INHIBITION OF THE OXIDOREDUCTASE ENZYME DPRE1 OF *MYCOBACTERIUM TUBERCULOSIS* BY TWO SMALL COMPOUNDS

and

INVESTIGATING THE ROLE OF THE CHAPERONINS OF *MYCOBACTERIUM MARINUM*

by

LAWRENCE BELL

A thesis submitted to the
University of Birmingham
for the degree of
MASTER OF RESEARCH

Institute of Microbiology and Infection
School of Biosciences
College of Life Sciences
University of Birmingham
September 2014

ABSTRACTS

Inhibition of the oxidoreductase enzyme DprE1 of *Mycobacterium tuberculosis* by two small compounds

DprE1 is an essential oxidoreductase enzyme found in the pathogen *Mycobacterium tuberculosis*, the causative agent of the disease tuberculosis. The recent emergence of multi-drug (MDR-) and extensively-drug resistant (XDR-TB) strains of TB has highlighted the need for new drugs to complement an outdated treatment regimen. This project details the purification of DprE1 and the efforts made to characterize its interaction with two small compounds thought to be inhibitors, 1326 and 1328. The K_d value of each compound was determined and the design of two plasmids for the overexpression of *Pseudomonas aeruginosa* and *Mycobacterium smegmatis* DprE1 homologues was detailed.

Investigating the role of the chaperonins of *Mycobacterium marinum*

The chaperonins are a distinct subset of molecular chaperones. Homologues of the essential *E. coli* chaperonin *GroEL* have been identified across the *Mycobacteria* family of bacteria. Gene knockout of Cpn60.1 (*cpn60.*1) has been achieved in *Mycobacterium marinum,* a fish pathogen closely related to the human pathogen *M. tuberculosis*. In this project efforts have been made to characterize this *cpn60.1*-knockout strain of *M. marinum*; determining growth curves, lipid profiles, response to heat-shock, ability to form biofilms and relative levels of expression of each *cpn* gene under standard and heat shock growth conditions.

INHIBITION OF THE OXIDOREDUCTASE ENZYME DPRE1 OF *MYCOBACTERIUM TUBERCULOSIS* BY TWO SMALL COMPOUNDS

ABSTRACT

In 2010, *Mycobacterium tuberculosis*, the causative agent of the disease tuberculosis (TB) was responsible for 8.8 million cases of TB of which 1.1 million were fatal. The courses of treatment for TB has remained relatively unchanged for the past 50 years allowing the development of multidrug-resistant TB (MDR-TB) and extensively drug-resistant TB (XDR-TB). Research has led to the identification of new inhibitors and potential drug targets against TB. One such target is the oxidoreductase enzyme DprE1, an essential enzyme in *Mycobacteria* that is required for the synthesis of cell wall arabinan. The dinitrobenzamides, benzothiazonines and azaindoles, are three classes of compounds that have all been identified as inhibitors of DprE1. As part of the ongoing drug-development process this paper outlines the efforts to characterise the binding of two known DprE1 inhibitors, 1328 and 1326. Using fluorescence spectroscopy, ligand-binding assays were performed to establish the dissociation constant (K_d) for each compound. Previous efforts to co-crystallise these inhibitors with DprE1 have been unsuccessful. In an attempt to overcome this, I detail the construction of two plasmids for the overexpression of the DprE1 homologues from *Mycobacterium smegmatis* and *Pseudomonas aeruginosa.* Efforts to purify *M. smegmatis* DprE1 have proven successful, whilst the *P. aeruginosa* enzyme is found to be mostly insoluble. Due to time constraints no ligand-binding studies were performed for the *M. smegmatis* and *P. aeruginosa* homologues of DprE1.

ACKNOWLEDGEMENTS

I would like to thank Dr. Klaus Fütterer for introducing me to the wonderful world of biophysics. To Dr. Sarah Batt for her patience and continued support. Finally, thanks to the rest of the Fütterer lab; Ali Kermani, Sarah Taylor, Chai Gopalasingam and Rabina Akhtar for providing daily fun and condolences on those days where things just didn't want to work.

Abbreviation/Acronym	
ACP	Acyl carrier protein
AIDS	Aquired Immune Deficiency Syndrome
AMP	Ampicillin
ATP	Adenosine triphosphate
BCG	Bacillus Calmette-Guérin
BCIP	5-bromo-4-chloro-3-indolyl phosphate
B_{max}	The total density (concentration) of receptors in a sample of tissue
BSA	Bovine serum albumin
BTZ	Benzothiazinone
CO	Carbon Monoxide
Cv	Column volume
Cys	Cysteine
DNB	Dinitrobenzamide
dNTP	Nucleoside triphosphate
DPA	decaprenylphosphoryl-D-arabinose
DPR	decaprenylphosphoryl-D-ribose
DprE1	decaprenylphosphoryl-d-ribose oxidase 1
DprE2	decaprenylphosphoryl-d-ribose oxidase 2
DPX	decaprenylphsphoryl-2-keto-β-D-arabinofuranose
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic Acid
DosR	Dormancy Survival Regulon

DOTS	Directly observed therapy, short course
DR-TB	Drug-resistant tuberculosis
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EMB	Ethambutol
FAD	Flavin adenine dinucleotide
FDA	Food and Drug Administration
FLD	First-line drugs
gDNA	Genomic DNA
Gly	Glycine
His	Histidine
HIV	Human Immunodeficiency Virus
IFN-γ	Interferon gamma
INH	Isoniazid or isonicotinylhydrazine
IPTG	Isopropyl β-D-1-thiogalactopyranoside
K_d	Dissociation constant
LAM	Lipoarabinomannan
LB	Luria-Bertani
MCS	Multiple cloning site
MDR-TB	Multidrug-resistant Tuberculosis
MIC	Minimum inhibitory concentration
NAD/NADH	Nicotinamide adenine dinucleotide
NO	Nitric Oxide
NTA	Nitrilotriacetic acid

OD	Optical density
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PZA	Pyrazinamide or pyrazinoic acid
QHP	Q Sepharose High performance
RIF	Rifampicin
RNA	Ribonucleic Acid
RNAP	RNA Polymerase
RPF	Resuscitation promoting factors
RPM	Revolutions per minute
SDS PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
Ser	Serine
SPEC	Spectinomycin
TB	Tuberculosis or terrific broth
tDNA	Template DNA
Tyr	Tyrosine
UV	Ultraviolet
WHO	World Health Organisation
XDR-TB	Extensively drug-resistant Tuberculosis

CONTENTS

1. INTRODUCTION

1.1 Tuberculosis

First described in 1882 by the German scientist Robert Koch (Fattorini, 2013) tuberculosis (TB) is a disease that has affected mankind for thousands of years (Hershkovitz *et al*., 2008). The disease is caused by certain members of the mycobacterium genus, most notably *Mycobacterium tuberculosis* and its related species; *Mycobacterium africanum, Mycobacterium bovis, Mycobacterium microti, Mycobacterium caprae, Mycobacterium pinnipedii* and *Mycobacterium canettii* (Zumla, Nahid and Cole, 2013). It is currently believed that many people are infected with latent tuberculosis. In 2010, 8.8 million incident cases of TB were reported with 1.1 million of these cases leading to fatalities (Mesfin et al, 2014).

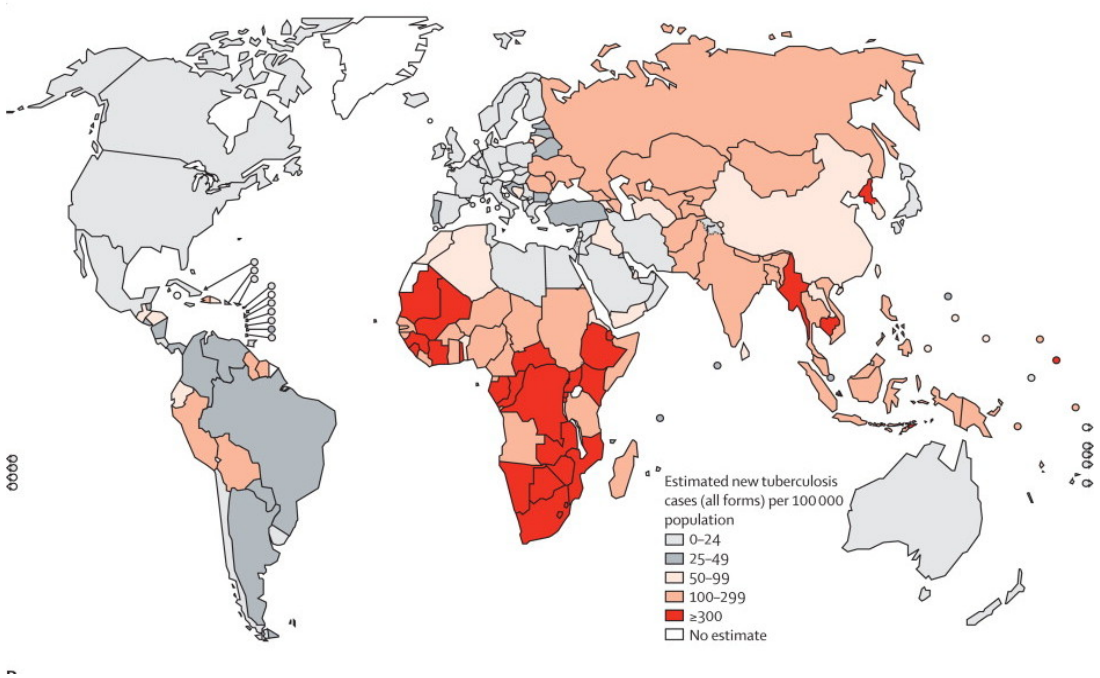


Fig. 1 Map showing prevalence of new TB cases across the world in 2010.

Taken from Lawn & Zumla, 2011

1.2 *M. tuberculosis* mode of infection

M. tuberculosis most commonly resides in the lungs, resulting in pulmonary TB but may also inhabit other areas of the body. Transmission of *M. tuberculosis* occurs via airborne droplets expectorated from infected individuals. The bacteria mostly reside in the phagosomal compartment of macrophage cells within the body (Ehrt & Rhee, 2013). By residing in these immune cells *M. tuberculosis* can effectively avoid the body's immune defence and persist indefinitely unless treated with an intense regime of drugs. Certain groups are at greater risk of developing active TB, such as individuals with diabetes mellitus (Baghaei et al, 2013) and HIV-positive individuals (Mesfin et al, 2014). Patients with HIV are 20-fold more likely to develop active TB than non HIV-TB co-infected patients, with the disease being "the most common cause of AIDS-related death" (Pawlowski et al, 2012). Tuberculosis is most common among third-world populations where overcrowding, malnutrition and cases of HIV are prevalent (Fig. 1; Lawn & Zumla, 2011).

M. tuberculosis exhibits a number of evolutionary adaptations that make it an excellent persister. Persistence refers to the ability of the bacteria to survive long-term within a host. *M. tuberculosis* has been shown to be able to persist for decades within the same host (Lillebaek *et al*., 2002). Once inhaled, *M. tuberculosis* cells travel to the small alveoli of the lungs where they are ingested by macrophages. Mobilisation of the host immune system in response to the pathogen creates an environment that would be inhospitable to most invading organisms. IFN-γ stimulates the macrophages to produce reactive oxygen species and increase acid stress (Saviola, 2013). The change in environmental conditions stimulates *M. tuberculosis* to elicit a stress response, altering the expression of key genes. 48 of these co-regulated key genes comprise the dormancy survival regulon (DosR), with their expression being controlled by the

response regulator DosR. This regulon is induced when the ability for aerobic respiration is inhibited, promoting the bacteria to enter a dormant, non-replicative state (Leistikow *et al.,* 2009). The regulon exists as a three-component system with DosR acting synergistically with the sensor histidine-kinases, DosT and DosS, which are activated in response to NO and CO (Leistikow *et al.,* 2010; Sivaramakrishnan & Ortiz de Montellano, 2013). These dormant, or latent cells can remain viable for many years. The subsequent host adaptive immune response results in the migration of other immune cells, e.g. lymphocytes, to the site of infection. These cells eventually form a clump of cells, dubbed a granuloma, also known as the primary lesion or Ghon's complex, which can become calcified (Delogu, Sali and Fadda, 2013). Reactivation of dormant TB cells requires stimuli by a family of molecules known as the resuscitation promoting factors (Rpf). The exact mode of re-emergence of dormant TB isn't entirely known, however it is suggested that when a high density of replicative cells is reached the host immune response is stimulated. This could lead to the liquefaction of the primary granuloma, releasing the tubercule bacilli (Fattorini et al, 2013).

1.3 Treatment of TB

Due to the aforementioned adaptations of *M. tuberculosis*, treatment of infection is a drawn-out multidrug process. The first drug discovered to show efficacy against TB was streptomycin, an aminogylcoside antibiotic derived from *Streptomyces griseus*. However, it was quickly discovered that single drug treatment of TB led to rapid development of antibiotic resistance (Zumla, Nahid & Cole, 2013). Current treatment of drug-susceptible pulmonary TB involves the use of four first-line drugs (FLDs); ethambutol, rifampicin, pyrazinamide and isoniazid. In areas affected by a high prevalence of antibiotic-resistant strains of TB, all four FLDs are administered in a Directly Observed Therapy, Short-Course

(DOTS) to ensure patient compliance with the drug regimen. This course of treatment takes at least six months – 2 months with daily doses of all 4 drugs, followed by a 4-month continuation phase consisting of isoniazid and rifampicin (Manca et al, 2013).

Ethambutol (EMB) was first observed to alter cell morphology and reduce clumping in the *M. tuberculosis*-related organism *Mycobacterium smegmatis*. This observation was hypothesised to be a result of cell wall disruption (Kilburn and Greenberg, 1977). EMB has since been shown to inhibit cell wall biosynthesis, specifically the synthesis of two key components; arabinogalactan and lipoarabinomannan (LAM) (Goude et al, 2009). EMB targets the proteins coded for by the *emb* operon consisting of the genes *embC, embA* and *embB*. This operon codes for three essential arabinosyltransferase enzymes EmbC, EmbA and EmbB.

Rifampicin (RIF) inhibits RNA transcription in *M. tuberculosis* cells by binding with high-affinity to the DNA-dependent RNA Polymerase (RNAP; Wehrli & Staehelin, 1971), more specifically to an area at the binding pocket within the DNA/RNA channel of the RNAP beta subunit (Campbell et al, 2001).

Pyrazinamide (PZA) is given to patients in the form of a prodrug. It only becomes toxic to *M. tuberculosis* after the bacteria take up the compound. Once ingested it is converted to the active form, pyrazinoic acid (PZA) by the enzyme pyrazinamidase (PncA) enzyme (Rajendran *et al*., 2014), which inhibits protein and RNA synthesis as well as disrupting membrane potential at low pHs (Zhang *et al*, 2003).

Isoniazid or **isonicotinylhydrazine** (INH), like PZA is administered as a prodrug. It enters *M. tuberculosis* by passive diffusion, at which point it is converted to active forms by the Mycobacterium enzyme KatG. The KatG-catalysed peroxidation reaction of INH leads to the formation of two radicals: isonicotinic hydrazyl radical and isonicotinoyl radical (Timmins & Deretic, 2006). These radicals have been proposed to bind to NAD forming an INH-NAD adduct, which inhibits the enzyme inhA – an NADH-dependent enoyl-ACP reductase crucial for mycolic acid biosynthesis (Vilchèze & Jacobs, Jr., 2007).

Although effective, these drugs can have side effects that increase morbidity from TB treatment, often leading to poor patient compliance (Yee *et al.*, 2003). Ethambutol has been directly shown to cause ocular neuritis (swelling of the optic nerve; Griffith *et al*., 2005). A common side effect of Isoniazid is hepatotoxicity, which has been shown to affect around 0.15% of patients, with risk of incidence increasing with age (Nolan, Goldberg & Buskin, 1999). Rifampicin has been shown to affect the processing of other drugs administered in conjunction with it. This is a result of RIF-induced stimulation of hepatic cytochrome P450 oxidase and glycoprotein transport systems, which can increase the metabolism of other administered medicines (Baciewicz *et al.*, 2013). For example, RIF administered to patients already prescribed the blood thinner warfarin has been shown to cause hematuria (Martins *et al.*, 2013).

All of these drugs fall under the World Health Organisation (WHO) Group 1 category of drugs for the treatment of TB. 4 other groups (Table. 1) have been categorised, which are less suitable for treating drug-susceptible TB due to decreased efficacy and/or toxicity.

Grouping	Drugs
Group1 First-line oral agents	Isoniazid (H); rifampicin (R); ethanbutol (E); pyrazinamide (Z) rifabutin (Rfb)
Group 2 Injectable agents	Kanamycin (Km); amikacin (Am); capreomycin (Cm); viomycin (Vm); streptomycin (S)
Group 3 Fluoroquinolones	Moxifloxacin (Mfx); levofloxacin (Lfx); ofloxacin (Ofx)
Group 4 Oral bacteriostatic second-line agents	Ethionamide (Eto); protionamide (Pto); cycloserine (Cs); teizidone (Trd); *p*-aminosalicylic acid (PAS)
Group 5 Agents with unclear role in DR-TB treatment (not recommended by WHO for routine use in DR-TB patients)	Clofazimine (Cfz); linezolid (Lzd); amoxicillin/clavulanate (Amx/Clv); thioacetazone (Thz); imipenem/cilastatin (Ipm/Cln); high-dose isoniazid (high-dose H); clarithromycin (Clr)

Table.1 WHO grouping of anti-TB drugs based on decreasing efficacy.

Table taken from WHO (2008)

1.4 Drug resistance in *M. tuberculosis*

The current course of treatment for tuberculosis has been in place for around 50 years. Multiple factors have led to the relatively recent emergence of multidrug-resistant TB (MDR-TB) and extensively drug-resistant TB (XDR-TB) strains. MDR-TB is defined as resistance to Rifampicin and Isoniazid, with second line (Group 2) drugs being employed for its treatment (WHO, 2013). XDR-TB strains are resistant to at least one Group 2 and Group 3 drug (Green & Garneau-Tsodikova, 2014). In 2012 there were an estimated 300,000 cases reported of MDR-TB, with 9.6% of these cases including XDR-TB (WHO, 2013).

Most resistance to Rifampicin arises from mutations in the *rpoB* gene encoding the beta-subunit of RNAP, inhibiting enzyme-ligand complex formation (Campbell *et al*., 2001). Mutations in the INH-activating enzyme, KatG as well as in the gene and promoter region of its protein target, InhA are the most common causes of INH resistance (Ando *et al*., 2013). Group 2 and 3 based treatments for MDR-TB and XDR-TB are lengthy, coupled with the potential toxicity of the compounds and inaccessibility to treatments in less-developed countries has led to an increase in the prevalence of these strains in recent years (Green & Garneau-Tsodikova, 2014; WHO, 2013). Pyrazinamide resistance is associated with mutations in the gene coding for PncA. Specific mutations in the enzyme, including D8G, S104R and C138Y have been shown to render the enzyme inactive by creating a rigid binding cavity, preventing conversion of PZA to pyrazinoid acid (Rajendran *et al.,* 2014).

Several new drugs have been identified as potential treatments for drug-resistant TB. The Food and Drug Administration (FDA) recently approved one such drug, Bedaquiline, an inhibitor of bacterial ATP Synthase (Mahajan, 2013). By binding to the C-subunit of the ATP synthase F_0 domain, Bedaquiline prevents the conformational change required for the production of ATP from ADP and PPi (Goel, 2014; Lakshmanan & Xavier, 2013). Another class of drugs, the Spectinamides, semisynthetic analogues of the antibiotic Spectinomycin, have recently been developed. Spectinamides show excellent antibacterial activity against nonreplicating *M. tuberculosis* by targeting 16S ribosomal RNA and are able to withstand efflux resistance due to the addition of a synthetic glycyl side chain (Lee et al, 2014). Other drug discovery efforts have focused on the high throughput screening of different compounds against *M. tuberculosis*, and have identified a class of compounds, the Dinitrobenzamides (DNBs) that target the oxidoreductase enzyme Decaprenyl-Phosphoribose 2' Epimerase (DprE1) coded for by the gene *dpre1* (*Rv3790*; Christophe et al, 2009).

1.5 DprE1

DprE1 is an essential enzyme for the survival of *M. tuberculosis*. Loss of this enzyme results in a phenotype characterised by swelling of the bacteria, leading to cell wall damage, lysis and death (Kolly et al, 2014). DprE1 is involved in cell wall biosynthesis, catalysing the conversion of decaprenylphosphoryl-D-ribose (DPR) to decaprenylphsphoryl-2-keto-β-D-arabinofuranose (DPX), a keto-intermediate, which is then converted to decaprenylphosphoryl-D-arabinose by a second enzyme DprE2 (Fig. 2). DPA is required for the synthesis of arabinogalactan and LAM. There is no alternative pathway for the production of arabinan in *Myobacterium*, making DprE1 and DprE2 excellent drug targets for the treatment of TB (Christophe et al, 2009). DprE1 in particular has been shown to have an active site pocket to which certain small molecules can bind to with high affinity and inhibit enzyme function (Batt et al., 2012).

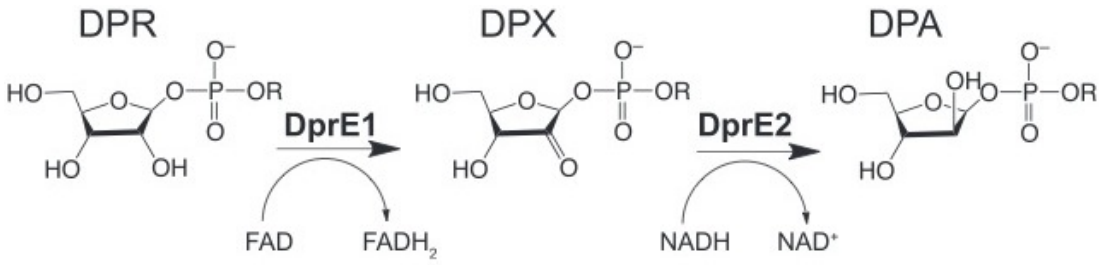


Fig. 2 DprE1- and DprE2-mediated conversion of DPR to DPA. (Taken from Batt et al, 2012)

1.6 Inhibitors of DprE1

To date, at least 4 classes of compounds have been identified as inhibitors of DprE1. They are discussed below:

Dinitrobenzamide (DNB) derivatives: Christophe *et al*., (2009), used confocal fluorescent microscopy to screen the effect of different compounds on the growth of GFP-tagged *M. tuberculosis* in macrophages. They identified a large group of compounds that showed efficacy against *M. tuberculosis*, which shared a benzamide scaffold (Fig. 2). The minimum inhibitory concentration (MIC) of these DNB compounds was found to be as low as 0.2 μM (for DNB1) and shown to be successful at killing drug-resistant strains of TB.

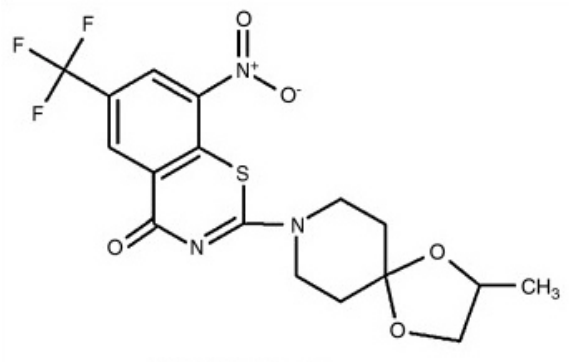
Benzothiazinones (BTZs): BTZ compounds (Fig. 3) were initially involved in a large-scale screen of compounds to identify those with antibacterial and antifungal properties (Makarov *et al*., 2006). One BTZ compound in particular, BTZ038, was selected for further study. The two enantiomers of BTZ038, named BTZ043 and BTZ044 were found to be equally as effective against *M. tuberculosis*, however BTZ043 showed the widest potency range across TB isolates, including drug-resistant strains (Makarov *et al*., 2009). BTZ043 was found to have a very low MIC against *M. tuberculosis*, around 1-4 ng/ml (Grover *et al*, 2014). In comparison, the MIC of Isoniazid, the most potent of the group 1 drugs, is around 0.03-0.06 μg/ml (Guo, 2006). The low toxicity of BTZ043, coupled with its efficacy against drug-resistant strains makes it a good candidate for treatment of MDR- and XDR-TB over current group 2 and 3 compounds (Pasca, 2010). BTZ043-resistant strains of Mycobacteria were all found to have amino acid substitutions at position 387 in DprE1, with the wild type cysteine residue replaced by Ser or Gly residues (Makarov *et al*., 2009). The x-ray crystal structure of

DprE1 in complex with another BTZ-derivative, CT325, provided evidence for the formation of a covalent bond between Cys387 and the compound (Batt *et al*., 2012), highlighting the importance of this residue for inhibitor binding.

TCA-1: A paper published by Wang et al., 2013 discusses the use of high-throughput assays to investigate the ability of compounds to inhibit biofilm formation in *Mycobacterium smegmatis.* One such molecule, TCA-1 (Fig.3C) was shown to inhibit this process in both *M. smegmatis* and *M. tuberculosis*. The same study used a TCA-1-resistant strain of *M. smegmatis* and a cosmid library to elucidate the target of TCA-1, which was found to be DprE1. The crystal structure of TCA-1 in complex with DprE1 was solved and a high level of similarity was observed between the binding nature of TCA-1 and BTZ043 to DprE1. TCA-1 resistant *M. tuberculosis* mutants were found to have an amino acid substitution at position 314 in *Rv3790*, with a Tyr replaced by Cys.

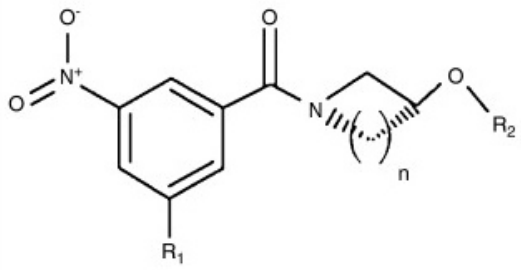
Azaindoles: The 1,4-azaindoles are a class of compounds that were developed by scaffold morphing experiments of a known non-DprE1 targeting anti-TB compound based around an imidazo-pyridine scaffold structure. Shirude et al, 2014 modified the scaffold molecule three distinct points to optimize the inhibitory effect of the compound; the amide side chain, the hydrophobic group, and substitutions of the core ring structure. A secondary amide ring structure acts essentially to maintain a low MIC (Fig. 3D). In contrast to DNBs and BTZs, the azaindoles have been shown to interact with DprE1 noncovalently. Spontaneous mutations of amino acid 314 of *Rv3790* (Y314H), the same residue implicated in TCA-1 resistance has also been shown to confer resistance to the azaindoles (Chatterji *et al.,* 2014).

A



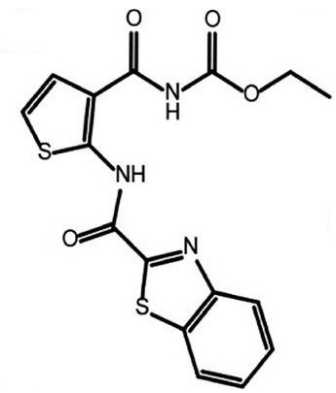
Benzothiazinone

B



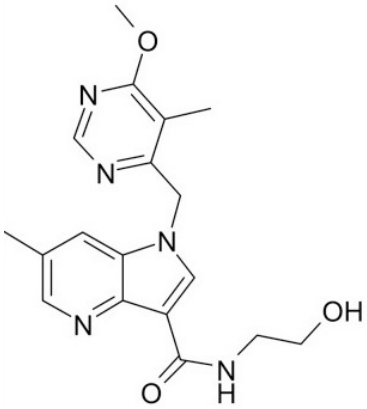
Dinitrobezamide

C



TCA-1

D



1,4-Azaindole

Fig. 3 Four of the known classes of DprE1 inhibitors

Structure of Benzothiazinone (A), Dinitrobenzamide (B) (Taken from Cole & Riccardi, 2011.), TCA-1 (C; taken from Wang *et al.*, 2013) and 1,4-Azaindole (D; taken from Chatterji *et al.*, 2014).

1.7 Previous work – DprE1 expression and purification

Previous work on DprE1 and its inhibitors has led to a specific protocol for the purification of the protein. A plasmid construct to express *M. tuberculosis* DprE1 was achieved using the vector pCDF. The gene encoding DprE1, *dprE1* (*Rv3790*) has been inserted into one of the multiple cloning sites (mcs) of pCDF. Upstream of the mcs there exists a T7 promoter - a "strong" promoter highly specific for T7 polymerase. T7 and therefore DprE1 expression is under the control of the *lac* operator, allowing expression of the gene to be switched on in response to the addition of Isopropyl β-D-1-thiogalactopyranoside (IPTG), an allolactose analog.

The use of pCDF has allowed for the expression of a modified protein; containing a C-terminal polyhistidine tag. The 6xHis tag expressed with DprE1 has a high affinity for nickel ions (Ni^{2+}). A chelating agent such as Nitrilotriacetic acid (NTA), immobilised on agarose beads is used to fix Ni^{2+} to a column. Cell lysate is then run through the column and the His-tagged protein will form bonds with the Ni^{2+} ions. Buffers containing a histidine analogue (imidazole) are then run through the column. A buffer containing a low concentration of imidazole can be used to elute proteins bound in a non-specific manner, i.e. those not expressing a His-tag, from the column. At sufficiently high concentrations, imidazole will compete with the Histidine tags for binding to Ni^{2+}, displacing the tagged protein and eluting it. This is an effective method of isolating protein to a purity of up to 95% (Bornhorst & Falke, 2000). A second chromatographic step, ion exchange chromatography is then used to mostly remove any remaining non-specifically bound proteins. This technique utilises the charge of the protein. The pH of the buffer can be altered to change the charge of the protein. When the

pH of the buffer is greater than the isoelectric point (pI), the protein achieves a net negative charge, and when pH<pI the net charge of the protein will be positive. *M. tuberculosis* DprE1 has a predicted pI of 7.769 (not including 6 His tag; Tuberculist; Lew et al, 2011), so a buffer of pH 8.5 has been used to establish a net negative charge on the protein. When passed through a column packed with strong anion quaternary ammonium (Q) ions immobolised on Sepharose beads, proteins in the solution will form ionic interactions with this stationary phase. Buffers containing increasing concentrations of NaCl are then passed through the column; the Na^+ and Cl^- ions disrupt protein-column interaction by competing with the charged functional group on the column. Different proteins will have varying isoelectric points and will therefore differ in the strength of their ionic interactions with the stationary phase. Weaker interactions will be broken at lower concentrations of NaCl, with more tightly bound proteins eluting at higher salt concentrations.

Work published in Batt *et al*., 2012, investigated the effect that co-expression of DprE1 with molecular chaperones had on the amount of DprE1 recovered. Expression of DprE1 with the *E. coli* chaperonin GroEL and the *M. tuberculosis* GroEL homologue - Cpn60.1 with the *E. coli* cochaperonin GroES both yielded small amounts of soluble protein, however the two protein products could not be readily isolated from each other by chromatography. Coexpression with the essential *M. tuberculosis* chaperonin Cpn60.2 and the *E. coli* cochaperonin GroES yielded protein products that could be separated from one another, achieving a yield of ~5mg of DprE1 per 2 litres of culture grown.

1.8 Aims and objectives

DprE1 will be overexpressed in order to study the interaction of the protein with two of its known inhibitors. Ligand binding assays will be performed using fluorescence spectroscopy. DprE1 has intrinsic fluorescence due to the presence of fluorophores - compounds that can absorb light, exciting them to a higher energy level. The light emitted when a fluorophores returns to ground energy state is of a different wavelength to the light absorbed. This is due to energy being lost to the surroundings as vibrational energy (Möller and Denicola, 2002). Aromatic amino acids such as tryptophan and tyrosine give proteins natural fluorescence. DprE1 also contains a flavin cofactor group, which also acts as a fluorophore. Within a fluorometer, a detector detects the fluorescence light emitted by a compound. Two monochromators allow the user to select the wavelengths of light the compound absorbs and the emitted wavelengths transferred to the detector. Different compounds exhibit different emission spectra. Fluorescence intensity is measured by the detector and can be used to determine the properties of ligand binding and the kinetics of the interaction. By performing titrations of the two known inhibitors with DprE1 we establish the dissociation constant (K_d) of the reaction, proportional to the affinity of the ligand for the protein.

Previous efforts to determine the structure of DprE1 in complex with the compounds 1328 and 1326 were unsuccessful. A homologue of DprE1 has been identified as a key component of the arabinofuranose pathway in *Pseudomonas aeruginosa* (Harvey et al, 2011). By establishing plasmid constructs for the overexpression of this homologue and the DprE1 homologue from the related organism *Mycobacterium smegmatis,* we hope to develop an efficient method for the purification of these two proteins. Once purified, attempts to develop crystals of these proteins in complex with the two ligands will be possible.

2. METHODS

2.1 Bacterial strains and plasmids

Strains:

Organism	Strain	Source
E. coli	BL21 (DE3)	Invitrogen™
	TOP10	Lab stock
Mycobacterium smegmatis	mc^2155	Dr. Sarah Batt (University of Birmingham)
Pseudomonas aeruginosa	PA7	Dr. Sarah Batt

Plasmids:

Plasmid name	Description	Source
pCDF-Rv3790	Plasmid for the expression of *Mycobacterium tuberculosis dpre1* (*Rv3790*). The gene is under the control of a T7 promoter and expressed with a C-terminal His-tag. An ampicillin resistance gene acts as a selectable marker.	Dr. Sarah Batt

pTrc-60.2-GroES	Plasmid containing the genes for the M. tuberculosis chaperonin Cpn60.2 (*cpn60.2*) and the *E. coli* cochaperonin GroES. Contains a spectinomycin-resistance gene.	Dr. Peter Lund (University of Birmingham)
pCDF-Rv3790ΔY314C	For the expression of a mutated form of *M. tuberculosis dpre1* (*Rv3790*) in which Cys387 has been substituted for Tyrosine. The gene is under the control of a T7 promoter and expressed with a C-terminal His-tag. An amp^R gene acts as a selectable marker.	Dr. Sarah Batt
pCDF_MS_6382	For the expression of *M. smegmatis dpre1* (*MSMEG_6382*). The gene is under the control of a T7 promoter and expressed with a C-terminal His-tag. An amp^R gene acts as a selectable marker.	This study

pCDF_psPA7_6248	Plasmid for the expression of *Pseudomonas aeruginosa dpre1* (*Rv3790*) homologue – PsPA7_6248. The gene is under the control of a T7 promoter and expressed with a C-terminal His-tag. An ampicillin resistance gene acts as a selectable marker.	This study

2.2 Primers

Primer name	Sequence (5' to 3')	Description
DprE1_SM_Sac_F	CATGCATGGAGCTCGATG GGCGCGGTACCCTCACTG	Forward primer for the PCR amplification of the *dprE1* (*Rv3790*) homologue from *M. smegmatis* (MSMEG_6382), with the introduction of an upstream *Sac*I restriction site.
DprE1_SM_Hind_RS	CATGCATGAAGCTTTCAG AGCAGTTGCAGGCGCCTG	Reverse primer for the PCR amplification of the *dprE1* (*Rv3790*) homologue from *M. smegmatis* (MSMEG_6382), with a downstream *Hind*III restriction site.

Primer name	Sequence (5’ to 3’)	Description
DprE1_PA_BamHI_FX	CATGCATGGGATCCGATGGAAGTCTTCGGCTGGGG	Forward primer for the PCR amplification of the *dprE1* (*Rv3790*) homologue from *P. aeruginosa* (PsPA7_6248), with the introduction of an upstream *Bam*HI restriction site.
DprE1_PA_Not_RS	CATGCATGGCGGCCGCTCAGTCCAGACCCAGGCGGC	Reverse primer for the PCR amplification of the *dprE1* (*Rv3790*) homologue from *P. aeruginosa* (PsPA7_6248), with the introduction of a downstream *Not*I restriction site.
pCDF-E1SM_4insF	GCGTCCGACATGGCCAGGCGCCTGCAACTGCTCTGAAAGGTTG	Forward primer for a site-directed mutagenesis reaction to introduce a single nucleotide base pair in order to correct a deletion that had occurred during the synthesis of pCDF_MSDPRE1

Primer name	Sequence (5' to 3')	Description
pCDF-E1SM_4insR	CAAGCTTTCAGAGCAGTT GCAGGCGCCTGGCCATGT CGGACGC	Reverse primer for a site-directed mutagenesis reaction to introduce a single nucleotide base pair in order to correct a deletion that had occurred during the synthesis of pCDF_MSDPRE1

2.3 Bacterial growth and media

Luria-Bertani (LB) broth and terrific broth (TB) were both used for the overexpression of proteins in *E. coli* during this study. They were prepared as below.

LB broth (1 L):

- 10 g tryptone
- 10 g NaCl
- 5 g yeast extract

Components were dissolved in 900 mL of deionized H_2O and the pH adjusted to 7.2 using NaOH. The final volume was made up to 1 L and autoclaved at 121 °C for 15 minutes.

LB agar (1 L):

LB agar was made by following the LB broth recipe with the addition of 15 g of agar before autoclaving.

TB broth (900 mL):

- 12 g tryptone
- 24 g yeast extract
- 4 mL glycerol

TB salts (100 mL)

- 2.6 g of KH_2PO_4
- 13.9 g of K_2HPO_4

Both TB broth and salts were autoclaved at 121 °C for 15 minutes. TB salts (100 mL) were added to the TB broth (900 mL) before use.

2.4 Heat shock transformation of *E. coli*

Competent cells were stored at -80 °C. Before use, a 50 µl aliquot of competent cells was allowed to thaw on ice for 5 minutes. Plasmids (1 µl each) were added directly to the competent cells and mixed gently by pipetting slowly up and down. The solution was left on ice for 30 minutes before being moved to a 42 °C water bath for 45 seconds. Immediately after heat-shock the tube was replaced on ice and left for 3 minutes. After 3 minutes, 250 µl of LB was added aseptically to the cells. Transformed cells were incubated at 37 °C, shaking at 180 RPM for 1 hour. After this time, 100 µl of cells were taken and plated on an LB agar plate containing the appropriate antibiotic. In case of poor transformation efficiency the remaining cells were pelleted by centrifugation (14000 RPM, 10 minutes), resuspended in 100 µl of LB and plated on LB agar supplemented with antibiotic.

2.5 Overexpression of *M. tuberculosis* DprE1

The plasmids pCDF_Rv3790 and pTrc-60.2-GroES were transformed into *E. coli* BL21 (DE3) competent cells (Invitrogen) by heat shock as described in section 2.4. Single colonies were picked from the plate using a sterile tip and used to inoculate 50 ml of LB media supplemented with 100 μg/ml Spectinomycin, 100 μg/ml Ampicillin and 1% glucose. The inoculated media was then incubated overnight at 37°C, 180 RPM in a shaker incubator. Overnight LB culture (10 ml) was added to a 2 L conical flask containing 1 L of terrific broth (TB) supplemented with 100 μg/ml Spectinomycin and 100 μg/ml AMP and grown at 37°C until an OD_{600nm} of 0.4-0.6 was reached. At this point the cultures were placed at 4°C for 1 hour, allowing them to cool before the addition of 0.5 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). Induced cultures were then grown at 16°C overnight in a shaker incubator before cells were harvested at 4°C by centrifugation (4000 RPM, 30 mins). Pellets were resuspended and washed in a 1X phosphate buffered saline (PBS) solution before further centrifugation (4000 RPM, 30 mins, 4°C). Pellets were then either frozen at -80°C or resuspended in 30 ml Buffer A/Lysis buffer (50 mM $NaH_2PO_4.H_20$, 300 mM NaCl, 10 mM Imidazole, pH8). An EDTA-free protease inhibitor mixture (Roche) and a small amount of FAD were added to each 30 ml sample. Cells kept on ice were then subject to lysis by sonication (Sonicator Ultrasonic Liquid Processor XL; Misonix) for 10 cycles of 20 s - on, 40 s - off. Crude cell extract was then centrifuged at 20,000 RPM, 4°C for 40 minutes to separate the soluble and insoluble cell fractions. Supernatant was then filtered through a 0.45 μm syringe filter tip.

2.6 Affinity chromatography

A Ni-NTA affinity chromatography column (1 ml Histrap HP; GE Healthcare Life Sciences) was set to a flow-rate of 1 ml/min. The column was washed with 5 column volumes (cv) of filtered deionised H_2O before being equilibrated with 5 cv of Buffer A. The filtered lysate was then loaded on to the column before 30 ml of wash buffer (50 mM Imidazole in Buffer A framework, pH8) was passed over the column. Protein was eluted using elution buffer (300mM Imidazole in Buffer A (see section 2.5, p.22), pH8). Fractions were analysed using SDS-PAGE electrophoresis (Mini-PROTEAN precast gel; BIO-RAD). Samples containing proteins of the same molecular weight as DprE1 were dialysed into buffer B (20 mM TrisHCl, 10mM NaCl, 10% Glycerol, pH 8.5). Alternatively, to avoid dialysis, after running the 30 mM Imidazole wash buffer, 10 ml of buffer B was run through the column before elution with buffer C (20 mM TrisHCl, 10 mM NaCl, 300 mM Imidazole, pH8). Glycerol (10%) was then added to the buffer C-DprE1 solution.

2.7 Ion-Exchange chromatography

Deionised H_2O (5 cv) was used to wash a strong anion exchange column (1 ml HiTrap Q HP; GE Healthcare Life Sciences) before equilibration with Buffer B (5 cv). Dialysed DprE1 solution was then loaded on to the column before buffers with varying NaCl concentrations of 20 mM steps in a gradient from 100-140 mM (Buffer B framework) were passed through the column. A final high salt buffer (500 mM NaCl, buffer B framework, pH8.5) was used to remove any remaining protein bound to the column. Fractions containing DprE1 (as determined by SDS-PAGE electrophoresis) were dialysed into Buffer B.

2.8 Concentration of protein

Protein concentration was measured by absorbance spectrophotometry at 280 nm (Nanodrop 2000 UV-Vis Spectrophotometer; Thermo Scientific) against a reference sample of buffer B. Increasing the concentration of protein in the solution was achieved by filter centrifugation (30 kDa cut-off ULTRA-4 filter tubes; Amicon). Double-deionised H_2O (4 ml) was added to each tube and centrifuged at 4000RPM for 10 minutes. Flow through was discarded and 4 ml of Buffer B was added to each tube and centrifuged again at 4000 RPM for 10 minutes. Finally dialysed protein solution (4 ml) was added to each tube and spun for 20 minutes, 4000 RPM at 4 C

2.9 Fluorescence Spectroscopy

DprE1 was diluted to a concentration of 30 mM using buffer B. Fluorescence was measured using a PTI fluorometer (PTI Quanta Master 40) with power set to 75 Joules. Samples were loaded into a quartz cuvette with a path distance of 1 cm. Temperature was maintained at 20°C by a liquid cooling system and a magnetic stirrer ensured homogeneity of the sample. Buffer B (600 μl; no glycerol) was used as a reference sample.

2.10 Design of plasmid constructs for the expression of *M. smegmatis* and *P. aeruginosa* DprE1 homologues

Primers were designed (see methods 2.2) for the amplification of DprE1-coding genes from *M. smegmatis* and *P. aeruginosa* to include flanking restriction endonuclease sites. For *M. smegmatis,* the forward primer (DprE1_SM_Sac_F) was designed to introduce a *SacI* site upstream of the M. smegmatis DprE1-coding gene (MSMEG_6382) and the reverse primer (DprE1_SM_Hind_RS) was designed to introduce a *Hind*III site downstream of the gene.

PCR reaction mixtures (50 μl) were composed of:

Component	Concentration
Deionised H_2O	variable
5X Phusion HF Buffer	1X
10 mM dNTPs	200 μM each
Forward primer (10 pM/ml)	0.5 μM
Reverse primer (10 pM/ml)	0.5 μM
Template DNA	variable
DMSO	3%
Phusion DNA Polymerase	0.02 U/μl

PCR program:

Step	Temperature (°C)	Time (min:sec)	No. of Cycles
Initial denaturation	98	05:00	x1
Denaturation	98	00:20	x40
Annealing	60 or 65	00:30	
Extension	72	00:30	
Hold	4	∞	∞

PCR products were run on 1% Agarose gel stained with Midori Green (5 µl per 100 ml Nippon Genetics Europe) to confirm the presence of amplified genes.

pCDF plasmid DNA (obtained from Dr. Sarah Batt; University of Birmingham) was transformed into *E. coli* TOP10 cells by heat shock treatment. Cells were plated on LB agar plates supplemented with 100 µg/ml SPEC and grown overnight at 37 °C. Colonies were picked and used to inoculate 5 ml of 100 µg/ml SPEC LB in cell culture tubes (BIO-RAD) and grown overnight at 37 °C in a shaker incubator. Plasmid DNA was purified from overnight cultures using a QIAprep Spin Miniprep Kit (QIAGEN).

Restriction double digest was performed by mixing 30 µl of purified plasmid or PCR product DNA with 5 Vector of 10x buffer (NEB), 2.5 µl of each restriction enzyme (NEB) and 10 µl of ddH_2O. Each sample was then incubated at 37 °C for 1 hour. Restriction products were run on 1% Agarose gel stained with midori green. The gel was visualised using near-UV (Helixx

LED Viewer) to avoid DNA damage and the bands of interest excised using a scalpel. DNA was purified using QIAquick Gel Extraction Kit (QIAGEN) before pCDF vector DNA was treated with Antarctic Phosphatase (NEB) to prevent re-circularisation. DNA concentration was measured by UV spectrophotometry at 260nm (Nanodrop 2000).

Insert and vector DNA were mixed in a ratio of 3:1, calculated using the following equation:

$$\text{Amount of insert (ng)} = \frac{\text{Amount of vector (ng)} \times \text{size of insert (kb)}}{\text{Size of vector (kb)}} \times \frac{3}{1}$$

Ligation reaction was performed using 1 μl of Quick-Stick DNA ligase, 5 μl of 4x Quick-Stick buffer (Bioline) with the final reaction volume adjusted to 20 μl with ddH_2O. A control-construct containing only vector DNA was also made using the same protocol. The ligation reaction was allowed to proceed for 10 minutes before constructs were transformed into *E. coli* One Shot TOP10 chemically competent cells (Invitrogen). 100 μl of each sample was plated onto separate 100 μg/ml SPEC LB agar before the remaining samples were centrifuged (14000 RPM, 10 minutes). Supernatant was discarded and pellets were then resuspended in 100 μl of LB and plated on 100 μg/ml SPEC LB agar. Colonies were picked and used to inoculate 5 ml of 100 μg/ml SPEC LB in 10 ml cell culture tubes and grown overnight at 37 °C in a shaker incubator. Plasmid DNA was purified from overnight cultures using a QIAprep Spin Miniprep Kit. A second restriction double digest was performed on the purified plasmids to confirm the composition of the constructs. Plasmids were sent for sequencing (Eurofins MWG Operon) using the primers from the initial amplification of the DprE1 genes.

2.11 Site-directed mutagenesis of *M. smegmatis* pCDF_MSMEG_6382 construct

Two oligonucleotide primers were designed for the introduction of the deleted base in *M. smegmatis* C4 by site-directed mutagenesis; Mutagenesis was performed using the Quickchange II XL kit (Angilent technologies). Reaction was set up as follows: 2.5 μl of 10X reaction buffer, 2.5 μl of a 1/10 dilution of pCDF-DprE1 DNA, 1 μl of each primer (1/10 diluted), 0.5 μl dNTPs, 1.5 μl of Quiksolution and 16 μl of ddH_2O. Finally 1 μl of PfuUltra HF DNA Polymerase was added. The sample was placed in a thermal cycler and run using the program:

Step	Temperature (°C)	Time (min:sec)	No. of Cycles
Initial denaturation	95	01:00	x1
Denaturation	95	00:50	x18
Annealing	60	00:50	
Extension	68	06:00	
Final extension	68	07:00	x1

Once the PCR program had completed, 1 μl of the enzyme DpnI was added to the sample. This was then centrifuged for 1 minute at 14000 RPM on a tabletop centrifuge before being incubated for 1 hour at 37 °C. The mutated construct was then transformed and cultured in *E. coli* TOP10 cells as before. Plasmids were purified and sent for sequencing using the primers for initial DprE1 amplification.

2.12 Small scale expression studies of pCDF_PsPA7_6248

Both PA7 pCDF-DprE1 and pTrc-60.2-GroES plasmids were transformed into *E. coli* BL21 (DE3) cells by heat-shock and plated on LB agar (100 μg/ml SPEC and 100 μg/ml AMP). A colony was then used to inoculate 50 ml of LB (100 μg/ml SPEC and 100 μg/ml AMP) and grown overnight at 37 °C, 180 RPM. From this, 500 μl was used to inoculate 50 ml of terrific broth (100 μg/ml SPEC and AMP) in 250 ml shake flasks. Six t-broth cultures were grown at 37 °C until reaching the required OD_{600nm} (0.4, 0.6 and 0.8), at which point they were transferred to 4 °C and cooled for 20 minutes. Each culture was induced with 0.5 mM IPTG and then one of each OD 0.4, 0.6 and 0.8 culture was grown at 16 °C and 25 °C overnight. Cultures were then transferred to a 50 ml falcon tube and centrifuged at 4000 RPM for 30 minutes at 4 °C. Supernatant was discarded and pellets were resuspended in 10 ml of Buffer A, pH 8.5. Cells were sonicated on ice (20-s on, 40-s off) for 4 cycles. 1 ml of each crude lysate was transferred to an eppendorf tube and centrifuged at 14000 RPM, 4 °C for 10 minutes. 10 μl of each lysate, insoluble and soluble fractions were mixed with 2.5 μl of 4X SDS and incubated at 100C for 5 minutes before being run on an SDS-PAGE gel.

2.13 Western blot analysis of BL21 (DE3) pCDF_PsPA7_6248 small-scale expression fractions

Using a BIO-RAD Mini Trans –Blot Cell, an SDS-PAGE gel from the small-scale expression study of PA7 was placed on a nitrocellulose filter and then sandwiched between two filters and sponges. The cell was then placed in an electrophoresis tank containing 1 L of Transfer Buffer (25 mM Tris, 200 mM Glycine, 10% (vol/vol) Methanol) and run for 1 hour at 20 V, 300 mA. The nitrocellulose membrane was then transferred to a gel tray and washed with a TBS-Tween (20 mM Tris, pH7.5, 150 mM NaCl, 0.05% (vol/vol) Tween) with 5% (w/v) milk powder solution for 30 minutes on a rocker. TBS-Milk solution was poured away and the membrane was washed with 50ml TBS-Tween for 5 minutes, 3 times, discarding the solution in between each. The membrane was then soaked in 50 ml of TBS-Tween solution with 0.1% milk powder and 50 μl (1:1000) of the primary antibody (monoclonal anti-polyHistidine antibody produced in mouse; Sigma-Aldrich). After 30 minutes the solution was discarded and 3 washes of 5 minutes each with TBS-Tween were performed. A 50 ml solution of TBS-Tween with 2 μl (1:20000) of secondary antibody (Anti-Mouse IgG; Sigma-Aldrich) was next added to the membrane and mixed for 30 minutes. Two further 5-minute washes with TBS-Tween were performed, followed by a wash with TBS (no Tween-20). Finally, a 10ml solution of 5-bromo-4-chloro-3-indolyl phosphate (BCIP; 1 tablet in 10 ml of ddH_2O; Sigma-Aldrich) was poured over the membrane. The reaction was left to proceed for 5 minutes, or until bands became visible, at which point the membrane was rinsed thoroughly with ddH_2O.

3. RESULTS

3.1 Overexpression and purification of *Mycobacterium tuberculosis* DprE1

Expression of *Mycobacterium tuberculosis* DprE1 had previously been shown to be improved by co-expression with the chaperone molecules GroEL from *E. coli* and Cpn60.2 from *M. tuberculosis (*Batt et al, 2012). Two plasmid constructs; pCDF-Rv3790 and pTrc-60.2-GroES were transformed in to (heat shock) competent *E. coli* BL21 (DE3) cells. Selected for cells containing the plasmid on Spectinomycin with Ampicillin LB agar plates. Cells were then grown, protein production induced and cultures harvested (see Methods), before affinity chromatography was performed. Fig. 4 shows different stages of the chromatography. An intense band visible in the lane with the flow-through suggests that not all DprE1 was successfully bound to the column.. The wash stage with Buffer A -50 mM imidazole (for buffer composition see page 24) led to slight elution of DprE1 as seen in lane 4, whilst the majority of protein is eluted with the 300mM imidazole buffer. Protein eluted from the Ni-NTA column was dialysed to reduce the salt concentration to ensure binding to the QHP column. This step appeared to result in a slight loss of protein (Fig. 4B, Lane 2). An alternative protocol was used after this, which negated the need for the dialysis step and the associated protein loss incurred with the additional protocol step. The eluate from the HisTrap column was then loaded through the QHP ion exchange column. All DprE1 bound to the QHP column as shown in the difference between lanes 2 and 3 in Fig 4B. Washing the column with 100 mM NaCl (in buffer B framework; Section 2.5) failed to elute any DprE1 (Fig. 4). Increasing the concentration to 120 mM NaCl initiated the elution of DprE1, albeit slowly (Fig. 4). The majority of DprE1 eluted at 140 mM NaCl with the remainder leaving the

column at 500 mM NaCl. Fractions containing DprE1 were combined and then dialyzed into a low NaCl buffer (based in buffer B framework; Section 2.5).

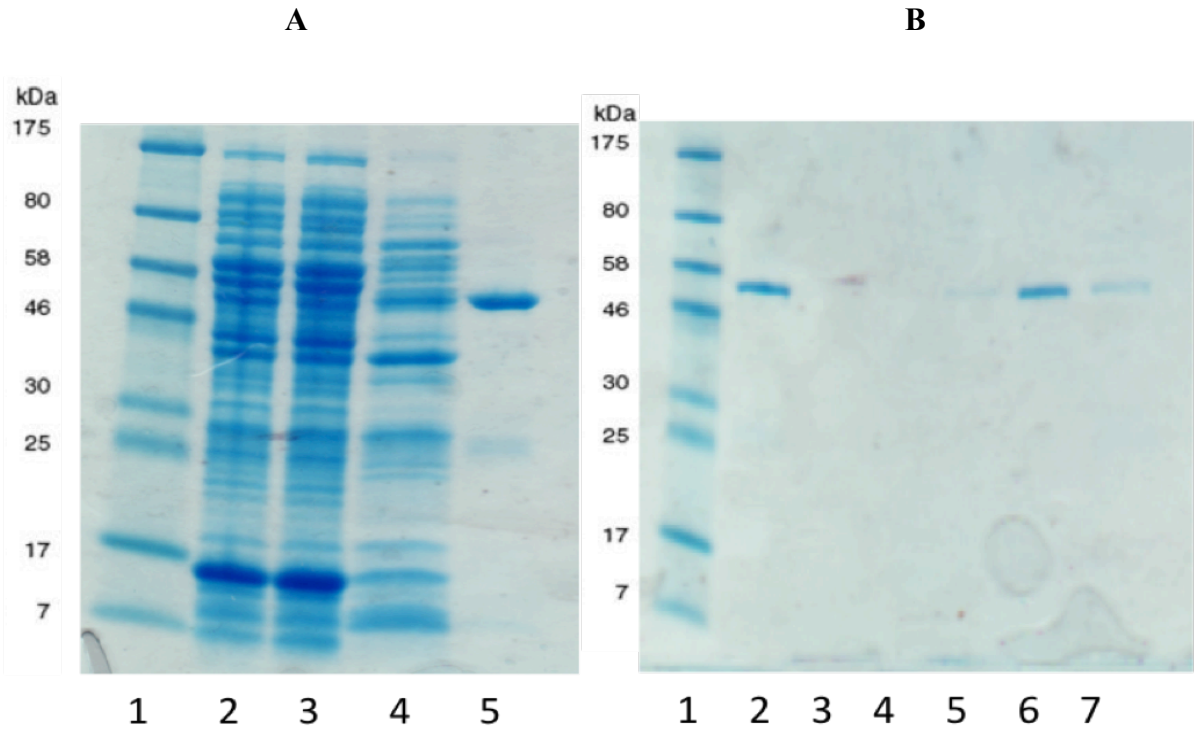


Fig. 4 (A) Ni-NTA chromatography of *M. tuberculosis* DprE1. Lanes from left to right contain mass standard (1), cleared lysate (2), initial flow through (3), 50 mM imidazole buffer eluate (4) and 300 mM imidazole buffer eluate (5). (B) Ion exchange chromatography of *M. tuberculosis* DprE1. Lanes from left to right contain mass standard (1), Ni-NTA eluate (2), initial flow through (3), 100 mM NaCl eluate (4), 120 mM NaCl eluate (5), 140 mM NaCl eluate (6), 500 mM NaCl eluate (7)

Protein was then concentrated by ultrafiltration, with a final concentration of 1.8 mg/ml. Successive purification and concentration attempts utilizing the alternative protocol resulted in a final concentration of ~11.5 mg/ml (Fig. 5) in roughly 0.5 ml of solution.

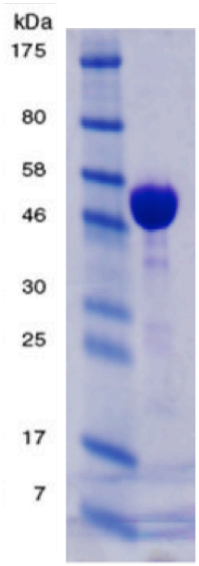


Fig. 5 DprE1 post-concentration. Lanes: mass standard (left), DprE1 post-concentration (right)

3.2 Ligand binding assays

3.2.1 Principles of Fluorescence

Fluorescence is the re-emission of light by a molecule in response to the absorption of a photon. Fluorescent compounds, called fluorophores initially exist in a base state where electrons are in an energy level within the ground state (Fig.6 S_0). When exposed to a photon these fluorophores become excited and move from the ground state to an energy level in the excited electronic state (Fig. 6 step 1). Over a period of time, known as the fluorescence lifetime, vibrational relaxation causes the fluorophore to lose energy in the form of heat to its surroundings or kinetic energy/vibration (Fig. 6 step 2, S_1' to S_1). Finally, transition to the ground state causes the release of energy in the form of a photon (Fig 6. step 3). Due to the loss of energy in step 2 the wavelength of the emitted photon ($h\nu_{EM}$) is longer than that of the incident photon that caused the initial excitation ($h\nu_{EX}$). The relationship between energy and wavelength is described by the equation: $E = hc/\lambda$, where E = energy, h = Planck's constant ($6.6260755 \times 10^{-34}$ Joule second), c = the speed of light (2.998×10^8 m/s) and λ = wavelength of the light. This shows that the energy of a photon is inversely proportional to its wavelength.

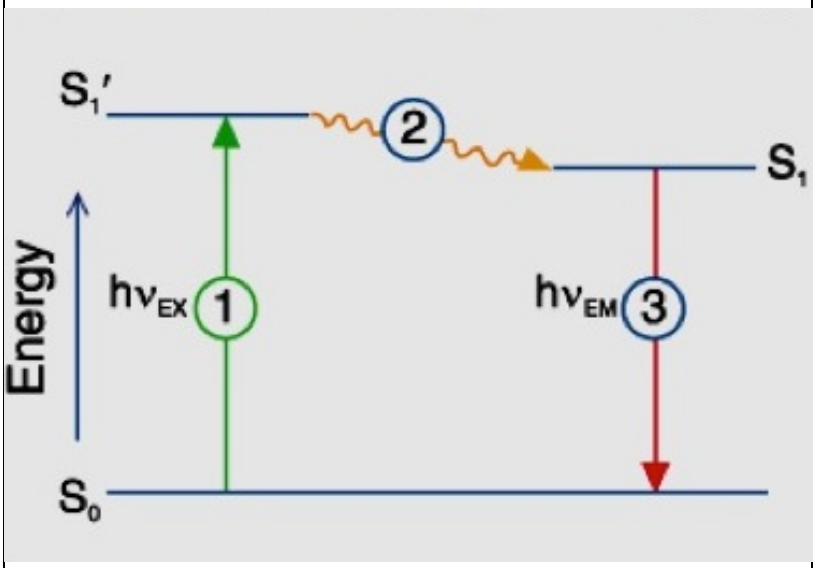


Fig.6 Jablonski diagram showing the three basic steps of electron movement during the process of fluorescence. (1) Photons ($h\nu_{EX}$) cause the excitation of the fluorophore to a higher energy level (S_1') from the ground state (S_0). (2) Vibrational relaxation causes the loss of a small amount of energy and the fluorophore moves to a lower energy level yet still remains within the excited state (S_1). (3) The fluorophore returns to the ground state releasing a photon of a different wavelength/energy ($h\nu_{EM}$) to that involved in step 1 ($h\nu_{EX}$). h = Planck's constant ($6.6260755 \times 10^{-34}$ Joule second), ν = frequency. (Johnson & Spence, 2010)

The photons involved in fluorescence can be detected using a machine known as a fluorimeter, which also measures their relative intensities. The concentration of a fluorophore is directly proportional to fluorescence intensity. Proteins can exhibit intrinsic fluorescence due to the presence of the aromatic amino acids; tryptophan, tyrosine and phenylalanine. Tryptophan fluorescence is the most commonly measured due to its sensitivity to its environment and high quantum yield (Kraft *et al.*, 2009; Φ = #photons emitted/#absorbed). The intensity of fluorescence can therefore be dependent on changes in fluorophore environment such as protein folding, conformational changes or ligand binding. This makes fluorescence spectroscopy a good tool for the study of protein interactions.

AstraZeneca provided two inhibitors of DprE1, named 1326 and 1328. The company did not disclose the chemical structures of these two compounds. To determine the binding affinity of these inhibitors fluorescence spectroscopy was performed.

Determining the maximum excitation and emission wavelengths of DprE1 and ligands

Initial scans were performed on a reference sample of buffer B. The results of an excitation scan of the reference sample revealed clear peaks with high intensities, indicative of contaminants. The glycerol was removed from the buffer and the aforementioned peaks disappeared, indicating the glycerol was responsible for the prior abnormal readings. Fig. 7 shows the excitation scan of DprE1 revealing a max excitation point of ~280 nm. Emission scans of a solution of buffer B (no glycerol; section 2.5) containing 1 µM DprE1 determined a max intensity at the wavelength 328 nm (Fig.8). An emission scan was also performed on a solution of buffer B and 1328 revealing a broad peak of intensity, overlapping somewhat with the emission peak of DprE1 (Fig.8).

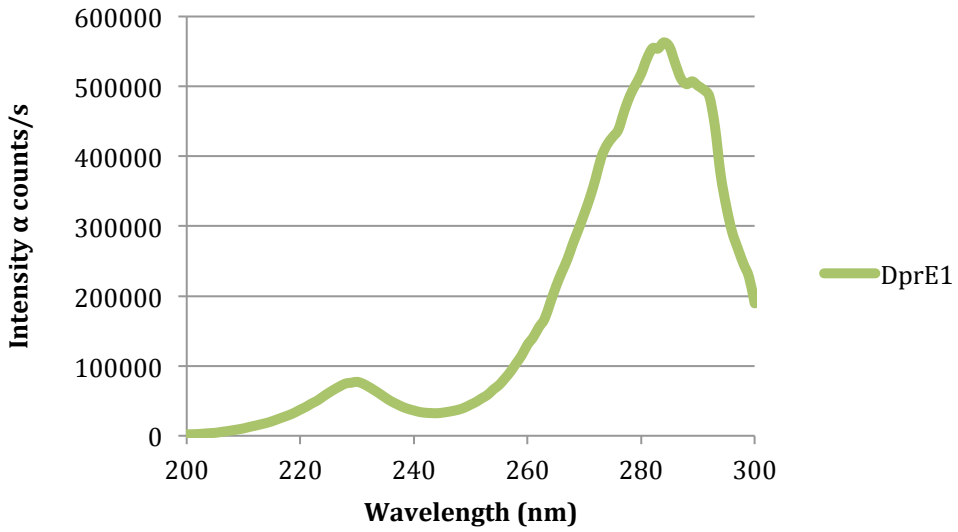


Fig. 7 Excitation scan of 1 µM DprE1 from 200-300 nm. Max intensity was determined to be at a wavelength of ~280 nm. Slit width = 0.6 mm

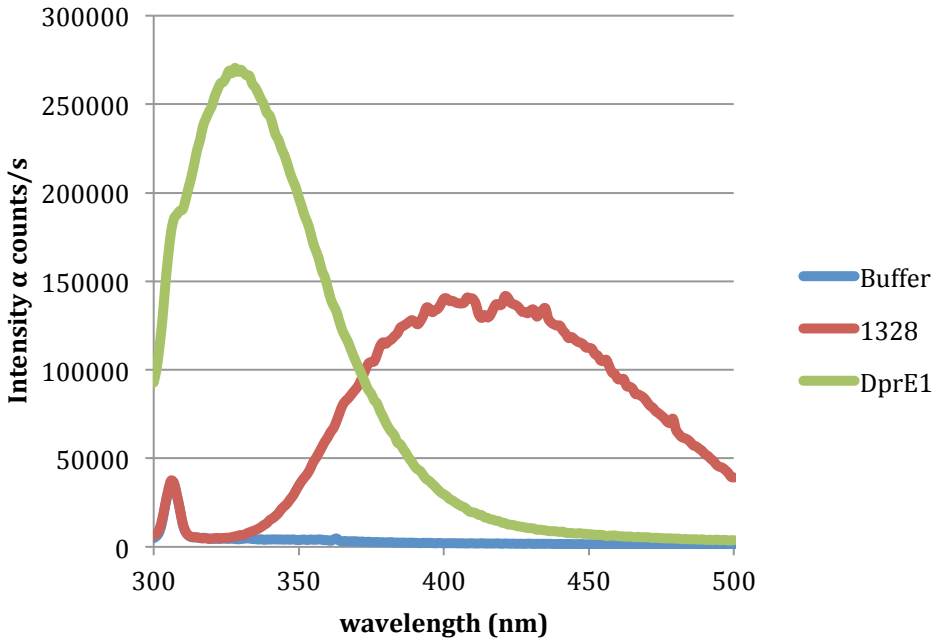


Fig. 8 Emission spectra (300-400 nm) of DprE1 (1 µM; green) and 1328 (8 µM; red) with a fixed excitation wavelength (280 nm). The maximum emission wavelength of DprE1 was found to be ~328 nm. 1328 emitted maximally at a wavelength of >415 nm. The overlapping of these two spectra meant that excitation at 280 nm would be unsuitable to ascertain a precise value of intensity if both DprE1 and 1328 were present in the same solution. Slit width = 0.6 mm

The excitation wavelength was decreased to 270 nm in an attempt to try and separate the ligand and protein peaks, but as the titration in Fig. 9 shows there was still a fair amount of interference between the two, making it hard to discern the intensity of emission at 425 nm particularly in samples with low concentrations of 1328. However, the increase in emission intensity at 425 nm was linear (Fig. 10).

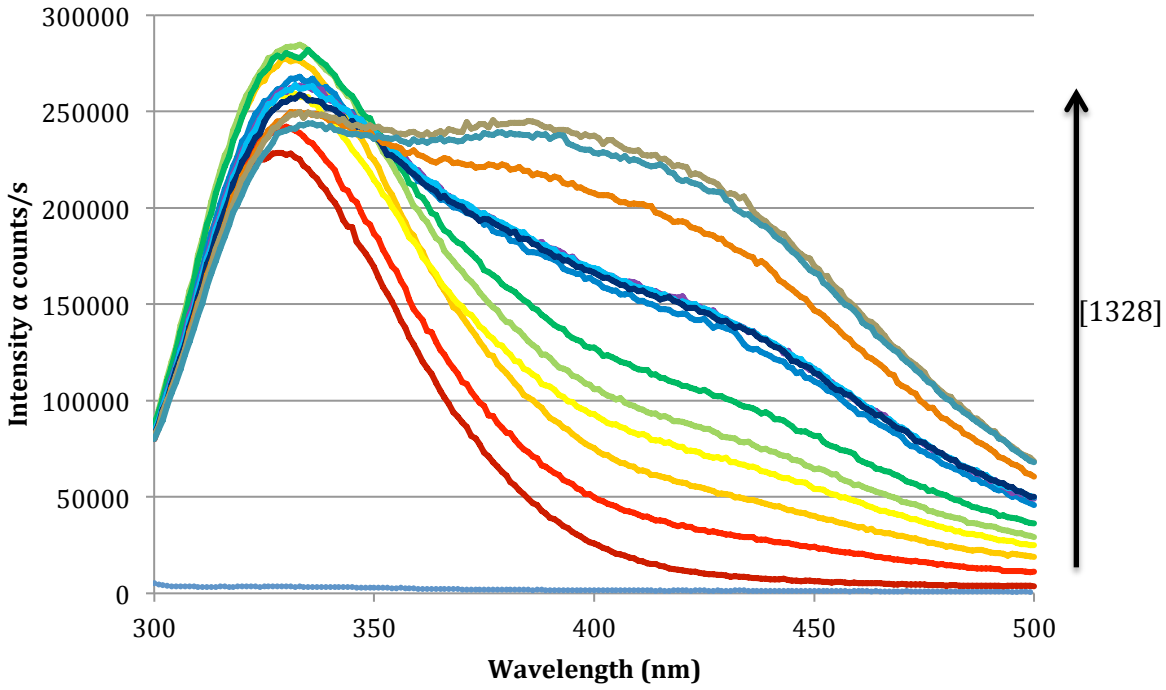


Fig. 9 Emission spectra (300-500 nm) obtained upon titration of 1328 into DprE1 with an excitation wavelength of 270 nm. At this wavelength the emission peaks of both enzyme and inhibitor still influence one another. DprE1 shows a peak emission at ~330 nm and 1328 at ~425 nm. Slit width = 0.6 mm

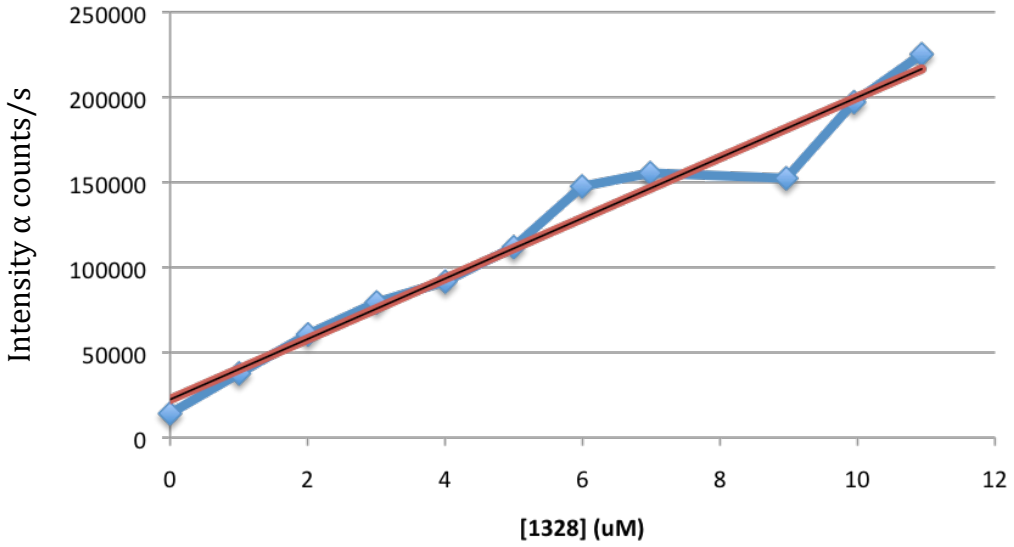


Fig. 10 The increase in intensity of peak emission wavelength observed at 425 nm with increasing concentration of 1328. Excitation wavelength = 270 nm.

In an attempt to separate the two intensity peaks at 425 nm and 328 nm new emission scans were performed for DprE1, this time exciting at a wavelength characteristic of the flavin co-factor, FAD. The excitation wavelength maximum for the flavin was determined to be 448 nm with peak emission intensity observed at 525 nm. Following on from this, an emission scan of the reference sample, buffer B excited at 448 nm revealed an unknown peak with a maximum intensity at 530 nm. Varying the excitation wavelength from 450 nm to 460 nm shifted this peak by a value proportional to the change in wavelength (Fig. 11), suggestive of the Raman effect. DprE1 (1 μM) was added and the excitation wavelength was again varied from 450 to 460 nm (Fig. 10). By increasing the excitation wavelength to 460 nm the two peaks were separated to an extent where the 520 nm peak was definable and the intensity value could be obtained (Fig.12).

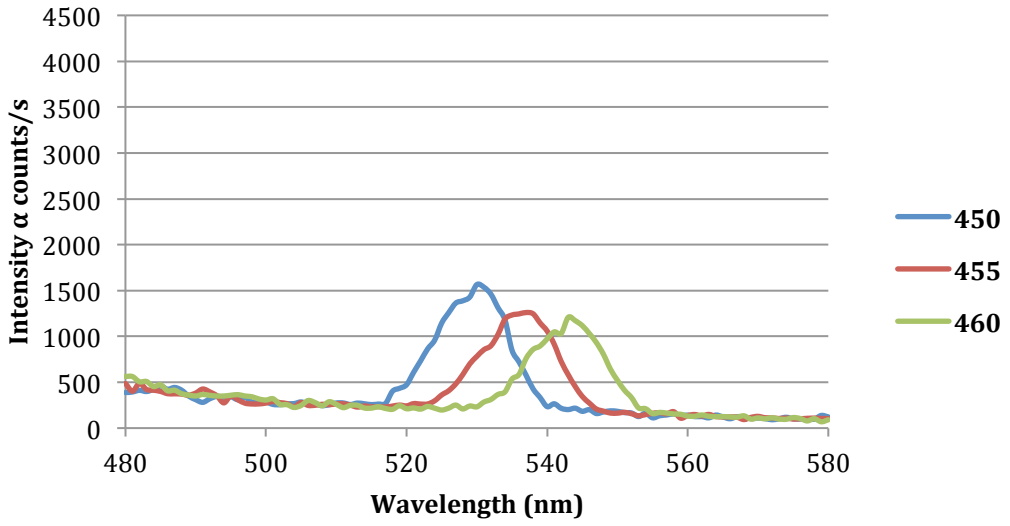


Fig. 11 The effect of varying excitation wavelength on the Raman peak observed in the emission spectra (480-580 nm) of a reference sample (Buffer B; Section 2.6) excluding ligand or protein. Excitation wavelength was found to be proportional to the emission wavelength of the peak observed. Key refers to the excitation wavelength. Slit width = 0.6 mm

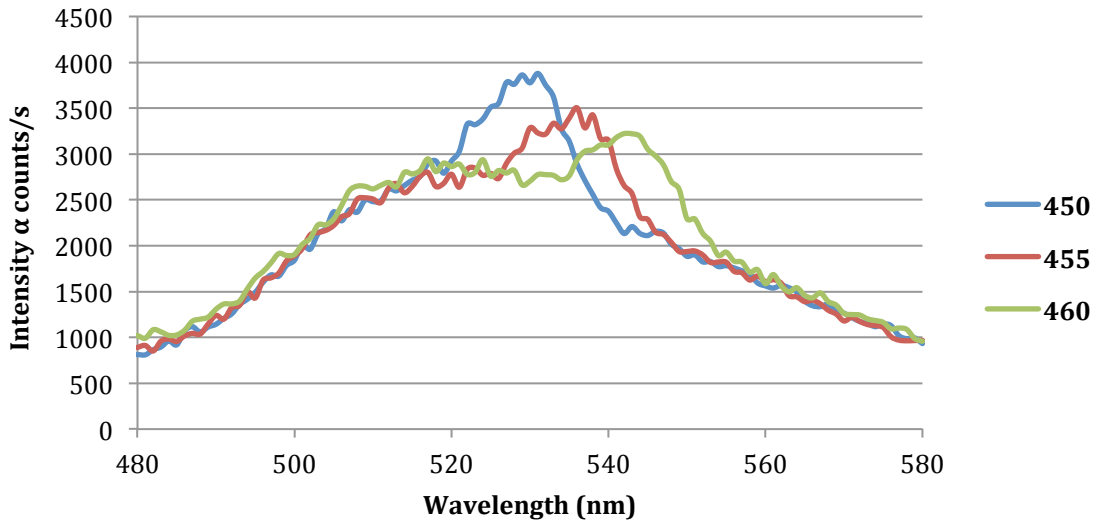


Fig. 12 The effect of varying excitation wavelength from 450 nm to 460 nm on the Raman peak observed in the emission spectra (480-580 nm) of 300 µl Buffer B + 1 µM DprE1. The shift in Raman peak emission wavelength maximum is proportional to the change in excitation wavelength. At an excitation wavelength of 460 nm a separation between Raman peak and FAD emission peak at ~520 nm is observed. Key refers to excitation wavelength. Slit width = 0.6 mm

Titrations of 1328 were performed at 460 nm. Fig. 13 shows the binding curve of a titration of 1328 into 2 µM DprE1 in increments of 2.5 µM. No saturation appears even at a ligand to protein ratio of 30:1. By this point it was expected that no further increase in fluorescence would be observed. Saturation occurs when all protein is bound to the ligand. The linear nature of the graph was suspected to be as a result of oxidation of flavin, which displays different fluorescence characteristics than reduced flavin (Ghisla, 1980). To test this, 2 µM DprE1 was excited at a wavelength of 460 nm at time intervals over a 4 hour periodcourse was measured. The intensity of the peak maximum increased linearly over a 4 hour time period, showing no sign of saturation (Fig. 14). In order to prevent this and continue working at this wavelength we would have to perform the experiments in a closed system, deprived of

oxygen. The limitations of the equipment available prevented this approach. Instead, it was proposed that rather than titrating ligand into protein, protein could be titrated into ligand, and therefore avoiding the high initial concentration of DprE1 that would be subject to oxidation.

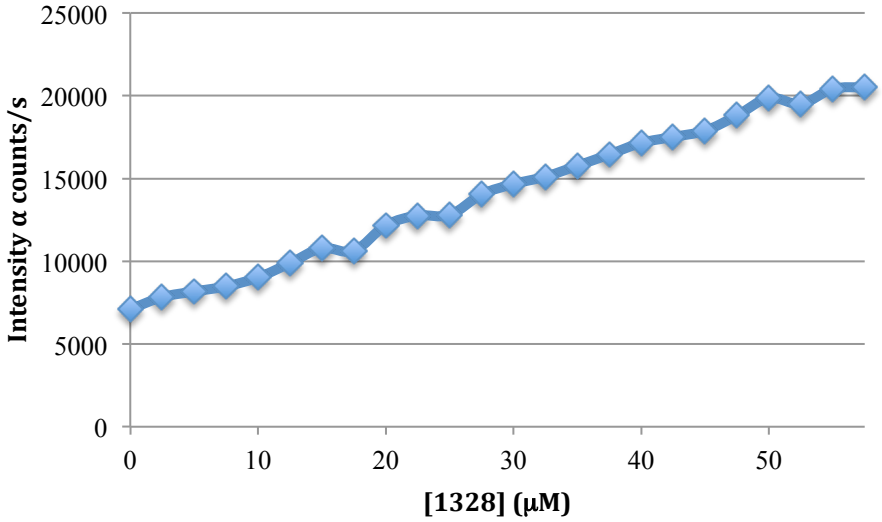


Fig. 13 Titration of 1328 from 0-55 μM in to an initial concentration of 2 μM of DprE1. Values correspond to the emission peak maxima recorded at the wavelength of 520 nm. Slit width = 0.6 mm

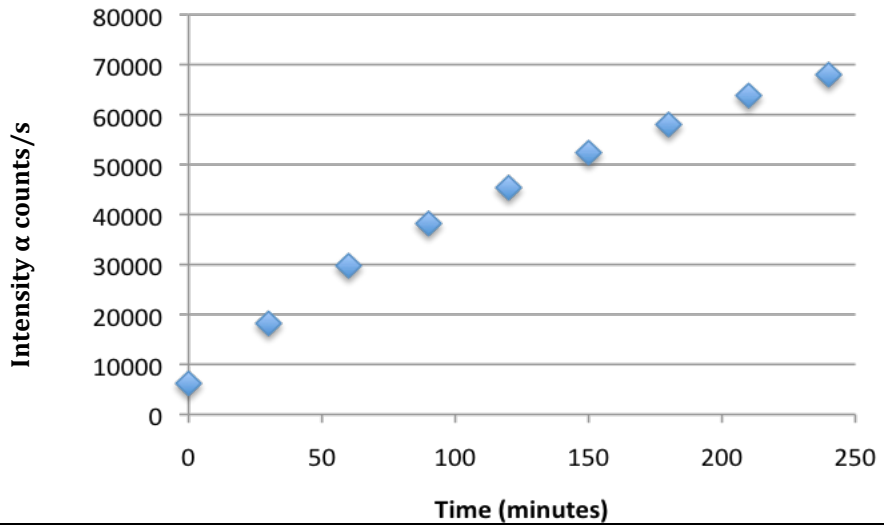


Fig. 14 The change of intensity observed in the FAD emission peak at wavelength 520 nm in a sample of DprE1 over time. A linear increase in intensity is observed proportional to an increase in time. This was most likely due to a change in the oxidation state of DprE1.

Previous emission scans of 1328 revealed an emission maximum at a wavelength of approximately 415 nm (as shown in Fig. 8) when excited at 280 nm. The excitation wavelength however, was increased to 297 nm to overcome the influence of the emission peak at 330 nm (from tryptophan residues). 1328 (10 μM) was established and DprE1 was titrated in to the sample in 0.2 μM increments. However, by the time 0.6 μM DprE1 had been added intensity values were over 600,000 counts, over the threshold for efficient detection by the fluorimeter. To circumvent this, the slit width of the fluorimeter monochromators was decreased from 0.6 mm to 0.4 mm, to allow less light to pass and therefore decrease the intensity of the detected fluorescence. Further readings were taken and intensity values well under 600,000 were recorded (Fig. 15).

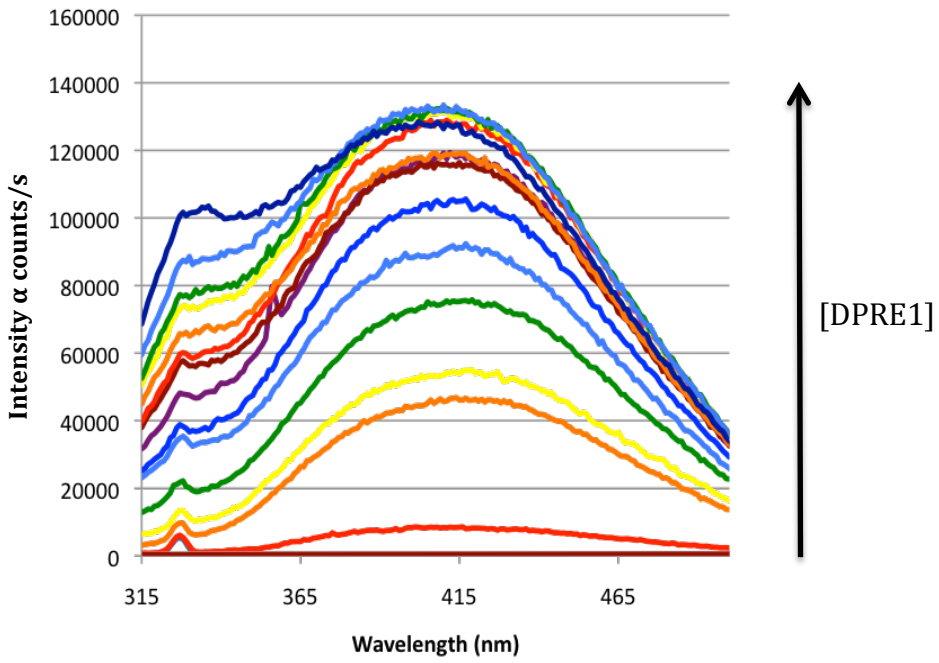


Fig. 15 Emission spectra (315-500 nm) of DprE1 (0-4.25 μM) titrated into initial 10 μM of 1328. Excitation wavelength for this experiment was 297 nm. The red line at the bottom of the graph is representative of the baseline 10 μM 1328. An emission wavelength peak is observed at a wavelength of ~415 nm. Slit width = 0.4 mm.

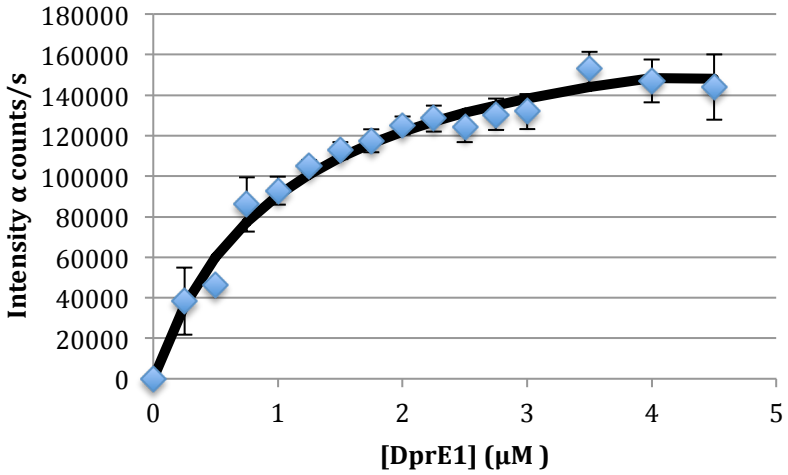


Fig. 16 Titration of DprE1 in to 10 μM 1328. By ~4 μM DprE1 peak emission signal intensity begins to saturate. Data has been corrected to account for the change in [1328] with increase in total sample volume. Mean K_d = 0.638 (±0.06) μM and a B_{max} = 154,581 (±10,316)

The interaction displayed single site saturation binding behavior (Fig. 16); with a calculated mean K_d = 0.638 (±0.06) μM and a B_{max} = 154,581 (±10,316). Titrations were repeated with 1326 under the same parameters. 1326 displayed a different emission spectra, with maximum intensity observed at a wavelength of ~365 nm (Fig. 17).

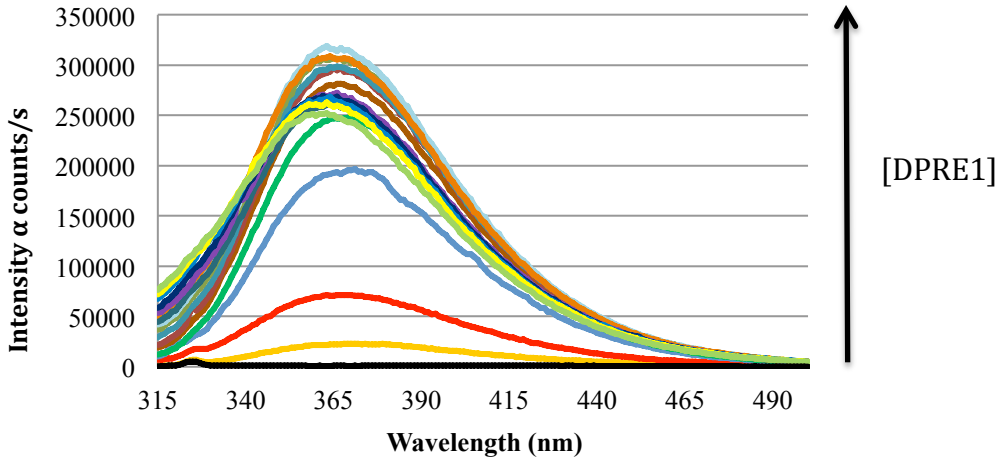


Fig. 17 Emission spectra (315-500 nm) of DprE1 (0-4.25 μM) titrated in to 10 μM 1326. Excitation wavelength was fixed at 297 nm. The yellow line at the bottom of the graph represents the emission spectra of 1326 maximum intensity at a wavelength of ~365 nm.

Titration of DprE1 in to 1326 (Fig. 17) displayed a similar single site binding characteristic as observed in titrations with 1328. The experiments were performed in triplicate and mean K_d of 1.24 (±0.479) µM and a B_{max} of 459763 (±61619) were calculated (Fig.18).

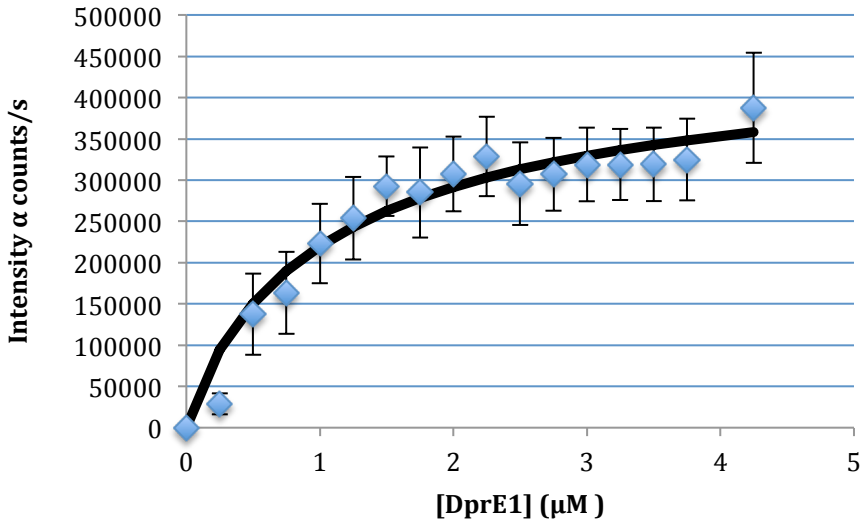


Fig. 18 Titration of DprE1 in to a concentration of 10 µM of 1326. Fluorescence intensity was recorded over the wavelength range 315-500 nm with peak intensity varying wavelength between 360 and 367 nm (see Fig. 17). Excitation wavelength was fixed at 297 nm. Mean K_d = 1.24 (±0.479) µM, B_{max} = 459763 (±61619), across 3 replicates.

Interestingly, the emission peak wavelength of the spectra shifted from 360 nm to 367 nm as concentration of DprE1 was decreased (Fig. 19).

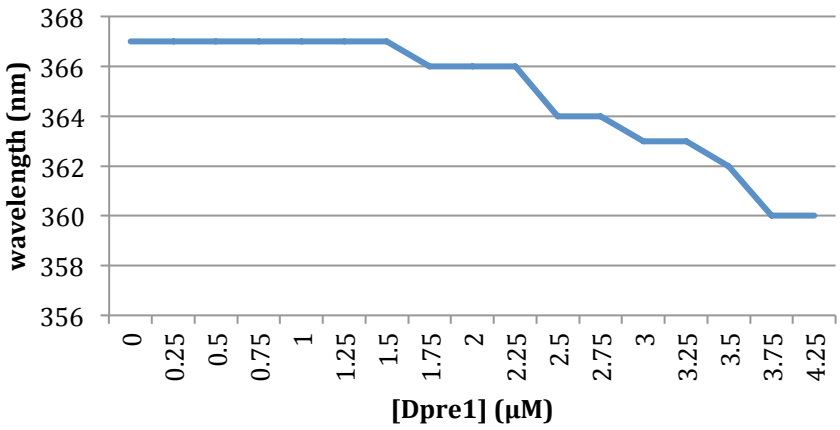


Fig. 19 Shift of emission peak wavelength observed with increasing concentration of DprE1.

Negative control titrations of bovine serum albumin (BSA) in to 1326 & 1328

A negative control for both 1326 and 1328 was performed using bovine serum albumin (BSA). Unlike DprE1, BSA was predicted to not display any interaction with the two compounds and therefore no increase in emission maxima at either 415 nm (for 1328) or 365 nm (for 1326). Fig 20 shows the result of titrating BSA in to 1328 (top) and 1326 (bottom).

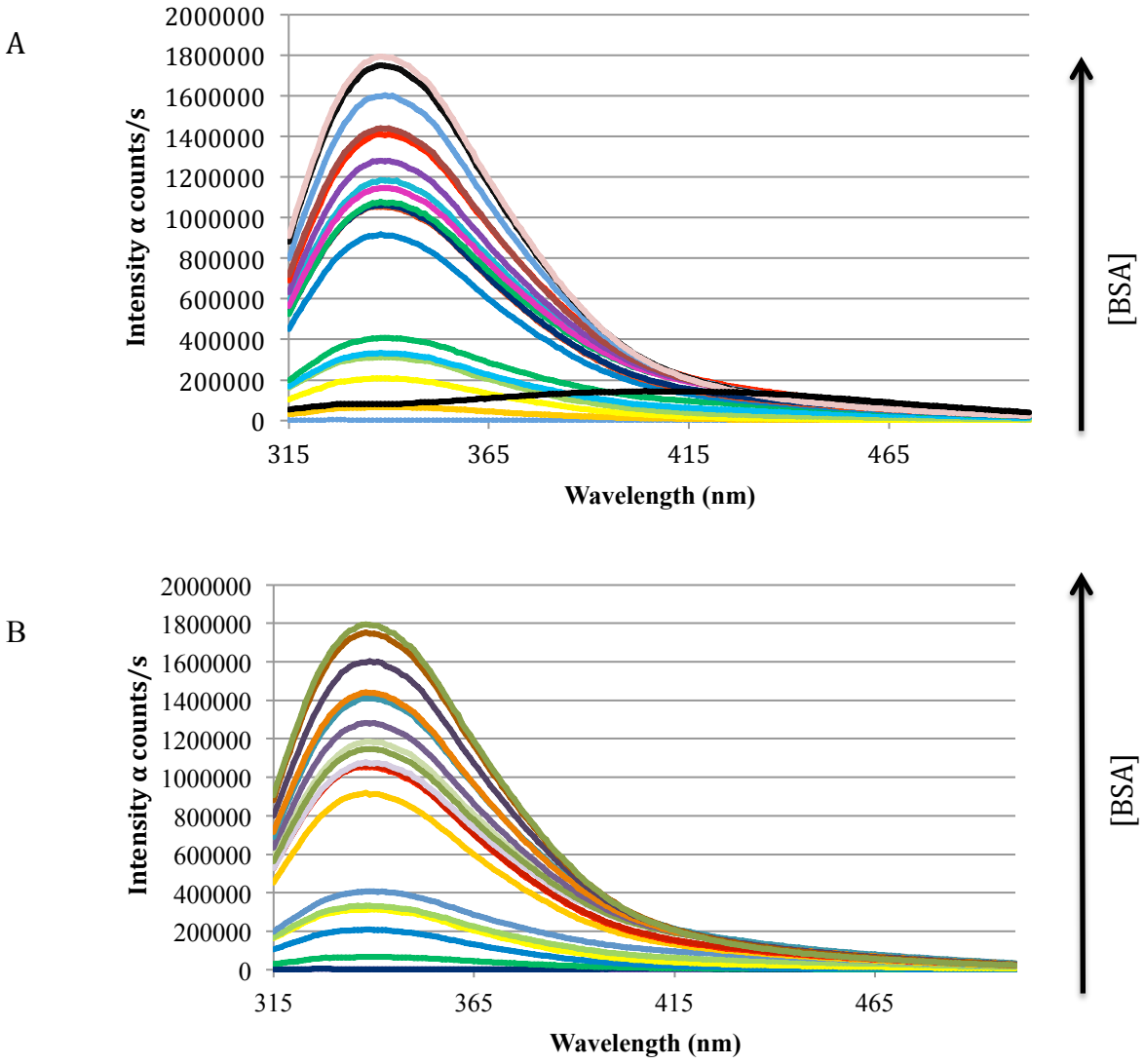


Fig. 20 The emission spectra of 0.25-5 µM bovine serum albumin (BSA) titrated in to 10 µM 1326. Excitation wavelength = 297 nm. An emission peak is observed at ~330 nm, representative of the fluorescence of tryptophan residues in BSA. Slit width = 0.4 mm

As Fig. 20 shows, with increasing concentrations of BSA an increase in fluorescence was observed at ~330 nm, but not at 415 nm (for 1328) or 365 nm (for 1326) as seen in figures 15 and 17, respectively. This increase in fluorescence at 330 nm is indicative of tryptophan residues found in BSA and is proportional to the overall increase in concentration of BSA in the solution.

Negative control titration of mutated DprE1 ΔY314C into 1326 & 1328

A further set of control titrations were performed substituting DprE1 for a mutated form of the protein in which the amino acid tyrosine at position 314 had been mutated to cysteine. This protein is found in *M. tuberculosis* mutants resistant to the inhibitors TCA1 and those of the azaindole class and is therefore highly implicated in DprE1 inhibitor resistance (Batt *et al.,* 2012). The titration experiments were repeated using DprE1 Y314C with 1326 and 1328 and showed similar single site binding saturation properties to the unmodified protein (Fig. 21). This suggested that 1326 and 1328 bind to DprE1 by a different mechanism than TCA1 and the azaindoles.

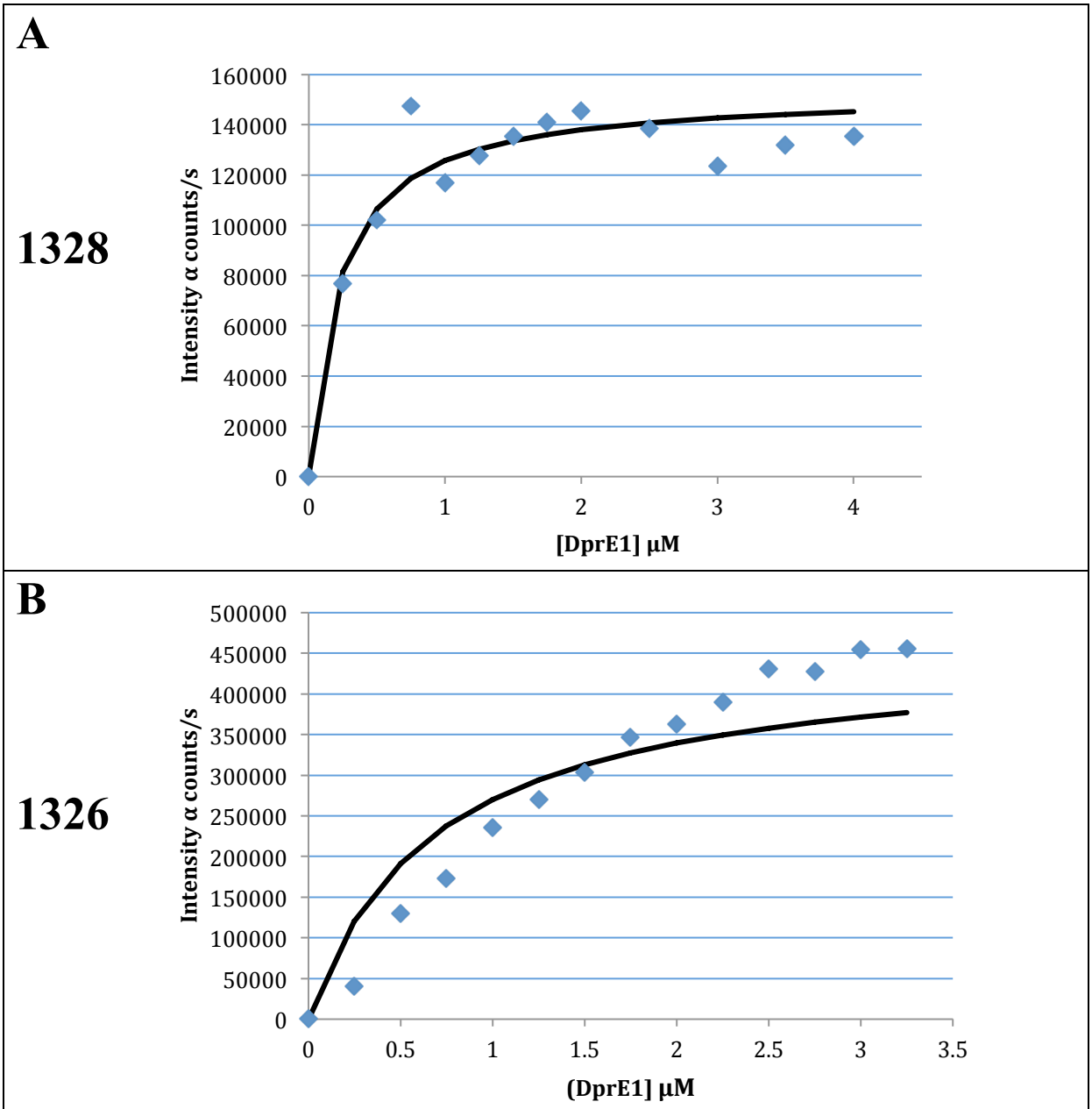


Fig. 21 (A) Titration of DprE1 Y314C (0 – 3.25 µM) in to 10 µM 1328. An excitation wavelength of 297 nm was used. Emission peak wavelength was recorded at ~420 nm. The experiment was performed in triplicate the and mean K_d = 0.21 µM, B_{max} = 152,958, (B) Saturation binding curve of 1326 with a titration of DprE1 Y314C. An excitation wavelength of 295 nm was used. Emission peak wavelength was recorded at a range between 360 and 367 nm. Mean values for K_d = 0.69 µM and B_{max} = 457,074 were established after the experiment was performed in triplicate (no error estimates available).

Although fluorescence spectroscopy is a useful tool in providing information as to whether protein and ligand interact it isn't always accurate and doesn't provide information on the exact mechanism of binding. If a conformational change in the protein is associated with ligand biding this can lead to the introduction of a quenching group in to the vicinity of the fluorophore and directly influence the intensity detected (Royer & Scarlata, 2008). This in turn would affect the calculated values for K_d and $B_{max.}$ In order to determine exactly if and how protein ligand interactions occur a structural technique such as x-ray crystallography could be employed.

3.3 Generation of DNA plasmid expression constructs encoding recombinant *Pseudomonas aeruginosa* and *Mycobacterium smegmatis* DprE1 homologues

Previous efforts to crystallise *M. tuberculosis* DprE1 with 1328 were successful. However, upon subjecting the crystals to x-ray scattering, the ligand was not detectable in the electron density map. In order to try and determine the binding mechanism of these inhibitors with DprE1 two homologues of the protein were chosen. From *Mycobacterium smegmatis*, MSMEG_6382, which shares 84.4% sequence identity to Rv3790, and from *Pseudomonas aeruginosa PA7*, PsPA7_6248, which shares 33% sequence identity to Rv3790. Oligonucleotide primers (obtained from Eurofins MWG Operon) were designed for the PCR amplification of the two homologues from gDNA isolated from the two species. Each primer was designed to include a different restriction endonuclease site to flank the amplified gene. These restriction enzymes were selected based on their presence in the multiple cloning site (MCS) of the plasmid pCDF. *Sac*I and *Hind*III were chosen for PA7 and *Bam*HI and *Not*I for *M. smegmatis*. PCR reactions were performed on gDNA from both species with two variable annealing temperatures, 60 °C and 65 °C to determine which was most effective for

amplification. For *M. smegmatis* gene amplification was observed at 60 °C, but not at 65 °C, whereas neither temperature resulted in amplification of the PA7 gene (Fig. 22).

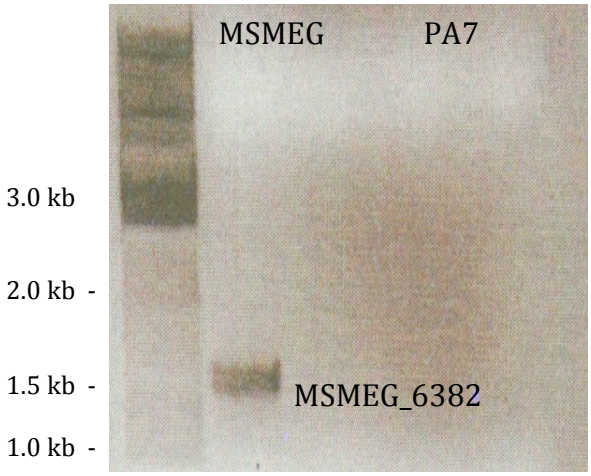


Fig. 22 Agarose gel (1%) showing the products of PCR amplification of MSMEG_6382 and PsPA7_6248**.** Lanes from left to right contain mass standard, *M. smegmatis* 60 °C, *M. smegmatis* 65 °C, PA7 60 °C, PA7 65 °C. (Given temperatures refer to annealing temperature of reaction used). Expected size of product band = ~1.5 kb

The PCR reaction was repeated with PA7 gDNA using a gradient thermal cycler (Touchgene Gradient; Techne Limited) varying the annealing temperature of the reaction from 50 °C to 70 °C across 12 steps. Product was formed at an annealing temperature of 63 °C (Fig. 23).

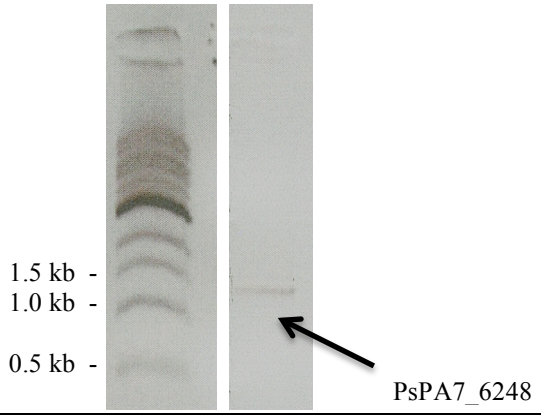


Fig. 23 Agarose gel analysis of PsPA7_6248 amplification from PA7 gDNA using an annealing temperature of 63 °C.

PCR products were then subjected to treatment with restriction enzymes (*Sac*I and *Hind*III for PsPA7_6248 and *Bam*HI and *Not*I for MSMEG_6382). The plasmid vector, pCDF was also digested with each pair of restriction enzymes. Digested DNA samples were run on agarose gel and viewed using a near-UV viewer to confirm digestion. Bands of interest were identified and then excised using a sterile scalpel. DNA was purified from the gel slices before the sample of vector DNA was treated with antarctic phosphatase to remove 5' phosphates, preventing self-ligation. Vector and insert DNA were mixed in a ratio of 3:1 and ligated.

A control sample with only vector DNA was produced using the same protocol. All samples were transformed into *E. coli* TOP10 cells and selected overnight on LB agar supplemented with Ampicillin. Transformation efficiency of PA7 constructs was low, with only 2 colonies appearing on the plate after 16 hours incubation time. Colonies were also present on the PA7 control plate suggestive of vector circularization. Colonies from both PA7 and *M. smegmatis* construct plates were cultured and plasmid DNA was purified. A second restriction digest was performed on the purified DNA to confirm the presence of both insert and vector DNA. Fig. 24 shows the result of this second digestion. Four out of 7 *M. smegmatis* (Lane 1-4) colonies digested to reveal vector and insert bands. Neither of the PA7 samples showed a band representative of insert DNA (Lanes 8 and 9).

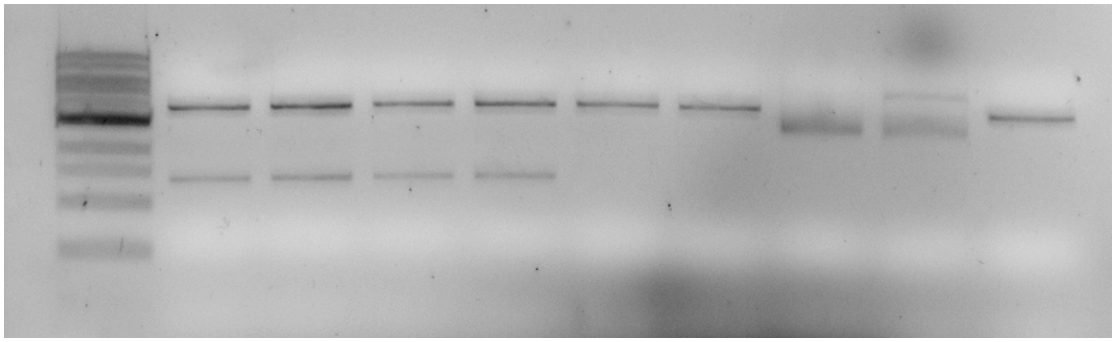


Fig. 24 Agarose gel analysis of second restriction digest to identify successful ligation of vector and insert DNA. Lanes from left to right: mass standard (M), *M. smegmatis* constructs (1-7) and PA7 constructs (8 and 9). Successfully ligated vector and insert constructs are shown in lanes 1-4.

The multiple band pattern displayed in lane 8 (Fig. 24) is typical of circular plasmid DNA. The other PA7 construct sample (lane 9) showed a single band for vector DNA but no band for insert DNA. Circular plasmid DNA and a single vector band suggested that full digestion hadn't been achieved. Homologous recombination of the plasmid DNA could explain these results by the removal of one (lane 9) or both (lane 8) restriction sites. Construction of the PA7 construct was repeated using two new samples of digested vector DNA but yielded similarly disappointing results. Repeating the amplification of DprE1 and increasing the insert to vector ratio to 9:1 also gave no result. All vector DNA previously used had been purified and digested from one sample of TOP10 cells transformed with pCDF. The issue of recombination would have likely arisen at this point, subsequently affected all following results. The protocol was repeated by a colleague (Dr. Sarah Batt), who was able to achieve the correct construct displaying clear vector and insert bands upon a second digestion with restriction enzymes. No explanation can be given for the previous failed attempts.

Both plasmids were sequenced against the PCR primers used to originally amplify the DprE1 genes. Both showed base substitutions toward the 3' end of sequence data. However, sequencing of both DNA strands using the forward and reverse primers revealed these to be artifacts of cloning. All *M. smeg* sequences revealed either one or two nucleotide deletions, 1391delC or 1386delC and 1402delC (Fig. 25).

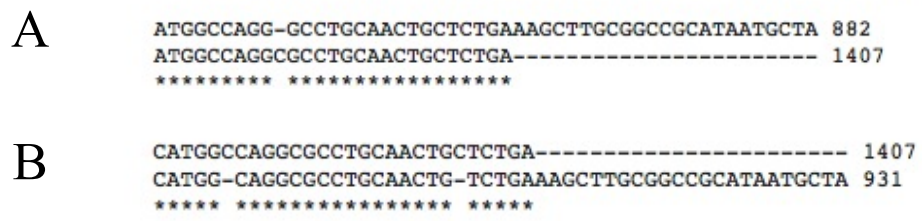


Fig. 25 Sequence alignment between sequenced *M. smegmatis* constructs and the wild-type MSMEG_6382 DNA sequence. (A) 1391delC, (B) 1386delC, 1402delC. In both alignments the stop codon of the wild type MSMEG_6382 gene is the TGA that ends at position 1407.

Primers were designed to reintroduce the deleted Cytosine base at position 1391 by site directed mutagenesis. Once carried out, mutated constructs were transformed into TOP10 cells, grown and purified. Sequencing confirmed successful reintroduction of the deleted base.

3.4 Overexpression and purification of PA7 and *M. smegmatis* DprE1 homologues from constructs

Initially, both constructs were transformed and grown in *E. coli* BL21 (DE3) cells using the same protocol as for *M. tuberculosis* DprE1 overexpression. The theoretical isoelectric point (pI) of both proteins was calculated using ProtParam (Gasteiger et al, 2005; http://web.expasy.org/protparam/). The pI of M. smegmatis DprE1 was predicted as 6.47 and the PA7 DprE1 pI as 8.19. Buffers used for the purification of *M. tuberculosis* DprE1 had been pH 8. Although suitable for the purification of the *M. smegmatis* protein, the pH of the buffers used to purify PA7 DprE1 was increased to 9.25 to avoid precipitation of the protein.

M. smegmatis DprE1:

Due to the high amino acid sequence similarity between *M. tuberculosis* and *M. smegmatis* DprE1 homologues it was decided that the protocol adapted for the purification of *M. tuberculosis* DprE1 would be suitable for the *M. smegmatis* homologue. *M. smegmatis* DprE1 behaved very similarly to *M. tuberculosis* DprE1, eluting at a concentration of 300 mM imidazole. However, a small amount of protein was also found in the 50 mM imidazole wash buffer fraction. Once loaded on to the QHP column, DprE1 was found initially in the 120 mM NaCl buffer fraction, and also in the fractions from the 140 mM and 500 mM NaCl buffer fractions. The protein was concentrated to ~ 6.7 μg/ml but lacked the intense yellow colour associated with the oxidised flavin cofactor, which is typical of the *M. tuberculosis* protein.

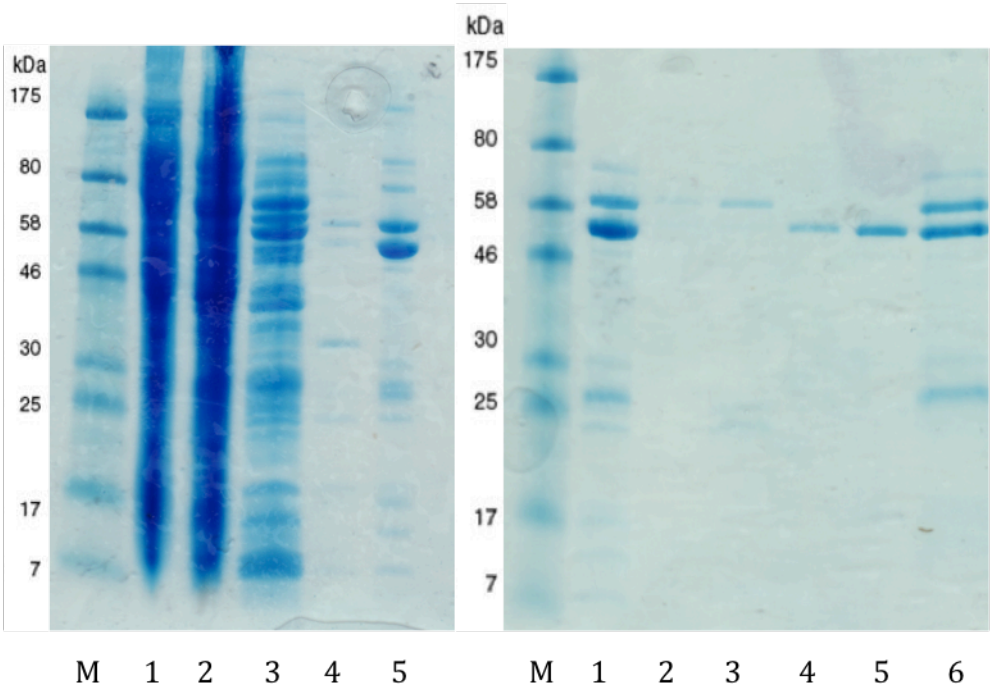


Fig. 25 SDS-PAGE gel of samples from the Ni-NTA (left) and QHP (right) chromatography of *M. smegmatis* DprE1 preperation. (Left) Lanes contain, from left to right; mass standard (M), cleared lysate (1), initial flow through (2), 50 mM imidazole buffer flow through (3), buffer B flow through (4), 300 mM imidazole buffer (low salt) flow through (5). (Right) Lanes contain, from left to right; Mass standard (M), Ni-NTA eluate (1), initial flow through (2), 100 mM NaCl buffer flow through (3), 120 mM NaCl buffer flow through (4), 140 mM NaCl buffer flow through (5), 500 mM NaCl buffer flow through (6).

PA7 DprE1:

BL21 (DE3) cells were transformed with the pCDF_psPA7_6248 plasmid and grown. Once induced, grown and harvested cells were lysed. Lysate was cleared and then loaded on to the Ni-NTA column. At this stage *M. tuberculosis* lysate appears yellow in colour but the lysate from the pCDF_psPA7_6248 cells was colourless. After passing the lysate through the column, buffers containing 50 mM and 300 mM imidazole were subsequently passed through. Analysis of the eluted fractions by SDS-PAGE indicated that all protein was eluted in the 50

mM imidazole buffer wash stage (Fig. 26, left, lane 3). After dialysis, the Ni-NTA eluate was loaded through the QHP column. Bands suspected to be DprE1 were observed initial flow through and thought to also be in the fractions of 120 mM and 140 mM NaCl buffers (Fig. 26, right, lanes 4 & 5). The eluate from 120 and 140 mM NaCl fractions were dialysed to a low-salt buffer before the protein was concentrated to ~ 6 μg/ml. Unlike the *M. smegmatis* protein purified, the PA7 protein displayed the typical yellow colour of DprE1. The purified protein was then aliquoted in to 96-well trays with the hope of forming protein crystals. However, after 2 weeks, no significant results were seen.

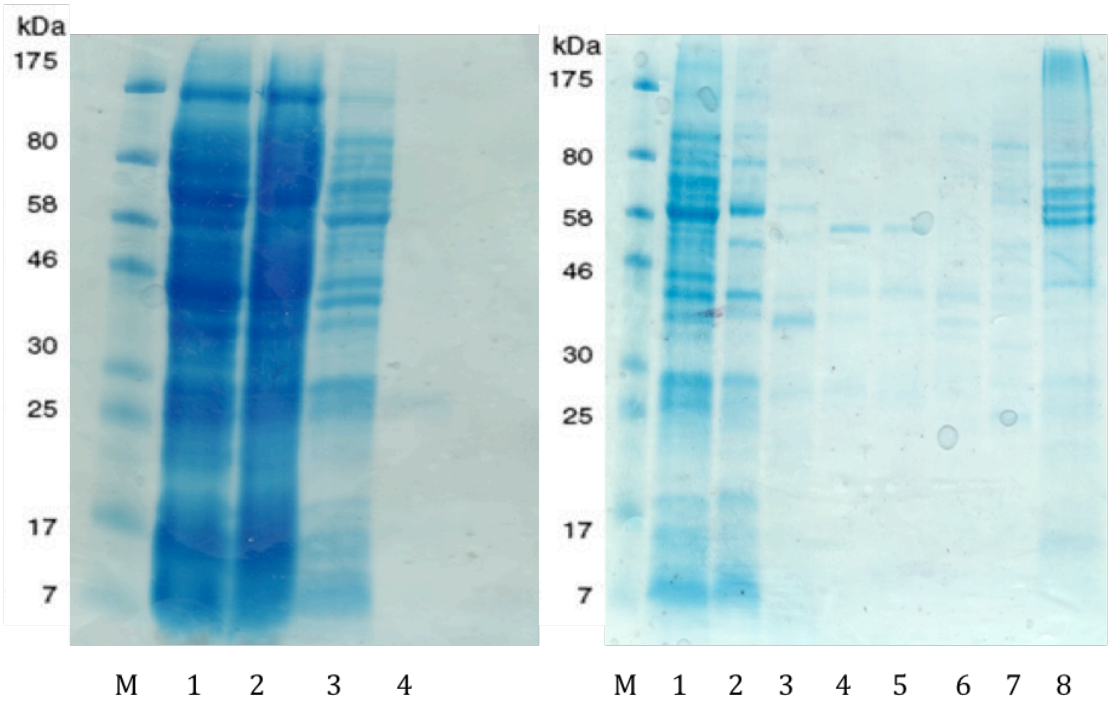


Fig. 26 SDS-PAGE gel of samples from the Ni-NTA (left) and QHP (right) chromatography of PA7 DprE1. (Left) Lanes contain, from left to right; mass standard (M), cleared lysate (1), initial flow through (2), 50 mM imidazole buffer flow through (3), 300 mM imidazole buffer flow through (4). (Right) Lanes contain, from left to right; Mass standard (M), Ni-NTA eluate (1), initial flow through (2), 100 mM NaCl buffer flow through (3), 120 mM NaCl buffer flow through (4), 140 mM NaCl buffer flow through (5), 160 mM NaCl buffer flow through (6), 200 mM NaCl buffer flow through (7), 500 mM NaCl buffer flow through (8).

3.5 Small-scale expression studies to optimize the expression of pCDF_PsPA7_6248

To try and optimise the expression of PA7 DprE1 a series of small-scale expression experiments were carried out. Two variables were set; time of induction and growth temperature post-induction. BL21 (DE3) cells transformed with pCDF_psPA7_6248 were induced at three different points; upon reaching either $OD_{600\ nm}$ 0.4, 0.6 or 0.8 and then grown overnight at either 16 °C or 25 °C. These temperatures were chosen as they are determined to be the best range for the overexpression of recombinant proteins in BL21 (DE3) (Rosano & Ceccarelli, 2014). Once harvested, cells were sonicated and a sample of crude lysate was centrifuged to separate soluble and insoluble fractions. Gel electrophoresis of these fractions showed that 25 °C appeared to be a more effective temperature for expression of DprE1 (Fig. 27).

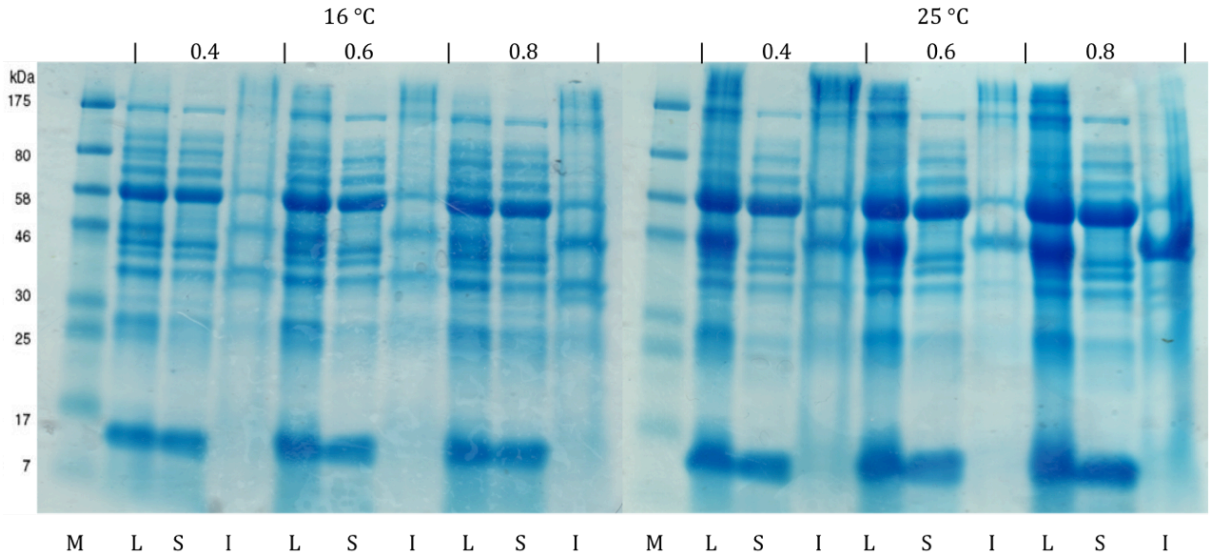


Fig. 27 SDS-PAGE analysis of lysate fractions induced at variable temperature; 16 °C and 25°C. Gels show samples of crude lysate (L), soluble fraction (S) and insoluble fraction (I) of different cultures of BL21 (DE3) pCDF_psPA7_6248 induced at different OD_{600nm} readings (0.4, 0.6 or 0.8) and then grown overnight at either 16 °C (left) or 25°C (right)

3.6 Western blot analysis of BL21 (DE3) + pCDF_PsPA7_6248 lysate fractions

Western blot analysis of the gels was then performed on gels identical to those shown in Fig. 28 (non-stained). An antibody specific to the His-tag located at the N-terminal end of the recombinant DprE1 was used to locate which fraction the protein was located in. The PA7 DprE1 homologue was revealed to be in both the lysate and insoluble fraction, but not the soluble fraction of the sample (Fig. 28).

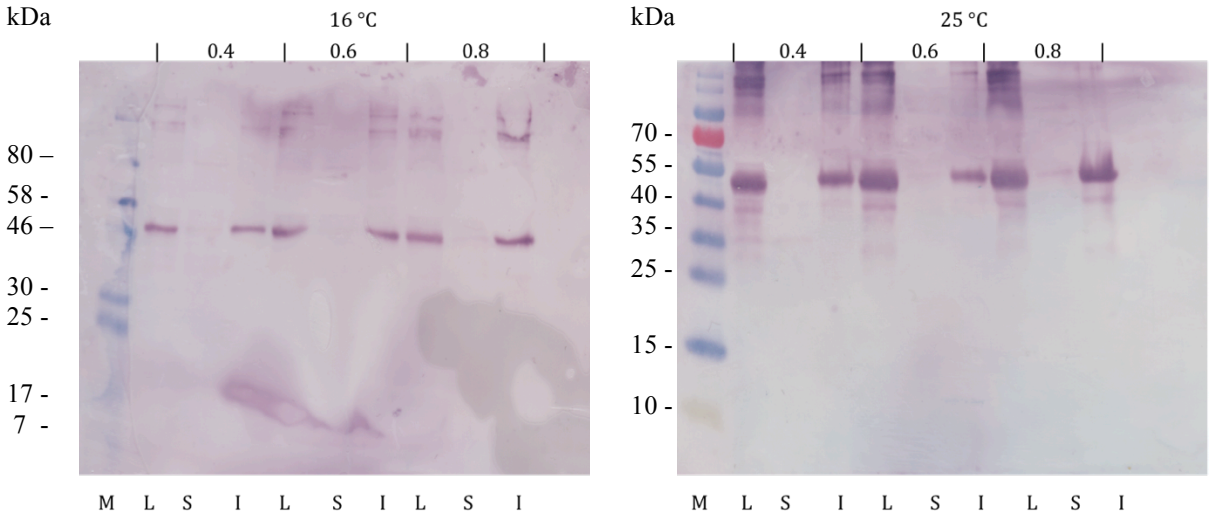


Fig. 28 Western blot analysis of lysate fractions induced at variable temperature; 16 °C and 25°C. Gels show samples of crude lysate (L), soluble fraction (S) and insoluble fraction (I) of different cultures of BL21 (DE3) pCDF_psPA7_6248 induced at different OD_{600nm} readings (0.4, 0.6 or 0.8) and then grown overnight at either 16 °C (left) or 25°C (right). Gels were then incubated with His-tag specific antibody to reveal the location of the PA7 DprE1 homologue.

4. DISCUSSION

Many studies have identified DprE1 as the target of a number of inhibitors effective at killing *M. tuberculosis* cells. As a target, DprE1 stands at the forefront of the drug-discovery effort for the treatment of TB, of which multiple drug-resistant strains now exist. DprE1 plays a key role in the growth of Mycobacteria, being required for the synthesis of arabinan, one of the components of the Mycobacterial cell wall, essential for the survival and persistence of *M. tuberculosis.* This study investigated the interaction between DprE1 and two of its inhibitors.

4.1 Purification of DprE1

The purification process of *M. tuberculosis* DprE1 was associated with some loss of the protein of interest (Page 30, Fig.4, lane 3). Western blot analysis could be used to confirm the presence of DprE1 in the flow-through sample and perhaps recycling of the flow-through back through the column would encourage more protein-column interactions.

4.2 Fluorescence spectroscopy of DprE1-inhibitor interactions

Ligand-binding assays were effective at determining the dissociation constants (K_d) of the two compounds tested. 1326 was found to have a mean K_d of 0.355 μM, whilst 1328 was found to have a higher mean K_d of 0.638 μM over 3 replicates. In contrast, the benzothiazinone compound BTZ043 has been previously shown to have a lower K_d value of 0.1462 μM (Batt *et al.,* 2012) in DprE1 titrations, therefore having a higher affinity for DprE1 than the two compounds tested in this study. BSA was used as a negative control to confirm that increase in fluorescence was due only to DprE1-inhibitor interactions and not as a result of increased

levels of protein *in vitro*. With increasing protein to ligand ratio a red shift in wavelength of the peak was observed. This effect has been well documented for the fluorescence of Tryptophan residues, which display solvatochromism, meaning the wavelength of their emission peak is highly dependent on the polarity of their environment (Park, Kim & Bark, 2002). The emission peak wavelength shift observed in the case of 1326 could be due to the changing of functional groups on residues in DprE1 stimulated by inhibitor binding (Gasymov & Glasgow, 2008). The fact this bathochromic shift (i.e. to a longer wavelength) is observed in titrations involving 1326 but not 1328 might suggest the specific binding mechanisms of each to DprE1 are different.

4.3 Purification of DprE1-homologues from *M. smegmatis* and *P. aeruginosa*

M. smegmatis DprE1 was purified with moderate success. Like *M. tuberculosis* DprE1, affinity chromatography resulted in a small loss of protein in the 50 mM Imidazole wash. Decreasing the imidazole concentration of the wash buffer might help to prevent this DprE1 loss, whilst still being effective at removing non-specific proteins from the column. Purified *M. smegmatis* DprE1 showed signs of FAD reduction, appearing colourless, however the exact reducing agent is not known. This reduction of may allow spectroscopic titrations to be performed at flavin emission wavelengths, as the linear increase in intensity over time associated with flavin oxidation may not be seen. Incubation of reduced-DprE1 with FAD could be tried in order to re-oxidise the flavin group.

PA7 DprE1 mostly resided in the insoluble fraction upon centrifugation of lysed cells. This is most likely due to its incorporation in inclusion bodies, dense aggregates of denatured protein, which cannot be purified under standard conditions (Singh & Panda, 2005). Incorporation of

protein into inclusion bodies is dependent on chaperone molecules (Hwang, Pan & Sykes, 2014). Expression of PA7 pCDF_PsPA7_6248 without the co-expression of the chaperonin Cpn60.2 and co-chaperonin GroES might prevent or limit the formation of these insoluble inclusion bodies. Lysis and fractioning of these cells followed by western blot analysis of each would reveal if this omission of chaperones has any effect. Conversely, co-expressing PsPA7_6248 with different chaperonin molecules, including GroL and GroS from *P. aeruginosa,* might yield better results. IPTG concentration has also been implicated in the formation of inclusion bodies during recombinant protein induction, with lower concentrations generally inducing protein at a slower rate and therefore generally increasing protein solubility (Tolia & Joshua-Tor, 2006).

Purification of inclusion bodies can be performed under denaturing conditions. Urea or similar is used to isolate the inclusion bodies before a strong denaturant and reducing agent e.g. β-mercaptoethanol or DTT are used to solubilise the protein (Lilie, Schwarz & Rudolph, 1998). Protein must then be refolded but this is not always achievable and often does not result in the natural state of the protein. (Peti & Page, 2007). The reducing agent would also change the redox state of the flavin cofactor and make performing fluorescence spectroscopy experiments and measuring flavin excitation difficult.

4.4 Further work

Due to the difficulties associated with the structure determination of DprE1 in complex with 1326 and 1328 and the potentially long time frame of this approach other experiments could be performed along side.

Determination of the MIC of the two compounds would allow their efficacy to be tested directly on *M. tuberculosis* cells either on their own, or in combination with other anti-TB compounds. This can be achieved using *in vitro* activity assays by incubating bacterial cultures of known cfu with the inhibitor(s) and monitoring their effect after a given period of time. Bacteria of different species could also be tested to determine the specificity of the compounds.

An accompanying set of experiments could be performed to force the evolution of inhibitor-resistant strains of Mycobacteria. As a preliminary, evolution could either be investigated in *M. smegmatis* or *M. bovis BCG* due to the highly similar DprE1 homologues expressed in this organisms. This would forgo the need for a category 3 lab, and should also be more efficient due to the generation time of *M. smegmatis* being significantly lower than that of *M. tuberculosis.* Generation of mutants requires selection by incubating cells with a concentration of inhibitor above that of the MIC of susceptible cells (MIC_{susc}) but below that of resistant bacteria (MIC_{res}). This is known as the mutant selective window hypothesis (Gullberg *et al.,* 2011). Inhibitor-resistant mutants can then be sequenced and any amino acid mutations in the DprE1-coding gene noted. Amino acid changes would most likely be indicative of key residues required for formation of enzyme-ligand complex. Expressing the altered gene sequence in a plasmid construct would then allow for the purification of any mutant DprE1 proteins identified. Purified mutant DprE1 could then be used to repeat the titration experiments described in this study. If resistance to the inhibitory compounds is as a result of the specific amino acid changes then no increase in fluorescence at the compound-specific emission wavelengths would be observed.

4.4 Conclusion

This study has used fluorescence spectroscopy to determine the K_d values of two inhibitors of the essential *M. tuberculosis* enzyme, DprE1. Titrations with an Y314C mutant form of the enzyme has showed that this amino acid change does not diminish the binding affinity of the inhibitors, indicating that they interact with DprE1 via a different mechanism to the azaindole and TCA-1 class of inhibitors. Mean K_d values of 0.355 μM and 0.638 μM were determined for 1326 and 1328, respectively. Efforts to produce efficient plasmid constructs for the expression of DprE1 homologues from *M. smegmatis* and *P. aeruginosa* were largely unsuccessful, therefore further work is required to achieve this goal.

5. REFERENCES

Alcaide, F., Pfyffer, G.E., Telenti, A., 1997. Role of embB in natural and acquired resistance to ethambutol in mycobacteria. Antimicrob Agents Chemother 41, 2270–2273.

Baciewicz, A.M., Chrisman, C.R., Finch, C.K., Self, T.H., 2013. Update on rifampin, rifabutin, and rifapentine drug interactions. Curr Med Res Opin 29, 1–12.

Baghaei, P., Marjani, M., Javanmard, P., Tabarsi, P., Masjedi, M.R., 2013. Diabetes mellitus and tuberculosis facts and controversies. J Diabetes Metab Disord 12, 58.

Batt, S.M., Jabeen, T., Bhowruth, V., Quill, L., Lund, P.A., Eggeling, L., Alderwick, L.J., Futterer, K., Besra, G.S., 2012. Structural basis of inhibition of Mycobacterium tuberculosis DprE1 by benzothiazinone inhibitors. Proc Natl Acad Sci U S A 109, 11354–11359.

Bornhorst, J.A., and Falke, J.J., 2000. Purification of Proteins Using Polyhistidine Affinity Tags. Methods In Enzymology 326, 245-254.

Brennan, P.J., 2003. Structure, function, and biogenesis of the cell wall of Mycobacterium tuberculosis. Tuberculosis (Edinb) 83, 91–97.

Campbell, E.A., Korzheva, N., Mustaev, A., Murakami, K., Nair, S., Goldfarb, A., Darst, S.A., 2001. Structural Mechanism for Rifampicin Inhibition of Bacterial RNA Polymerase. Cell 104, 901–912.

Chatterji, M., Shandil, R., Manjunatha, M.R., Solapure, S., Ramachandran, V., Kumar, N., Saralaya, R., Panduga, V., Reddy, J., Kr, P., Sharma, S., Sadler, C., Cooper, C.B., Mdluli, K., Iyer, P.S., Narayanan, S., Shirude, P.S., 2014. 1,4-Azaindole, a Potential Drug Candidate for Treatment of Tuberculosis. Antimicrob. Agents Chemother. 58, 5325–5331. doi:10.1128/AAC.03233-14

Cole, S.T., Riccardi, G., 2011. New tuberculosis drugs on the horizon. Current Opinion in

Microbiology 14, 570–576.

Daniel, T.M., 2004. The impact of tuberculosis on civilization. Infectious Disease Clinics of North America 18, 157–165.

Delogu, G., Sali, M., Fadda, G., 2013. The Biology of Mycobacterium Tuberculosis Infection. Mediterr J Hematol Infect Dis 5.

Ehrt, S., Rhee, K., 2013. Mycobacterium tuberculosis metabolism and host interaction: mysteries and paradoxes. Curr. Top. Microbiol. Immunol. 374, 163–188.

Fattorini, L., Piccaro, G., Mustazzolu, A., Giannoni, F., 2013. Targeting Dormant Bacilli to Fight Tuberculosis. Mediterr J Hematol Infect Dis 5.

Gagneux, S., 2012. Host-pathogen coevolution in human tuberculosis. Philos Trans R Soc Lond B Biol Sci 367, 850–859.

Gasymov, O.K., Glasgow, B.J., 2007. ANS Fluorescence: Potential to Augment the Identification of the External Binding Sites of Proteins. Biochim Biophys Acta 1774, 403–411. doi:10.1016/j.bbapap.2007.01.002

Ghisla, S., 1980. Fluorescence and Optical Characteristics of Reduced Flavins and Flavoproteins. First publ. in: Methods in Enzymology 66, (1980), pp. 360-373.

Goel, D., 2014. Bedaquiline: A novel drug to combat multiple drug-resistant tuberculosis. J Pharmacol Pharmacother 5, 76–78.

Goude, R., Amin, A.G., Chatterjee, D., Parish, T., 2009. The Arabinosyltransferase EmbC Is Inhibited by Ethambutol in Mycobacterium tuberculosis. Antimicrob Agents Chemother 53, 4138–4146.

Green, K.D., Garneau-Tsodikova, S., 2013. Resistance in tuberculosis: what do we know and where can we go? Front Microbiol 4.

Griffith, D.E., Brown-Elliott, B.A., Shepherd, S., McLarty, J., Griffith, L., Wallace, R.J., 2005. Ethambutol Ocular Toxicity in Treatment Regimens for *Mycobacterium avium*

Complex Lung Disease. American Journal of Respiratory and Critical Care Medicine 172, 250–253.

Grover, S., Alderwick, L.J., Mishra, A.K., Krumbach, K., Marienhagen, J., Eggeling, L., Bhatt, A., Besra, G.S., 2014. Benzothiazinones Mediate Killing of Corynebacterineae by Blocking Decaprenyl Phosphate Recycling Involved in Cell Wall Biosynthesis. J. Biol. Chem. 289, 6177–6187.

Gullberg, E., Cao, S., Berg, O.G., Ilbäck, C., Sandegren, L., Hughes, D., Andersson, D.I., 2011. Selection of Resistant Bacteria at Very Low Antibiotic Concentrations. PLoS Pathog 7, e1002158. doi:10.1371/journal.ppat.1002158

Guo, H., Seet, Q., Denkin, S., Parsons, L., Zhang, Y., 2006. Molecular characterization of isoniazid-resistant clinical isolates of Mycobacterium tuberculosis from the USA. J Med Microbiol 55, 1527–1531.

Harvey, H., Kus, J.V., Tessier, L., Kelly, J., Burrows, L.L., 2011. Pseudomonas aeruginosa d-Arabinofuranose Biosynthetic Pathway and Its Role in Type IV Pilus Assembly. J Biol Chem 286, 28128–28137.

Hershkovitz, I., Donoghue, H.D., Minnikin, D.E., Besra, G.S., Lee, O.Y.-C., Gernaey, A.M., Galili, E., Eshed, V., Greenblatt, C.L., Lemma, E., Bar-Gal, G.K., Spigelman, M., 2008. Detection and Molecular Characterization of 9000-Year-Old Mycobacterium tuberculosis from a Neolithic Settlement in the Eastern Mediterranean. PLoS ONE 3, e3426. doi:10.1371/journal.pone.0003426

Hwang, P.M., Pan, J.S., Sykes, B.D., 2014. Targeted expression, purification, and cleavage of fusion proteins from inclusion bodies in Escherichia coli. FEBS Letters 588, 247–252.

Johnson, I., Spence, M.T.Z., 2010. Molecular Probes Handbook, A Guide to Fluorescent Probes and Labelling Technologies, 11th Edition, Carlsbad California, Invitrogen.

Kilburn, J.O., Greenberg, J., 1977. Effect of Ethambutol on the Viable Cell Count in

Mycobacterium smegmatis. Antimicrob Agents Chemother 11, 534–540.

Kolly, G.S., Boldrin, F., Sala, C., Dhar, N., Hartkoorn, R.C., Ventura, M., Serafini, A., McKinney, J.D., Manganelli, R., Cole, S.T., 2014. Assessing the Essentiality of the Decaprenyl-phospho-D-arabinofuranose Pathway in Mycobacterium tuberculosis using Conditional Mutants. Mol. Microbiol.

Koma, Y., Goto, K., Yoshida, C., Kimura, K., Matsumoto, Y., Koyama, M., Nakashima, N., Masuya, D., Matsuoka, H., Yoshimatsu, H., Suzuki, Y., 2013. Pneumonitis induced by rifampicin: a case report and literature review. Intern. Med. 52, 473–477.

Kraft, C.A., Garrido, J.L., Leiva-Vega, L., Romero, G., 2009. Quantitative Analysisof Protein-Lipid Interactions Using Tryptophan Fluorescence. Science Signalling 2 (99), pp.4

Lakshmanan, M., Xavier, A.S., 2013a. Bedaquiline - The first ATP synthase inhibitor against multi drug resistant tuberculosis. J Young Pharm 5, 112–115.

Lakshmanan, M., Xavier, A.S., 2013b. Bedaquiline – The first ATP synthase inhibitor against multi drug resistant tuberculosis. Journal of Young Pharmacists 5, 112–115.

Lawn, S.D., Zumla, A.I., 2011. Tuberculosis. The Lancet 378, 57–72.

Lee, R.E., Hurdle, J.G., Liu, J., Bruhn, D.F., Matt, T., Scherman, M.S., Vaddady, P.K., Zheng, Z., Qi, J., Akbergenov, R., Das, S., Madhura, D.B., Rathi, C., Trivedi, A., Villellas, C., Lee, R.B., Rakesh, Waidyarachchi, S.L., Sun, D., McNeil, M.R., Ainsa, J.A., Boshoff, H.I., Gonzalez-Juarrero, M., Meibohm, B., Böttger, E.C., Lenaerts, A.J., 2014. Spectinamides: a new class of semisynthetic antituberculosis agents that overcome native drug efflux. Nat Med 20, 152–158.

Leistikow, R.L., Morton, R.A., Bartek, I.L., Frimpong, I., Wagner, K., Voskuil, M.I., 2010. The Mycobacterium tuberculosis DosR Regulon Assists in Metabolic Homeostasis and Enables Rapid Recovery from Nonrespiring Dormancy. J. Bacteriol. 192, 1662–

1670.

Lew, J.M., Kapopoulou, A., Jones, L.M., Cole, S.T., 2011. TubercuList – 10 years after. Tuberculosis (Edinb) Jan 91(1), 1-7.

Lilie, H., Schwarz, E. & Rudolph, R. (1998) Advances in refolding of proteins produced in E. coli. Current Opinion Biote, n.d.

Mahajan, R., 2013. Bedaquiline: First FDA-approved tuberculosis drug in 40 years. Int J Appl Basic Med Res 3, 1–2.

Makarov, V., Manina, G., Mikusova, K., Mollmann, U., Ryabova, O., Saint-Joanis, B., Dhar, N., Pasca, M.R., Buroni, S., Lucarelli, A.P., Milano, A., De Rossi, E., Belanova, M., Bobovska, A., Dianiskova, P., Kordulakova, J., Sala, C., Fullam, E., Schneider, P., McKinney, J.D., Brodin, P., Christophe, T., Waddell, S., Butcher, P., Albrethsen, J., Rosenkrands, I., Brosch, R., Nandi, V., Bharath, S., Gaonkar, S., Shandil, R.K., Balasubramanian, V., Balganesh, T., Tyagi, S., Grosset, J., Riccardi, G., Cole, S.T., 2009. Benzothiazinones Kill Mycobacterium tuberculosis by Blocking Arabinan Synthesis. Science 324, 801–804. doi:10.1126/science.1171583

Manca, C., Koo, M.-S., Peixoto, B., Fallows, D., Kaplan, G., Subbian, S., 2013. Host Targeted Activity of Pyrazinamide in Mycobacterium tuberculosis Infection. PLoS One 8.

Martins, M.A.P., Reis, A.M.M., Sales, M.F., Nobre, V., Ribeiro, D.D., Rocha, M.O.C., Ribeiro, A.L.P., 2013. Rifampicin-warfarin interaction leading to macroscopic hematuria: a case report and review of the literature. BMC Pharmacol Toxicol 14, 27.

Mesfin, Y.M., Hailemariam, D., Biadglign, S., Kibret, K.T., 2014. Association between HIV/AIDS and Multi-Drug Resistance Tuberculosis: A Systematic Review and Meta-Analysis. PLoS One 9.

Möller, M., Denicola, A., 2002. Study of protein-ligand binding by fluorescence.

Biochemistry and Molecular Biology Education 30, 309–312.

Nolan, C.M., Goldberg, S.V., Buskin, S.E., 1999. Hepatotoxicity associated with isoniazid preventive therapy: a 7-year survey from a public health tuberculosis clinic. JAMA 281, 1014–1018.

Norton, B.L., Holland, D.P., 2012. Current management options for latent tuberculosis: a review. Infect Drug Resist 5, 163–173.

Park, H.-R., Kim, T.H., Bark, K.-M., 2002. Physicochemical properties of quinolone antibiotics in various environments. European Journal of Medicinal Chemistry 37, 443–460.

Pasca, M.R., Degiacomi, G., Ribeiro, A.L. de J.L., Zara, F., Mori, P.D., Heym, B., Mirrione, M., Brerra, R., Pagani, L., Pucillo, L., Troupioti, P., Makarov, V., Cole, S.T., Riccardi, G., 2010. Clinical Isolates of Mycobacterium tuberculosis in Four European Hospitals Are Uniformly Susceptible to Benzothiazinones. Antimicrob. Agents Chemother. 54, 1616–1618.

Peti, W., Page, R., 2007. Strategies to maximize heterologous protein expression in Escherichia coli with minimal cost. Protein Expr. Purif. 51, 1–10. doi:10.1016/j.pep.2006.06.024

Protein Purification - Extraction and Clarification - In vitro denaturation and refolding - EMBL [WWW Document], n.d. URL http://www.embl.de/pepcore/pepcore_services/protein_purification/extraction_clarification/invitro_denaturation_refolding/ (accessed 3.1.14b).

Rajendran, V., Sethumadhavan, R., 2014. Drug resistance mechanism of PncA in Mycobacterium tuberculosis. Journal of Biomolecular Structure and Dynamics 32, 209–221. doi:10.1080/07391102.2012.759885

Ramaswamy, S.V., Amin, A.G., Goksel, S., Stager, C.E., Dou, S.-J., El Sahly, H., Moghazeh,

S.L., Kreiswirth, B.N., Musser, J.M., 2000. Molecular Genetic Analysis of Nucleotide Polymorphisms Associated with Ethambutol Resistance in Human Isolates of Mycobacterium tuberculosis. Antimicrob Agents Chemother 44, 326–336.

Rosano, G.L., Ceccarelli, E.A., 2014. Recombinant protein expression in Escherichia coli: advances and challenges. Front. Microbiol. 5, 172. doi:10.3389/fmicb.2014.0017

Royer, C.A., and Scarlata, S.F., 2008. Fluorescent Approaches to Quantifying Biomolecular Interactions. Methods in Enzymology; 450, pp. 80-103

Saviola, B., 2013. Mycobacterium tuberculosis Adaptation to Survival in a Human Host, in: Mahboub, B. (Ed.), Tuberculosis - Current Issues in Diagnosis and Management. InTech.

Sharma, S.K., Mohan, A., 2013. Tuberculosis: From an incurable scourge to a curable disease - journey over a millennium. Indian J Med Res 137, 455–493.

Sivaramakrishnan, S., de Montellano, P.R.O., 2013. The DosS-DosT/DosR Mycobacterial Sensor System. Biosensors (Basel) 3, 259–282. doi:10.3390/bios3030259

Takayama, K., Kilburn, J.O., 1989. Inhibition of synthesis of arabinogalactan by ethambutol in Mycobacterium smegmatis. Antimicrob Agents Chemother 33, 1493–1499.

Takayama, K., Wang, C., Besra, G.S., 2005. Pathway to Synthesis and Processing of Mycolic Acids in Mycobacterium tuberculosis. Clin. Microbiol. Rev. 18, 81–101.

Tessema, B., Beer, J., Emmrich, F., Sack, U., Rodloff, A.C., 2012. Analysis of gene mutations associated with isoniazid, rifampicin and ethambutol resistance among Mycobacterium tuberculosis isolates from Ethiopia. BMC Infect Dis 12, 37.

Timmins, G.S., Deretic, V., 2006. Mechanisms of action of isoniazid. Molecular Microbiology 62, 1220–1227.

Tolia, N.H., Joshua-Tor, L., 2006. Strategies for protein coexpression in Escherichia coli. Nat Meth 3, 55–64. doi:10.1038/nmeth0106-55

Vilchèze, C., Jacobs, Jr., W.R., 2007. The Mechanism of Isoniazid Killing: Clarity Through the Scope of Genetics. Annual Review of Microbiology 61, 35–50.

Wehrli, W., and Staehelin, M., 1971. Actions of the rifamycins. Bacteriology Reviews, 35(3), pp.290-309

WHO | Guidelines for treatment of tuberculosis, fourth edition [WWW Document], n.d. WHO. URL http://www.who.int/tb/publications/2010/9789241547833/en/ (accessed 2.23.14).

Yee, D., Valiquette, C., Pelletier, M., Parisien, I., Rocher, I., Menzies, D., 2003a. Incidence of serious side effects from first-line antituberculosis drugs among patients treated for active tuberculosis. Am. J. Respir. Crit. Care Med. 167, 1472–1477.

Yee, D., Valiquette, C., Pelletier, M., Parisien, I., Rocher, I., Menzies, D., 2003b. Incidence of Serious Side Effects from First-Line Antituberculosis Drugs among Patients Treated for Active Tuberculosis. American Journal of Respiratory and Critical Care Medicine 167, 1472–1477.

Zhang, Y., Mitchison, D., 2003. The curious characteristic of pyrazinamide. Int J Tuberc Lung Dis 7(1), 6-21

Zhang, Y., 2014. Persisters, persistent infections and the Yin-Yang model. Emerg Microbes Infect 3, e3.

Zumla, A., Nahid, P., Cole, S.T., 2013. Advances in the development of new tuberculosis drugs and treatment regimens. Nature Reviews Drug Discovery 12, 388–404.

s

INVESTIGATING THE ROLE OF THE CHAPERONINS IN *MYCOBACTERIUM MARINUM*

ABSTRACT

The chaperonins are a distinct class of molecular chaperones. Mycobacteria contain multiple chaperonins, coded for the *cpn* genes; *cpn60* and *cpn10* homologous to *E. coli groEL* and *groES*, respectively. Both *cpn10* and *cpn60.2* have been found to be essential, whilst *cpn60.1* is dispensable. Mutants lacking *cpn60.1* have been shown to be deficient in biofilm formation and the ability to establish granulomas in *Mycobacterium smegmatis* and *Mycobacterium tuberculosis*, respectively. *Mycobacterium marinum* has been established as an excellent model system for the study of TB infection due to its relative safety, ease of use and fast growth rate compared to *M. tuberculosis*. *M. marinum* also has the advantage of available animal infection systems in *Danio rerio* for the study of *in vivo* phenotype. Here we investigate the phenotype of a Δ*cpn60.1* mutant of *M. marinum*. Loss of Cpn60.1 appears to convey an advantage to heat-shock; possibly due to an up-regulation of Cpn60.2. Δ*cpn60.1* mutants also show an altered lipid profile with increased expression of one cell wall component. Overall, these results support the evidence that Cpn60.1 is a non-essential chaperonin with a probable moonlighting function.

ACKNOWLEDGMENTS

Many thanks to Dr. Pete Lund for giving me the opportunity to work with such an interesting organism and learn new, exciting techniques. Nikhil, Hrishi, Thippesh & Albel, for all your help throughout the project – thank you.

Abbreviation/Acronym	
2D-TLC	Two dimensional thin layer chromatography
ATP	Adenosine triphosphate
AUC	Analytical ultracentrifugation
BCG	Bacillus Calmette-Guérin
BLAST	Basic Local Alignment Search Tool
BSL	Biosafety level
cDNA	Complementary DNA
CFU	Colony-forming units
CIRCE	Controlling inverted repeat of chaperone expression
Cpn	Chaperonin
DAG	Diacyl Glycerol
DEPC	Diethyl pyrocarbonate
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dsDNA	Double stranded DNA
EDTA	Ethylenediaminetetraacetic acid
EMB	Ethambutol
FAS	Fatty acid synthase
FLD	First line drug
gDNA	Genomic DNA
GOI	Gene of interest
HAIR	HspR-associated inverted repeats
HIV	Human Immunodeficiency Virus
Hsp	Heat shock protein
IL- 12	Interleukin-12
IFN-γ	Gamma interferon
INH	Isoniazid
mAGP	Mycolyl-arabinogalactan-peptidoglycan
MAMT	Mycolic acid methyltransferase
MDR-TB	Multidrug-resistant tuberculosis
mRNA	Messenger RNA
MTBC	*Mycobacterium tuberculosis* complex
NAG	*N*-acetyl glucosamine
NAM	N-acetyl muramic acid
OD	Optical density
PCR	Polymerase chain reaction
PZA	Pyrazinamide
qRT-PCR	Quantitative real-time polymerase chain reaction
RIF	Rifampicin
RNA	Ribonucleic acid

TB	Tuberculosis
TDR-TB	Totally drug-resistant TB
TLC	Thin-layer chromatography
WHO	World Health Organisation
XDR-TB	Extensively drug-resistant tuberculosis

CONTENTS

1. INTRODUCTION

1.1. Mycobacteria

Mycobacteriaceae is a vast family of bacteria belonging to the order Actinobacteria. The Mycobacteria are non-motile and Gram-positive based on their phylogeny and lack of outer cell membrane. However, unlike other gram-positive bacteria, the mycobacteria do not retain the crystal violet dye used in gram staining and are therefore defined as being acid-fast. This is due in large part to their thick, lipid-rich cell walls. The morphology of Mycobacteria varies from rod-like to filamentous (Bruijnesteijn van Coppenraet, 2009).

Of the *Mycobacteriaceae* family, three species are known to be human pathogens: *Mycobacterium tuberculosis*, *Mycobacterium ulcerans* and *Mycobacterium leprae*, the etiological agents behind tuberculosis (TB), Buruli ulcers and leprosy, respectively. A number of other *Mycobacteria a*re non-human pathogens such as *Mycobacterium bovis*, which is the causative agent in bovine tuberculosis. Most of the pathogenic *Mycobacteria* are classified as slow growing, taking 7 days or more to form colonies visible to the naked eye. However, there is no direct correlation between growth rate and pathogenicity (Hett & Rubin, 2008). Another group, termed the fast growing Mycobacteria, including *Mycobacterium smegmatis* and *Mycobacterium marinum*, tend to be non-pathogenic or opportunistic pathogens (Forrellad *et al,* 2013).

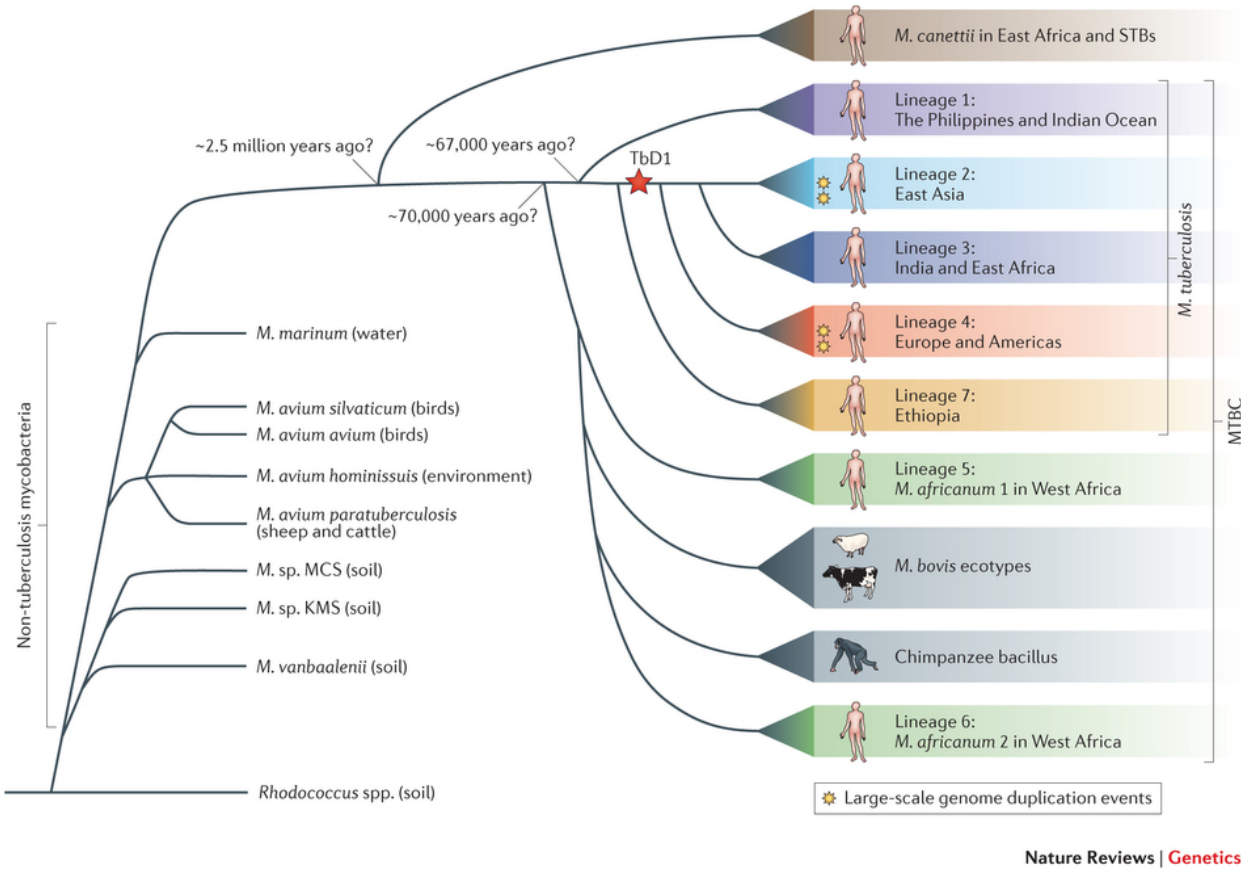


Fig. 1.1 Phylogeny and distribution of the Mycobacteria family**.**

(taken from Galagan, 2014.)

1.1.1. *Mycobacterium tuberculosis*

It's estimated that one third of the world's population is a carrier of *M. tuberculosis*. The World Health Organisation (WHO) estimates that in 2012, 8.6 million people developed active tuberculosis, of which 1.3 million cases were fatal (WHO, 2014). *Mycobacterium tuberculosis* and its sub-species form the *Mycobacterium tuberculosis* complex (MTBC), which are responsible for tuberculosis across a number of host species. Members of the MTBC are evolutionarily related, originating from a common ancestor, estimated to be ~2.5 million years old, based on evolutionary data (Fig. 1; Galagan, 2014).

1.1.2. *M. tuberculosis* mode of infection

M. tuberculosis is an air-borne pathogen, entering the host via the respiratory system in air droplets called droplet-nuclei. If these cells then reach the alveoli of the lungs they activate the release of interleukin-12 (IL-12) by dendritic cells and become ingested by macrophages, a response triggered by gamma-interferon (IFN-γ) release from $CD4^+$ T helper 1 (Th1) cells (Flynn, 2004). This immune response leads to the formation of a granuloma, resulting in the stage of *M. tuberculosis* infection known as latent tuberculosis. Bacteria can persist inside the macrophage for decades (Lillebaek et al., 2003), by preventing the development of macrophage phagosome to a bacterial phagolysosome. This is achieved by the inhibition of macrophage Ca^{2+} signaling - *M. tuberculosis* inhibits the enzyme sphingosine kinase, a key component of the Ca^{2+}/calmodulin signaling pathway (Kushner, 2005).

The switch from latent to active TB is known as reactivation and can be brought about by a number of factors including, but not limited to; immune failure, HIV, or in response to antibiotic treatment (Flynn & Chan, 2001). During reactivation, dormant bacteria are released from the granuloma as the result of macrophage apoptosis. From here the bacteria can establish pulmonary tuberculosis (67% of U.S. cases in 2011; CDC, 2013) or enter the bloodstream, where they can establish infection in another part (extrapulmonary TB) or multiple parts (miliary TB) of the body. *M. tuberculosis* relocation to the central nervous system results in a form of TB named tuberculous meningitis.

1.1.3. Survival and virulence of *M. tuberculosis*

The survivability and virulence of *M. tuberculosis* is in a large part due to its complex cell wall. All Mycobacteria share a common waxy, lipid-rich cell wall, with varying degrees of

thickness, contributing to their acid-fast staining nature. Antibiotic-resistant strains of *M. tuberculosis* have significantly larger cell walls (Veleyati *et al*., 2009), highlighting the importance of these structures in cell survival. Multiple components comprise the Mycobacterial cell wall. Peptidoglycan, arabinogalactan and mycolic acids make up the cell wall core mycolyl-arabinogalactan-peptidoglycan (mAGP) complex. The peptidoglycan layer is composed of repeating units of β-(1-4)-linked *N*-acetyl glucosamine-*N*-acetyl muramic acid (NAG-NAM) bonded together with peptide bridges. This layer is the target of antibiotics such as penicillin, but these have no efficacy in mycobacteria due to the impermeability of the mycolic acid layer. Arabinogalactan surrounds the peptidoglycan layer, comprised of arabinan-tagged galactan. The arabinan is branched and can be subjected to modifications such as the addition of non-N-acetylated galactosamine, which may have a role in pathogenesis (Škovierová *et al*, 2010). The arabinogalactan acts as an anchor for the mycolic acid layer to which it is attached via ester bonds. These are long-chain (C_{60}-C_{90}) α-alkyl β-hydroxy fatty acids formed of a saturated α-branch chain attached to a longer meromycolate chain ($\geq C_{56}$). The mycolic acids can be subdivided into three classes according to the different modifications present on the meromycolate chain; alpha-, methoxyl- and keto-mycolates. Fatty acid synthase-I (FAS-I) and FAS-II are the key components in the mycolic acid biosynthesis pathway. FAS-I synthesizes the initial short chain (C_{26}) fatty acid chain from acetyl-CoA, whilst FAS-II synthesizes the meromycolate (C_{56}) chain (Takayama, Wang & Besra, 2005; Bhatt et al., 2007). One of the components of FAS-II, the elongation-condensing enzyme, KasB has been shown to be essential for virulence. Δ*kasB* knockout strains of M. tuberculosis have been shown to persist in, but did not display pathogenicity toward immunocompetent mice (Bhatt et al., 2007). Mycolic acid modifications are also important for dampening the host immune response. Strains deficient in the certain mycolic acid methyltransferases (MAMTs) responsible for mycolate cyclopropanation show an increased

sensitivity to the host immune response, becoming attenuated one or two weeks post infection (Barkan *et al.,* 2012). Another key factor, Trehalose 6,6′-dimycolate (cord factor) has been shown to induce a granulomatous response when injected into rat corneas (Saita et al., 2000). Overall the Mycobacterial cell wall is fundamental to the survival and pathogenicity of *M. tuberculosis*, and any factors that affect this key structure can have drastic consequences.

1.1.4. Drug resistance in *M. tuberculosis*

Mycobacterium tuberculosis infection is currently treated with a combination of 4 first-line drugs (FLD); Isoniazid (INH), Rifampicin (RIF), Ethambutol (EMB) and Pyrizinamide (PZA). A number of second and third line anti-tubercular drugs are also used in combination when a poor patient response to the first-line drugs is observed. However, these show decreased efficacy against TB and often have side effects leading to poor patient compliance. The drug regimen of TB is very strict, with a minimum treatment time course of a 6-month. Development of the disease is more prolific amongst immunocompromised patients, such as those who are HIV-positive and are 12-20 times more at risk (WHO, 2014). These factors, coupled with non-diagnosed cases and poor access to anti-TB treatment have led to the emergence of antibiotic-resistant strains of TB. Multi drug-resistant tuberculosis (MDR-TB) strains are resistant to one or more of the FLDs. Strains that are non-responsive to treatment with all FLDs and at least one second-line drug are known as extensively drug-resistant (XDR-TB). Reports of totally drug-resistant tuberculosis (TDR-TB) were first noted in Italy in 2004. Since then, numerous cases have been reported, with patients showing no response to all first and second-line drugs (Parida *et al,* 2014). The majority of current TB research has therefore been focused on identifying anti-TB compounds and their targets.

1.1.5. *Mycobacterium marinum* as a model system for TB-study

Mycobacterium marinum is primarily a pathogen of ectotherms, in which it causes a TB-like infection with the establishment of granulomas. It also has a well-documented opportunistic pathogenicity towards humans. Due to its low optimal growth temperature, infection in humans is usually limited to extremities, such as hands and feet, but has also been shown to establish infection in other areas (Benton & Karkanevatos, 2007). *M. marinum* is the closest evolutionary relative to the members of the *M. tuberculosis* complex (see Fig. 1.1; Cronan & Tobin, 2014). 3000 orthologous genes are shared between *M. marinum* and *M. tuberculosis* with an average amino acid identity of 85% (Stinear *et al*., 2008). Many of these orthologues are virulence genes, with a number of *M. tuberculosis* genes able to complement a mutant gene in *M. marinum* (El-Etr et al., 2004). Unlike *M. tuberculosis*, a biosafety level 3 (BSL-3) pathogen requiring specialist equipment and laboratories, *M. marinum* is a BSL-2 pathogen and can be worked with safely on the bench top. *M. marinum* also has the advantage of being a relatively fast growing member of the Mycobacteria genus with a doubling time of 4-6 hours (without antibiotics), compared to the 20 hour doubling time of *M. tuberculosis* (Barker *et al.*, 1999). Numerous animal model systems are available for *M. marinum* infection assays. Hosts such as *Drosophila melanogaster* and *Dario rerario* (zebrafish) have been used for the study of anti-tubercular drugs and disease pathogenesis (Takaki *et al.,* 2013; Oh *et al.,* 2013). These animal systems have the advantage of being cheap, easily genetically manipulated, having a rapid life cycle, and generating large numbers of offspring. Zebrafish also have the added benefit of being translucent at the embryonic stage, allowing *in vivo* imaging of infecting bacteria in real time (Davis et al., 2002).

1.2. The chaperonins of *E. coli*

The chaperonins are a sub-class of molecular chaperones, and can be divided into two types; type-I and type-II. Type-II chaperonins are found in archaea and the eukaryotic cytosol and are self sufficient, not requiring a co-chaperonin to function (Horwich *et al.,* 2007). Type-I chaperonins are homologous to GroEL in *E. coli.* GroEL, along with the co-chaperonin GroES, are coded for by the *groE* genes. Discovery of the *groE* genes in the early 1970s resulted from inhibition of bacteriophage growth as a direct result of their mutation (Friedman, 1992). Although classically thought of as heat-shock proteins, GroEL and GroES were also shown to be essential for growth at low temperatures in *E. coli* (Fayet *et al*., 1989). The function of these proteins was initially brought to light upon discovery that they are homologous to proteins essential for the correct folding of Rubisco, a key part of the photosynthetic process in plant chloroplasts (Hemmingsen *et al*., 1988). This was confirmed when overexpression of GroEL and GroES were shown to aid in the *in vitro* refolding of un-folded Rubisco (Gouloubinoff *et al.,* 1989). Around 5% of *E. coli* proteins interact with GroEL (Azia *et al.,* 2012). Roughly 300 proteins are substrates of GroEL, of which 85 are GroEL-dependent and 15 are essential for cell viability (Lund, 2009).

The induction of GroEL-GroES during the heat-shock response is dependent on the gene *rpoH*, which encodes the alternative sigma factor, σ^{32} (Guisbert, *et al*., 2004). Increases in temperature lead to a destabilization of an RNA structural element, which under normal conditions serves to inhibit σ^{32} transcription (Morita *et al*, 1999). The concentration of σ^{32} under normal conditions is also dependent on negative regulation by the proteins DnaK, DnaJ and GrpE. Strains mutated for *dnaK, dnaJ* or *grpE* show increased levels of σ^{32} at 30°C (Straus *et al*., 1990). These three proteins collectively form the DnaK chaperone system

responsible for promoting ATP-dependent protease degredation of σ^{32} (Suzuki *et al*., 2012). σ^{32} acts as a transcription factor to direct RNA polymerase to specific promoters, including the *groESL* operon (Cowing & Gross, 1989). A second σ^{70}-dependent promoter also lies upstream of the *groESL* operon leading to low-levels of transcription of GroEL and GroES under normal conditions, in line with their secondary housekeeping role (Zhou *et al.,* 1988).

1.2.1. Structure and the GroEL ATPase cycle

GroEL (also known as hsp60) comprises fourteen 57 kDa protomers arranged as two symmetrical seven-member rings, referred to as the *cis* and *trans* rings ,(Saibil *et al*., 2013). Each protomer can be divided into three domains; an equatorial domain and an apical domain connected by a smaller intermediate domain. Each protomer contains an ATP-binding pocket located in the equatorial domain, essential for function (Roseman *et al*., 1996). The apical domains form a large hydrophobic ring essential for binding of polypeptides and GroES, with which GroEL forms a complex (Fenton *et al*., 1994). The GroEL reaction is initiated by the movement of the individual protomers between the apical and intermediate domains in response to ATP-binding (Braig *et al*., 1995). This is followed immediately by the binding of nonnative polypeptide to the *cis* ring. GroES displaces the polypeptide by binding to the *cis* ring, forming a "lid" like structure and forcing the polypeptide into the large 85 000 $Å^3$ central cavity (Lund, 2009). GroES is comprised of seven 10 kDa subunits, within which β-strands of individual subunits interact with one another to stabilize a dome-like ring structure (Ikeda-Kobayashi *et al*., 2012). A mobile loop at the opposite end of each subunit makes a 1:1 contact with the hydrophobic region on the apical domain of an ATP-containing GroEL protomer (Saibil, *et al*., 2013). Binding of GroES stimulates a conformational change in GroEL as the hydrophobic surface on the apical domains is shifted and replaced with a

hydrophilic surface (Xu *et al*., 1997). The environment created by the binding of GroES allows the nonnative protein to refold, unhindered. Binding of ATP to the *trans* ring releases the newly folded protein and ADP from the *cis* portion of GroEL (Lund, 2009). Overall, the GroEL-GroES-ATP_7 complex promotes refolding of denatured proteins by isolating them from non-permissive conditions and introducing them to an environment without external hindrances (Lin & Rye, 2006). This model of chaperonin-mediated refolding is known as the Anfinsen cage, named for the nobel prize-winning biochemist Christian B, Anfinsen.

1.3 The multiple chaperonins of Mycobacteria

Around 30% of all bacteria have more than one copy of GroEL (Lund, 2009), suggestive of a divergence in function. The Mycobacteria contain genes homologous to *groEL* and *groES*, named *cpn60* and *cpn1*0, respectively. Most Mycobacteria have more than homologue of *cpn60.* Initial phylogenetic analysis of *groEL* genes across all bacteria revealed two distinct clades (Fig. 1.3; Rao & Lund, 2010; Goyal et al., 2006). These are given the nomenclature *GroEL1 (cpn60.1)*, and *GroEL2 (cpn60.2).* In *Mycobacterium smegmatis* a further gene named *cpn60.3* is present. BLAST searches revealed the best-matched *cpn60.3* homologue exists in *Rhodococcus jostii,* likely making it the result of a more recent horizontal gene transfer event (Rao & Lund, 2010). In all recorded cases *cpn60.2* is present, whilst *cpn60.1* is occasionally absent, hinting at its dispensable nature. This hypothesis was confirmed in *M. tuberculosis* (Hu *et al*., 2008), *M. smegmatis* (Ojha *et al*., 2005) and *M. bovis BCG* (Wang *et al*., 2011) in which Δ*cpn60.2* mutants were non-viable. In the same three species Δ*cpn60.1* mutants were viable but displayed altered phenotypes (discussed later). Phylogenetic analysis reveals that the GroES homologue, *cpn10* is found in all Mycobacteria and is also indispensable (Hu *et al*., 2008).

In species with more than one *cpn60* gene, *cpn10* is found in an operon with *cpn60.1,* with *cpn60.2* located elsewhere on the genome. When *cpn60.2* is the only *cpn60* gene, *cpn10* exists on its own. The arrangement of the *cpn* genes in Actinobacteria is shown in Fig. 1.4 (Lund, 2009).

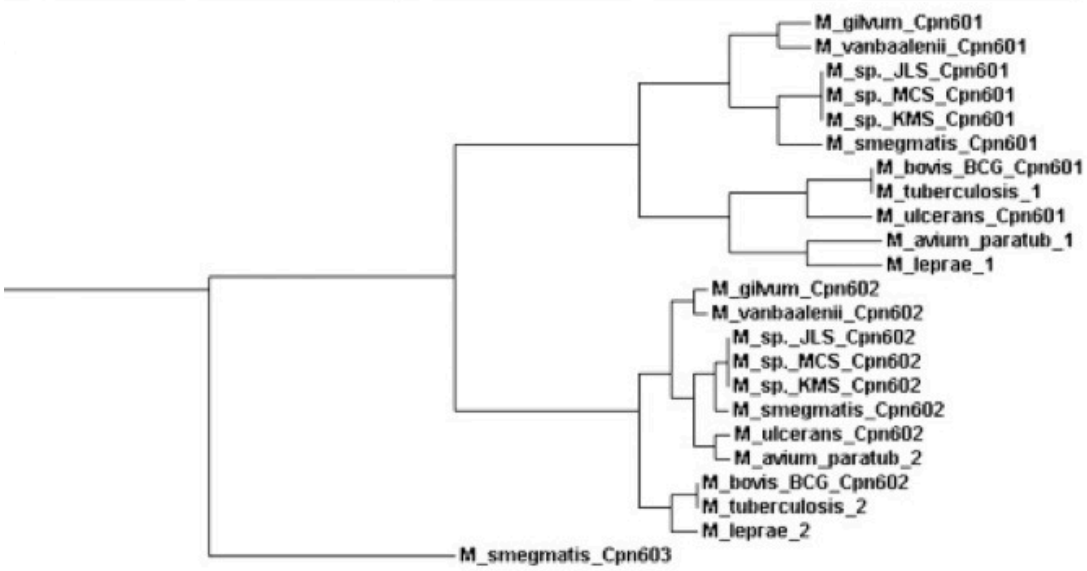


Fig. 1.2. Phylogenetic tree showing divergence of the Mycobacterial chaperonins (taken from Rao & Lund, 2010).

The Cpn60 proteins can be characterised based on C-terminal amino acid consensus sequences. Cpn60.1 has a Histidine-rich C-terminus, and Cpn60.2 C-termini always contain a glycine and methionine-rich motif known as a GGM repeat. The relevance of these C-terminal features is unknown, however Colaco & MacDougall, 2013 suggest a role for the differential recognition of the two homologues by target proteins.

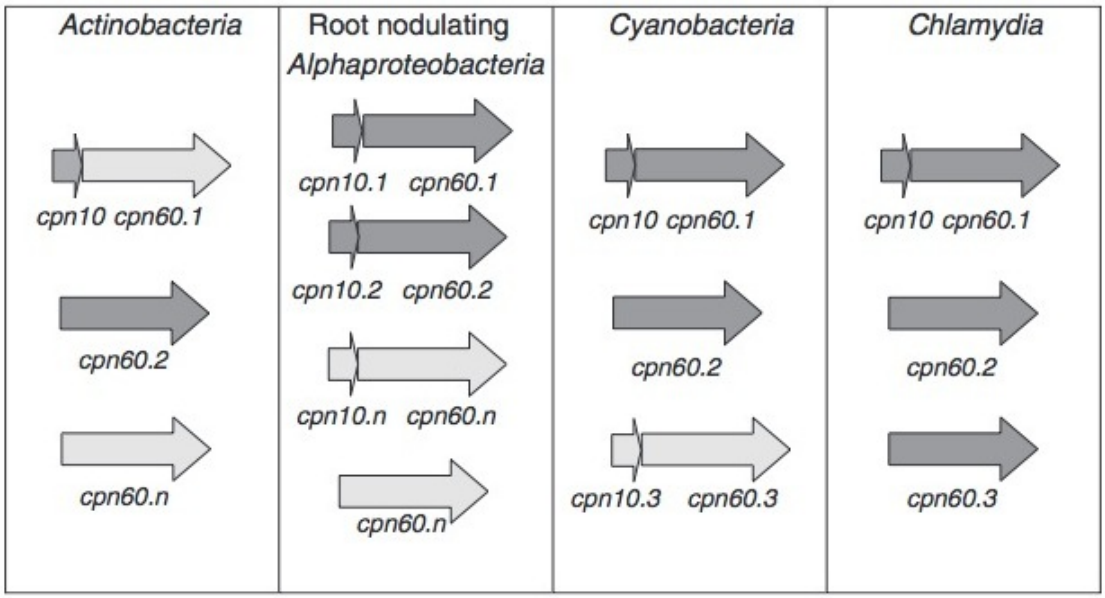


Fig.1.3. The genomic organization of chaperonin genes across different genera of bacteria (taken from Lund, 2009).

Unlike *E. coli* GroEL, Cpn60.1 and Cpn60.2 do not form large multimeric complexes under standard conditions, but rather form lower oligomers (Qamra, Srinivas & Mande, 2004). Fan *et al*., 2012, used analytical ultracentrifugation (AUC) and showed that the molecular mass of eluted Cpn60.1 and Cpn60.2 proteins were much lower than that of GroEL, and confirmed this was the case using gel electrophoresis. They then went on to show that at high concentrations of salt and nucleotide Cpn60.2 formed high-order structures reminiscent of the tetradecameric structure of GroEL. This, coupled with the finding that Mycobacterial Cpn60.2 proteins can replace *E. coli* GroEL *in vivo* (Hu *et al*., 2008; Fan *et al*., 2012), confirm that Cpn60.2 is the Mycobacterial equivalent of *E. coli* GroEL.

1.3.1 Induction of Mycobacterial chaperonins in the heat-shock response

The Mycobacterial chaperonin genes all show an inverted repeat region upstream, known as a Controlling Inverted Repeat of Chaperone Expression (CIRCE) sequence (TTAGCACTC-N_9-GAGTGCTAA-5'; Zuber & Schumann, 1994; Narberhaus, 2002). The regulatory protein, HrcA binds to CIRCE sequences, repressing the expression of CIRCE-regulated genes. In Δ*hspR*Δ*hrcA* double mutants of *M. tuberculosis* the expression of the chaperonin genes are upregulated. In the Δ*hspR M. tuberculosis* single mutant, only *cpn10* is up-regulated, indicating *cpn60.1* and *cpn60.2* are regulated by HrcA, and HspR plays a role in the regulation of *cpn10* (Stewart *et al*., 2002). HspR binds to a consensus sequence called HspR-associated inverted repeats (HAIR). It has been shown that *cpn60.1* and *cpn60.2* (with DnaK) are required for the activation of HspR (Das Gupta *et al.*, 2008), highlighting a possible positive feedback response for the expression of *cpn10*. The HspR-activated expression of *cpn10* is consistent with the hypothesis that *cpn60.1* and *cpn10* might not be co-transcribed and their expression possibly regulated by two different mechanisms (Kong *et al*., 1993).

1.3.2. The role of *cpn60.1* in biofilm formation

The multiple copies of *cpn60* found in Mycobacteria have led to questioning of the potential moonlighting roles of the Cpn60 proteins. Moonlighting is a term used to describe the multiple roles that Cpn60.1 appears to play within different bacteria – a divergence of function away from the standard protein-folding role played by GroEL and Cpn60.2. Due to its non-essential nature, the role of Cpn60.1 has been studied through the creation of gene-specific knockout strains. In a paper published in 2005, Ojha *et al.*, found that in *M. smegmatis*, *cpn60.1* is required for the formation of mature biofilms. The ability to form of a

biofilm is an evolutionary characteristic shared by many bacteria. By secreting an exopolysaccharide matrix, the bacteria are able to form self-containted communities protecting themselves from harmful external conditions. *M. tuberculosis* is thought to form biofilms *in vivo* adding to its already high tolerance to antibiotic treatment (Islam et al., 2012). However, Hu et al., 2008 showed that Cpn60.1 was not required for the formation of biofilms in *M. tuberculosis*. In *M. smegmatis* Δ*cpn60.1* mutants, synthesis of the C_{56}-C_{68} long chain fatty acids typical of the Mycobacteria mycolic acids is absent, especially during biofilm formation. Cpn60.1 was shown to physically associate with KasA, the essential protein that forms part of the FAS-II complex discussed previously. This evidence suggests a key role for Cpn60.1 in the switch to biofilm formation by promoting mycolate synthesis.

1.3.3 Immunomodulatory effect of the chaperonins

Mycobacterial chaperonins have been implicated in the establishment of host immune response. Both Cpn60.2 and Cpn10 from *M. tuberculosis* have been shown to have immunomodulatory effects in guinea pigs, promoting the production of certain inflammatory cytokines (Friedland *et al*., 1993; van Eden *et al.,* 2005). Likewise, Hsp65 (Cpn60.2) was shown to act as an antigen for T cell recognition in healthy humans, with the ability to react with Hsp65 from *M. tuberculosis*, *M. bovis BCG* and *M. lepraea* (Munk *et al*., 1989; Mustafa *et al*., 1993). *M. tuberculosis* Δ*cpn60.1* mutants fail to elicit an immune response in guinea pig infection models (Hu *et al*., 2008), preventing the formation of granulomas. This appears to be a result of the multinucleated giant cell (MGC) -promoting activity of Cpn60.1 (Cehovin *et al*., 2010), highlighting the role of Cpn60.1 as a virulence factor essential for the establishment of TB. Cpn60.1 and Cpn10 have both been detected as secreted products in the supernatant of *M. tuberculosis* cultures (Cehovin *et al.*, 2010; Fossati *et al*., 2003). In the

same study, Cehovin *et al*., show that Cpn60.2 was not detected in culture supernatant, but has been reported to be located on the cell surface along with the Hsp70, DnaK (Hickey *et al*., 2009). Here, Cpn60.2 has been shown to bind to the glycoprotein, CD43, located on the surface of murine macrophages (Hickey *et al.,* 2010). By binding to CD43, Cpn60.2 might act as a signal for phagocytosis of cells (as seen in *Streptococcus gordonii*; Urano-Tashira *et al.,* 2008).

1.4 Aims of this project

Previous work in our lab has looked into the role of the chaperonin genes in *Mycobacterium smegmatis*, in particular their role in stress responses (Rao & Lund, 2009). However, as a non-pathogenic organism, there exists no infection model for *M. smegmatis* for *in vivo* study of phenotypes. The work of a previous MSc student, Farah Islam, led to the generation of a Δ*cpn60.1* mutant strain of *Mycobacterium marinum* by transposon phase mutagenesis (phAE181; Islam, unpublished).

The main aim of this project will be to characterize the phenotype of this mutant strain. As Cpn60.1 appears to have a key role in mycolate synthesis, the lipid profile of Δ*cpn60.1* strain will be compared to wild type to see the effect of Cpn60.1 loss. *M. smegmatis cpn* genes were shown to be upregulated in response to heat shock and of this will be investigated in *M. marinum* using heat-killing plating assays and qRT-PCR. If phenotype is observed in any experiment, attempts will be made to generate a complemented Δ*cpn60.1* strain.

2. MATERIALS AND METHODS

2.1Bacterial strains and Plasmids

2.1.1 Strains

Organism	Strain	Source
E. coli	DH5α	Lab stock
	MGM100 + pTrcESmmcpn60.1	This lab
	MGM100 + pTrcESmmcpn60.2	This lab
M. smegmatis	Mc^2155	(Snapper *et al*., 1990)
	Mc^2155 Δcpn60.1	(Ojha *et al*., 2005)
M. marinum	WT + pMSP-dsRed	This lab
	Δcpn60.1	This lab
	Δcpn60.1 + pMSP-dsRed	This lab

2.1.2 Plasmids

Name	Description	Source
pMSGroEL1	*M. smegmatis cpn10* + *cpn60.1* -expressing plasmid derived from pMV361	(Ojha *et al*., 2005)
pMSGroEL2	*M. smegmatis cpn10* + *cpn60.2* -expressing plasmid derived from pMV361	(Ojha *et al*., 2005)
pMS10.cpn60.2	pMSGroEL1-derivative containing M. smegmatis *cpn60.2*	(Rao & Lund, 2010)
pMS10.cpn60.3	pMSGroEL1-derivative containing M. smegmatis *cpn60.3*	(Rao & Lund, 2010)
pMS10.groEL	pMSGroEL1-derivative containing *E. coli GroEL*	(Rao & Lund, 2010)

2.2 Oligonucleotide primers

Name	Sequence (5’ to 3’)	Description
Mm_cpn10_F	GCGAAGGTGAACATCAAGCCA	Forward primer for qPCR-detection of *M. marinum cpn10*
Mm_cpn10_R	AGGAATGACCAGACCGGACG	Reverse primer for qPCR-detection of *M. marinum cpn10*
Mm_cpn60_1_F	CTGCCGTTGCTGGAGAAGGT	Forward primer for qPCR-detection of *M. marinum cpn60.1*
Mm_cpn60_1_R	CGTCTTGCGAATGGAGTTCACG	Reverse primer for qPCR-detection of *M. marinum cpn60.1*
Mm_cpn60_2_F	GGTGGTCAGGTCATCAGCGAA	Forward primer for qPCR-detection of *M. marinum cpn60.2*
Mm_cpn60_2_R	ATGGTGGTCTCGTCCTTGGTG	Reverse primer for qPCR-detection of *M. marinum cpn60.2*
Mm_gapdh_F	CCTACACCCAGGACCAGAAC	Forward primer for qPCR-detection of *M. marinum gap*
Mm_gapdh_R	ATCACCAGACCGATGGCTTT	Reverse primer for qPCR-detection of *M. marinum gap*
MS_cpn10_R	AGGTACTCCTCGCCGTTGTA	Reverse primer with affinity to *M. smegmatis cpn10* for upstream sequencing of plasmids

2.3 Bacterial growth media

For *Mycobacteria* strains:

Liquid Media: Middlebrook 7H9 broth (Difco; BD) supplemented with 10% OADC and 0.5% Tween-80

Solid Media: Middlebrook 7H10 agar (Difco; BD) supplemented with 10% OADC

Antibiotics: wild type strains were grown in the presence of 50 μg/ml Kanamycin. Δcpn60.1 strains were grown in the presence of both 50 μg/ml Kanamycin and 100 μg/ml of Hygromycin.

For *E. coli* :

LB was used for both solid and liquid culturing.

Antibiotics: *E. coli* transformants were grown in the presence of 50 μg/ml Kanamycin.

Supplements: MGM100 strains containing the pBAD promoter were grown in the presence of 0.2% arabinose

2.4 *M. marinum* growth curves

Frozen -80 stocks of *M. marinum* strains were streaked on 7H10 agar plates supplemented with OADC and appropriate antibiotics and grown at 30 °C until single colonies formed. In 25 mL universal tubes, 5 mL 7H9 growth media was inoculated with a single colony of *M. marinum* and grown until saturated. Cultures were diluted to an OD_{600} of 0.05 in 50 mL of media and grown in 250 mL shake flasks. At 3 time points, with 3-hour intervals, throughout the working day, 1 mL samples were taken and the OD_{600} measured using a spectrophotometer (Ultrospec 2100 Pro; GE Healthcare).

2.5 Heat-killing assay

5 mL cultures of M. marinum were grown to an OD_{600} of 0.2 at 30 °C before a serial dilution of each culture from 10^{-1} to 10^{-6} was performed in 1.5 mL eppendorf tubes using 7H9 supplemented with OADC and 0.05 % Tween-80 to give a final volume of 900 µl per tube. Tubes were placed at 30 °C, 100 RPM for 4 hours. An 8 µl sample from each tube was spotted onto a square 7H10 OADC antibiotic agar plate. Tubes were then placed in a heated water bath and samples taken at 30, 60, 120 and 180 minutes and spotted onto agar plates as previously described. Agar plates were then incubated at 30 °C and grown for 7 days.

2.6 Plasmid DNA extraction

Plasmid containing strains were grown in LB supplemented with appropriate antibiotics (and 0.2% arabinose for MGM100 strains) overnight at 37 °C, 180 RPM in a shaking incubator. Plasmid DNA was then extracted using either a QIAprep Spin Mini Kit (QIAGEN) or GeneJET Plasmid Miniprep Kit (Thermo Scientific). Plasmid DNA was quantified using a nanodrop 2000.

2.7 Heat shock transformation of *E. coli*

E. coli DH5α cells were made competent according to the standard protocol:
A 5 mL overnight culture of DH5α was pelleted by centrifugation (4000 RPM, 20 minutes, 4 °C). Cells were resuspended in 10 mL of ice-cold 0.1 M $CaCl_2$ and incubated on ice for 20 minutes. Another centrifugation step (with the same settings as before) was performed and cells were resuspended in 5 mL of ice-cold 0.1 M $CaCl_2$/15 % Glycerol (vol/vol) and separated in to 200 µl aliquots in sterile 1.5 mL eppendorf tubes. Cells were stored at -80 °C for future use.

In a sterile eppendorf tube 2 µl of plasmid DNA (H_2O used as a negative control) was added to 100 µl of competent cells, mixed gently by tapping the tube and then incubated on ice for 30 minutes. Tubes were then heat-shocked at 42 °C for 60 seconds and immediately replaced on ice. After 3 minutes, 800 µl of LB was added to each tube and mixed by gently pipetting before incubation at 37 °C, 180 RPM for 1 hour. After incubation 100 µl of cells were pipetted onto an LB agar plate supplemented with 50 µg/ml Kanamycin and spread using sterile glass beads. Plates were incubated at 37 °C overnight.

2.8 Transformation of *M. marinum*

M. marinum strains were grown in 100 mL of 7H9 supplemented with OADC, 0.05% Tween-80 and appropriate antibiotic to mid-log phase. Cells were harvested by centrifugation at 5000 rpm for 15 minutes. The supernatant was discarded before pelleted cells were resuspended in 25 mL of ice-cold 10% (w/v) glycerol solution. The resuspended cells were centrifuged (5000 rpm, 15 minutes). Two more centrifugation steps were performed under the same conditions with cells resuspended in 12.5 mL and then 7.5 mL of 10% glycerol solution. After the fourth centrifugation step the cell pellet was resuspended in 1 mL of 10% glycerol. Cells were then divided into 200 µl aliquots and stored at -20 °C.

Cells were thawed on ice shortly before use. Up to 1 µg of plasmid DNA was added to the cells, mixed gently by pipetting and incubated on ice for 10 minutes. Samples were then transferred to 0.1 mm electroporation chambers and subjected to a pulse of 2.5 kV using an electroporator (Eppendorf Eporator). 1 mL of 7H9 (+ OADC & 0.05% Tween-80) was immediately added to the transformed cells. The sample was transferred to a sterile 1.5 mL eppendorf tube and grown at 30 °C for 24 hours. After incubation, 100 µl of cells were transferred to a 7H10 agar plate (+ OADC, appropriate antibiotics), and spread using glass beads before being incubated at 30 °C until colonies appeared.

2.9 Preparation of glycerol stocks

Cultures of bacteria were grown to late-log phase (overnight for *E. coli* and for 1 week for *M. marinum*). Cultures (1.5 mL) were centrifuged for 5 minutes at 14000 RPM to pellet cells. Cells were then resuspended in 1.2 mL of 15% (vol/vol) glycerol and 300 µl of the standard growth media of the bacteria (7H9 for *M. marinum*, LB for *E. coli*) and stored at -80 ºC.

2.10 RNA Extraction

5 M GTC Solution:

- 5 M guanidine thiocyanate
- 0.5 % Sarkosyl
- 25 mM tri-sodium citrate, pH7
- 0.1 M β-mercaptoethanol
- 0.5 % Tween-80

A 150 mL culture of *M. marinum* (7H9 + OADC + 0.05 % Tween-80) was grown at 30 °C to mid-log phase. A 10 mL sample of culture was taken and transferred to a sterile 50 mL Falcon tube to which 40 mL of 5 M GTC solution was added. The sample was mixed by inverting the tube 4-6 times. The remaining cultures were incubated at either 30 °C (normal conditions) or 45 °C (heat-shock conditions), with 10 mL samples taken at 10 minute, 30 minute and 1 hour time points. All samples were prepared as before with the addition of 40 mL 5 M GTC solution. Samples were then centrifuged at 3000 RPM for 30 minutes to pellet the cells. The supernatant was discarded and pellets were resuspended in 1 mL of 4M GTC solution and transferred to a 1.5 mL eppendorf tube. Tubes were centrifuged for 1 minute at 13000 RPM and the supernatant was discarded. Pellets were resuspended in 1.2 mL of TRI reagent (Sigma Aldrich) and transferred to a 2 mL screw-cap tube containing 0.1 mm ceramic beads (Lysing Matrix B, MP Bio). Tubes were then placed in a reciprocal shaker (Ribolyser Homegeniser, Hybaid), which was run at 6.5 speed for 45 seconds. Tubes were then left at room temperature for 10 minutes before the addition of 200 µl of chloroform and then mixed by vortexing for 30 seconds. Samples were incubated at room temperature for 10 minutes before centrifugation at 13,000 RPM, 4 °C for 15 minutes. The newly formed upper phase of

each sample was transferred to a fresh eppendorf tube and 1 vol of chloroform was added before vortexing and incubation at room temperature for 10 minutes. Centrifugation (13,000 RPM, 4 °C, 15 minutes) separated the samples into two phases. The upper phase to a new eppendorf and 0.8 vols of isopropanol was added. Samples were incubated at -20 °C overnight and then centrifuged (13,000 RPM, 4 °C, 15 minutes). Supernatant was discarded and the pellet was washed with 70% ethanol before re-pelleting (13,000 RPM, 4 °C, 15 minutes) and air-dried to remove excess ethanol. Pellets were resuspended in 100 µl of nuclease-free water. RNA samples were subject to RNA clean-up using an RNeasy mini kit (QIAGEN). RNA was quantified using a UV-vis spectrophotometer (Nanodrop 2000, Thermo Scientific).

2.11 cDNA synthesis

cDNA was created from RNA samples using a Tetro cDNA Synthesis Kit (Bioline) according to manufacturers instructions. Random hexamers were used as the primers. RNA volume was adjusted for each sample to normalize the concentration across all reactions. For DMSO gradient experiments volume of DEPC-treated water was adjusted to accommodate for any additional DMSO ensuring a total reaction volume of 20 µl was constant.

Incubation steps were performed in a thermal cycler (Mastercycler Pro-S, Eppendorf) using the following programme:

Step	Temperature (°C)	Time (minutes)	Cycles
Annealing	25	10	1
Reverse Transcription	45	30	1
Reverse Transcriptase denaturation	85	5	1
Hold	4	∞	n/a

cDNA concentration was quantified using the dsDNA HS Assay Kit on a Qubit 2.0 fluorimeter (Life Technologies).

2.12 gDNA extraction

GTE solution:

- 50 mM glucose
- 25 mM Tris-HCl, pH 8
- 10 mM EDTA

A culture of *M. marinum* was grown to mid-log phase. Cells (2 mL) were harvested by centrifugation (14,000 RPM, 5 minutes). Cell pellet was resuspended in 450 µl GTE + 50 µl lysozyme (10 mg/ml) and incubated at 37 °C overnight. After incubation 100 µl of 10 % SDS and 50 µl of proteinase K was added to the tubes. Samples were then incubated at 55 °C for 30 minutes. NaCl (5M; 200 µl) was then added to each sample before the addition of 1 mL chloroform, vortexing and centrifugation at 8000 RPM for 6 minutes. The newly formed upper phase was transferred to a fresh eppendorf tube and 0.7 vol of isopropanol was added before incubation at room temperature for 10 minutes. Tubes were then spun at 8000 RPM for 10 minutes and the supernatant was discarded. The pellet was washed with 70% ethanol and centrifuged at 8000 RPM for 10 minutes. Pellets were air dried at 65 °C for 15-20 minutes before resuspension in 100 µl of nuclease-free water.

gDNA concentration was quantified using the dsDNA HS Assay Kit on a Qubit 2.0 fluorimeter (Life Technologies).

2.13 Polymerase Chain Reaction (PCR) amplification of DNA

PCR was performed using either cDNA or gDNA. Initial primer annealing temperature optimization was performed using Phusion polymerase (NEB), combined with the provided 5x High GC Buffer and 3% DMSO. Reactions were prepared in 1.5 mL eppendorf tubes and aliquoted into 0.2 mL PCR tubes.

Step	Temperature (°C)	Time (minutes)	Cycles
Denaturation	95	05:00	x1
Denaturation	95	00:20	x40
Annealing	50-70	00:30	
Extension	60	00:30	
Hold	4	∞	n/a

PCR products were run on a 2% agarose gel stained with Midori green.

Subsequent reactions were performed using SYBR green master mix according to the qPCR protocol.

2.14 Quantitative Real-Time PCR (qRT-PCR)

Primer design:

Primers for qRT-PCR are integral to the success of the experiment. Non-specific binding of primers will amplify unwanted regions of DNA giving a fluorescent signal non-representative of GOI amplification. Dimerisation and secondary structures of primers will also bind SYBR green and therefore increase fluorescence.

Following the primer design protocol outlined in (Thornton & Basu, 2010), Primer3 primer design software (available at: http://frodo.wi.mit.edu/primer3/; Rozen *et al,* 2000) was used to identify the best primer pair for each gene sequence (obtained through MarinoList at mycobrowser (available at: http://mycobrowser.epfl.ch/), Kapopolulou, Lew & Cole, 2011). Obtained primer pairs were then inputted into Beacon Designer™ Free Edition (available at: http://free.premierbiosoft.com) to check for potential primer self-dimerisation, cross-dimerisation and secondary structure. Any primers that formed any structure with ΔG values <-3.5 kcal/mol were rejected. To check for amplicon secondary structure the program, mFold (available at: http://www.idtdna.com?Scitools/Applications/mFold;), was used. Due to the relatively high GC content of the *M. marinum* genome and the short length of *cpn10* gene, amplicon secondary structures were unavoidable without compromising on other, more essential primer parameters. Finally nucleotide BLAST (available at: *www.ncbi.nlm.nih.gov/BLAST*; Lobo, 2008) searches were performed against the *M. marinum* genome to check primer specificity. Primers with non-specific priming targets were discarded and redesigned.

Primers for qPCR were designed by following the protocol described by Thornton and Basu, 2010, and were synthesized by Illumina. Reaction master mix for each primer pair were prepared on ice in a sterile eppendorf tube, with the following composition:

Total volume per well = 20 µl

- 10 µl Brilliant II SYBR Master Mix (Agilent Technologies)
- 1 µl forward primer
- 1 µl reverse primer
- 1 µl DMSO
- x amount of DNA

Volume adjusted to 20 µl using nuclease-free H_2O

Master mix was aliquoted into wells on a 96-well plate, non-skirted (Agilent Technologies) and kept at 4 °C using a PCR-cooler (Eppendorf). DNA was then added to each well. Each reaction was performed in duplicate with no template controls (NTCs) for each primer pair. The 96-well plate was sealed using optical caps (8x strip; Agilent Technologies). Plates were centrifuged (Sigma 4-15C, QIAGEN) at 4000 RPM for 3 mins. Reactions were performed in an MX3005P qPCR system (Agilent Technologies).

Thermal profile was conFig.d as below:

Step	Temperature (°C)	Time (minutes)	Cycles
Denaturation	95	05:00	x1
Denaturation	95	00:20	x40
Annealing	56	00:30	
Extension	60	00:30	
Denaturation	95	01:00	x1
Annealing	56	00:30	
Denaturation	95	00:30	

2.15 Restriction digests

Restriction digests were performed at both 20 µl and 50 µl reaction volumes. All restriction enzymes used were NEB. Reactions were performed according to manufacturer's protocol.

Component	20 µl reaction	50 µl reaction
10x Buffer	2 µl	5 µl
Restriction Enzyme	1 µl	2 µl
DNA	~400 ng	~1 µg
dH_2O	To 20 µl	To 50 µl

2.16 Gel electrophoresis

Depending on the expected size of the PCR product either 1% or 2% agarose gels were run:
For products as a result of qPCR primers, 2% gels were achieved by dissolving 2 g of agarose in 100 ml of 1x TAE buffer and heating until dissolved. Midori green direct stain (5 µl per 100 ml of agarose; Nippon Genetics Europe) was used for Phusion polymerase based reactions. For SYBR green based reactions no stain was added to the gel. DNA samples were prepared by mixing 3 µl 5x DNA loading buffer blue (Bioline) with 5 µl nuclease-free H_2O and 7 µl of PCR product. Hyperladder IV (100 bp; Bioline) was used as a marker.

For restriction digest products 1% gels were made by dissolving 1 g of agarose in 100 ml of 1x TAE buffer and stained with Midori green.
Gels were run at 120 V, 500 mA for 40 mins (small gels) and 1 hour (large gels).

2.17 Gel-Purification of DNA

Gels were visualized using a near-UV LED viewer (Helixx UltraSlim LED viewer; Helixx). Bands of interest were excised using a sterile scalpel and transferred to a pre-weighed eppendorf tube. Tubes were re-weighed and the weight of the gel slice determined. DNA was then purified using a QIAquick Gel Extraction Kit (Qiagen) according to manufacturer's protocol. Purified DNA was quantified using a nandrop 2000.

2.17 Lipid extraction

A single colony of *M. marinum* was used to inoculate 5 mL of 7H9 media supplemented with OADC and antibiotic and grown for 7 days or until late-log phase was reached. Smaller cultures were then used to inoculate 200 mL of 7H9 media supplemented with OADC and antibiotic in a 1 L shake flask. Large cultures were grown at 30 °C , 100 RPM until saturated. Culture (50 mL) was transferred to a 50 mL falcon tube and cells were harvested by centrifugation at 4000 RPM for 30 minutes. Supernatant was discarded and more culture was added to the same tube to 50 mL before re-harvesting. This was repeated until all cells had been harvested. Cells were resuspended in 10 mL 1x Phosphate Buffered Saline (PBS), and transferred to a 15 mL borosilicate screw neck, rounded base tube (125 mm x 16 mm; Pyrex, Appleton Woods). Tubes were spun at 3000 RPM for 15 minutes before discarding supernatant and drying pellets for 1.5 hours in a heating block set to 50 °C under a stream of nitrogen gas. 50-100 mg of dried biomass was transferred to a fresh borosilicate tube (referred to onwards as tube A) and resuspended in a solution of Methanol and 0.3% NaCl mixed in the ratio 10:1 followed by 1 ml of petroleum ether (60-80 C). Tubes were mixed on a rotor and placed in an ultrasonic cleaner (need model info) to ensure complete resuspension. Tubes

were then centrifuged at 3000 RPM for 15 mins and the newly formed upper layer transferred to a pre-weighed borosilicate tube (referred to onwards as tube B). 1 mL of petroleum ether (60-80 C) was added to the remaining lower layer of Tube A, mixed on a rotator for 15 minutes and then centrifuged for 5 mins at 3000 RPM. The upper layer was transferred to tube B, which was then dried on a heating block at 50 °C under nitrogen gas. The weight of tube B was recorded and the mass of non-polar lipid weight was calculated by subtracting the empty weight of tube B previously recorded. The dried non-polar lipids were resuspended in 200 µl of 2:1 chloroform:methanol

2.3 mL of a 9:10:3 chloroform:methanol:0.3% NaCl solution was added to the remaining contents of tube A, which was then mixed on a rotator for 60 minutes and then centrifuged for 5 minutes at 3500 RPM. The supernatant was then transferred to another pre-weighed borosilicate tube (named tube C). A 5:10:4 solution of chloroform:methanol:0.3% NaCl (750 µl) was added to tube A, mixed on a rotator for 30 minutes and centrifuged for 3 minutes at 3500 RPM. The supernatant from tube A was then transferred to tube C. The previous 5:10:4 chloroform:methanol:0.3% NaCl step was repeated again and the supernatant added to tube C. To tube C 2.6 mL of a 1:1 chloroform:0.3% NaCl solution was added, mixed on a rotator and centrifuged for 5 mins at 3500 RPM. The bottom phase consisting of the purified polar lipids was transferred to a clean pre-weighed borosilicate tube (tube D) and dried at 50 °C under nitrogen gas before resuspending dried material in 200 µl of a 2:1 chloroform:methanol solution. Tube D was weighed and mass of polar lipids calculated by subtracting the empty weight of the tube.

2.19 Two-Dimensional Thin Layer Chromatography (2D-TLC)

Aluminium TLC plates (20x20 cm; aluminium oxide matrix; Sigma Aldrich) were cut into 6x6 cm size squares (36 cm^2). A small pencil mark was placed 1 cm from the horizontal and vertical edges in one corner of the plate. Plates were labeled using a paper along the edge of the plate below the pencil mark. Four plates were required per sample of non-polar lipids and 2 plates for each sample of polar lipid.Six solvent systems were run in total, 4 for non-polar and 2 for polar lipids.

Lipid type	System	Direction	Constituents
Non-polar	A	1	Petroleum ether:ethyl acetate (98:2)
		2	Petroleum ether:acetone (98:2)
	B	1	Petroleum ether:acetone (92:8)
		2	Toluene:acetone (95:5)
	C	1	Chloroform:methanol (96:4)
		2	Toluene:acetone (80:20)
Non-polar & Polar	D	1	Chloroform:methanol:water (100:14:0.8)
		2	Chloroform:acetone:methanol:water(50:60:2.5:3)
Polar	E	1	Chloroform:methanol:water (60:30:6)
		2	Chloroform:acetic acid:methanol:water(40:25:3:6)

Purified lipids were normalized by diluting to a concentration of 10 µg/µl using 2:1 chloroform:methanol. A 10 µl sample of lipid was drawn into a pipette and spotted onto the pencil mark on the plate by capillary action, with the spot diameter not exceeding ~5 mm. Spots were allowed to dry before the sample was re-spotted and this was repeated until all 10 µl of sample was spotted on the plate.

Solvent solutions were freshly made before use. Solvent solution (50 mL) was placed in a glass TLC tank, which was then closed with a lid and allowed to equilibrate for 15 mins. After 15 mins the lid was opened slightly and the TLC plates placed in the tank with ~0.5 cm of the plate submerged in the solution. The tank lid was replaced and the TLC plates removed once the solvent front reached the top of the plate. At this point the plates were removed and left to dry at room temperature. The solution was discarded in the appropriate manner before the tank was washed with acetone and dried. Certain solvent systems required more than one "direction 1" or "direction 2" run using the same solution. (see Table for details). The next solvent solution "direction 2" was placed into the tank and this was again allowed to equilibrate with the lid on. Once dried the plates were then placed in the tank, rotated about 90 degrees with the original lipid spot nearest the solvent solution.

System	Direction	Number of Runs
A	1	3
	2	1
B	1	3
	2	1
C	1	1
	2	1
D	1	1
	2	1
E	1	1
	2	1

This method applied to all solvent systems with the exception of System E, which required the baking of the TLC plates for 15 minutes at ~90 °C prior to running the direction 1.

TLC plates were stained with (need to find out) and developed by heating with a heat gun.

2.20 Biofilm formation assay (from Rao, 2010)

Biofilm base media:

- 100 mM Monopotassium Phosphate KH_2PO_4
- 15 mM $(NH_4)_2SO_4$, pH 7.0 (adjusted using KOH)
- 0.5 mg Iron Sulphate Heptahydrate $FeSO_4.7H_2O$
- 5 g Casamino acid

Biofilm media:

- 94 ml Biofilm base media
- 5 ml 40% glucose
- 1 ml 0.1 M Calcium Chloride $CaCl_2$
- 100 µl 1 M Magnesium Sulphate $MgSO_4$

10 ml of Biofilm media was aliquoted into a small petri dish (60x15 mm). 10 µl of a mid-log phase culture of *M. marinum* or *M. smegmatis* was added to the centre of the petri dish and incubated at 30°C for 7 days.

2.21 Pellicle formation assay

To sterile test tubes, 5 mL of 7H9 media with OADC (no Tween-80) was aliquoted. Samples (1 mL) of mid-log phase cultures of *M. marinum* and *M. smegmatis* were transferred to autoclaved eppendorf tubes and centrifuged for 2 minutes at 1000 RPM to pellet clumped cells. Supernatant (10 µl) was taken from each sample and pipetted into their respective test tube. Tubes were incubated at 30 °C for 7 days, after which they were photographed and length of pellicle formation was measured in mm using a ruler.

2.22 Investigation of cell clumping characteristics

Centrifugation, needle-pass and further centrifugation test:

Cultures of *M. marinum* wild type and Δ*cpn60.1* were grown to either early or late log phase. Samples (1.5 mL) were taken from each culture and transferred to a sterile eppendorf tube. Tubes were centrifuged for 2 minutes at 1000 RPM to pellet cell clumps. Supernatant (100 µl) was used to establish a serial dilution from 10^{-1} to 10^{-4} of which 100 µl were plated on 7H10 agar and spread using sterile glass beads. The OD_{600} of the remaining supernatant was measured for each sample. Using a 23G needle and a sterile 5 mL syringe, 5 mL of the original culture was passed in and out of the syringe 3 times. Another serial dilution using the same range as before was established and plated on 7H10. The OD_{600} of the syringed cells was recorded. Samples of the syringed cells (1.5 mL) were taken and transferred to eppendorf tubes. Tubes were centrifuged for 2 mins at 1000 RPM and 100 µl of supernatant taken to establish a serial dilution, which was then plated as previously described. The OD_{600} of the supernatant was recorded. All plates were incubated at 30 °C for 7 days or until colonies were visible.

Needle-pass experiment:

Two 5 mL cultures of wild-type and Δ*cpn60.1 M. marinum* were grown to mid-log phase. Using a 23G needle and 5 mL syringe the samples were passed in and out of the syringe 4 times. The OD_{600} was taken after each pass and 100 µl of cells were used to create a serial dilution to 10^{-6} from each pass. Three samples of each strain (10^{-4} to 10^{-6}) were plated on 7H10, spread using sterile glass beads and incubated at 30 °C for one week, or until colonies formed.

Hemocytometetry:

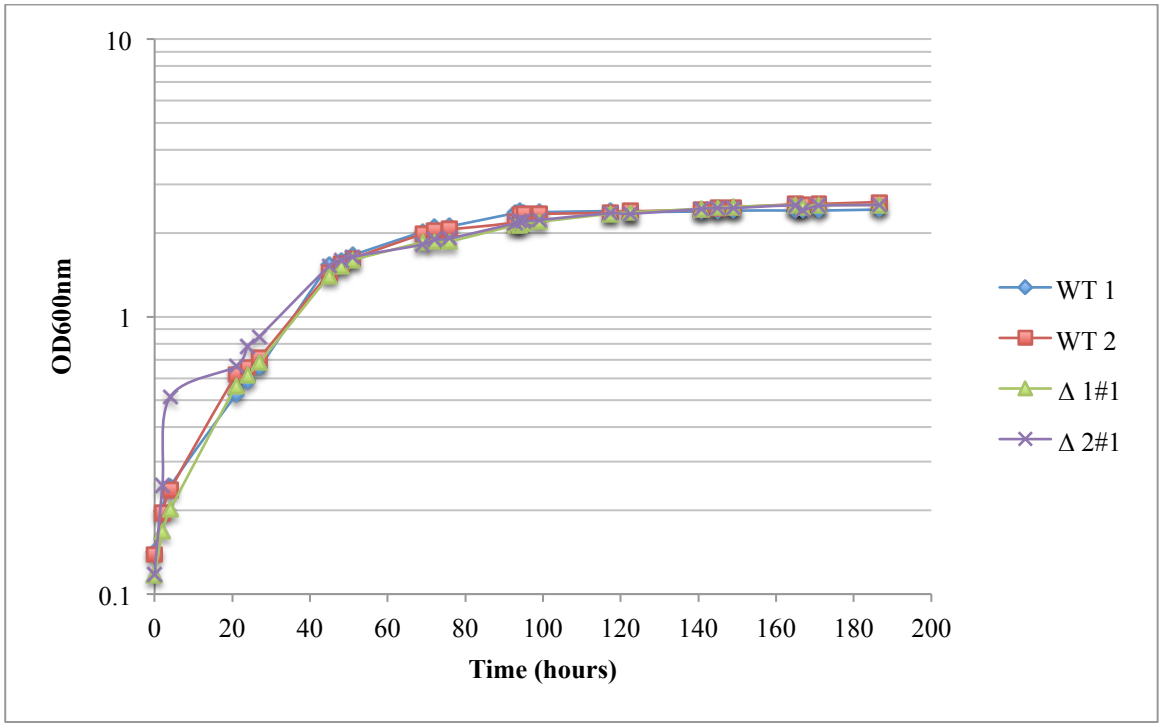
Two 5 mL cultures of wild-type and Δ*cpn60.1 M. marinum* were grown to mid-log phase. An initial 100 µl of each culture was taken and placed in a sterile eppendorf tube. Using a 23G needle and 5 mL syringe the remaining cultures were passed through a syringe needle. Samples (100 µl) were taken in between each needle pass. Each sample was used to create a serial dilution to 10^{-6} from each pass. Samples (20 µl) were pipetted into the hemocytometer under the cover slip and observed using the x40 objective lense of a microscope (Alphaphot YS2-T; Nikon) using the phase contrast setting Ph4. Total cells found in each corner quadrant of the hemocytometer were counted using a hand tally counter clicker.

3. RESULTS

3.1 Growth rates of wild-type and Δ*cpn60.1 M. marinum*

The first experiment performed was to establish growth curves for Δ*cpn60.1* strains of and wild type *M. marinum*. This was required to ensure that difference in growth rate didn't impact any future growth-based experiments. Two isolates of each strain were used; wild type 1 (wt1), wild type 2 (wt2), Δ*cpn60.1* 1#1 and Δ*cpn60.1* 2#1. Both of these strains had previously been transformed with the plasmid pMSP-dsRed (kanaR) for fluorescent detection in zebrafish (*Danio rerario*) embryo infection models. All strains were streaked on to 7H10 agar from frozen -80 °C glycerol stocks and selected for using the appropriate antibiotics. Single-colony liquid cultures were then grown and used to inoculate a larger liquid culture to a starting OD_{600} of 0.05. The effect of diluting *M. marinum* on OD_{600} was not linear -values calculated using the equation $c_1v_1=c_2v_2$ were incorrect by a factor of 4. To achieve the correct dilution, calculated values were divided by 4 and this volume of cell culture was used to inoculate. This effect was due to the high frequency of cell clumping of *M. marinum* at higher OD_{600} values (discussed later). Readings were taken three times a day with ~3 hour intervals.

Fig.3. compares the growth rates of all *M. marinum* strains used. No significantly different rate in growth can be observed between the strains meaning there is no growth bias that would affect growth-based experiments. Reproducibility of this experiment was an issue. The rates of growth appeared to vary between repeats, with the wild-type cells entering log-phase slightly sooner than



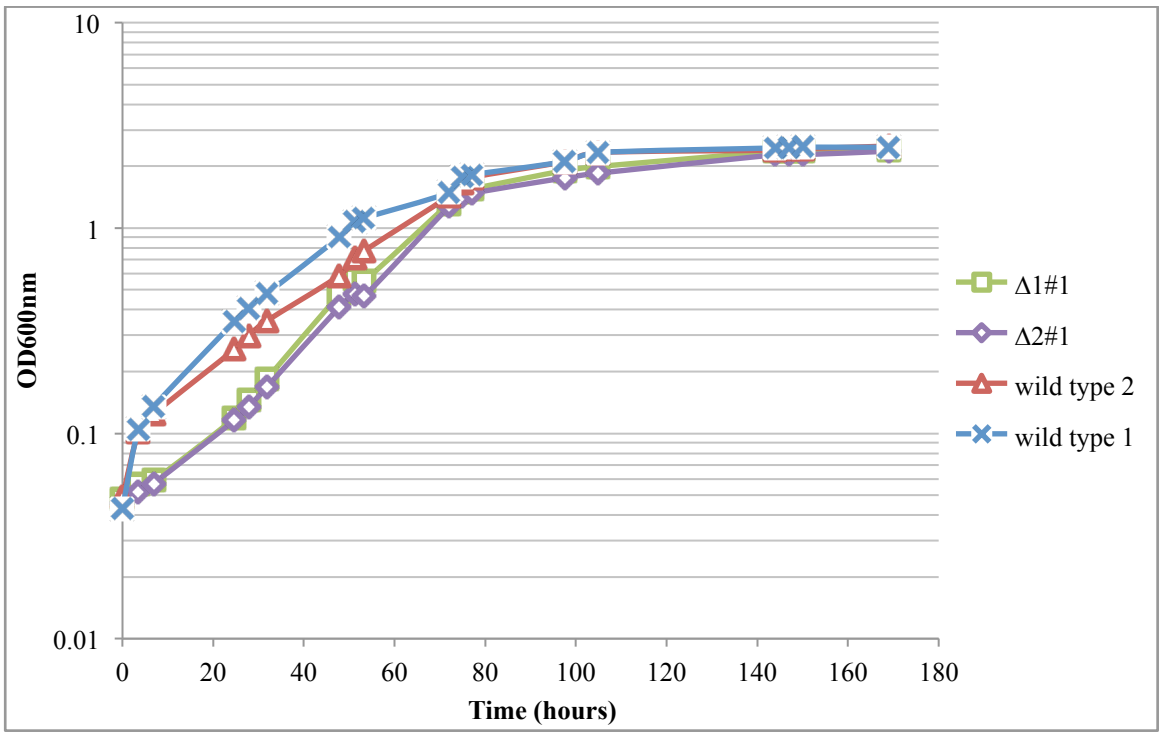


Fig. 3.1 Growth rates of *M. marinum* wild type and Δ*cpn60.1* strains. (Top) Run 1, issues were had achieving the correct starting $OD_{600\ nm}$, therefore this run started from an $OD_{600\ nm}$ of ~0.1 for each strain. (Below) Run 2, all strains were diluted to a starting $OD_{600\ nm}$ of ~0.05. Readings were taken at 3 time points roughly 3 hours apart per day.

3.2 Heat-killing experiments

Previous experiments have documented the effect of cpn60.1 loss on temperature sensitivity. Δ*cpn60.1* mutants of *Mycobacterium tuberculosis* were shown to have an increased sensitivity at 55 °C, with this temperature killing the knockout faster than the wild-type strain (Hu *et al.,* 2008). In *M. smegmatis* there is a >100-fold up-regulation of *cpn60.1* expression in response to heat shock at 42 °C (Rao and Lund, 2010). To investigate the role of *cpn60.1* during the heat shock response, both wild-type and Δ*cpn60.1 M. marinum* were exposed to three different temperatures above that of their optimal growth temperature of 30 °C; 37 °C, 45 °C and 50 °C. Samples were taken at 30 minutes, 1 hour, 2 hours, 3 hours and grown on 7H10 media for one week at 30 °C.

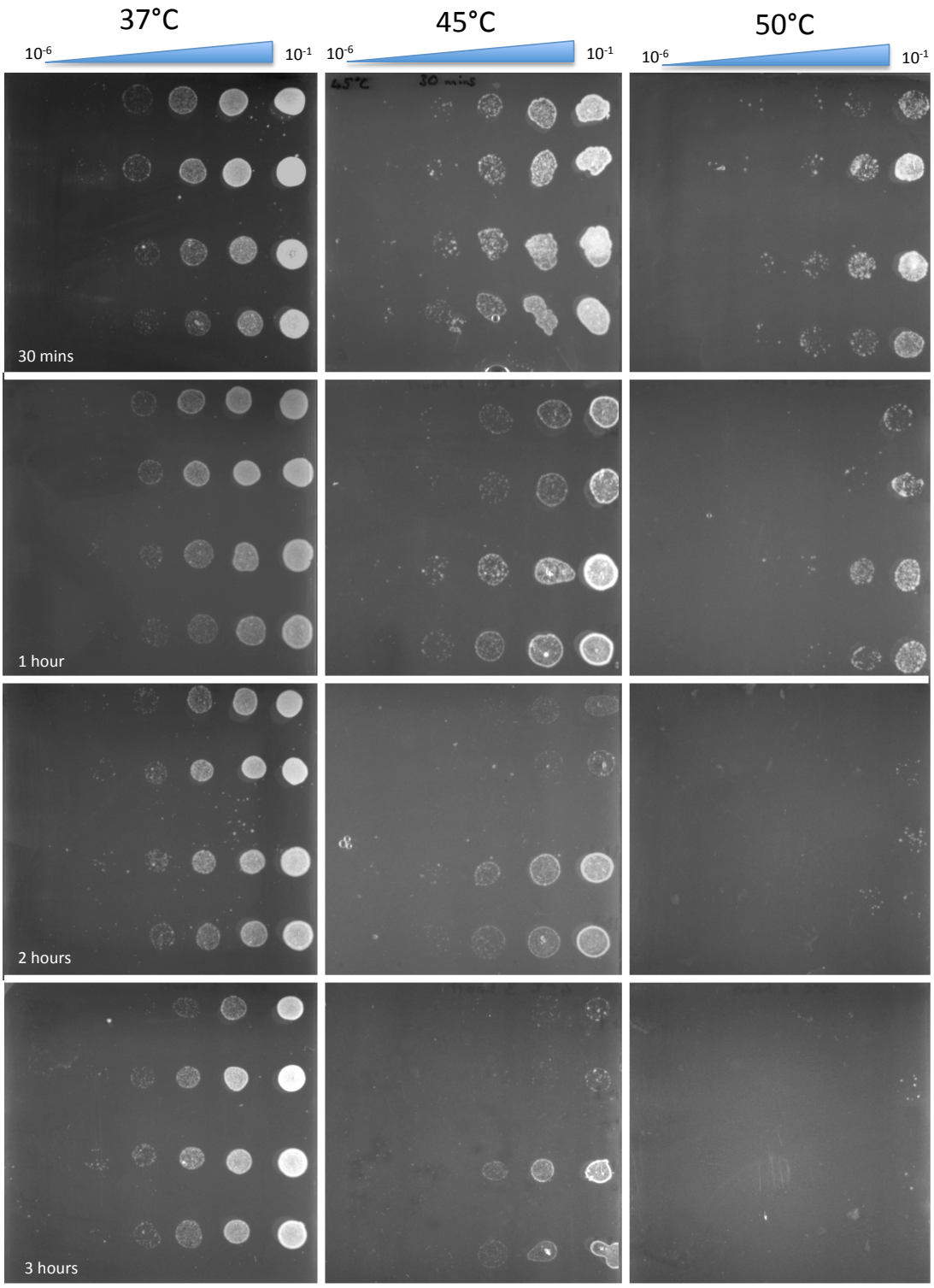


Fig.3.2 Effect of heat shock on the survival of *M. marinum* wild type and Δ*cpn60.1* strains. On each plate top row = wild type 1, second row from top = wild type 2, third row = Δ*cpn60.1* 1#1, bottom row = Δ*cpn60.1* 2#1

Fig.3.2 shows the survival of *M. marinum* strains at different heat shock temperatures. At 37 °C no change is observed between wild type and Δ*cpn60.1* strains after 3 hours. At 45 °C the Δ*cpn60.1* strains show a clear survival advantage against heat shock. After 30 minutes at 45 °C there is little difference between the strains. In the 1-hour samples we start to see an increase in the killing of the wild type strains and after 3 hours nearly all wild type cells have been killed whilst the Δ*cpn60.1* strains have survived. This trend is repeated in the results from the 50 °C exposure experiment except over a shorter timeframe with all strains completely killed after 2 hours.

3.3 Quantitive Real-Time PCR analysis of *cpn* gene expression

Quantitative real-time PCR is a method of measuring the level of transcription of genes in a given sample. It requires a reverse transcription step to convert RNA isolated from a sample to complementary DNA (cDNA). Like standard PCR DNA is then amplified using gene-specific primers and *Taq* polymerase. However, in qPCR the use of certain fluorescent dyes allows the reaction to be quantified in real-time. For this set of experiments SBYR green dye was used which emits a fluorescent signal in the presence of double-stranded DNA. As target DNA is amplified the fluorescence signal increases. Therefore by measuring fluorescence the relative amount of cDNA of the gene of interest (GOI) can be determined, and therefore it's level of transcription.

Following on from the heat-killing experiments we wanted to identify how the transcription of all three *cpn* genes (*cpn10, cpn60.1* and *cpn60.2*) changes in response to heat shock and their difference in expression between wild-type and Δ*cpn60.1* mutant. Expression of a housekeeping gene, in this experiment *gap* (MMAR_2239, Glyceraldehyde 3-phosphate dehydrogenase), needs to be measured in every sample to normalize the data between each.

3.3.1 Optimising qRT-PCR conditions

PCR was performed to check the validity of designed primers. Due to the unavoidable amplicon secondary structure, reaction conditions were varied to determine the optimal setup. The first experiment performed involved varying the primer annealing temperature from 55 °C to 75 °C. This experiment was performed using Phusion polymerase and included 3% DMSO. Fig. 3.3 shows the results of the annealing temperature variation experiment. Products were run on 1% agarose gels (*cpn10, cpn60.1* and *cpn60.2*) or 2% agarose (*gap*).

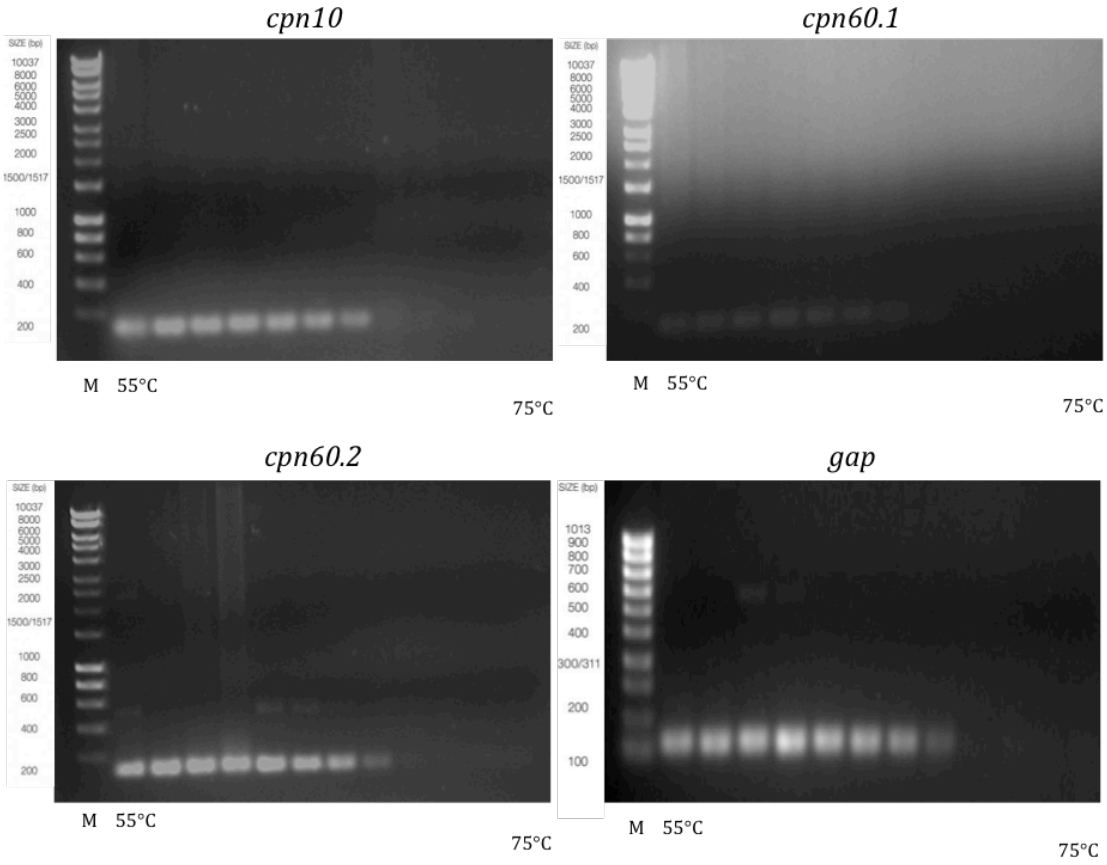


Fig.3.3 PCR products of qPCR primer amplification, varying annealing temp. A gradient of annealing temperatures of 55 °C to 75 °C was used. Expected product sizes: for *cpn10* – 87 kb, *cpn60.1* = 111 kb, *cpn60.2* = 107 kb, *gap* = 121 kb.

Primers for *cpn10* annealed at temperatures from 54.8 °C to 66 °C, but not above. Annealing temperatures from 54.8 °C to 68.6 °C were sufficient to amplify *cpn60.1*. However, both cpn60.2 and gap suffered from non-specific amplification at certain annealing temperatures. At 54.8 °C and between 60.7 and 63.3 °C primers for *cpn60.2* show a second band indicating amplification of a secondary product. Likewise, at 56.5 and 58.3 °C a second product is seen in the *gap* primer-amplified samples. An annealing temperature of 56 °C was chosen for future experiments, avoiding the production of secondary products produced by *cpn60.2* and *gap* primers.

3.3.2 Testing qRT-PCR primer efficiency

Genomic DNA (gDNA) was purified from a culture of *M. marinum* wild type 2. A serial dilution from 1x to 10^{-10} of gDNA was established and a run of qPCR performed to create standard curves of each primer pair.

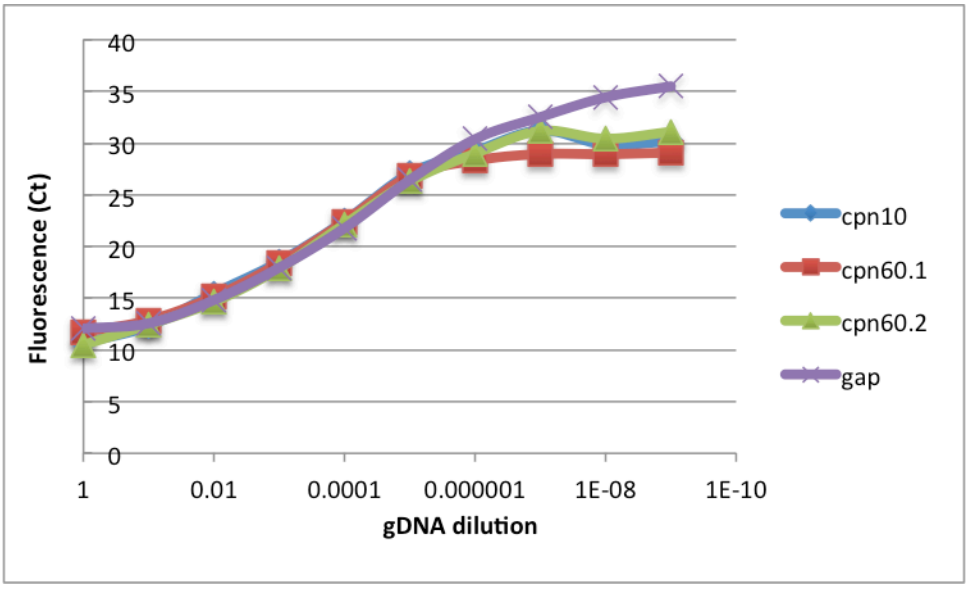


Fig.3.4 Fluorescence signal output of qPCR reaction on serial dilution of gDNA to test efficiency of each primer pair

Fig. 3.4 shows that the efficiencies of all primer pairs were very similar. DNA concentrations that fall within the linear portion of the standard curve are deemed ideal for further experiments, in this case a fluorescence value (Ct) in the range of ~15 to ~27. QRT-PCR products were run on 2% agarose (Fig. 3.5).

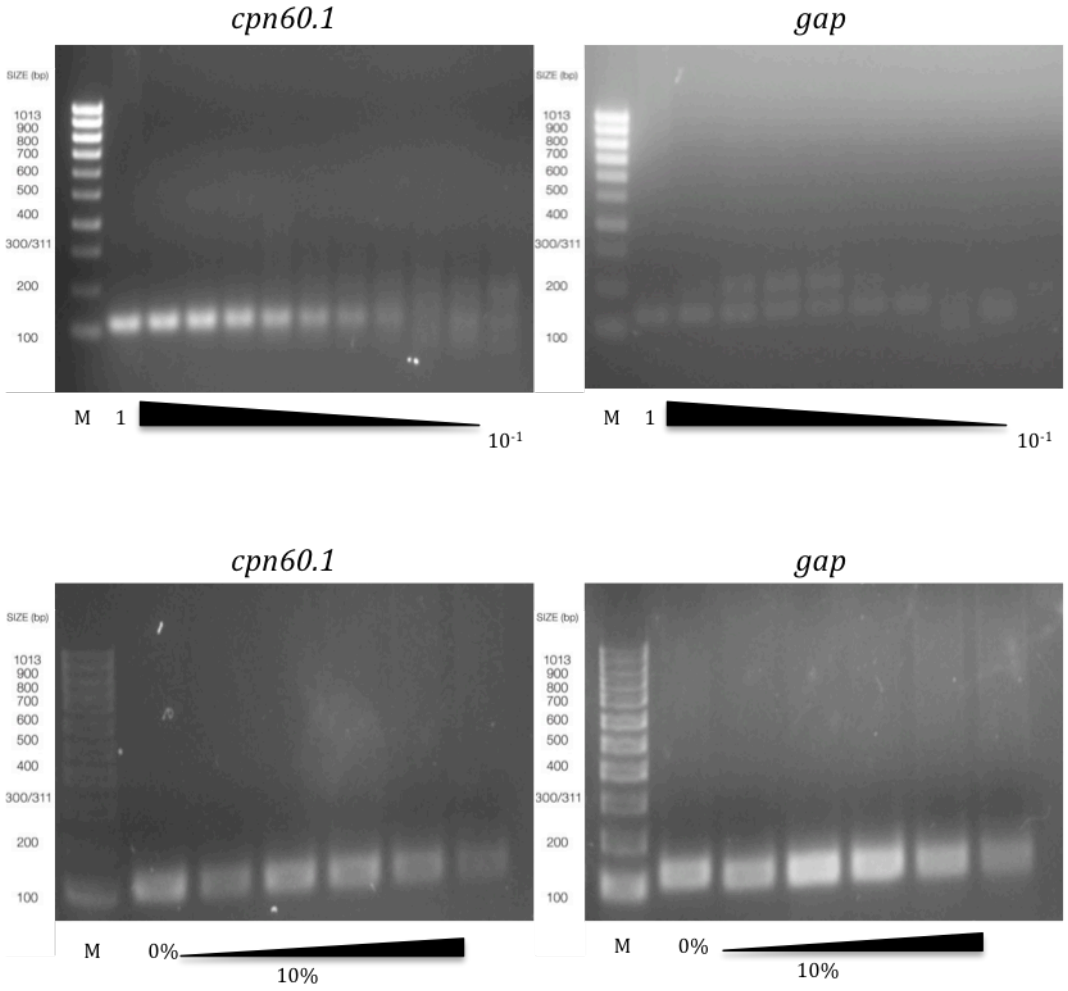


Fig. 3.5 (top) qPCR products from *cpn60.1* and *gap* primer pairs run on 2% agarose. (Bottom) the effect of varying DMSO concentration on *cpn60.1* and *gap* PCR product formation

Products formed from the *cpn60.1* primers (Fig.3.5, top) showed some smearing indicative of the presence of secondary structure in the amplicon. A non-specific product was formed in the amplification using *gap* primers, but not at all gDNA concentrations. The use of DMSO to reduce secondary structure in high GC content DNA templates has been well documented (Jensen, 2010; Mammedov, 2008). In the original primer optimization experiments DMSO

had been included in the reaction mixture at 3%. However, DMSO was not added to the qPCR reaction. To investigate whether addition of DMSO alleviated the smearing and non-specific amplification observed in *cpn60.1* and *gap* qPCR products, the PCR was repeated using a gradient of DMSO from 0% to 10% in 2% increments. Fig. 3.5 (bottom) shows that DMSO appeared to alleviate non-specific amplification in *gap* and no smearing is observed in the *cpn60.1* products.

3.3.3 RNA isolation from heat-shocked *M. marinum* strains

The results of the initial heat-shock experiments indicated that Δ*cpn60.1* mutants of *M. marinum* appeared to have a higher rate of survival under heat-shock conditions. At 50 °C the cells were killed between the 1 and 2 hour mark, but at 45 °C the wild type survived until the 3 hour mark, therefore 45 °C was chosen as the working temperature to investigate the change of expression of the *cpn* genes under heat-shock. Single cultures of wild type 1 and Δ*cpn60.1* 1#1 were grown to mid-log phase and split into two 50 mL cultures, each. One culture of each strain was kept at 30 °C to act as a control whilst the other was moved to 45 °C. Samples of the 30 °C cultures were taken after 1 hour to compare to the 0 minute sample to ensure there is no change in gene expression at the control temperature. Samples were taken from the 45 °C cultures at 0, 10, 30 and 60 minutes after exposure to heat-shock. GTC solution was immediately added to each sample after sampling. The main component of GTC, guanidine thiocyanate is a potent inhibitor of RNase when present at high concentrations, in this case 4 molar. This acts to prevent the degradation of the mRNA, providing the transcriptome of the culture at that specific point in time. RNA was purified (see methods) and cDNA synthesized adjusting RNA volume to normalize the concentration of starting template.

3.3.4 qRT-PCR of cDNA samples

A serial dilution of each cDNA sample was made ranging from 10^{-1} to 10^{-7}. As the efficiency of each primer pair had previously been established as roughly equal, preliminary experiments were performed with only the primer pair for *gap* amplification. Serial dilutions of cDNA were assayed to create a standard curve. Initially concentrations ranging from 1x to 10^{-7} were run, but were found to be too low, with most concentrations not lying within the linear portion of the curve, where Ct < ~27. A further run with 3 µl of each cDNA sample per well was performed, and although the Ct values obtained were within the acceptable range (Fig. 3.6), but the remaining volume of cDNA stock was too low for replicates of each primer pair to be carried out.

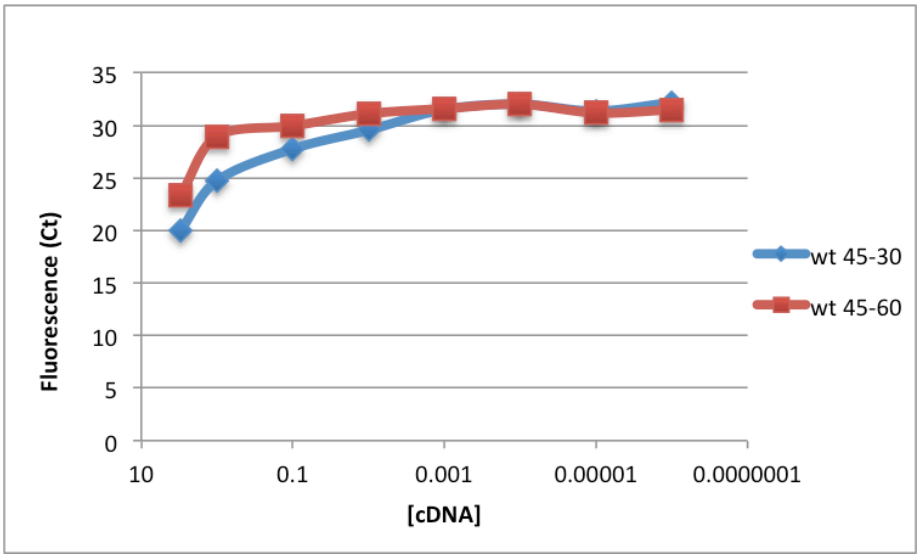


Fig.3.6 Standard curves of cDNA samples amplified using *gap* qPCR primers

3.3.5 Optimising cDNA synthesis

Due to the low concentration of cDNA, the reverse transcriptase (RT) protocol was revisited. By increasing the length of the reverse transcription step from 30 minutes to 60 minutes it was hoped that amount of product would be increased. The temperature of the RT step was also increased from 45 °C to 47 °C to help decrease any RNA template secondary structures. Resulting cDNA was diluted to 10^{-1} concentration and quanitified using a Qubit dsDNA high-specificity assay kit and compared to the cDNA previously created, which acted as a control. Fig. 3.7 shows that cDNA created using the altered protocol was much lower than the cDNA created using the standard protocol.

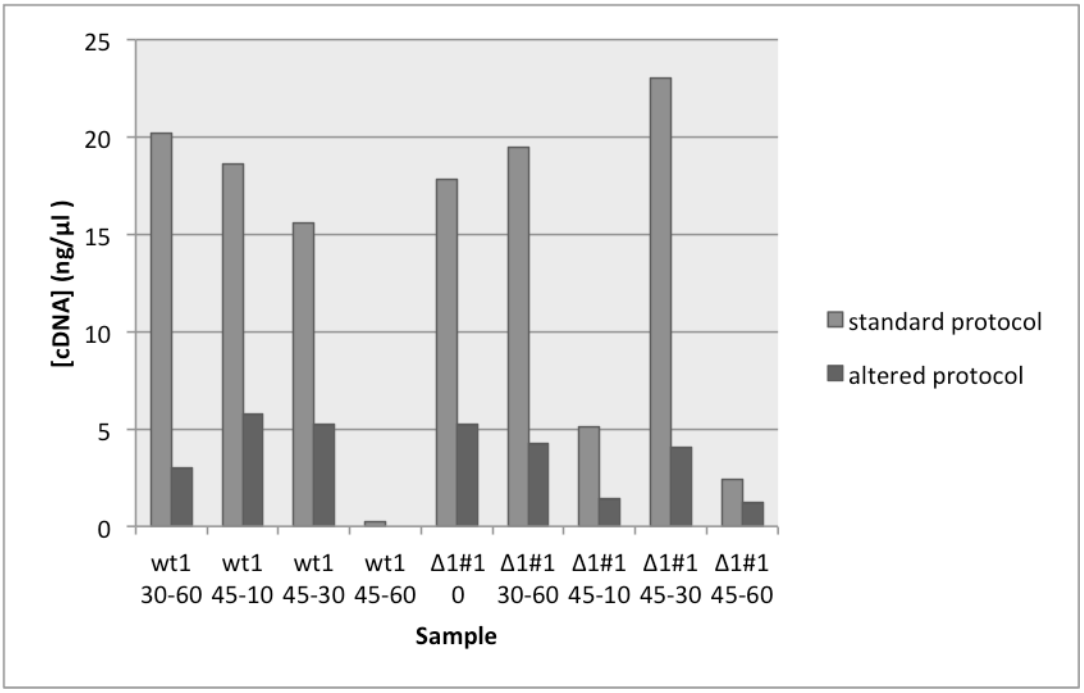


Fig.3.7 The effect of increasing the Reverse Transcriptase step duration and temperature on cDNA yield. From left to right *M. marinum* wild-type 1 samples: after 60 minutes at 30 °C (wt 1 30-60), after 10 minutes at 45 °C (wt1 45-10), after 30 minutes at 45 °C (wt1 45-30), after 60 minutes at 45 °C (wt1 45-60), Δ*cpn60.1* 1#1; at t= 0 minutes (Δ1#1 0), after 60 minutes at 30 °C (Δ1#1 30-60), after 10 minutes at 45 °C (Δ1#1 45-10), after 30 minutes at 45 °C (Δ1#1 45-30), after 60 minutes at 45 °C (Δ1#1 45-60).

Next the effect of adding DMSO and varying its concentration on cDNA yield was studied. RNA isolated from Δ*cpn60.1* 1#1 from 45 °C after 60 minutes was used and the reaction was run according to standard thermal profile settings. DMSO was varied at 3%, 6% and 9% (with a 0% control), DMSO concentration above 10% was avoided as at this concentration and above it starts to lower the annealing temperature of the reaction and inhibits *Taq* polymerase activity by 47% (Sun, 1993). Unlike before, cDNA was quanitified using the UV-vis spectrophotometer. Although not an entirely accurate measure due to the constituents of the cDNA synthesis reaction, the values are relative to one another allowing a simple comparison to be drawn.

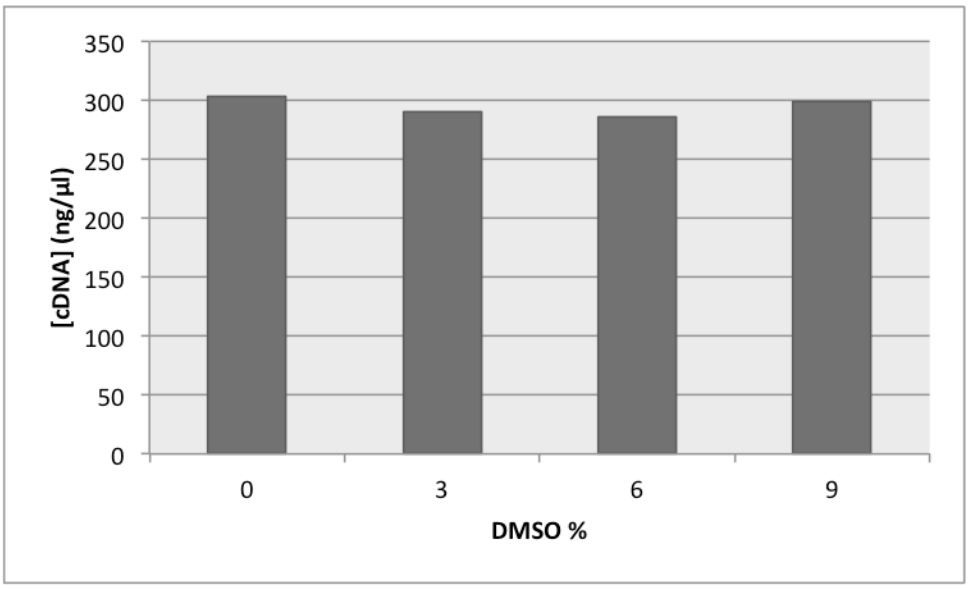


Fig. 3.8 The effect of increasing DMSO concentration on cDNA yield. Concentration of DMSO was varied in the reaction mixture of the cDNA protocol. Volume of ddH_2O was adjusted to maintain a constant final reaction volume.

The results (Fig.3.8) appear to indicate that varying DMSO has no real effect on the yield of cDNA achieved.

3.3.6 qRT-PCR of cDNA samples (cont.)

Due to time constraints it was decided that qRT-PCR would be performed using the cDNA synthesized initially using the standard cDNA synthesis protocol. Concentrations of cDNA used were normalized to the fluorescence of the weakest sample (wild type 1 45°C, 60 minutes) at 1x concentration by identifying the dilutions with Ct values closest to 29 (using the standard curves generated previously). All samples were run in duplicate for each primer pair along with duplicate no template controls.

Fold change of induction of genes were calculated using the ΔΔct method:

$$\text{Fold change} = 2^{-\Delta\Delta ct}$$

where ΔΔct = [(Ct of GOI – Ct internal control)standard condition –

(Ct of GOI – Ct internal control) variable condition]

(taken from Schmittgen & Livak, 2008)

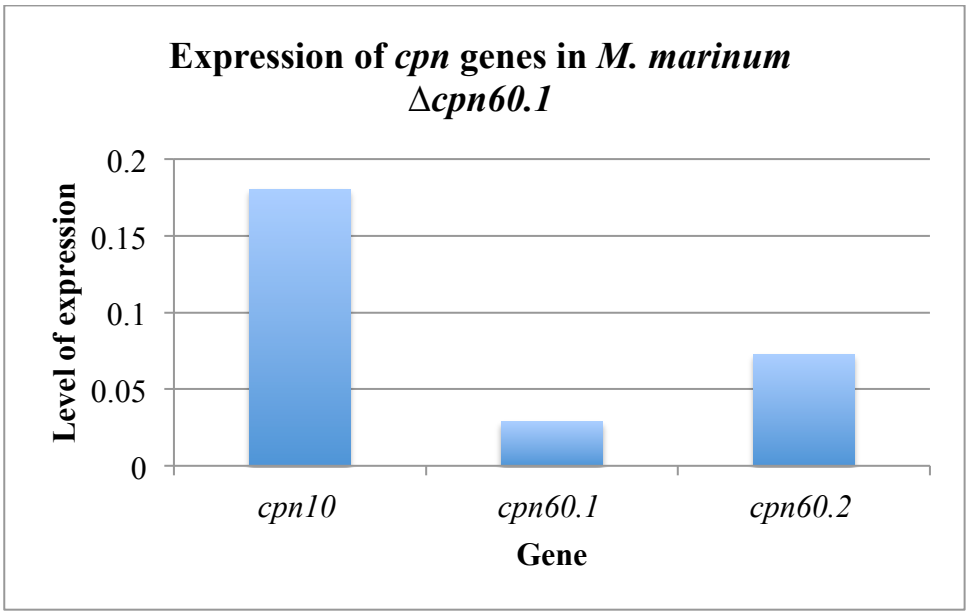


Fig. 3.9 qRT-PCR based detection of the expression of *cpn* genes in *M. marinum* Δ*cpn60.1* strain with respect to levels of *cpn* expression observed in *M. marinum* wild type strain, both grown under normal growth conditions.

Overall poor quality of cDNA had a negative impact on the quality of the results achieved, with fluorescence intensity values of the individual reactions rarely falling within the acceptable 15-27 Ct range.

Fig. 3.9. shows the relative expression of the *cpn* genes observed in the Δ*cpn60.1* strain of *M. marinum* under normal growth conditions. In consensus with previous findings in *M. smegmatis* (Rao & Lund, 2010), *cpn10* is the most highly expressed of the *cpn* genes followed by *cpn60.2*. Expression of *cpn60.1* was also detected, at levels higher than that of those found in wild-type strain samples, bringing again into question the validity of these results.

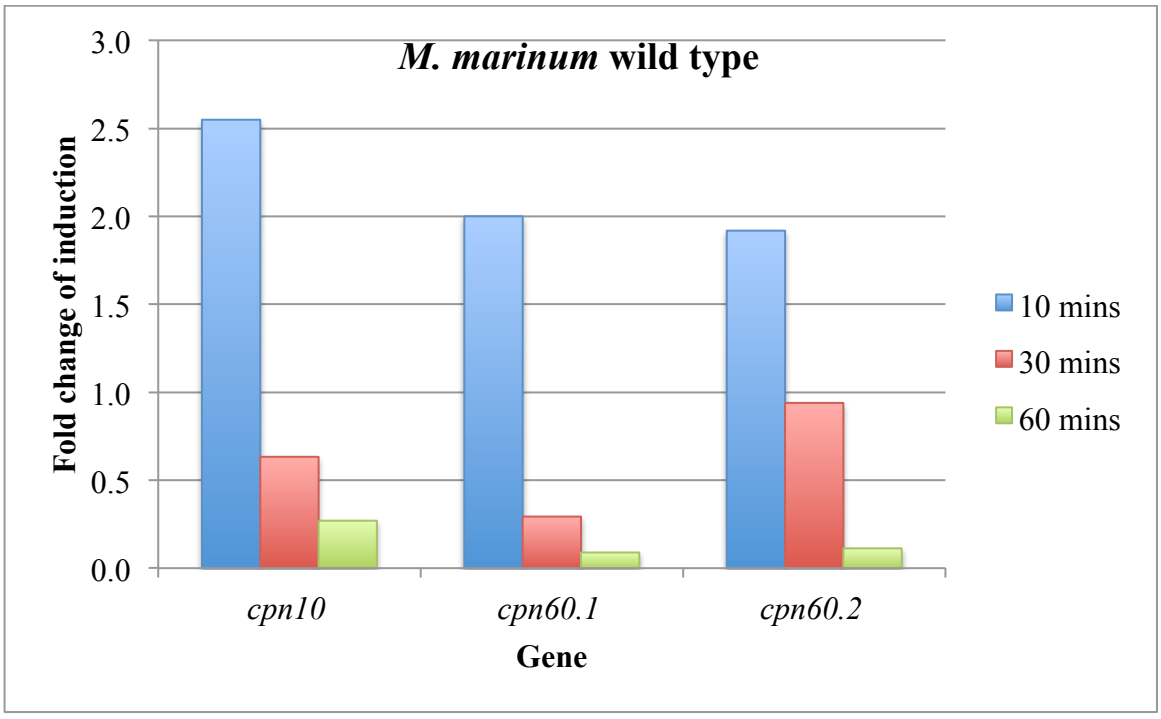


Fig. 3.10 qRT-PCR based detection of the change in expression of the *cpn* genes during heat shock (45 °C) in wild type *M. marinum*. Values have been normalized to the expression of control gene *gap*.

Fig. 3.10 shows that in wild type *M.* marinum, fold induction of the expression of *cpn60.1* and *cpn60.2* increases by roughly 2 fold after 10 minutes after heat-shock (45 °C), with a 2.5-fold induction observed in *cpn10* expression. After 30 minutes, fold induction of genes drops to less than a 1 fold increase, with *cpn60.2* showing highest fold change of expression. After 60 minutes expression of genes are almost at the levels observed under standard growth conditions.

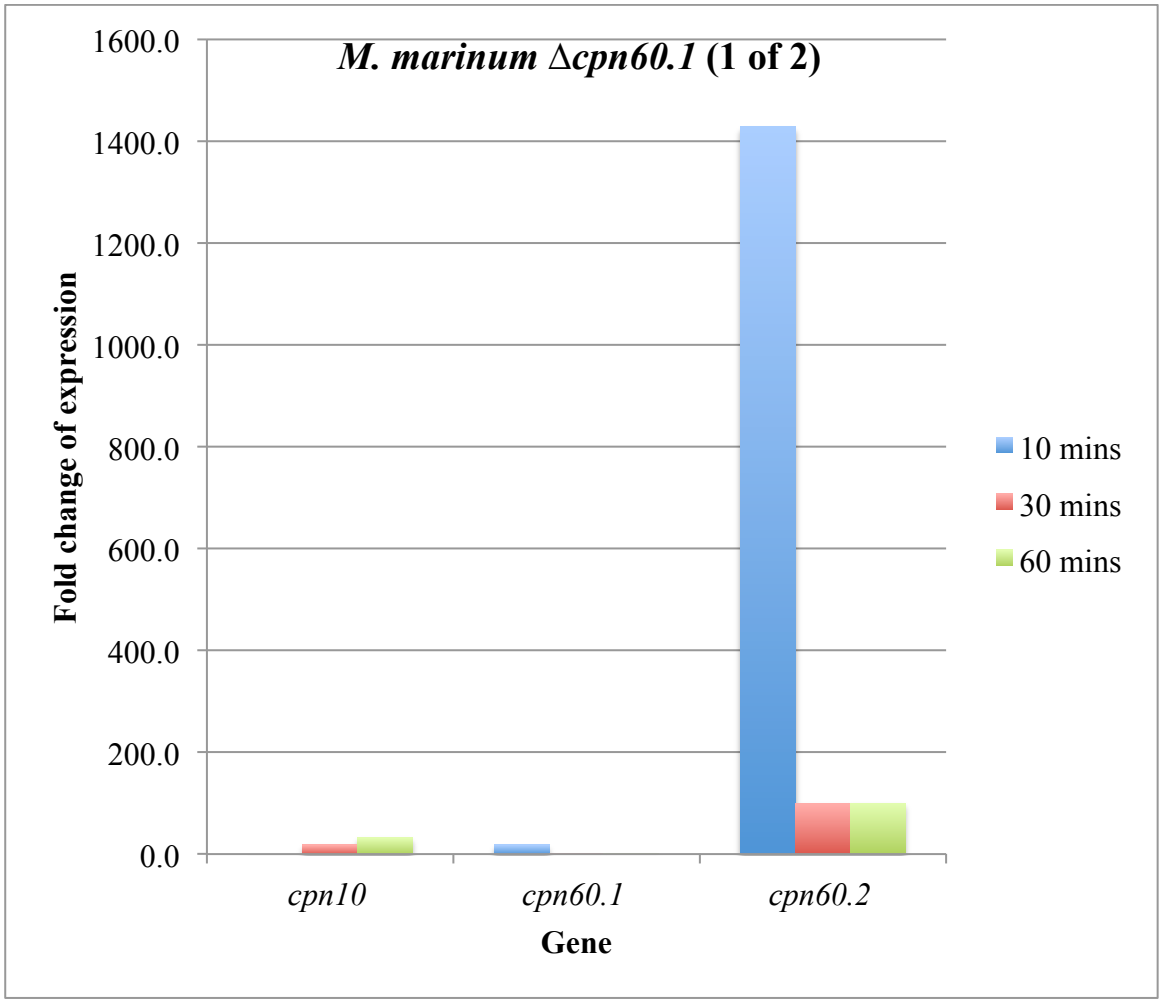


Fig.3.11 (Replicate 1 of 2) QRT-PCR based detection of the change in expression of the *cpn* genes during heat shock (45 °C) in Δ*cpn60.1 M. marinum* with respect to a 0 minute time point sample grown at 30 °C. Values have been normalized to the expression of control gene *gap*.

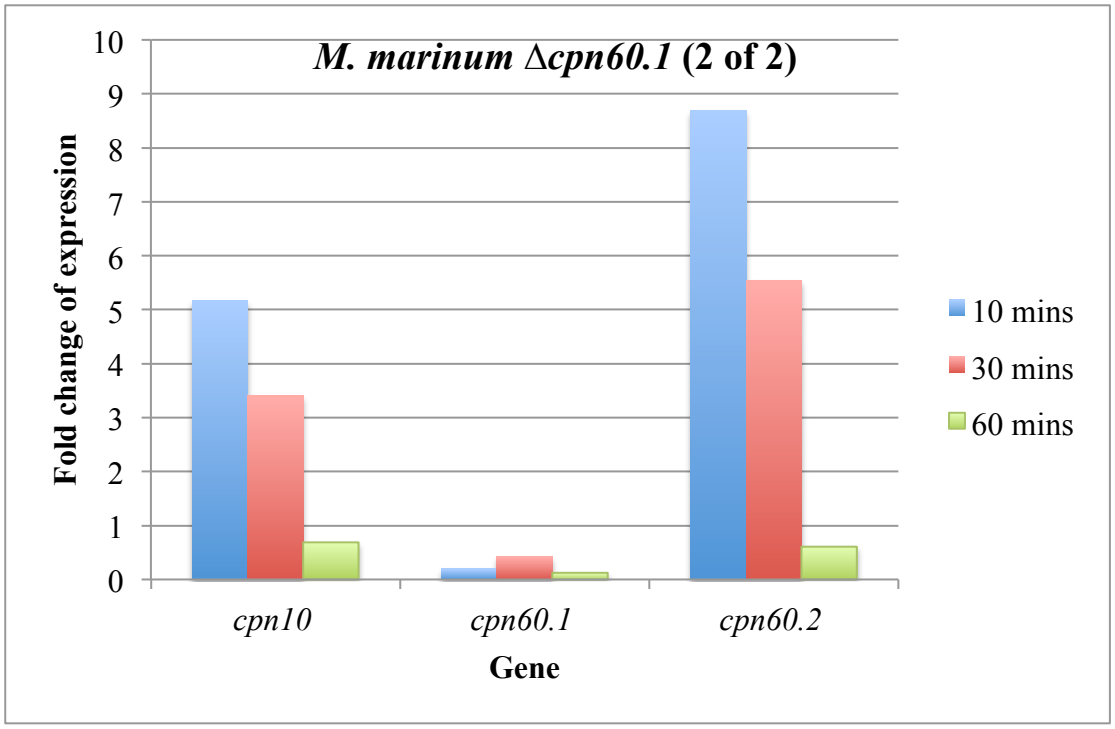


Fig. 3.12 (Replicate 2 of 2) qRT-PCR based detection of the change in expression of the *cpn* genes during heat shock (45 °C) in Δ*cpn60.1 M. marinum* with respect to the expression of the same genes under standard conditions. Values have been normalized to the expression of control gene *gap*. All values are derived from the mean results of technical duplicates.

Fig. 3.11 shows the results obtained from the first replicate of the qRT-PCR investigation into the expression of the *cpn* genes in the mutant *M. marinum* Δ*cpn60.1* strain under heat-shock at 45 °C. A massive increase in the fold induction of *cpn60.2* is observed. Whilst this was a possibility hypothesized in response to the finding that heat-shock resistance is conferred by *cpn60.1* loss, this level of induction was surprising. Repetition of this experiment was the next task performed. New cDNA samples were synthesized from fresh biological replicate cultures. The values displayed in Fig. 3.12 are the result of technical duplicates at the qRT-PCR stage. Quality of cDNA obtained from this repeat was much higher, and this was reflected in the values of fluorescence obtained from the qRT-PCR reaction, making this data

more valid than that shown in Fig. 3.11. Displaying a similar pattern of induction to wild type *M. marinum*, the fold change induction of expression of *cpn10* and *cpn60.2* is both increased after 10 minutes under heat-shock condition. This then decreases after 30 minutes, and again after 60 minutes at 45 °C. Low levels of expression of *cpn60.1* were detected but this does not appear to be significant.

3.4 Lipid profiling of wild type and Δ*cpn60.1 M. marinum* strains

Previous experiments have identified key differences between the lipid profiles of wild type and Δcpn60.1 mutants of *M. smegmatis* and *M. bovis BCG*. In *M. bovis* BCG Δcpn60-1 mutants, Pthiocerol dimycocerosates are absent (Wang, 2011) and in *M. smegmatis* GroEL1 (cpn60.1) is required for the induction of C_{56}-C_{68} fatty acids during biofilm formation (Ojha *et al*, 2005).

Lipids were extracted from both wild type and Δ*cpn60.1 M. marinum* and the equivalent strains of *M. smegmatis*. 2D-TLCs were then performed of all samples using solvent sytems A-E (described in Minnikin *et al,* 1985) in order to separate out the multiple different classes of lipids found in *Mycobacteria*. The resulting lipid patterns formed on each plate where then compared between wild type and Δ*cpn60.1* mutants to identify any differences.

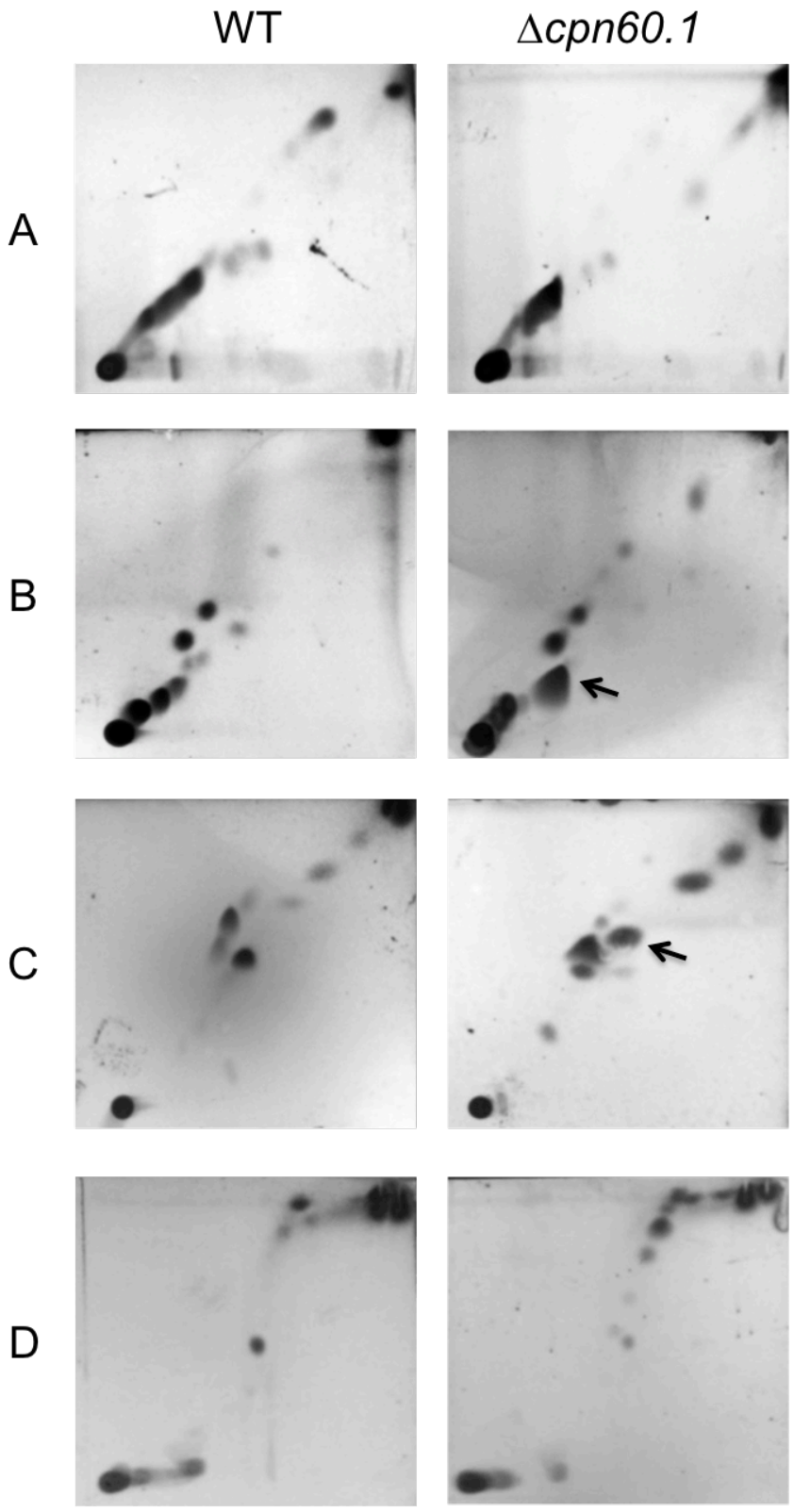


Fig. 3.13 2D-TLCs of *M. marinum* wild type and Δ*cpn60.1* apolar lipids**.** (A) solvent system A, (B) solvent system B, (C) solvent system C, (D), solvent system D (solvent systems are described in Materials and Methods section 2.18. Key differences are highlighted with arrows.

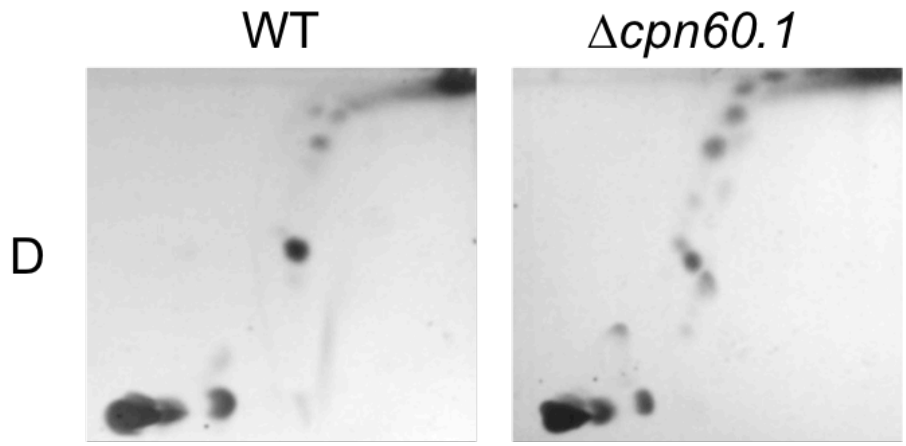


Fig. 3.14 2D-TLCs of *M. marinum* wild type and Δ*cpn60.1* polar lipids separated using solvent system D (see Materials and Methods section 2.18 for full details on solvent systems).

Fig. 3.13 shows the results of apolar lipid separation using solvent systems A to D. In the TLC plates from both systems B and C there appears to be two differences observed in the Δ*cpn60.1* mutant compared to the wild type (Fig.3.13; indicated by arrows). Both system B and C resolve the same lipid class, the free fatty acids.

The resulting TLC plates from samples run using solvent system E unsatisfactory (see supplementary data, S1) and due to time constrains could not be repeated. Even still, no visible difference was seen (although the validity of this statement is dependent on a repeat of the experiment). No obvious differences were observed in the polar lipids between wild type and mutant strains run using solvent system D (Fig. 3.14).

3.5 Generation of complemented strains of *M. marinum* Δ*cpn60.1*

To determine whether the phenotype observed in the lipid profiles of *M. marinum* Δ*cpn60.1* was as a result of *cpn60.1* loss or another unknown factor, complemented strains were generated.

E. coli DH5α cells were transformed with each of the following plasmids;

- pMSGroEL1, expressing GroEL1 (*cpn60.1)* from *M. smegmatis*
- pMSGroEL2, expressing GroEL2 (*cpn60.2*) from *M. smegmatis*
- pMS10.cpn60.2, a different plasmid expressing *M. smegmatis cpn60.2*
- pMS10.cpn60.3, expressing *M. smegmatis cpn60.3*
- pMS10. groEL, expressing *E. coli GroEL*

All of these genes are in an operon with *M. smegmatis cpn10* under the control of the Hsp60 promoter from *M. bovis* BCG with Kan^R as a selectable marker.

Plasmids were purified and sequenced against the qPCR primer MM_cpn10_F to determine the sequence downstream of *cpn10* and confirm the identity of the plasmids. All plasmids were correct with the exception of pMSGroEL2, which was found to contain *cpn60.1*, most likely as a result of mislabeling of the eppendorf tube.

Due to these plasmids containing a Kanamycin resistance gene the previously used Δ*cpn60.1* strains already transformed with Kan^R pMSP-dsRed meant that selection of complemented transformants could not be achieved using Kanamycin. Therefore, the Δ*cpn60.1* mutant strain not transformed with pMSP-dsRed was used for these transformations.

Plasmids were initially eluted from the spin columns (as part of Qiagen kit) using the elution buffer provided, and thus were not de-salted. Therefore, the first attempt at transformation used only 2 µl of plasmid DNA was used in each reaction. The protocol used was based on that found in Rao & Lund, 2010 for the transformation of *M. smegmatis*, using 400 µl of competent cells and an electroporation voltage of 1.8 kV. However, after 1-week incubation on solid media no distinct colonies had formed. Cultures of recombinant *E. coli* DH5α strains were grown and Plasmid DNA was again purified but instead of elution buffer, nuclease-free water was used for elution. Electroporation was performed according to the *M. marinum*-specific protocol from El-Etr *et al*., 2004 (200 µl competent cells, 2.5 kV). Ahigh transformation efficiency was observed across all plates, with the exception of the pMScpn10.cpn60.3 strain, which only formed 1 colony. Glycerol stocks were created of each new strain and stored at -80 °C.

3.5.1. Generation of a plasmid vector for the expression of *M. marinum cpn60.1*

To act as a positive control in the complementation experiments - work was started to achieve a plasmid to express *cpn60.1* from *M. marinum.* A strain of MGM100 containing the plasmid pTrcESmmcpn60.1 was grown and the plasmid purified. The exact details on the development of the pTrcESmmcpn60.1 were not clear. However, Rao (2010) had used pTrcES-X plasmids in the creation of the pMS10-X plasmids, where X refers to *M. smegmatis cpn60.2, cpn60.3* or *E. coli groEL* genes. Generation of these plasmids is detailed in Rao (2010). All of the pTrcES-X plasmids contained at least one *Hind*III site flanking the *cpn* or *groEL* genes. Therefore digestion of pTrcESmmcpn60.1 with *Hind*III was performed but analysis of the product formed by gel electrophoresis revealed only 1 band of around 6 kb

in size, suggesting that only partial digest had been achieved. Due to time constraints this portion of the study was not followed up.

3.5.2. Generation of a *cpn10*-expressing control plasmid for complementation

All of the *cpn60* genes available in the complementation plasmids are co-expressed with *cpn10*, efforts were made to develop a control plasmid containing only the *cpn10* gene. The plasmid pMS10.cpn60.2 contains *cpn60.2* flanked by *Hind*III restriction sites (Rao, 2010).

Digestion was performed with HindIII and the product run on a 1% agarose gel. Two bands were observed when visualized using near-UV, the larger of which was excised and purified. The large DNA fragment, thought to be representative of pMS10.cpn60.2 with *cpn60.2* removed was left at room temperature to re-circularize. Purification yield was low so sequencing couldn't be performed to confirm identity. Instead, the plasmid was re-transformed into *E. coli* DH5α and grown in presence of arabinose and kanamycin. However, after 24 hours no colonies were formed.

3.6 2D-TLC analysis of lipid content of complemented *M. marinum* Δ*cpn60.1* strains

Large cultures (200 mL) of complemented strains were established and grown until saturated. Lipids were purified from pelleted cells according to the protocol described in section 2.17. Due to time constraints only solvent systems B and C (see section 2.18) were repeated using these complemented strains.

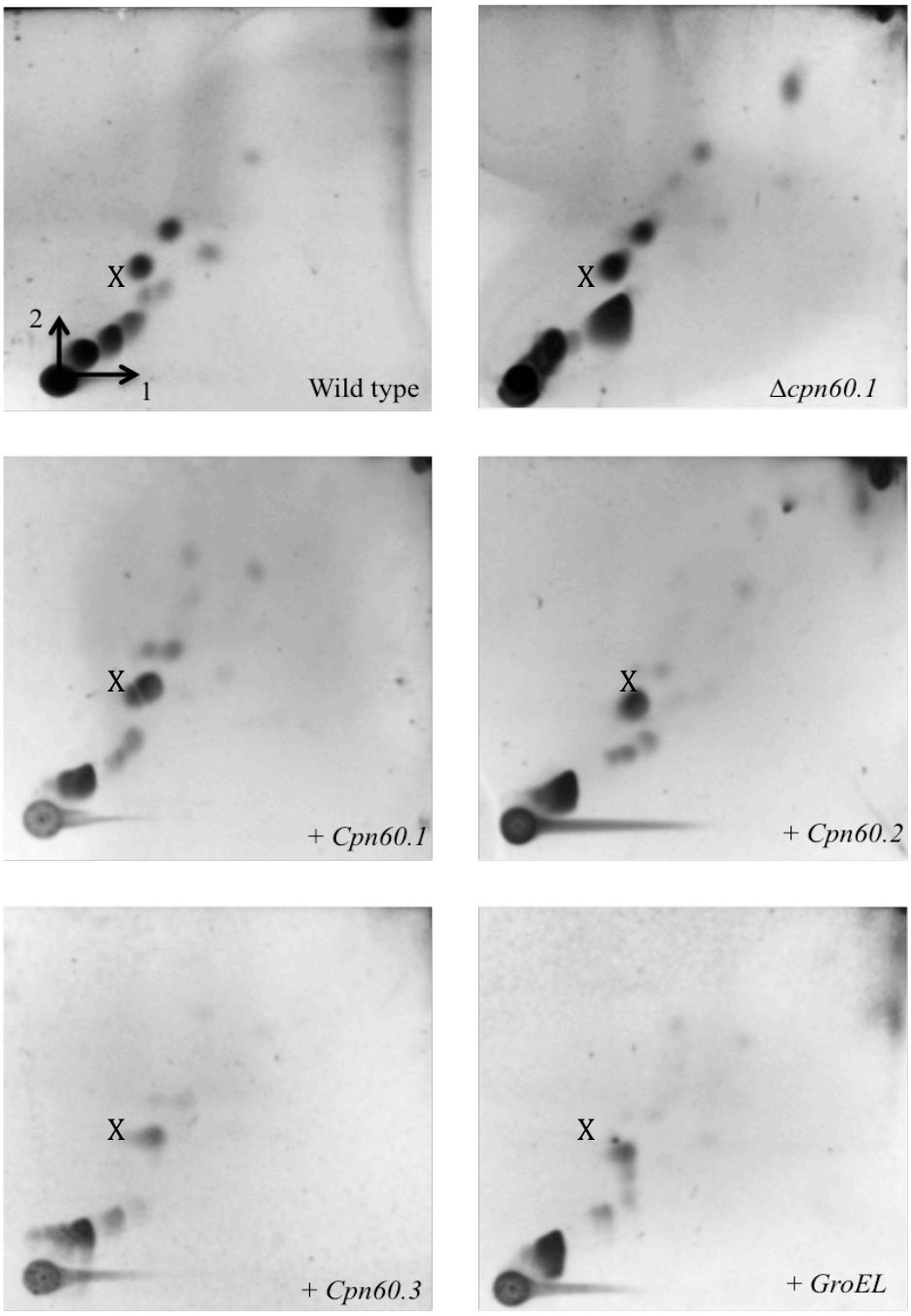


Fig.3.15. 2D-TLC separation of lipids using system B (see section 2.18). (Top left) wild-type *M. marinum*, (top right) Δ*cpn60.1 M. marinum*, (centre left) Δ*cpn60.1* + *M. smegmatis cpn60.1,* (centre right) Δ*cpn60.1* + *M. smegmatis cpn60.2,* (bottom left) Δ*cpn60.1* + *M. smegmatis cpn60.3,* (bottom right) Δ*cpn60.1* + *E. coli groEL.* Plates were developed using MPA. X has been added to each TLC to act as a reference point.

Fig.3.15 shows the results of the 2D-TLCs of *M. marinum* Δ*cpn60.1* strain complemented with the different plasmids, run using system B. Separation of the lipids wasn't as effective as seen in the previous experiment. Due to time constraints this experiment couldn't be repeated. It appears that *M. smegmatis* Cpn60.1, Cpn60.2, Cpn60.3 and *E. coli* GroEL all restored the phenotype of the wild type *M. marinum* strain as no intense spot is observed (beneath spot marked with X in Fig.3.15). However, without repetition of this experiment and the previous 2D-TLC experiments it would not be accurate to make such assumptions due to the poor separation observed. The same samples were run using solvent system C but due to experimental error no data was obtained.

3.7 Standardization of growth protocols

An issue was called to our attention by Dr. Astrid M. van der Sar at the VU University Medical Centre in Amsterdam who is performing the zebrafish embryo infection assays of our strains. Her group noted some issues with the correlation between OD of grown strains and the number of colonies formed when plated. In order to assess this a literature search was performed. Gao *et al.,* 2005 state that during log-phase >98% of bacteria are aggregated in clusters of >100. The same paper provides protocols specific to *M. marinum* growth and maintenance. By adopting these techniques we aimed to standardize our methods and provide our Dutch collaborators with a set of standard protocols.

Two practices suggested by the Gao *et al,* 2005 paper are the low speed centrifugation of cultures to pellet clumped cells and the use of a syringe needle to shear cell aggregates.

The first experiment was to investigate the effect of centrifugation compared to syringe shearing and syringe shearing followed by a centrifugation step. We expected that centrifugation would result in a large decrease in colony formation compared to un-treated cultures, whilst passing culture through a needle would increase number of colony forming units. Centrifugation after syringe-pass should then have little effect on number of colony forming units. The OD_{600} was also taken after each step.

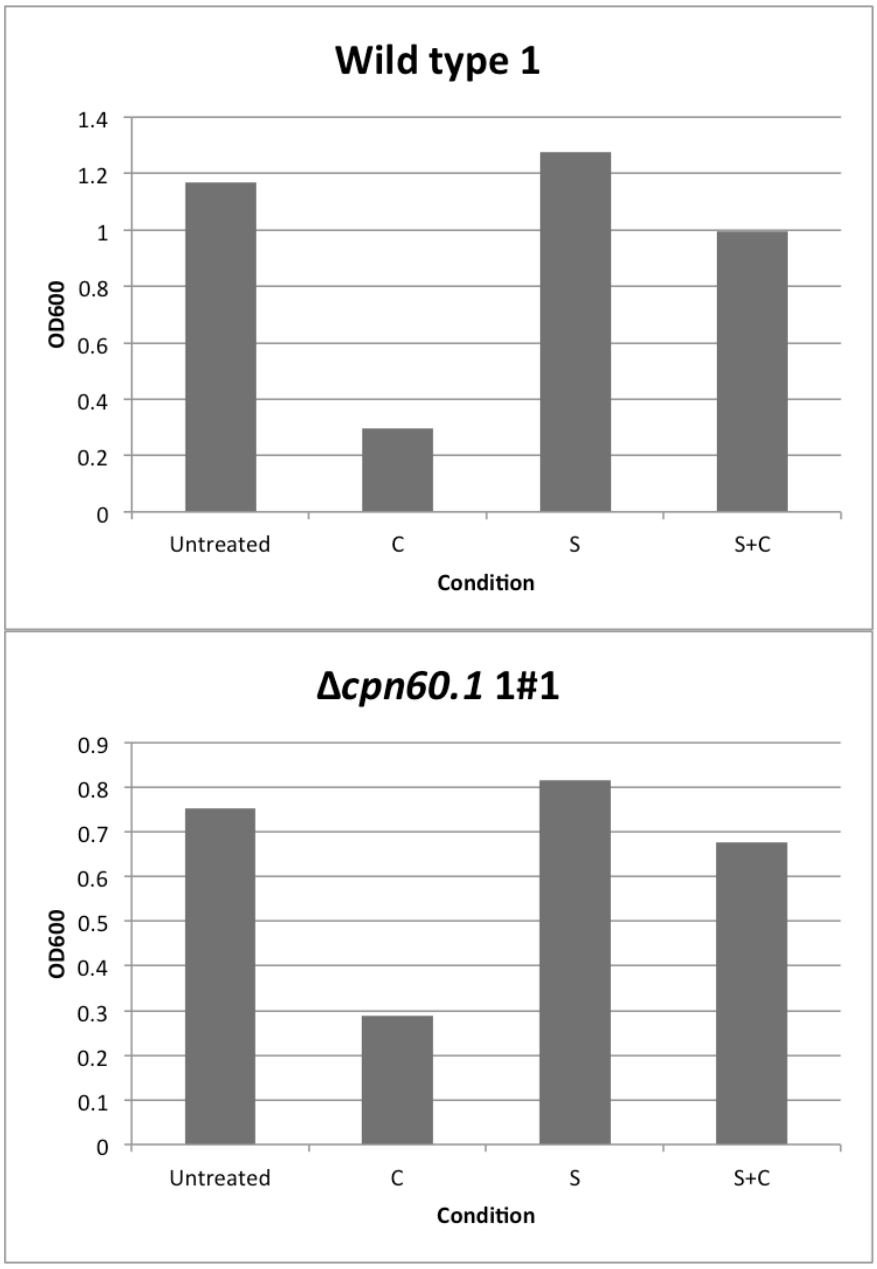


Fig. 3.16 Effect of centrifugation (C), syringing (S) and syringing + centrifugation (S+C) on the OD_{600} of *M. marinum* wild type (top) and Δ*cpn60.1* (bottom) cultures.

OD_{600} values of the cells taken after each amendment followed the expected trend (Fig 3.16.). Centrifuged cultures show a large decrease in OD_{600}. Passing culture through a syringe showed a small increase in OD_{600} with a subsequent centrifugation of the same sample decreasing OD_{600}.

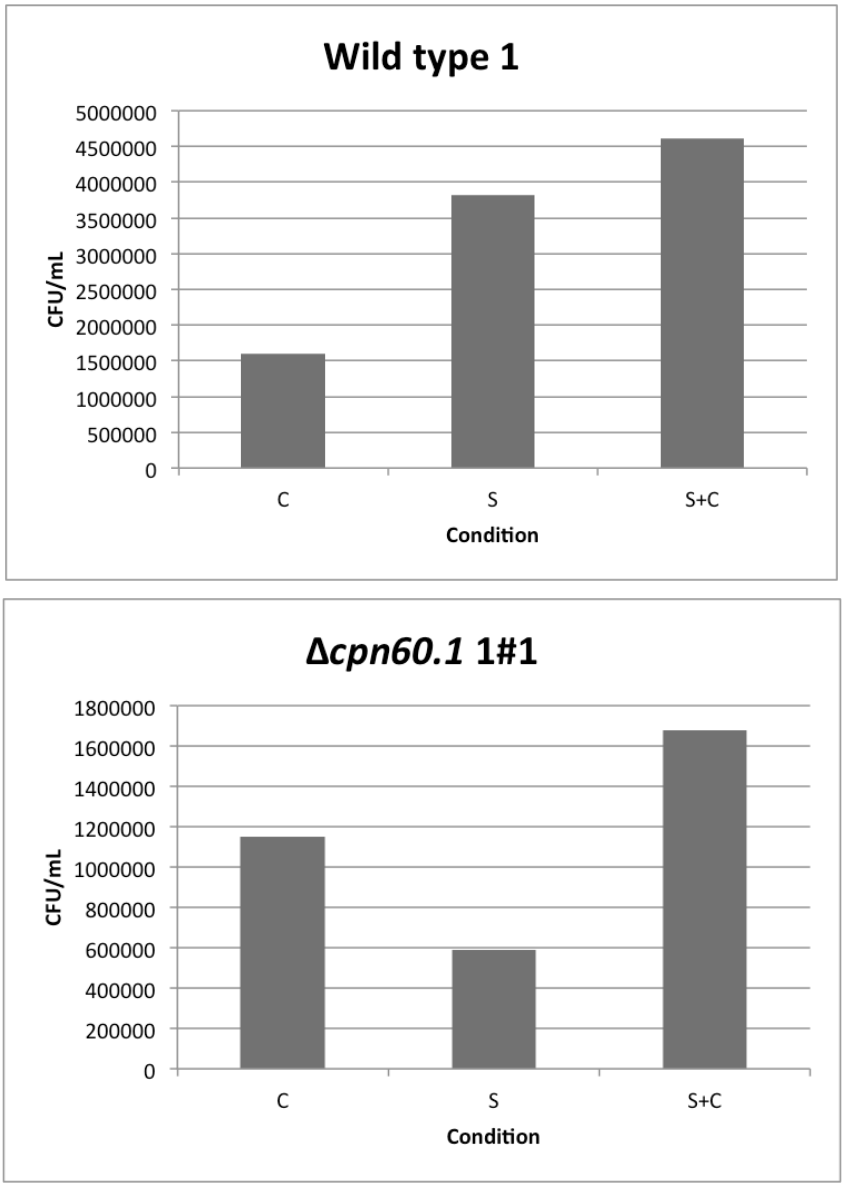


Fig. 3.17 Effect of centrifugation (C), syringing (S) and syringing + centrifugation (S+C) on the number of CFUs present in *M. marinum* wild type (top) and Δ*cpn60.1* (bottom) cultures.

Number of colonies formed from each sample after 1 week's incubation was recorded. In both the wild type and Δ*cpn60.1* strains, combined action of syringing and centrifugation yielded the greatest CFU. Syringing alone increased CFU in the wild type sample but not in the Δ*cpn60.1* sample (most likely due to pipetting error due to S+C result).

For shearing of cell aggregates, cultures were passed through the needle six times (3 x in and out). A second experiment was carried out, plating samples of a culture after each subsequent needle-pass to study the effect this had on CFU to determine the optimal number of passes. The OD_{600} was also taken after each pass.

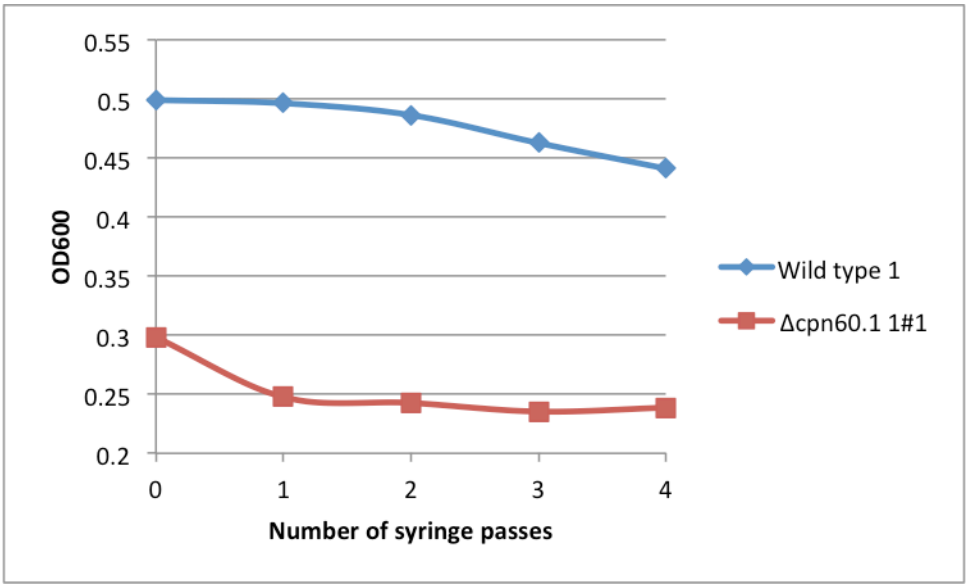


Fig. 3.18 The change in $OD_{600\ nm}$ observed upon subsequent passes through a syringe-needle in both wild type (blue) and Δ*cpn60.1* mutant strains of *M. marinum*. Culture (5 mL) of each strain was passed through a syringe needle 5 times. In between each pass, 1 mL of culture was removed and the $OD_{600\ nm}$ of the sample recorded using a reference of sterile 7H9 + OADC + Tween-80.

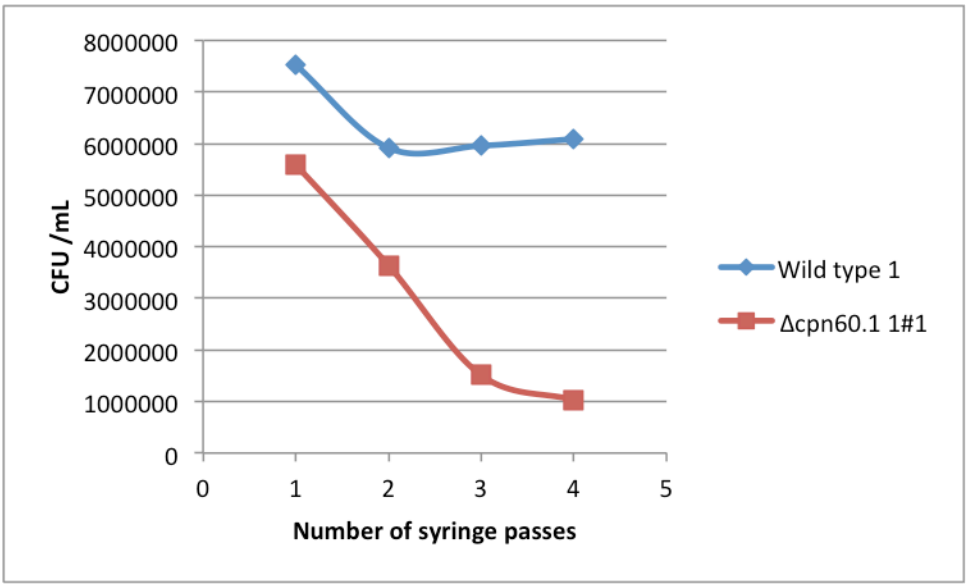


Fig. 3.19 The change in CFU observed after subsequent passes through a syringe-needle in both wild type (blue) and Δ*cpn60.1* mutant strains of *M. marinum*. Culture (5 mL) of each strain was passed through a syringe needle 5 times. In between each pass, 1 mL of culture was removed, a serial dilution established and plated on 7H10 agar. CFU was recorded 1 week later and the mean value is represented above.

Fig. 3.18 shows the effect of needle passes on OD_{600} wild type and knockout mutants. There was a decrease of OD_{600} with each subsequent pass through the needle. This is converse to the previous evidence that OD_{600} increases with needle shearing. We also saw a decrease in the CFU of both samples with each needle pass (Fig. 3.19). Again, this is opposite to the previously observed result that CFU increases when culture is passed through a needle.

To determine the actual effect of passing culture through a needle, cells were observed under a microscope and counted using a hemocytometer. Due to the acid-fast nature of *M. marinum,* and the impracticality associated with performing Ziehl-Neelson staining on a vast array of samples under time constraints, phase contrast microscopy was used to visualize cells.

Cultures of both wild type and Δ*cpn60.1* strains of *M. marinum* were grown to the same OD_{600}. A manageable density of cells was determined by creating a serial dilution of cells in standard media and observing them on the hemocytometer. Once the dilution was established number of cells were counted, with clumps of cells counted as 1 colony-forming unit (cfu). A large difference was observed between the cfu of wild type and Δ*cpn60.1 M. marinum* strains. The wild type samples contained around twice as many cfus compared to Δ*cpn60.1* (636 compared to 342). However, this was mostly due to Δ*cpn60.1* cells existing in larger clusters of cells compared to wild-type. After one pass through the syringe needle the number of cfu doubled in both cases (from 636-1183 in wild-type and 342 to 855 for Δ*cpn60.1*). After the second pass the number of cfu increased further, to an uncountable level, with the majority of cells existing in clusters of 10 cells or less.

3.8 Assessing biofilm/pellicle formation of *M. marinum* strains

In our lab extensive previous work had been carried out to assess the biofilm-forming ability of Δ*GroEL1* mutants and wild type *M. smegmatis* strains. This was as a result of the observation that Δ*GroEL1* mutants of *M. smegmatis* lose the ability to form mature biofilms (Ojha *et al.,* 2005).

To assess whether this phenotype was also characteristic of the M. marinum Δ*cpn60.1* mutant the same assay was performed with this strain. Wild type *M. marinum* was also assayed as a control and the two corresponding strains from *M. smegmatis* were also assessed.

Assays were performed according to the method outlined in Rao, 2010. However after 1 week results were contradictory and non-reliable (Data not shown). An alternative protocol was

received from a PhD student working on *M. smegmatis* biofilm formation (Nandita Keshevan, Imperial College London). This new protocol looked specifically at pellicle formation, i.e. cell proliferation at the air/water interface. Two-day old cultures of *M. marinum* were used to inoculate 7H9 media (with no Tween-80) and incubated for 1 week. After incubation tubes were photographed without moving from the incubator so as to not disturb any pellicle formed. All but one sample showed no pellicle formation with cells settled at the bottom of each tube (Fig. 19). The one sample in which pellicle had formed, *M. marinum* wild type 2, was anomalous as replicates of the same strain showed no pellicle.

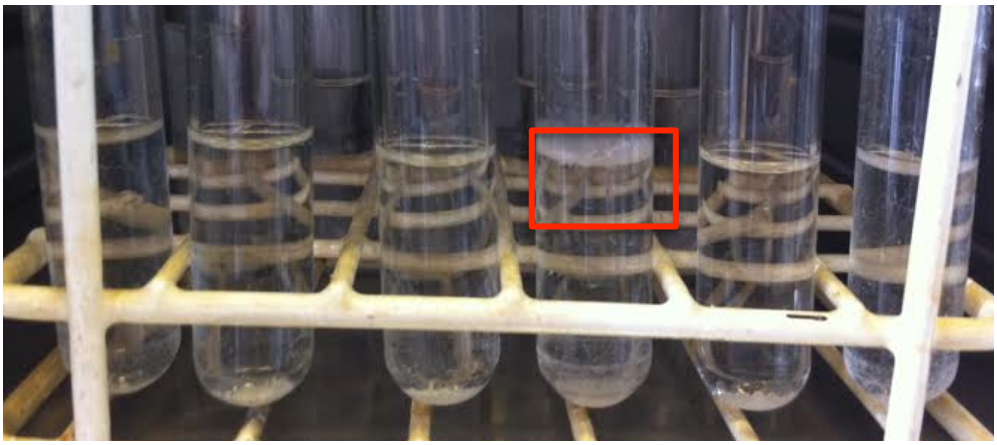


Fig. 3.19 Pellicle formation in *M. marinum* wild type 2

4. DISCUSSION

4.1 Loss of *cpn60.1* from *M. marinum* appears to convey resistance to heat-shock

The finding that loss of Cpn60.1 appears to convey resistance upon exposure to high temperatures was surprising. The Δ*cpn60.1* mutant *M. marinum* strain was able to survive exposure to higher temperatures for a longer time period than wild type (Fig. 3.2; >3 hours at 45 °C for mutant compared to >3 hours for 45 °C). The loss of *cpn60.1* also appeared to confer resistance to even higher temperatures (50 °C). GroEL and it's homologues have long been known to a play a role in the heat shock response, even earning them the alternative name of heat shock proteins (HSPs). However, in most documented cases, loss of the dispensable chaperonin - Cpn60.1 in Mycobacteria tends to have an adverse effect on response to heat shock. In *M. tuberculosis* Δ*cpn60.1* mutants were found to be more sensitive to 55 °C (Hu *et al.,* 2008). Alternatively, in *M. bovis BCG*, loss of Cpn60.1 has no effect upon cell viability at 42 °C (Wang *et al.,* 2011). Although 42 °C is not as radical a temperature deviation as that tested in the *M. tuberculosis* experiments, it is still above the 37 °C optimal physiological growth temperature of *M. bovis BCG*. No difference in viability was observed in *M. marinum* strains grown at 37 °C, despite an optimal growth temperature of 30 °C. Therefore a possible role for Cpn60.1 in *M. bovis BCG* might be elucidated if Δ*cpn60.1* mutant strains were exposed to higher temps.

The evidence gathered in this study appears to suggest that unlike the findings of other studies into Mycobacterial Cpn60.1-knockouts, in *M. marinum,* Cpn60.1 loss actually appears to be advantageous, increasing viability at higher temperatures. It is unclear as to whether this is

due to the function of an alternative protein compensating for Cpn60.1 loss or an example of another moonlighting function of Cpn60.1 specific to *M. marinum.*

In order to confirm this result is definitely a consequence of CPn60.1 knockout, the heat-killing experiments could be repeated with a complemented strain of *M. marinum* Δ*cpn60.1*. Alternatively, the genome of both wild type 1 and 2, Δ*cpn60.1* 1#1 and 2#1 strains have been sequenced, so genome-wide alignments could be used to identify any single nucleotide polymorphisms (SNPs), or potentially larger regions of sequence difference could be identified.

4.2 Induction of the *M. marinum cpn* genes is increased in response to heat-shock

As a result of the heat-killing experiment findings the expression of the *cpn* genes was investigated using qRT-PCR. Previous studies have shown that the expression of the chaperonin genes is responsive to different stresses, especially as a result of heat-shock (Stewart *et al*., 2002; Rao & Lund, 2010). It is also worth noting that the expression of individual *cpn* genes is differential under stress conditions, i.e. not equal, with certain genes more highly expressed than others, indicative of a divergence of function.

The increased temperature-resistance of the Δ*cpn60.1* strain of *M. marinum* was hypothesized to be as a result of Cpn60.2 compensation for Cpn60.1 loss. Both wild type and mutant *M. marinum* strains were subjected to heat shock at 45 (°C). In wild type samples all *cpn* genes were shown to have an increase in expression upon heat shock, in agreement with the findings of Stewart *et al.,* (2002)*,* and Rao and Lund, (2010) in *M. tuberculosis* and *M. smegmatis*, respectively. Highest levels of expression were observed at 10 minutes post temperature

increase. Smaller time-points (5, 10, 15, 30 minutes) were used by Rao & Lund (2010) when investigating the change in *M. smegmatis cpn* expression during heat shock, but the overall pattern of expression (initial large increase of fold induction followed by slow decline) was the same as observed in *M. marinum* (Fig. 3.10). Results obtained for expression in Δ*cpn60.1* strains were conflicting, mostly due to issues surrounding RNA quality. A spectrophotometric measurement had been the only method of detecting RNA yield from purification, however this is by no means representative of the quality of purified RNA. Use of technology such as the Bioanalyser 2100 (Agilent technologies) could have provided this information that would have informed quantity of RNA used for cDNA synthesis. If RNA quality were found to not be the issue then poor cDNA yields would mostly likely be as a result of poor amplification of DNA by the random hexamer primers. Efforts were made to optimize this cDNA synthesis protocol by increasing the time of the extension step and by the use of betaine or DMSO but all were found to have no effect. An alternative approach could involve the use of gene-specific primers to create cDNA of the genes of interest rather than of the whole transcriptome of the cell. Assuming the rate of amplification of each gene was identical between primer pairs used then qRT-PCR could still be used to detect relative expression of each of the genes.

For Δ*cpn60.1 M. marinum* cDNA samples, expression of Cpn60.1 was detected, despite the absence of the *cpn60.1* gene. Expression of *cpn60.2* was shown to significantly increase in heat-shocked Δ*cpn60.1* samples (Fig. 3.12). Replicate runs of the same experiment (Fig 3.13) on new cDNA samples showed the same trend of an increase *cpn60.2* expression, although not to such a high degree as shown in Fig. 3.12. The result from run 1 of 2 was most likely anomalous due to the low concentration of reaction template obtained from the first cDNA synthesis attempt. However, it does appear that *cpn60.2* expression is increased in the

knockout strain to compensate for *cpn60.1* loss, showing a ~9-fold change (Fig. 3.13) in induction compared to ~2-fold change observed in wild-type samples (Fig. 3.11)

To determine the larger effects *cpn60.1* loss has on the overall transcriptome, a larger scale, genome-wide transcriptome analysis could be used. Microarrays and RNA-Seq are nboth widely used technologies that allow the comparison between the relative levels of transcription between many genes under different physiological conditions. These technologies could be used to contrast the transcriptomes of both wild type and Δ*cpn60.1* strains grown under standard and heat-shock conditions.

4.3 Further evidence that Cpn60.1 is most likely required for proper establishment of the Mycobacterial cell wall

Cpn60.1 is important for cell wall synthesis. In *M. smegmatis* the synthesis of mycolates during the formation of biofilms requires Cpn60.1 (Ohja *et al.,* 2005). In *M. bovis* BCG, Δ*cpn60.1* mutants fail to synthesise the phthiocerol dimycocerosate class of lipids, important for virulence of both *M. bovis* BCG (Beatty *et al.,* 2002) and *M. tuberculosis* (Asterie-Dequeker *et al.,* 2009). These findings led to the investigation of the lipid profiles of our *M. marinum* strains. Two key differences were observed in the lipids from the Δ*cpn60.1* strain separated by solvent system B and C. The position of these altered lipids can be compared to Fig 4.1, which shows the position and identity of *M. tuberculosis* lipids separated using the same method (Bhatt *et al.,* 2007a). This suggests that the overexpressed lipids observed in the Δ*cpn60.1* strain of *M. marinum* are either free fatty acids or free mycolic acids. Efforts to restore phenotype by complementation were unsuccessful mostly due to time constraints of the project. Although various complemented mutant strains were achieved only one run of

2D-TLC could be performed and the results were likely unreliable. Mass spectrometry or high-performance liquid chromatography could be used to validate the identity of altered lipid.

Altered expression of mycolic acids has been shown to have significant effects on morphology and immune response to *M. tuberculosis.* Deletion of one of the genes responsible for mycolic acid synthesis, *kasB,* in *M. tuberculosis* drastically altered colony morphology of cells (Bhatt *et al.,* 2007a). Δ*cpn60.1 M. marinum* strains displayed distinct phenotypic characteristics different to those observed in wild-type strains. Therefore, the change in cell wall composition observed in Δ*cpn60.1 M. marinum* could be as a result of the altered mycolic acid expression.

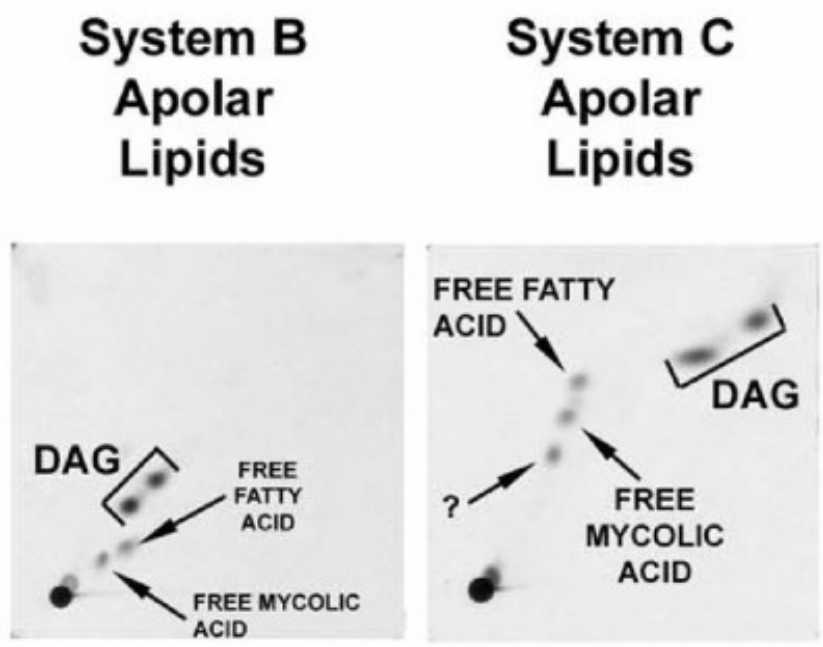


Fig. 4.1 The position and identity of specific apolar lipids of M. tuberculosis separated using 2D-TLC solvent systems B and C. Taken from Bhatt *et al*., 2007a.

4.4 Conclusion

Wild-type cells have the propensity to exist in clusters (up to 98% of cells during log-phase; Gao & Manoranja, 2005). Microscopy was used to observe the cell-clumping tendency of Δ*cpn60.1* cells and this was found to be much higher than in wild type, with clusters comprised of twice the number of cells forming in Δ*cpn60.1* cultures. This increased clumping was also visible to the naked eye and was associated with increased adherence of mutant cells to glassware used in culturing.

Methods for overcoming clusters of cells are described in Gao & Manoranjan, 2005. Use of a syringe needle in combination with a low-speed centrifugation step to pellet any larger, heavier clusters was found to be most effective. Three passes through a 23G syringe needle was sufficient to shear the majority of cell clusters in both wild type and Δ*cpn60.1* strains of *M. marinum.*

The findings of this study appear to indicate that the role of Cpn60.1 in *M. marinum* is typical of the moonlighting functions associated with this dispensable chaperonin. Loss of *cpn60.1* appears to convey a resistance to higher temperatures that are lethal to wild type cells. This is either due to a suicide-promoting role of Cpn60.1 under heat shock or more likely due to an increased expression of Cpn60.2 to compensate for the loss of Cpn60.1 protein. The correct establishment of the *M. marinum* cell wall also appears to require the action of Cpn60.1, however the mechanism remains unknown. Overall, *cpn60.1* shows promise as a potential target for the development of anti-TB drugs and *M. marinum* remains one of the most promising model organisms for the study of *M. tuberculosis.*

5. REFERENCES

Apetri, A.C., Horwich, A.L., 2008. Chaperonin chamber accelerates protein folding through passive action of preventing aggregation. PNAS 105, 17351–17355. doi:10.1073/pnas.0809794105

Astarie-Dequeker, C., Le Guyader, L., Malaga, W., Seaphanh, F., Chalut, C., Lopez, A., Guilhot, C., 2009. Phthiocerol Dimycocerosates of *M. tuberculosis* Participate in Macrophage Invasion by Inducing Changes in the Organization of Plasma Membrane Lipids. PLOS Pathogens 5(2), e1000289. doi: 10.1371/journal.ppat.1000289

Azia, A., Unger, R., Horovitz, A., 2012. What distinguishes GroEL substrates from other *Escherichia coli* proteins? FEBS Journal 279(4), 543-550

Barkan, D., Hedhli, D., Yan, H.-G., Huygen, K., Glickman, M.S., 2012. Mycobacterium tuberculosis Lacking All Mycolic Acid Cyclopropanation Is Viable but Highly Attenuated and Hyperinflammatory in Mice. Infect. Immun. 80, 1958–1968. doi:10.1128/IAI.00021-12

Barker, L.P., Porcella, S.F., Wyatt, R.G., Small, P. l. ., 1999. The Mycobacterium marinum G13 promoter is a strong sigma 70-like promoter that is expressed in Escherichia coli and mycobacteria species. FEMS Microbiology Letters 175, 79–85. doi:10.1111/j.1574-6968.1999.tb13604.x

Beatty, W.L., Rhoades, E.R., Ullrich, H-J., Chatterjee, D., Heuser, J.E., Russell, D.G., 2002. Trafficking and Release of Mycobacterial Lipids from Infected Macrophages. Traffic 1(3), 235-247.

Benton, J., & Karkanevatos, A., 2007. Preseptal cellulitis due to Mycobacterium marinum. The Journal of Layrngology and Otology 121(6), 606-8

Bhatt, A., Fujiwara, N., Bhatt, K., Gurcha, S.S., Kremer, L., Chen, B., Chan, J., Porcelli, S.A., Kobayashi, K., Besra, G.S., Jacobs, W.R., 2007a. Deletion of kasB in Mycobacterium tuberculosis causes loss of acid-fastness and subclinical latent tuberculosis in immunocompetent mice. Proc Natl Acad Sci U S A 104, 5157–5162. doi:10.1073/pnas.0608654104

Bhatt, A., Molle, V., Besra, G.S., Jacobs, W.R., Kremer, L., 2007b. The Mycobacterium tuberculosis FAS-II condensing enzymes: their role in mycolic acid biosynthesis, acid-fastness, pathogenesis and in future drug development. Molecular Microbiology 64, 1442–1454. doi:10.1111/j.1365-2958.2007.05761.x

Bruijnesteijn van Coppenraet, L.E.S., LUMC, 2009. Diagnostics of non-tuberculous mycobacteria [WWW Document]. URL https://openaccess.leidenuniv.nl/handle/1887/13665 (accessed 7.18.14).

CDC - Core Curriculum: What the Clinician Should Know - TB [WWW Document], n.d. URL http://www.cdc.gov/tb/education/corecurr/default.htm (accessed 7.15.14).

Colaco, C.A., MacDougall, A., 2014. Mycobacterial chaperonins: the tail wags the dog. FEMS Microbiol Lett 350, 20–24. doi:10.1111/1574-6968.12276

Cowing, D.W., Gross, C.A., 1989. Interaction of Escherichia coli RNA polymerase holoenzyme containing sigma 32 with heat shock promoters. DNase I footprinting and methylation protection. J. Mol. Biol. 210, 513–520.

Cronan, M.R., Tobin, D.M., 2014. Fit for consumption: zebrafish as a model for tuberculosis. Dis Model Mech 7, 777–784. doi:10.1242/dmm.016089

El-Etr, S.H., Subbian, S., Cirillo, S.L.G., Cirillo, J.D., 2004. Identification of Two Mycobacterium marinum Loci That Affect Interactions with Macrophages. Infect Immun 72, 6902–6913. doi:10.1128/IAI.72.12.6902-6913.2004

Fan, M., Rao, T., Zacco, E., Ahmed, M.T., Shukla, A., Ojha, A., Freeke, J., Robinson, C.V.,

Benesch, J.L., Lund, P.A., 2012. The unusual mycobacterial chaperonins: evidence for in vivo oligomerization and specialization of function. Molecular Microbiology 85, 934–944. doi:10.1111/j.1365-2958.2012.08150.x

Fayet, O., Ziegelhoffer, T., Georgopoulos, C., 1989. The groES and groEL heat shock gene products of Escherichia coli are essential for bacterial growth at all temperatures. Journal of Bacteriology 171, 1379–1385.

Fenton, W.A., Kashi, Y., Furtak, K., Norwich, A.L., 1994. Residues in chaperonin GroEL required for polypeptide binding and release. Nature 371, 614–619. doi:10.1038/371614a0

Flynn, J.L., Chan, J., 2001. Tuberculosis: Latency and Reactivation. Infect. Immun. 69, 4195–4201. doi:10.1128/IAI.69.7.4195-4201.2001

Flynn, J.L., 2004. Immunology of tuberculosis and implications in vaccine development. Tuberculosis 84(1-2), 93-101. doi:10.1016/j.tube.2003.08.010

Forrellad, M.A., Klepp, L.I., Gioffre, A., Sabio y Garcia, J., Morbidoni, H.R., Santangelo, M. de la P., Cataldi, A.A., Bigi, F., 2013. Virulence factors of the Mycobacterium tuberculosis complex. Virulence 4, 3–66. doi:10.4161/viru.22329

Fossati, G., Izzo, G., Rizzi, E., Gancia, E., Modena, D., Moras, M.L., Niccolai, N., Giannozzi, E., Spiga, O., Bono, L., Marone, P., Leone, E., Mangili, F., Harding, S., Errington, N., Walters, C., Henderson, B., Roberts, M.M., Coates, A.R.M., Casetta, B., Mascagni, P., 2003. Mycobacterium tuberculosis chaperonin 10 is secreted in the macrophage phagosome: is secretion due to dissociation and adoption of a partially helical structure at the membrane? J. Bacteriol. 185, 4256–4267.

Friedland, J.S., Shattock, R., Remick, D.G., Griffin, G.E., 1993. Mycobacterial 65-kD heat shock protein induces release of proinflammatory cytokines from human monocytic cells. Clin Exp Immunol 91, 58–62.

Friedman, D.I., 1992. Interaction between bacteriophage λ and its Escherichia coli host. Current Opinion in Genetics & Development 2, 727–738. doi:10.1016/S0959-437X(05)80133-9

Gao, L.Y. & Manoranjan, J., 2005. Laboratory maintenance of Mycobacterium marinum. Curr Protoc Microbiol, Chapter 10: Unit 10B.1 doi: 10.1002/9780471729259.mc10b01s00

Guisbert, E., Herman, C., Lu, C.Z., Gross, C.A., 2004. A chaperone network controls the heat shock response in E. coli. Genes Dev. 18, 2812–2821. doi:10.1101/gad.1219204

Gupta, T.D., Bandyopadhyay, B., Gupta, S.K.D., 2008. Modulation of DNA-binding activity of Mycobacterium tuberculosis HspR by chaperones. Microbiology 154, 484–490. doi:10.1099/mic.0.2007/012294-0

Hemmingsen, S.M., Woolford, C., van der Vies, S.M., Tilly, K., Dennis, D.T., Georgopoulos, C.P., Hendrix, R.W., Ellis, R.J., 1988. Homologous plant and bacterial proteins chaperone oligomeric protein assembly. Nature 333, 330–334. doi:10.1038/333330a0

Hett, E.C., Rubin, E.J., 2008. Bacterial Growth and Cell Division: a Mycobacterial Perspective. Microbiol. Mol. Biol. Rev. 72, 126–156. doi:10.1128/MMBR.00028-07

Hickey, T.B.M., Thorson, L.M., Speert, D.P., Daffé, M., Stokes, R.W., 2009. Mycobacterium tuberculosis Cpn60.2 and DnaK are located on the bacterial surface, where Cpn60.2 facilitates efficient bacterial association with macrophages. Infect. Immun. 77, 3389–3401. doi:10.1128/IAI.00143-09

Hickey, T.B.M., Ziltener, H.J., Speert, D.P., Stokes, R.W., 2010. Mycobacterium tuberculosis employs Cpn60.2 as an adhesin that binds CD43 on the macrophage surface. Cellular Microbiology 12, 1634–1647. doi:10.1111/j.1462-5822.2010.01496.x

Horwich, A.L., Fenton, W.A., Chapman, E., Farr, G.W., 2007. Two families of chaperonin: physiology and mechanism. Annu. Rev. Cell Dev. Biol. 23, 115–145.

doi:10.1146/annurev.cellbio.23.090506.123555

Horwich, A.L., Low, K.B., Fenton, W.A., Hirshfield, I.N., Furtak, K., 1993. Folding in vivo of bacterial cytoplasmic proteins: role of GroEL. Cell 74, 909–917.

Hu, Y., Henderson, B., Lund, P.A., Tormay, P., Ahmed, M.T., Gurcha, S.S., Besra, G.S., Coates, A.R.M., 2008. A Mycobacterium tuberculosis Mutant Lacking the groEL Homologue cpn60.1 Is Viable but Fails To Induce an Inflammatory Response in Animal Models of Infection. Infect Immun 76, 1535–1546. doi:10.1128/IAI.01078-07

Ikeda-Kobayashi, A., Taniguchi, Y., Brockwell, D.J., Paci, E., Kawakami, M., 2012. Prying Open Single GroES Ring Complexes by Force Reveals Cooperativity across Domains. Biophys J 102, 1961–1968. doi:10.1016/j.bpj.2012.03.046

Islam, M.S., Richards, J.P., Ojha, A.K., 2012. Targeting drug tolerance in mycobacteria: a perspective from mycobacterial biofilms. Expert Rev Anti Infect Ther 10, 1055–1066. doi:10.1586/eri.12.88

Kieser, K.J., Rubin, E.J., 2014. How sisters grow apart: mycobacterial growth and division. Nature Reviews Microbiology 12, 550–562. doi:10.1038/nrmicro3299

Kong, T.H., Coates, A.R., Butcher, P.D., Hickman, C.J., Shinnick, T.M., 1993. Mycobacterium tuberculosis expresses two chaperonin-60 homologs. Proc Natl Acad Sci U S A 90, 2608–2612.

Kusner, D.J., 2005. Mechanisms of mycobacterial persistence in tuberculosis. Clinical Immunology, Renovations for an intracellular life style 114, 239–247. doi:10.1016/j.clim.2004.07.016

Lillebaek, T., Dirksen, A., Vynnycky, E., Baess, I., Thomsen, VØ, Anderssen, A.B., 2003. Satbility of DNA patterns and evidence of Mycobacterium tuberculosis reactivation occurring decades after the initial infection. The Journal of Infectious Diseases 188(70), 1032-9

Lin, Z., Rye, H.S., 2006. GroEL-Mediated Protein Folding: Making the Impossible, Possible. Crit Rev Biochem Mol Biol 41, 211–239. doi:10.1080/10409230600760382

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

Morita, M., Kanemori, M., Yanagi, H., Yura, T., 1999. Heat-Induced Synthesis of ς32 inEscherichia coli: Structural and Functional Dissection ofrpoH mRNA Secondary Structure. J. Bacteriol. 181, 401–410.

Multiple Gene Duplication and Rapid Evolution in the groEL Gene: Functional Implications - Springer, n.d. doi:10.1007/s00239-006-0037-7

Munk, M.E., Schoel, B., Modrow, S., Karr, R.W., Young, R.A., Kaufmann, S.H., 1989. T lymphocytes from healthy individuals with specificity to self-epitopes shared by the mycobacterial and human 65-kilodalton heat shock protein. J. Immunol. 143, 2844–2849.

Mustafa, A.S., Lundin, K.E., Oftung, F., 1993. Human T cells recognize mycobacterial heat shock proteins in the context of multiple HLA-DR molecules: studies with healthy subjects vaccinated with Mycobacterium bovis BCG and Mycobacterium leprae. Infect. Immun. 61, 5294–5301.

Parida, S.K., Axelsson-Robertson, R., Rao, M.V., Singh, N., Master, I., Lutckii, A., Keshavjee, S., Andersson, J., Zumla, A., Maeurer, M., 2014. Totally drug-resistant tuberculosis and adjunct therapies. J Intern Med n/a–n/a. doi:10.1111/joim.12264

Qamra, R., Srinivas, V., & Mande, S.C., 2004. Mycobacterium tuberculosis GroEL homologues unusually exist as lower oligomers and retain the ability to suppress aggregation of substrate proteins. J Mol Biol 342(2), 605-17

Rao, T., Lund, P.A., 2010. Differential expression of the multiple chaperonins of Mycobacterium smegmatis. FEMS Microbiology Letters 310, 24–31. doi:10.1111/j.1574-6968.2010.02039.x

Roseman, A.M., Chen, S., White, H., Braig, K., Saibil, H.R., 1996. The Chaperonin ATPase Cycle: Mechanism of Allosteric Switching and Movements of Substrate-Binding Domains in GroEL. Cell 87, 241–251. doi:10.1016/S0092-8674(00)81342-2

Saibil, H.R., Fenton, W.A., Clare, D.K., Horwich, A.L., 2013. Structure and Allostery of the Chaperonin GroEL. Journal of Molecular Biology, Allosteric Interactions and Biological Regulation (Part I) 425, 1476–1487. doi:10.1016/j.jmb.2012.11.028

Saita, N., Fujiwara, N., Yano, I., Soejima, K., Kobayashi, K., 2000. Trehalose 6,6′-Dimycolate (Cord Factor) of Mycobacterium tuberculosis Induces Corneal Angiogenesis in Rats. Infect. Immun. 68, 5991–5997. doi:10.1128/IAI.68.10.5991-5997.2000

Škovierová, H., Larrouy-Maumus, G., Pham, H., Belanová, M., Barilone, N., DasGupta, A., Mikušová, K., Gicquel, B., Gilleron, M., Brennan, P.J., Puzo, G., Nigou, J., Jackson, M., 2010. Biosynthetic Origin of the Galactosamine Substituent of Arabinogalactan in Mycobacterium tuberculosis. J. Biol. Chem. 285, 41348–41355. doi:10.1074/jbc.M110.188110

Stewart, G.R., Wernisch, L., Stabler, R., Mangan, J.A., Hinds, J., Laing, K.G., Young, D.B., Butcher, P.D., 2002. Dissection of the heat-shock response in Mycobacterium tuberculosis using mutants and microarrays. Microbiology 148, 3129–3138.

Stinear, T.P., Seemann, T., Harrison, P.F., Jenkin, G.A., Davies, J.K., Johnson, P.D.R., Abdellah, Z., Arrowsmith, C., Chillingworth, T., Churcher, C., Clarke, K., Cronin, A., Davis, P., Goodhead, I., Holroyd, N., Jagels, K., Lord, A., Moule, S., Mungall, K., Norbertczak, H., Quail, M.A., Rabbinowitsch, E., Walker, D., White, B., Whitehead,

S., Small, P.L.C., Brosch, R., Ramakrishnan, L., Fischbach, M.A., Parkhill, J., Cole, S.T., 2008. Insights from the complete genome sequence of Mycobacterium marinum on the evolution of Mycobacterium tuberculosis. Genome Res 18, 729–741. doi:10.1101/gr.075069.107

Straus, D., Walter, W., Gross, C.A., 1990. DnaK, DnaJ, and GrpE heat shock proteins negatively regulate heat shock gene expression by controlling the synthesis and stability of sigma 32. Genes Dev. 4, 2202–2209. doi:10.1101/gad.4.12a.2202

Suzuki, H., Ikeda, A., Tsuchimoto, S., Adachi, K., Noguchi, A., Fukumori, Y., Kanemori, M., 2012. Synergistic Binding of DnaJ and DnaK Chaperones to Heat Shock Transcription Factor ?32 Ensures Its Characteristic High Metabolic Instability. J Biol Chem 287, 19275–19283. doi:10.1074/jbc.M111.331470

Takayama, K., Wang, C., Besra, G.S., 2005. Pathway to Synthesis and Processing of Mycolic Acids in Mycobacterium tuberculosis. Clin. Microbiol. Rev. 18, 81–101. doi:10.1128/CMR.18.1.81-101.2005

Urano-Tashiro, Y., Yajima, A., Takashima, E., Takahashi, Y., Konishi, K., 2008. Binding of the Streptococcus gordonii DL1 surface protein Hsa to the host cell membrane glycoproteins CD11b, CD43, and CD50. Infect. Immun. 76, 4686–4691. doi:10.1128/IAI.00238-08

Van Eden, W., van der Zee, R., Prakken, B., 2005. Heat-shock proteins induce T-cell regulation of chronic inflammation. Nat Rev Immunol 5, 318–330. doi:10.1038/nri1593

Velayati, A.A., Farnia, P., Ibrahim, T.A., Haroun, R.Z., Kuan, H.O., Ghanavi, J., Farnia, P., Kabarei, A.N., Tabarsi, P., Omar, A.R., Varahram, M., Masjedi, M.R., 2009. Differences in cell wall thickness between resistant and nonresistant strains of Mycobacterium tuberculosis: using transmission electron microscopy. Chemotherapy

55, 303–307. doi:10.1159/000226425

WHO | Tuberculosis [WWW Document], n.d. WHO. URL http://www.who.int/mediacentre/factsheets/fs104/en/ (accessed 7.15.14).

WHO | Tuberculosis and HIV [WWW Document], n.d. WHO. URL http://www.who.int/hiv/topics/tb/en/ (accessed 7.15.14).

Xu, Z., Horwich, A.L., Sigler, P.B., 1997. The crystal structure of the asymmetric GroEL–GroES–(ADP)7 chaperonin complex. Nature 388, 741–750. doi:10.1038/41944

Zhou, Y.N., Kusukawa, N., Erickson, J.W., Gross, C.A., Yura, T., 1988. Isolation and characterization of Escherichia coli mutants that lack the heat shock sigma factor sigma 32. J Bacteriol 170, 3640–3649.

Zuber, U., Schumann, W., 1994. CIRCE, a novel heat shock element involved in regulation of heat shock operon dnaK of Bacillus subtilis. J. Bacteriol. 176, 1359–1363.

6. SUPPLEMENTARY INFORMATION

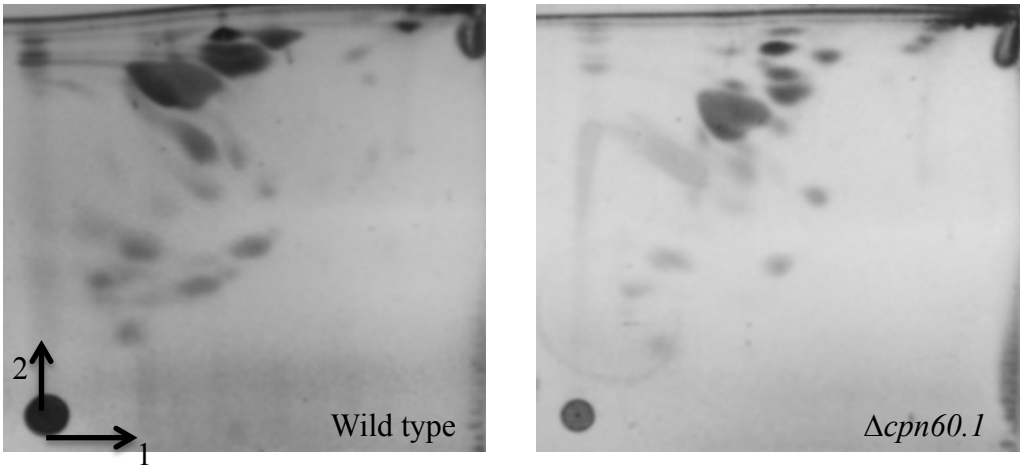
6.1 Further information regarding qRT-PCR primers

Name	Length (bp)	T_m (°C)	GC content (%)
Mm_cpn10_F	21	63.95	52.38
Mm_cpn10_R	20	64.09	60.00
Mm_cpn60_1_F	20	65.22	60.00
Mm_cpn60_1_R	22	64.74	54.55
Mm_cpn60_2_F	21	64.87	57.14
Mm_cpn60_2_R	21	64.57	57.14
Mm_gapdh_F	20	61.88	60.00
Mm_gapdh_R	20	61.89	50.00

Supplementary Table 1. Key properties of qRT-PCR primers. T_m refers to the melting temperature of the primer.

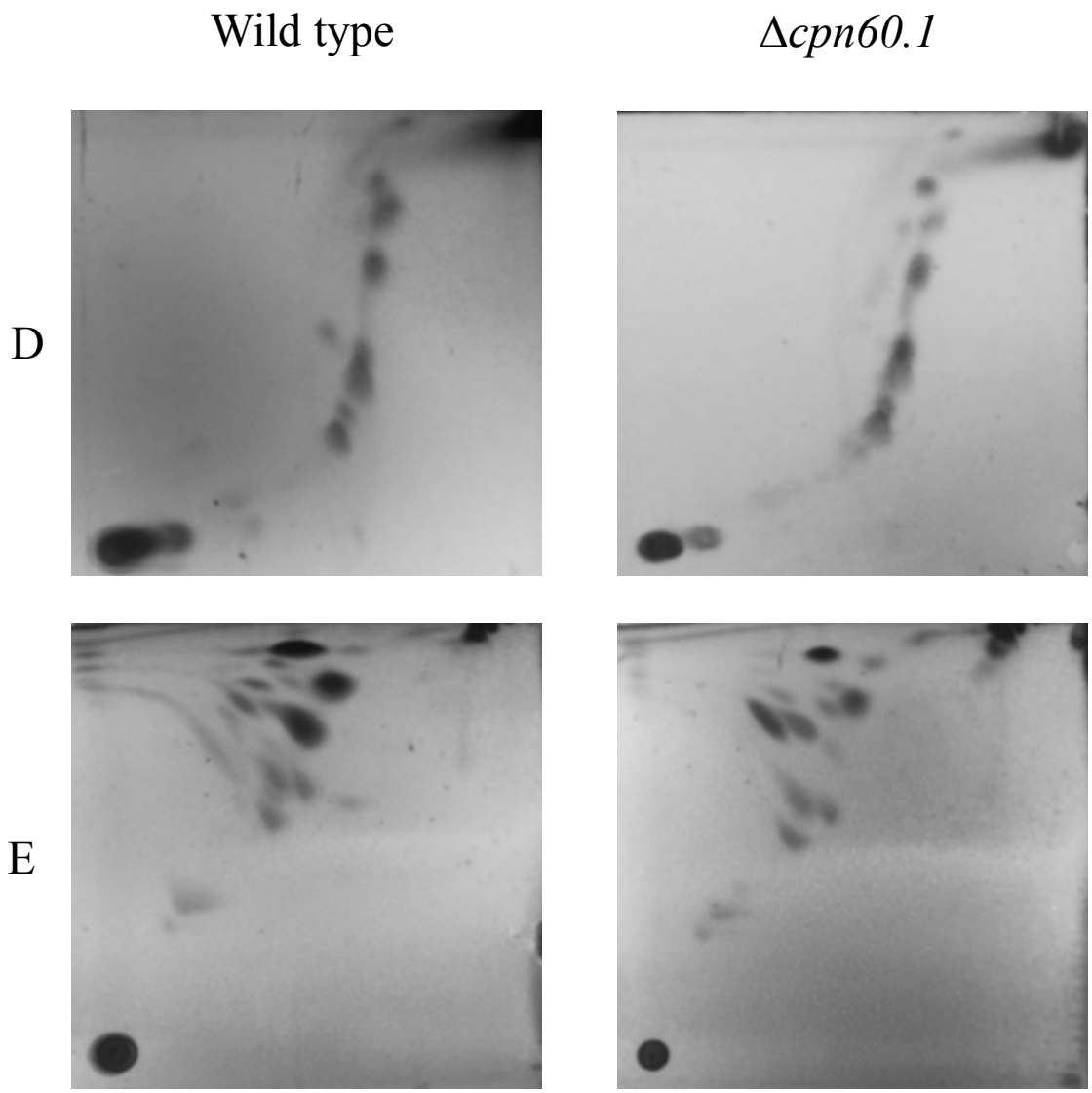
6.2 Additional TLC data

6.2.1 *M. marinum* polar lipids separated using system E



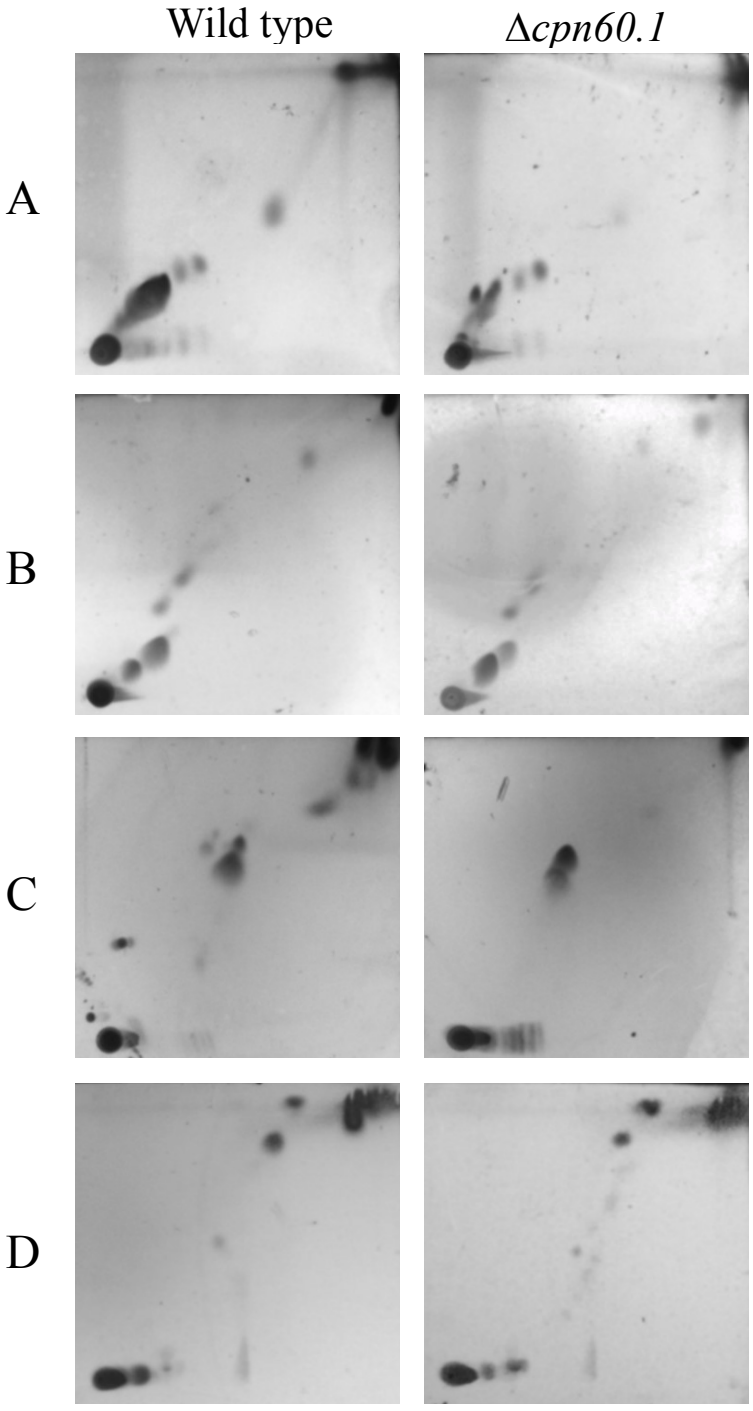
Supplementary Fig. 1 2D-TLC of polar lipid**s** purified from *M. marinum* wild-type and Δ*cpn60.1* strains

6.2.2 2D-TLC of Polar lipids from *M. smegmatis* wild type and Δcpn60.1 strains



Supplementary Figure 2. 2D-TLC of Polar lipids from *M. smegmatis* wild type (left) and Δcpn60.1 (right) strains using solvent systems D (top) and E (bottom).

6.2.2 2D-TLC of Apolar lipids from *M. smegmatis* wild type and Δcpn60.1 strains



Supplementary Figure 3. 2D-TLC of Apolar lipids from *M. smegmatis* wild type (left) and Δcpn60.1 (right) strains using solvent systems A-D (referred to by letters on left hand side of diagram).