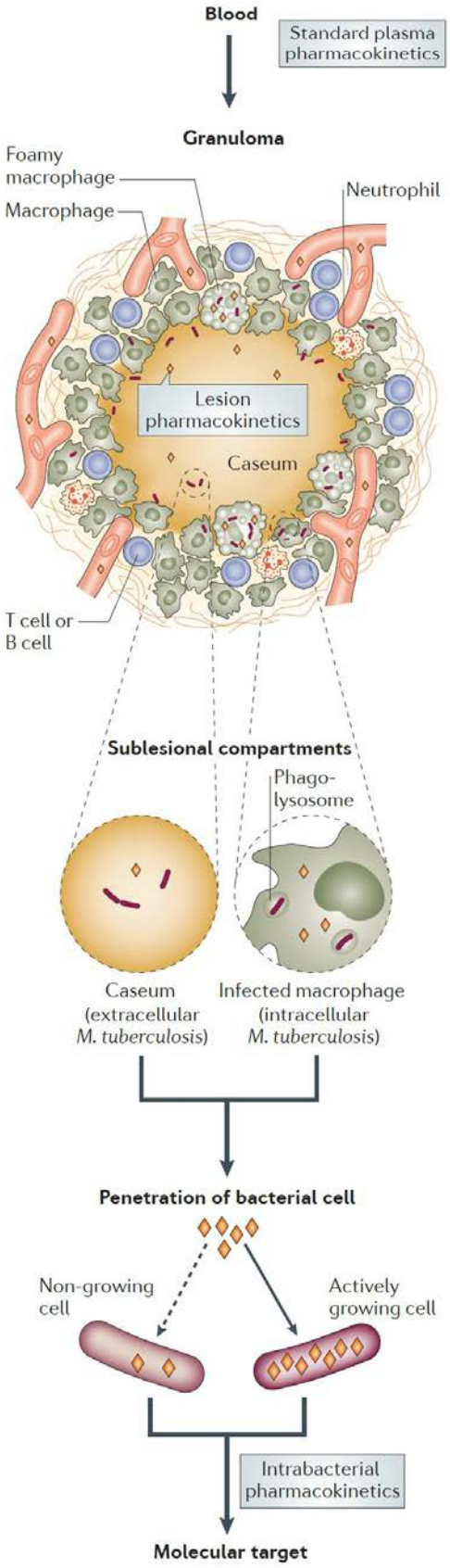


Figure 1.5. Path of antituberculotic drugs from blood to molecular target (Dartois, 2014). Orange rhombuses represent the drug compound.

1.2 Characteristics of the mycobacterial cell wall

Acid-fast bacilli are different from Gram-positive and Gram-negative bacteria regarding cell envelope structure. Briefly, in mycobacteria peptidoglycan is covalently linked to the cell wall core termed mycolyl-arabinogalactan-peptidoglycan (mAGP) (Abrahams and Besra, 2018). Glycolipids, like lipomannan (LM) and lipoarabinomannan (LAM) are integral parts of the cell wall (Abrahams and Besra, 2018). Covalently linked mycolic acids of mAGP together with free lipids and waxes form an outer membrane establishing a pseudo-periplasmic space like Gram-negatives (see **Figure 1.6.** for comparison). The outer membrane is covered by a capsule-like material of polysaccharides and proteins (Abrahams and Besra, 2018).

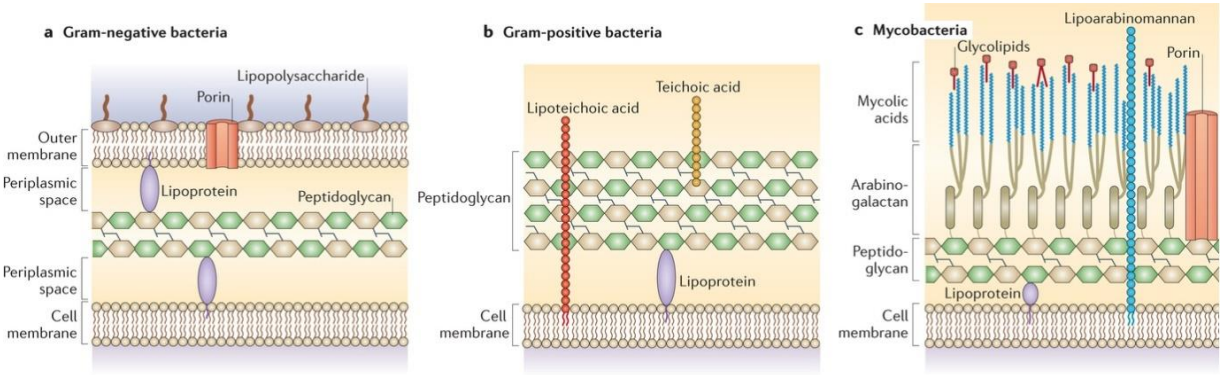


Figure 1.6. Comparison of bacterial cell walls (Brown et al., 2015)**.** Capsule material is not illustrated.

A unique cell wall structure also means highly distinguished set of biosynthetic enzymes – many of these enzymes are essential and have no homologues outside the Corynebacterium-Mycobacterium-Nocardia branch of bacteria (Jankute et al., 2012).

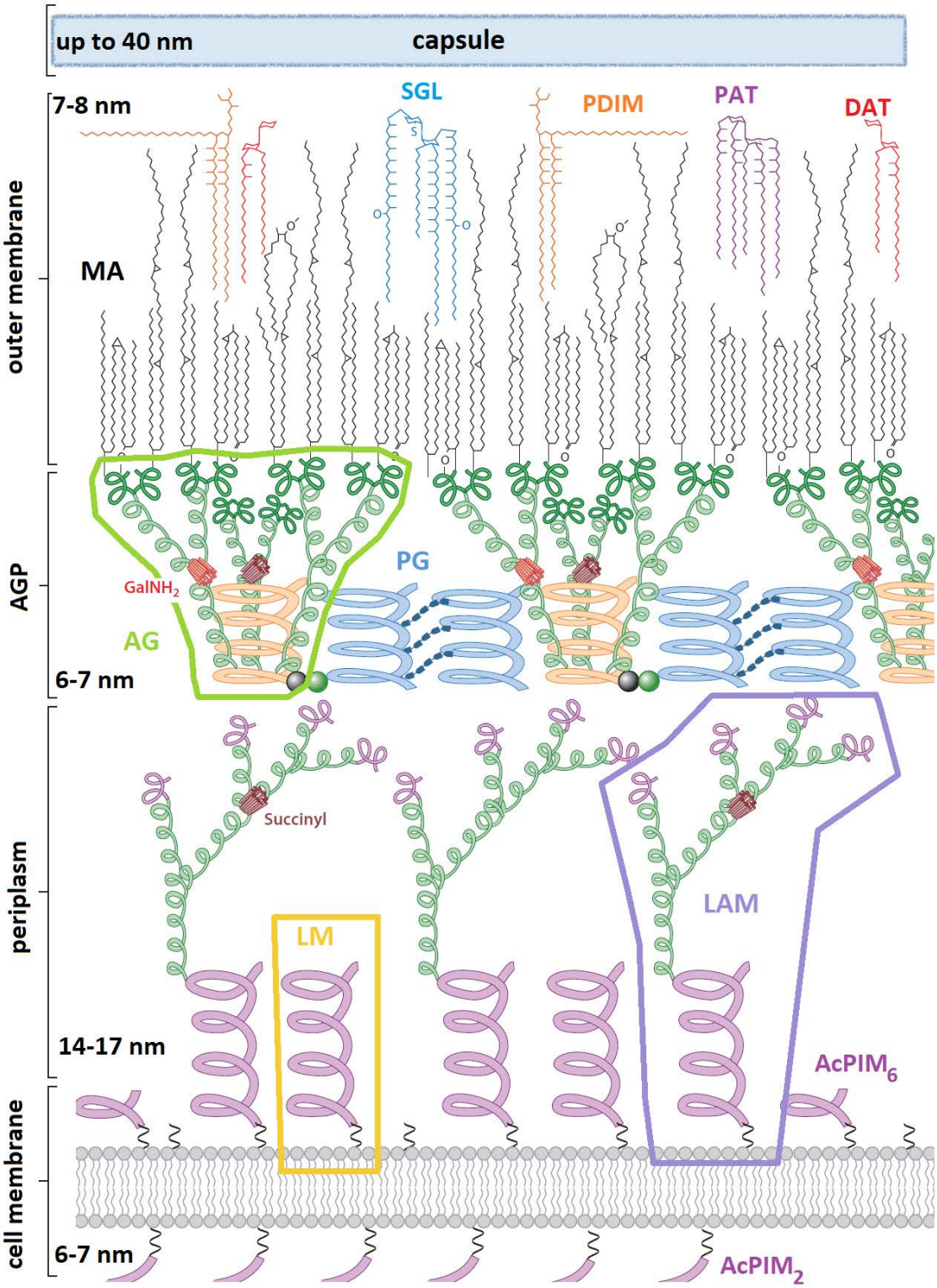


Figure 1.7. Simplified *M. tuberculosis* cell wall model highlighting discussed elements in the text, original picture (Jankute et al., 2015) **with edited elements.** AGP is arabinogalactan-peptidoglycan, MA mycolic acid, AG arabinogalactan, PG peptidoglycan, LM lipomannan, LAM lipoarabinomannan, AcPIM acylated phosphatidyl-*myo*-inositol mannoside. Extractable lipids: SGL sulfoglycolipid, PDIM phthiocerol dimycocerosate, PAT polyacyltrehalose, DAT diacyltrehalose.

1.2.1 Lipoarabinomannan (LAM), lipomannan (LM) and phosphatidyl-*myo*-inositol mannosides (PIMs)

By the latest cell wall models (Kaur et al., 2009; Minnikin et al., 2015) these structures are non-covalently immersed into the leaflets of both the cell membrane and the mycobacterial outer membrane (only the former represented in **Figure 1.7.**). In fact, PIMs modify the stability of the inner cell membrane so much, that it has been suggested to distinguish the structure as mycobacterial inner membrane (Minnikin et al., 2015).

The anchor for all three listed macromolecules is phosphatidyl-*myo*-inositol. In PIMs the anchor is typically bound to two or six mannosyl residues (PIM_2 and PIM_6), while maximum two more acyl groups may be added, one to the *myo*-inositol and another to the mannopyranose part. In *M. tuberculosis* the most common acyl chains are palmitic acid and tuberculostearic acid substitutions (Jankute et al., 2015).

LM and LAM are also linked to phosphatidyl-*myo*-inositol extended through a mannan homopolymer (α(1→6) Man*p*), with species-dependent typical monomannose residues bound to the main polymer chain (α(1→2) Man*p*), in total containing 20-30 mannosides (Jankute et al., 2015). The difference between LM and LAM is the addition of an arabinan chain (α(1→5) Ara*f*), approximately 50-70 residues in LAM, which is also branched (α(3→5) Ara*f*) with a tetra-arabinoside or a (further branched and succinylated) hexa-arabinoside (Jankute et al., 2015). These sidechains are frequently capped with dimannoside residues (or mono and tri-mannosides less frequently), forming ManLAM (Jankute et al., 2012). This mannoside capping motif of LAM has an important role in immune reactions, however results in this topic are contradictory and need further assessment (the effect of purified ManLAM as model for cell wall component ManLAM has been questioned (Kaur et al., 2009)).

Synthesis of these three molecules are similar until tri- or tetra-acylated-PIM_2 synthesis, aided by PimA and PimB (Rv2188c) (Guerin et al., 2009). These two enzymes utilize GDP-Man*p* as the sugar donor, but interestingly, the enzyme Ppm1 synthesises a polyprenyl-phosphorylmannose, which is utilised by transmembrane glycotransferases (Rana et al., 2012). C_{50} isoprenyl carriers involved in this reaction are similar to the prenyl chain of decaprenyl-phosphorylarabinose (DPA), the substrate of Mt-DprE1 (discussed in **Chapter 1.4.**). Also, these polyprenyl substrate utilising steps take place in the periplasm (similar to DPA synthesis (Brecik et al., 2015)).

Enzymes MptA and MptB synthesise the linear mannan polymer chain, MptC finishing the mannan side chains of LM (Abrahams and Besra, 2018). LAM arabinan side chain is synthesised by some of the enzymes involved in arabinan synthesis of AG (see **Chapter 1.2.3.**), involving AftB, AftC, EmbC (Alderwick et al., 2005, 2006; Escuyer et al., 2001), however several enzymes of the pathway have not yet been determined.

1.2.2 Peptidoglycan (PG)

PG is part of the mAGP, it is a rigid biopolymer responsible for the structural integrity of the cell wall (Minnikin et al., 2015). It is located on the outer side of the cell membrane and it is considered to be similar to other bacterial PGs (Brennan and Nikaido, 1995). The polymer has a backbone structure of alternating N-acetylglucosamine and N-acetyl or N-glycolyl-muramic acid residues (β(1→4) linkage), which are highly cross-linked (70-80%) by tetra-peptide chains (Alderwick et al., 2015).

The synthesis of PG starts with the formation of uridine-diphosphate-N-acetylglucosamine (UDP-GlcNAc), which is modified by a sequence of Mur ligases to form Park's nucleotide, a soluble building block that is than coupled to decaprenyl-phosphate and anchored to the cell membrane (Jankute et al., 2015). This structure is called Lipid I and reacts with another

UDP-GlcNAc molecule to form Lipid II, which is the base component of this polymer. Lipid II is flipped to the outer surface of the cell membrane, and penicillin binding proteins (PBPs, the targets of β-lactam antibiotics), which are transglycosylases and transpeptidases, link the moieties to form the PG polymer.

Several TB drugs target the PG pathway. D-cycloserine prevents Park's nucleotide formation, capuramycins block the formation of Lipid I (Muramatsu et al., 2003), vancomycin prevents polymer crosslinking (Meadow et al., 1964). It is important to point out that vancomycin and β-lactams are not potent against MTB on their own, instead only have an effect if the cell wall structure is already compromised by other drugs (Soetaert et al., 2015).

1.2.3 Arabinogalactan (AG)

AG is the central component of mAGP and it is formed of galactose and arabinose homopolymers (Jankute et al., 2012). About 10-12% of N-glycolylmuramic acid (MurNGlyc) of PG is covalently linked to AG. According to the latest model by Minnikin *et al.* (2015), AG adopts a helical structure. The model featured in **Figure 1.7.** shows PG helices interspersed by the galactan part of AG. Other models consider PG as a web-like base layer below AG (Kaur et al., 2009). The terminal hexa-arabinoside of AG arabinan is covalently linked to mycolic acids, which forms the basis of the inner leaflet of the outer membrane (Alderwick et al., 2015).

AG biosynthesis starts with the formation of a linker unit, which is a disaccharide consisting of GlcNAc and *L*-rhamnose, shown as green and black spheres in **Figure 1.7.** GlcNAc of the linker unit is added by WecA, while rhamnose by WbbL (Jin et al., 2010; Mills et al., 2004). Galactose monomers are added on the cytosolic side of the membrane by enzymes Glft1 and Glft2, with alternating β(1→5) and (β1→6) linkages (Mikušová et al., 2006).

Meanwhile, the synthesis of the arabinan chain is expected to take place on the periplasmic side (Jankute et al., 2012). AftA primes the galactan with the initial arabinose and the chain is elongated by the Emb proteins, while Aft proteins catalyse branch additions of Ara*f* (Alderwick et al., 2006; Escuyer et al., 2001). The arabinan part of AG is further decorated with galactosamine and succinic acid (see in **Figure 1.7.**). Ultimately, the synthesised macromolecule is ligated to PG by the enzyme Lcp1 (Harrison et al., 2016).

GlcNAc of the linker unit is anchored to the membrane by decaprenyl-phosphate, while the *L*-rhamnose is transferred from desoxythymidine diphosphate rhamnopyranose (dTDP-Rha*p*) (McNeil et al., 1990; Jin et al., 2010). The donor for galactose is the soluble uridine-diphosphate-galactofuranose (UDP-Gal*f*)(Weston et al., 1997), while arabinose residues are transferred from decaprenyl-phosphorylarabinose (DPA)(Mikušová et al., 2005).

Inhibition of the AG synthesis has been highlighted by the discovery of several compound classes. Currently, the enzyme WecA involved in this synthesis steps is the target of the experimental drug CPZEN-45 (capramycins) (Clinical pipeline, 2018). DPA formation is prevented by Mt-DprE1 inhibitors, such as benzathiazinones and dinitrobenzamides in pre-clinical studies (Makarov et al., 2009; Christophe et al., 2009). The front-line drug EMB targets the Emb proteins of AG biosynthesis (Telenti et al., 1997).

1.2.4 The outer membrane of *M. tuberculosis*

Mycolic acids are long chain fatty acids (C_{70}-C_{90}) mostly located attached to AG or unbound as free trehalose-mycolates (Jankute et al., 2012). Together with other trehalose derivatives they form the basis of the mycobacterial outer membrane (see **Figure 1.7.** in various colours, names not listed for simplification). These molecules render *M. tuberculosis* with its hydrophobic outer membrane and result in exceptionally low permeability against staining, drugs and endurance in the intracellular microenvironment (Angala et al., 2014).

The “free lipids” of the outer membrane are not covalently bound and can be extracted and separated with the use of different organic solvents (Minnikin et al., 1985). (In **Figure 1.7.** sulfoglycolipid, phthiocerol dimycocerosate and the trehalose-derivateives, polyacyltrehalose and diacyltrehalose are shown with the sulfur-containing sulfoglycolipid.

Trehalose dimycolate termed the ‘cord factor’ is exceedingly important in immune responses, particularly in causing inflammation (Angala et al., 2014). The transport of hydrophilic molecules is possible via porin proteins immersed in this outer membrane structure (not shown on **Figure 1.7.**) (Angala et al., 2014).

Trehalose is a non-reducing disaccharide absent from human cells, yet mycobacteria possess three different biosynthetic pathways for its synthesis (Kaur et al., 2009). Mycolic acids are synthesised in the cytoplasm via the fatty acid synthase pathways FAS-I and FAS-II (Quémard, 2016). Enzymes of the FAS-II cycle in mycolic acid synthesis are especially favourable targets because this pathway is completely absent from humans (Chetty et al., 2017). Due to the abundance of enzymes involved in their synthesis, here only proteins relevant to drug mechanism of action are discussed further.

The front-line drug INH is activated by KatG and targets InhA (Islam et al., 2017), which is also the target of ethionamide and prothionamide (Marcinkeviciene et al., 1995). InhA is an enzyme in FAS-II (Abrahams and Besra, 2018). Pks13 participates in one of the final steps of mycolic acid synthesis and it is the target of some preclinical drugs (Islam et al., 2017). MmpL3 is a transport protein which is through to transport mycolic acids to the periplasm, which operates under the control of proton motive force (Li et al., 2014). Therefore, structurally diverse inhibitors e.g. delamanid and SQ109 (Islam et al., 2017) have one common feature: they disrupt the electrochemical gradient which drives MmpL3 (Cole, 2016). The most promising such inhibitor, SQ109 is in phase 2 clinical trials (Clinical

pipeline, 2018). Finally, the antigen 85 complex, which is responsible for linking mycolates to AG and catalysing cord factor synthesis, is the target of preclinical drug I3-AG85 (Warrier et al., 2012).

1.2.5 Capsule material

The outer layer of the *M. tuberculosis* is a composite of polysaccharides and proteins, and its extension is highly variable and dependent on the growth environment of the bacilli (Angala et al., 2014). 70% of capsular material is α-D-glucan, a polymer of glucose similar to glycogen (Ortalo-Magne et al., 1995). In addition, D-mannan and D-arabinomannan are components of this structure, which are identical to the relevant features of LM and LAM, sharing enzymes in the biosynthetic pathway. Enzymes of the capsule synthesis may be interesting drug targets, for example GlgE (Rv1327c) maltose-transferase needed for α-glucan synthesis has been validated as potential drug target (Kalscheuer et al., 2010).

1.3 Present state of tuberculosis treatment

1.3.1 Current set of anti-tuberculosis drugs

Modern tuberculosis therapy applies exclusively drugs in combinations to minimise the risk of drug-resistance. Combinational chemotherapy has been advised since 1948 because STM and PAS resistance occurred quite frequently when administered in monotherapy (Fox et al., 1999). Some elements of the regimen were replaced because other drugs were more effective in lower dosages (PAS replaced by EMB), and some were held back for severe side-effects, such as STM was replaced by PZA, thioacetazone was withdrawn (Murray et al., 2015). Some of these unfavourable drugs are still second-line alternatives in tuberculosis therapy.

Currently, drug-susceptible tuberculosis treatment is according to the directly observed therapy short-course (DOTS) programme (Iseman, 2002). In the initial two months the

combination of four front-line drugs is administered (INH/RIF/PZA/EMB), followed by four months of treatment with INH and RIF only. This 6-months course is the shortest treatment period with acceptable relapse rates (Fox et al., 1999).

For drug resistant tuberculosis therapy, the earlier WHO guidelines (WHO, 2010) required four and the latest guidelines a minimum of five effective drugs administered in combination in the intensive phase (8 months). The WHO 2016 guidelines (WHO, 2016) regrouped drugs involved in MDR-TB therapy and also introduced the 'shorter MDR-TB regimen'. This is a standardised treatment based on the Bangladesh regimen (TBfacts.org, 2018) for pulmonary TB with suspected/confirmed drug resistance, regardless of the HIV status of the patient and it may be concluded in only 9 months.

1.3.2 Adjunct therapies in TB treatment

Host directed therapies (HDTs) follow an often overlooked approach in tuberculosis research: they explore the possibility to improve the reaction of the human host to the disease. This may be achieved by either enhancing adequate defensive immune reactions, lowering negative immunological mechanisms such as inflammation, and hindering interactions and processes in the host that are exploited by the bacterium (Palucci and Delogu, 2018). The containment of the infection and an optimal immune response is important both in primary infection, and also in the case of reactivation of the active disease from a latent phase (see **Figure 1.4.**). The main problem with modulating the immune response is the consequences that it may hold for TB patients with HIV-coinfection or diabetes, who could react negatively to these drugs or simply show no improvement (Lee et al., 2018). With co-morbidities present, it is also more and more difficult to weigh the cost-benefit of these drugs added to the treatment and how they would interact with each other.

Similarly, compounds targeting the virulence of the microbe may improve anti-tuberculosis chemotherapy. *M. tuberculosis* is known for its ability of forming biofilms with altered drug susceptibility (Ojha et al., 2008). Also, increased transcription and presence of drug efflux pumps have been shown to cause drug resistance during early drug treatment *in vivo*, which was suppressed by the inhibition of these efflux pumps by reserpine for example (Adams et al., 2011).

1.3.3 Typical drug targets of *M. tuberculosis*

TB drugs target various physiological processes of *M. tuberculosis*. Nevertheless, it is not without precedent that the mechanism of action of a drug is poorly understood, yet it gains approval for clinical use, such as clofazimine (Yano et al., 2011). In general, TB drugs can be grouped into three categories: nucleic acid synthesis inhibitors, protein synthesis inhibitors or other mechanism of actions.

Nucleic acid synthesis inhibitors include the front-line drug RIF (and derivatives rifapentine, rifabutin) that inhibit RNA polymerisation (Chetty et al., 2017). Fluoroquinolons (ciprofloxacin, gatifloxacin, levofloxacin, moxifloxacin, ofloxacin) that inhibit DNA replication and supercoiling (Chetty et al., 2017).

Numerous TB drugs target protein synthesis. PZA, although it has multiple mechanisms to kill MTB (for example causes intracellular acidification) also hinders protein translation (Zhang and Mitchison, 2003). STM and aminoglycosides (kanamycin, amikacin) target the 30S ribosomal subunit in translation, just like the experimental compound, spectinamide 1599 (Lee et al., 2014). Macrolids (like clarithromycin) and oxazolidinones (sutezolid and linezolid (Chetty et al., 2017; Islam et al., 2017) and the compound delpazolid (LCB01-0371) (Kaul et al., 2018)) all target the 50S subunit (Chetty et al., 2017). Capreomycin and

viomycin are cyclic peptides that also inhibit protein synthesis, they impede the interaction between the two ribosomal subunits (Chetty et al., 2017).

Many compounds in development attempt to inhibit novel pathways, that clinically approved drugs do not exploit. Q203, an imidazopyridine amide in phase 1 clinical trials, targets the protein QcrB of the cytochrome bc1 complex, essential to ATP synthesis (Pethe et al., 2013). Meanwhile, compound GSK 070 (GSK 3036656) is in phase 1 clinical studies targets leucyl-tRNA synthetase to attack protein synthesis from a different angle (Chetty et al., 2017).

Several clinically approved drugs target the cell wall – proteins involved in cell wall synthesis (some discussed in Chapter 1.2.), or because they are located in the cell wall. The front-line drugs INH and EMB both target mycolic acid and AG synthesis (Marcinkeviciene et al., 1995; Telenti et al., 1997), while bedaquiline targets transmembrane ATP synthase AtpE (Goldman, 2013). QcrB is also located in the cell membrane, since the electron transport chain complexes of prokaryotic cells are situated there (instead of the mitochondrion).

One of the reasons why cell wall targets are widely studied is that they have appeared in multiple phenotypic screens (Cole, 2016). It is yet to be explained why they are highly druggable, but several proteins of the cell wall are promiscuous targets and they can be inhibited by different chemical scaffolds unrelated to each other. QcrB is one of these proteins, among MmpL3 (both integral membrane proteins) and DprE1 (Cole, 2016).

Another reason for drugs targeting membrane proteins is the hydrophobic nature of both compounds and the cell wall. Even highly hydrophobic compounds can reach their target in the lipid-rich environment, while hydrophilic compounds may never reach their intracellular target exactly because of this lipid-rich, acid-fast mycobacterial cell wall.

1.4 Drug targets Mt-DprE1 and Mt-DprE2

1.4.1 Mt-DprE1, validated drug target

1.4.1.1 Physiological function of DprE1/2

M. tuberculosis has one arabinose donor in cell wall synthesis, DPA (Wolucka et al., 1994). It is used as substrate by the arabinofuranosetransferase families Aft and Emb to synthesise AG, LAM (see **Chapter 1.2.**). Enzymes involved in DPA synthesis (see **Figure 1.8.**) are essential except for Rv3807c (Kolly et al., 2014). DprE epimerase subunits 1 and 2 (DprE1 and DprE2) catalyse the final steps of DPA biosynthesis (Mikušová et al., 2005).

Although, the biosynthetic pathway of DPA is essential for *M. tuberculosis*, the related species *Corynebacterium glutamicum* does not need arabinan to grow (Alderwick et al., 2005). Viable mutants that lack *ubiA* have been produced that have no arabinan incorporated into the cell wall. This has been best explained by Grover *et al.* as *dprE1* is a vulnerable drug target not because of the essentiality of arabinan, but because DPR accumulates in toxic levels inside the cell without *dprE1*. Furthermore, other biological processes which would use the same prenyl-carrier as ribose in DPR, are limited due to the withdrawal of the carrier molecule (Grover et al., 2014).

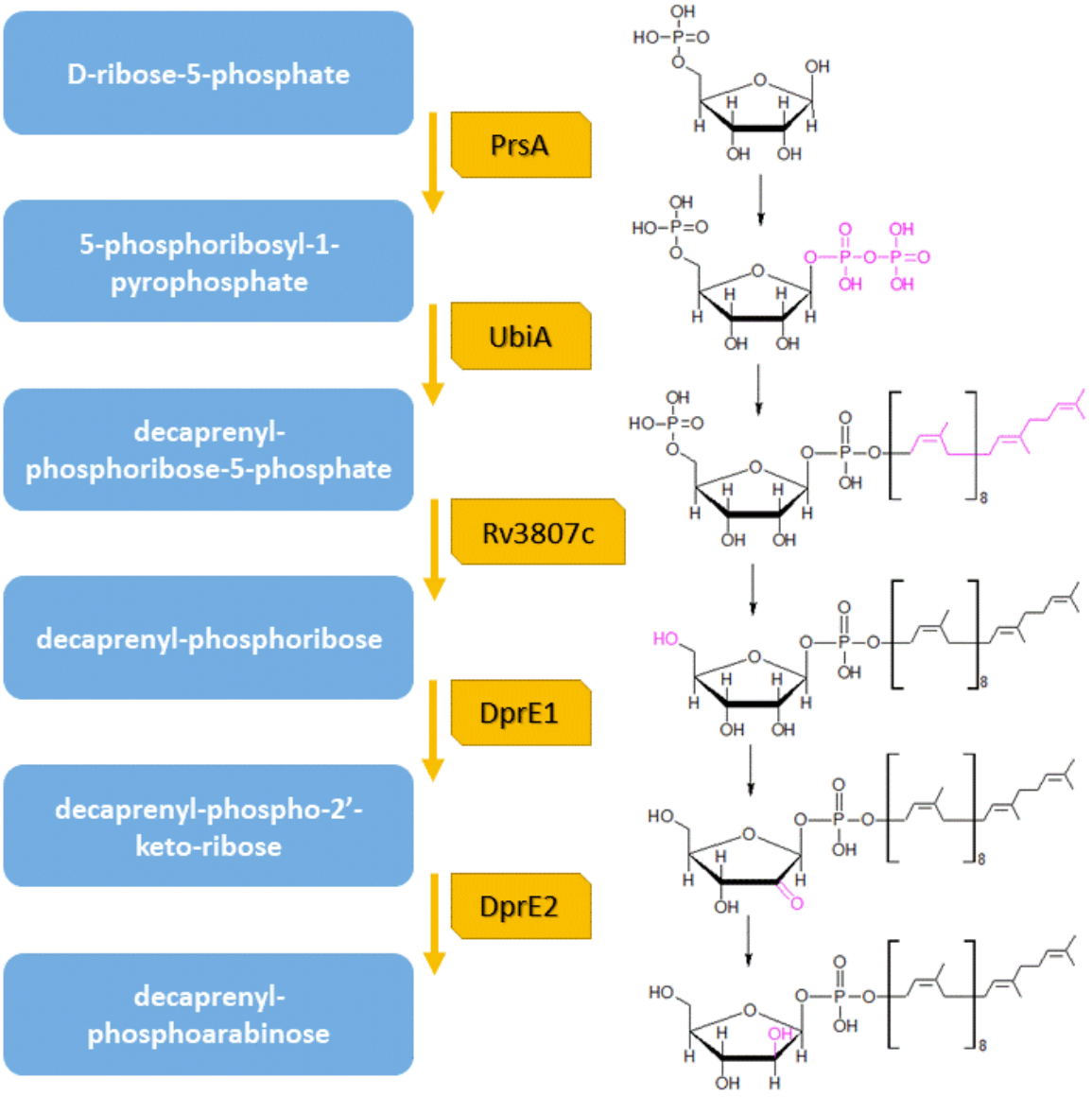


Figure 1.8. The DPA biosynthetic pathway. The blue boxes show the long name of intermediate sugar moieties, yellow boxes indicate the enzyme catalysing each step. Pink highlight signals the molecular changes in the structure at each step (Jankute et al., 2015; Abrahams and Besra, 2018)

The glycotransferases that utilise prenylated sugar donors were postulated to operate on the extracellular side of the membrane, while nucleotide diphosphate-bound sugar reactions take place on the intracellular side of the membrane (Berg et al., 2007). The prenyl-sugar moiety is immersed in the membrane, and most enzymes that use them as substrate, are also integral membrane proteins (Berg et al., 2007; Abrahams and Besra, 2018).

The epimerisation reaction of DPR to DPA is a two-step reaction. Firstly, DprE1 oxidises DPR to the intermediate decaprenylphosphoryl-2-keto-β-D-*erythro*-pentofuranose (DPX), which is made possible by the concomitant reduction of a non-covalently bound FAD co-factor (Batt et al., 2012; Mikušová et al., 2005). Secondly, DprE2 reduces DPX to DPA with the assistance of reduced nicotinamide adenine dinucleotide (NADH) or nicotinamide adenine dinucleotide phosphate (NADPH) as a co-factor (Trefzer et al., 2012). Both enzymes are reported to be periplasmic, membrane associated but not integrated, therefore possibly this epimerisation reaction takes place on the periplasmic side of the cell membrane (Brecik et al., 2015).

1.4.1.2 Structure and inhibition of Mt-DprE1

Mt-DprE1 belongs to the vanillyl alcohol oxidase/*para*-cresol methylhydroxylase family of proteins (Ewing et al., 2017). These proteins are structurally similar, all possessing an FAD-binding domain, some bind the co-factor covalently (Ewing et al., 2017). Mt-DprE1 does not, however, in the absence of the co-factor it is structurally unstable that has been an issue during early purification studies of the protein, before the use of chaperons and codon-optimised plasmids (Batt et al., 2012). Earlier, FAD was added in excess during the purification process to keep the protein folded (Neres et al., 2012). Also, members of this protein family often form homodimers or octamers (Ewing et al., 2017), but pure recombinant Mt-DprE1 is monomeric in solution (Batt et al., 2012).

Mt-DprE1 has a highly conserved DNA sequence among related taxa (Batt et al., 2012). Indeed, Mt-DprE1 (gene *rv3790*) shares 82% sequence identity with DprE1 in *M. smegmatis* (gene *MSMEG_6382*) on the amino acid level. As a surrogate the protein the *M. smegmatis* protein has been used in inhibitor studies with encouraging results (de Jesus Lopes Ribeiro et al., 2011). *Mycobacterium avium* and *Mycobacterium aurum* however are poor models:

they are intrinsically resistant to covalent inhibitors (like benzothiazinones) because their Cys387 is substituted with other amino acids (de Jesus Lopes Ribeiro et al., 2011), the target residue of these compounds. Mt-DprE1 is a validated drug target and got to the centre of attention due to its high druggability and known 3D structure (Batt et al., 2012). So far 15 different compound classes have been identified to inhibit Mt-DprE1 specifically (Piton et al., 2017).

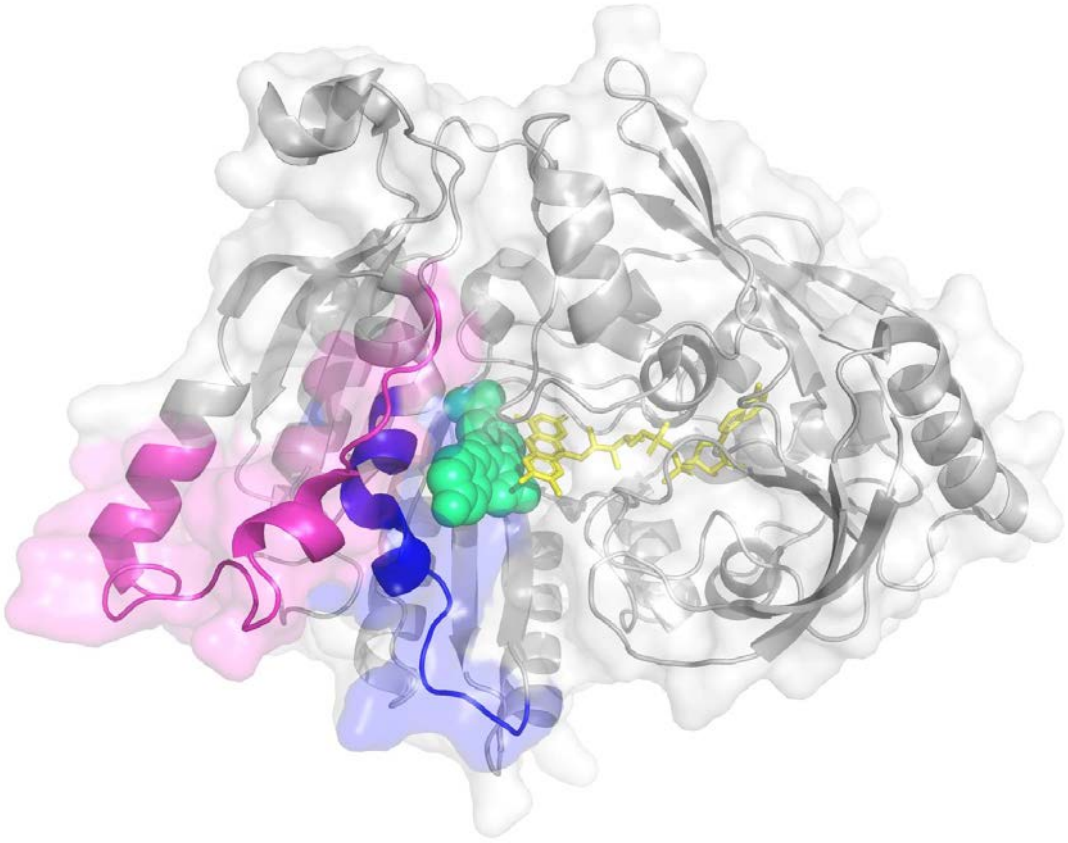


Figure 1.9. Three-dimensional structure of Mt-DprE1. Surface and gray cartoon shows Mt-DprE1 (pdb:4P8L) with the short and large loop region highlighted in blue and pink (316-330, 269-297) respectively. The ligand is Ty36C (Neres et al., 2015) in green spheres, FAD is in yellow sticks. Made in PyMol Version 2.0.6. (DeLano, 2002)

The most notable covalent inhibitors are benzothiazinones (BTZs) (Makarov et al., 2009) and dinitrobenzamides (Christophe et al., 2009). All covalent inhibitors are aromatic nitro-compounds that are activated *via* FADH by reducing the nitro-group to the active nitroso form (Trefzer et al., 2012) and form an adduct with Cys387 of Mt-DprE1. Nitro compounds

are prone to mutagenicity (Piton et al., 2017) which makes optimisation difficult, but PBTZ169 which is a clinical candidate molecule, recently termed macozinone (Clinical pipeline, 2018) is in phase 2 clinical studies with one of the lowest minimal inhibitory concentrations in the field (0.6 nM) (Makarov et al., 2014; Piton et al., 2017).

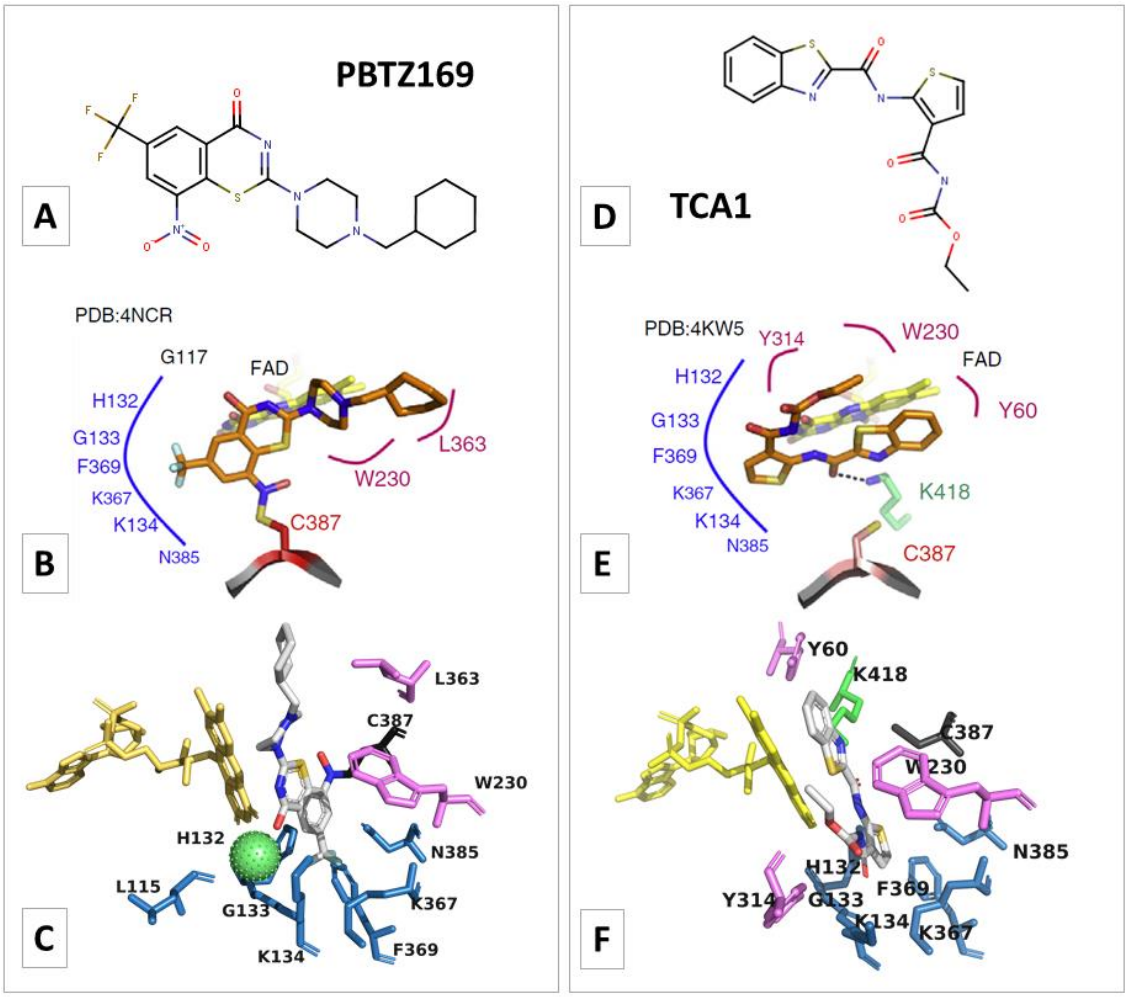


Figure 1.10. Mt-DprE1 binding a covalent and non-covalent inhibitor example. (A-C) showcase covalent inhibitor PBTZ169, (D-F) the non-covalent inhibitor TCA1. (A) and (D) show compound structure, (B) and (E) are interaction maps by Piton et al. (2017), (C) and (F) the same interaction excerpt in 3D model, with the ligand coloured by element. FAD in yellow sticks. In C) the green ball represents the water molecule that interacts with both PBTZ169 and L115.

Non-covalent inhibitors without nitro-groups, like TCA1 (Wang et al., 2013) or azaindoles (Shirude et al., 2014) offer a safer but less potent alternative to covalent inhibitors. Lead compounds of the different classes usually have an MIC below 1 μM (Piton et al., 2017).

The azaindole TBA-7371 is in the most advanced stage of clinical studies among these compounds (in phase 1) (Clinical pipeline, 2018). The number of inhibitors is still rising with the help of docking studies with the solved crystal structures and the development of *in vitro* enzymatic assay (Batt et al., 2016; Neres et al., 2012).

Difference between covalent and non-covalent binding is illustrated in **Figure 1.10.** The involvement of Cys387 is less in binding with non-covalent inhibitors, also different amino acid side chains interact with the ligand. For example, the involvement of Tyr60 in TCA1 binding, but not with PBTZ169 (Piton et al., 2017). The colouring in **Figure 1.10.** follows the example of Piton *et al.* (2017) with recognising a more constant, hydrophobic surface (residues in blue) interacting with various ligands. Other residues join the interaction only with certain ligands and with van der Waals bonds (in pink) or with hydrogen bonds (in green). These interactions are more difficult to predict based on *in silico* docking studies because these residues include more flexible elements.

Mt-DprE1 has two loop regions, a long (amino acids 269-297) and a short disordered region (amino acids 316-329) as shown in **Figure 1.9.** These structures do not appear on apo enzyme crystal structures, because of their flexibility and have been a topic of speculation as to their function and importance in protein-ligand interactions. Co-crystals with inhibitors have demonstrated that certain ligands can stabilise one or both of these loops (Piton et al., 2017).

The short disordered region is placed over the active site of the protein (see **Figure 1.10.**), and became ordered in the presence of ligand CT319 (BTZ derivative), while Ty36c and QN129 (2-carboxyquinoxalines) stabilised both loops (Neres et al., 2015).

Azaindoles, TCA1, pyrazolopyridones and aminoquinolones resistant strains can be cross-resistant due to a common Tyr314His mutation (Piton et al., 2017). This residue is near the

short loop and in direct interaction with these inhibitors, but even these mutant strains remained susceptible to BTZs (Piton et al., 2017). Besides different point mutations, BTZ resistance was also demonstrated in *M. smegmatis* by overexpression of the nitroreductase NfnB (Manina et al., 2010). This enzyme has no confirmed homologue in MTB, but similar mechanisms are possible with similar inactivating enzymes.

1.4.1.3 Biochemical assay of Mt-DprE1

Enzyme activity of Mt-DprE1 was first monitored with radiolabelled substrates (Mikušová et al., 2005). The method used the membrane fraction of *M. smegmatis* incorporating an active enzyme mixture and p[^{14}C]Rpp, the substrate of UbiA (see **Figure 1.7.**). Radioactive products DPR, DPA and DPX were separated by thin layer chromatography.

A big step towards miniaturisation and screening was to introduce farnesyl-phosphorylribose (FPR), a truncated form of DPR to perform enzymatic reactions in solutions. The reaction could be followed by oxidising the reduced co-factor, FAD (Makarov et al., 2009). FAD can be re-oxidised by Amplex Red or dichlorophenolindophenol (DCPIP) electron acceptors (Neres et al., 2012) or resazurin (Batt et al., 2016) (illustrated in **Figure 2.3.** in Chapter 2). This assay type is useful in that it can be used to measure enzyme kinetics and to test inhibitors. It has been found that DprE1 exhibits allosteric enzyme kinetics (Neres et al., 2012; Batt et al., 2012), a phenomenon yet to be explained.

1.4.2 Mt-DprE2, potential drug target

Mt-DprE2 is an essential enzyme (Kolly et al., 2014) in *M. tuberculosis* and as most enzymes of the DPA synthetic pathway (**Figure 1.8.**), it is a potential drug target.

Mt-DprE2 is located in the same genetic cluster as Mt-DprE1 on the mycobacterial chromosome. Upstream of *dprE1* (*rv3790*) is *rv3789* coding Mt-GtrA, that has been identified as a transmembrane protein with suggested flippase (Larrouy-Maumus et al.,

2012) and/or recruiting function for AftA (Kolly et al., 2015). *dprE2* (*rv3791*) is just downstream of *dpre1*, followed by *aftA* and the Emb operon (see **Figure 1.11.**) (Kolly et al., 2015). AftA is the transmembrane enzyme that primes the galactan polymer of AG with the first arabinofuranose residue (Alderwick et al., 2006), while the EmbCAB proteins are glycotransferases that participate in arabinan biosynthesis (Jankute et al., 2012).

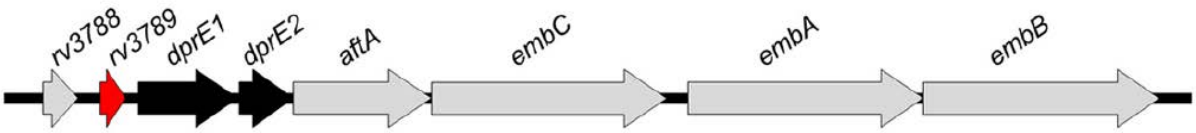


Figure 1.11. *dprE1/2* locus in *M. tuberculosis* by Kolly *et al.* (2015). The genes *rv3789*, *dpre1*, *dpre2* and *aftA* are co-transcribed to the same mRNA (Kolly et al., 2015).

DprE2 has a wider interactome than DprE1 based on bacterial two-hybrid system experiments in *C. glutamicum*. While DprE1 was indicated to bind itself and DprE2, DprE2 was reported to bind DprE1, DprE2, AftA, AftB, AftC and GlfT1 (Jankute et al., 2014). This implicates further that DprE2 inhibition could impair severely MTB physiology.

The solution structure of Mt-DprE2 has not yet been solved, but an *ab initio* computer model based on amino acid sequence has been produced (Bhutani et al., 2015). The model predicts that the protein belongs to the short-chain dehydrogenase/reductase family performing NAD(P)H-dependent redox catalysis. Presumably, Mt-DprE2 takes the form of a typical Rossmann fold and has three disordered regions (Bhutani et al., 2015). The only enzymatic assay established to date for Mt-DprE2 is a radiolabelled assay developed by Trefzer *et al.* (Trefzer et al., 2012). It has been applied to confirm NAD(P)H-dependence and to prove that BTZs do not inhibit Mt-DprE2 (Trefzer et al., 2012).

1.5 Phenotypic screens in TB early stage drug discovery

1.5.1 Early stage drug discovery

Early stage drug discovery focuses on finding new compounds active against *M. tuberculosis* that can be improved in potency and physicochemical properties to become drug candidates in the future. It is an iterative process of constantly re-evaluating each derivative compound in the general setting of the 4Ms: compounds tested with the Molecular target, at whole-cell level for MIC, in Mice and finally in Man (Ballell et al., 2013). The compound as it progresses signifies higher and higher financial investment which makes it crucial to identify and eliminate unsuitable molecules as early as possible. Compounds follow the typical evolution to drug illustrated in **Figure 1.12.**

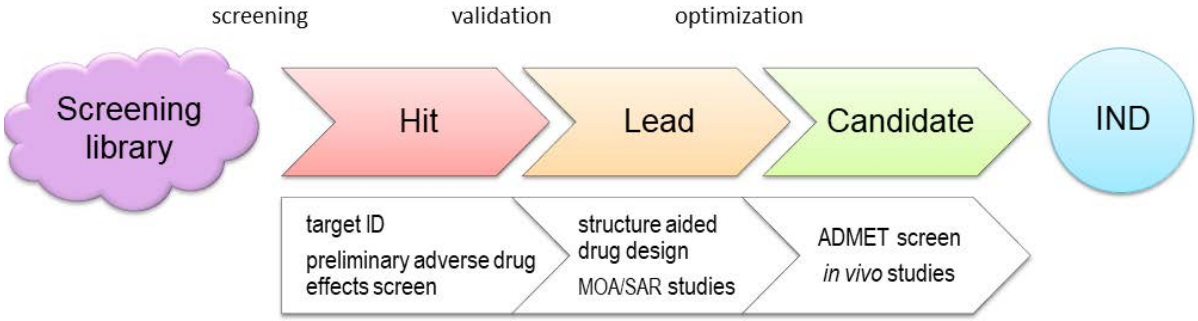


Figure 1.12. The progression of a compound in early stage drug discovery. ID – identification; MOA – mechanism-of-action; SAR – structure-activity relationship; ADMET – absorption, distribution, metabolism, excretion, toxicity; IND – investigational new drug.

Drug attrition is high in most antibiotic discovery projects – it has been estimated that about 2066 high-throughput screens deliver one antibiotic with a new mechanism of action (So et al., 2011). As R. H. Baltz stated "the success of pharmaceutical discovery programs is

dependent on two things: (1) having appropriate high quality screens… and (2) having adequate numbers of high quality compounds to screen."

From the beginning of the drug development process some projects lean towards target-to-drug, while others to the drug-to-target approach. The target-to-drug strategy focuses on the potential target first, a vulnerable pathway or enzyme, develops an assay to test inhibition of this isolated target and screens compounds for inhibition or binding. Meanwhile, the drug-to-target approach firstly identifies an antimicrobial compound, then identifies and validates its target and proceeds with lead optimisation (Sala and Hartkoorn, 2011). Both have their short-comings, but the most crucial one is that target-to-drug hit compounds are prone to inactivity *in vivo* because they never engage their target through the numerous barriers (see **Figure 1.5.**), an issue not addressed in assays with purified targets. Purely target-to-drug concepts are considered a failure and have never delivered any drug candidate against *M. tuberculosis* to date (Cole, 2016).

Minimal inhibitory concentrations (MICs) are the most important initial indicators of phenotypic screens in the drug-to-target approach. It is often specified as MIC50 or MIC99 referring to the level of growth reduction (by 50% or 99%) caused by the compound (Franzblau et al., 2012), and it is a preliminary reporter on compound potency.

Although, drug-to-target approach dominates in *M. tuberculosis* research, it is possible to include some aspects from the target-to-drug approach in screening. Hypomorph bacteria, that express less of a protein by repressing the promoter with the TET-OFF system for example (Sala and Hartkoorn, 2011) have an increased sensitivity towards inhibitors of that enzyme or pathway (Cole, 2016). Overexpression of a target results in the opposite effect and cells become tolerant to inhibitors of that specific target. This phenomenon is called the MIC shift. Mt-DprE1 overexpressing *M. bovis* BCG in high-throughput screens (HTS) has

identified a new inhibitor solely based on the shift in MIC comparing overexpressor and wild type strains (Batt et al., 2016), an exemplary strategy conducted in this work.

1.5.2 Whole cell screening design in TB drug development

1.5.2.1 Basic assay requirements

The phenotypic whole cell assay measures *M. tuberculosis* growth inhibition in the presence of a compound. Therefore, essentially the assay requires a suitable bacterium, a medium that promotes cell growth, and a test compound.

The choice of tested strain is an important determinant of whole cell screens. The laboratory strain MTB H37Rv is a widely applied standard and it has its whole genome sequenced (Cole et al., 1998). Unfortunately, it is a slow growing mycobacterium species with an average generation time 24 hours (Cole et al., 1998) that needs handling at biosafety level 3. Clinical isolates require similar treatment and may exhibit unique phenotypes. On the other hand, it may have its benefits to screen against drug resistant strains to narrow down the active compounds that are unaffected by prior resistance mechanisms. To standardise and rationalise, surrogate organisms are often opted to lower the hazard and to speed up culture growth. *M. bovis* BCG has a cell wall structure and similar antibiotic-susceptibility profile that mimics *M. tuberculosis* and needs only biosafety level 2, but unfortunately another slow growing species. *M. smegmatis* and *M. aurum* are fast growing mycobacteria and often substitute *M. tuberculosis* in research, but their sensitivity results are often different from *M. tuberculosis* and *M. bovis* BCG (Franzblau et al., 2012).

Whole cell *M. tuberculosis* screens vary greatly in assay properties depending on which pathophysiological niche is simulated. *M. tuberculosis* survival can be assessed aerobically, which models the actively growing bacteria, while oxygen-depleted and nutrient-starved models aim to model susceptibility in the dormant/NRP subpopulations (Sala and Hartkoorn,

2011). Macrophage assays carry *M. tuberculosis* intracellularly to imitate the microenvironment and drug-permeation in that stage of infection, however the additional advantage of it compared to the extracellular assays is uncertain (Franzblau et al., 2012).

HTS assay media are typically liquid and based on Middlebrook 7H9 with supplements (Franzblau et al., 2012), however solid medium may be used as well. Supplements may be glycerol (0.2%), oleic acid, palmitic acid, albumin, dextrose and Tween-80, amongst others (Franzblau et al., 2012). The surfactant Tween-80 is added to prevent clumping of the cells and keep the culture in planktonic solution. However, it has been shown that Tween-80 can attenuate or even potentiate compounds (Franzblau et al., 2012). The same is true for the carbon source – glycerol is a common supplement, but false positive hits can emerge which are not potent *in vivo*, only due to an artefact of the active glycerol metabolism pathway, which is not the real target of these molecules (Franzblau et al., 2012). Therefore, alternative carbon sources (oleic acid, palmitic acid, acetate) are recommended to supplement the growth medium.

Screening compounds are usually stem from compound libraries. Such libraries are the property of the Molecular Libraries Screening Center Network, the National Institute of Allergy and Infectious Disease, and private pharmaceutical companies like GVKbio (Ekins et al., 2010) or GlaxoSmithKline (GSK). Their libraries are curated by medicinal chemists that assess compounds by physicochemical factors (logP, Lipinski rule of 5, molecular weight) and biological data at service (MICs against different species, models etc.), and also weed out compounds that are typically false hits like the pan assay interference compound structures (Ekins et al., 2010). TB drugs tend to differ from other small drug compounds setting off alarms of usual filters, therefore special evaluation is required.

Current methods attempt to minimise compound amount necessary to assess potency and to speed up the assay process. Therefore, microtiter plates with low culture volumes are generally implemented. For precision compound dispensing special technologies, like the inkjet technology (like HP D300 Dispenser) or acoustic dispensing (like Labcyte Echo) are used for HTS (Marx, 2014). The former is cheaper, but cannot aliquot aqueous solutions or heat sensitive samples (Marx, 2014).

Finally, an often overlooked aspect of assay design is the format in which the HTS is conducted. Assay miniaturisation is desired for little compound and reagent consumption and assaying a multitude of compounds simultaneously. However, MIC values can be different depending on the culture volume and format. In the case of this work the MICs seemed higher in small volumes than in slightly larger format (see **Chapter 4**). Solid and liquid media produce inherently different MICs as well (substance diffusion is different) therefore they cannot be directly compared.

1.5.2.2 Indicators of cell viability

As to the read-out of the whole cell assay, there are various options to determine cell growth. Absorbance, redox reactions and ATP assays are the most common (Franzblau et al., 2012). Optical density measurements need no reagent addition but have low signal to noise ratio. Redox reagents like resazurin and Alamar Blue have a higher redox potential than FAD, NAD^+, $NADP^+$ or even the respiratory chain enzymes so any of them can reduce the blue, barely fluorescent oxidised forms to a reduced, highly fluorescent and pink coloured form (Primm and Franzblau, 2007). Tetrazolium dyes, like MTT ((3-(4,5-dimethyl-2-thiazolyl)-2,5-diphenyl-2H-tetrazolium bromide), are applicable in HTS but do not offer any advantage compared to resazurin, however water-soluble ones are more expensive. Due to their larger size compared to resazurin they need a shuttle for electron transport as well, like menadione

or phenazine methyl sulphate (Primm and Franzblau, 2007). For intracellular ATP measurement commercial kits are available, such as BacTiter-Glo™, which is considered to be sensitive and has the shortest incubation time needed for reading (5 minutes compared to 24 hours with resazurin) (Sala and Hartkoorn, 2011). The limiting factor of this reagent is its price.

Reporter genes (like green fluorescent protein expression) can be used for real-time monitoring of *M. tuberculosis* survival (Franzblau et al., 2012). It has the additional benefit to produce a reliable signal in the intracellular macrophage assay as well and indicate inhibition kinetics.

Another special read-out method is the combination of resazurin read-out with solid media, called charcoal agar resazurin assay (CARA), developed to HTS application (Gold et al., 2015). This method uniquely distinguishes bacteriostatic and bactericidal compounds, and despite the solid media platform, it can be performed in 96-well microplates. The idea is to incubate bacteria cells in liquid culture with the tested compounds, but only part of the test volume is assessed *via* resazurin addition (or optical density). Another part of the culture is transferred to charcoal agar which adsorbs the compound from the culture and allows renewed cell growth if the compound was only bacteriostatic.

1.5.3 Biological hit compound characterisation

1.5.3.1 Toxicity screening of hit compounds

In addition to activity studies hit compounds are early on tested for toxicity. To prevent drug attrition at an early stage of the drug development process hit compounds are profiled *in vitro* against a number of targets that are indicative of adverse drug reactions in the future (Bowes et al., 2012). If a hazard is identified, then the hit can get eliminated or the hazard circumvented during the lead optimisation process and drive hit to lead selection (see **Figure**

1.12. for evolution of a compound in the process). Furthermore, *in vitro* diagnostics can identify specific issues related to human physiology which could pass unnoticed even in animal models.

In vitro toxicity profiling is usually required at the lead optimisation phase. The panel is different for each pharmaceutical research organisation, but Bowes *et al.* summarized a list of 44 targets that should be a minimum requirement for compound safety measurement (Bowes et al., 2012). These targets are receptors, ion channels and transporters. For example, the enzyme monoamine oxidase (MAO) is part of most panels because MAO inhibitors can cause hypertensive crisis and serotonin syndrome.

1.5.3.2 Target validation

A novel mechanism of action is highly desirable from a new drug candidate and target validation can help with lead prioritisation during development. A biochemical assay can confirm that the compound inhibits a specific enzyme function, or spontaneous mutation against the hit occurring in the target gene can indicate target of the inhibitor. Transcription profiling is another option to identify the pathway or pathways affected by the compound. For example, in the elucidation of mechanism of action of BTZ it was helpful to conduct comparative transcriptome studies. They resembled the transcriptome profile during EMB exposure and narrowed down the possible targets to the cell wall arabinan synthesis (Makarov et al., 2009).

It is also possible to label the inhibitor and locate it within the cell by fluorescence or to purify the target in a pull-down method. However, without a known SAR it cannot be excluded that the label will change the protein-inhibitor binding. BTZ043 and PBTZ169 has been linked to the red fluorescent dye TAMRA and was used to verify cell penetration of the compound and target engagement (Neres et al., 2012). Moreover, it was also the basis to

determine the periplasmic localisation of Mt-DprE1 and classify it as a highly vulnerable drug target (Brecik et al., 2015).

1.6 Aims

This thesis aims to support drug development research and basic research of *Mycobacterium tuberculosis* physiology.

The first goal is to gain more information about the drug target Mt-DprE1, in particular how it interacts with substrates and inhibitors. The question is whether the short disordered region of Mt-DprE1 is involved in inhibitor or substrate binding? To this end, a systematic approach was used to substitute the residues of the shorter disordered region and observe the alterations caused in enzyme activity and inhibitor binding. This can be achieved by applying available biochemical assay methods and implement them with novel Mt-DprE1 muteins. Also, in ligand binding experiments by intrinsic tryptophan fluorescence decay measurements. Studying disordered regions is a follow-up to the observation that certain inhibitors showed higher potency in biochemical assays yet interact nearly identically with Mt-DprE1 based on X-ray co-crystallography data.

The second goal is to gain more information about drug target Mt-DprE2. There is no available crystallographic data about the structure of Mt-DprE2, and purification of the enzyme has been challenging. Crystallisation experiments failed mainly because of aggregation, especially in highly concentrated protein solutions which are generally used for crystal growth. Io overcome these problems a ‘tag-fusion’ – developing a 6xHis-SUMO-tagged and MBP-tagged fusion protein was generated of Mt-DprE2. Potential benefits can result from size exclusion chromatography (removing seeds of aggregation) and exploring mild concentration methods instead of high-pressure filtration.

The third goal, with Mt-DprE2 was to design a biochemical enzyme activity assay. Testing Mt-DprE2-targeted inhibitors is currently restricted to assays with radiolabelled substrates and membrane preparations from a surrogate organism (*M. smegmatis*). A direct biochemical assay could utilise absorbance changes of the co-factor NAD and measure inhibition in an end-point or kinetic manner.

Finally, a whole cell high-throughput screen was performed against a compound library to find new hits targeting Mt-DprE2. For this a *M. bovis* BCG surrogate system was used and to test susceptibility towards the TB Box library of GlaxoSmithKline, a collection of 10 000 compounds specifically selected against *M. tuberculosis*.

2 Biochemical studies with DprE1 site-specific mutants

2.1 Introduction

M. tuberculosis DprE1 (Mt-DprE1) is the target protein of intensive new drug development studies (Piton et al., 2017). A better understanding of the physiological function and interaction of the protein with inhibitors may be highly informative in evaluating hits and leads, and potentially improving their affinity to Mt-DprE1. Currently there are 19 crystal structures of *M. tuberculosis* (MTB) Mt-DprE1 deposited in the Protein Data Bank (PDB) (Protein Data Bank, n.d.). The most important protein-ligand co-crystals have recently been summarised by Piton *et al.* (2017). These models have also been implemented in digital docking studies (Pore et al., 2015; Haribabu et al., 2015), however in each case the authors acknowledged the issue of two loop regions, which are disordered and therefore do not appear in the electron density maps, only when certain ligands are also present that they stabilise these regions (see crystal structure in **Figure 1.9.**).

The two disordered regions of Mt-DprE1 are located above the active site, therefore they have the potential to influence protein-ligand interactions. Furthermore, the short disordered region, termed loop II by Piton *et al.* (2017) between amino acids 316 and 330 is highly conserved among mycobacteria (Batt et al., 2012), which might suggest that these residues are important to its physiological function.

To test the exact extent of how these residues may affect enzyme function, in this thesis single-site directed mutants of Mt-DprE1 were constructed, focussing on the short disordered region, and the protein was characterised in enzymatic assays and in ligand binding studies.

2.2 Materials and methods

2.2.1 Site-specific mutant generation

Single amino acid substitutions were introduced into plasmid pCDFDuet-6xHis-Mt-DprE1 *via* the Quikchange protocol (Agilent Technologies, 2011). Complementary mutagenesis primers carried the desired mutations in the central part of the DNA oligos, the plasmid in its whole was amplified by PCR, then the parental (non-mutant) plasmids were digested based on their methylation, leaving only the unmethylated newly synthesised strands. Finally, the mutations were confirmed by DNA sequencing with specific sequencing primers. All primer sequences are listed in the Appendix.

2.2.1.1 Primer design

Mutagenesis primers were designed based on the guidelines in the Quickchange II site-directed mutagenesis kit instruction manual (Agilent Technologies, 2011). The guideline suggests using completely complementary primers, however some PCR failed due to high melting temperatures of the primer pair. Therefore, for some mutants (see example mutant N385A below) partially complementary primer pairs were designed and also the concentration ratio between the forward and reverse primers were altered to prevent primers annealing to each other. The nomenclature of primers and resulting plasmid constructs was based on which amino acid is replaced at which position and by which amino acid.

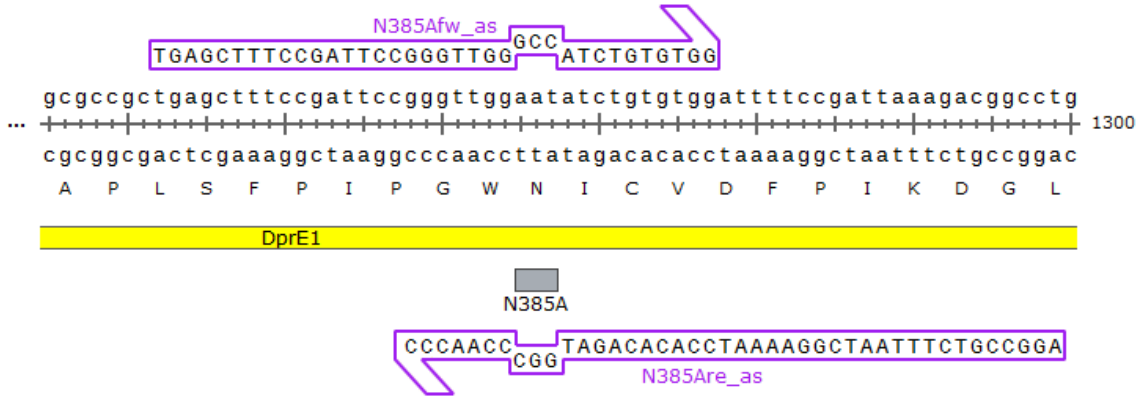


Figure 2.1. Primer design for Mt-DprE1 N385A mutagenesis. Both primers (in purple) carry the substitution, but only partially overlap. Illustration made in SnapGene 1.1.3.

Primer pairs for the mutagenesis were designed with the aid of online programs OligoAnalyzer 3.1 (OligoAnalyzer 3.1, n.d.) and PrimerX (PrimerX, n.d.), and synthesized by Eurofin Genomics Ltd.

2.2.1.2 Polymerase Chain Reaction

PCR amplification was performed in sterile 200 µl tubes. The basic reaction composition worked in most cases, but troubleshooting was necessary, and a modified version implemented in some cases. For example, with DprE1 M319A mutagenesis asymmetric primer concentration ratio and a different polymerase resolved this issue.

Basic PCR composition				Modified PCR composition		
component	final conc.	volume		component	final conc.	volume
template DNA	50 ng	0.1 µl		template DNA	15 ng	3 µl
forward primer	1 µM	2 µl		forward primer	0.1 µM	2.5 µl
reverse primer	1 µM			reverse primer	1 µM	
Pfu DNA polymerase	2.5 unit	1 µl		Phusion® high-fidelity DNA polymerase	~1 unit	0.4 µl
dNTP mixture	0.2 mM each	2 µl		dNTP mixture	0.2 mM each	2.5 µl
10x *Pfu* buffer with $MgSO_4$	1x	2 µl		5x Phusion GC buffer	1x	5 µl
-	-	-		DMSO	3%	0.75 µl
MilliQ water up to final volume: 20 µl				MilliQ water up to final volume: 25 µl		

Table 2.1. Mt-DprE1 mutagenesis PCR conditions.

2.2.1.3 Clone evaluation

PCR products were visualised on agarose gel, then transformed into *E. coli Top10* cells, based on Hanahan method, heat shock at 37 °C for 5 min (Green and Sambrook, 2012) and grown on Luria-Bertani (LB) solid and liquid media supplemented with 100 µg/ml spectinomycin. Plasmids were isolated using a Qiagen Miniprep Kit (Cat No./ID: 27106) and sent to Eurofin Genomics Ltd. for sequencing. The received DNA sequences were analysed with Vector NTI ContigExpress program (Invitrogen).

2.2.2 Protein production of Mt-DprE1 and muteins

Proteins were expressed and purified based on the protocol by Batt et al. (2012), which was optimised for the full length wild type (wt) protein. Modifications were introduced to avoid an additional dialysis step to shorten purification time and to prevent protein denaturation.

The plasmid was heat transformed into *E. coli BL21 (DE3)* cells and co-transformed with chaperon-carrier plasmid pTrc99a-GroES-Cpn60.2 expressing a GroES chaperonin from *E.*

coli and a GroEL chaperonin encoded by *M. tuberculosis* (Batt et al., 2012). Cells were grown as preculture overnight, then diluted 100-fold and grown to log phase (OD_{600} = 0.6) in Terrific Broth (Sigma, #T0918), supplemented with both 100 μg/ml ampicillin and spectinomycin, with constant shaking at 200 rpm. The cultures were then cooled to 20 °C and induced with 0.5 mM isopropyl-β-D-thiogalactoside (IPTG) and incubated overnight. Cells were harvested by centrifugation and washed with 0.85% saline solution and stored at -20 °C. Cells were then resuspended in ice cold lysis buffer (50 mM potassium phosphate pH 8.0, 300 mM NaCl, 10 mM imidazole, 1 Roche EDTA-free cOmplete protease inhibitor tablet/40 ml solution) and sonicated on ice (10 cycles of 20 seconds of treatment and 40 seconds pause between cycles). The soluble fraction was separated by centrifugation at 21,800 *g* for 30 min at 4 °C, filtered through a 0.22 μm syringe filter disk where necessary.

The supernatant was loaded on pre-equilibrated GE HisTrap FF nickel-sepharose column and washed with lysis buffer containing 50 mM imidazole. Then the buffer was changed to storage buffer (50 mM Tris-HCl pH 8.5, 20 mM NaCl, 10% glycerol) and at the same time connected to a GE Healthcare HiTrap DEAE FF weak anion exchanger column directly (pre-equilibration step for the second column). Mt-DprE1 was eluted from the HisTrap column to the HiTrap column by storage buffer containing 300 mM imidazole. The HiTrap column with the sample was further washed and fractions collected using the storage buffer with increasing saline content. Mt-DprE1 typically eluted around 100 mM NaCl concentration. Fractions were evaluated by SDS-PAGE and Mt-DprE1 fractions with acceptable contaminations were concentrated by filtration (Amicon Ultra centrifugal unit, MWCO 10 kDa) to about 10 mg/ml final protein concentration, measured by NanoDrop, ε = 63495 [1/M*cm] and estimated by ProtParam (ProtParam, n.d.) for the wild type protein sequence and stored at -20 °C.

2.2.3 Ligand binding studies with Mt-DprE1 muteins

Ligand-protein interaction was monitored by intrinsic tryptophan fluorescence quenching. In this method the fluorescence signal is altered due to the ligand interacting with the protein and changing the microenvironment of the tryptophan and tyrosine residues (Callis, 2014). The protein in dialysis buffer was kept at 20 °C. Ligand in the same buffer (stock solutions of 50 or 75 μM ligand) were added stepwise to achieve a 0-18 μM final concentration. The quartz cuvette contained 600 μl protein solution at the start of the reaction, volume expansion was corrected for in the fluorescence intensity calculations. Since the ligand stock solutions are in pure DMSO, the working stock may contain DMSO up to 10 % (v/v). DMSO content and volume expansion were the limiting factors in the highest tested concentrations.

Fluorescence spectra were recorded with a PTI QuantaMaster 40 spectrometer (Photon Technology International) set to 0.5 mm monochromator slits and 1 nm steps through reading. The excitation wavelength was fixed through the experiment to 280 nm, the absorbance maximum of tryptophan, and emission spectra scanned from 300 to 400 nm, excitation maximum of proteins usually ranges between 305 and 352 nm (Callis, 2014).

Data of three independent repeat experiments were then analysed using GraphPad Prism (version 6.00 for Windows, GraphPad Software, La Jolla California USA, www.graphpad.com) and data plotted with non-linear regression, 'One-site – specific binding', defining K_d and B_{max} values.

2.2.3.1 Selection of mathematical model for ligand binding

Mt-DprE1 is a monomeric enzyme (Batt et al., 2012) with one expected ligand binding site. Therefore, the expected binding was a specific binding mechanism of enzyme and ligand at a 1:1 molar ratio.

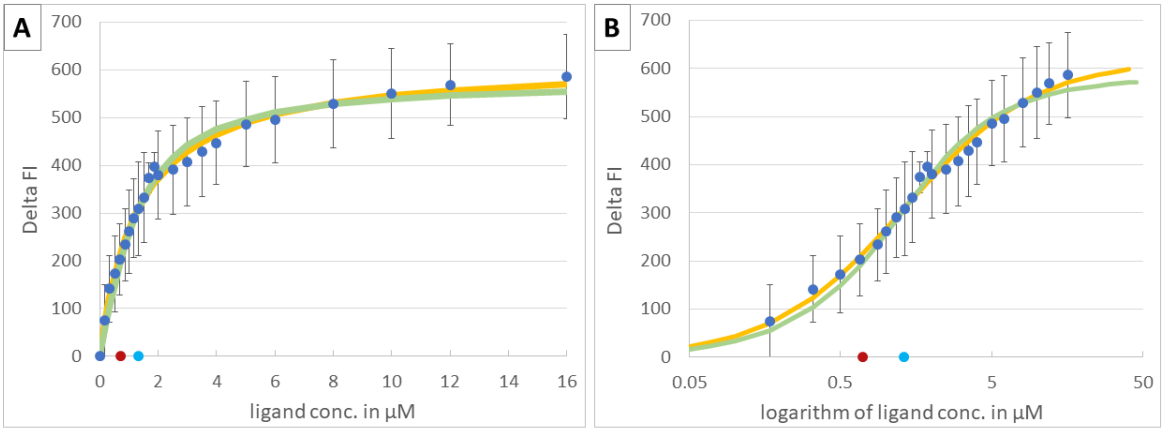
Calculation of the dissociation constant (K_d [μM]) and overall binding capacity (B_{max} [relative fluorescence units]) were conducted in Microsoft Excel with the Solver extension, and were first modelled on the equation below (based on Langmuir isotherm and Lambert-Beer law) (Epps et al., 1999):

$$\Delta F_{emit} \sim [EL] = \frac{B_{max} \times [L]}{K_d + [L]}$$

However, based on the results for K_d the assumption of this equation, that free ligand and total ligand concentrations are approximately the same, are incorrect, because the K_d value is comparable with the enzyme concentration. In this case, the more general quadratic formula for 1:1 complex formation is applicable (Van De Weert and Stella, 2011):

$$\Delta F_{emit} \sim [EL] = \frac{[P]_t + [L]_t + K_d - \sqrt{([P]_t + [L]_t + K_d)^2 - 4 * [P]_t * [L]_t}}{2 * [P]_t}$$

Both calculations were made for comparison, and in the end the simpler equation used for ease of calculation, better use in global analysis with GraphPad Prism, and because in most cases it gave a better fit (sum of square errors was lower) for the measurement points. To illustrate the little difference between the two methods considering the large standard deviation between measurements see **Figure 2.2.** below.



C

	Quadratic	Langmuir
Bmax	581	618
Kd	0.70	1.31
SSE	9528	4559

Figure 2.2. Difference between the simplified Langmuir isotherm-based equation and the quadratic equation calculation. This example is with wt Mt-DprE1 binding 93a3. Blue dots are the averages of four independent measurements with standard deviation bars, green curve is the quadratic equation best fit, the relevant K_d [µM] as a red dot on the abscissa. Orange curve is for the Langmuir-based equation best fit, with the relevant K_d [µM] in light blue on the abscissa. SSE is sum of squared errors. (A) Linear concentration values, (B) with logarithmic concentration values to show that enough measurement points were taken, and (C) summarises characteristic values (B_{max} is in [relative fluorescence units]).

2.2.4 Enzyme activity assays with Mt-DprE1 muteins

Mt-DprE1 is a FAD-dependent oxidase-dehydrogenase enzyme and its activity can be monitored by measuring the reoxidation of $FADH_2$ coupled to the enzyme reaction. The reoxidation event is linked to the redox dye resazurin, illustrated in **Figure 2.3.** This assay was first published by Neres *et al.* (2012). The resazurin assay was further optimised and implemented in a high-throughput screening as an end-point inhibitor assay using the substrate GGPR in GSK Stevenage (unpublished work by Argyrides Argyrou).

Initial efforts aimed to reproduce the assay using DCPIP as an electron acceptor (Neres et al., 2012), using a published protocol measuring Mt-DprE1 enzyme activity. However, this was not successful, as absorbance values were not indicative of any measurable reaction

taking place. DCPIP is expected to turn from blue, with absorbance maximum at 600 nm, to colourless by reduction.

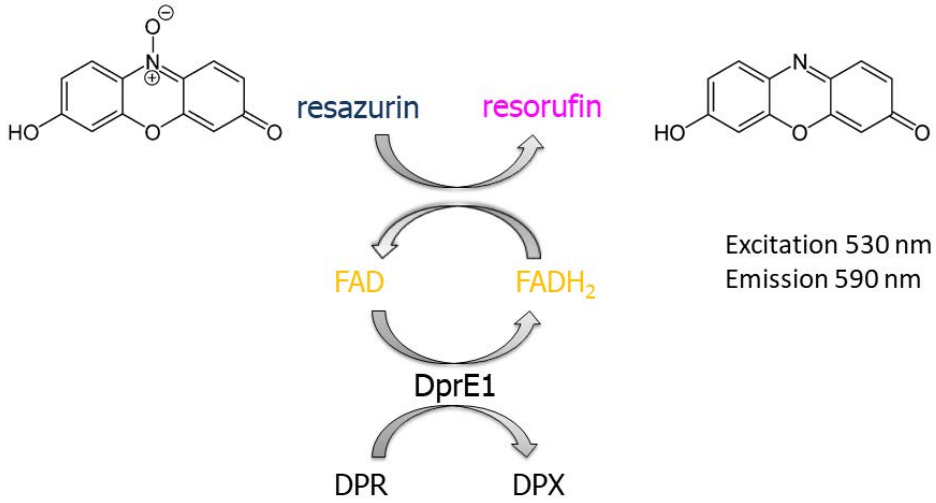


Figure 2.3. Coupling of the Mt-DprE1 enzyme activity to fluorescence read-out signal. Resazurin is blue and shows low fluorescence, while the reduced form resorufin is highly fluorescent and pink in colour. Colorimetric determination is also possible but results lower sensitivity. DPR – decaprenyl-phosphorylribose, DPX – keto-intermediate of the epimerisation reaction.

The native substrate DPR is difficult to obtain, highly amphipathic and forms micelles in a single phase aqueous buffer. Instead, substrates with shorter prenyl chains, carrying the same phosphorylribose moiety as DPR were prepared and evaluated. The shortened prenyl chains were farnesyl (3 isoprenes) and geranylgeranyl (4 isoprenes), with geranylgeranyl being reported to be already long enough to anchor a protein to a membrane surface (Resh, 2015). Both substrate analogues were provided by GSK.

Reaction protocols were based on the published protocol of Batt *et al.* (2016). The reaction volume was 50 µl in 96-well plate format (Greiner Bio-One, black flat-bottom microplate, M9936) in an assay buffer containing 50 mM HEPES pH 7.5, 100 mM NaCl, 1.5 v/v % DMSO, 50 µM Brij-35, 5 µM FAD, 50 µM resazurin, 4 µM bovine serum albumin, substrate and Milli-Q water. Substrates farnesyl- and geranylgeranyl-phosphorylribose (FPR and GGPR) were added from 0 to 1.1 mM and 0 to 180 µM respectively, with reactions

performed in triplicates on the same plate. The plates were then incubated for 15 minutes on 37 °C before measurement, initiating the reaction by the addition of the enzyme, Mt-DprE1 in 2 μM final concentration.

The reaction mixture was incubated at 37 °C throughout and monitored for 35 minutes with GGPR and 40 minutes with FPR. The fluorescence signal was indicative of first order reaction with a linear correlation between signal and time usually in the first six minutes of the measurement. Since the enzyme was added manually with ten measurement points in triplicates with two control reactions, resulting in altogether 26 reactions, with the first approximately two minutes of the reaction not measured.

Each assay included two control reactions, one without enzyme and with 85 μM GGPR or 400 μM FPR concentration (negative controls), and a positive control with wild type Mt-DprE1 and 150 μM GGPR or 700 μM FPR (positive controls).

Data were analysed by Microsoft Excel to calculate reaction rates, and GraphPad Prism for non-linear regression to fit one-site specific binding Hill-slopes to the reaction rates and calculate V_{max} and K_{half} values. For assays where automatic fit was not possible (deemed ambiguous by the software), were processed manually in Microsoft Excel. The Hill equation applied by GraphPad Prism software:

$$v = \frac{v_{max} \times [S]^h}{{K_{half}}^h + [S]^h}$$

V is reaction rate, v_{max} is maximal reaction rate, S substrate concentration, h the Hill slope, and K_{half} the substrate concentration at which the reaction rate is half of the maximal reaction rate.

2.3 Results and discussion

2.3.1 Ligand binding assays

A total of 11 single-site mutants and wild type Mt-DprE1 were expressed and purified. Intrinsic tryptophan fluorescence quenching for these proteins with non-covalent ligands 93a3 and 8287, which are unpublished inhibitor compounds by GSK and bind the active site of the enzyme, were also evaluated. These ligands were chosen because they are known inhibitors of Mt-DprE1 by overexpression MIC shift in whole cell assay (unpublished), and the crystal structure of the enzyme-inhibitor complexes are also solved (unpublished work by Sarah Batt and Klaus Fütterer, University of Birmingham). The inhibitors bind near the FAD co-factor, near the hypothetic substrate-binding site. If the amino acid substitutions affected the microenvironment of the aromatic sidechains of the protein, it would imply conformational changes (Callis, 2014), and calculate parameters from the quenching, such as K_d and B_{max} values. An example of recorded fluorescence spectra is shown in **Figure 2.4.**

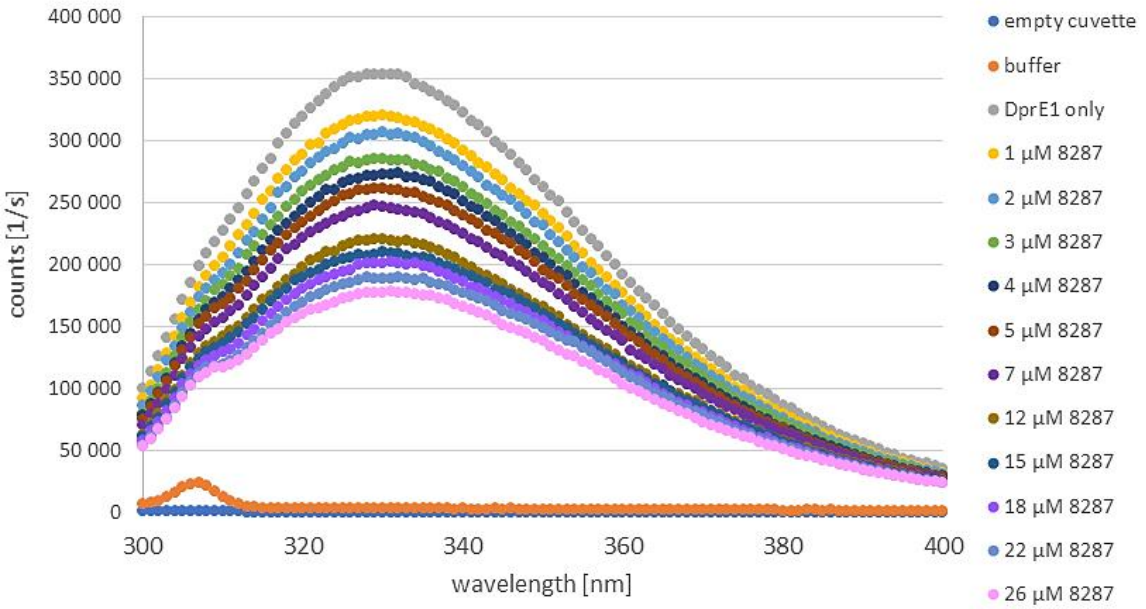


Figure 2.4. Typical intrinsic tryptophan fluorescence spectra collected in ligand binding studies. In the example wt Mt-DprE1 is titrated with ligand 8287. Excitation wavelength: 280 nm, emission maximum at 330 nm. Highest emission curve (in grey) corresponds to the ligand-free Mt-DprE1, decreasing signals to the addition steps of the ligand. Buffer background signal can be seen in orange, near the baseline with the solvent Raman peak (emission maximum around 307 nm).

An interesting observation was that the wavelength of the fluorescence emission peak (λ_{max}) was mostly influenced by not the replacement of the tryptophan at position 323, but the mutants of methionine and phenylalanine at positions 319 and 320 (see **Figure 2.5.**). Tryptophans and tyrosines are considered to contribute the most to the fluorescence intensity, and their immediate environment to influence the most λ_{max} (Callis, 2014). There are eight tryptophans and thirteen tyrosines in Mt-DprE1, and one of each in the short disordered region 316-329 (full enzyme sequence in Appendix).

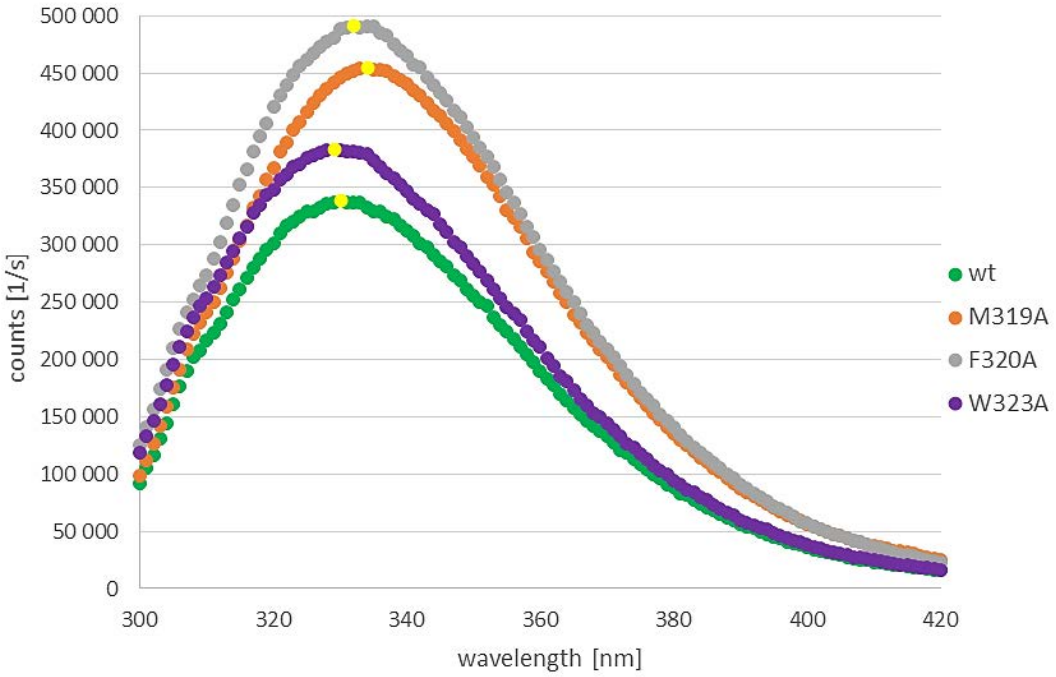


Figure 2.5. Maximum emission peak differences among wt and mutant Mt-DprE1. Bright yellow point labels the maximum value in all measurements. All proteins are 1 μM in the same buffer composition, before any ligand addition.

Ligand titration experiments were analysed to decide if any of the mutations significantly changed ligand affinity of the enzyme. Two different ligands were tested, and only the wild type and Mt-DprE1 W323A with both ligands, in the rest of the experiments only one ligand was measured. The primary aim of the experiment was to identify mutants with significantly different K_d values from the wild type, which would suggest an active participation from that residue in ligand binding and indicate the importance of the short disordered region in ligand binding.

construct	K_d	K_d SE	B_{max}	B_{max} SE
wt	1.28	0.39	580	63
P316G	2.02	0.23	672	34
L317A	1.02	0.24	504	41
D318A	1.07	0.14	628	28
M319A	0.93	0.13	590	26
F320A	0.78	0.11	479	21
E322A	0.91	0.14	495	26
W323A	1.70	0.43	450	43

Table 2.2. Ligand binding analysis for Mt-DprE1 carrying point mutations in the mobile loop straddling across the active site with 93a3. Data were analysed by GraphPad Prism one-site ligand binding regression. K_d is dissociation constant [μM], B_{max} is maximal change in fluorescence [relative fluorescence unit], SE is standard error.

construct	K_d	K_d SE	B_{max}	B_{max} SE
wt	2.78	0.72	393	44
W323A	1.94	1.02	314	66
N324A	2.52	0.63	304	30
Y327A	1.73	0.72	303	44
G328P	4.87	-	362	-
P329A	4.47	1.08	374	50

Table 2.3. Ligand binding analysis for Mt-DprE1 carrying point mutations in the mobile loop straddling across the active site with 8287. Data were analysed by GraphPad Prism one-site ligand binding regression. K_d is dissociation constant in [μM], B_{max} is maximal change in fluorescence [relative fluorescence unit], SE is standard error.

The measurements revealed no obvious changes in affinity, the largest deviations were still within range of analytical error (three-times the standard error of the wild type K_d, which means a K_d greater than 2.45 μM for 93a3, and 4.94 μM for 8287).

It can be deduced that between the two ligands 93a3 has a higher affinity to the enzyme (lower K_d values) than 8287, and it is possible, that the higher B_{max} values indicate a larger change in protein conformation. Although the latter is a speculation, since the fluorescence quenching could be due to both spatial protein backbone changes, as well as other qualitative changes in the conformation.

2.3.2 Enzyme activity assays

To test the importance of the short disordered region (residues 316-329) of Mt-DprE1 enzyme function, an activity assay based on the published protocol of Batt *et al.* (2016) was used. The physiological substrate DPR, would form micelles in aqueous solution, and could only be examined in a well-controlled membrane-solution interface. Alternatively, the strategy pursued in these studies takes advantage of truncated substrates (FPR and GGPR).

Mt-DprE1 catalytic activity does not follow Michaelis-Menten kinetics. Instead, it shows allosteric enzyme kinetics. This has been observed before by Neres *et al.* (2012), and also in our research laboratories in Birmingham and at GSK (Batt et al., 2016).

With both substrates the enzyme kinetics shaped a sigmoidal curve (see **Figure 2.6**), however turnover of the two differed: GGPR added to the reaction seemed to result in much higher resorufin formation than FPR. However, when continuing the experiments with a second batch of FPR, there was an 8-fold increase in fluorescence values and change in K_d values as well between the FPR batches. Considering the scarcity of the substrates available (only one batch of GGPR, and two batches of FPR) fluorescence values were used in the data analysis (instead of calculating turn-over numbers), compared to enzyme characteristic values K_{half}, V_{max} and h only relative to the wild type values performed with the exact same substrate batch.

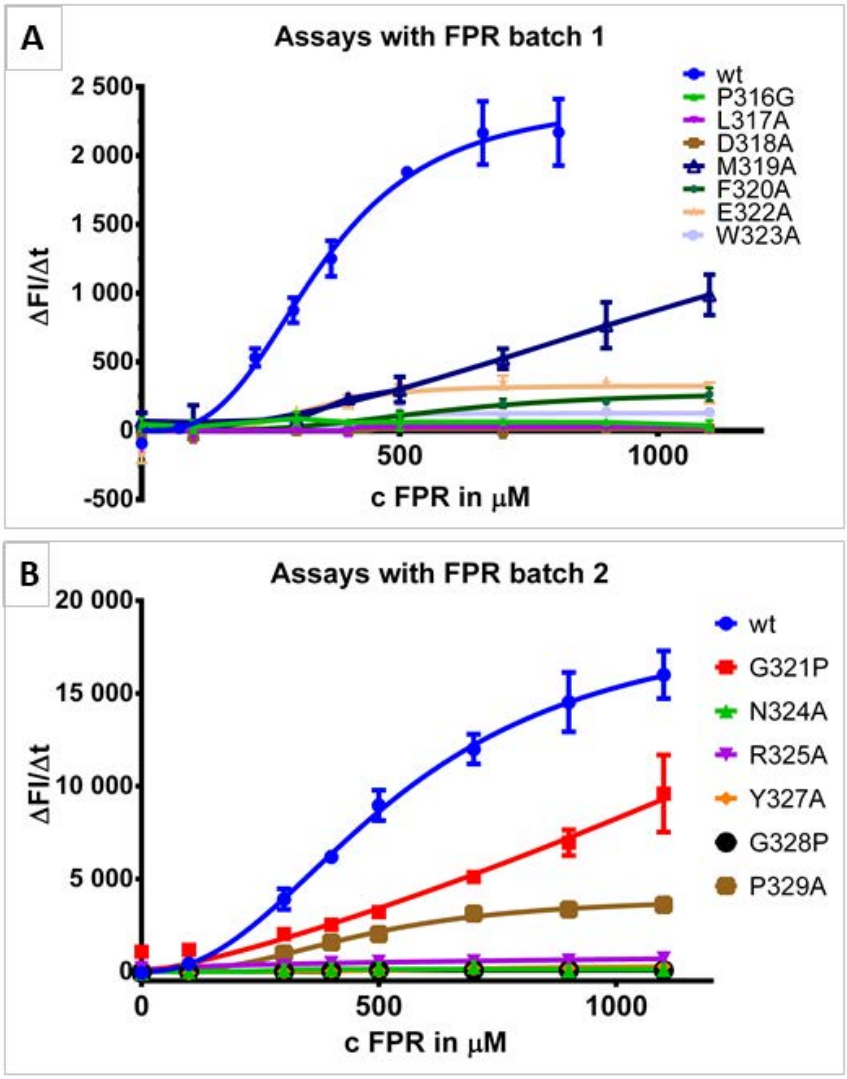


Figure 2.6. Kinetic allosteric sigmoidal fit curves of FPR batch 1 with wt and mutant Mt-DprE1. Error bars are the standard errors for three repeat measurements.

The mutants grouped into three categories regarding activity. Complete inactivation was an arbitrary distinction with more than a 90% decrease in V_{max} and highlighted in red. Active mutants have V_{max} more than 50% of V_{max} of wt Mt-DprE1 and are highlighted in green, with others termed 'low activity' highlighted in yellow (**Tables 2.4. and 2.5.**).

	wt	P316G	L317A	D318A	M319A	F320A	E322A	W323A
Best fit values		*	Ambiguous					
V_{max}	2406		27.92	12.62	2715	281.1	327.4	126.3
h	3.034		~ 86,84	101	1.92	3.371	4.209	10.82
K_{half}	347.9		~ 407,9	446.3	1469	571.9	332	414.8
Std. Error								
V_{max}	136.2		18.85	10.58	3038	40.38	35.91	18.71
h	0.4207		~ 2,898e+008	834367	0.6139	0.9258	2.571	11.83
K_{half}	19.19		~ 2,655e+007	407532	1445	72.5	39.75	40.43
Goodness of Fit								
Degrees of Freedom	21		20	21	21	21	21	21
R square	0.9803		0.1205	0.1166	0.9271	0.9273	0.8106	0.6793

Table 2.4. Allosteric sigmoidal curve fit values with FPR batch 1 in GraphPad Prism. V_{max} is maximal reaction rate [fluorescence intensity unit/second], K_{half} is substrate concentration at 50% of active sites occupied [μM], h is unitless. Active enzymes were highlighted in green (V_{max} > 1/2 of V_{max} wt), inactive in red (V_{max} < 1/10 of V_{max} wt), and low enzyme activity in yellow (1/10 < V_{max} <1/2 of wt). K_{half} is highlighted for mutant M319A to emphasise the lower affinity to FPR compared to any other mutant. Mutant P316G was completely inactive. 'Ambiguous' labels curve fits that have very wide confidence intervals.

	wt	G321P	N324A	R325A	Y327A	G328P	P329A
Best fit values		Ambiguous		Ambiguous		Ambiguous	
V_{max}	19 316	~ 1,794e+007	134.5	~ 71 233	375.7	74.16	4 019
h	2.247	1.271	7.882	~ 0.4192	3.041	~ 100.7	2.636
K_{half}	549.8	~ 421 462	290.2	~ 6.850e+007	796.1	~ 400.4	473.4
Std. Error							
V_{max}	1720	~ 2.227e+010	24.64	~ 1.636e+007	266.9	12.93	278.7
h	0.2978	0.4852	19.23	~ 0.6696	2.014	~ 8.190e+009	0.3952
K_{half}	54.5	~ 4.125e+008	58.88	~ 3.909e+010	439.6	~ 2.870e+007	33.98
Goodness of Fit							
Degr. of Freedom	21	21	21	21	21	21	21
R square	0.9848	0.9223	0.4952	0.5563	0.7509	0.5261	0.9788

Table 2.5. Allosteric sigmoidal curve fit values with FPR batch 2. Highlighting as in **Table 2.4.** Mutant R325A appears to have a high V_{max}, but actual activity is very low. Active enzymes were highlighted in green (V_{max} > 1/2 of V_{max} wt), inactive in red (V_{max} < 1/10 of V_{max} wt), and low enzyme activity in yellow (1/10 < V_{max} <1/2 of wt).

In summary, all mutants are less active than the wild type (V_{max} is lower) in the FPR assay, most of them are inactive. In the case of Mt-DprE1 M319A the K_{half} shifted significantly, while this parameter is hardly comparable in most other cases, with curve fitting barely possible (R square measures were below 0.6 with 6 mutants).

To confirm substrate inhibition by GGPR, a reaction with and without the addition of 400 μM geranylgeranyl-pyrophosphate (GGPP) was set up, and reaction rates measured. It showed lower enzyme activity in presence of GGPP (see **Figure 2.7.**)

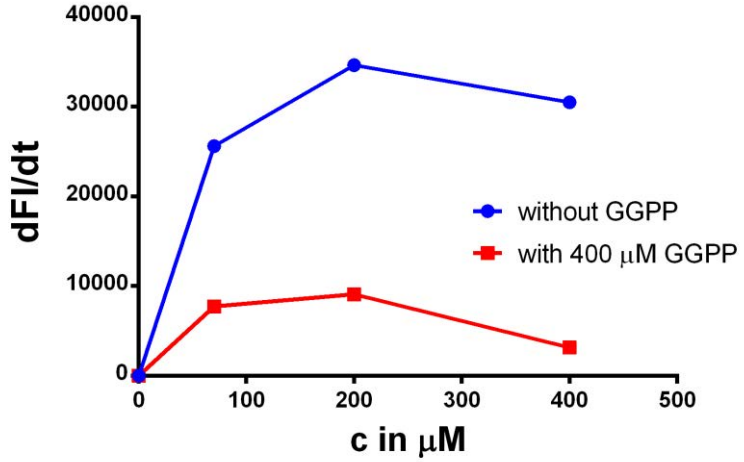


Figure 2.7. The effect of additional GGPP on 6xHis-Mt-DprE1 activity reaction rates with GGPR. Single shot experiment, graph connects measurement values, not a fitted curve. FI is relative fluorescence intensity. The reagent GGPP was kindly provided by James Harrison, University of Birmingham.

Reaction rates were also determined with GGPR. However, it is important to point out that GGPR due to its amphipathic character unfolds DprE1 above a certain concentration, about 220 μM substrate concentration. (This could be further examined and confirmed with NMR experiments in the future.) Therefore, enzyme activities were only measured below this critical concentration range, only until 180 μM and V_{max}, K_{half} extrapolated from this data. This is a limitation to all the results of these characteristic values describing DprE1 in the GGPR assay, therefore the listed V_{max} and K_{half} are all apparent values true to these measurement conditions.

	wt	P316G	L317A	D318A	M319A	F320A	G321P	E322A	W323A	N324A	R325A	Y327A	G328P	P329A
Best-fit values											*			
V_{max}	41 095	410.3	933.2	200.8	37 805	49 243	50 541	25 763	8 047	5 186		748	561.2	13 871
h	4.927	2.699	6.005	13.73	5.103	5.272	8.128	7.646	7.391	8.011		8.609	8.271	5.886
K_{half}	101.5	69.41	97.07	73.89	92.96	139.2	117.6	72.89	131.4	102.1		100.7	96.16	108.4
Std. Error														
V_{max}	1965	92.69	71.42	21.59	807.1	3743	3044	555.8	456.6	304.3		31.41	37.98	643.3
h	0.587	1.548	1.539	11.2	0.332	0.474	1.479	1.074	0.835	1.732		1.433	2.291	0.698
K_{half}	3.094	16.21	4.659	4.933	1.298	5.22	3.404	1.177	3.277	3.248		2.259	3.63	2.892
Goodness of Fit														
Degr. of freedom	21	21	19	21	21	21	21	21	21	21	21	21	21	21
R square	0.9795	0.651	0.917	0.58	0.994	0.99	0.959	0.9815	0.9842	0.935		0.96	0.896	0.979

Table 2.6. Allosteric sigmoidal curve fit values with GGPR in GraphPad Prism. Asterix signs that fitting was not possible with DprE1 R325A, the enzyme had low activity (but not inactive). Active enzymes were highlighted in green (V_{max} > 1/2 of V_{max} wt), inactive in red (V_{max} < 1/10 of V_{max} wt), and low enzyme activity in yellow (1/10 < V_{max} <1/2 of wt).

Accepting the limitations of the GGPR assay and considering its results comparable with the FPR assay reaction rates, we can observe that the wild type enzyme has higher affinity to GGPR than to FPR (lower K_{half} values). However, K_{half} values are to be handled here with caution, because based on the two FPR batches would have different affinities to wt Mt-DprE1. Also, the fluorescence values are higher with GGPR than any of the FPR batches, which is difficult to elucidate theoretically, but as a more practical aspect the GGPR assay has a better signal-to-noise ratio than the FPR assay, which is advantageous in miniaturisation of the high-throughput screen.

Another minor difference between FPR and GGPR enzyme activity tests is that with GGPR only five mutants were inactive, while with FPR there were eight (**table 2.7.**). This can be explained by the aforementioned better signal-to-noise ratio. Also, since the activity scale was arbitrarily chosen, other thresholds could possibly deliver the same activity results with the two substrates.

In general, it seems that the GGPR assay is more robust, but it would be presumptuous to conclude that it is a more relevant model of the enzyme mechanism than the FPR assay.

However, for inhibitor testing it is preferable, because GGPR is more stable in buffer solutions, FPR once dissolved in buffer could not be frozen again or reused later, while GGPR could withstand many freeze-thaw cycles.

To better understand the consequences of single amino-acid substitutions in Mt-DprE1 **Table 2.7** showcases the assigned enzyme activities – active, low activity, inactive with colour-scale with both substrates, along with the level of evolutionary conservation of the specific residues.

activity	P316G	L317A	D318A	M319A	F320A	G321P	E322A	W323A	N324A	R325A	A326	Y327A	G328P	P329A
with FPR														
with GGPR														
sequence alignment														
M. tuberculosis P9WJF1	P	L	D	M	F	G	E	W	N	R	A	Y	G	P
M. smegmatis A0R607	P	L	D	M	F	G	E	W	N	R	A	Y	G	S
M. leprae Q9CDA4	P	L	D	M	F	G	E	W	N	R	A	Y	G	A
Corynebacterium glutamicum A0A161JNA3	P	L	D	L	I	G	E	W	N	R	G	Y	G	S
Rhodococcus sp. PBTS 2 A0A143QI70	P	L	D	M	F	S	E	W	N	R	A	Y	G	S
Nocardia terpenica A0A291RC15	P	L	D	M	M	G	E	W	N	R	G	Y	G	S
Gordonia phthalatica A0A0N9N880	P	L	D	L	F	G	N	W	N	R	A	Y	G	S
Streptomyces regensis A0A0J8AGH2	M	L	D	L	V	G	E	W	N	R	G	Y	G	S
Pseudomonas fluorescens A0A0B7DH09	P	L	D	R	I	E	H	W	N	R	I	Y	G	R

Table 2.7. Similarities between enzyme activity results and residue conservation among example species of the order of *Actinobacteria*. *Pseudomonas fluorescens* is a Gram-negative non-related species with a published probable *dpre1*. Below the name of the species UniProt (UniProt Consortium, 2018) accession numbers are shown. Highlighting as in T**able 2.4.** Grey colour depicts amino acids that are identical with the residues in *M. tuberculosis*. Active enzymes were highlighted in green (V_{max} > 1/2 of V_{max} wt), inactive in red (V_{max} < 1/10 of V_{max} wt), and low enzyme activity in yellow (1/10 < V_{max} <1/2 of wt).

It can be deduced that the substitutions in Mt-DprE1 that caused enzyme inactivation are the most conserved residues between related species. Also, since alanine 326 was not mutated to any other amino acid, it can be predicted that a mutation to glycine on that position would

maybe lower enzyme activity, but most likely it would not inactivate the enzyme, since some orthologs of Mt-DprE1 have that amino acid substitution.

It is interesting to note, that in the case of proline 329, it is often replaced by other amino acids, even alanine in the orthologous enzymes (**Table 2.7.**), yet single amino acid substitution to alanine lowered enzyme activity significantly.

In conclusion it can be said that there is a correlation between Mt-DprE1 mutant enzyme activity and evolutionary conservation in residues, but one does not strictly determine the other.

A possible explanation to sigmoidal kinetics in monomeric enzymes is a gateway function of some flexible parts of the protein. A lid-like motion of mobile loops near the active site could limit active site access. When the time-scale of lid motion is comparable to the enzyme turn-over, it can without actual cooperativity exhibit allosteric enzyme kinetics (Porter and Miller, 2012). It is possible, that the short disordered region 316-329 limits enzyme activity in a lid-like manner, causing the sigmoidal substrate-response kinetic curve. Following this hypothesis, the enzyme activity results with single site mutants highlight two distinct hinge-like parts to the "lid", residues 316-318 and 327-328, which cannot be replaced without major hindrance of enzyme function.

It is possible that the loop interacts with the prenyl-chain of the substrate, and that the geranylgeranyl group contributes to the interaction more like the native substrate than the farnesyl substrate. To this end, crystallisation trials were set up with inactive mutants Mt-DprE1 K367A published as the inactive mutant of ortholog K425A DprE1 *M. smegmatis* (Neres et al., 2012) and Mt-DprE1 K418A (unpublished inactive mutant by Sarah M. Batt, Universityy of Birmingham) with GGPR. Unfortunately, the achieved crystals did not

diffract, but it is an experiment of interest that could be repeated or performed with other mutants in the future, that could give an insight into enzyme-substrate interaction.

3 Expression and purification of Mt-DprE2

3.1 Introduction

To characterise and perform enzymatic assays with Mt-DprE2, large quantities of enzymologically active, purified protein is needed. Literature on Mt-DprE2 reports genetic analyses (Sridhar et al., 2016), whole cell experiments (Kolly et al., 2014) and *in silico* modelling (Bhutani et al., 2015). However, the production of pure, enzymatically active recombinant Mt-DprE2 has not yet been reported. Attempts have been made to produce recombinant Mt-DprE2 in *E. coli*, but purified Mt-DprE2 forms large aggregates, as is evident from ultracentrifugation experiments, and it did not crystallise (unpublished work by Sarah M. Batt, University of Birmingham). This body of work aims at developing a biochemical assay reporting on Mt-DprE2 activity and identification of inhibitors.

3.2 Materials and methods

3.2.1 Cloning

3.2.1.1 pMal-p2x-MBP-DprE2

Commercially available pMal-p2x was inserted with *Mt-dprE2* utilising XmnI and HindIII restriction endonuclease sites. The plasmid design utilises XmnI blunt-end cleaving restriction endonuclease, using any other site in the multicloning site would add additional amino acids into the overexpressed fusion protein between the Xa protease cleaving site in the MBP-tag and the target protein, Mt-DprE2. Therefore, the type IIS restriction enzyme SapI was used to create a sticky end at the N-terminus of the insert with the required amino acid code, producing the required blunt end by Klenow-fragment treatment.

Prior to inserting *Mt-dprE2* in the pMal-p2x vector, the innate SapI recognition site was silenced in the pET-6xHis-SUMO-DprE2 vector by Quikchange mutagenesis (Agilent Technologies, 2011), then *Mt-dprE2* gene was amplified by PCR. The amplicon was cleaved with SapI and purified using the QIAquick PCR Purification Kit (Qiagen, ID: 28106), then treated with Klenow-fragment of DNA polymerase according to manufacturer's protocol (New England Biolabs, #M0210S). The amplicon was cleaved with HindIII and ligated with XmnI-HindIII double digested pMal-p2x vector with T4 DNA ligase (New England Biolabs, #M0202S). The cloning process is depicted in **Figure 3.1.**

The final construct was transformed into *E. coli* Top10, isolated using the Qiagen Miniprep Kit (ID: 27106), performed diagnostic digest with HindIII and NdeI. Selected clone sequences were confirmed by Sanger sequencing by Eurofin Genomics Ltd.

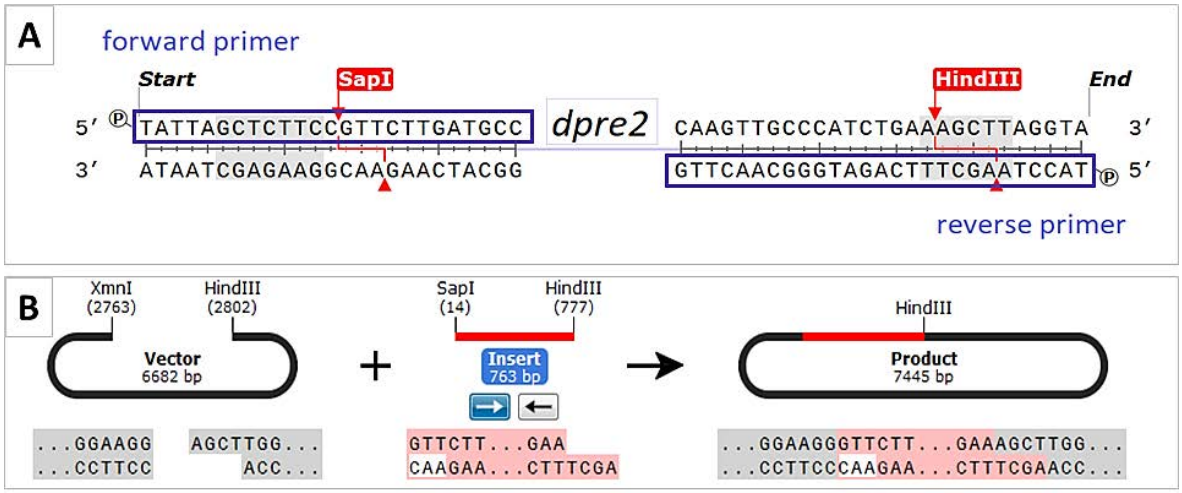


Figure 3.1. The cloning process of pMAL-p2x-MBP-Mt-DprE2 plasmid. A) The PCR amplicon of *Mt-dprE2* gene. In blue frame forward and reverse primers (with complementary strands), enzyme recognition sites in grey, overhangs after enzyme cut are signalled in red. GTT base triplet of the SapI overhang is the first code of *Mt-dprE2*. B) Ligation of vector and insert. The final product does not contain any SapI or XmnI restriction endonuclease recognition sites.

3.2.1.2 6xHis-SUMO-DprE2

The gene *Mt-dprE2* was amplified from pET-6xHis-Mt-DprE2 (construct made by Sarah M. Batt, University of Birmingham) and inserted into pET-6xHis-SUMO vector (modified plasmid provided by Dr. Patrick Moynihan, University of Birmingham) using the BamHI and HindIII restriction endonuclease sites. A two-step PCR was performed with 52 °C and 63 °C annealing temperatures, both vector and insert digested and visualised on an agarose gel. Corresponding bands were extracted using the QIAquick Gel Extraction Kit (Qiagen, #28706) and ligated with T4 DNA ligase. The resulting plasmid was tested by restriction digest with NotI and verified by Sanger sequencing by Eurofin Genomics Ltd.

3.2.1.3 Tag-free Mt-DprE2

In order to perform co-expression studies with Mt-DprE1 from separate plasmids, tag-free Mt-DprE2 was required. Consideration of selectable markers suggested the use of a pET-6xHis-SUMO-Mt-DprE2 vector, as it carries kanamycin resistance. The 6xHis-SUMO-tags were deleted by Quikchange mutagenesis (Agilent Technologies, 2011), primers are listed in Appendix.

The resulting clone was digested with NotI and BglI restriction endonucleases and visualised on an agarose gel, clones confirmed by Sanger sequencing by Eurofin Genomics Ltd.

3.2.2 Mt-DprE2 protein purification

3.2.2.1 Periplasmic MBP-Mt-DprE2

This experiment aimed at producing Mt-DprE2 with a chaperonin-like tag that could be separated from contaminating proteins in fewer steps than with cytosolic expression. This protocol is based on the periplasmic extraction described by Chen *et al.* (2004). A new transformation of pMAL-p2x-Mt-DprE2 plasmid into *E. coli* BL21 was performed, colonies selected on 100 μg/ml ampicillin supplemented LB agar. Bulk cultures (1 L) were inoculated

with a 10 ml overnight culture and grown to an OD_{600}=0.8 at 37 °C while shaking at 180-200 rpm. The cultures were than cooled to 18 °C and 0.3 mM final concentration of IPTG was added to induce overexpression and incubated overnight. Cells from a 1.25 L culture were harvested by centrifugation, washed with 0.85% ice-cold saline solution, and pellets stored at -20 °C until use.

Pre-treatment of cells consisted of resuspension in 30 ml of ice-cold 5 mM $CaCl_2$ solution, incubated on ice for 10 min, cells were pelleted by centrifugation and the supernatant discarded. Then in the high osmotic pressure step the pellet was resuspended in 40 ml shrinker buffer, 20 w/v% sucrose, 50 mM Tris-HCl pH 8.0, 5 mM ethylenediaminetetraacetic acid (EDTA), and incubated on ice for 10 min, cells were pelleted and the supernatant discarded. Finally, in the osmotic shock step the pellet was resuspended in 50 ml 5 mM $MgSO_4$ solution and incubated on ice for 10 min. Cell debris was separated by high speed centrifugation (15,000 rpm, 4 °C, 30 min) and the supernatant kept, and the pellet discarded. The supernatant was adjusted to the column buffer composition, 50 mM Tris-HCl pH 7, 400 mM NaCl for affinity chromatography.

An MBP-Trap column (GE Healthcare, #28-9187-78) was pre-equilibrated with column buffer and the sample loaded at 1 ml/min by a peristaltic pump. The column was washed with 40 ml buffer, 50 mM Tris-HCl pH 7.5, 200 mM NaCl, and the fusion protein eluted with 15 ml of wash buffer with 10 mM maltose, then dialysed to 50 mM Tris-HCl pH 7.8, 150 mM NaCl, 10% glycerol using a 15 kDa dialysis tube (Sigma #D0530), and then concentrated to 1 ml through ultrafiltration (Amicon® Ultra-15 PLGC Ultracel-PL membrane, MWCO 10 kDa) and stored fat -20 °C.

3.2.2.2 6xHis-Mt-DprE2 production

E. coli BL21 cells were heat-transformed with pET-6xHis-SUMO-Mt-DprE2 plasmid and grown overnight at 37 °C in LB broth supplemented with 50 μg/ml kanamycin. Bulk cultures (1 L) were inoculated with 10 ml overnight culture and grown at 37 °C with shaking to OD_{600}=0.4-0.8. Then cultures were cooled to 20 °C, overexpression induced with 0.5 mM IPTG and continued overnight. Cells were harvested by centrifugation, washed with 0.85% saline solution, and pellets stored at -20 °C until further use.

Protein was extracted from a 2.5 L culture cell pellet by resuspending in 60 ml of ice cold lysis buffer: 100 mM NaH_2PO_4 pH 8.0, 250 mM NaCl, 0.05% (3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate, (CHAPS, Sigma #C3023), 10 mM imidazole pH 8.0, 10 mM (tris(2-carboxyethyl)phosphine hydrochloride, (TCEP, Sigma #C4706), 1 tablet of cOmplete Mini Protease Inhibitor Cocktail (Roche #11836153001) and sonicated (10 s input, 20 s pause on ice, for 10 cycles). Following cell lysis all extraction steps were conducted on ice. Cell debris was separated by high-speed centrifugation (21,800 x *g*, 4 °C, 30 min), the supernatant collected and filtered through a 0.22 μm syringe filter disk where necessary.

A HisTrap column (1 ml, GE Healthcare #17524701) was pre-equilibrated with lysis buffer and loaded with the protein sample at 1 ml/min *via* a peristaltic pump. Alternatively, 1.5 ml of Talon resin slurry (Clontech, #635503) was used instead of a prepacked column, washing steps performed similarly, separating the resin from supernatant by low speed centrifugation. The column was washed with 20 ml of lysis buffer, followed by 20 ml of lysis buffer with 40 mM imidazole, then the buffer changed to 20 ml of column buffer (100 mM Tris-HCl pH 8.5, 200 mM NaCl, 5% glycerol) and eluted in approximately 15 ml of elution buffer (column buffer containing 300 mM imidazole, pH 8). Overnight dialysis was avoided to

prevent possible protein aggregation and/or degradation by eliminating imidazole by concentrating and diluting the sample repeatedly three times into storage buffer (50 mM Tris-HCl pH 8.5, 50 mM NaCl, 10% glycerol). The final volume was 1 ml to 1.5 ml, and the sample was stored at -20 °C until further use. A 2.5 L culture of bacteria yielded approx. 6-10 mg of pure protein using this methodology.

It was observed that ultrafiltration partially precipitated the protein (visible aggregate membrane formed over the filter), therefore a possible alternative would be to elute the protein from Talon or a Ni-NTA bead resin in a small elution volume and replace the buffer by PD-10 desalting columns (GE Healthcare #17085101 or similar).

For size exclusion chromatography (SEC) the eluted sample was dialysed (Sigma #D0530, MWCO 15 kDa) overnight at 4 °C in SEC buffer (60 mM Tris-HCl pH 8, 200 mM NaCl). Overnight dialysis with lower salt concentrations lead to a complete precipitation of the protein, repeatedly.

3.2.2.3 Small-scale protein expression trials

Small-scale trials were conducted to confirm protein expression levels and select preferable co-expression partners, for example whether to include chaperon proteins. Chemically competent *E. coli* cells were transformed by heat shock (Green and Sambrook, 2012) with selected plasmids. Clones were selected on LB agar supplemented with appropriate antibiotics (100 µg/ml of streptomycin or ampicillin, kanamycin at 50 µg/ml final concentration). Plates were incubated at 37 °C in a static thermostat, single colonies spotted and used for inoculation of 5 ml LB broth aliquots. These small batches were incubated at 37 °C while shaking (180-200 rpm) until OD_{600} = 0.6, then overexpression induced with 1 mM IPTG and continued at 18 °C overnight.

Cells were harvested by centrifugation and washed with ice cold 0.85% saline solution, then transferred to clean 2 ml test tubes. Cells were then pelleted and resuspended in base buffer (50 mM Tris-HCl pH 8, 50 mM NaCl, 5% glycerol). 0.5 g of glass beads were added, then cell lysis achieved using a bead beater FastPrep-24 (MP Biomedicals), with the recommended program for *E. coli*, the program repeated three times. Cell debris was separated *via* centrifugation at a maximum speed in a bench-top centrifuge (13,000 rpm, 2 min, room temperature). 40 µl of a pre-equilibrated Ni^{2+} or Co^{2+}-covered bead slurry was added (Ni-NTA agarose Qiagen #30210, Talon metal affinity resin Clonetech #635502) and samples were incubated on ice for 30 minutes with manual gentle mixing, periodically. After binding the beads were separated by centrifugation at low speed (5000 rpm, 2 minutes, room temperature) and washed with base buffer supplemented with 20 mM imidazole. The supernatant was removed and washing step repeated twice. Finally, the beads were collected at low speed centrifugation and all samples were kept on ice until use and analysed by SDS-PAGE.

The affinity chromatography resin beads were treated as regular protein samples for heating with loading dye and loaded onto a precast gel. The SDS content of the loading dye (Green and Sambrook, 2012) elutes the proteins from the beads. When loading on the gel, the 20 µl pipette tips needed to be cut with a scissor for enlarging the tip for the slurry. Beads were visible on the gel but remained in the pocket, however the bound proteins penetrated the acrylamide matrix for separation.

3.2.2.4 Mt-DprE1-Mt-DprE2 co-purification

Co-purification of Mt-DprE1 and Mt-DprE2 have been attempted in three different setups that are summarised in **Table 3.1.** All constructs share the principle of utilising the 6xHis-tag fused to Mt-DprE1 and co-purify a tag-free Mt-DprE2 that associates with Mt-DprE1 as

hypothesised and observed elsewhere (two-hybrid system in *C. glutamicum* in (Jankute et al., 2014), preliminary studies in theses by Alice Moorey and Luke Broadbent, University of Birmingham).

Protein	**6xHis-Mt-DprE1**	**Mt-DprE2**	**chaperons**	**6xHis-Mt-GtrA-6xHis**
coding gene	*rv3790*	*rv3791*	*groEL2 (MTB)* *groES (E. Coli)*	*rv3789*
	plasmid that carries this protein:			
construct 1	pCDFDuet (2 cassettes)		pTrc99a	pET
construct 2	pCDFDuet (2 cassettes)		pTrc99a	-
construct 3	pCDFDuet	pET	pTrc99a	-

Table 3.1. Mt-DprE1-Mt-DprE2 co-expression constructs grouped by plasmids carrying targeted proteins. Plasmids are labelled by commercial vector name without inserted target protein. For pCDFDuet the selection marker is streptomycin/spectinomycin, for pET kanamycin, for pTrc99a ampicillin.

3.2.3 Analytical size exclusion chromatography (SEC)

Size exclusion chromatography was performed using a Superdex 16/600 200 pg (GE Healthcare, 28-9893-35), with column volume 120 ml and maximum recommended purification sample of 4 ml and 22 mg protein, according to manufacturer's manual. The column was washed with one column volume MilliQ (doubly deionised) water and equilibrated with one column volume of buffer (40 mM Tris-HCl pH 8, 200 mM NaCl for 6xHis-SUMO-Mt-DprE2, and 40 mM Tris-HCl pH 8.5, 20 mM NaCl, 10% glycerol for Mt-DprE1-Mt-DprE2 co-expresssion).

The size exclusion run used a ÄKTA pure chromatography system (GE Life Sciences). The sample was filtered (0.22 μm) before use and 3.6 ml injected in the sample loop, which was emptied with continuous pressure and 5 ml buffer onto the column. Flow rate was set to 1 ml/min, the process took place in a cold room (4-6 °C). Fractions of 1.8 ml were collected

by the system in a deep well 96-well plate (after 0.15 CV delay for collection). Absorbance chromatograms were recorded at 280 nm (protein maximum absorbance) and 340 nm (NADH maximum absorbance). Fractions were supplemented with 0.2 ml of glycerol for cryoprotection and stored at -20 °C until use.

Calibration of SEC measurement was performed with bovine serum albumin (Sigma #A3803). The mature protein is 66.463 kDa (UniProt Consortium, 2018) and in solution forms monomers and dimers (Levi and González Flecha, 2002). The void volume (V_{void}) was measured with blue dextran (Sigma #D5751, molecular weight approx. 2 MDa) and found to be 45 ml. Based on the manufacturers' manual of the column, the relationship between apparent distribution coefficient (K_{av}) and the logarithm of the molecular weight of the protein is linear.

$$K_{av} = \frac{V_{elution} - V_{void}}{V_{column} - V_{void}}$$

$V_{elution}$ is the volume at which a given protein sample elutes, thus every protein of a given size has a relevant K_{av} value. For better illustration the calibration curve is showcased in the results section (**Figure 3.10.**).

3.2.4 Thermal shift assay (TSA)

Also called differential scanning fluorimetry, thermal shift assay (TSA) is a technique that has been used to test protein stability in ligand-binding screens, including in high-throughput settings (Sorrell et al., 2010). This assay measures the protein melting temperature as a function of solvent composition. Protein unfolding in monitored by addition of a highly hydrophobic fluorescent dye, such as Sypro Orange. Melting temperature (T_m) is defined as the temperature at which 50% of the protein is unfolded. This method is favourable because

it requires small quantities of pure protein, inexpensive reagents and a real-time PCR thermal cycler.

The measurements were performed on a Applied Biosciences 7500 Fast Real-Time PCR System, recorded on a 7500 Software v2.0.6, data collected from the TAMRA filter ('Filter 3', emission peak around 580 nm) and data analysed with the freely available DMAN software (Wang et al., 2012). The reactions were set up in a MicroAmp Fast Optical 96-well reaction plate, sealed with MicroAmp Optical Adhesive Film (#4311971). The first and last row and end-columns on the plates contained only water to prevent evaporation of the reaction mixtures. The dye Sypro Orange used was from Sigma (#S5692-50UL), a 5000x stock solution in DMSO, always diluted freshly in MilliQ water to a 200-fold solution for the experiment. The final reaction volume was 20 µl that contained 0.5 µl dye (5x final concentration). Final protein concentration was approximately 0.2-0.5 mg/ml, all reactions were in triplicates on the same plate.

3.2.5 Mt-DprE2 enzyme activity assay

The assay is loosely based on the Mt-DprE1 enzyme activity assay developed in GSK. The main constituents of the general reaction mixture are the enzymes Mt-DprE1 and Mt-DprE2, either as a co-expressed and co-purified mixture, or added separately as 6xHis-Mt-DprE1 and 6xHis-SUMO-Mt-DprE2. The reaction mixture contained the substrate GGPR, provided by GSK as a 10 mM stock solution in 50 mM HEPES pH 7.5. The assay buffer was kept near physiological pH, most experiments conducted in GSK utilised bis-tris propane (BTP), whereas experiments at the University of Birmingham were conducted in HEPES buffer for comparability with previous studies. A minimum of 40 mM NaCl was added to reactions for osmotic pressure. NADH (Sigma, #N8129) and NADPH (Sigma, #N0411-15VL) stock solutions were prepared in 20 mM Tris-HCl buffer pH 8.5 and small aliquots kept at -20 °C,

and freeze-thaw cycles limited to three. Final reaction volumes were 10 µl for each reaction, in a 384-well small volume microplates (Greiner, #784076). Exact reaction mixtures are always showcased with specific experiment in the results section.

The assays at GSK were performed with SpectraMax M2 and SpectraMax M5e instruments, fluorescence measurements with a monochromator set to 340 nm excitation and 462 nm emission wavelength. The reaction mixture was kept at 26 °C for 5 min, before reaction was initiated with manual GGPR substrate addition. Assays were monitored as kinetic readout for 40-90 minutes, 20 reads/well, and 15 sec/cycle.

The assay at the University of Birmingham was performed using a PHERAstar FS microtitre plate reader, with fixed fluorescence filters at 350/450 nm, and assays run for a minimum of 50 minutes.

Control reactions included buffer and salt solution, the co-factor in buffer and salt solution, at least one negative control lacking either GGPR or one of the enzymes and a positive control with all necessary reagents added for enzyme activity.

3.3 Results

3.3.1 Initial purification methods of recombinant Mt-DprE2

Expression and purification of MBP-Mt-DprE2 has been attempted to the cytoplasm (unpublished work by Sarah M. Batt, University of Birmingham) and had no significant benefit to protein production. However, the plasmid that fuses MBP-tag to the target protein was also available with the signal peptide for periplasmic translocation, which had the prospect of purifying Mt-DprE2 in fewer steps from the periplasm of E. coli (without the

contaminating proteins within the cytoplasm), and a possible disadvantage of producing low amounts of protein.

The purification of MBP-Mt-DprE2 from the periplasm was attempted twice and produced highly pure protein with possible contaminations only from the self-cleaved proteins (based on protein size by SDS-PAGE, **Figure 3.2.**). Unfortunately, out of 5 L of bacteria culture this method produced only a total of 1.5 mg protein, which assessed by ultracentrifugation sedimented as a large aggregate. Therefore, this protein production strategy was abandoned.

1	osmotic shock supernatant	4 µl
2	osmotic shock pellet	-
3	MBPTrap column flow-through	10 µl
4	MBPTrap column washing step	10 µl
5	MBPTrap column elution	20 µl
6	BioRad #1610373	7 µl

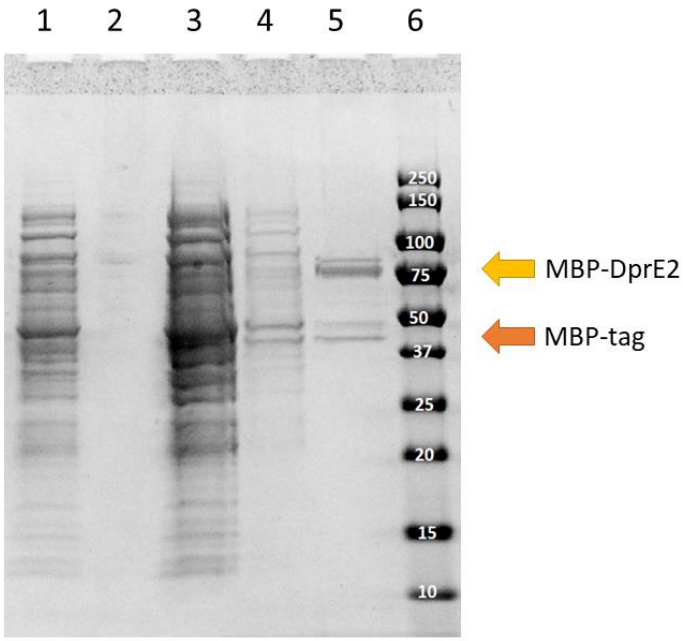


Figure 3.2. SDS-PAGE analysis of MBP-Mt-DprE2 periplasmic protein purification. Sample 5 is eluted and concentrated, approximately 10 µg protein loaded on SDS-PAGE gel. The column binds both the fusion protein and spontaneously cleaved MBP-tag. Pellet sample 2 floated out of the well. MBP-Mt-DprE2 is 72.4 kDa, MBP-tag 42.9 kDa, Mt-DprE2 27.5 kDa.

The co-expression of His-Mt-DprE1 with Mt-DprE2, chaperons and Mt-GtrA (construct 1 in **Table 3.1.**) was unsuccessful due to the low levels of target proteins observed after purification. As captured in **Figure 3.3.**, the small-scale expression trial conducted with this set-up and also large scale purification attempts (not illustrated) produced insufficient amounts of Mt-DprE2 for further investigation.

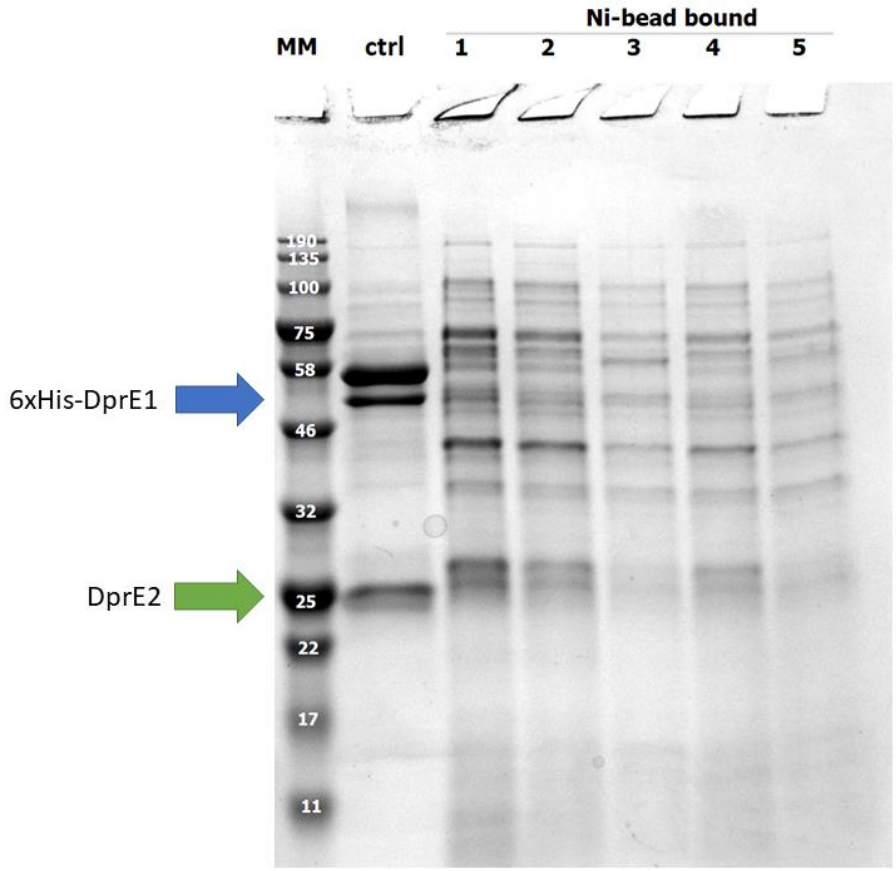


Figure 3.3. Small-scale purification trial with co-expression of 6xHis-Mt-DprE1, Mt-DprE2, chaperones GroES/GroEL, and Mt-GtrA. Samples are 15 µl slurry samples of Ni-NTA beads and proteins bound to them after a 40 mM imidazole wash, different clones (1 to 5) of the same construct. Ctrl sample is large scale purified clone after metal affinity and ion exchange chromatography, reference sample from Luke Broadbent (University of Birmingham). Mt-GtrA is 13.3 kDa integral membrane protein, not expected in the purified samples. MM is 7 µl molecule marker (New England Biolabs, #P77065). 6xHis-Mt-DprE1 is 51.8 kDa, Mt-DprE2 is 27.5 kDa.

What was more concerning about the Mt-GtrA co-expression profile is that it produced high amounts of a protein with slightly higher size based on SDS-PAGE than Mt-DprE1, easily confused with Mt-DprE1 in SDS-PAGE analysis, complicating both identification and quantification of Mt-DprE1. As illustrated in **Figure 3.4.**, co-expression of Mt-DprE1 with Mt-DprE2 and chaperons without Mt-GtrA did not produce this contaminant. Ion exchange chromatography was also implemented to separate Mt-DprE1 or more likely a complex of Mt-DprE1 and Mt-DprE2 from this abundant contaminating protein but was not successful

(see **Figure 3.5.**). The complex eluted at much higher salt concentration than 6xHis-Mt-DprE1 alone and eluted together with the contaminant.

	Expressed proteins	Purification	Volume
1	6xHis-Mt-DprE1 GroES/GroEL	IMAC + IEX	7 µl
2	6xHis-Mt-DprE1 Mt-DprE2 GroES/GroEL 6xHis-Mt-GtrA-6xHis	IMAC	8 µl
3	6XHis-Mt-DprE1 Mt-DprE2 GroES/GroEL	IMAC	1.5 µl

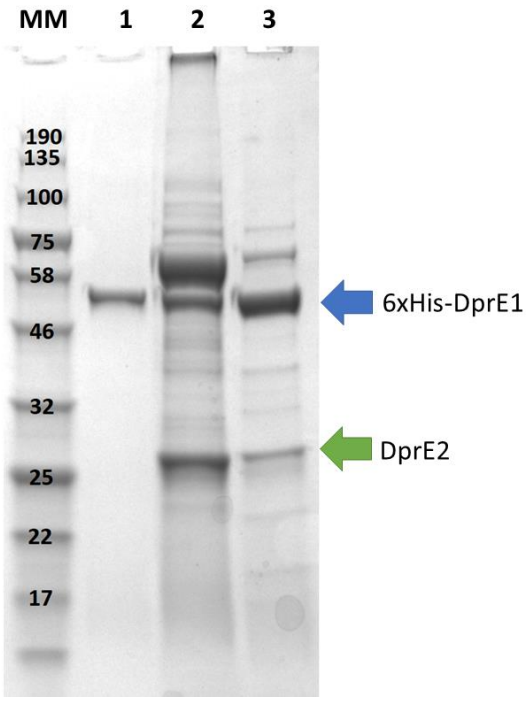


Figure 3.4. Protein products of Mt-DprE1 alone and Mt-DprE1-Mt-DprE2 co-expressions. IMAC is immobilised metal affinity chromatography, IEX is ion exchange chromatography. MM is 5 µl molecule marker (New England Biolabs, #P77065). 6xHis-Mt-DprE1 is 51.8 kDa, Mt-DprE2 is 27.5 kDa.

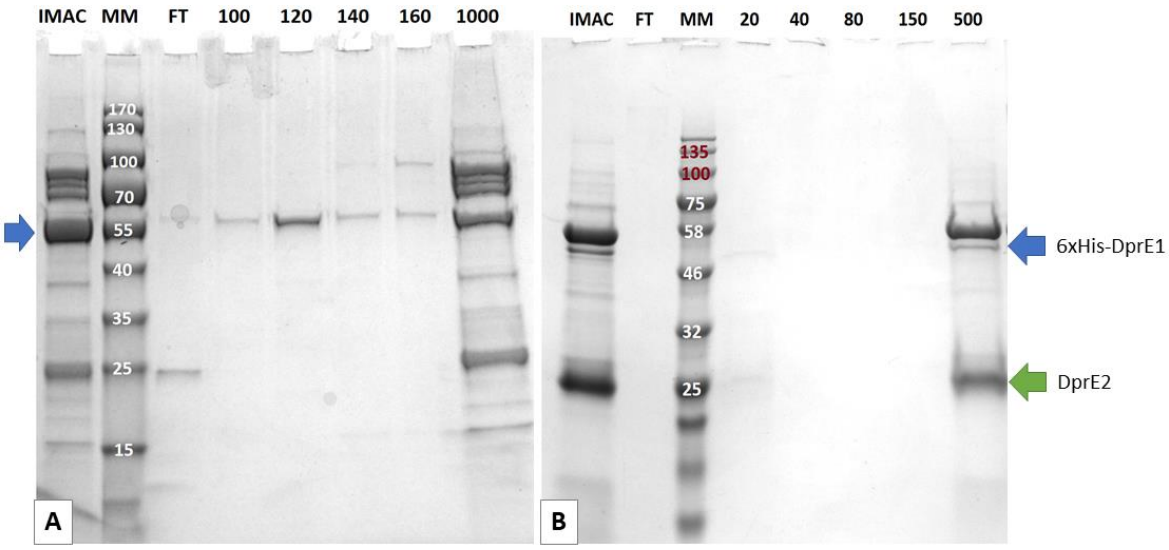


Figure 3.5. Ion exchange chromatography purification of Mt-DprE1 constructs. (A) 6xHis-Mt-DprE1 with GroES/GroEL, (B) 6xHis-Mt-DprE1 co-expression with Mt-DprE2, GroES/GroEL, and Mt-GtrA. IMAC is immobilised affinity chromatography sample eluted from the HisTrap column, FT is flow-through (protein unbound to the anion exchange column), numbers depict NaCl concentration [mM] of the eluted sample. MM is 7 µl molecule marker (A) ThermoFisher #26616 and (B) New England Biolabs, #P77065. 15 µl samples loaded on the gel. 6xHis-Mt-DprE1 is 51.8 kDa, Mt-DprE2 is 27.5 kDa.

3.3.2 6xHis-SUMO-Mt-DprE2 production

6xHis-SUMO-Mt-DprE2 fusion protein expressed from a pET-based plasmid produced high amounts of protein via metal affinity chromatography (see **Figure 3.6.**).

In order to enhance protein expression two additives were tested during protein expression. Benzyl-alcohol (BA) induces membrane fluidisation in *E. coli* and elicits increased native chaperone expression (De Marco et al., 2005). Together with the osmolyte betaine especially high levels of recombinant proteins have been achieved (De Marco et al., 2005).

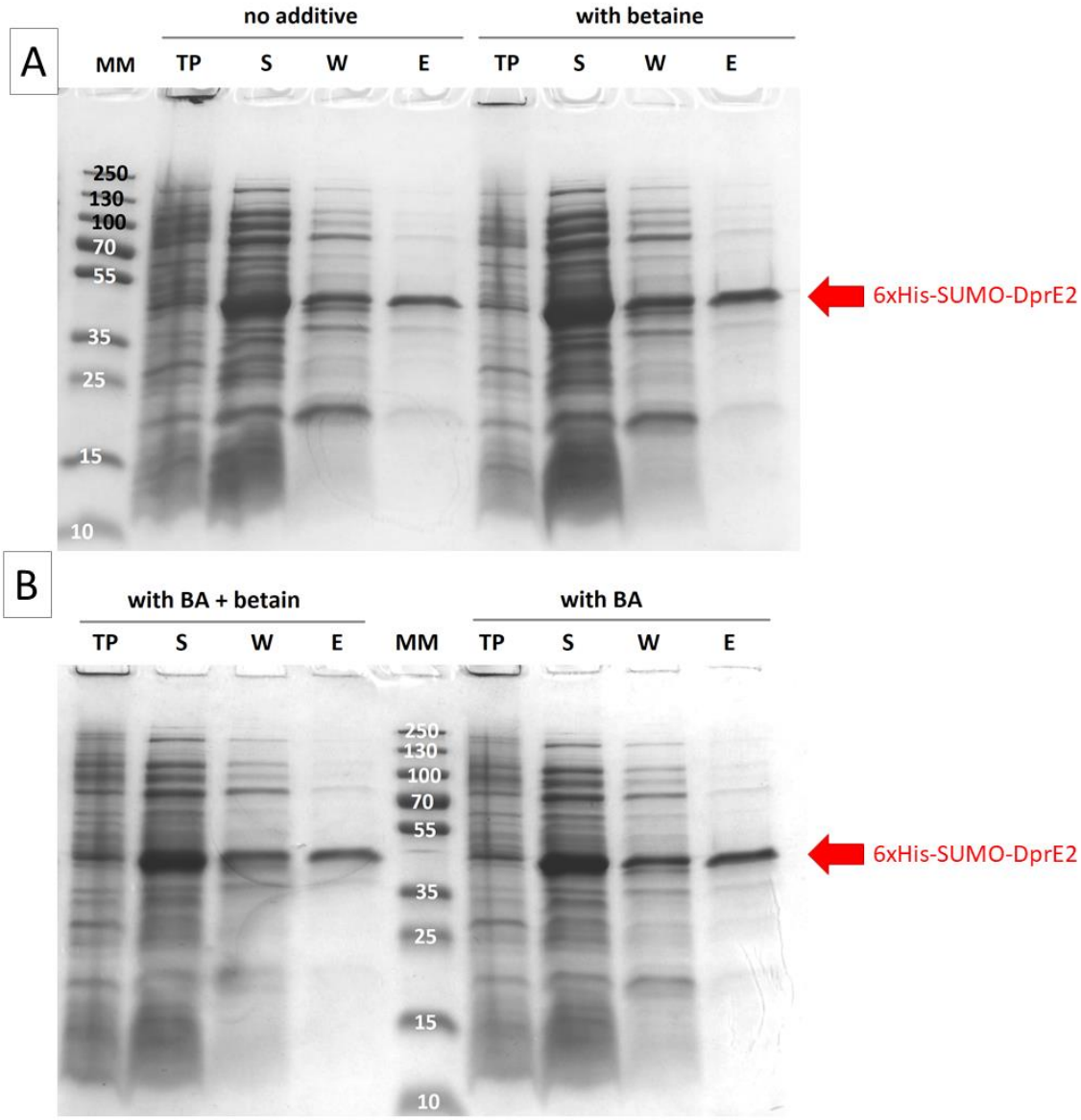


Figure 3.6. Additive optimisation of 6xHis-SUMO-Mt-DprE2 expression. BA is benzyl alcohol, TP total protein after cell lysis, S supernatant after centrifugation of lysed cells, W washing step with 50 mM imidazole, E elution step with 300 mM imidazole, MM molecule marker (ThermoFisher, #26619). 3 µl of TP and S, 15 µl of W and E samples were loaded on the gels. 6xHis-SUMO-Mt-DprE2 is 40.9 kDa.

The chosen additives did not visibly alter expression levels of 6xHis-SUMO-Mt-DprE2 (see **Figure 3.6.**). However, as the osmotic stabiliser betaine is relatively inexpensive, it was included in the liquid broth during overexpression of this construct.

3.3.3 6xHis-Mt-DprE1 and Mt-DprE2 co-expression

Co-expression of the two proteins, Mt-DprE1 and Mt-DprE2 was successful aided by chaperon proteins. In **Figure 3.7.** the expression levels are compared when including or omitting chaperon, co-expression with the chaperons showed increased levels of Mt-DprE1 and consequently Mt-DprE2.

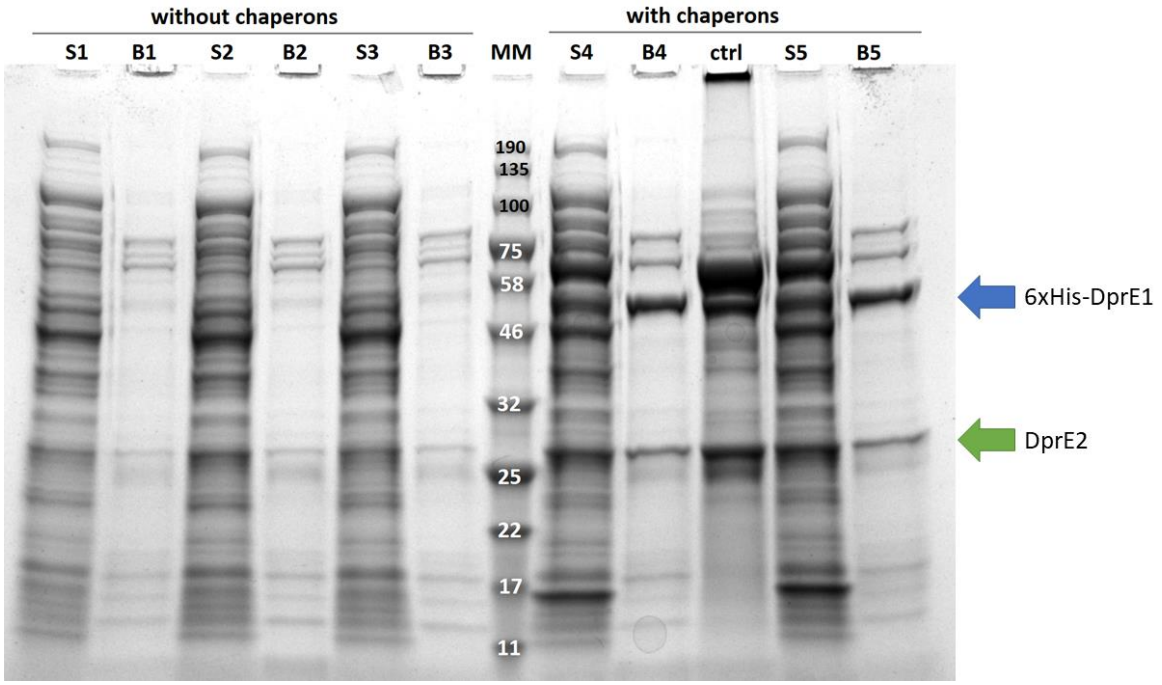


Figure 3.7. Small-scale expression trials with 6xHis-Mt-DprE1 and Mt-DprE2 co-expression assessing chaperone necessity. S signifies soluble fraction after lysis, B labels Ni-bead bound protein samples. Clones 1 to 5 were tested. MM is molecule marker (New England Biolabs, #P77065). Chaperons are GroES from *E. coli* and GroEL (*groEL2*) from *M. tuberculosis*. Mt-DprE2 is expressed from a pET plasmid, separate from 6xHis-Mt-DprE1 (construct 3 in **table 3.1.**). 6xHis-Mt-DprE1 is 51.8 kDa, Mt-DprE2 is 27.5 kDa.

It was tested whether ion exchange chromatography improved protein purity following metal affinity chromatography. Overall protein yields were low and Mt-DprE2 production barely visible on SDS-PAGE (**Figure 3.8.**). Hence, ion exchange chromatography was omitted in following studies.

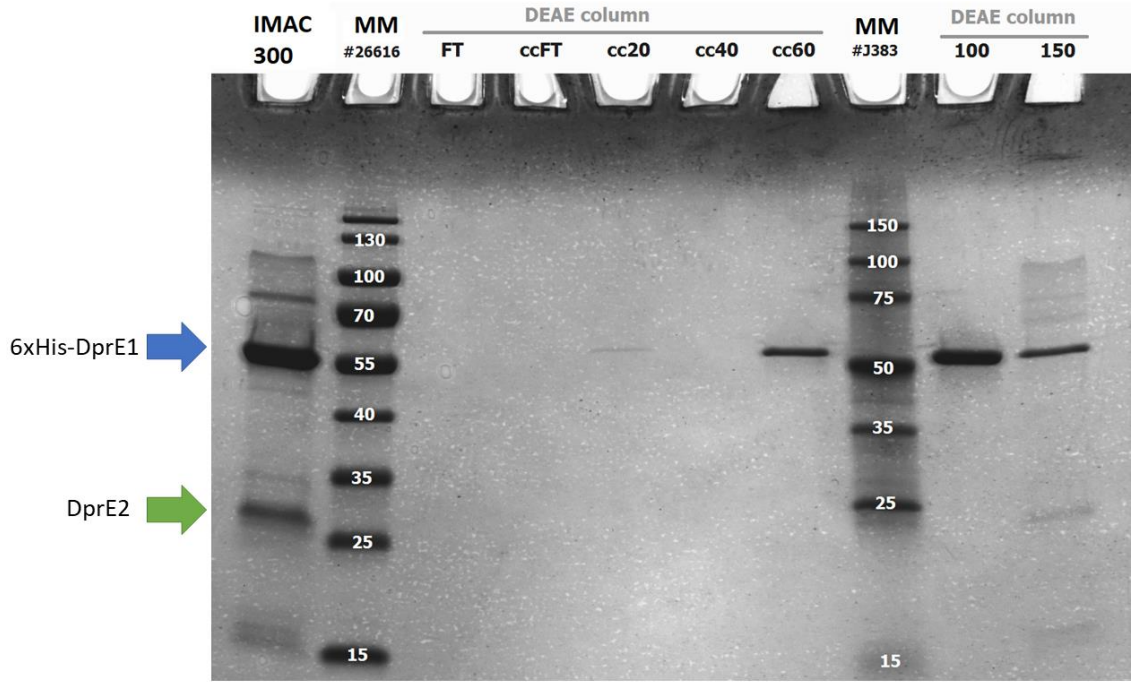


Figure 3.8. Ion exchange chromatography of Mt-DprE1-Mt-DprE2 co-expression. The construct co-expressed 6xHis-Mt-DprE1, Mt-DprE2 and GroES/GroEL proteins. 'IMAC 300' is the sample after metal affinity chromatography, eluted in 300 mM imidazole and dialysed. FT is flow-through (proteins unbound to the anion exchange column), numbers label the NaCl concentration in [mM] in the eluted fraction from DEAE anion exchange column. 'cc' signifies that the sample was concentrated via filtration. MM is molecule marker, here ThermoFisher #26616 was compared to Amresco #J383, the latter producing unusual running pattern indicating higher protein size than true. 6xHis-Mt-DprE1 is 51.8 kDa, Mt-DprE2 is 27.5 kDa.

The expression of Mt-DprE2 from a separate plasmid (**Table 3.1.** construct 3) also produced Mt-DprE2, notably the highest relative Mt-DprE2 concentration compared to Mt-DprE1 (see **Figure 3.9.**). Based on the SDS-PAGE analysis this construct contained an approximately 1:1 ratio of the two proteins.

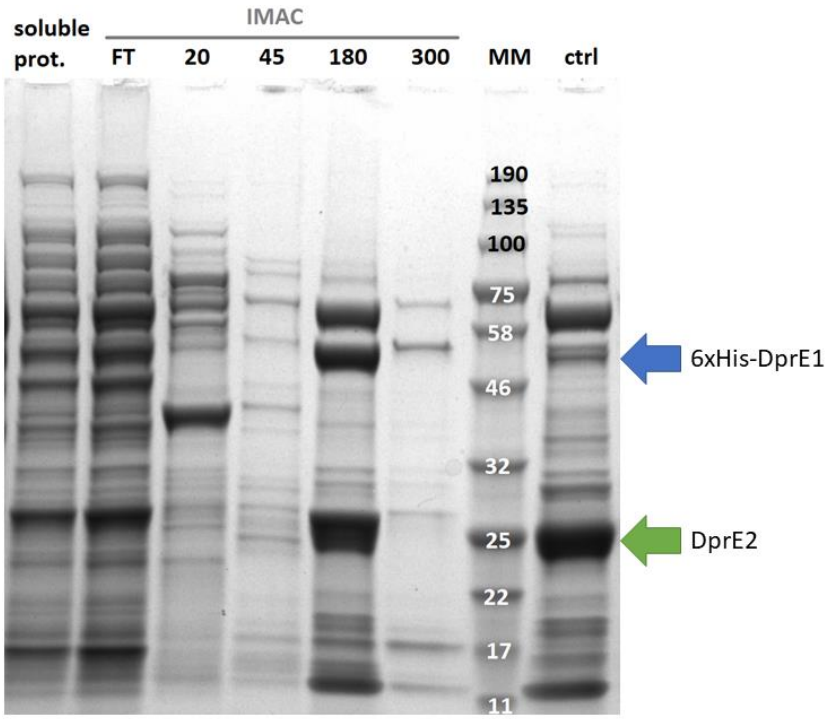


Figure 3.9. Co-expression of Mt-DprE1 and Mt-DprE2 transcribed from separate plasmids. The construct co-expressed 6xHis-Mt-DprE1, Mt-DprE2, GroES/GroEL (construct 3 in **Table 3.1.**). Soluble protein is sample after cell lysis and centrifugation, IMAC is immobilised metal affinity chromatography, FT is flow-through (proteins unbound to HisTrap column), numbers signify imidazole concentration in [mM] in eluted fraction. Ctrl is reference sample from Luke Broadbent (University of Birmingham). MM is molecule marker (New England Biolabs, #P77065). 6xHis-Mt-DprE1 is 51.8 kDa, Mt-DprE2 is 27.5 kDa.

3.3.4 Analytical size exclusion chromatography (SEC)

Size exclusion chromatography (SEC) separates proteins or macromolecules by hydrodynamic radius. Here I analysed if SUMO-tag-fused Mt-DprE2 was prone to aggregation, furthermore if 6xHis-DprE1 formed a complex with DprE2.

All SEC experiments with 6xHis-SUMO-Mt-DprE2 resulted in low concentration protein samples. This is due to the aggregation in the IMAC eluted sample, because concentration *via* ultrafiltration promotes aggregation on the filter, and aggregates need to be removed by filtration (0.22 μm) before loading onto a SEC column. Thus, the highest protein

concentration achieved was only 5 mg/ml (based on an absorption at 280 nm). Also, the column had a limit of 4 ml for loaded sample. These limits resulted in low protein concentration fractions after chromatography, that in biochemical assays performed more poorly than the more contaminated but concentrated sample before SEC purification.

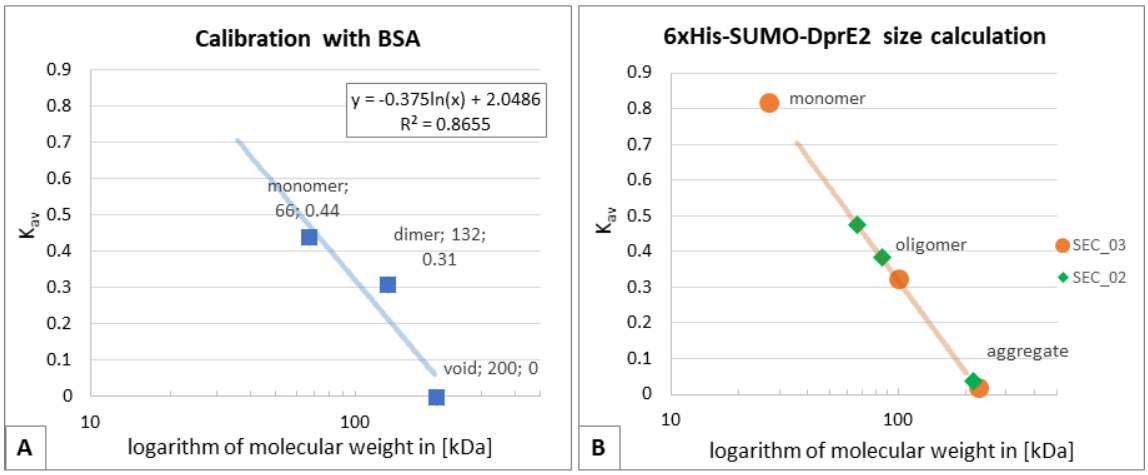


Figure 3.10. Calibration and size estimation of 6xHis-SUMO-Mt-DprE2 in SEC. (A) Calibration with bovine serum albumin (monomer and dimer) and blue dextran (void volume). Equation for linear regression and coefficient of determination (R^2) is in box on top right. (B) Peaks of representative species from two SEC measurements (SEC_02 and SEC_03). K_{av} is apparent dissociation constant.

The measurement was repeated three times with similar results. Most of the protein eluted as a large aggregate around 45 ml (void volume), while a small fraction of the protein eluted at near 70 ml elution volume. This fraction contained a 6xHis-SUMO-Mt-DprE2 oligomer with degrading contaminant according to the SDS-PAGE visualisation (see **Figure 3.11.**). Based on the calibration with bovine serum albumin, the oligomer is about 100 kDa in size. Since the fusion protein is 40.88 kDa, taken together with results of the repeat experiment, where the oligomer eluted as a double peak at 73 and 81 ml (calculated weights 84 and 66 kDa, respectively, **Figure 3.11.B**), the oligomer is likely a dimer. This is reproduced in **Figure 3.12.**, where a second small peak (see fraction eluting at 80.5 ml) also appeared, which was likely the degradation product of this dimer. The contaminant proteins present in

these samples are visualised on SDS-PAGE and appear to be also between 70 and 100 kDa, which further supports the conclusion of it being a dimer (see fraction eluting at 64 ml in **Figure 3.11.C**).

It is worthy to note, that certain samples were sent for analysis to Cellzome GmbH, (GSK, Heidelberg) where Western blot with anti-His-tag antibody and mass spectrometry was performed for fractions eluting at 46.3, 48.1, 73.3 and 80.5 ml (shown in **Figure 3.12.B**). The presence of Mt-DprE2 amino-acid sequence was confirmed in all samples by mass spectrometry. Also, the His-tag was detected on some smaller size proteins, possible degradation products of the fusion protein in sample eluted at 80.5 ml (**Figure 3.13.**).

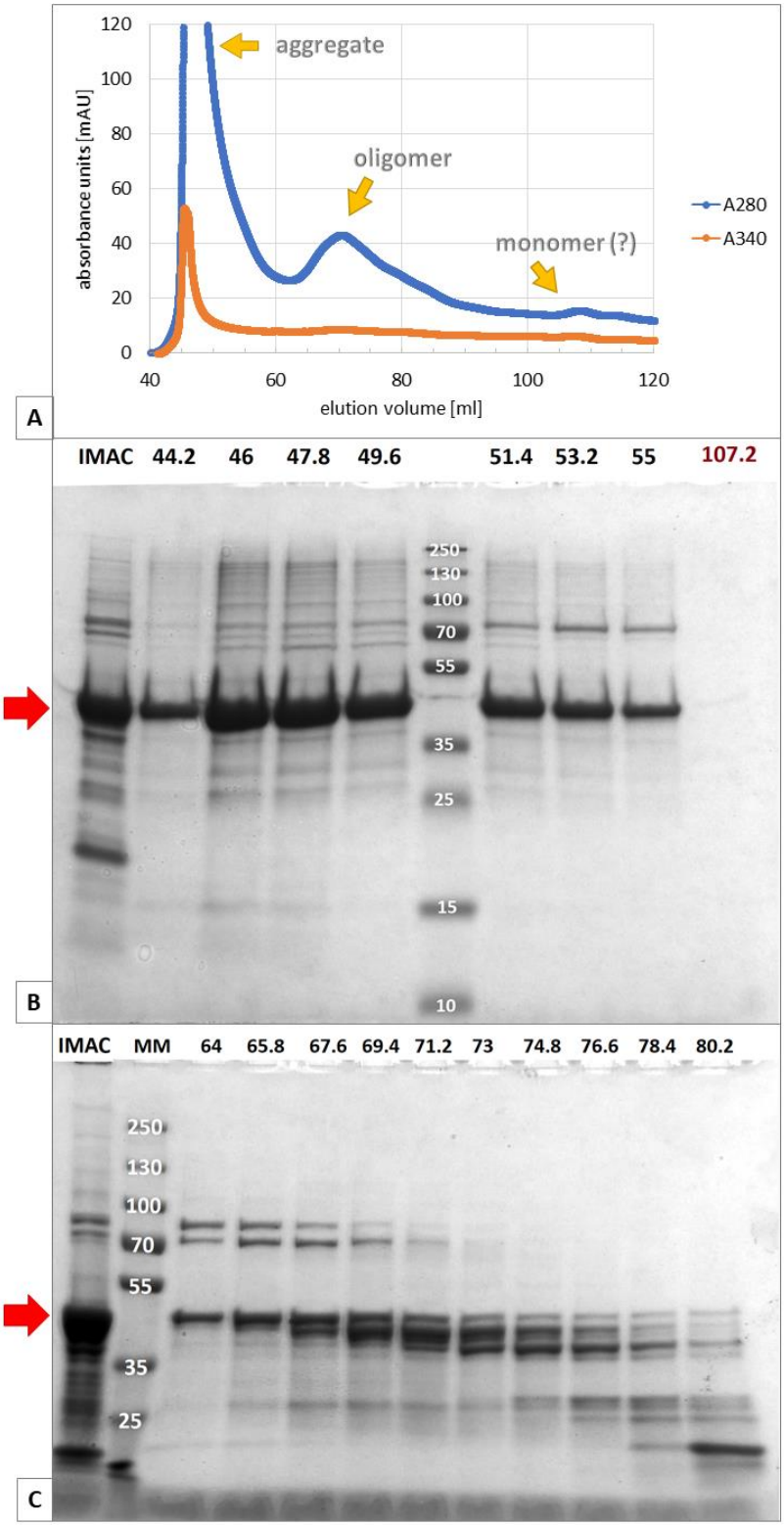


Figure 3.11. Size exclusion chromatography of 6xHis-SUMO-Mt-DprE2. (A) Elution chromatogram at two absorbance wavelengths. (B) and (C) SDS-PAGE analysis of collected protein fractions labelled by elution volume in [ml], 20 µl samples loaded on SDS-PAGE gel. IMAC is the sample loaded on the SEC column, MM molecule marker (ThermoFisher, #26619). 6xHis-SUMO-Mt-DprE2 is 40.9 kDa (red arrows).

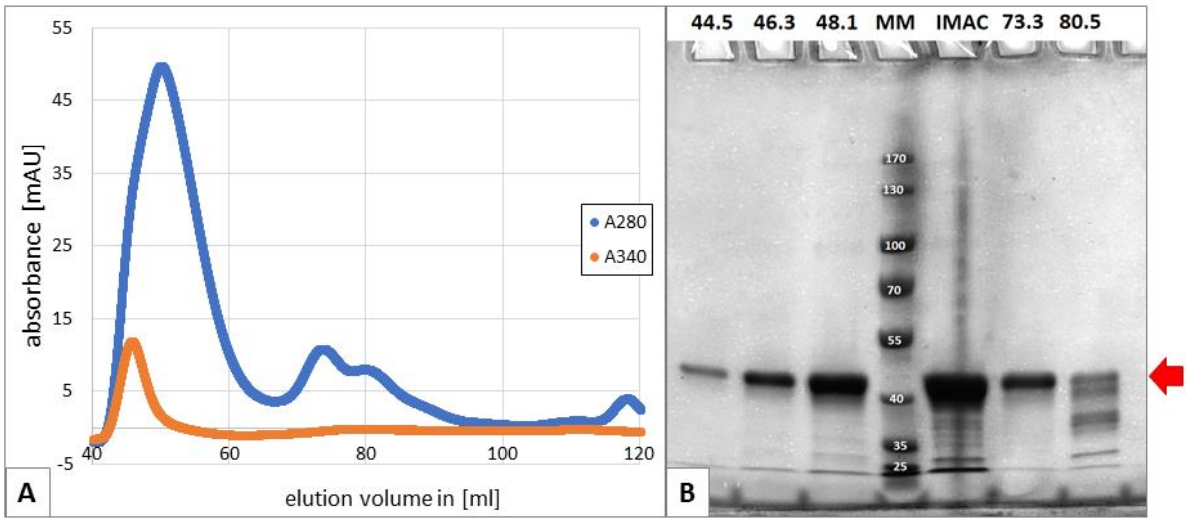


Figure 3.12. Size exclusion chromatography of 6xHis-SUMO-Mt-DprE2 (repeat experiment). (A) Elution chromatogram at two absorbance wavelengths. (B) SDS-PAGE analysis of collected protein fractions labelled by elution volume in [ml], 20 µl samples loaded on gel. MM molecule marker (ThermoFisher, #26616), IMAC is the sample loaded on SEC column. 6xHis-SUMO-DprE2 is 40.9 kDa (red arrow).

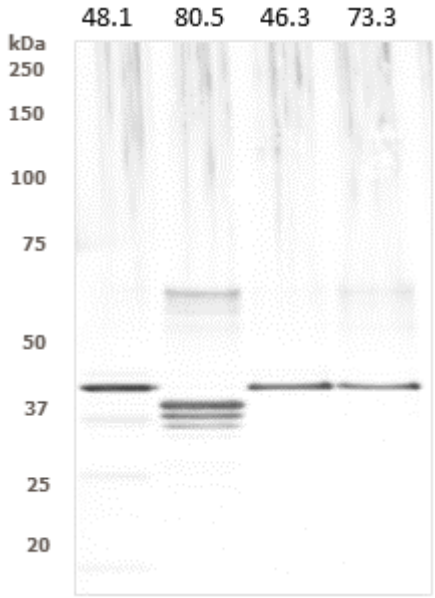


Figure 3.13. Western blot analysis of 6xHis-SUMO-Mt-DprE2. Samples are from size exclusion chromatography (**Figure 3.12.**), labels refer to elution volumes in [ml]. 1 µg proteins were loaded, Abcam anti-6xHis-tag antibody (ab9108) 1:500 applied. Fusion protein is 40.9 kDa, 6xHis-SUMO-tag 13.5 kDa, Mt-DprE2 27.3 kDa. Work performed by Sonja Ghidelli-Disse (Cellzome GmbH).

SEC of co-expressed 6xHis-Mt-DprE1 with Mt-DprE2 was performed following metal affinity chromatography. The buffer was similar to the published dialysis buffer for Mt-DprE1 purification (40 mM Tris-HCl pH 8.5, 20 mM NaCl, 10% glycerol), which is higher pH and lower salt concentration than used for the fusion protein experiments. Once again, due to the necessary removal of aggregate proteins by filtration, only very low amounts of protein (approx. 2 mg/ml, 8 mg in total) were loaded on the column and resulted in diluted protein fractions after SEC. The elution volume at which the possible oligomers eluted, around 80 ml, corresponds to approx. 65 kDa with the bovine serum albumin calibration standard (**Figure 3.10.A**).

There is a possible monomer species eluting at 127 ml, with a calculated size of 13 kDa (**Figure 3.11.A**), which is beyond the range of this size exclusion column to determine the size accurately, hence the difference of calculated (27.3 kDa) and measured size values.

The oligomer could be the 1:1 complex of 6xHis-Mt-DprE1 and Mt-DprE2 (78 kDa), however the bands are not equally broad, Mt-DprE1 seems to be present at a higher concentration then Mt-DprE2 (see **Figure 3.14.B** fraction eluted at 82 ml). On the SDS-PAGE analysis of these samples some faint bands of unknown proteins can be observed as well between the size of the suspected Mt-DprE1 and Mt-DprE2 species. Furthermore, the protein approximately at the size of Mt-DprE2 is present in both aggregate and oligomer fractions, however, run differently on SDS-PAGE (see **Figure 3.14.B**, fractions eluted at 46 vs 80.2 ml).

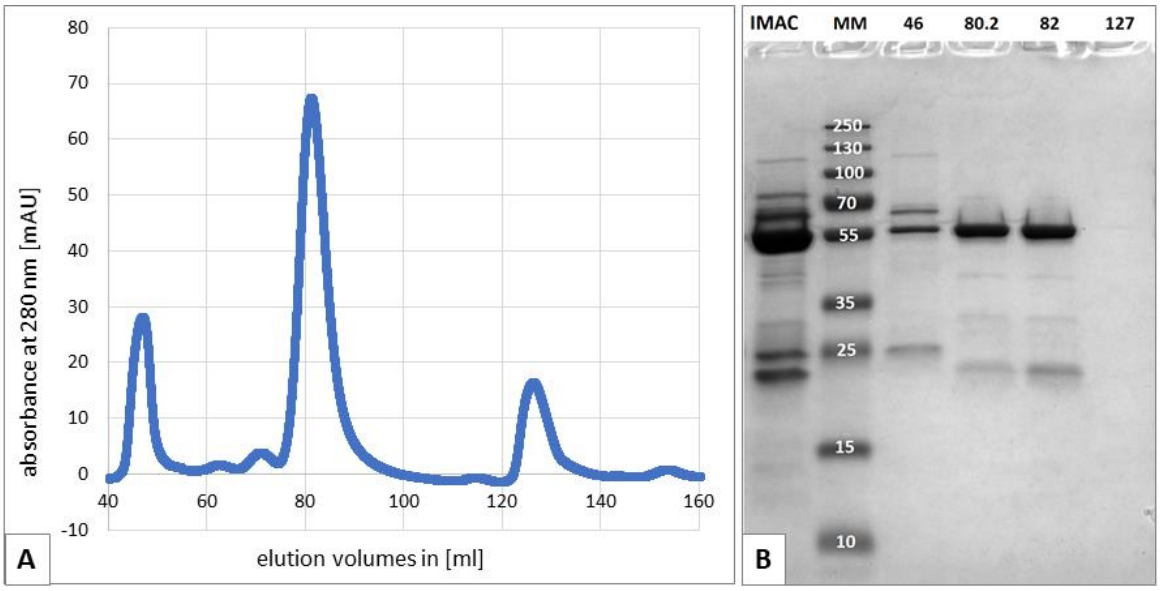


Figure 3.14. Size exclusion chromatography of His-Mt-DprE1 and Mt-DprE2 co-expression. (A) Elution chromatogram, (B) SDS-PAGE analysis of representative peak fractions. IMAC is the sample that is loaded on the SEC column, labels correspond to elution volumes in [ml]. MM molecule marker (ThermoFisher, #26619). 6xHis-Mt-DprE1 is 51.8 kDa, Mt-DprE2 is 27.5 kDa. Co-expressed in the presence of GroES/GroEL.

3.3.5 Buffer optimisation with thermal shift assay

To optimise the buffer for biochemical assay and protein purification, the pH preference was assessed with 6xHis-SUMO-Mt-DprE2 in a thermal shift assay. The fusion protein has a calculated isoelectric point of 6.33 calculated by ProtParam (ProtParam, n.d.). A high melting temperature (T_m) was achieved at around pH 7 (see **Figure 3.15.**).

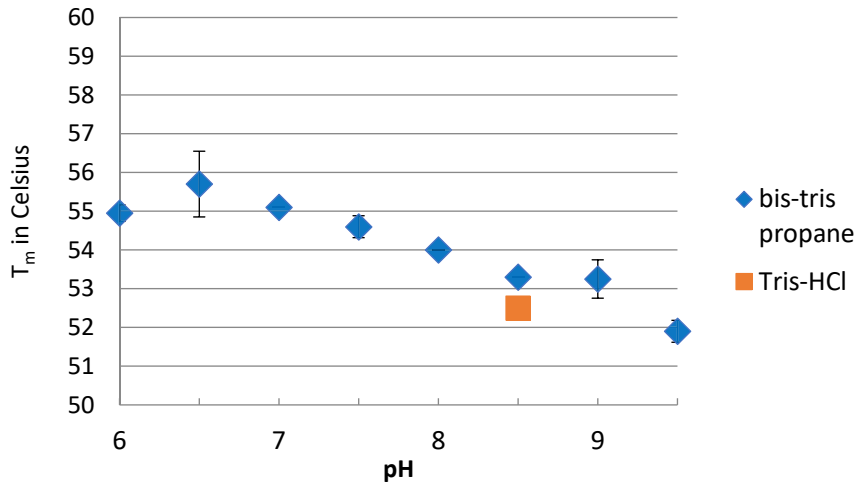


Figure 3.15. Buffer pH optimisation with 6xHis-SUMO-Mt-DprE2 in thermal shift assay. Error bars are standard deviations for duplicates measured on the same plate. Protein is 0.4 mg/ml of the 300 mM imidazole elution sample of betaine additive test (in **Figure 3.6.A**).

In order to find the best buffer for both Mt-DprE2 and Mt-DprE1, buffer preferences for 6xHis-Mt-DprE1 were evaluated. Interestingly, Mt-DprE1 had a much lower T_m than the fusion-protein, highest stability could be achieved by keeping the pH relatively high (pH 8.5) and adding higher salt concentrations than the published storage buffer which is 20 mM Tris-HCl (pH 8.5), 10 mM NaCl and 10 v/v% glycerol (Batt et al., 2012).

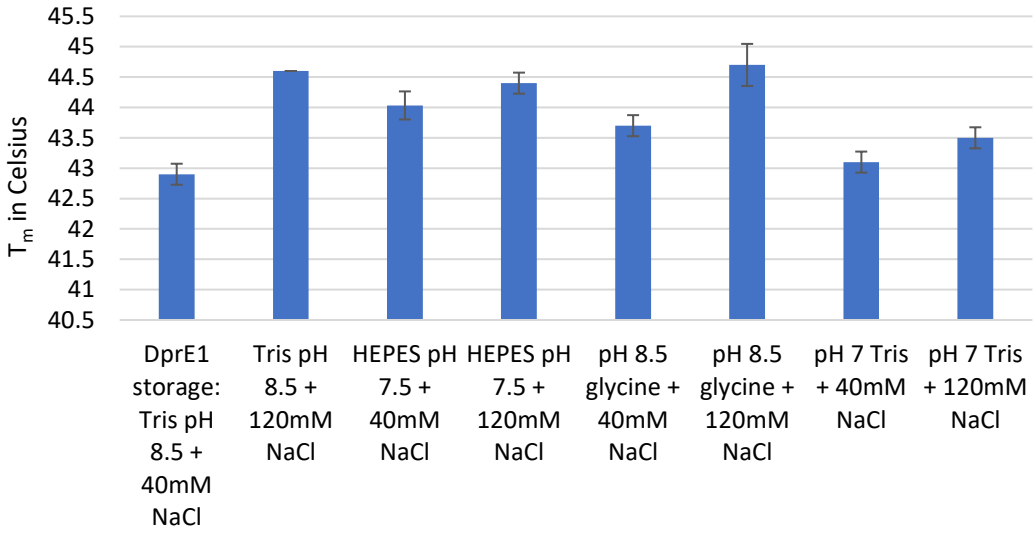


Figure 3.16. Thermal shift assay measurements with 6xHis-Mt-DprE1 for buffer optimisation. Error bars are for triplicate measurement standard deviations.

As a conclusion, for enzyme activity assay the pH 7.2 was chosen in BTP (bis-tris propane) and minimum 100 mM salt concentration.

To evaluate additives that may prevent aggregation of Mt-DprE2, screening advice from Lebendiker *et al.* (2014) was folloed. Osmolytes and salts are reported to stabilise proteins in general (Lebendiker et al., 2014), also the putative native ligand, NADH or NADPH may increase the T_m in solution. GGPR was also assessed, but it decreased the T_m for the fusion protein already at 100 μM final concentration, which further implies that the substrate acts as a surfactant and may unfold the protein at higher concentrations (see **Figure 2.7.**).

Among beneficial additives, the salt $(NH_4)_2SO_4$ had the most stabilising effect. This salt has been used historically in high concentrations to reversibly precipitate proteins for purification purposes, but in these tested concentrations stabilised Mt-DprE2 in solution.

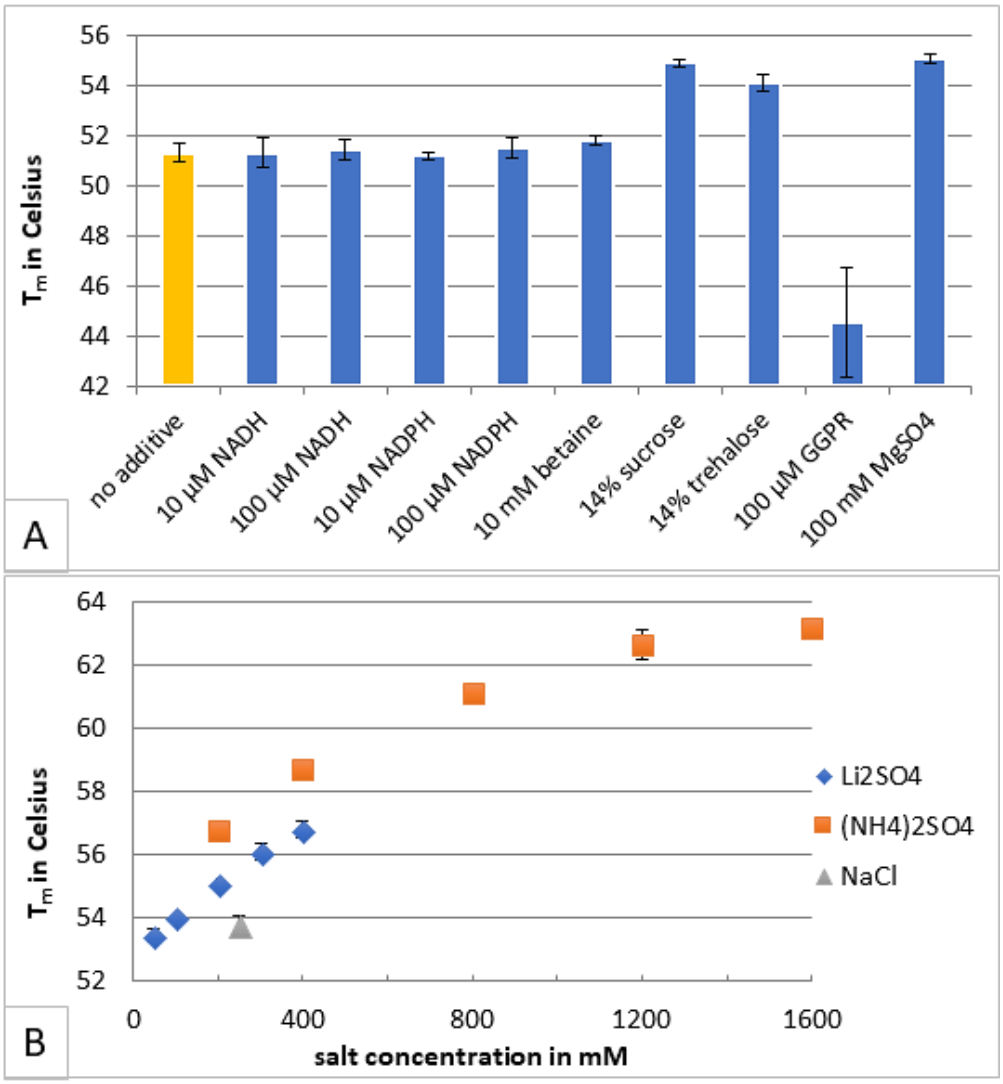


Figure 3.17. Additive screen with 6xHis-SUMO-Mt-DprE2 in thermal shift assay. 14 w/v % of both sugars relate to approx. 410 mM. Buffer for samples in (A) is Tris-HCl pH 8.5, in (B) Na-phosphate buffer pH 8. The control 'no additive' has 50 mM NaCl from storage buffer. Error bars are standard deviations for triplicates from the same plate. Protein as in **Figure 3.15.**

3.3.6 Storage conditions optimisation by enzyme activity assay

Highly similar protein samples produced for expression improvement were utilised (**Figure 3.6.**) for testing the impact of different cryoprotectant additions. Based on the Lebendiker aggregation optimisation (Lebendiker et al., 2014), sucrose, glycerol, PEG-400 and Tween-80 were used for testing. Sucrose is an osmolyte, glycerol a widely used

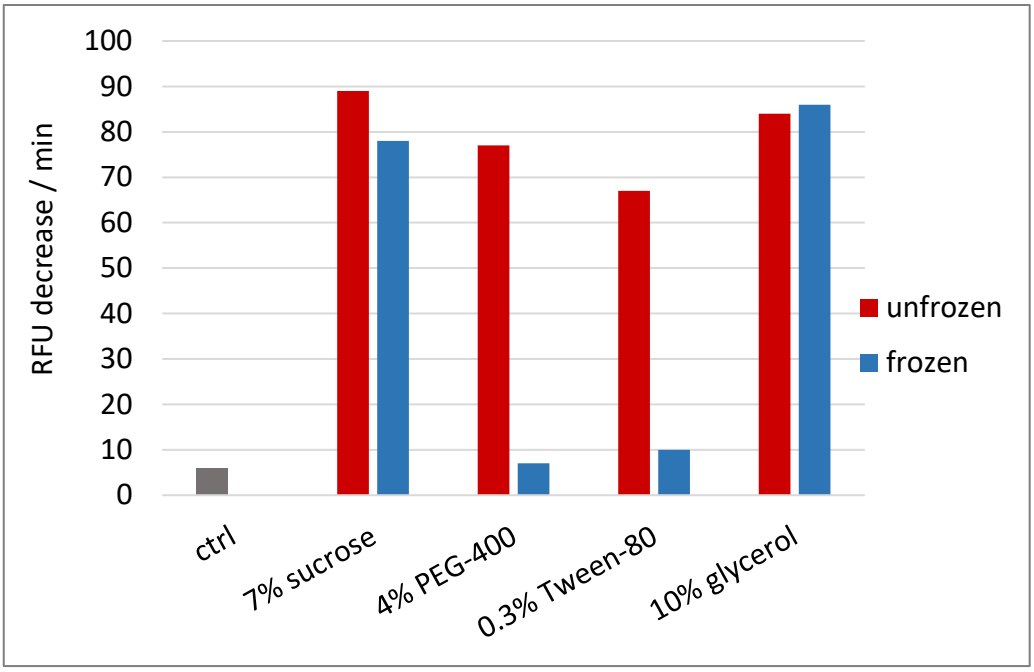
cryoprotectant in protein crystallography, hence possibly advantageous for freeze-thaw cycles to keep the protein soluble. PEG-400 is a polyol with protein stabilising properties, and Tween-80 is a detergent that lowers surface-induced aggregation that can bind hydrophobic parts of the protein, increasing water solubility (Lebendiker et al., 2014).

Cryoprotectants were assessed by ability to keep the protein in solution after overnight freezing at -20 °C, compared to unfrozen samples without additive kept on 4 °C. In this experiment, as shown in **Table 3.2**. all additives performed similarly with the exception of PEG-400, which separated from the solution and most likely had a fluorescence absorption by itself at 280 nm, hence reporting false protein concentrations. Also, the overnight frozen samples were centrifuged at high speed to separate precipitated protein, which was the most prominent in the PEG-supplemented sample.

expression batch	protein conc. (unfrozen)	conc. after volume expansion	additive final conc.	protein conc. (frozen)
ctrl (no additive)	0.9	0.81	7% sucrose	0.68
BA	0.65	0.55	4% PEG-400	0.69*
betaine	1	0.9	0.3% Tween-80	0.75
BA+betaine	1	0.85	10% glycerol	0.6

Table 3.2. Additive test for protein cryoprotection. Protein samples as in **Figure 3.6.** Conc. is concentration, frozen and unfrozen samples measured at 280 nm absorbance with Nanodrop (ThermoFisher), given in [mg/ml]. Concentrations after volume expansion are calculated, dilution (volume expansion) was 10-15 v/v% in the frozen samples. Asterisk shows misleading concentration, most of the protein visibly precipitated in the sample.

Furthermore, all these protein samples of 6xHis-SUMO-Mt-DprE2 were investigated in enzyme activity assay and compared unfrozen enzymes. PEG-400 and Tween-80 frozen samples had low enzyme activity, while glycerol and sucrose retained enzyme activity well (**Figure 3.18.**).



Buffer composition:
40 mM HEPES pH 7.5 (1 M)
100 mM NaCl (5 M)
180 µM NADH (3 mM)
10 µM DprE1 (7.2 mg/ml ≈ 113 µM)
0.4 mg/ml 6xHis-SUMO-DprE2
150 µM GGPR (1.5 mM)
MilliQ water addition to 10 µl

Figure 3.18. Cryoprotectant testing with 6xHis-SUMO-Mt-DprE2 enzyme activity assay. Background is NADH-oxidation by Mt-DprE1 alone (no GGPR, no Mt-DprE2 added). Frozen samples incubated at -20 °C overnight, unfrozen samples were kept at 4 °C (see **table 3.2.**). DprE2 was purified by IMAC only (without subsequent SEC purification). Control reaction lacked GGPR. All reaction samples were centrifuged before activity measurement. Single shot assay, stock solution concentrations listed in parentheses, RFU is relative fluorescence unit.

3.3.7 Enzyme activity assay optimisation

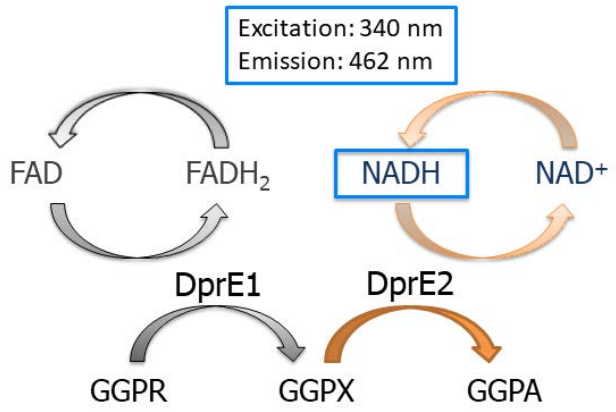


Figure 3.19. Basis of Mt-DprE2 enzyme activity measurement: fluorescence-decrease by NADH oxidation. GGPR is geranylgeranyl phosphorylribose, GGPA geranylgeranyl phosphorylarabinose, GGPX the keto-intermediate. The fluorescent co-enzyme is highlighted in blue.

To measure Mt-DprE2 activity it is plausible to measure the oxidation of NADH resulting in fluorescence signal decrease (**Figure 3.19.**). GGPX production and separation from GGPR and Mt-DprE1 was unsuccessful. Therefore, Mt-DprE1 also had to be present in the Mt-DprE2 activity assay supplying the substrate, GGPX. As a result, the Mt-DprE2 assay is a two-step coupled enzyme activity assay.

NADH (and NADPH) reduction can be observed *via* change in intrinsic absorption or fluorescence of the molecule. However, none of the absorbance studies showed enzyme activity. Therefore, all biochemical assays presented here were monitored by fluorescence spectroscopy.

Firstly, the fluorescence spectrum of NADH was measured and the range of measurement, where co-factor concentration is directly proportional to fluorescence signal, established. (See **Figure 3.20.**) With current setup the maximum NADH concentration was set to 300 µM or less to remain in the linear range, and the general emission wavelength for assays, to measure the highest signal, was 462 nm.

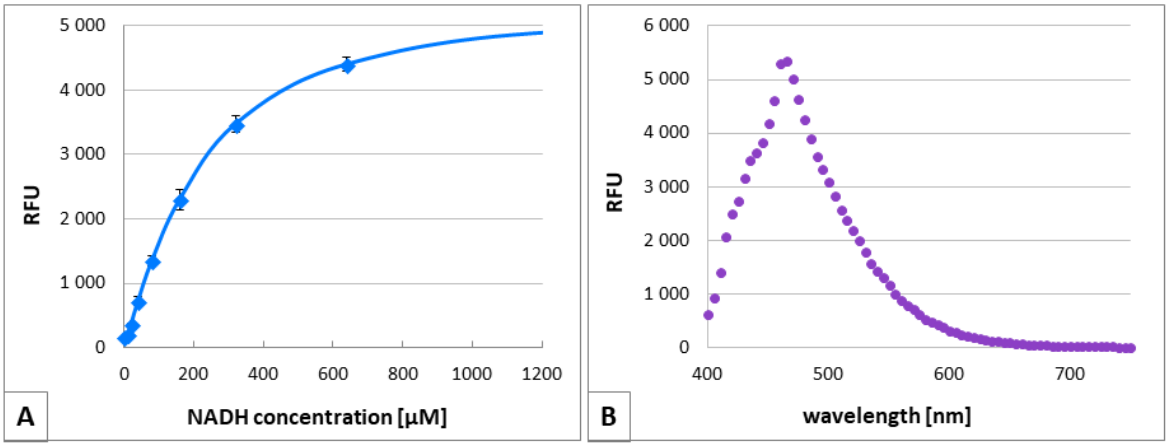
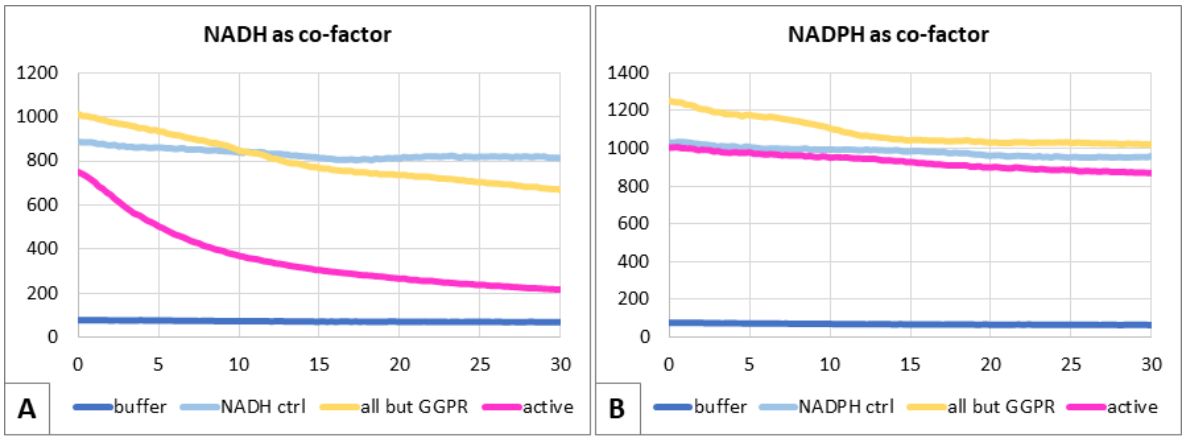


Figure 3.20. NADH fluorescence signal test. (A) Linear range of signal to concentration with SpectraMax M5 spectrophotometer, in 10 µl final volume, on black small volume plates (Greiner #784076). Measurement in triplicates. (B) Fluorescence emission spectrum of sample 2.5 mM NADH, at excitation wavelength 340 nm. RFU is relative fluorescence unit.

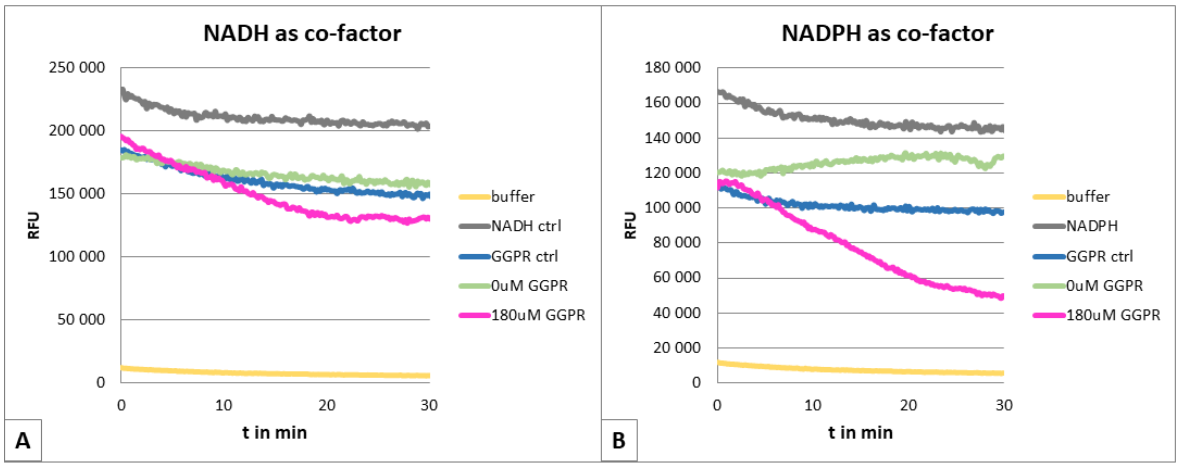
Addressing the question of whether NADH or NADPH may be better co-factor for the Mt-DprE2 assay, assays with both in different protein construct assays were utilised. Interestingly, only NADH performed with tag-free and SUMO-tagged Mt-DprE2 similarly, with a slow constant decrease in background signal compared to pronounced signal drop in an active reaction. However, NADPH signal was undetectable with 6xHis-SUMO-Mt-DprE2 (**Figure 3.21.**), while the Mt-DprE1-Mt-DprE2 co-expression construct expressed from separate plasmids (construct 3 in **Table 3.1.**) gave a solid signal with a prolonged linear phase in co-factor oxidation (**Figure 3.22.**). Previous studies (conducted by Alice Moorey, MRes thesis work in University of Birmingham) utilised co-expression of Mt-DprE1-Mt-DprE2-GroES/GroEL-Mt-GtrA in which the assay monitored NADPH oxidation. Reproducing those results in this work of the same construct was unsuccessful, no enzymatically active preparation of the proteins was obtained.



	buffer	NAD(P)H ctrl	all but GGPR	active
80 mM HEPES pH 7.5 (1M)	+	+	+	+
100 mM NaCl (5M)	+	+	+	+
160 µM NADH/NADPH (1 mM)		+	+	+
22 µM DprE1 (7.2 mg/ml = 113 µM)			+	+
0.32 mg/ml SUMO-DprE2 (1 mg/ml)			+	+
150 µM GGPR (1.5 mM)	+	+		+
MilliQ water addition for 10 µl	+	+	+	+

Figure 3.21. Co-factor testing with 6xHis-SUMO-Mt-DprE2 enzyme activity assay. (A) NADH and (B) NADPH as co-factor in the same reaction settings. Stock solution concentrations in parentheses. DprE2 was purified with IMAC.

It is important to note, that due to the different detectors and instruments there is a large difference in fluorescence values between experiments conducted in GlaxoSmithKline Tres Cantos DDW (as in Figure 3.21.) and University of Birmingham (as in Figure 3.22.). The former reactions usually started from 1000 fluorescence intensity units, while the latter well over 100,000 fluorescence intensity units.

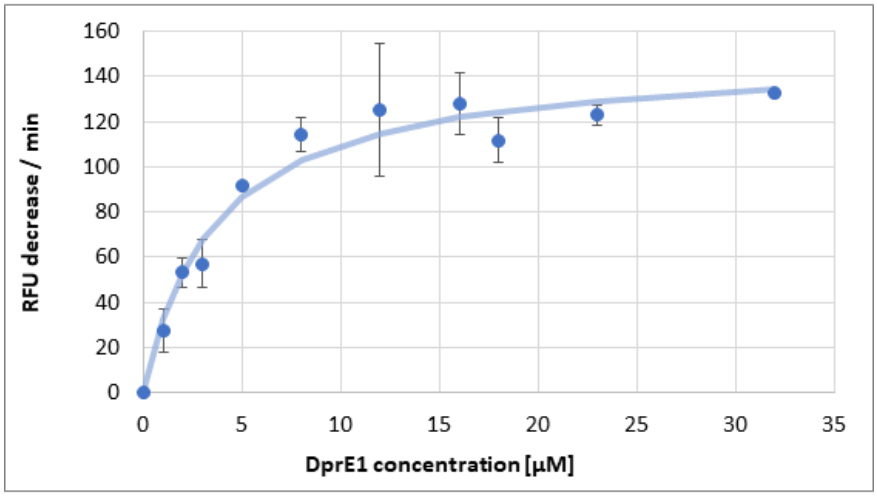


	buffer	NAD(P)H ctrl	GGPR ctrl	0 µM GGPR	180 µM GGPR
40 mM HEPES pH 7.25 (1M)	+	+	+	+	+
80 mM NaCl (2M)	+	+	+	+	+
300 µM NAD(P)H (5 mM)		+	+	+	+
7.2 mg/ml His-Mt-DprE1 +Mt-DprE2 (10.8 mg/ml)				+	+
180 µM GGPR (1.5mM)	+		+		+
MilliQ addition for 15 µl	+	+	+	+	+

Figure 3.22. Co-factor testing with co-expressed 6xHis-Mt-DprE1 and Mt-DprE2 in enzyme activity assay. Expression as in construct 3 in **Table 3.1. Enzymes purified with IMAC.** (A) NADH and (B) NADPH in the similar reaction setup. RFU is relative fluorescence unit.

Assays with purified 6xHis-SUMO-Mt-DprE2 required additional Mt-DprE1 for the coupled reaction to take place (see **Figure 3.19.**). As a general guideline, to minimise the effect of Mt-DprE1 limiting Mt-DprE2 activity, an excess of the first enzyme is needed. With an assay trial (**Figure 3.23.**) it was determined as minimum 10 µM DprE1 for reaction with 0.5 mg/ml 6xHis-SUMO-Mt-DprE1.

Also, Mt-DprE1 oxidised NADH by itself in some reactions, therefore only fresh, exceptionally pure protein samples could be applied in the assay. Purity was enhanced by larger volumes in washing steps than usual.



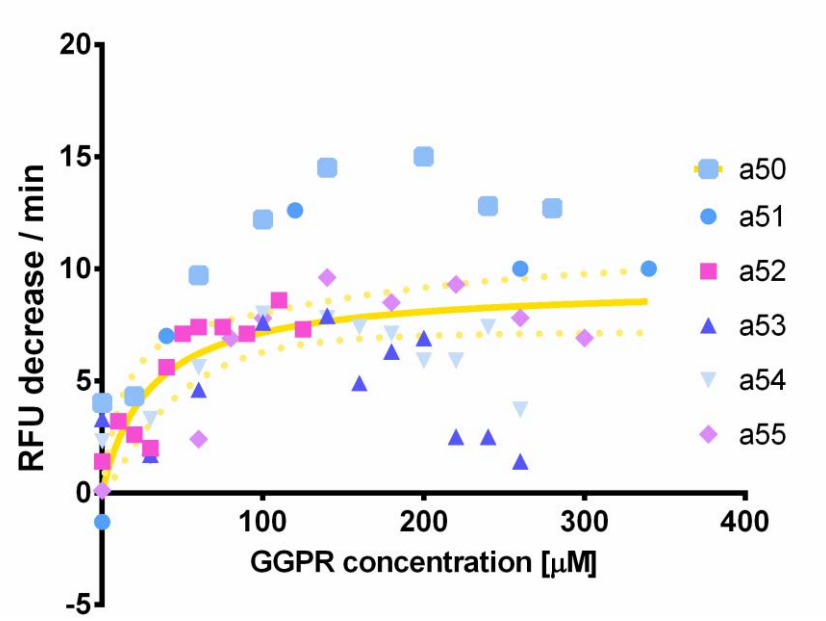
Reaction mixture:
40 mM HEPES pH 7.5 (1M)
80 mM NaCl (5M)
150 µM NADH (3 mM)
Mt-DprE1 (2.24 mg/ml = 35.2 µM)
0.5 mg/ml 6xHis-SUMO-DprE2 (3 mg/ml)
180 µM GGPR (2 mM)
MilliQ water addition for 10 µl

Figure 3.23. Optimisation of Mt-DprE1 concentration for enzyme activity assay with 6xHis-SUMO-Mt-DprE2. DprE1 was purified with IMAC and IEX, while DprE2 only with IMAC. Stock solution concentrations in parentheses. RFU is relative fluorescence unit, error bars are standard deviations of triplicate measurements. Bright blue curve represents the non-linear regression best fit of the Michaelis-Menten equation to the data, V_{max} = 149 RFU/min, K_m = 3.6 µM.

Another rate limiting factor of the Mt-DprE2 enzyme activity assay is the concentration of GGPR. On the one hand, it has to be high to allow GGPX formation and supply excess substrate for Mt-DprE2 to measure accurate maximum reaction velocity. But also, it cannot be too high, because it would denature Mt-DprE1. Partial denaturation of Mt-DprE1 at GGPR concentrations above 200 µM were observed (**Chapter 2.3.2.**).

Surprisingly, measurements even with the same enzyme and similar reaction mixture (different only in GGPR concentration) delivered highly different results (**Figure 3.24.**). The general V_{max} value was less than 10 RFU/min, at a standard error of 1 RFU/min, while the coefficient of determination (R^2) was 0.349 for global data. Consequently, this assay setup

does not have the sensitivity to report precise enzyme kinetic data, but it does indicate enzyme activity and inactivity with optimised conditions.

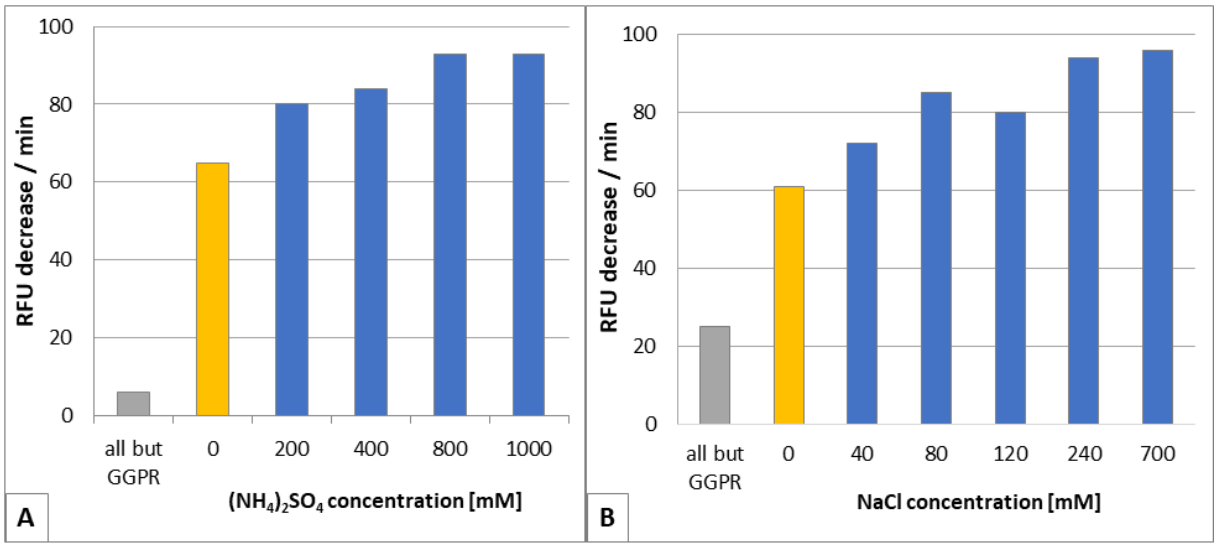


Reaction mixture:		
50 mM BTP pH 7.2 (0.5M)		
100 mM $(NH_4)_2SO_4$ (3.6M)		
180 µM NADH (2 mM)		
0.24 mg/ml 6xHis-DprE1+DprE2 (0.6 mg/ml)		
varying GGPR (0.5 mM)		
MilliQ water addition for 10 µl		
Michaelis-Menten global curve fit		
		Standard Error
Vmax	9.276	1.065
Km	29.7	15.39
Degrees of Freedom	54	
R square	0.359	

Figure 3.24. GGPR dependence of enzyme activity assay with co-expressed 6xHis-Mt-DprE1 and Mt-DprE2. Enzymes were co-purified with IMAC. Yellow curve is the global fit for all data presented from assays 50-55. Dotted curves show the 95% confidence interval for the fit. Measurements over 300 µM GGPR showed decreased/no activity. K_m in [µM], V_{max} in [relative fluorescence unit / min].

It is worthy to note that GGPR oxidises some NADH by itself (see **Figure 3.22.** assay 'GGPR ctrl' versus 'NADH', 'NADPH' controls). This is likely due to hydrolysis of GGPR to ribose and prenylphosphate, with ribose capable of reacting with NAD(P)H.

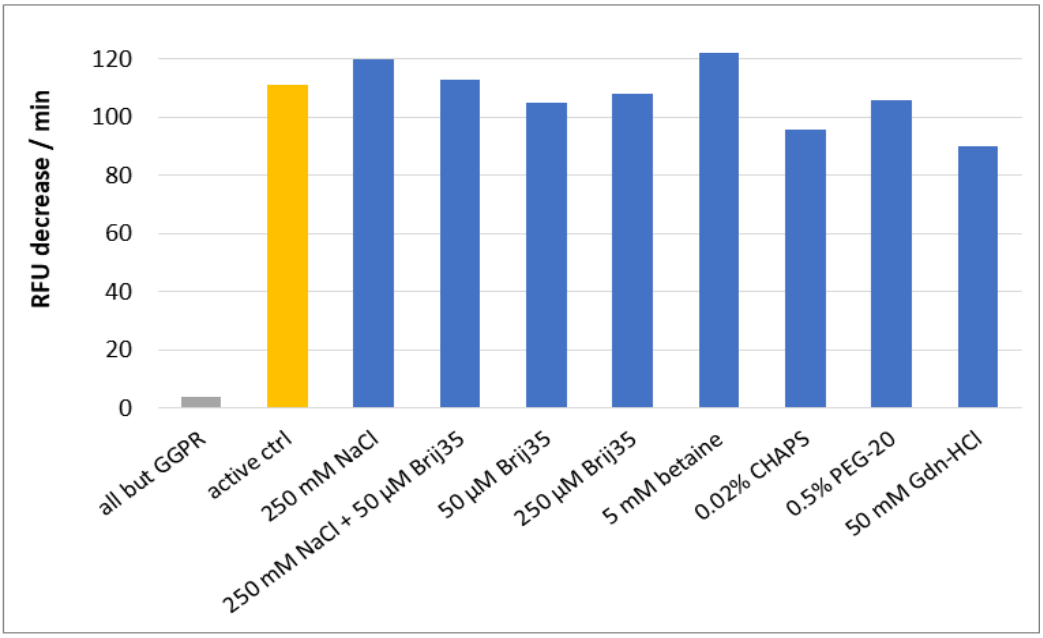
The addition of ammonium sulphate stabilised 6xHis-SUMO-Mt-DprE2 in the thermal shift assay (**Figure 3.17.B**). Enzyme activity assay observations confirmed that ammonium sulphate supported enzyme stability and hence enzyme activity increased. Both NaCl and ammonium sulphate delivered similar increase in reaction rates (**Figure 3.25.**).



Reaction mixture in (A)	Reaction mixture in (B)
80 mM BTP pH 7 (0.5M)	40 mM HEPES pH 7.5 (1M)
ammonium sulphate (4M)	NaCl (1M)
50 mM NaCl (2M)	50 mM NaCl (2M)
180 µM NADH (3 mM)	180 µM NADH (3 mM)
12 µM 6xHis-Mt-DprE1 (7.2 mg/ml = 113 µM)	10 µM 6xHis-Mt-DprE1 (7.2 mg/ml = 113 µM)
0.34 mg/ml 6xHis-SUMO-Mt-DprE2 (1 mg/ml)	0.75 mg/ml 6xHis-SUMO-Mt-DprE2 (3 mg/ml)
150 µM GGPR (2.5 mM)	150 µM GGPR (1.5mM)
MilliQ water addition for 10 µl	MilliQ water addition for 10 µl

Figure 3.25. Salt addition testing in enzyme activity assay with 6xHis-SUMO-Mt-DprE2. (A) With ammonium sulphate and (B) with NaCl. Concentrations of stock solutions in parentheses. RFU is relative fluorescence unit. DprE1 was purified with IMAC and IEX, while DprE2 with IMAC.

Potential benefit of other additives was tested similarly to the salt addition tests. Detergents, salts, osmolytes are known for stabilising certain proteins, keeping them in solution and hinder aggregation (Lebendiker et al., 2014). This screen was different to the storage condition experiments (**Chapter 3.3.6.**) in the way that no freeze-thaw of the protein was conducted, only the possible benefit of additives in the 6xHis-SUMO-Mt-DprE2 enzyme activity assay was evaluated.

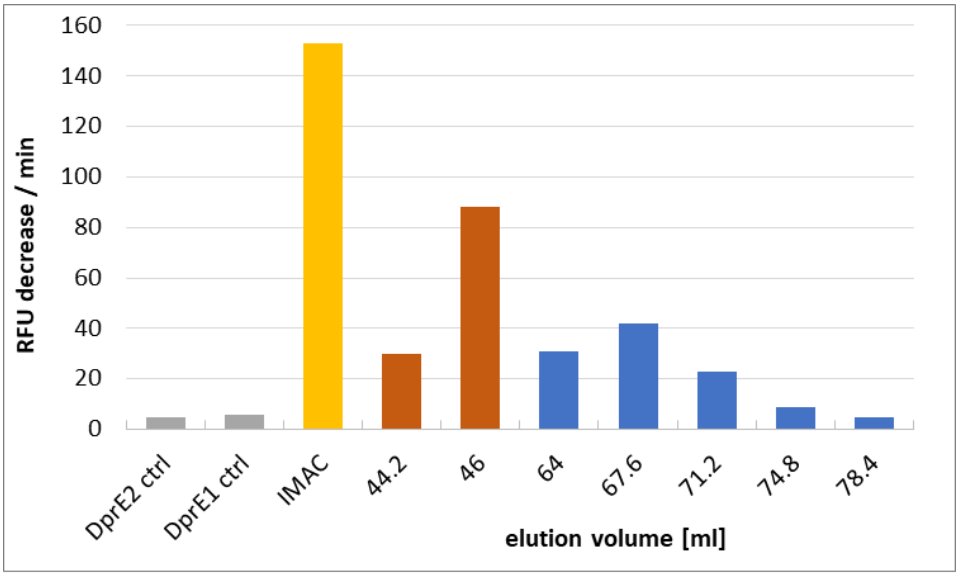


Reaction mixture:
40 mM HEPES pH 7.5 (1M)
80 mM NaCl (1M)
180 µM NADH (3 mM)
10 µM 6xHis-DprE1 (7.2 mg/ml = 113 µM)
0.75mg/ml 6xHis-SUMO-DprE2 3 mg/ml)
150 µM GGPR (2.5 mM)
+ additive from stock:
Brij-35 (0.012%=1 mM) (Sigma, #B4184)
betaine (100 mM) (Sigma, #B2629)
PEG-20 (4%) (Sigma, #81300)
Gdn-HCl pH 7.5 (0.5M) (Sigma, #50950)
CHAPS (0.1%) (Sigma, #C3023)
MilliQ water addition for 10 µl

Figure 3.26. Additive screen in 6xHis-SUMO-Mt-DprE2 enzyme activity assay. Stock solutions concentration and product numbers in parentheses. RFU is relative fluorescence unit. DprE1 was purified with IMAC and IEX, while DprE2 with IMAC.

The results show (**Figure 3.26.**) that salt addition and the osmolyte betaine increased the rate of reaction, detergents Brij-35 and CHAPS lowered it. PEG-20 had no additional benefit. Guanidine hydrochloride (Gdn-HCl) lowered enzyme activity as expected and it served in the experiment as a negative control.

3.3.8 Evaluation of SEC fractions in enzyme activity assay



Reaction mixture:
40 mM HEPES pH 7.5 (1M)
80 mM NaCl (1M)
180 µM NADH (3 mM)
10 µM 6xHis-DprE1 (7.2 mg/ml = 113 µM)
5 µl of SEC fraction sample
180 µM GGPR (2 mM)
MilliQ water addition for 10 µl
Storage buffer of SEC samples:
55 mM HEPES pH 8
180 mM NaCl
10% glycerol

Figure 3.27. Enzyme activity measurement of 6xHis-SUMO-Mt-DprE2 size exclusion chromatography fractions. 'IMAC' is protein sample after IMAC purification, before loaded on SEC column, A_{280} = 5. Controls contain only that one enzyme from the coupled reaction. RFU is relative fluorescence unit.

SEC of both 6xHis-SUMO-Mt-DprE2 and Mt-DprE1-Mt-DprE2 co-expression produced very pure, low concentration preparations of the target protein(s) (**Chapter 3.3.4.**).

SEC fractions of 6xHis-SUMO-Mt-DprE2 were used in enzyme activity assay and showed that the aggregate protein fraction was active (see **Figure 3.27.**), but lower than the preparation before the SEC purification, which is best explained by the low concentration of the protein present in pure SEC fractions.

Samples of the size exclusion chromatography fractions with 6xHis-Mt-DprE1 and Mt-DprE2 co-expression were tested in enzyme activity assay but were all inactive, likely due to low protein concentrations.

3.4 Discussion

3.4.1 Overcoming aggregation of Mt-DprE2

Previous experiments purifying Mt-DprE2 from *E. coli* cell extracts provided only an amount of protein insufficient for crystallography. His-tagged Mt-DprE2 fusion protein expressed from a pET vector aggregated during purification and required high concentrations of glycerol for stabilisation (unpublished work by Sarah M. Batt, University of Birmingham).

In this thesis we have explored the possibilities of using other tags for both purification and chaperon-like purposes. ‘Maltose binding protein’ (MBP) and ‘small ubiquitin-like modifier’ (SUMO) proteins have been reported to enhance soluble expression levels of certain proteins (Arbing et al., 2013; Peroutka Iii et al., 2011). Both are commercially available, also their protein constructs are designed to be cleavable by proteases, Factor Xa and SUMO protease respectably, so the tags are removable after purification.

The fusion of Mt-DprE2 with 6xHis-SUMO was successful in a way that it overexpressed the protein in sufficient concentrations for enzyme activity assay, studies by size exclusion chromatography and thermal shift assays.

The size exclusion chromatography demonstrated that most of the protein formed an aggregate (**Figure 3.11.**), which was despite aggregation enzymatically active (**Figure 3.27.**). Fractions attributed to an oligomer, most likely a dimer, were active as well, performing with worse reaction rates due to lower enzyme concentrations. These experiments indicate that the tag-fusion with 6xHis-SUMO did not prevent aggregation altogether, however, the obtained enzyme aggregate has the potential to report on enzyme activity and therefore to be applied in inhibition assays.

Another strategy revolved around the idea that Mt-DprE1, as a part of Mt-DprE2 interactome (Jankute et al., 2014), may stabilise the structure of Mt-DprE2. Therefore, co-expression and co-purification of Mt-DprE2 with Mt-DprE1 could prove to be beneficial. Similar work has been done with the result of low-level expression of either of the proteins (unpublished work by Sarah M. Batt, University of Birmingham). Also, it has been further investigated with the co-expression of Mt-GtrA protein (*rv3789*), which is part of the same operon as *dprE1* and *dprE2* (see **Figure 1.11.**). It has been reported that Mt-GtrA is likely to act as an anchor for AftA, advancing DPA incorporation into the cell wall (Kolly et al., 2015) and/or as a DPA flippase (Larrouy-Maumus et al., 2012) .

In this work the co-expression experiment and also compared to a new construct producing Mt-DprE2, His-Mt-DprE1 and chaperon proteins in the same cell simultaneously, but from separate plasmids (see constructs in **Table 3.1.**).

Both co-expression of five target proteins in the same cell (construct 1 in **Table 3.1.**) and the MBP-tagged periplasmic expression failed for the reason of resulting in excessively small

amounts of Mt-DprE2 (see **Chapter 3.3.1.**). This result was anticipated since periplasmic extractions are prognosed to result in one tenth of the amount of MBP-fusion protein expressed cytoplasmically (Inc, n.d.). Logically, fine tuning of expressions levels of five different proteins is also a difficult task that was expected as a set-back in Mt-DprE2 production.

Simplification of the co-expression construct, omitting Mt-GtrA, lead to the production of sufficient amounts of protein for further studies. Mt-DprE2 was observed to co-purify with the His-tagged Mt-DprE1 in affinity chromatography (**Figure 3.7.**) and samples were active in enzymatic assay (**Figure 3.22.**). Size exclusion chromatography of the proteins (**Figure 3.14.**) showed that the aggregate and the oligomer (65 kDa in size, suspected 1:1 complex of 6xHis-Mt-DprE1 and Mt-DprE2) contained different additional protein species. In fact, the oligomer on SDS-PAGE showcased two unidentified bands at molecular weight between the suspected Mt-DprE1 and Mt-DprE2 proteins (**Figure 3.14.**). Unfortunately, all fractions were so diluted that enzyme activity assessment or further studies were not rational to pursue. Repeat experiments resulted also in small protein quantities. Therefore, it remains to be solved how Mt-DprE1 forms a complex with Mt-DprE2, to confirm that it is Mt-DprE2 in the complex and to identify the co-eluting proteins.

Co-expression of 6xHis-Mt-DprE1 and Mt-DprE2 are yet the physiologically most relevant preparations since only a small peptide (6xHis-tag is 841 Da) is fused to Mt-DprE1, and Mt-DprE2 is produced in its physiological amino acid sequence. However, even in the presence of chaperons (**Figure 3.7.**) the co-expression produced low levels of the heterologous proteins, along with typical contaminating proteins that are co-purified in metal affinity chromatography with the construct in high abundance (**Figure 3.9.**). This is detrimental to

the yield of Mt-DprE2 and constrains additional protein separation steps to gain homogenous protein.

One observation with implication for further studies is that the 6xHis-SUMO-fused Mt-DprE2 protein in size exclusion chromatography displayed an array of degradation products (**Figure 3.11.**), which retained the His-tag on their N-terminus (**Figure 3.13.**). Although, since a high molecular weight protein band was also labelled with the anti-His antibody in the Western blot, one cannot rule out that these are not exclusively degradation products, but also include some non-specifically detected proteins by the antibody. But if the C-terminus of the protein is prone to degradation, it can be truncated or fused to a tag in future studies to prevent degradation.

Another finding of the 6xHis-SUMO-Mt-DprE2 experiments was that the oligomer state did not contain NADH/NADPH based on absorbance at 340 nm (see size exclusion chromatogram in **Figure 3.11.A**). This could be due to loss of co-factor during purification. Taken together with the results of the thermal shift assay in which NADH and NADPH did not stabilise the protein (**Figure 3.17.**), and the findings of **Chapter 4**, which resulted in spontaneous resistant mutants of DprE2 hit compounds affecting F_{420} co-factor production, it cannot be ruled out that F_{420} is an alternative or native co-factor of Mt-DprE2. This co-factor is not present in the *E. coli* expression system, and should be investigated in alternative expression systems, such as *M. smegmatis* that natively produce and utilise F_{420}.

In conclusion, none of the assessed methods in this work prevented completely the aggregation of Mt-DprE2, therefore for protein crystallography, in the future the focus of research has to be on producing high amounts of protein and attempt to concentrate the sample with minimal additional aggregation in the process. Unfortunately, it is also possible that further chromatography steps would be necessary to eliminate the contaminating

proteins remaining after metal affinity chromatography. A concentrated, but soluble pure preparation may be subjected to size exclusion chromatography and oligomeric fractions may be used for crystal growth. This purification step separating oligomers or a complex with Mt-DprE1 from large aggregates seems to be the next milestone towards protein preparations for crystallography.

3.4.2 Analysis of enzyme activity assays

The enzyme activity assay was optimised regarding several factors. The conditions supporting the highest reaction rate were pH value near 7, 80 mM or higher salt concentrations, either NaCl or Li_2SO_4 resulted a reaction stable and a well distinguishable signal difference between active and inactive reaction (see **Figures 3.15., 3.17., 3.25.,** respectively). The reaction tolerated glycerol well up to 10% (**Figure 3.18.**). For the instruments range of signal linearity was established. Furthermore, GGPR reaction substrate concentration maximum was investigated and found to be similar to the one observed in the Mt-DprE1 assay studies (**Figures 3.24.** and **2.7.**). The 6xHis-SUMO-Mt-DprE2 at 0.5 mg/ml protein preparation required approx. 0.65 mg/ml = 10 μM 6xHis-Mt-DprE1 to reach optimal reaction rate (**Figure 3.23.**).

An unexpected finding of the enzyme activity assays was the high variability and differences between the performance of enzyme preparations. One reason for this phenomenon was the varying Mt-DprE2 content in Mt-DprE1-Mt-DprE2 co-expression studies, where Mt-DprE2 could appear to be present in a 1:1 ratio to Mt-DprE1 (**Figure 3.9.**) or even barely noticeable on SDS-PAGE analysis (**Figure 3.14.**). This unpredictability poses a huge problem for enzyme activity assays intended for large sample sizes (like an HTS campaign). One option to ensure active enzymes for assays is to produce large amounts of enzyme at the same time and run a quality test enzyme activity assay, after which smaller aliquots could be frozen in

the explored preferable storage buffers, that contained high osmotic pressure agents, such as sucrose, trehalose or glycerol (**Chapter 3.3.6.**). This way a standard enzyme preparation could be ensured for a limited number of assays.

An additional concern about the enzyme activity assay originates in its lack of refinement in GGPR dependence. GGPR is the substrate of the coupled reaction of Mt-DprE1-Mt-DprE2, and results showed that even with the same enzyme preparation and set-up significant deviations are common between re-runs (**Figure 3.24.**). The apparent K_m of the co-expressed protein activity assay was measured as 30 μM±15 μM with DprE1-DprE2 co-expressed preparation. In general, it is recommended (Acker and Auld, 2014) that inhibitor screens should be performed with a set substrate concentration near the K_m to detect competitive and uncompetitive inhibitors without bias. Lower substrate concentrations are biased to recognise competitive inhibitors that bind to the active site and compete with the substrate, and miss uncompetitive inhibitors, which bind to the enzyme-substrate complex. Meanwhile, higher substrate concentrations achieve the opposite effect (Acker and Auld, 2014). With the current low signal-to-noise ratio between active and inactive reaction even at higher substrate concentrations (approx. 200 μM), it is unwise to lower the GGPR concentration to the K_m value. Consequently, a high enough GGPR concentration that produces a good signal also favours uncompetitive inhibitors and may underrepresent competitive inhibitors. Non-competitive inhibitors, that bind allosteric sites of the enzyme, are unaffected by the substrate concentration in the reaction.

Another matter that needs to be addressed at the level of assay design is that the read-out of the developed enzyme activity assay in this work is fluorescence based with excitation at 340 and emission at 460 nm. The compounds of interest need to be tested for fluorescence at these values to rule out false results in the screen, which may be challenging with large

compound libraries and render the assay not applicable for compounds that are fluorescent or quench at the same wavelength ranges.

As for the different preference of enzyme constructs towards co-factors, 6xHis-SUMO-Mt-DprE2 was found to perform only with NADH, while tag-free Mt-DprE2 had a better signal-to-noise ratio with NADPH, although it could utilise NADH as well (**Figures 3.21** and **3.22.**). This imposes a difference of the 6xHis-SUMO fusion protein compared to the tag-free protein. For lack of 3D conformational data, it is impossible to tell how and at what extent does the tag fusion affect enzyme function. Therefore, Mt-DprE1-Mt-DprE2 co-purifications are to be promoted to be used in inhibitor assays until the fusion protein is not proven to function equally.

Finally, a flaw in assay design regarding inhibitor screens is that Mt-DprE1 inhibitors are not distinguished from Mt-DprE2 in the assay read-out. Results have to be compared to Mt-DprE1 enzyme activity assays with the same compounds for reliable reporting.

Considering all these results, and the low success rate of target-based enzyme activity assays in drug discovery (Cole, 2016), it can be concluded that the enzyme activity assay developed herein is only meaningful and applicable as an auxiliary assay for a Mt-DprE2 targeting whole-cell assay to test potential hits against Mt-DprE2, because it is limited by many constraints.

4 High-throughput screening against Mt-DprE2 and hit assessment

4.1 Introduction

The protein Mt-DprE2 is an essential enzyme of *M. tuberculosis* (Kolly et al., 2014) and a potential target of new drugs against mycobacteria. In this Chapter a target-focused whole-cell screen against the TB box compound library (unpublished set by GSK), a library of 11,096 compounds with activity against mycobacterial species was performed to identify potential hit molecules.

The set-up of the experiment is based on the approach that both strategies of target-based and phenotypic-based discovery could be implemented simultaneously, combining their advantages. The approach of selecting compounds with activity against mycobacteria and screening for activity against strains that overexpress a particular target aims to minimise compound attrition due to inactivity (failure to penetrate the cell wall for example). Target overexpression at the same time gives an early indication whether the inhibitor targets specifically Mt-DprE2. When there is a shift in cell growth during compound exposure due to target overexpression it may be due to the enhanced target concentration in the cell. The shift is further confirmed as a shift in MIC in a larger assay format thereby validating target engagement, with selected hit compounds further tested in thermal shift assays and in a Mt-DprE2 enzyme activity assay, which was developed in **Chapter 3.** This set of experiments has the potential to select among whole-cell active compounds those that target Mt-DprE2.

4.2 Materials and methods

4.2.1 Bacterial cultures

M. bovis BCG was grown at 37°C in 5% CO_2 atmosphere, in Middlebrook 7H9 broth with 0.025% Tween-80 and 10% albumin-dextrose-catalase (ADC) supplement, or on Middlebrook 7H10 plates with 10% oleic acid-ADC (OADC) enrichment. Both media were supplemented with kanamycin at a 25 μg/ml final concentration. Liquid cultures were passaged for a maximum of two months, then started a new from frozen stocks (1 ml aliquot of liquid culture with 30% glycerol cryoprotectant, stored at -80 °C).

Overexpression of Mt-DprE2 was achieved by transforming *M. bovis* BCG cells with plasmid pMV261-Mt-DprE2 (Parish and Roberts, 2015). The plasmid was cloned and kindly provided by Sarah M. Batt, University of Birmingham. Plasmid pMV261 is a shuttle vector stably maintained in both *E. coli* and mycobacteria *in vivo*, which can constitutively express proteins in excess from an extrachromosomal copy of the selected gene, under the control of *hsp60* stress protein regulatory region, with a kanamycin-resistance selection marker (Stover et al., 1991).

4.2.2 Primary screening

The collection of TB Box compounds was requested from GSK Harlow UK in a 384-well plate format (Greiner, #781091) in two concentrations and in duplicate. Selected screening concentrations were 1 and 5 μM (50 nl and 250 nl compound/well). On each plate column 6 served as live cell control (only DMSO normalisation), and column 18 as a dead cell control by adding rifampicin (same as in **Figure 4.1.**). The plates were stored overnight at -80 °C before the screen to confirm sterilisation.

M. bovis BCG transformed with pMV261 (control, ctrl) and pMV261-Mt-DprE2 (overexpressor, oe) were grown to mid-log phase (OD_{600} = 0.4-0.8), diluted to 10^5 CFU/ml (considered OD_{600} = 0.125 equal to 10^7 CFU/ml). 50 µl of culture was aliquoted to each well by MultiDrop Combi (Thermo Scientific), covered with a sterile lid without condensation rings (#656161 Greiner), wrapped in parafilm and aluminium foil, packed in plastic boxes along with tissue soaked into water to ensure humidity, and kept at 37 °C in a 5% CO_2 incubator for 7 days.

After incubation with compounds the published REMA protocol (Sharma et al., 2014) was followed and scaled down to small volume plates. In short, a 1:1 ratio of 0.02% resazurin (in MilliQ water) and 10% Tween-80 was mixed and filter-sterilised, then 10 µl added to each well on the plate by MultiDrop Combi. Plates were sealed with the lid and wrapped in aluminium foil and incubated overnight at 37 °C in a 5% CO_2 incubator. Plates were sealed with adhesive film (A5596 EASYSeal) and read with EnVision spectrophotometer (Perkin Elmer), with fluorescence filters at 531/595 nm (excitation/emission wavelength). Collected data were analysed in Microsoft Excel.

The complete screen with the TB Box was performed twice, because the Z' values of the first screen were very low.

4.2.3 Secondary screening

Initial hits were tested again in a 384-well plate format, only in a dose-response manner as a 2-fold dilution series from 100 to 0.098 µM. Each plate contained the dilution series of 32 different compounds, column 6 and 18 dedicated to live and dead cell controls, respectively (**Figure 4.1.**) with DMSO content normalised to 1%.

Two repeats were tested against each strain, replicas were on different plates. These plates were dispensed and shipped from GSK Harlow, UK. Assay protocol exactly as in HTS.

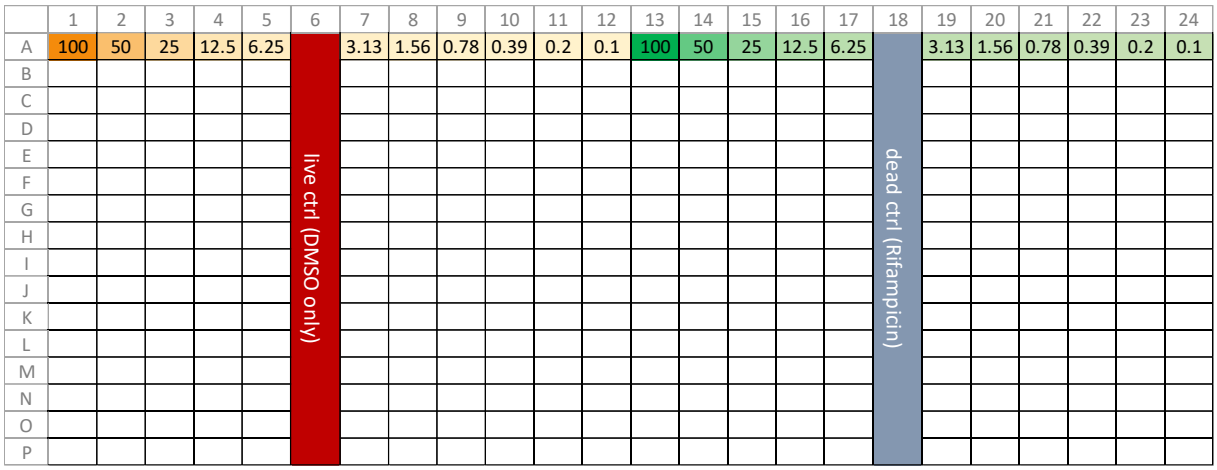


Figure 4.1. The plate layout for dose-response testing in 384-well plate format. Concentration values in [µM]. Orange and green colouring depicts the two different compounds in the same row.

Hits that reproduced an MIC shift in a 2-fold dilution series in 50 µl in 384-well plate format were further evaluated in a 96-well format with 100 µl culture volume. Compounds were distributed as 2-fold dilution series by hand from 10-fold master plates or by HP D300 Digital Dispenser, leaving an evaporation frame empty of compound all around the plate (first and last row, first and last column on the plate). DMSO concentration was maximised at 2.8% to avoid cell toxicity by the compound solvent. The cell growth inhibition experiment protocol was very similar to the HTS above, only scaled up based on the REMA protocol (Taneja and Tyagi, 2007). The main difference was that for 100 µl of culture a 30 µl mixture of 0.02% resazurin and 12.5 µL 10% Tween-80 were added, and the plate readings were performed using a SpectraMax spectrophotometer, with monochromators set to 530 nm excitation and 590 nm emission wavelength.

4.2.4 Data evaluation

Firstly, raw fluorescence data were imported from raw data files and plate layouts annotated manually in Microsoft Excel. Then fluorescence data were converted to percentage of inhibition (%inh):

$$\%inh = \frac{FI - dead\ ctrl}{live\ ctrl - dead\ ctrl} * 100$$

where FI is fluorescence intensity of the sample well. Standards for dead and live cell in 384-well plate layouts data evaluation controls were taken from the same row on the plate as the analysed sample well to compensate for the extensive edge effect. In 96-well plate layouts, where evaporation frames were included, the average of dead and live control values was calculated and applied in the formula.

As quality assessment to each plate the Z' value (Zhang et al., 1999) was calculated (also called Zprime or Z-factor):

$$Z' = 1 - \frac{3 * SD\ of\ live\ ctrl + 3 * SD\ of\ dead\ ctrl}{mean\ of\ live\ ctrl - mean\ of\ dead\ ctrl}$$

where SD is standard deviation. Based on guidelines adopted at GSK this value should be above the significance threshold 0.4, or else a repeat measurement is recommended. For better representation the Z' calculations of two examples are portrayed in **Figure 4.2.**

MICs recorded in liquid culture assays were MIC50 values, where values refer to the compound concentration that elicited 50% percentage of growth inhibition (%inh=50%). However, on solid media this value is not measurable, hence the concentration of compound that allowed no colony growth was taken as MIC. This concentration value does not directly translate to percentage of inhibition, but sometimes in literature is referred to as MIC95 (%inh=95%) (Blanco-Ruano et al., 2015).

4.2.5 Mt-DprE2 inhibition assay

In order to assay enzymatic activity of purified Mt-DprE2, oxidation of the redox co-factor NAD(P)H was monitored by measuring fluorescence. The co-enzyme was exclusively

NADH in GSK experiments, and NADH and NADPH in University of Birmingham. Exact reagent mixtures are portrayed along specific reported reaction curves.

The general inhibitor assays in GSK DDW contained 50 mM bis-tris propane at pH 7.0, 100 mM ammonium-sulfate, 50 μM Brij-35 surfactant, 180-250 μM NADH, 100 μM compound or 3% DMSO and MilliQ water. Reaction mixtures were split from the master mix to wells of a small-volume microplate (Greiner, #784076), and enzymes were added to the mixture at maximum 50% of the assay volume (total assay volume 10 μl). His-Mt-DprE1 and Mt-DprE2 co-purification samples were used, also 6xHis-SUMO-Mt-DprE2 purification samples (see **Chapter 3** for detailed method), in which case His-Mt-DprE1 was added in excess. Plates were kept at room temperature (25-26°C) for 5 minutes inside the spectrophotometer. The substrate GGPR was added last and initiated the reaction which was monitored by fluorescence measurement at 340/462 nm using a SpectraMax M5 spectrophotometer, with measurement points taken at every 15 or 30 seconds for 50 minutes.

The general inhibitor assays in University of Birmingham contained enzyme at 50% assay volume (approx. 7.5 mg/ml total protein concentration of Mt-DprE1-Mt-DprE2 co-purification), 40 mM HEPES pH 7.5, 80 mM NaCl, 400 μM NADPH, 333 μM inhibitor or 3.3% DMSO, 280 μM GGPR added last to initiate reaction. The reaction was performed in a 15 μl volume, kept at 25 °C, and fluorescence monitored at 350/450 nm using a PHERAstar FS microtiter plate reader for 50 minutes, with signal linearity in the first 12 minutes.

4.2.6 Spontaneous resistant mutant generation

Previous attempts for mutant generation have failed against hit compound 15 (in **Table 4.3.**, hit compound of PhD thesis work by Giacomo Chiodarelli, University of Birmingham). No colonies were observed on 7H10-OADC solid media supplemented with compound 15 at 5-

10-times the MIC95 (unpublished). A liquid culture-based protocol was performed, where cells are first grown in liquid culture at near MIC concentrations before plating on compound-containing plates, which did not deliver any mutant colonies either.

A different approach on the other hand generated mutants against certain derivatives of compound 15 and other hits of the Mt-DprE2 HTS. This method is based on the serial passaging technique of Leeds *et al.* (2014) scaled to the 24-well SPOTi assay format (Gupta and Bhakta, 2012).

Firstly, 24-well plates (Corning, #CLS3524) were made for solid MIC measurement as in the SPOTi assay, but instead of the generally applied 2-fold dilution series, I opted for 1.25-fold and 1.5-fold dilution series. Therefore, cells were exposed to several near-MIC concentrations of the compounds. Each well contained 1 ml of Middlebrook 7H10 solid media with 10% OADC supplement, 0.5% glycerol, 0.025% Tween-80 and the tested compound (stock solution in neat DMSO, 0.1% DMSO final concentration in solid media), live control contained 2% DMSO. Made three replicate plates of each layout (plate 1, plate 2 and plate 3), and kept the unused plates at 4 °C until handling.

Compounds were selected from the HTS hits (see **Chapter 4.3.2.** for selection criteria) based on the highest shift of MIC50 in the dose response experiments. Derivatives of compound 15 were selected based on solubility and potency in dose response assay (compound series generated by Giacomo Chiodarelli, University of Birmingham).

	1	2	3	4	5	6	
A	0.09	0.14 MIC	0.21	0.31	0.47	0.70	cmp 15
B	1.05	1.58	2.37	3.56	5.33	8.00	
C	0.17	0.26	0.39	0.59	0.88	1.32	cmp 27
D	1.98	2.96 MIC	4.44	6.67	10.00	live ctrl	

Table 4.1. Example of mutant generation layout in 24-well plate format. 1.5-fold dilution series, concentration values are in [µM].

Liquid culture of wild type *M. bovis* BCG was grown in Middlebrook 7H9 broth with 10% ADC supplement, 0.5% glycerol and 0.025% Tween-80 to log phase (OD_{600}=0.4-0.8), then 2 µl of culture were spotted on each well. Plate lids were sealed with parafilm and aluminium foil, and plates kept at 37 °C in a 5% CO_2 incubator. The first aggregate colonies normally appeared after two weeks on the plate at the location of the spot, then MICs could be determined. Cells growing at the highest concentration with the given compounds were then inoculated from plate 1, and grown in compound-free liquid media to log phase.

At this stage plate 1 was still not discarded, colony appearance was observed up to 6 weeks after incubation. With some compounds already from plate 1 single colonies were found on higher than MIC conditions, these were also inoculated and grown in liquid media. If there was a mutant with high compound tolerance, then that culture was spotted on to plate 2. If there was no mutant colony, then the culture with the highest tolerated concentration from plate 1 was spotted on the relevant compound series. This strategy was repeated from plate 2 to plate 3 in a similar manner.

Mutant strains were archived to -80 °C in a 30% glycerol stock, and four mutant genomic DNA samples were sent for whole genome sequencing to Genomic Facilities of University of Birmingham.

4.2.7 Genomic DNA extraction

50 ml of *M. bovis* BCG culture was grown to log phase (OD_{600}=0.4-0.8) in Middlebrook 7H9 broth with 10% ADC supplement, 0.5% glycerol and 0.025% Tween-80. The culture was pelleted by centrifugation on 4 °C, then resuspended in 450 µM Qiagen P1 buffer (#19051, 50 mM Tris-HCl pH 8.0, 10 mM EDTA, 100 µg/ml RNaseA (Qiagen, 2005)) supplemented with 50 µl 10 mg/ml lysozyme solution. Samples were incubated at 37 °C overnight, then 100 µl 10% SDS solution and 50 µl 20 mg/ml proteinase K was added for 30 min incubation at 55 °C. After the addition of 200 µl 5M NaCl solution the samples were heated to 65 °C and kept for 15 min. A 1 ml of chloroform:isoamylalcohol (24:1) mixture was added and gently mixed. The aqueous layer was separated by centrifugation (10 min, room temperature, 13,000 rpm on a bench-top centrifuge), transferred to a sterile tube and the washing step with chloroform:isoamylalcohol repeated. Then the aqueous layer was transferred again to a sterile tube, and 700 µl ice-cold 100% isopropanol was added and gently mixed. DNA was then pelleted by centrifugation at 4 °C, 13,000 rpm for 1 hour, and the supernatant gently removed, the pellet washed with 700 µl of 70% ice-cold ethanol. The sample was pelleted again at 4 °C, 13,000 rpm for 1 hour. Ethanol was gently removed, and sample was left at room temperature in a fume hood until all the solvent evaporated. Finally, the pellet was resuspended in 25 µl MilliQ water and stored at -20 °C, while 1 µl of each gDNA preparation sample was loaded onto 0.8% agarose gel for visualisation. Samples with satisfactory amount of gDNA were sent for whole genome sequencing.

4.2.8 Thermal Shift Assay

As described in **Chapter 3.2.4.** with the difference, that in some cases Filter 4 output of the real time PCR thermocycler (recommended with ROX dye according to manufacturers' manual) resulted in less noisy signal than with filter 3 (recommended with TAMRA dye). Therefore, filter 4 output data were analysed.

4.3 Results

4.3.1 Whole-cell screens against Mt-DprE2 in *M. bovis* BCG

The first results from the HTS showed a much less robust screen than anticipated. Z' values, that are expected to be 0.4 or higher for a successful screen, were well below this value. From the total of 132 plates only 67 passed this limit. As a lower limit when Z'>0 was accepted, then 94 plates passed the quality assessment, which means 29% of the plates failed even this permissive evaluation. See **Figure 4.2.** with examples for acceptable and unacceptable plates by Z' analysis. Plates with failed Z' values were not excluded from the hit selection process, but a high number of false positive hits were expected.

Low Z' values can be explained by the edge effect (Carralot et al., 2012). In this specific case, this observation may be explained by high evaporation with a typical central pattern on almost every plate. Since the 384-well plates are fitted with a lid without condensation rings, visible amounts of vapour accumulated in the central part of the plates, likely originating from the wells in the middle. As the protocol describes, these lids are discarded and replaced by sealing films for the read, and at this scale of experiment a few microliters of sample missing makes for a large impact on results.

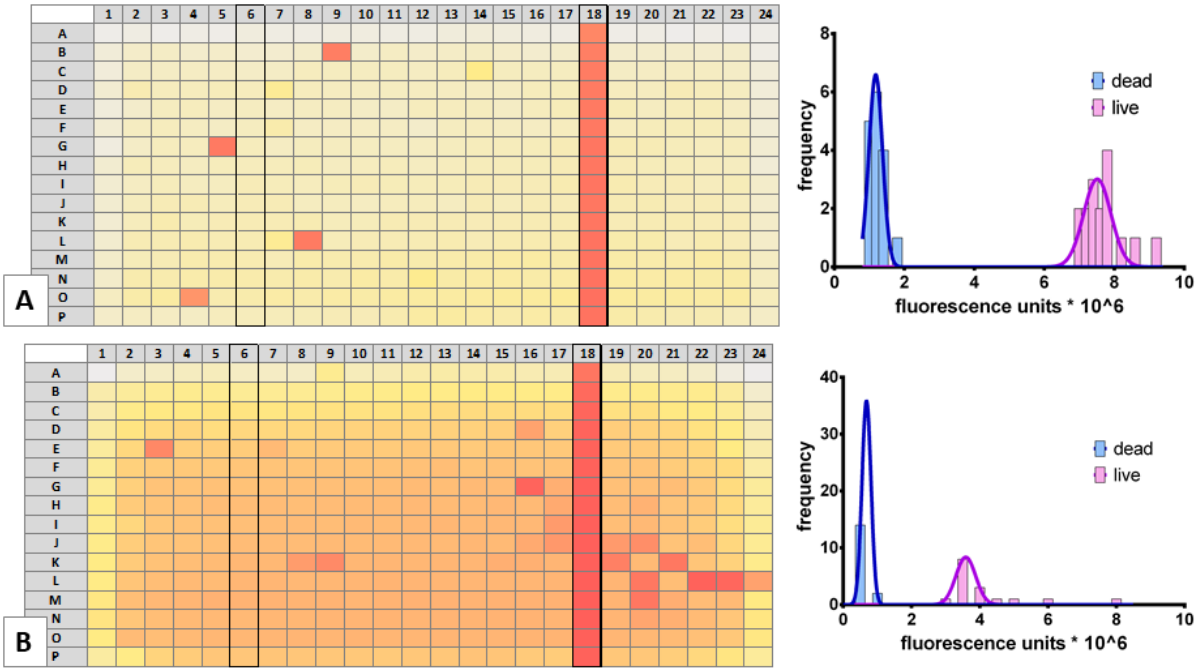


Figure 4.2. Examples for plate quality assessment by Z' values. On the left two plate layouts with fluorescence intensities as gradients (yellow high, red low) are shown. Wells towards the plate edges retained more culture sample than the central, more evaporated wells. On the right histograms of control wells of the example plate are shown. (A) Plate with Z' = 0.64 passes, (B) plate with Z' = -0.23 fails HTS quality requirements.

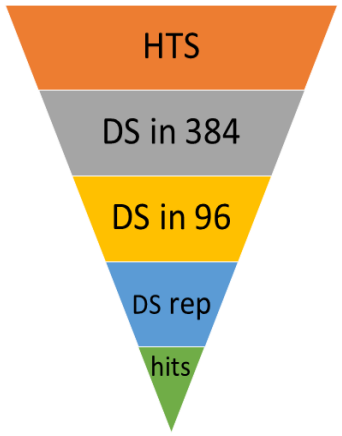
This pattern between central and peripherial wells lead to the decision of instead of using average values for dead and live controls when calculating the percentage of inhibition (and consequently the MIC shift) for each well, the single values of controls in the corresponding row were used. This gave systematic correction for the evaporation effect, but also simpler than other published method with systematic correction algorithm (Carralot et al., 2012).

As an attempt to minimise evaporation, the entire HTS was repeated with more attention to slow heating from room temperature to 37 °C, and also incubated all plates in a different thermostat. This resulted in better quality of the screen, 94 plates passed the Z'>0.4 criterium (71%). However, a suspiciously low amount of hits emerged, and a disproportionate amount of the failed plates belonged to the overexpressing strain (for example at 1 µM compound concentration against overexpressing strain 23 plates out of 33 failed quality assessment). Therefore, it is likely that the bacteria cultures were not optimal for the screen and the second

HTS was only considered as a complement to the first HTS, and finally produced merely 8 new preliminary hits to test in dose response.

4.3.2 Hit compound selection process

A compound was selected from the HTS for dose response testing in two scenarios. Either compounds that showed a shift or compounds that were too potent at the set compound concentrations. Compound causing a shift was defined as the compound that is potent against the control strain at 5 µM (at least 40% growth inhibition) and shows a shift of minimum 40% in percentage of inhibition compared to the overexpressing strain on either concentration. A compound too potent for evaluation was defined as the compound by which both ctrl and overexpressing strains were inhibited already at 1 µM (at least 80% growth inhibition).



	HTS1		HTS2	
	cmps	selection criteria	cmps	selection criteria
HTS	11 096	TB Box	11 096	TB Box
DS in 384	340	possible hit in HTS	8	additional possible hits in HTS
DS in 96	235	available in 10 mM stock	7	available in 10 mM stock
DS rep	24	shows MIC50 shift in 384-well	-	-
hits	16	MIC50 shift reproducible in 96-well	3 +2*	MIC50 shift reproducible in 96-well
	2	reproducible in both laboratories	N/A	reproducible in both laboratories

Table 4.2. Compound attrition during whole-cell screen and the influence of scale-up and location. Asterix labels two compounds that were questionable due to solubility issues. HTS2 hits were only confirmed in 96-well plate format. HTS is high-throughput screening, DS dose-response in 2-fold dilution series, DS rep means dose-response reproducibility, MIC50 minimum inhibitory concentration of growth by 50%, cmps compounds. N/A not applicable, HTS2 hits were tested in dose response in one laboratory only.

Initial dose response was performed in 384-well plates, then compounds with at least 2-fold MIC50 shift were selected from the 96-well plate (Corning, #CLS3896) repeat dose response experiment. Compounds with the largest shift were applied in enzymatic assay against Mt-DprE2.

During hit selection and confirmation, it was taken into account which compounds were reproducible in both laboratories, which resulted in surprisingly low number of hits (**Tables 4.2.** and **4.3.**).

cmp	MW	Results in GSK		Results in UoB	
		MIC50	shift	MIC50	shift
1	401.453	0.75	**4-6x**	6	no
2	351.402	1	**2-3x**	3	reverse
3	300.7	3	**4x**	9	no
4	497.548	1.3	**4x**	8	?
5	415.46	3	**large**	2.5	**4x**
6	364.249	0.375	**large**	3	reverse
7	406.482	5	2x	4.5	no
8	377.458	>16	?	13	reverse
9	552.917	10	reverse	10	no
10	257.334	2.5	**large**	unavailable	
11	262.354	>16	?	>100	?
12	403.439	12	small	12	2x
13	349.32	4.5	small	untested	
14	368.43	>16	?	>20	reverse
15	425.39	0.6	**8x**	0.6	**4x**
16	487.385	>16	?	60	reverse
17	317.7	3	**2-3x**	5-12	reverse
18	442.554	1	**large**	6-8	no
19	394.349	5	2x	9	no
20	451.416	>16	?	17	2x
21	504.96	8	small	15-24	no
22	297.18	10	?	12-17	no
23	506.62	15	reverse	15	reverse
24	379.456	1	reverse	1	reverse
25	336.429	-	-	0.75	2x
26	303.359	-	-	6	reverse
27	434.94	-	-	0.75-2	**large**
28	271.318	-	-	2	reverse
29	341.408	-	-	11	no
31	483.904	-	-	0.7	**large**
32	325.276	-	-	0.75	**4x**

Table 4.3. Refined hit selection of the HTS in various formats, on two locations. Compound 15 and 29 had different stock sources in UoB and GSK. Reverse shift means that overexpressing strain had lower MIC than control. Compounds 25-32 are from HTS2, all other from HTS1. MICs in [μM]. MW is molecular weight in [g/mol], 'small' shift means less than 2-fold, 'large' means over 8-fold shift in MIC. Question marks signal inconsistent dose response curves (possible solubility issue) and possible hit for compound 10. Yellow/green highlighting is for confirmed hits.

4.3.3 Proof of concept for Mt-DprE2 enzyme inhibition assay

As proof of concept Mt-DprE1 enzyme inhibitors were cross-referenced in Mt-DprE2 enzyme activity assay. By design, the Mt-DprE2 assay does not distinguish between Mt-DprE1 and Mt-DprE2 inhibitors, since Mt-DprE1 activity is a prerequisite to DPX formation which is the substrate of Mt-DprE2. Attempts to separate DPX from DPA/DPRs have been unsuccessful. Therefore, all Mt-DprE1 inhibitors appear as false hits with the Mt-DprE2 assay, that can be assorted by testing the false hits in Mt-DprE1 assay (Batt et al., 2016), where only Mt-DprE1 inhibitors would test positive.

TCA1 (Wang et al., 2013) and 93a3 (studied in **Chapter 2** Mt-DprE1 ligand binding studies) are highly potent Mt-DprE1 inhibitors and were selected for this experiment. TCA1 was tested with both enzyme preparations, 6xHis-SUMO-Mt-DprE2 and 6xHis-Mt-DprE1+Mt-DprE2 co-expression preparations and exhibited the anticipated constant read-out fluorescence signal, instead of a decreasing signal indicative of active enzyme reaction (**Figure 4.3.**). 93a3 also functioned similarly with the co-expression protein preparation. Compounds were incubated for 20 min with the enzymes for these results.

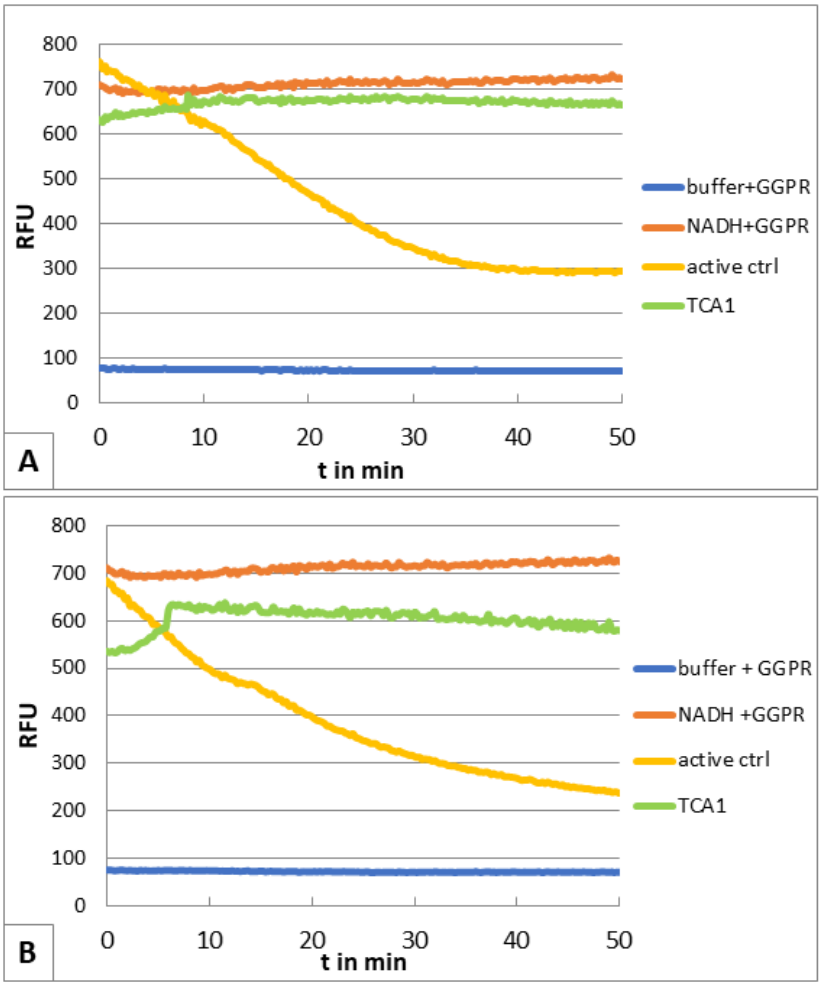


Figure 4.3. Mt-DprE2 inhibition assay proof-of-concept with TCA1 and (A) 6xHis-Mt-DprE1 + Mt-DprE2 co-expression and (B) 6xHis-SUMO-Mt-DprE2. RFU is relative fluorescence unit. Assays run in GSK.

			6xHis-Mt-DprE1 + Mt-DprE2		6xHis-SUMO-Mt-DprE2	
Reaction mixture:	buffer +GGPR	NADH +GGPR	active	TCA1	active	TCA1
BTP pH 7.2 45 mM	+	+	+	+	+	+
100 mM ammonium sulfate	+	+	+	+	+	+
180 µM NADH		+	+	+	+	+
DprE1 appr. 9.5 µM					+	+
0.2 mg/ml 6xHis-SUMO-Mt- DprE2					+	+
0.3 mg/ml 6xHis-Mt-DprE1 + Mt-DprE2			+	+		
50 µM TCA1				+		+
180 µM GGPR	+	+	+	+	+	+
MilliQ water addition for 10 µl	+	+	+	+	+	+

Table 4.4. Reaction composition of assays in Figure 4.3.

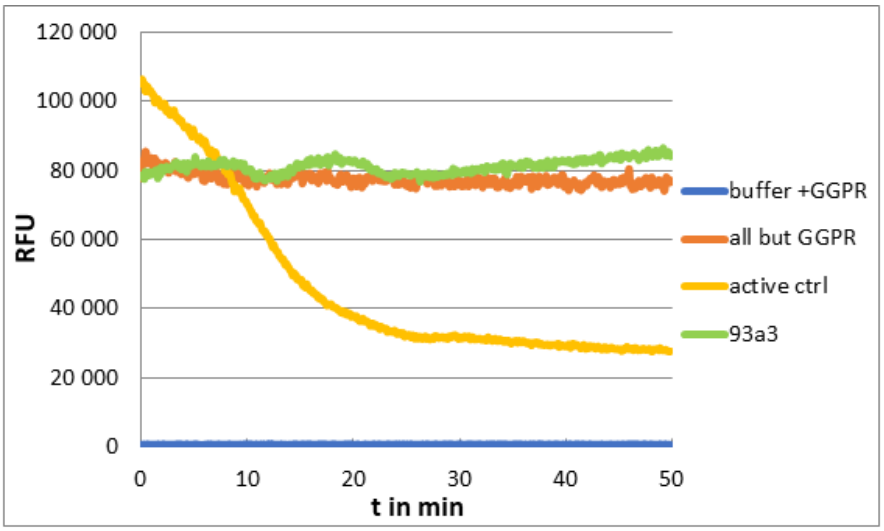


Figure 4.4. Mt-DprE2 inhibition assay proof-of-concept with 93a3 and 6xHis-Mt-DprE1 and Mt-DprE2 co-expression. RFU is relative fluorescence unit.

Reaction mixture:	buffer +GGPR	all but GGPR	active ctrl	93a3
40 mM HEPES pH 7.25	+	+	+	+
80 mM NaCl	+	+	+	+
400 µM NADPH		+	+	+
333uM 93a3				+
7 mg/ml 6xHis-DprE1 + DprE2		+	+	+
280 µM GGPR	+		+	+
MilliQ addition for 15 µl	+	+	+	+

Table 4.5. Reaction mixture for Figure 4.4.

4.3.4 DprE2 enzyme inhibition assays

Compound 15 was the first Mt-DprE2 hit compound from the published collection of 177 compounds (Ballell et al., 2013), with extensive reproducibility in whole-cell assays (**Table 4.3.**), hence evaluate in the Mt-DprE2 enzyme activity assay in the hope of developing the first Mt-DprE2 enzyme inhibitor control.

Unfortunately, compound 15 and four of its derivatives had no observable effect on enzyme activity. This result suggested that either compound 15 may be a pro-drug that is converted *in vivo* to the active form, or Mt-DprE2 is not the target of this compound series.

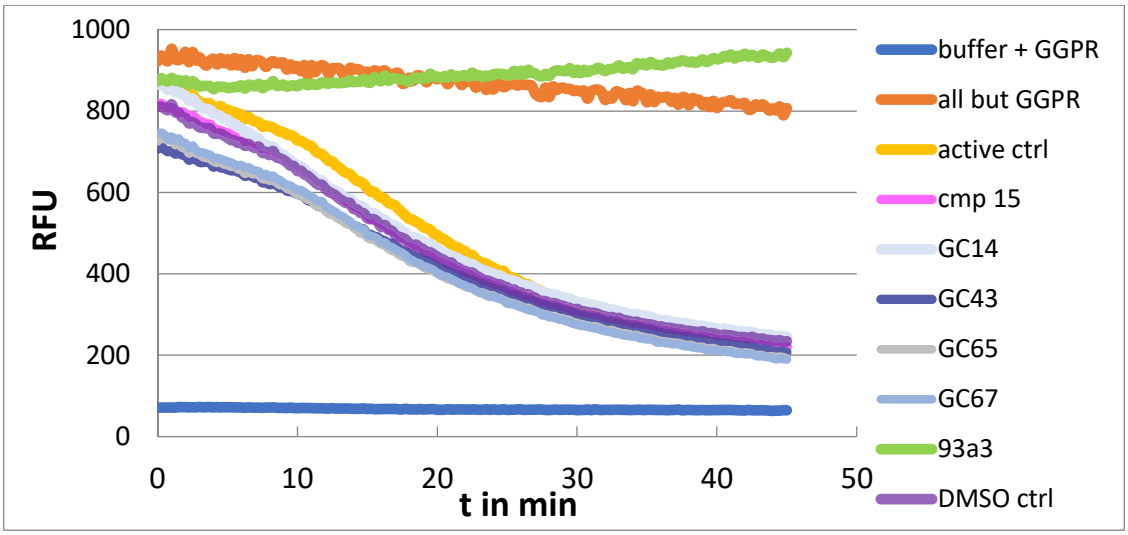


Figure 4.5. Inhibition screen of compound 15 derivatives. RFU is relative fluorescence units.

Reaction mixture:	buffer +GGPR	all but GGPR	active ctrl	cmp 15	GC14	GC43	GC65	GC67	93a3	DMSO ctrl
80 mM BTP pH 7	+	+	+	+	+	+	+	+	+	+
50 mM NaCl	+	+	+	+	+	+	+	+	+	+
600 mM ammonium sulphate		+	+	+	+	+	+	+	+	+
180 µM NADH		+	+	+	+	+	+	+	+	+
12 µM DprE1		+	+	+	+	+	+	+	+	+
0.3 mg/ml 6xHis-SUMO-DprE2		+	+	+	+	+	+	+	+	+
50 µM inhibitor or 3% DMSO				+	+	+	+	+	+	+
150 µM GGPR	+		+	+	+	+	+	+	+	+
MilliQ water addition to 10 µl	+	+	+	+	+	+	+	+	+	+

Table 4.6. Reaction mixture of assays in Figure 4.5. Assays run in GSK.

Hits form the HTS (see **Table 4.3.**) were tested in the Mt-DprE2 inhibition assay to explore if any of them directly inactivated Mt-DprE2. Since these compounds were not related in structure to compound 15, therefore it was possible that some may have been targeting Mt-

DprE2 without prior activation. Of the 18 from 34 compounds listed in **Table 4.3.** the following have been evaluated: compounds 1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 15, 17, 18, 19, 20, 27, 31, 32.

Interestingly, none of the selected compounds inhibited enzyme activity similar to the Mt-DprE1 inhibitor controls. Compound 3 oxidised NADH and NADPH without enzymes present, which gave a low starting level of fluorescence and little difference due to enzyme activity. Compound 20 gave a slightly prolonged reaction start compared to active ctrl reaction, which in coupled reactions may be an indication of low activity of the first enzyme, Mt-DprE1. This phenomenon could be repeated with lower substrate concentrations, at 240 µM GGPR as well, hence the lag of the reaction was not caused by the relatively high GGPR concentration near the maximum limit of substrate tolerance of the assay (**Figure 3.24.**).

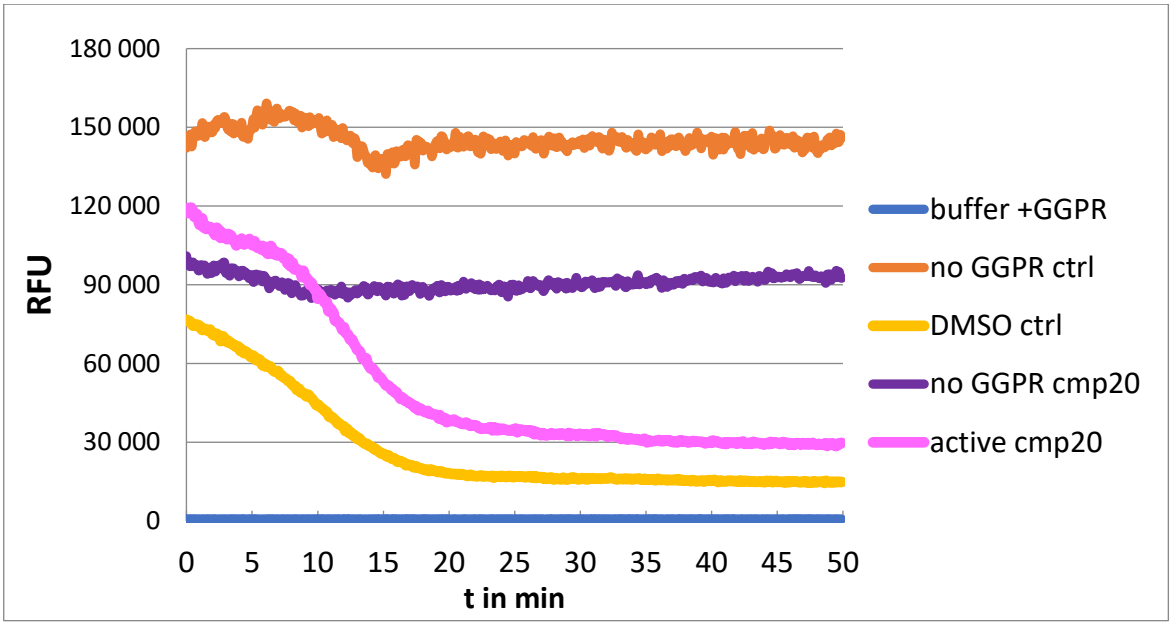


Figure 4.6. Mt-DprE2 inhibition assay with compound 20 (Table 4.3) HTS hit compound. Caused reproducible prolonged lag phase compared to DMSO control.

Reaction mixture:	buffer +GGPR	no GGPR ctrl	DMSO ctrl	no GGPR cmp20	active cmp 20
40 mM HEPES pH 7.5	+	+	+	+	+
80 mM NaCl	+	+	+	+	+
400 µM NADPH		+	+	+	+
333 µM inhibitor				+	+
3.3% DMSO		+	+		
7 mg/ml coexpressed 6xHis-DprE1+DprE2		+	+	+	+
280 µM GGPR	+		+		+
MQ water to 15 µl	+	+	+	+	+

Table 4.7. Reaction mixture of assays in Figure 4.6. Assays run in University of Birmingham.

4.3.5 Thermal shift assay (TSA) with 6xHis-SUMO-Mt-DprE2 and compound 15

In theory, TSA is an adequate method to show enzyme-ligand interaction in cases when the bound ligand stabilises the enzyme protein conformation. With Mt-DprE1 several of these interactions have been investigated (in GSK DDW, unpublished) and as a general rule of thumb ligands that increased the melting temperature (T_m) of Mt-DprE1, were inhibitors, but not all inhibitors stabilised Mt-DprE1. For example, a well known inhibitor BTZ043 did not cause a Mt-DprE1 T_m increase.

Fractions of 6xHis-SUMO-Mt-DprE2 with metal affinity chromatography followed by size exclusion chromatography purification seemed pure and concentrated enough to conduct ligand binding studies using TSA (**Figure 4.7.**). This technique requires only 0.2 mg/ml protein for detectable signal, but protein homogeneity is essential for sharp transition changes.

Firstly, as proof-of-concept NADH co-factor was tried as a ligand. NADH is not supplied during protein purification in excess, therefore it is possible that some co-factor is lost in the

process that may be reincorporated in this assay. Interestingly, NADH addition did not impact T_m significantly (**Table 4.8.**).

Compound 15 was tested in TSA with fraction and resulted the same small increase in T_m as the DMSO control (**Table 4.9.**).

	T_m in °C	SD
no ligand	60.1	0.4
+ 4 µM NADH	60.9	0.3
+ 10 µM NADH	60.7	1.3

Table 4.8. NADH as ligand of 6xHis-SUMO-Mt-DprE2 in thermal shift assay. Protein as '75.1' in **Figure 4.7.** Melting curves generated and analysed with filter 4 output data.

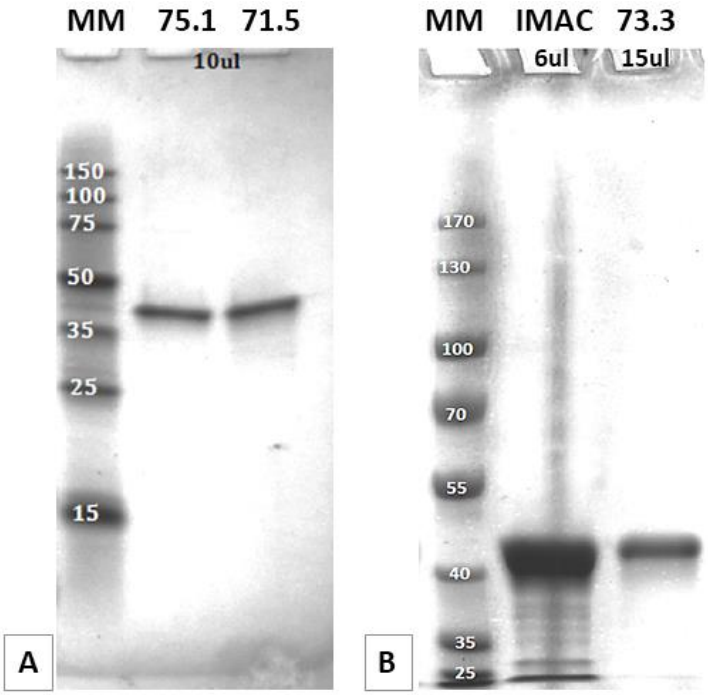


Figure 4.7. SDS-PAGE analyses of protein fractions of 6xHis-SUMO-Mt-DprE2, after size exclusion chromatography, used in thermal shift assay. (A) '75.1' fraction used in NADH binding. MM molecule marker Amresco #J383. (B) Fraction '73.3' used in compound 15 binding TSA. IMAC sample is the protein before size exclusion chromatography. MM is molecule marker is ThermoFisher, #26616. Fraction labels refer to elution volumes in [ml].

	T_m in °C	SD
no ligand	61.4	1.05
+ 0.25% DMSO	64.8	0.52
+1 µM cmp 15	65.2	0.56
+ 5 µM cmp 15	63.5	0.29
+ 10 µM cmp 15	64.5	0.92
+ 15 µM cmp 15	63.7	0.36
+ 25 µM cmp 15	62.0	1.33

Table 4.9. Compound 15 titration study with 6xHis-SUMO-Mt-DprE2 in thermal shift assay. 0.25% DMSO is as the final DMSO concentration with 25 µM compound 15 addition. Protein as '73.3' in **Figure 4.7.** Melting curves generated and analysed with filter 4 output data.

In conclusion, the thermal shift assays did not confirm that compound 15 binds DprE2, because there is no significant increase in melting temperature, but it does not exclude the possibility either, since this technique only measures the protein stability which might not be altered by the ligand.

4.3.6 Spontaneous mutant generation against hit compounds

Three compounds, the hit compound 15 and two derivatives, GC14 and GC46 (referred to as compound 20 and 31 in thesis work currently in publishing by Giacomo Chiodarelli, University of Birmingham) were tried in the modified mutant generation protocol, detailed in **Chapter 4.2.6.** Each mutagenesis delivered one clone that had a 5-fold higher MIC than wt *M. bovis* BCG (**Table 4.10.**). These clones were highly crossresistant to the other compounds with similar structures. Known structures are listen in **table 4.12.**

	solid MIC in SPOTi [µM]	clone:	**liquid MIC90 [µM]**				
			wt BCG	**Mt-DprE2 oe**	**p1C6**	**p2A3**	**p4B5**
GC46	< 1.5		2.99	9.14	16	16	N/A
GC14	< 0.24		0.56	5.22	5.22	5.22	N/A
cmp 15	< 0.6		0.56	1.71	2.99	2.99	2.99
cmp 31	< 0.6		0.97	> 28	-	-	> 28

Table 4.10. Cross-resistance shown as similar shift in MIC between generated spontaneous resistant mutant clones. Compounds labelled GC46 and GC14 are potent derivatives of compound 15 (produced by Giacomo Chiodarelli, University of Birmingham).

Resistant mutant ‘p1C6’ originates from mutagenesis against compound GC46, p2A3 against compound GC14, p4B5 against compound 31. Mt-DprE2 oe labels the strain overexpressing Mt-DprE2.

Four spontaneous mutants were subjected to whole genome sequencing (performed by Genomics Facilities at the University of Birmingham) and compared to wt *M. bovis* BCG genomic DNA sequence. **Table 4.11.** details which mutant colonies appeared against which compound during mutagenesis (see ‘highest tolerated concentration’). The results highlighted two key genes affected by mutation: *fgd1* and *fbiC*. Mutant p1C6 and p2A3 had missense mutations in *fgd1* leading to a L321R amino acid replacement, while p3A4 acquired a missense mutation leading to a S196P change. The protein encoded by *fgd1* is identical in both *M. tuberculosis* and *M. bovis* BCG on the amino acid level (P9WNE1 and A0A0H3M1G9 on UniProt (UniProt Consortium, 2018)). Mutant p4B5 had a missense mutation in *fbiC* that resulted in a R847H amino acid replacement. Again, this gene codes identical proteins in both species (P9WP77 and Q7U0G9 in *M. tuberculosis* and *M. bovis* BCG, respectively (UniProt Consortium, 2018)).

compound	comment	wt MIC	highest tolerated conc.	mutant ID	pointmutation
cmp 15	initial hit	0.14	0.31	p3A4	*fgd1* gene S196P
GC46	derivatives of cmp 15	< 1.54	4.69	p1C6	*fgd1* gene L321R
GC14		< 0.24	0.53	p2A3	*fgd1* gene L321R
cmp 31	HTS hit, cmp 15 derivative	0.7	8	p4B5	*fbiC* gene R847H

Table 4.11. Genomic sequencing results of resistant mutant *M. bovis* BCG clones. Concentration values in [μM].

In conclusion, these spontaneous mutations are suggestive of similarity to delamanid and pretomanid activation in the case of compound 15 and its derivatives, which is detailed in **Chapter 4.4.**

4.4 Discussion

The compound library ‘TB Box’ was investigated as a pool of potential Mt-DprE2 inhibitors in this HTS. Details about the library are not published, but it contains about 11,000 compounds that in previous phenotypic assays inhibited MTB growth. The library excludes drugs with known mechanisms of action, furthermore only compounds with an acceptable level of cytotoxicity were included that were detailed in the TB Box subset ‘177 TB set’ published by Ballell *et al.* (2013).

The compound concentrations of the single-shot HTS is a crucial decision in hit discovery. Earlier published (Cox et al., 2016) and unpublished screens explored in similar set-ups against other targets the concentration range between 0.5 to 20 μM. At the selected 1 and 5 μM single-shot HTS in HTS1 986 compounds were potent (inhibited growth in ctrl strain by 40% or more), in HTS2 580 compounds (which is 11.3% and 5.2% of all compounds, respectively). This implies that the TB Box has a large pool of compounds that were beyond the reach of this screen and may be of interest to test at higher concentrations in the future.

It is worth mentioning, that there are alternatives to single concentration testing. If there were information regarding MICs in whole-cell assays with similar assay conditions on each compound of the library, then it would be more informative to test each compound around its MIC or slightly above that value to observe elevated MICs.

One of the most important findings of the HTS and following hit confirmation experiments was the unexpectedly high compound attrition rate (**Figure 4.2.**). In the case of preliminary hits not confirmed in dose response assay, it was an expected consequence of low initial screen quality. In the first screen (HTS1) 29% of the 384-well plates failed quality assessment by Z’ values (**Chapter 4.3.1.**), which could easily translate to several initial hits

resulting to be false positives. A more complex task was to interpret how the majority of 384-well dose response screen results became irreproducible in 96-well plates.

There are several differences between the two dose response assay formats. The 384-well plate format had 50 µl of bacteria culture grown and evaluated compared to the 100 µl reaction volume of the 96-well format. A larger bacterial culture may give more robust response to compound exposure, the survival of the cells is based on a larger statistical sample leading to a better average than with small cell numbers. Also, plate well geometry is different: the 384-well plates' chimneys by Greiner are rectangular cuboids, while 96-well plates by Corning are cylindrical with a V-bottom (**Figure 4.8.**). Bacterial cells accumulate at the bottom of the well at rest, so from the aspect of oxygen diffusion, it is important how tall the liquid layer is above the cells. In the 384-well plate chimney well it is approximately 4.6 mm, while in the V-bottom 96-well plate chimney approximately 3.1 mm (calculated with well parameters of the manufacturer's brochure and reaction volumes). The relative surface area in contact with the culture is bigger in the case if 384-well plate format. MTB is known to be capable to form biofilms and that this feature confers higher drug tolerance to the cells (Sambandan et al., 2013). The detergent Tween-80 is included in the media precisely to prevent biofilm formation and cell clumping. Taken these information together it is likely, that the small size inoculum and the geometry more supportive of cell adherence, and hence deviance from average drug susceptibility for cell subpopulations adhered to the surface, the 384-well plate format is less robust than the 96-well plate format, more responsive to variation in controlled and uncontrolled factors.

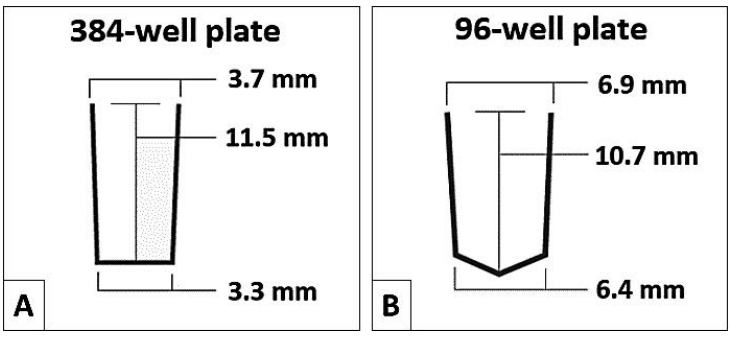


Figure 4.8. Well chimney geometry of (A) the Greiner 384-well plate used in HTS, (B) the Corning 96-well plate used in dose response measurements. Well parameters from the manufacturers' brochures (Gbo.com, n.d.)(Corning, n.d.).

Difficulties reproducing results between laboratories proved to be an issue as well with hit compounds. Out of 16 hit compounds investigated in GSK DDW only 2 were reproducible hits in University of Birmingham laboratories (**Tables 4.2.** and **4.3.**). Experiments were conducted, with two exceptions (see **Table 4.3.**), from the same compound stock – two different aliqouts from the exact same compound preparation in GSK Harlow compound library. Compounds 1 to 8 were deposited in the same 96-well plate, in the same column during the GSK DDW experiments, but the cross-contamination of compounds in-between wells is unlikely for two reasons: if the only shift-causing compound among these, compound 5 was to spill to the surroundings, it makes no sense that it would affect only some neighbouring cells, but not others. Furthermore, the MIC50 values were not representative to compound 5, in GSK the compound 4 and 6 were more potent than compound 5 (**Table 4.3.**). Therefore, it is more likely that the differences in MIC values between laboratories was somehow the consequence of using different *M. bovis* BCG stocks. Only the plasmids were transferred between laboratories, locally available wild type *M. bovis* BCG bacteria were transformed with pMV261-Mt-DprE2 for overexpressing strain, and with pMV261 empty plasmid for control strain. These theoretically identical strains responded similarly to compound 15, the main hit compound, however with seven other hits

(compound 1, 2, 3, 4, 6, 17, 18 see **Table 4.3.**) the MIC values were equally high for the control strain, as it was only with the overexpressing strain in GSK.

One of the most useful and pursuable outcome of the Mt-DprE2 screen were the few hit compounds, that proved to be reproducibly elicit a shift in MIC depending on Mt-DprE2 overexpression in *M. bovis* BCG. These compounds are compound 5, 15, 27, 31, 32 and potentially compound 10 (which was not tested in University of Birmingham). Compound 10 and 27 are unpublished structures of GSK, but the structures of compound 15, 31 and 32 are not only published, but share structural similarities: they are all nitrofurans (**Table 4.12.**).

Nitrofurans have been reported previously as potent inhibitors of *M. tuberculosis* (Tangallapally et al., 2007). Some derivatives demonstrated potency in the nanomolar range (Yempalla et al., 2015). The main challenge of compound optimisation was then to keep the compound metabolically stable.

Also, erroneously the target of nitrofurans was firstly identified as Glf (Scherman et al., 2003; Tangallapally et al., 2007), based on enzyme activity screens where whole cell screens indicated it as the target, but enzyme activity assays did not. Glf contributes to the arabinogalactan synthesis, as does Mt-DprE2, but is involved in the synthesis of the UDP-Gal*f* residue, the building block of the galactan homopolymer (Weston et al., 1997). To date no mechanism of action have been published to these compounds.

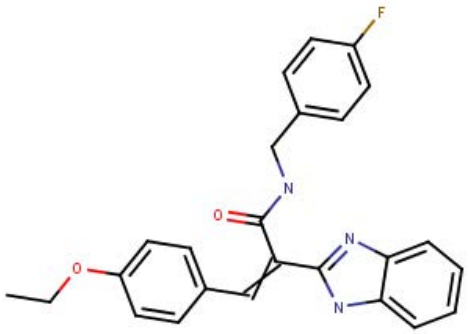
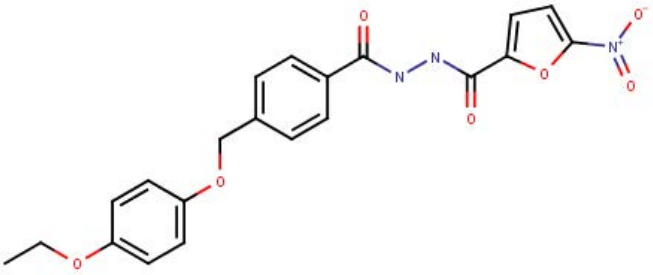
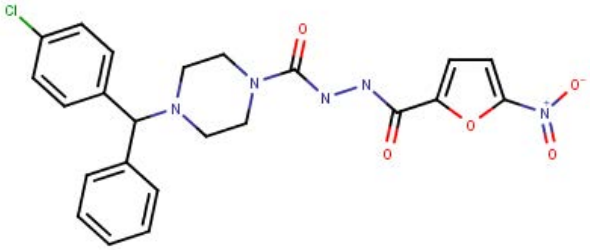
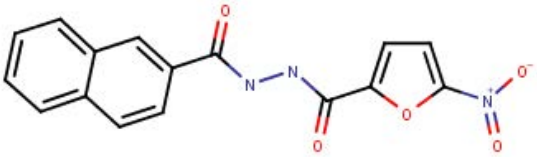
GSK1301131A (cmp 5) 415.46 g/mol HTS hit compound	
GSK426032A (cmp 15) 425.397 g/mol HTS hit compound	
GW432086X (cmp 31) 483.909 g/mol HTS hit compound	
GSK347301A (cmp 32) 325.279 g/mol HTS hit compound	

Table 4.12. Published compound structures with antimycobacterial activity, hits in the Mt-DprE2 screen against the TB Box compound library. Available from PubChem (Kim et al., 2016).

The spontaneous mutant generation studies pointed to two genes that lead to resistance against Mt-DprE2 screen hit compounds: genes *fgd1* and *fbiC*. These two genes have been previously reported to contribute to pretomanid (PA-824) (Haver et al., 2015) and delamanid resistance (Fujiwara et al., 2018), the latter is a WHO approved drug for treating MDR-TB since 2014 (WHO, 2017). These compounds are similar (**Table 4.13.**), both are nitroimidazoles and pro-drugs that require activation (Islam et al., 2017). Also, both

inhibitors have a complex mechanism of action. In the case of pretomanid, a combination of respiratory effects changing the redox status of the cell, releasing nitric oxide and hindering mycolic acid synthesis (Manjunatha et al., 2009).

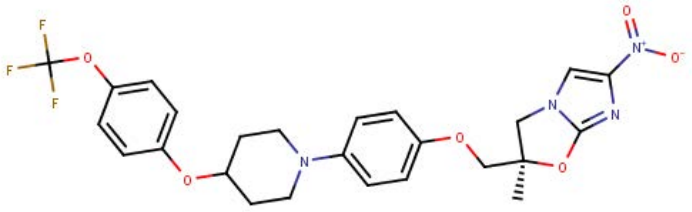
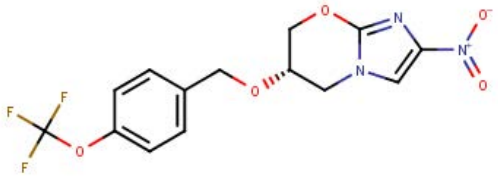
Delamanid	
MW = 534.492 g/mol WHO approved drug for MDR-TB therapy	
Pretomanid	
MW = 359.261 g/mol in phase 3 clinical studies	

Table 4.13. Structures of delamanid and pretomanid, drugs in clinical testing.

Why mutations within these two genes are often encountered against pretomanid and delamanid, is explained by the bioactivation of these compounds. The activator of delamanid is the protein deazaflavin-dependent nitroreductase (Ddn, Rv3547), however this enzyme functions only in the presence of the reduced F_{420} co-enzyme. The genes *fbiA/B/C* synthesise F_{420} while *fgd1* codes a glucose-6-phosphate dehydrogenase which utilises the oxidised co-factor and completes the catalytical cycle of F_{420} (Fujiwara et al., 2018) (**Figure 4.9.**).

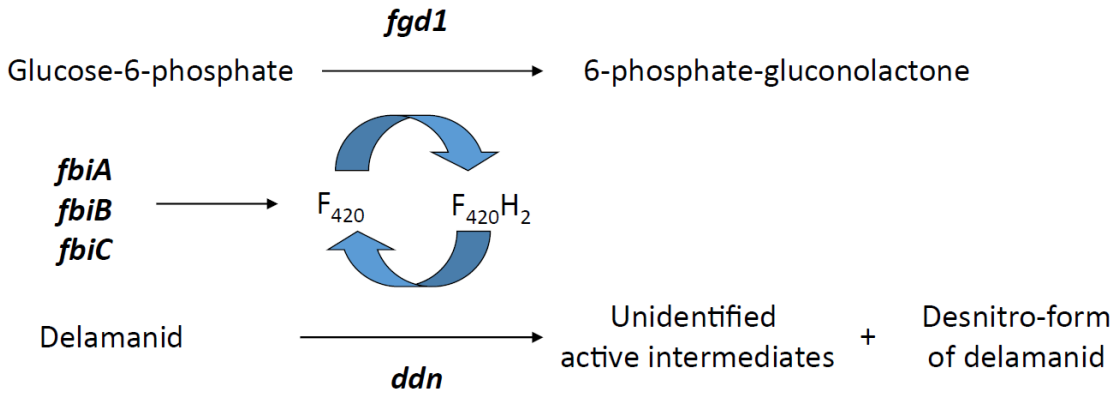


Figure 4.9. *In vivo* activation of delamanid by (Fujiwara et al., 2018)**.**

Based on structure similarity and spontaneous mutations in these specific genes together suggest that compounds 15, 31 and 32 may be similar in bioactivation and mechanism of action to pretomanid and delamanid. That would explain the failure of Mt-DprE2 enzyme activity assay proving compound 15 inhibiting Mt-DprE2 (**Chapter 4.3.4.**). Firstly, if it is a pro-drug, then the activator is not included in the assay setup, unless it is Mt-DprE1. Secondly, Mt-DprE2 may not be a unique and/or specific target of these compounds. The latter could be elucidated in the future with *in vitro* bioactivation of compound 15 with Ddn protein and reduced F_{420} and subsequently preforming the Mt-DprE2 enzymatic assay.

A further implication of similarities between hit compounds and other nitroimidazoles is that if their mechanism of action is similar, the hit compounds may also be highly active against dormant *M. tuberculosis*. Delamanid is a sterilising drug (Fujiwara et al., 2018). Pretomanid has been confirmed to inhibit *M. tuberculosis* growth by NO release, a phenomenon that is insufficient to kill aerobically grown bacteria, but gains importance in anaerobic conditions (Manjunatha et al., 2009). In the future it is worth looking into the potential sterilising qualities of the hit compounds (**Table 4.12. and 4.3.**) because it would influence the possible drug combinations in therapy and even their application in latent tuberculosis treatment and prophylaxis.

Notwithstanding the distance between species, these compounds might be active even against parasites. Nitroimidazoles have been used against infectious agents unrelated to *M. tuberculosis*, in particular against trichomoniasis, leishmaniasis and Chagas-disease (Ang et al., 2017). These diseases are caused by kinetoplastid protozoa, eukaryotic species (Ang et al., 2017). If the Mt-DprE2 screen hit compounds could be repurposed for the treatment of these neglected diseases, it could be an interesting new area of investigation.

5 Conclusions

Experiments conducted in this thesis were all aimed at investigating the decaprenyl-phosphorylribose epimerase subunits Mt-DprE1 and Mt-DprE2 proteins, that are both promising targets for developing new inhibitors against mycobacteria. Inhibitors directed against novel targets can limit proliferation of drug-resistant strains of *M. tuberculosis*, which is a prime ambition of new drug discovery research currently.

Mt-DprE1 is an established and thoroughly studied drug target of mycobacterial cell wall synthesis. Mt-DprE1 has emerged as a target from several independent phenotypic screening campaigns, suggesting it is highly druggable, while the availability of various protein structures *via* crystallography make docking studies and structure activity relationship optimisation possible. The proximity of two disordered regions to the active site and frequent ligand binding site raised the question if these disordered regions participated in ligand binding.

Results of ligand binding experiments with intrinsic tryptophan fluorescence measurement revealed that with two investigated ligands the affinity for binding did not change significantly by introducing pointmutations to the short disordered region. Therefore, from the drug developmental aspect, the lack of this disordered region in crystal structures causes a level of uncertainty to how ligands interact with the protein in solution, but in some cases (such as with these tested ligands) do not interfere with ligand affinity.

Based on the enzyme activity assay studies with Mt-DprE1 muteins, it became clear that some amino acid residues in the short disordered region are vital for normal enzyme function. The enzyme follows an allosteric kinetics which is uncommon for monomeric

enzymes, together with the pattern in residue essentiality in the short loop corroborate the theory that the enzyme likely follows a lid-like enzyme dynamic (Porter and Miller, 2012). Mutations of the lid 'hinges', residues at both ends of the loop affect enzyme activity the most, and these are also evolutionally highly conserved amino acids.

A notable observation of the enzyme activity assay with Mt-DprE1 was the direct comparison of the published assay with substrate FPR, and the not yet published assay of GSK with GGPR. Using the latter substrate analogue led to a much more robust and better signal-to-noise assay design which is beneficial to inhibitor testing at any scale. However, at higher concentrations GGPR appeared to inhibit enzyme activity, which makes enzyme characterisation with this substrate impossible beyond a certain concentration limit.

With the generated Mt-DprE1 single site muteins in the future some interesting experiments can continue. The small disordered region may or may not interact with Mt-DprE2. Native Mt-DprE2 may be produced for example from 6xHis-SUMO-Mt-DprE2 expression and purification, followed by tag cleavage. Furthermore, the inactive muteins of Mt-DprE1 could also participate in crystallisation trials, since the substrate binding interaction with the enzyme has never been solved.

The second main goal of this work was to produce pure, non-aggregated but concentrated Mt-DprE2. There is still no experimental protein structure solved for this potential drug target and there are no biochemical studies published to date due to the difficulty of obtaining non-aggregated protein. There are only whole cell, genetic or *in silico* modelling studies with it, because its purification has proven to be difficult. Crystallisation failed likely because the aggregation impedes protein crystal formation. It is undetermined, why Mt-DprE2 self-associates so readily, perhaps it functions natively also as an oligomeric enzyme. The most similar protein structures listed in an *in silico* modelling study (Bhutani et al., 2015)

paralleled Mt-DprE2 to proteins that all form homo-tetramers. Multiple native protein-protein interaction sites present in a highly concentrated protein solution may produce large aggregates, especially in the artificial environmental to which the protein is exposed to during purification, including large changes in pressure, the presence of a filtration membrane and shear by the solvent during ultrafiltration, which could further propagate aggregation (Kim et al., 1993). Nevertheless, it has been hypothesised that periplasmic proteins in *E. coli* are more resistant to aggregation, than other proteins of the cell (Liu et al., 2004), which by similarity between species either adds to the unique properties of Mt-DprE2, or adds to the question whether it is located in the periplasm as reported (Brecik et al., 2015).

In the third chapter various conditions have been explored to improve protein solubility and homogeneity in solution. Firstly, tag-fusion with 6xHis-SUMO improved enzyme quantities and led to enzymatically active, adequately pure preparations. However, protein degradation is likely taking place in fresh preparations already, and the fusion protein kept its tendency towards aggregation, as seen with size exclusion chromatography. Therefore, without solving the issue of protein concentration causing further aggregation by filtratioin, the method is not effective to produce protein preparation for crystallography. As for enzyme activity assay, the fusion protein did not show activity in combination with the redox co-factor NADPH, the most probable co-factor of the enzyme, therefore its usefulness in inhibitor screening is ambiguous.

Co-expression of Mt-DprE2 with His-tagged Mt-DprE1 lead to useful amounts (near 1 mg/ml stock solutions) of the two proteins for enzyme activity measurements, especially when expressed from different plasmids. The preparation was enzymatically active with both NADH and NADPH, and activity was reversible by Mt-DprE1 inhibitors. The main concern regarding the assay that needs further assessment is the difficulty to standardise the enzyme

activity across batches. Each enzyme preparation had different activity profile which did not strictly correlate with total protein concentration. One clue to this behaviour is that when co-expressing Mt-DprE2 with Mt-DprE1, the molar ratio was often distinctly less than one. Also, depending on the expression design, expressing the proteins from one single or two separate plasmids led to different types and amounts of characteristic contaminants co-purified with the mixture. Therefore, in the future different purification techniques could be included in the protein production to achieve consistent ratio and eliminate the remaining contaminants. One main contaminating protein stood out, which has to be separated to measure precise concentrations. Since this contaminant was very similar in size on SDS-PAGE, and ion exchange chromatography required unusually high salt concentrations for elution, it is sensible to replace the second chromatography step to hydrophobic interaction chromatography in the future.

Size exclusion chromatography studies revealed, that Mt-DprE1 and Mt-DprE2 in solution form a complex. Interestingly, unidentified proteins were also present in the complex fraction of the experiment, which when reproduced with higher enzyme quantity, may be informative to assess in mass spectrometry. These complexes, if separated in higher concentration, could also serve as sample in crystallography. High salt concentrations helped both enzymes to stabilise in solution, which already led to improved crystallisation with difficult *M. tuberculosis* proteins in the past (Cohen-Gonsaud et al., 2002).

High-throughput screening with Mt-DprE2 overexpressing *M. bovis* BCG aimed at identifying Mt-DprE2-specific inhibitor compounds in the TB Box compound library. The whole cell assay in 50 µl final volume and resazurin read-out lead to an unexpectedly low signal to noise ratio and a systematic error of edge effect. For this reason, sample location on a microplate greatly influenced the signal-to-noise ratio of the reaction and hit

identification of the compound. This issue was most reliably solved by doubling the assay volume to 100 µl, which in secondary screens delivered first approximately 16 hit compounds, followed by 5 more after the repeat of the total high-throughput screen. Interestingly, only seven compounds were confirmed in multi-laboratory testing (tested in both University of Birmingham and GSK DDW Tres Cantos), and none of the tested hit compounds inhibited the developed in Mt-DprE2 enzyme activity assay. Which also proves that they are not Mt-DprE1 inhibitors, since the enzyme activity assay is coupled to and dependent on Mt-DprE1 activity.

One of the biggest successes of the Mt-DprE2 inhibitor screen was the generation of spontaneous resistant mutants against some potent inhibitors. As the negative result in enzyme activity assay suggested, none of the resistance-conferring mutations arose in *dprE2*. Instead, resistance was linked to point mutations in two enzymes, *fbiC* and *fgd1*, involved in co-enzyme F_{420} synthesis and recycling to its reduced form, respectively. These mutations are reminiscent of mutations that provided resistance against delamanid and pretomanid, nitroimidazole antituberculotic drugs in development and approved for drug resistant tuberculosis therapy. Spontaneous mutations against these drugs focus on the activator enzyme, Ddn, and enzymes producing the reduced co-factor of Ddn, the deazaflavin F_{420}, in particular the enzymes FbiA, FbiB, FbiC and Fgd1 (Fujiwara et al., 2018).

The screening revealed one specific chemical scaffold: 2-nitrofuran compounds were confirmed hits in several different derivatives. These compounds may be prodrugs similar to the nitroimidazoles, whose activation has been studied in detail. Pretomanid mechanism of action is complex, in aerobic environment it inhibits the mycolic acid synthesis (Stover et al., 2000), while anaerobically cell death is attributed to NO production (Manjunatha et al., 2009), which hinders the function of multiple native enzymes. Mt-DpE2 may be one of

several targets impaired by reactive nitrogen species formed during the reduction of the hit compound.

Target identification of the newly found hits is of paramount importance. Since **Chapter 4** elaborated on a phenotypic compound screen, a change in susceptibility due to overexpressing Mt-DprE2 may be only a minor contributor to compound resistance and may play a small part in the mechanism of action. To determine, if Mt-DprE2 is a direct target of an activated compound, in the future a fully functioning enzyme inhibition assay needed to be conducted. Assay reaction should include the activator enzyme, which could be Ddn-F_{420} or a different enzyme, which has to be identified and then added to the assay reaction. Pull-down methods may be beneficial as well in the search. Lastly, the identified activator either has to be expressed and purified in a system that uses F_{420}, like *M. smegmatis*, or the co-factor has to be added to the sample throughout purification and enzyme reaction.

Involving the "omics" technologies in future experiments may help in narrowing down the group of potential targets of the hits. Transcriptome profiling for pretomanid elucidated how anaerobic environments propagated cell death by a different mechanism of action, through respiratory poisoning (Manjunatha et al., 2009). Metabolome profiling of pretomanid was informative regarding the diversified mechanism of action that did not resemble any compared clinical drugs (Baptista et al., 2018), which is advantageous in combinational drug therapy inhibiting multiple metabolic pathways to eliminate drug-resistant strains of *M. tuberculosis*.

In conclusion, the high-throughput screen in search of new inhibitors of Mt-DprE2 resulted in compounds with a more complex mechanism of action than anticipated. Although Mt-DprE2 may be only one of multiple targets of these new compounds, but this implies once more the vulnerability of the decaprenyl-phosphorylarabinose biosynthetic pathway of *M.*

tuberculosis, and encourages the study of other identified, essential enzymes involved in mycobacterial cell wall biosynthesis as drug targets.

6 References

Abrahams, K.A. and Besra, G.S. (2018) Mycobacterial cell wall biosynthesis: A multifaceted antibiotic target. *Parasitology*. doi:10.1017/S0031182016002377.

Acker, M.G. and Auld, D.S. (2014) Considerations for the design and reporting of enzyme assays in high-throughput screening applications. *Perspectives in Science*. doi:10.1016/j.pisc.2013.12.001.

Adams, K.N., Takaki, K., Connolly, L.E., et al. (2011) Drug tolerance in replicating mycobacteria mediated by a macrophage-induced efflux mechanism. *Cell*. doi:10.1016/j.cell.2011.02.022.

Agilent Technologies (2011) QuikChange II Site-Directed Mutagenesis Kit: Instruction Manual. *Europe*. doi:PMID:12449398.

Alderwick, L.J., Harrison, J., Lloyd, G.S., et al. (2015) The mycobacterial cell wall—peptidoglycan and arabinogalactan. *Cold Spring Harbor Perspectives in Medicine*. doi:10.1101/cshperspect.a021113.

Alderwick, L.J., Radmacher, E., Seidel, M., et al. (2005) Deletion of Cg-emb in corynebacterianeae leads to a novel truncated cell wall arabinogalactan, whereas inactivation of Cg-ubiA results in an Arabinan-deficient mutant with a cell wall galactan core. *Journal of Biological Chemistry*. doi:10.1074/jbc.M506339200.

Alderwick, L.J., Seidel, M., Sahm, H., et al. (2006) Identification of a novel arabinofuranosyltransferase (AftA) involved in cell wall arabinan biosynthesis in Mycobacterium tuberculosis. *Journal of Biological Chemistry*. doi:10.1074/jbc.M600045200.

Alnimr, A.M. (2015) Dormancy models for mycobacterium tuberculosis: A minireview. *Brazilian Journal of Microbiology*. doi:10.1590/S1517-838246320140507.

Andersen, P. and Doherty, T.M. (2005) The success and failure of BCG - Implications for a novel tuberculosis vaccine. *Nature Reviews Microbiology*. doi:10.1038/nrmicro1211.

Ang, C.W., Jarrad, A.M., Cooper, M.A., et al. (2017) Nitroimidazoles: Molecular Fireworks That Combat a Broad Spectrum of Infectious Diseases. *Journal of Medicinal Chemistry*. doi:10.1021/acs.jmedchem.7b00143.

Angala, S.K., Belardinelli, J.M., Huc-Claustre, E., et al. (2014) The cell envelope glycoconjugates of Mycobacterium tuberculosis. *Critical Reviews in Biochemistry and Molecular Biology*. doi:10.3109/10409238.2014.925420.

Arbing, M.A., Chan, S., Harris, L., et al. (2013) Heterologous expression of mycobacterial Esx complexes in Escherichia coli for structural studies is facilitated by the use of maltose binding protein fusions. *PLoS ONE*. doi:10.1371/journal.pone.0081753.

Baer, C.E., Rubin, E.J. and Sassetti, C.M. (2015) New insights into TB physiology suggest untapped therapeutic opportunities. *Immunological Reviews*. doi:10.1111/imr.12267.

Ballell, L., Bates, R.H., Young, R.J., et al. (2013) Fueling Open-Source Drug Discovery: 177 Small-Molecule Leads against Tuberculosis. *ChemMedChem*. doi:10.1002/cmdc.201200428.

Baptista, R., Fazakerley, D.M., Beckmann, M., et al. (2018) Untargeted metabolomics reveals a new mode of action of pretomanid (PA-824). *Scientific Reports*. doi:10.1038/s41598-018-23110-1.

Batt, S.M., Cacho Izquierdo, M., Castro Pichel, J., et al. (2016) Whole Cell Target Engagement Identifies Novel Inhibitors of Mycobacterium tuberculosis Decaprenylphosphoryl-β- d -ribose Oxidase. *ACS Infectious Diseases*. doi:10.1021/acsinfecdis.5b00065.

Batt, S.M., Jabeen, T., Bhowruth, V., et al. (2012) Structural basis of inhibition of Mycobacterium tuberculosis DprE1 by benzothiazinone inhibitors. *Proceedings of the National Academy of Sciences*. doi:10.1073/pnas.1205735109.

Berg, S., Kaur, D., Jackson, M., et al. (2007) The glycosyltransferases of Mycobacterium tuberculosis - Roles in the synthesis of arabinogalactan, lipoarabinomannan, and other glycoconjugates. *Glycobiology*. doi:10.1093/glycob/cwm010.

Bhutani, I., Loharch, S., Gupta, P., et al. (2015) Structure, dynamics, and interaction of Mycobacterium tuberculosis (Mtb) DprE1 and DprE2 examined by molecular modeling, simulation, and electrostatic studies. *PLoS ONE*. doi:10.1371/journal.pone.0119771.

Blanco-Ruano, D., Roberts, D.M., Gonzalez-Del-Rio, R., et al. (2015) Antimicrobial susceptibility testing for Mycobacterium sp. *Methods in Molecular Biology*. doi:10.1007/978-1-4939-2450-9_15.

Bowes, J., Brown, A.J., Hamon, J., et al. (2012) Reducing safety-related drug attrition: The use of in vitro pharmacological profiling. *Nature Reviews Drug Discovery*. doi:10.1038/nrd3845.

Brecik, M., Centárová, I., Mukherjee, R., et al. (2015) DprE1 Is a Vulnerable Tuberculosis Drug Target Due to Its Cell Wall Localization. *ACS Chemical Biology*. doi:10.1021/acschembio.5b00237.

Brennan, P.J. and Nikaido, H. (1995) The Envelope of Mycobacteria. *Annual Review of Biochemistry*. doi:10.1146/annurev.bi.64.070195.000333.

Brown, L., Wolf, J.M., Prados-Rosales, R., et al. (2015) Through the wall: Extracellular vesicles in Gram-positive bacteria, mycobacteria and fungi. *Nature Reviews Microbiology*. doi:10.1038/nrmicro3480.

Callis, P.R. (2014) Binding phenomena and fluorescence quenching. II: Photophysics of aromatic residues and dependence of fluorescence spectra on protein conformation. *Journal of Molecular Structure*. doi:10.1016/j.molstruc.2014.04.051.

Calmette, A. (1931) Preventive Vaccination against Tuberculosis with BCG. *Journal of the Royal Society of Medicine*. doi:10.1177/003591573102401109.

Cambier, C.J., Falkow, S. and Ramakrishnan, L. (2014) Host evasion and exploitation schemes of Mycobacterium tuberculosis. *Cell*. doi:10.1016/j.cell.2014.11.024.

Carralot, J.-P., Ogier, A., Boese, A., et al. (2012) A novel specific edge effect correction method for RNA interference screenings. *Bioinformatics*. doi:10.1093/bioinformatics/btr648.

Chan E.D. and Iseman M.D. (2008) Multidrug-resistant and extensively drug-resistant tuberculosis: A review. *Current Opinion in Infectious Diseases*.

Chen, Y.C., Chen, L.A., Chen, S.J., et al. (2004) A modified osmotic shock for periplasmic release of a recombinant creatinase from Escherichia coli. *Biochemical Engineering Journal*. doi:10.1016/j.bej.2004.03.001.

Chetty, S., Ramesh, M., Singh-Pillay, A., et al. (2017) Recent advancements in the development of anti-tuberculosis drugs. *Bioorganic and Medicinal Chemistry Letters*. doi:10.1016/j.bmcl.2016.11.084.

Christophe, T., Jackson, M., Hee, K.J., et al. (2009) High content screening identifies decaprenyl-phosphoribose 2′ epimerase as a target for intracellular antimycobacterial inhibitors. *PLoS Pathogens*. doi:10.1371/journal.ppat.1000645.

Clinical pipeline (2018).

Cohen-Gonsaud, M., Ducasse, S., Hoh, F., et al. (2002) Crystal structure of MabA from Mycobacterium tuberculosis, a reductase involved in long-chain fatty acid biosynthesis. *Journal of Molecular Biology*. doi:10.1016/S0022-2836(02)00463-1.

Cole, S.T. (2016) Inhibiting Mycobacterium tuberculosis within and without. *Philosophical Transactions of the Royal Society B: Biological Sciences*. doi:10.1098/rstb.2015.0506.

Cole, S.T., Brosch, R., Parkhill, J., et al. (1998) Deciphering the biology of mycobacterium tuberculosis from the complete genome sequence. *Nature*. doi:10.1038/31159.

Comas, I., Coscolla, M., Luo, T., et al. (2013) Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans. *Nature Genetics*. doi:10.1038/ng.2744.

Corning (n.d.) *Microplates*.

Cox, J.A.G., Mugumbate, G., Del Peral, L.V.G., et al. (2016) Novel inhibitors of Mycobacterium tuberculosis GuaB2 identified by a target based high-throughput phenotypic screen. *Scientific Reports*. doi:10.1038/srep38986.

Dartois, V. (2014) The path of anti-tuberculosis drugs: From blood to lesions to mycobacterial cells. *Nature Reviews Microbiology*. doi:10.1038/nrmicro3200.

Davis, A.L. (2000) "A Historical Perspective on Tuberculosis and Its Control." In Reichman,

L.B. and Hershfield, E.S. (eds.) *Tuberculosis : A Comprehensive International Approach*.

DeLano, W.L. (2002) The PyMOL Molecular Graphics System. *Schrödinger LLC wwwpymolorg*. doi:citeulike-article-id:240061.

Ehrt, S. and Rhee, K. (2013) Mycobacterium tuberculosis metabolism and host interaction: Mysteries and paradoxes. *Current Topics in Microbiology and Immunology*. doi:10.1007/82-2012-299.

Ekins, S., Kaneko, T., Lipinski, C.A., et al. (2010) Analysis and hit filtering of a very large library of compounds screened against Mycobacterium tuberculosis. *Mol. BioSyst.* doi:10.1039/C0MB00104J.

Epps, D.E., Raub, T.J., Caiolfa, V., et al. (1999) Determination of the affinity of drugs toward serum albumin by measurement of the quenching of the intrinsic tryptophan fluorescence of the protein. *The Journal of pharmacy and pharmacology*. doi:10.1211/0022357991772079.

Escuyer, V.E., Lety, M.A., Torrelles, J.B., et al. (2001) The Role of the embA and embB Gene Products in the Biosynthesis of the Terminal Hexaarabinofuranosyl Motif of Mycobacterium smegmatis Arabinogalactan. *Journal of Biological Chemistry*. doi:10.1074/jbc.M102272200.

Ewing, T.A., Fraaije, M.W., Mattevi, A., et al. (2017) The VAO/PCMH flavoprotein family. *Archives of Biochemistry and Biophysics*. doi:10.1016/j.abb.2017.06.022.

Fairchild, A.L. and Oppenheimer, G.M. (1998) Public health nihilism vs pragmatism: History, politics, and the control of tuberculosis. *American Journal of Public Health*. doi:10.2105/AJPH.88.7.1105.

Fleming, A. (1929) On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B.influenzae. *British journal of experimental pathology*. doi:10.1038/146837a0.

Fox, W., Ellard, G.A. and Mitchison, D.A. (1999) Studies on the treatment of tuberculosis undertaken by the British Medical Research Council Tuberculosis Units, 1946-1986, with relevant subsequent publications. *International Journal of Tuberculosis and Lung Disease*.

Franzblau, S.G., Degroote, M.A., Cho, S.H., et al. (2012) Comprehensive analysis of methods used for the evaluation of compounds against Mycobacterium tuberculosis. *Tuberculosis*. doi:10.1016/j.tube.2012.07.003.

Fujiwara, M., Kawasaki, M., Hariguchi, N., et al. (2018) Mechanisms of resistance to delamanid, a drug for Mycobacterium tuberculosis. *Tuberculosis*. doi:10.1016/j.tube.2017.12.006.

Galagan, J.E. (2014) Genomic insights into tuberculosis. *Nature Reviews Genetics*. doi:10.1038/nrg3664.

Ganbat, D., Seehase, S., Richter, E., et al. (2016) Mycobacteria infect different cell types in the human lung and cause species dependent cellular changes in infected cells. *BMC*

Pulmonary Medicine. doi:10.1186/s12890-016-0185-5.

Gbo.com (n.d.) *Greiner Bio-One Microplate Dimensions Guide*.

Getahun, H., Matteelli, A., Chaisson, R.E., et al. (2015) Latent Mycobacterium tuberculosis Infection. *New England Journal of Medicine*. doi:10.1056/NEJMra1405427.

Gold, B., Roberts, J., Ling, Y., et al. (2015) Rapid, semiquantitative assay to discriminate among compounds with activity against replicating or nonreplicating Mycobacterium tuberculosis. *Antimicrobial Agents and Chemotherapy*. doi:10.1128/AAC.00803-15.

Goldman, R.C. (2013) Why are membrane targets discovered by phenotypic screens and genome sequencing in Mycobacterium tuberculosis? *Tuberculosis*. doi:10.1016/j.tube.2013.09.003.

Green, M. and Sambrook, J. (2012) *Molecular cloning*. doi:10.3724/SP.J.1141.2012.01075.

Grover, S., Alderwick, L.J., Mishra, A.K., et al. (2014) Benzothiazinones mediate killing of Corynebacterineae by blocking decaprenyl phosphate recycling involved in cell wall biosynthesis. *Journal of Biological Chemistry*. doi:10.1074/jbc.M113.522623.

Guerin, M.E., Kaur, D., Somashekar, B.S., et al. (2009) New insights into the early steps of phosphatidylinositol mannoside biosynthesis in mycobacteria: PimB′ is an essential enzyme of Mycobacterium smegmatis. *Journal of Biological Chemistry*. doi:10.1074/jbc.M109.030593.

Gupta, A. and Bhakta, S. (2012) An integrated surrogate model for screening of drugs against mycobacterium tuberculosis. *Journal of Antimicrobial Chemotherapy*. doi:10.1093/jac/dks056.

Haribabu, J., Subhashree, G.R., Saranya, S., et al. (2015) Synthesis, crystal structure, and in vitro and in silico molecular docking of novel acyl thiourea derivatives. *Journal of Molecular Structure*. doi:10.1016/j.molstruc.2015.03.035.

Harrison, J., Lloyd, G., Joe, M., et al. (2016) Lcp1 is a phosphotransferase responsible for ligating arabinogalactan to peptidoglycan in mycobacterium tuberculosis. *mBio*. doi:10.1128/mBio.00972-16.

Haver, H.L., Chua, A., Ghode, P., et al. (2015) Mutations in Genes for the F 420 Biosynthetic Pathway and a Nitroreductase Enzyme Are the Primary Resistance Determinants in Spontaneous In Vitro -Selected PA-824-Resistant Mutants of Mycobacterium tuberculosis. *Antimicrobial Agents and Chemotherapy*. doi:10.1128/AAC.00308-15.

Hershkovitz, I., Donoghue, H.D., Minnikin, D.E., et al. (2015) Tuberculosis origin: The Neolithic scenario. *Tuberculosis*. doi:10.1016/j.tube.2015.02.021.

Inc, N.E.B. (n.d.) *pMAL™ Protein Fusion and Purification System*. In . Version 5.

Iseman, M.D. (2002) Tuberculosis therapy: past, present and future. *European Respiratory Journal*. doi:10.1183/09031936.02.00309102.

Islam, M.M., Hameed, H.M.A., Mugweru, J., et al. (2017) Drug resistance mechanisms and novel drug targets for tuberculosis therapy. *Journal of Genetics and Genomics*. doi:10.1016/j.jgg.2016.10.002.

Jankute, M., Byng, C. V., Alderwick, L.J., et al. (2014) Elucidation of a protein-protein interaction network involved in Corynebacterium glutamicum cell wall biosynthesis as determined by bacterial two-hybrid analysis. *Glycoconjugate Journal*. doi:10.1007/s10719-014-9549-3.

Jankute, M., Cox, J.A.G., Harrison, J., et al. (2015) Assembly of the Mycobacterial Cell Wall. *Annual Review of Microbiology*. doi:10.1146/annurev-micro-091014-104121.

Jankute, M., Grover, S., Rana, A.K., et al. (2012) Arabinogalactan and lipoarabinomannan biosynthesis: structure, biogenesis and their potential as drug targets. *Future Microbiology*. doi:10.2217/fmb.11.123.

Jeon, C.Y. and Murray, M.B. (2008) Diabetes mellitus increases the risk of active tuberculosis: A systematic review of 13 observational studies. *PLoS Medicine*. doi:10.1371/journal.pmed.0050152.

Jeon, C.Y., Murray, M.B. and Baker, M.A. (2012) Managing tuberculosis in patients with diabetes mellitus: Why we care and what we know. *Expert Review of Anti-Infective Therapy*. doi:10.1586/eri.12.75.

de Jesus Lopes Ribeiro, A.L., Degiacomi, G., Ewann, F., et al. (2011) Analogous mechanisms of resistance to benzothiazinones and dinitrobenzamides in Mycobacterium smegmatis. *PLoS ONE*. doi:10.1371/journal.pone.0026675.

Jin, Y., Xin, Y., Zhang, W., et al. (2010) Mycobacterium tuberculosis Rv1302 and Mycobacterium smegmatis MSMEG___4947 have WecA function and MSMEG__4947 is required for the growth of M. smegmatis. *FEMS Microbiology Letters*. doi:10.1111/j.1574-6968.2010.02045.x.

Kalscheuer, R., Syson, K., Veeraraghavan, U., et al. (2010) Self-poisoning of Mycobacterium tuberculosis by targeting GlgE in an α-glucan pathway. *Nature Chemical Biology*. doi:10.1038/nchembio.340.

Kaul, G., Dasgupta, A. and Chopra, S. (2018) Delpazolid. Oxazolidinone antibiotic, Treatment of tuberculosis. *Drugs of the Future*, 43 (4): 233. doi:10.1358/dof.2018.043.04.2795153.

Kaur, D., Guerin, M.E., Škovierová, H., et al. (2009) Chapter 2 Biogenesis of the Cell Wall and Other Glycoconjugates of Mycobacterium tuberculosis. *Advances in Applied Microbiology*. doi:10.1016/S0065-2164(09)69002-X.

Kazda, J., Falkinham, J.O., Pavlik, I., et al. (2009) *The ecology of mycobacteria: Impact on animal's and human's health*. doi:10.1007/978-1-4020-9413-2.

Kim, K.J., Chen, V. and Fane, A.G. (1993) Some factors determining protein aggregation during ultrafiltration. *Biotechnology and Bioengineering*. doi:10.1002/bit.260420216.

Kim, S., Thiessen, P.A., Bolton, E.E., et al. (2016) PubChem substance and compound databases. *Nucleic Acids Research*. doi:10.1093/nar/gkv951.

Koch, R. (1884) Die Aetiologie der Tuberkulose. *Mittheilungen aus dem Kaiserlichen Gesundheitsamte*. doi:10.1007/BF01765224.

Koch, R. (1890) I. Weitere Mittheilungen über ein Heilmittel gegen Tuberculose. *Deutsche Medizinische Wochenschrift*. doi:10.1055/s-0029-1207546.

Kolly, G.S., Boldrin, F., Sala, C., et al. (2014) Assessing the essentiality of the decaprenyl-phospho-d-arabinofuranose pathway in Mycobacterium tuberculosis using conditional mutants. *Molecular Microbiology*. doi:10.1111/mmi.12546.

Kolly, G.S., Mukherjee, R., Kilacskovï¿½, E., et al. (2015) GtrA protein Rv3789 is required for arabinosylation of arabinogalactan in Mycobacterium tuberculosis. *Journal of Bacteriology*. doi:10.1128/JB.00628-15.

Kramnik, I. and Beamer, G. (2016) Mouse models of human TB pathology: roles in the analysis of necrosis and the development of host-directed therapies. *Seminars in Immunopathology*. doi:10.1007/s00281-015-0538-9.

Larrouy-Maumus, G., Skovierova, H., Dhouib, R., et al. (2012) A small multidrug resistance-like transporter involved in the arabinosylation of arabinogalactan and lipoarabinomannan in mycobacteria. *J Biol Chem*. doi:10.1074/jbc.M112.400986.

Lebendiker, M., Maes, M. and Friedler, A. (2014) "A screening methodology for purifying proteins with aggregation problems." In *Insoluble Proteins: Methods and Protocols*. doi:10.1007/978-1-4939-2205-5_14.

Lee, R.E., Hurdle, J.G., Liu, J., et al. (2014) Spectinamides: A new class of semisynthetic antituberculosis agents that overcome native drug efflux. *Nature Medicine*. doi:10.1038/nm.3458.

Lee, Y.-J., Han, S.K., Park, J.H., et al. (2018) The effect of metformin on culture conversion in tuberculosis patients with diabetes mellitus. *The Korean journal of internal medicine*, 33 (5): 933–940. doi:10.3904/kjim.2017.249.

Leeds, J.A., Sachdeva, M., Mullin, S., et al. (2014) In vitro selection, via serial passage, of clostridium difficile mutants with reduced susceptibility to fidaxomicin or vancomycin. *Journal of Antimicrobial Chemotherapy*. doi:10.1093/jac/dkt302.

Lehmann, J. (1946) PARA-AMINOSALICYLIC ACID IN THE TREATMENT OF TUBERCULOSIS. *The Lancet*. doi:10.1016/S0140-6736(46)91185-3.

Levi, V. and González Flecha, F.L. (2002) Reversible fast-dimerization of bovine serum albumin detected by fluorescence resonance energy transfer. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1599 (1–2): 141–148. doi:10.1016/S1570-9639(02)00414-4.

Li, W., Upadhyay, A., Fontes, F.L., et al. (2014) Novel insights into the mechanism of inhibition of MmpL3, a target of multiple pharmacophores in Mycobacterium tuberculosis.

Antimicrobial Agents and Chemotherapy. doi:10.1128/AAC.03229-14.

Liu, Y., Fu, X., Shen, J., et al. (2004) Periplasmic proteins of Escherichia coli are highly resistant to aggregation: Reappraisal for roles of molecular chaperones in periplasm. *Biochemical and Biophysical Research Communications*. doi:10.1016/j.bbrc.2004.02.125.

Makarov, V., Lechartier, B., Zhang, M., et al. (2014) Towards a new combination therapy for tuberculosis with next generation benzothiazinones. *EMBO Molecular Medicine*. doi:10.1002/emmm.201303575.

Makarov, V., Manina, G., Mikusova, K., et al. (2009) Benzothiazinones Kill Mycobacterium tuberculosis by blocking Arabinan synthesis. *Science*. doi:10.1126/science.1171583.

Manina, G., Bellinzoni, M., Pasca, M.R., et al. (2010) Biological and structural characterization of the Mycobacterium smegmatis nitroreductase NfnB, and its role in benzothiazinone resistance. *Molecular Microbiology*. doi:10.1111/j.1365-2958.2010.07277.x.

Manjunatha, U., Boshoff, H.I.M. and Barry, C.E. (2009) The mechanism of action of PA-824: Novel insights from transcriptional profiling. *Communicative & integrative biology*. doi:10.4161/cib.2.3.7926.

Marcinkeviciene, J.A., Magliozzo, R.S. and Blanchard, J.S. (1995) Purification and characterization of the Mycobacterium smegmatis catalase-peroxidase involved in isoniazid activation. *J. Biol. Chem.* doi:10.1074/jbc.270.38.22290.

De Marco, A., Vigh, L., Diamant, S., et al. (2005) Native folding of aggregation-prone recombinant proteins in Escherichia coli by osmolytes, plasmid- or benzyl alcohol-overexpressed molecular chaperones. *Cell Stress and Chaperones*. doi:10.1379/CSC-139R.1.

Marx, V. (2014) Pouring over liquid handling. *Nature Methods*. doi:10.1038/nmeth.2785.

McNeil, M., Daffe, M. and Brennan, P.J. (1990) Evidence for the nature of the link between the arabinogalactan and peptidoglycan of mycobacterial cell walls. *Journal of Biological Chemistry*.

Meadow, P.M., Anderson, J.S. and Strominger, J.L. (1964) Enzymatic polymerization of UDP-acetylmuramyl-L-ala·L-glu-L-lys-D-ala·D-alaand UDP-acetylglucosamine by a particulate enzyme from Staphylococcus aureus and its inhibition by antibiotics. *Biochemical and Biophysical Research Communications*. doi:10.1016/S0006-291X(64)80014-0.

Mikušová, K., Beláňová, M., Korduláková, J., et al. (2006) Identification of a novel galactosyl transferase involved in biosynthesis of the mycobacterial cell wall. *Journal of Bacteriology*. doi:10.1128/JB.00489-06.

Mikušová, K., Huang, H., Yagi, T., et al. (2005) Decaprenylphosphoryl arabinofuranose, the donor of the D-arabinofuranosyl residues of mycobacterial arabinan, is formed via a two-step epimerization of decaprenylphosphoryl ribose. *Journal of Bacteriology*. doi:10.1128/JB.187.23.8020-8025.2005.

Mills, J.A., Motichka, K., Jucker, M., et al. (2004) Inactivation of the mycobacterial rhamnosyltransferase, which is needed for the formation of the arabinogalactan-peptidoglycan linker, leads to irreversible loss of viability. *Journal of Biological Chemistry*. doi:10.1074/jbc.M407782200.

Minnikin, D.E., Dobson, G. and Parlett, J.H. (1985) "Extraction and Chromatographic Analysis of Characteristic Mycobacterial Lipids." In Habermehl, K.-O. (ed.). *Rapid Methods and Automation in Microbiology and Immunology*. Berlin, Heidelberg, 1985. Springer Berlin Heidelberg. pp. 274–282.

Minnikin, D.E., Lee, O.Y. and Wu, H.H.T. (2015) "Pathophysiological Implications of Cell Envelope Structure in Mycobacterium tuberculosis and Related Taxa." In *Tuberculosis - Expanding Knowledge*. doi:dx.doi.org/10.5772/59585.

Mitchison, D. and Davies, G. (2012) The chemotherapy of tuberculosis: past, present and future. *The International Journal of Tuberculosis and Lung Disease*. doi:10.5588/ijtld.12.0083.The.

Monin, L. and Khader, S.A. (2014) Chemokines in tuberculosis: The good, the bad and the ugly. *Seminars in Immunology*. doi:10.1016/j.smim.2014.09.004.

Muramatsu, Y., Ishii, M.M. and Inukai, M. (2003) Studies on novel bacterial translocase I inhibitors, A-500359s. II. Biological activities of A-500359 A, C, D and G. *J Antibiot (Tokyo)*.

Murray, J.F., Schraufnagel, D.E. and Hopewell, P.C. (2015) Treatment of tuberculosis: A historical perspective. *Annals of the American Thoracic Society*. doi:10.1513/AnnalsATS.201509-632PS.

Myllymäki, H., Niskanen, M., Oksanen, K.E., et al. (2015) Animal models in tuberculosis research – where is the beef? *Expert Opinion on Drug Discovery*. doi:10.1517/17460441.2015.1049529.

Neelsen, F.K.A. (1892) *Grundriss der pathologisch-anatomischen Technik fur praktische Arzte und Studierende*.

Neres, J., Hartkoorn, R.C., Chiarelli, L.R., et al. (2015) 2-carboxyquinoxalines kill mycobacterium tuberculosis through noncovalent inhibition of DprE1. *ACS Chemical Biology*. doi:10.1021/cb5007163.

Neres, J., Pojer, F., Molteni, E., et al. (2012) Structural basis for benzothiazinone-mediated killing of Mycobacterium tuberculosis. *Science Translational Medicine*. doi:10.1126/scitranslmed.3004395.

Ojha, A.K., Baughn, A.D., Sambandan, D., et al. (2008) Growth of Mycobacterium tuberculosis biofilms containing free mycolic acids and harbouring drug-tolerant bacteria. *Molecular Microbiology*. doi:10.1111/j.1365-2958.2008.06274.x.

OligoAnalyzer 3.1 (n.d.).

Ortalo-Magne, A., Dupont, M.A., Lemassu, A., et al. (1995) Molecular composition of the

outermost capsular material of the tubercle bacillus. *Microbiology*. doi:10.1099/13500872-141-7-1609.

Palucci, I. and Delogu, G. (2018) Host Directed Therapies for Tuberculosis: Futures Strategies for an Ancient Disease. *Chemotherapy*. doi:10.1159/000490478.

Parikka, M., Hammarén, M.M., Harjula, S.K.E., et al. (2012) Mycobacterium marinum Causes a Latent Infection that Can Be Reactivated by Gamma Irradiation in Adult Zebrafish. *PLoS Pathogens*. doi:10.1371/journal.ppat.1002944.

Parish, T. and Roberts, D.M. (2015) *Mycobacteria protocols: Third edition*. doi:10.1007/978-1-4939-2450-9.

Peroutka Iii, R.J., Orcutt, S.J., Strickler, J.E., et al. (2011) SUMO fusion technology for enhanced protein expression and purification in prokaryotes and eukaryotes. *Methods in molecular biology (Clifton, N.J.)*. doi:10.1007/978-1-61737-967-3_2.

Pethe, K., Bifani, P., Jang, J., et al. (2013) Discovery of Q203, a potent clinical candidate for the treatment of tuberculosis. *Nature Medicine*. doi:10.1038/nm.3262.

Piton, J., Foo, C.S.Y. and Cole, S.T. (2017) Structural studies of Mycobacterium tuberculosis DprE1 interacting with its inhibitors. *Drug Discovery Today*. doi:10.1016/j.drudis.2016.09.014.

Pore, V.S., Divse, J.M., Charolkar, C.R., et al. (2015) *Design and synthesis of 11α-substituted bile acid derivatives as potential anti-tuberculosis agents*. doi:10.1016/j.bmcl.2015.08.006.

Porter, C.M. and Miller, B.G. (2012) Cooperativity in monomeric enzymes with single ligand-binding sites. *Bioorganic Chemistry*. doi:10.1016/j.bioorg.2011.11.001.

PrimerX (n.d.).

Primm, T.P. and Franzblau, S.G. (2007) Recent Advances in Methodologies for the Discovery of Antimycobacterial Drugs. *Current Bioactive Compounds*. doi:10.2174/157340707781695550.

Protein Data Bank (n.d.).

ProtParam (n.d.).

Qiagen and Qiagen (2005) QIAGEN® Plasmid Purification Handbook. *Plasmid*.

Quémard, A. (2016) New Insights into the Mycolate-Containing Compound Biosynthesis and Transport in Mycobacteria. *Trends in Microbiology*. doi:10.1016/j.tim.2016.04.009.

Rana, A.K., Singh, A., Gurcha, S.S., et al. (2012) Ppm1-Encoded Polyprenyl Monophosphomannose Synthase Activity Is Essential for Lipoglycan Synthesis and Survival in Mycobacteria. *PLoS ONE*. doi:10.1371/journal.pone.0048211.

Resh, M.D. (2015) "Lipid Modification of Proteins." In *Biochemistry of Lipids, Lipoproteins*

and Membranes: Sixth Edition. doi:10.1016/B978-0-444-63438-2.00013-4.

Sala, C. and Hartkoorn, R.C. (2011) Tuberculosis drugs: new candidates and how to find more. *Future Microbiol*. doi:10.2217/fmb.11.46.

Salgame, P., Geadas, C., Collins, L., et al. (2015) Latent tuberculosis infection - Revisiting and revising concepts. *Tuberculosis*. doi:10.1016/j.tube.2015.04.003.

Sambandan, D., Dao, D.N., Weinrick, B.C., et al. (2013) Keto-Mycolic acid-dependent pellicle formation confers tolerance to drug-sensitive Mycobacterium tuberculosis. *mBio*. doi:10.1128/mBio.00222-13.

Scherman, M.S., Winans, K.A., Stern, R.J., et al. (2003) Drug targeting Mycobacterium tuberculosis cell wall synthesis: Development of a microtiter plate-based screen for UDP-galactopyranose mutase and identification of an inhibitor from a uridine-based library. *Antimicrobial Agents and Chemotherapy*. doi:10.1128/AAC.47.1.378-382.2003.

Sharma, S., Gelman, E., Narayan, C., et al. (2014) Simple and rapid method to determine antimycobacterial potency of compounds by using autoluminescent mycobacterium tuberculosis. *Antimicrobial Agents and Chemotherapy*. doi:10.1128/AAC.03205-14.

Shirude, P.S., Shandil, R.K., Manjunatha, M.R., et al. (2014) Lead optimization of 1,4-azaindoles as antimycobacterial agents. *Journal of Medicinal Chemistry*. doi:10.1021/jm500571f.

Smith, N.H., Hewinson, R.G., Kremer, K., et al. (2009) Myths and misconceptions: The origin and evolution of Mycobacterium tuberculosis. *Nature Reviews Microbiology*. doi:10.1038/nrmicro2165.

So, A.D., Gupta, N., Brahmachari, S.K., et al. (2011) "Towards new business models for R&D for novel antibiotics." In *Drug Resistance Updates*. 2011. doi:10.1016/j.drup.2011.01.006.

Soetaert, K., Rens, C., Wang, X.M., et al. (2015) Increased vancomycin susceptibility in mycobacteria: A new approach to identify synergistic activity against multidrug-resistant mycobacteria. *Antimicrobial Agents and Chemotherapy*. doi:10.1128/AAC.04856-14.

Sorrell, F.J., Greenwood, G.K., Birchall, K., et al. (2010) Development of a differential scanning fluorimetry based high throughput screening assay for the discovery of affinity binders against an anthrax protein. *Journal of Pharmaceutical and Biomedical Analysis*. doi:10.1016/j.jpba.2010.02.024.

Sridhar, S., Dash, P. and Guruprasad, K. (2016) Comparative analyses of the proteins from Mycobacterium tuberculosis and human genomes: Identification of potential tuberculosis drug targets. *Gene*. doi:10.1016/j.gene.2015.12.054.

Stover, C.K., De La Cruz, V.F., Fuerst, T.R., et al. (1991) New use of BCG for recombinant vaccines. *Nature*. doi:10.1038/351456a0.

Stover, C.K., Warrener, P., VanDevanter, D.R., et al. (2000) A small-molecule nitroimidazopyran drug candidate for the treatment of tuberculosis. *Nature*.

doi:10.1038/35016103.

Taneja, N.K. and Tyagi, J.S. (2007) Resazurin reduction assays for screening of anti-tubercular compounds against dormant and actively growing Mycobacterium tuberculosis, Mycobacterium bovis BCG and Mycobacterium smegmatis. *Journal of Antimicrobial Chemotherapy*. doi:10.1093/jac/dkm207.

Tangallapally, R., Yendapally, R., Daniels, A., et al. (2007) Nitrofurans as Novel Anti-tuberculosis Agents: Identification, Development and Evaluation. *Current Topics in Medicinal Chemistry*. doi:10.2174/156802607780059772.

TBfacts.org (2018) *Bangladesh MDR TB treatment regimen – Francophone Study, STREAM*.

Telenti, A., Philipp, W.J., Sreevatsan, S., et al. (1997) The emb operon, a gene cluster of Mycobacterium tuberculosis involved in resistance to ethambutol. *Nature Medicine*. doi:10.1038/nm0597-567.

Tiemersma, E.W., van der Werf, M.J., Borgdorff, M.W., et al. (2011) Natural history of tuberculosis: Duration and fatality of untreated pulmonary tuberculosis in HIV negative patients: A systematic review. *PLoS ONE*. doi:10.1371/journal.pone.0017601.

Trefzer, C., Škovierová, H., Buroni, S., et al. (2012) Benzothiazinones are suicide inhibitors of mycobacterial decaprenylphosphoryl-β-d-ribofuranose 2′-oxidase DprE1. *Journal of the American Chemical Society*. doi:10.1021/ja211042r.

UniProt Consortium, T. (2018) UniProt: the universal protein knowledgebase. *Nucleic Acids Research*. doi:10.1093/nar/gky092.

Villemin, J.A. (2015) On the virulence and specificity of tuberculosis. *The International Journal of Tuberculosis and Lung Disease*, 19 (3): 256–266. doi:10.5588/ijtld.06.0636-v.

Wang, C.K., Weeratunga, S.K., Pacheco, C.M., et al. (2012) DMAN: A java tool for analysis of multi-well differential scanning fluorimetry experiments. *Bioinformatics*. doi:10.1093/bioinformatics/btr664.

Wang, F., Sambandan, D., Halder, R., et al. (2013) Identification of a small molecule with activity against drug-resistant and persistent tuberculosis. *Proceedings of the National Academy of Sciences*. doi:10.1073/pnas.1309171110.

Warrier, T., Tropis, M., Werngren, J., et al. (2012) Antigen 85C inhibition restricts Mycobacterium tuberculosis growth through disruption of cord factor biosynthesis. *Antimicrobial Agents and Chemotherapy*. doi:10.1128/AAC.05742-11.

Van De Weert, M. and Stella, L. (2011) Fluorescence quenching and ligand binding: A critical discussion of a popular methodology. *Journal of Molecular Structure*. doi:10.1016/j.molstruc.2011.05.023.

Weston, A., Stern, R.J., Lee, R.E., et al. (1997) Biosynthetic origin of mycobacterial cell wall galactofuranosyl residues. *Tubercle and Lung Disease*. doi:10.1016/S0962-8479(98)80005-1.

WHO (2010) “Treatment of tuberculosis: guidelines.” In *4Th Edition*. doi:10.1164/rccm.201012-1949OC.

WHO (2016) WHO treatment guidelines for drug-resistant tuberculosis : 2016 update. *Who*. doi:WHO/HTM/TB/2016.04.

WHO (2017) *Global Tuberculosis Report 2017*. doi:10.1001/jama.2014.11450.

Wolucka, B.A., McNeil, M.R., De Hoffmann, E., et al. (1994) Recognition of the lipid intermediate for arabinogalactan/arabinomannan biosynthesis and its relation to the mode of action of ethambutol on mycobacteria. *Journal of Biological Chemistry*.

Yano, T., Kassovska-Bratinova, S., Teh, J.S., et al. (2011) Reduction of Clofazimine by Mycobacterial Type 2 NADH:Quinone Oxidoreductase. *Journal of Biological Chemistry*. doi:10.1074/jbc.M110.200501.

Yempalla, K.R., Munagala, G., Singh, S., et al. (2015) Nitrofuranyl Methyl Piperazines as New Anti-TB Agents: Identification, Validation, Medicinal Chemistry, and PK Studies. *ACS Medicinal Chemistry Letters*. doi:10.1021/acsmedchemlett.5b00141.

Zhang, J., Chung, T. and Oldenburg, K. (1999) A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *Journal of biomolecular screening*. doi:10.1177/108705719900400206.

Zhang, Y. and Mitchison, D. (2003) The curious characteristics of pyrazinamide: A review. *International Journal of Tuberculosis and Lung Disease*. doi:10.5588/ijtld.16.0360.

7 Appendix

7.1 DNA oligonucleotides

7.1.1 Mt-DprE1 mutagenesis primers

E1_P316G_fw	GAACCTGACGCAATTTTACCATGGTCTGGATATGTTCGGTGAATGG
E1_P316G_re	CCATTCACCGAACATATCCAGACCATGGTAAAATTGCGTCAGGTTC
E1_L317A_fw	GACGCAATTTTACCATCCGGCCGATATGTTCGGTGAATGG
E1_L317A_re	CCATTCACCGAACATATCGGCCGGATGGTAAAATTGCGTC
E1_D318A_fw	GCAATTTTACCATCCGCTGGCCATGTTCGGTGAATGGAATCG
E1_D318A_re	CGATTCCATTCACCGAACATGGCCAGCGGATGGTAAAATTGC
E1_M319A_fw_as	TTTACCATCCGCTGGACGCCTTCGGTG
E1_M319A_re_as	ATACGCGCGATTCCATTCACCGAAGGCGTCC
E1_F320A_fw	CCATCCGCTGGATATGGCCGGTGAATGGAATCGCG
E1_F320A_re	CGCGATTCCATTCACCGGCCATATCCAGCGGATGG
E1_E322A_fw	GCTGGATATGTTCGGTGCCTGGAATCGCGCGTATGGC
E1_E322A_re	GCCATACGCGCGATTCCAGGCACCGAACATATCCAGC
E1_W323A_fw	GCTGGATATGTTCGGTGAAGCCAATCGCGCGTATGGCCCG
E1_W323A_re	CGGGCCATACGCGCGATTGGCTTCACCGAACATATCCAGC
E1_N324A_fw	GATATGTTCGGTGAATGGGCCCGCGCGTATGGCCCGGCTG
E1_N324A_re	CAGCCGGGCCATACGCGCGGGCCCATTCACCGAACATATC
E1_Y327A_fw	GTGAATGGAATCGCGCGGCCGGCCCGGCTGGTTTTC
E1_Y327A_re	GAAAACCAGCCGGGCCGGCCGCGCGATTCCATTCAC
E1_G328P_fw	GTGAATGGAATCGCGCGTATCCGCCGGCTGGTTTTCTGCAGTAC
E1_G328P_re	GTACTGCAGAAAACCAGCCGGCGGATACGCGCGATTCCATTCAC
E1_P329A_fw	GGAATCGCGCGTATGGCGCCGCTGGTTTTCTGCAGTAC
E1_P329A_re	GTACTGCAGAAAACCAGCGGCGCCATACGCGCGATTCC

7.1.2 Sequencing primers for Mt-DprE1 pointmutants

E1seq_fw	GACGGGTGAAGATGCGGAAC
E1seq_re	CAGCTGTTCAACGGTTGCCAG

7.1.3 Periplasmic MBP-Mt-DprE2 cloning primers

E2_SapI_removal_fw	GGCGACGCCGAAGAATTGTGGCAGAACCAG
E2_SapI_removal_re	CTGGTTCTGCCACAATTCTTCGGCGTCGCC
E2_pMAL_SapI_fw	TATTAGCTCTTCCGTTCTTGATGCC
E2_pMAL_HindIII_re	TACCTAAGCTTTCAGATGGGCAACTTG

7.1.4 6xHis-SUMO-Mt-DprE2

E2_BamHI_fw	TGGTGGATCCGTTCTTGATGCCGTAGG
E2_HindIII_re	ACCTAAGCTTTCAGATGGGCAGCTTGC

7.1.5 Tag-free Mt-DprE2 cloning primers

sumotag_deletion_fw	GTTTCCTACGGCATCAAGAACCATATGTATATCTCCTTCTT
sumotag_deletion_re	AAGAAGGAGATATACATATGGTTCTTGATGCCGTAGGAAAC

7.2 Protein sequences

7.2.1 6xHis-Mt-DprE1

His-tag, DprE1 uniprot ID: P9WJF1. Molecular weight: 51 775.84 Da fusion protein. Theoretical pI: 6.98.

MGSSHHHHHHSQDPMLSVGATTTATRLTGWGRTAPSVANVLRTPDAEMIVKAVA
RVAESGGGRGAIARGLGRSYGDNAQNGGGLVIDMTPLNTIHSIDADTKLVDIDAG
VNLDQLMKAALPFGLWVPVLPGTRQVTVGGAIACDIHGKNHHSAGSFGNHVRSM
DLLTADGEIRHLTPTGEDAELFWATVGGNGLTGIIMRATIEMTPTSTAYFIADGDV
TASLDETIALHSDGSEARYTYSSAWFDAISAPPKLGRAAVSRGRLATVEQLPAKLR
SEPLKFDAPQLLTLPDVFPNGLANKYTFGPIGELWYRKSGTYRGKVQNLTQFYHPL
DMFGEWNRAYGPAGFLQYQFVIPTEAVDEFKKIIGVIQASGHYSFLNVFKLFGPRN
QAPLSFPIPGWNICVDFPIKDGLGKFVSELDRRVLEFGGRLYTAKDSRTTAETFHA
MYPRVDEWISVRRKVDPLRVFASDMARRLELL

7.2.2 MBP-Mt-DprE2

MBP-DprE2. Features: Xa protease recognition site, periplasmic translocation signal. Molecular weight: 72 350.73 Da fusion protein, MBP-tag is 40 208.69 Da, DprE2 is 27 338 Da. Theoretical pI: 5.87.

MKIKTGARILALSALTTMMFSASALAKIEEGKLVIWINGDKGYNGLAEVGKKFEK
DTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAF
QDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKA
KGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFL
VDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPT
FKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVAL
KSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDE

ALKDAQTNSSSNNNNNNNNNNLGIEGRVLDAVGNPQTVLLLGGTSEIGLAICERY
LHNSAARIVLACLPDDPRREDAAAAMKQAGARSVELIDFDALDTDSHPKMIEAAF
SGGDVDVAIVAFGLLGDAEELWQNQRKAVQIAEINYTAAVSVGVLLAEKMRAQG
FGQIIAMSSAAGERVRRANFVYGSTKAGLDGFYLGLSEALREYGVRVLVIRPGQV
RTRMSAHLKEAPLTVDKEYVANLAVTASAKGKELVWAPAAFRYVMMVLRHIPR
SIFRKLPI

7.2.3 6xHis-SUMO-Mt-DprE2

6xHis-SUMO-DprE2. Molecular weight: 40 766 Da fusion protein, DprE2 is 27 338 Da. Theoretical pI: 6.33.

MGSSHHHHHHGSGLVPRGSASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEI
FFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQADQTPEDLDMEDNDIIEA
HREQIGGGSVLDAVGNPQTVLLLGGTSEIGLAICERYLHNSAARIVLACLPDDPRR
EDAAAAMKQAGARSVELIDFDALDTDSHPKMIEAAFSGGDVDVAIVAFGLLGDA
EELWQNQRKAVQIAEINYTAAVSVGVLLAEKMRAQGFGQIIAMSSAAGERVRRA
NFVYGSTKAGLDGFYLGLSEALREYGVRVLVIRPGQVRTRMSAHLKEAPLTVDKE
YVANLAVTASAKGKELVWAPAAFRYVMMVLRHIPRSIFRKLPI

7.2.4 Tag-free Mt-DprE2

Molecular weight: 27 469 Da. Theoretical pI: 7.78.

MVLDAVGNPQTVLLLGGTSEIGLAICERYLHNSAARIVLACLPDDPRREDAAAAM
KQAGARSVELIDFDALDTDSHPKMIEAAFSGGDVDVAIVAFGLLGDAEELWQNQR
KAVQIAEINYTAAVSVGVLLAEKMRAQGFGQIIAMSSAAGERVRRANFVYGSTKA
GLDGFYLGLSEALREYGVRVLVIRPGQVRTRMSAHLKEAPLTVDKEYVANLAVT
ASAKGKELVWAPAAFRYVMMVLRHIPRSIFRKLPI