Molecular mechanisms of DNA repair in *Mycobacterium tuberculosis*

Sophia Johanna Gessner

Division of Medical Microbiology
Department of Pathology

A thesis submitted to the Faculty of Health Sciences, University of Cape Town, in fulfilment of the requirements for the degree of Doctor of Philosophy. March 2016.

The search for reason ends at the shore of the known;
on the immense expanse beyond it
only the sense of the ineffable can glide.
It alone knows the route to that
which is remote from experience and understanding.
Neither is amphibious:
reason cannot go beyond the shore,
and the sense of the ineffable
is out of place where we measure, where we weigh…….

Citizens of two realms, we must all sustain dual allegiance:
we sense the ineffable in one realm;
we name and exploit reality in another.

Between the two we set up a system of references,
but can never fill the gap.
They are as far and as close to each other

As time and calendar, as violin and melody,
as life and what lies beyond the last breath.

The tangible phenomena we scrutinize with our reason,

The sacred and indemonstrable we overhear

with the sense of the ineffable.

Heschel A. J. (1971)

Declaration

I declare that this thesis is my own unaided work, it is being submitted for the degree of Doctor of Philosophy at the University of Cape Town. It has not been submitted for any degree of examination at any other university.

Signed by candidate

Signature Removed

Sophia J van Coller

2017-03-09

Date

Abstract

The mycobacterial DNA damage and repair pathways involved in the emergence of drug-resistance during host infection remain poorly understood, yet are critical to any efforts to develop novel "anti-evolution" drugs aimed at reducing the capacity of *Mycobacterium tuberculosis* to adapt genetically during tuberculosis (TB) treatment. The thesis presented here aimed to investigate the contribution of the DNA damage (SOS) response in adaptive mutagenesis, and focused on two specific components: the role of the specialist translesion synthesis DNA polymerase, DnaE2, in mutagenesis under stress and, secondly, the function of the mycobacterial homologue of a putative *SOS response associated peptidase* (SRAP) protein which has been identified in comparative genomics analyses of organisms possessing a DnaE2-type C family DNA polymerase. This work focused on the putative SRAP protein which was predicted to form part of the mycobacterial DNA damage response as a functional switch by binding DNA in an autoproteolytic dependent manner. To this end, SRAP deletion mutants were generated for both *M. smegmatis* (*MSMEG_1891*) and *M. tuberculosis* (*Rv3226c*). Despite the fact that SRAP was upregulated in both *M. smegmatis* and *M. tuberculosis* following genotoxic stress, no DNA damage phenotype was detected in any SRAP deletion mutant using a variety of DNA damaging agents. In parallel, an eGFP-tagged *M. smegmatis* SRAP allele was constructed to enable visualisation of SRAP upregulation and sub-cellular recruitment using fluorescent microscopy; however no eGFP expression could be visualised after MMC treatment. It was not clear whether this was due to faulty eGFP expression in the fusion protein, or to low-level induction of SRAP. In a biochemical approach to elucidate SRAP function, soluble *M. smegmatis* SRAP protein was expressed and purified using a N-terminal hexa-histidine tag. No proteolytic activity was detected in gelatine or casein zymography, perhaps indicating that SRAP has a very specific substrate. Moreover, while it was predicted that autocatalytic cleavage of the C-terminus was required for activation of SRAP, no such cleavage was detected using hexa-histidine tag staining, possibly pointing to a set of very specific conditions for activation. In combination, therefore, neither microbiological nor biochemical assays could elucidate a definitive role for SRAP in the mycobacterial DNA damage response.

DnaE2 has been directly implicated in induced mutagenesis to rifampicin (Rif) resistance in *Mycobacterium tuberculosis* following exposure of bacilli to genotoxic stress. In previous work in our group, a vitamin B_{12}-sensitive Δ*metH* strain was found to form "B_{12}-resistant" suppressor mutants at a frequency higher than could be explained by spontaneous mutagenesis alone. The first part of this thesis investigated the potential role of DnaE2 in the high-frequency emergence of B_{12}-resistance by mutating DnaE2 in the Δ*metH* background. Whereas elimination of polymerase function in a DnaE2^{AIA} mutant abrogated DNA damage-induced mutagenesis to Rif resistance, no change in B_{12} sensitivity was detected in a Δ*metH dnaE2*^{AIA} double mutant. PCR sequencing of spontaneous B_{12}-resistant mutants revealed mutations in genes previously associated with the suppressor phenotype; moreover, there was no apparent difference in the nature of mutations observed in both parental and *dnaE2*^{AIA} mutant strains. Instead, these results suggest that an alternative mechanism must exist to enable adaptive mutagenesis in methionine-starved mycobacteria.

Acknowledgements

I would like to thank my **Almighty Lord** for the grace and strength I require every day and for blessing me so richly.

I am grateful to, the National Research Foundation (**NRF**), the University of Cape Town (**UCT**) and the Molecular Mycobacteriology Research Unit (**MMRU**) for financial support and for contributing to local and international travel expenses.

To my supervisor, **Assoc. Prof Digby Warner**, thank you for your constant support and encouragement throughout my Ph.D. After changing research fields, I felt like I was lost at sea and through the challenging waters you helped change my way of thinking and inspired me to persevere. Thank you for making the journey exciting and fun.

To my co-supervisor, **Prof. Valerie Mizrahi**, your drive and knowledge of your field is inspirational. Thank you for your constant scientific guidance and advice. It is a privilege to be part of your research group.

My colleagues at the **MMRU**, I thank you for the practical assistance in the lab and the thought provoking conversations.

I would like to acknowledge **Dr. Česlovas Venclovas** and **Dr. Natasha Wood** for generating the SRAP protein models as well **Prof. Edward Sturrock** for his advice and input on SRAP structure and function. Thank you to **Prof. Trevor Sewell** and **Dr. Brandon Weber** for their invaluable guidance and help in attempting crystallisation as well as Mohamed Jaffer for TEM assistance. Lastly, I want to thank **Dr. Thomas Ioerger** for WGS done to confirm generated mutants.

To my wonderful father, **At**, thank you for your unwavering support, you are my rock and I none of this would have been possible without you. To my dearest mother, **Corli**, thank you for being my best friend and always showering me with love and being my comfort when things are though. I thank my beloved step mother, **Erika** and step father,

Rian, for your love and support. My brother, **Gert**, thank you for always believing in me, you are my inspiration. **Richard**, thank you for bringing balance to my life. Your easy-going nature has helped me through stressful and challenging times. Thank you for your love and believing in me.

Publications from this thesis

van Coller, S., J., Mizrahi, V., Warner, D., F. (2017) Investigating the role of SOS response associated peptidase (SRAP) in the mycobacterial DNA damage response. In preparation

van Coller, S., J., Ioerger, T., R., Mizrahi, V., Warner, D., F. (2017). DnaE2 is not required for starvation-induced adaptive mutagenesis in *Mycobacterium tuberculosis*. In preparation.

Table of Contents

List of Figures

List of Tables

Chapter 1: General Introduction

1.1 Tuberculosis

1.1.1 Disease burden

Mycobacterium tuberculosis (Mtb) is the most common causative agent of tuberculosis (TB) and the main infectious killer worldwide (WHO, 2016). TB is spread through the inhalation of Mtb-contaminated droplets from an infected individual's throat or lungs (WHO, 2016). Following inhalation, the bacilli are phagocytosed by alveolar macrophages as part of the host's innate immune response (Hmama *et al.*, 2015; Stallings & Glickman, 2010). The pathophysiology of Mtb is greatly dependent on its ability to evade the immune system, with specific reference to its ability of finding alternative routes of entry into macrophages, preventing macrophage maturation and interfering with immune recognition (Hmama *et al.*, 2015). This allows Mtb to replicate within the macrophage, causing primary infection (Harding & Boom, 2010). One of the most important aspects of Mtb virulence (Russell, 2001) is its ability to prevent macrophage maturation which is facilitated by blocking fusion between the infected macrophage and lysosomes, avoiding macrophage acidification (Hmama *et al.*, 2015). By doing so, Mtb evades proteolytic degradation and prevents antigen presentation that is required to activate the host adaptive immune response (Hmama *et al.*, 2015). Furthermore, Mtb improves its chances of survival by actively blocking apoptosis (Briken & Miller, 2008) and uses the necrotic fate of the macrophage to spread to neighbouring cells (Martin *et al.*, 2014). Infected cells trigger a pro-inflammatory response and granulomas are formed through the recruitment of lymphocytes, monocytes and fibroblasts (Gengenbacher & Kaufmann, 2012). However, in a healthy, immunocompetent host, infection is asymptomatic and could be maintained without progressing to active disease unless favourable conditions for reactivation are encountered. This type of manipulation of the host immune system allows latent TB infection (LTBI) of approximately one third of the world population (Gengenbacher & Kaufmann, 2012; Philips & Ernst, 2012).

Disease burden is heaviest in the developing world with 95% of deaths occurring in Third World countries. The death rate dropped a staggering 47% between 1990 and 2014, with 49 million lives saved between 2000 and 2015 through effective treatment (WHO, 2016). The global TB control policy, DOTS (*d*irectly *o*bserved *t*herapy, *s*hort course) employs combination chemotherapy consisting of a two month "intensive

phase" of four first-line drugs - isoniazid, rifampicin (Rif), pyrazinamide and ethambutol - followed by a four-month "continuation phase" comprising isoniazid and Rif treatment (Cox *et al.*, 2008). Treatment of latently infected individuals is even longer with a nine month isoniazid preventative treatment (Cardona, 2006). Due to the extensive nature of treatment, a great attempt was made as part of the DOTS program to improve patient adherence by offering support on different levels from socioeconomic and emotional to managing non-adherent patients (Yin *et al.*, 2016). Even though this picture seems promising in controlling drug-sensitive TB, the increase in emergence and spread of drug-resistant TB poses a major threat to current treatment regimens. The emergence of resistance is mainly due to extensive treatment duration, making patients more prone to non-adherence (Warner & Mizrahi, 2006). An estimated 48 000 people developed multidrug-resistant TB (MDR-TB) in 2015 with resistance against isoniazid and Rif. Treatment of MDR-TB relies on second-line drugs, which are less effective, more costly and can cause severe side effects. Further resistance to second-line antibiotics such as fluoroquinolones and at least one injectable (amikacin, kanamycin or capreomycin) is referred to as extensively drug-resistant TB (XDR-TB), which becomes extremely difficult to treat (Abubakar *et al.*, 2013; Dheda *et al.*, 2017; Eldholm & Balloux, 2016).

TB poses a major threat to individuals infected with the human immunodeficiency virus (HIV) in the developing world (Bruchfeld *et al.*, 2015). In 2015, TB was determined to be the leading killer of HIV-infected people, with one in every three deaths caused by TB (WHO, 2016). Susceptibility to TB was shown to arise very soon after HIV infection (Sonnenberg *et al.*, 2005) and, once co-infected, individuals require co-treatment (Lawn *et al.*, 2013; Pepper *et al.*, 2007). A variety of complications are associated with co-treatment, challenging treatment regimens further. Not only does co-treatment have a high pill burden (Pepper *et al.*, 2007) but drugs for TB and HIV show an overlap in toxicity profiles and are often associated with drug-induced liver injury, renal impairment and neuropathy (Lawn *et al.*, 2013). Also, co-treatment is commonly associated with immune reconstitution inflammatory syndrome (IRIS) (Bruchfeld *et al.*, 2015; Pepper *et al.*, 2007). There are certain important interactions between anti-TB and anti-retroviral treatment (ART) that need to be considered (Bruchfeld *et al.*, 2015). Rif was found to interfere with ART by influencing ART metabolism, causing heterogeneity in ART plasma levels of patients (Cabrera *et al.*,

2009). Co-treatment is of great concern since it could lead to drug resistance against both anti-TB drugs and ART (Pepper *et al*., 2007). A recent study showed no correlation between co-infection and the development of TB drug resistance; however whether HIV infected individuals were on ART is not clear (Eldholm *et al*., 2016).

1.1.2 New developments in treatment and prevention

It is apparent that new chemotherapeutic agents are needed to treat MDR-TB and XDR-TB as well as shorten treatment to limit future resistance acquisition (Borsari *et al.*, 2016; Hoagland *et al.*, 2016; Mikušová & Ekins, 2016). Unfortunately, it does not seem that the current pipeline for drug discovery and development holds significant promise in changing the current treatment regimen drastically in terms of treatment shortening (Mdluli *et al.*, 2015). New agents should not only have high bactericidal activity to shorten treatment, but should ideally also have a novel mechanism of action to prevent cross-resistance, good pharmacokinetic and pharmacodynamics properties to allow a single daily administration, and should not show any interactions with other drugs that can form part of combination therapy (Hoagland *et al*., 2016). Despite efforts by government organisations, non-government organisations and pharmaceutical companies, only a handful of alternatives – including bedaquiline (Cohen, 2013) and delamanid (Gler *et al*., 2012) - have been provisionally approved (Borsari *et al.*, 2016). However, delamanid has been associated with severe side-effects (Hoagland *et al*., 2016). By combining lessons learned from previous and ongoing drug discovery endeavours, improved drug discovery technology and a better understanding of TB biology and disease, there is much reason for hope of successful outcomes (Hoagland *et al*., 2016; Mdluli *et al*., 2015).

The emerging pandemic of drug-resistant Mtb stresses not only the desperate need for new drugs but also alternative routes to combat TB such as vaccine development (Kaufmann *et al*., 2014). BCG (bacillus Calmette-Guerin) is an attenuated strain of *M. bovis* and is administered as part of the childhood vaccine program in many countries (Mangtani *et al*., 2014; Scriba *et al*., 2016). BCG has been shown to provide partial protection in humans but its efficacy in preventing pulmonary TB, the greatest load from TB, varies greatly (Mangtani *et al*., 2014; Scriba *et al*., 2016; Trunz *et al.*, 2006). BCG showed the highest level of protection when administered to infants that have not been exposed to Mtb (Mangtani *et al*., 2014), which speaks to the desperate need for

developing new and improved vaccines (Kaufmann *et al*., 2014). More than 12 vaccine candidates are currently under clinical study (Evans *et al.*, 2016; Kaufmann *et al*., 2014). Vaccine candidates are categorised based on their target population as being either therapeutic or preventative. Therapeutic vaccines are targeted at patients with severe cases of TB such as MDR, XDR and HIV co-infected patients and are aimed at enhancing current drug treatment strategies (Kaufmann *et al*., 2014; Prabowo *et al*., 2013). Preventative vaccines are aimed at replacing BCG and must therefore show an increase in safety and/or efficacy (Kaufmann *et al*., 2014). Live attenuated and killed whole-cell vaccines (WCVs) have been shown to be advantageous in TB vaccine development, mostly because of their broad antigen composition and long lasting immune response (Scriba *et al*., 2016). An example of a therapeutic WCV is RUTI which is a polyantigenic liposomal vaccine targeted at LTBI to prevent progression to active disease (Cardona, 2006; Kaufmann *et al*., 2014; Scriba *et al*., 2016). A variety of preventative WCVs candidates are also in the process of being developed (Scriba *et al*., 2016). Although single vaccines are used in clinical trials, it might also be worth investigating combination vaccination strategies (Kaufmann *et al*., 2014).

There has been great interest in developing strategies to help infected individuals respond to Mtb appropriately to shorten treatment. This is reinforced by the fact that attempts to stop TB infection by means of vaccination and drug development are hampered by the lack of our understanding of early infection pathogenesis, host protection, host susceptibility and development of active disease (Kaufmann *et al*., 2014; Kiran *et al.*, 2016). Host-directed therapy (HDT) is not only aimed at shortening the treatment period, but also at preventing Mtb resistance acquisition and reducing lung damage (Wallis & Hafner, 2015). HDTs might have the added advantage for HIV and TB co-infected individuals of reducing the risk of developing IRIS (Zumla *et al*., 2015). There are a broad range of candidates with a variety of different targets such as stimulating autophagy (Ravikumar *et al*., 2004), promoting antimicrobial peptide synthesis (Rivas-Santiago *et al*., 2013), reducing inflammation (Mayanja-Kizza *et al*., 2005), and antibody based therapies (Balu *et al*., 2011). Furthermore, current drugs have limited entry into granulomas which limits treatment of LTBI (Wallis & Hafner, 2015), therefore granuloma formation and resolution are also targets for HDT (Kiran *et al*., 2016). However, the debate about whether granulomas are protective in terms of pathogen containment or whether they are destructive due to the permanent damage

caused to lung tissue from sustained inflammation (Pasipanodya *et al*., 2010) is steering research rather to develop HDT strategies that play a more protective role (Kiran *et al*., 2016). Regardless, development of HDTs does not bypass the current need for effective and efficient drugs and future research should be aimed at complementing new and current treatment regimens (Zumla *et al*., 2015).

1.1.3 Drug resistance acquisition

The extensive use of antibiotics has led to the exposure of microbes to sub-lethal concentrations of drugs which could be mutagenic (Andersson & Hughes, 2014). A gradient of antibiotic concentrations is established in the environment due to antibiotics intended for animal farming, plant production and human therapy that end up in rivers and soil (Andersson & Hughes, 2012; Taylor *et al*., 2011). Furthermore, low levels of antibiotics are also produced in natural environments as ecological weapons (Andersson & Hughes, 2014; Galan *et al.*, 2013). A similar gradient is established in the human body due to suboptimal dosage, poor pharmacokinetic properties and patient non-adherence (Andersson & Hughes, 2009). The minimum inhibitory concentration (MIC) is defined as the lowest concentration of an antibiotic that inhibits growth of a specified proportion of a bacterial population *in vitro* (Mouton *et al*., 2012); so, inhibition of 99% of the population is termed the MIC_{99} value, and this is commonly used in determining therapeutic doses. Exposure to sub-MIC levels within the concentration gradient of antibiotics can cause the expansion of antibiotic resistant subpopulations *in vivo* (McVicker *et al.*, 2014), selecting resistant bacteria which can be directly transmitted.

Not only does sub-MIC exposure cause expansion of resistant subpopulations but constant exposure has also been shown to generate *de novo* resistance (Gullberg *et al.*, 2011). Sub-MICs cause errors during translation reducing the growth rate (Davies *et al.*, 1964) that can lead to the appearance of drug tolerant/persister cells. These cells are slow or non-growing cells and are important in clinical settings since they showed reduced treatment efficiency (Brauner *et al*., 2016; Helaine *et al.*, 2014; Levin-Reisman *et al*., 2017). An important distinction should be made between tolerance, a population effect, and persistence which only applies to a subpopulation (Brauner *et al*., 2016). Not only do sub-MICs cause an increase in tolerant/persister cells when exposed to bactericidal concentrations of antibiotics (Cohen *et al.*, 2013; Dörr *et al.*, 2009), they

also render them less sensitive to antibiotics of other classes (Cohen *et al*., 2013; Johnson & Levin, 2013). Drug resistance developed from a completely susceptible population of *E. coli* cells in micro-environments that established an antibiotic gradient (Bos *et al.*, 2015; Zhang *et al*., 2011). However, it has also been established that, even when exposed to cidal drug concentrations, persister cells survive (Cohen *et al.*, 2013; Wakamoto *et al*., 2013). In *M. smegmatis* (Msm) it was shown that the cells do not only survive in the presence of isoniazid but divide, characterised by a steady number of cells which suggests a fine balance between cell division and cell death (Wakamoto *et al*., 2013). Lastly, antibiotics can act as signalling molecules between cells influencing functions such as gene expression and virulence (Davies, 2013; Dietrich *et al.*, 2008; Goh *et al*., 2002). It is not yet clear if the antibiotic itself is the signalling molecule or if it triggers the release of true signalling molecules. The latter was found to be true in *E. coli* where resistant cells exposed to antibiotics secreted indole which signalled the susceptible cells to increase efflux pump expression (Lee *et al.*, 2010), increasing resistance at a population level.

It has been widely accepted that a combination of intrinsic and genetically encoded resistance contribute to treatment failure of Mtb. Intrinsic resistance and genetic resistance are two very different phenomena that decrease drug susceptibility (Warner & Mizrahi, 2006). Intrinsic resistance is attributed to features such as the mycobacterial cell wall which acts as a permeability barrier to both hydrophilic and hydrophobic compounds (Niederweis *et al.*, 2010). Bacterial thiols, specifically mycothiol in Mtb, are redox systems that act as substrates for detoxification enzymes and have been implicated in tolerance to oxidative stress and survival in macrophages (Bhaskar *et al*., 2014; Newton *et al*., 2008). Furthermore, resistance to drugs has been directly linked to chemical modification by means of phosphorylation, glucosylation, ribosylation and methylation (Dabbs *et al*., 1995; Tanaka *et al*., 1996; Warrier *et al*., 2016). Membrane-associated proteins known as efflux pumps actively remove most classes of antimicrobials, reducing drug efficacy (Chim *et al.*, 2015; Milano *et al.*, 2009; Szumowski *et al.*, 2013). Mistranslation, an error that causes mRNA to be translated incorrectly (Reynolds *et al.*, 2010), could lead to phenotypic drug resistance (Javid *et al*., 2014). Another contributor could be transcriptional mutagenesis where the mRNA contains an error, especially if the encoded protein is involved in replication or DNA repair (Saxowsky & Doetsch, 2006). These are in contrast to drug resistance which is a

genetic hereditary trait (Warner & Mizrahi, 2006). Drug resistance in Mtb seems to arise exclusively through the acquisition of chromosomal mutations in the form of single nucleotide polymorphisms (SNPs), chromosomal rearrangements and gene amplification (Almeida Da Silva & Palomino, 2011; Farhat *et al*., 2016; Hershberg *et al*., 2008; Sandgren *et al*., 2009; Witney *et al*., 2016); there is no evidence for horizontal transfer and plasmids in Mtb (Galagan, 2014).

Mtb was long thought to be clonal (Achtman, 2012), therefore variation observed post infection was attributed to host genetics and environmental factors (Warner *et al*., 2015). For this reason, bacterial genetics and genotypic variation was not deemed important in treatment efficiency (Warner *et al*., 2015). However, whole-genome sequencing (WGS) has transformed the field of Mtb genomics, revealing not only strain diversity (Gagneux, 2013; Stucki & Gagneux, 2013), but also differences in host adaptation between lineages (Gagneux, 2012; Hershberg *et al*., 2008), suggesting a major role for bacterial genomics in heterogeneous outcomes. Furthermore, microvariation within a host (Casali *et al*., 2014; Walker *et al*., 2013) was observed which was associated with mixed infections especially in high burden settings (Bryant *et al*., 2013). However, there is also evidence to suggest that the diversity developed from a single infection (Lieberman *et al.*, 2016; Pérez-Lago *et al*., 2014; Sun *et al*., 2012). This substantiates previous evidence that resistance to different drugs can develop in a lesion infected with a single susceptible strain (Kaplan *et al*., 2003). In fact, more diversity was observed within a patient (Pérez-Lago *et al*., 2014) than between strains from different epidemiological clusters (Walker *et al*., 2013). Although contradictory results were obtained (Saunders *et al*., 2011) the methods used might have limited the detection of low frequency SNPs (McGrath *et al*., 2014; Warner *et al*., 2015). Substantiated evidence for microdiversity suggests a pivotal role for diversification in disease progression (Warner *et al*., 2015). Importantly, high levels of heterogeneity in drug-resistance conferring mutations were also reported; however, since only cultured bacteria were considered, a metagenomics approach will be required to reveal the true diversity (Koch *et al.*, 2014; Sun *et al*., 2012).

An important, unanswered question remains whether spontaneous mutagenesis alone can account for the high frequency of drug resistance observed. The first studies that calculated the Mtb mutation rate *in vitro* estimated it to be similar to other bacteria

(Boshoff *et al*., 2003; David, 1970; Ford *et al*., 2011; McGrath *et al*., 2014). Other studies have subsequently used WGS to attempt the estimation of the *in vivo* mutation rate. In a study by Ford *et al*. (2011), the estimated *in vivo* mutation rate in macaques was comparable to the *in vitro* mutation rate, suggesting no state of hypermutation in the host (Ford *et al*., 2011). This was substantiated by a retrospective observational study where the *in vivo* mutation rate was estimated to be 0.5 SNPs/genome/year, which is similar to the *in vitro* estimation as well as the *in vivo* estimation in macaques (Walker *et al.*, 2013). In addition to this, results from Saunders *et al.* supported the idea of a low mutation rate *in vivo* but, since the strategy they used might not have detected low frequency mutations, a more sensitive approach had to be employed (Saunders *et al*., 2011). By increasing the sensitivity, a high degree of diversity and heterogeneity was observed in sputum samples over the course of treatment (Sun *et al.*, 2012). Furthermore, the generation time to mutation rate was similar between active and latent disease which suggests that either both generation time and mutation rate are the same during latent and active disease or, if the generation time decreases during latent disease, as expected, an increased mutation rate would be required (Ford *et al*., 2011). In contrast to what was observed in macaques, Colangeli *et al.* (2014) reported a higher generation time to mutation rate ratio during active disease than latent disease in humans. However, a major limitation of this work was that only two LTBI individuals were studied (Colangeli *et al.*, 2014). It therefore remains largely unclear whether the mutation rate is increased during infection to account for the high rate of drug resistance acquisition, as well as whether this rate differs between disease stages (McGrath *et al.*, 2014). This reiterates the importance of studying mutagenesis not only *in vivo* but also *in vitro* to gain better insight into mechanisms that could contribute to adaptation.

Discrepancies between mutation rate *in vitro* and *in vivo* could be indicative of factors that either increase the mutation rate during infection or promote diversity within the host (Warner *et al.*, 2015). It is difficult to calculate mutation rate modulation *in vivo* during infection due to difficulty in distinguishing between the rate at which mutations are generated and the rate at which resistance gets selected (McGrath *et al.*, 2014). Constitutive mutators could provide a short-term selective advantage even though these mutators are expected to disappear over time (McGrath *et al*., 2014). Furthermore, it still remains unclear if certain Mtb strains have an elevated mutation rate. Beijing strains were found to have mutations in genes that can cause a mutator effect (Ebrahimi-

Rad *et al*., 2003). In another study, mutations were also identified in genes involved in replication, DNA repair and recombination which suggests an increased mutation rate (Mestre *et al*., 2011). Furthermore, it was revealed that East Asian strains that are associated with a higher frequency of drug resistance showed a higher basal mutation rate than the Euro-American lineages, even in the absence of antibiotics (Ford *et al.*, 2013). However, contrasting results were obtained *in vitro*, where no differences were observed between the Beijing and non-Beijing lineages (Werngren *et al.*, 2003). These inconsistencies could be because the strains used were not representative of the lineages studied, and so more research is required to address these differences (Mokrousov, 2014). Some evidence suggests that resistance conferring mutations increase the basal mutation rate, for example, the S522L mutation in *rpoB* (Anthony *et al.*, 2005). In addition to permanent genetic (heritable) mutators, transient mutagenesis could lead to a temporary increase in mutation rate (McGrath *et al*., 2014). The DNA damage response (section 1.3) could be a major contributor to adaptive evolution of Mtb since it was shown that the deletion of DnaE2 reduced Rif resistance (Rif^R) acquisition (Boshoff *et al.*, 2003), as will be discussed later in section 1.2.3. This becomes particularly important since Mtb encounters a range of DNA damaging agents upon infection such as reactive oxygen species and reactive nitrogen intermediates produced by host macrophages as a defence response (Ehrt & Schnappinger, 2009). These can cause DNA damage, triggering the SOS response which can lead to base mispairings (Gorna *et al*., 2010; Kurthkoti & Varshney, 2010). A better understanding of these components could yield drug targets that limit bacterial evolution (anti-evolution drugs), protecting current drugs (Smith and Romesberg, 2011; Warner, 2010). There is some evidence that suggests promise to the approach, this includes disruption of DNA repair (Wigle, 2009) and eliminating induction of SOS response (Georgescu, 2008). DnaE2 has been suggested as an excellent candidate for such an approach in Mtb (Warner, 2010).

1.2 Maintenance of genome integrity

Mutations, although imperative for adaptation and evolution, are expected to be mostly neutral or deleterious (Sherman & Gagneux, 2011). The fidelity of genome replication is greatly influenced by the accuracy at which replicative polymerases incorporate nucleotides (Arana & Kunkel, 2010). Mismatches due to polymerase errors and DNA

damage that remain unfixed can lead to mutations or could be lethal if they block the replication fork from passing (Warner *et al*., 2013). The repair pathway employed depends on the type of DNA damage encountered (Davis & Forse, 2009). In general, base modifications are targeted by either base excision repair (BER) or nucleotide excision repair (NER), whereas strand breaks are fixed by homologous recombination (HR) or non-homologous end joining (NHEJ) (Davis & Forse, 2009; Dos Vultos *et al*., 2009). Like *E. coli*, Mtb uses NER, BER and recombination; unlike *E. coli*, Mtb also uses NHEJ and, until recently, was thought to lack a post-replication mismatch repair (MMR) system (Mizrahi and Andersen, 1998; Warner *et al.*, 2014). Loss of MMR is typically associated with an increased occurrence of frame-shift mutations, especially in simple repeat sequences (Davis & Forse, 2009; Springer *et al*., 2004). MMR is critical in rejecting recombination between divergent sequences (Tham *et al*, 2016). In contrast to *Helicobacter pylori*, which lacks MMR, the Mtb genome is quite stable and without major diversity (Dos Vultos *et al.*, 2009; Kang & Blaser, 2006) showing compensation through strict control of HR initiation (Springer *et al.*, 2004) and selective codon use (Wanner *et al.*, 2008). This suggests that there is a nonorthologous system to MMR (McGrath *et al*., 2014; Mizrahi & Andersen, 1998) or that an alternative system such as the recently discovered novel MMR specific endonuclease in archaea (Ishino *et al*., 2016) contributes to the fidelity of the Mtb genome. In fact, it was recently reported that a putative endonuclease, NucS, is involved in maintaining the genome integrity of Mtb. Elimination of NucS function showed a similar phenotype to canonical MMR system deletion mutants (Castañeda-García *et al*., 2017). Further research is required to establish definitively whether this system functions as a *de facto*, non-canonical MMR system in Mtb.

An important distinction should be made between DNA damage repair and DNA damage tolerance. Whereas DNA repair refers to accurate repair and restoration to the original Watson-Crick base pairing, DNA damage tolerance does not always restore the original base pair but rather allows lesions to be tolerated by bypassing replication blocking lesions with the help of translesion synthesis (TLS) polymerases (Warner *et al*., 2013). However, this process can be mutagenic due to nucleotide misincorporations and strand slippage and thus requires tight control (Davis & Forse, 2009; Warner *et al*., 2013).

1.2.1 Replication fidelity

The major replicative polymerase in Mtb is DNA polymerase III (Boshoff *et al*., 2003; Warner *et al.*, 2014). It consists of a Pol III holoenzyme, a β_2 sliding clamp (processivity factor) and the clamp-loader complex (Gu *et al*., 2016; McHenry, 2011; O'Donnell, 2006; Warner *et al*., 2014). In turn, the holoenzyme consists of three subassemblies of *dnaE*-encoded α, *dnaQ* encoded ε exonuclease and θ, which is responsible for stabilising the ε-subunit (McInerney *et al.*, 2007; Rock *et al*., 2015; Warner *et al*., 2014). The level of fidelity at which the polymerase replicates the genome depends on the selectivity of nucleotides incorporated, and the removal of misincorporated nucleotides by either the intrinsic exonuclease activity of the polymerase or exonuclease activity of an interacting partner (Arana & Kunkel, 2010; Warner *et al*., 2014). With an estimated mutation rate of 10^{-10} (Boshoff *et al*., 2003; Ford *et al*., 2011; McGrath *et al*., 2014), a mutation is expected to occur only once every 10^{10} replication events. In *E.coli* the nucleotide selectivity of the polymerase is thought to be the major contributor to replication fidelity (Fijalkowska *et al.*, 2012).

Major replicative polymerases have high nucleotide selectivity and tightly fit only the nucleotide with correct base pairing in the nucleotide-binding pocket (Arana & Kunkel, 2010). This is in contrast to specialist polymerases – for example, TLS polymerases - that have high error rates due to the relaxed geometry of the nucleotide binding pocket (Donigan *et al.*, 2014; Yang, 2003; Yang & Woodgate, 2007). Some TLS polymerases have even been found to incorporate a certain nucleotide preferentially, discounting correct base pairing (Bebenek *et al.*, 2001). Incorrect base pairing hinders strand elongation that causes the polymerase to activate the exonuclease activity (Reha-Krantz, 2010). The Mtb replicative polymerase, DnaE1, has a polymerase and histidinol phosphatase (PHP) domain that is responsible for proofreading, and it appears that the *dnaQ* encoded ε subunit does not contribute to replication fidelity in Mtb (Rock *et al.*, 2015).

1.2.2 DNA damage repair

Mtb encounters extremely challenging and hostile environments during infection due to the host immune response and possible antibiotic treatment (Gorna *et al*., 2010; Warner, 2010; Warner *et al*., 2014). A suite of different endogenous and exogenous radicals and oxidants, in the form of reactive oxygen and nitrogen species are

encountered (Darwin & Nathan, 2005; Gorna *et al*., 2010). These compounds can compromise the integrity of bacterial lipids (Rubbo *et al.*, 1994), proteins (Alvarez & Radi, 2003; Ischiropoulos *et al.*, 1992) and nucleic acids (Yermilov *et al.*, 1995) as part of the host defence response (Darwin & Nathan, 2005; Gorna *et al*., 2010; Szabó, 2003). Mtb has to withstand the genotoxic stresses encountered in order to establish infection by employing a complement of DNA repair pathways which may act according to a functional hierarchy (Dos Vultos *et al*., 2009). Transcriptional analyses suggest active DNA repair throughout the course of infection (Gorna *et al*., 2010; Warner, 2010) in mouse infection models (Talaat *et al.*, 2004) and clinical samples (Rachman *et al.*, 2006), stressing the importance of DNA repair in Mtb survival and proliferation.

1.2.2.1 Direct reversal

Direct reversal is the simplest form of DNA repair, and requires only one protein and acts without incising the backbone (Yi & He, 2013). However, only certain types of lesions can be fixed by means of direct reversal (Friedberg *et al.*, 1995). The protein Ogt is an O^6-alkylguanine-DNA alkyltransferase that transfers an alkyl group from the O^6-position of a guanine to a cysteine and is present in both Mtb and Msm (Davis & Forse, 2009). In addition to Ogt, a similar protein is also present in Msm with a different alkyltransferase function (Davis & Forse, 2009).

1.2.2.2 Excision repair

BER is executed in four steps (Krokan & Bjøra, 2013; van der Veen & Tang, 2015). The damaged base is removed by cleavage of the glycosylic bond which links the base to the sugar phosphate backbone by an appropriate glycosylase, leaving an abasic site. The sugar phosphate backbone is then cleaved, adjacent to the abasic site, by an endonuclease which leaves DNA polymerase I to fill the gap and DNA ligase to seal the nick. Different glycosylases act on different types of DNA damage and mispaired bases; for example, in Mtb (Figure 1.1), TagA function is substrate specific whereas AlkA has a broad substrate range (Davis & Forse, 2009; Dos Vultos *et al*., 2009; O'Brien & Ellenberger, 2004; Kurthkoti & Varshney, 2011). Expression of *alkA* is induced by H_2O_2 (Boshoff *et al*., 2003) and to a lesser degree in macrophages (Schnappinger *et al.*, 2003). However, deletion of *alkA* did not influence the ability of the bacterium to replicate in a mouse model, stressing the importance of functional redundancy of repair enzymes (Durbach *et al*., 2003). Loss of AlkA function can be compensated for by TagA and/or Mpg (Warner *et al*., 2013). Glycosylases that remove

oxidised DNA fall within two families: Fpg/Nei family and Nth superfamily (Guo *et al*., 2010). Mtb has four homologs of the Fpg/Nei family of glycosylases (*fpg1*, *fpg2*, *nei1* and *nei2*) as well as one Nth homolog (Guo *et al*., 2010; Moolla *et al.*, 2014; Warner *et al*., 2013). The glycosylase, Fpg, repairs 8-oxo-7,8-dihydroguanine (8-oxo-G), commonly known to mispair with adenine and cause G -> T transversions (Jain *et al.*, 2007), whereas MutY removes the adenine once mispaired (Guo *et al*., 2010; Jain *et al*., 2007). Uracil is often found in DNA due to deamination of cytosine causing G -> A transitions. Uracil DNA glycosylase (Ung) is responsible specifically for repairing these mismatches (Venkatesh *et al*., 2003) although this function can also be fulfilled by UdgB. An additional homolog of Mpg, similar to the eukaryotic 3-methyladenide DNA glycosylases, is found in mycobacteria (Davis & Forse, 2009; Dos Vultos *et al*., 2009; Kurthkoti & Varshney, 2011). An interplay between Msm Nth and Fpg/Nei homologs was confirmed to contribute to spontaneous mutagenesis by showing that combined deletion of Fpg/Nei and Nth increased the rate of spontaneous mutagenesis (Moolla *et al.*, 2014). Following removal of the damaged base by the appropriate glycosylase, two endonucleases End (endonuclease IV) and Xth (endonuclease III) process the lesion prior to gap filling. Homologs of both endonucleases are present in the Mtb complex, with Msm having an additional homolog of XthA, the endonuclease with 3' exonuclease activity (Davis & Forse, 2009; Dos Vultos *et al*., 2009). Endonuclease IV was shown to play a dominant role in protecting Mtb against oxidative damage (Puri *et al.*, 2013).

Crosslinks caused by cyclobutane pyrimidine dimers form as a result of UV and interstrand crosslinks are caused by bifunctional chemical reagents such as mitomycin C (MMC) forming bulky lesions (Reardon & Sancar, 2005). These are typically removed by NER, in three steps, using four proteins (Kisker *et al.*, 2013). Firstly, an UvrA dimer recognises the damage and associates with UvrB for verification. Upon UvrB binding, UvrA is released and UvrC is recruited, facilitating incision of the DNA backbone upstream and downstream of the damage. UvrD, a superfamily I helicase, aids in the release of the incised oligomer, after which the gap is filled by DNA polymerase I and sealed by DNA ligase (Kisker *et al*., 2013; Truglio *et al.*, 2006). Homologs of all four proteins are present within the Mtb complex (Davis & Forse, 2009; Kurthkoti & Varshney, 2011). An *in vitro* study identified UvrB, a component of

the NER pathway, as vital for survival in these toxic environments (Darwin *et al*., 2003; Ehrt & Schnappinger, 2009). Furthermore, a potential fusion protein with homology to the N-terminus of the UvrC protein was identified in Mtb (Rv2191) and is thought to play a role in reactivation in a rabbit model (Kesavan *et al*., 2009; Kurthkoti & Varshney, 2011). NER is also often coupled with transcription, removing lesions specifically from the transcribed strand. This coupling is mediated by Mfd, which recognises stalled RNA polymerases and recruits NER machinery (Kamarthapu & Nudler, 2015; Roberts & Park, 2004). Unusual properties of the Mtb Mfd homolog suggest a species-specific role during transcription coupled repair (Prabha *et al.*, 2011).

1.2.2.3 Recombination repair

HR is used to repair single- and double-stranded gaps that are left due to repair post replication (Bell & Kowalczykowski, 2016; Pagès, 2016; Wigley, 2013). Recombination repair uses the second undamaged strand in the helix as a template, allowing error-free repair (Smith, 2004). The protein complex, RecBCD, combines helicase (RecB and RecD) and nuclease (RecB) activities to create 3' single-stranded tails, a substrate for RecA, from blunt-ended double-strand breaks (DSBs) (Bell & Kowalczykowski, 2016; Wigley, 2013; Wilkinson *et al.*, 2016). It does so by recognising a Chi site close to the damage and preferentially degrades the 5' end, leaving a 3' tail (Bell & Kowalczykowski, 2016; Wigley, 2013). RecA is then actively loaded onto the 3' tail to allow strand invasion and subsequent strand exchange (Kowalczykowski, 2000; Wigley, 2013). In contrast to RecBCD, RecFOR preferentially repairs single-stranded breaks (SSBs); however, in the absence of RecBCD, RecFOR can repair DSBs as well (Pagès, 2016). In this pathway, the proteins function sequentially rather than in a complex like RecBCD (Rangarajan *et al.*, 2002). In short, RecF and RecR bind at the break while RecO displaces the single-stranded binding protein (SSBP), after which RecFOR loads RecA onto the single-stranded tail (Bell & Kowalczykowski, 2016). RecFOR is thought to play a major role in mycobacterial recombination since Δ*recO* was shown to be extremely sensitive to a range of clastogens (Glickman, 2014; Gupta *et al*., 2013). Roles for both Mtb RecF and RecR were experimentally confirmed (Gupta *et al.*, 2015). Both RecBCD and RecFOR form complex structures known as D-loops and Holiday junctions that remain to be resolved (Nautiyal *et al.*, 2016; Pagès, 2016). The proteins RuvABC and the helicase RecG complete the repair process by resolving these complex structures. RuvAB and

RecG migrate the Holiday junction to facilitate base pair exchange between the two homologous duplexes and RuvC, an endonuclease, resolves and separates the helices again (McGlynn & Lloyd, 2002; West, 1995). Homologs of RuvABC and RecG are present in Mtb, and are thought to form part of the minimal functional pathway (Davis & Forse, 2009; Singh *et al.*, 2016). Mtb also has an additional homology dependent error-free alternative to HR known as single strand annealing (SSA) which is independent of RecA (Glickman, 2014; Gupta *et al.*, 2011; Warner & Mizrahi, 2011). Furthermore, RecBCD does not contribute to HR in Mtb but forms part of the SSA pathway with AdnAB providing the helicase-nuclease function required for HR instead of RecBCD (Glickman, 2014; Gupta *et al.*, 2016; Sinha *et al*., 2009; Warner & Mizrahi, 2011).

NHEJ is used as an alternative for DSB repair and requires only Ku and DNA ligase to fuse ends without significant homology (Deriano & Roth, 2013; Rodgers & Mcvey, 2016). Ku binds to the broken ends and not only limits degradation and processes the ends, but also recruits additional elements required for NHEJ (Gong *et al*., 2005; Rodgers & Mcvey, 2016). Ku can bind both blunt ends and overhangs and exclusively stimulates the ligase activity of LigD through direct interaction (Weller *et al.*, 2002). In addition to a ligase domain, LigD has a nuclease and a polymerase domain, capable of utilising both ribonucleotides (rNTPs) and deoxyribonucleotides (dNTPs), making it important for DNA repair during dormancy and dNTP limiting conditions (Gong *et al.*, 2005; Weller *et al.*, 2002). NHEJ can be mutagenic and is commonly associated with simple insertions and deletions (Deriano & Roth, 2013; Rodgers & Mcvey, 2016). Ku and LigD functions were confirmed in Mtb when joining of DNA ends was observed *in vitro* (Weller *et al*., 2002) and was found to play a crucial role in the repair of DSBs in late stationary phase (Glickman, 2014; Stephanou *et al*., 2007). LigD activity was also confirmed to contribute to the efficiency of NHEJ *in vitro*; however, when the ligase activity of LigD was inactivated, LigC could seal the nicks (Aniukwu *et al.*, 2008). The *Rv3730c* gene, located adjacent to LigC has been identified as a standalone homolog of the LigD polymerase domain, however, no specific function has been assigned (Zhu *et al*., 2012).

1.2.3 DNA damage tolerance - Translesion synthesis (TLS)

Specialised TLS polymerases most commonly belong to the Y-family of DNA polymerases; for example, DNA damage-inducible UmuD'$_2$C (Pol V) and DinP (Pol IV) in *E. coli*. In Mtb, DinB1 (Rv1537) and DinB2 (Rv3056) (Pol IV homologs) are not DNA damage inducible (Kana *et al*., 2010). Msm also harbours copies of DinB1 (MSMEG_3178) and DinB2 (MSMEG_1002) as well as an additional Y family polymerase, DinB3 (MSMEG_6405). DinB2 seems to have a unique ability to utilise rNTPs which may be vital for survival during dNTP limiting conditions (Ordonez *et al.*, 2014). It was suggested that DinB2 might contribute to mycobacterial mutagenesis (Ordonez & Shuman, 2014), even though a *dinB2* deletion mutant did not show any DNA damage sensitivity or change in spontaneous mutagenesis (Kana *et al*., 2010). The major TLS polymerase in mycobacteria is a C family polymerase, DnaE2. This enzyme is DNA damage inducible and has been shown to contribute to inducible mutagenesis (Boshoff *et al*., 2003). This polymerase lacks key residues in the PHP domain, obliterating the exonuclease and proofreading activity (Ditse *et al.*, 2015; Rock *et al.*, 2015). In addition to DnaE2, an operon consisting of *Rv3395c-Rv3394c* was shown to be damage inducible and required for induced mutagenesis (Warner *et al*., 2010). These two genes are found in the same operon as DnaE2 in *Caulobacter cresecentus* (Galhardo *et al.*, 2005). The function of the cassette was anticipated since the absence of this cassette in organisms with an UmuDC system suggested an analogous function. The *Rv3394c* gene encodes ImuB, a putative Y-family polymerase which lacks the highly conserved carboxylates required for polymerase function. The C-terminus was shown to play a pivotal role in protein-protein interactions with itself (ImuB), DnaE1, DnaE2 and Rv3395c (ImuA'). Furthermore, a DinB3-type motif was identified in ImuB, facilitating interaction with the β-clamp, making it crucial for DnaE2 function through the interaction with the replication fork (Warner *et al*., 2010). Homologs of Rv3395c are found only in a limited amount of species and the function of its product, ImuA', remains unknown. It loosely resembles RecA but appears not to interact with itself to allow filamentation like RecA. It does, however, interact with ImuB through the C-terminus and so is thought to perform a specialist role that remains to be defined (Warner *et al*., 2010). Functional disruption of any one of the three components of the mutasome cassette (DnaE2, ImuA' and ImuB) was shown to eliminate induced mutagenesis in Mtb (Warner *et al*., 2010). This is in contrast to what was seen in *Pseudomonas putida* where elimination of DnaE2 led to an increase in

mutation frequency, indicative of a role as an anti-mutator (Koorits *et al*., 2007). Expression of *dnaE2*, *imuA'* and *imuB* forms part of the Mtb DNA damage response, perhaps indicating a role in survival under genotoxic conditions (Boshoff *et al*., 2003; Warner *et al*., 2010).

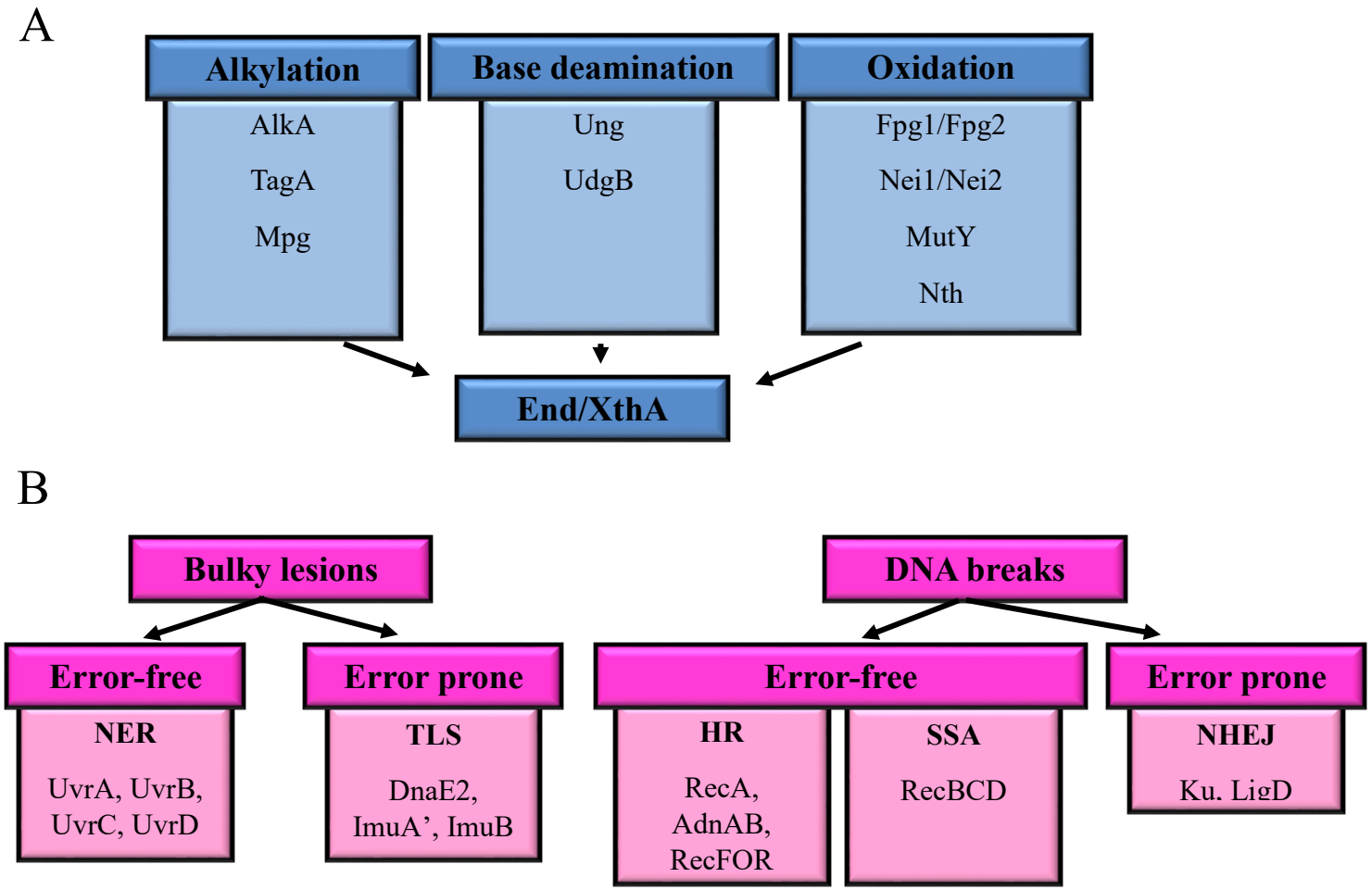


Figure 1.1 Proteins involved in Mtb DNA repair. A. Proteins involved in base excision repair (BER). Different types of lesions are recognised by different proteins followed by endonuclease (End/XthA) processing. **B.** DNA repair pathways involved in the repair of bulky lesions and DNA strand breaks. Figure adapted from Warner *et al.*, 2013.

1.3 DNA damage response

1.3.1 SOS response in mycobacteria

The SOS response (Figure 1.2) is best characterised in *E.coli*, the organism in which it was first discovered (Radman, 1974; Baharoglu & Mazel, 2014). Regulation of the SOS response depends primarily on two proteins, RecA and LexA. LexA plays a central role in SOS response regulation by acting as a repressor for itself and the rest of the SOS inducible genes (Butala *et al.*, 2009). Under normal growth conditions, a LexA dimer (Erill *et al.*, 2007) blocks transcription of the SOS-regulated genes by binding a DNA sequence termed the SOS-box in the operator of those genes (Baharoglu & Mazel, 2014; Butala *et al*., 2009). The N-terminus of LexA is crucial for SOS-box recognition, whereas the C-terminus allows dimerization (Neher *et al.*, 2003). LexA-mediated inhibition can be obliterated by the accumulation of single-stranded DNA (ssDNA) that accumulates during both normal growth and stress conditions (Baharoglu & Mazel, 2014; Foster, 2007). The ssDNA tails assemble RecA in an ATP-dependent manner. The co-protease activity of RecA is activated upon binding the single-stranded tails which then interacts with LexA, causing autocatalytic cleavage of the LexA dimer (Courcelle *et al*., 2001). Cleavage of LexA reduces the affinity of LexA for DNA and exposes vulnerable residues for protease digestion (Neher *et al.*, 2003). This alleviates LexA-mediated repression, allowing expression of SOS-regulated genes. The timing and expression level of a gene depends on the affinity of LexA for the SOS-box, the strength of the promoter, and the location of the operator in relation to the promoter (Qin *et al*., 2015; Simmons *et al*., 2008; Williams & Schumacher, 2016). To prioritise error-free repair, only a weak interaction exists between LexA and the SOS-box of genes required for NER and HR when the SOS response is weak (Erill *et al*., 2007; Janion, 2008; Williams & Schumacher, 2016). However, if DNA damage persists and the error-free mechanisms of repair do not eliminate damage sufficiently, RecA nucleoprotein filaments accumulate, reducing the LexA pool further, allowing expression of more strongly repressed genes such as *sulA* and the TLS polymerases. SulA inhibits septum formation, thereby preventing cell division and segregation of the damaged chromosomes (Simmons *et al.*, 2008). Expression of the specialist TLS polymerases allows for mutagenic repair, potentially sacrificing fidelity for survival (Erill *et al*., 2007; Qin *et al*., 2015). As DNA damage is repaired, the co-protease

activity of RecA is lost, allowing the accumulation and binding of LexA to the SOS-boxes once again (Qin *et al*., 2015).

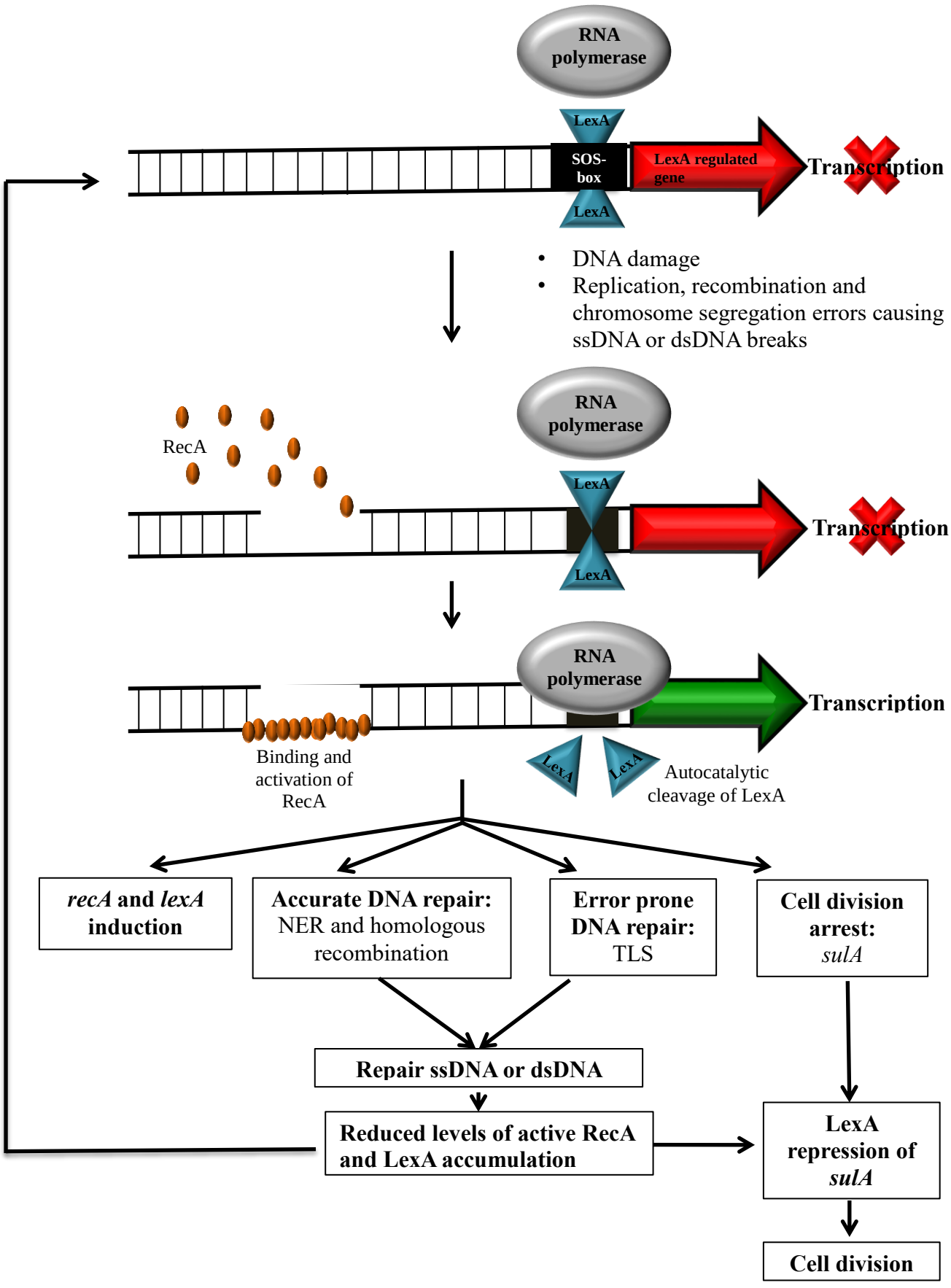

RNA polymerase
LexA
SOS-box
LexA regulated gene
LexA
Transcription
• DNA damage
• Replication, recombination and chromosome segregation errors causing ssDNA or dsDNA breaks
RecA
RNA polymerase
LexA
LexA
Transcription
RNA polymerase
Transcription
Binding and activation of RecA
LexA
LexA
Autocatalytic cleavage of LexA
recA and lexA induction
Accurate DNA repair: NER and homologous recombination
Error prone DNA repair: TLS
Cell division arrest: sulA
Repair ssDNA or dsDNA
Reduced levels of active RecA and LexA accumulation
LexA repression of sulA
Cell division

Figure 1.2 Schematic representation of the SOS response in *E. coli* and Mtb. Under normal cellular growth a LexA dimer binds the SOS-box in the operator upstream of the regulated genes, suppressing expression. Upon DNA damage activated RecA causes autocatalytic cleavage of the LexA dimer, relieving suppression followed by expression of SOS response genes. As DNA damage gets repaired the level of activated RecA is reduced which allows accumulation and subsequent binding of LexA.

The presence in mycobacteria of a classical, RecA/LexA SOS response, was first suspected upon the identification of the *recA* gene and a full complement of genes required for a SOS response except *umuD* and *polB* (Durbach *et al.*, 1997; Mizrahi & Andersen, 1998). Mtb *recA* is larger, producing a precursor protein that requires post-translational splicing to provide mature RecA (Davis *et al.*, 1991; Davis *et al*., 1992; Kumar *et al.*, 1996). In addition to this, *recA* is controlled by two promoters, a distal $RecAP_2$ with a typical SOS-box (Davis *et al*., 2002; Movahedzadeh *et al.*, 1997) and a proximal $RecAP_1$ that has no LexA binding site (Gopaul *et al.*, 2003). LexA is also functionally conserved (Durbach *et al*., 1997) and the SOS response in Mtb, like *E. coli*, can be mutagenic; in Mtb, though, DnaE2 contributes to induced mutagenesis and not the Y-family polymerases (Boshoff *et al*., 2003). Only ~25 genes are regulated by the SOS response (Smollett *et al.*, 2012) as opposed to the *E. coli* SOS response that includes more than 40 genes (Qin *et al*., 2015).

The LexA-binding site was first identified in *Bacillus subtilis* and named the Cheo-box (Cheo *et al.*, 1991). The consensus sequence of the Mtb SOS-box was subsequently defined as $TCGAAC(N_4)GTTCGA$ using a site-directed mutagenesis approach which allowed the identification of a putative set of RecA/LexA-regulated genes (Davis *et al.*, 2002). SOS-boxes associated with DNA damage-inducible genes were identified upstream of open reading frames (ORFs), overlapping translational start sites as well as within the coding region. This implies that LexA is capable not only of stopping transcription initiation, but also of aborting transcription once it has started (Davis *et al.*, 2002). In a subsequent study, combining chromatin immunoprecipitation and high-throughput sequencing, all LexA binding sites across the Mtb genome were identified (Smollett *et al.*, 2012). These SOS-box sequences are listed in Table 1.1. The majority of SOS-inducible genes belong to the 13E12 repeat family indicative of mobile elements. These mobile elements can take advantage of the SOS response to escape

threatening DNA damaging conditions (Davis *et al.*, 2002). Furthermore, during mobile element propagation, DNA breaks are formed that need to be repaired and, by coupling propagation with the DNA damage response mobile elements can overcome this threat and make use of the bacterial repair systems (Gamulin *et al.*, 2004). It is striking that very few of the DNA damage-inducible genes have well defined functions since the DNA damage response was found crucial in drug resistance acquisition in Mtb (Boshoff *et al*., 2003) and *E. coli* (Bos *et al*., 2015; Zhang *et al*., 2011).

Table 1.1 Genes with an SOS-box in Mtb

Gene[a]	**Product**	**SOS-box sequence**[b]	**Location**[c]
Rv0058	DnaB	tcgaaTatgcgttcgG	-44
Rv0071	Possible maturase	tcgaaTatgagttcga	-6
Rv0095	Conserved hypothetical protein	tcgCTAacatgttcga	internal
Rv0336	Conserved 13E12 repeat family protein	tcgaaAgtatgttcga	-32
Rv0515	Conserved 13E12 repeat family protein	tcgaaAgtatgttcga	-32
Rv1000c	Conserved hypothetical protein	tcgaacgaatgtGcga	-8
Rv1057	Conserved hypothetical protein	tcgaaAatatAttcga	-535
Rv1378c	Conserved hypothetical protein	tcgaacacatgttcga	+129
Rv1588c	Partial REP13E12 repeat protein	tcgCTAacatgttcga	-6
Rv1702c	Conserved hypothetical protein	tcgaacatgtAttcga	-6
Rv2100	Conserved hypothetical protein	tcgCacacatgttcga	+106
Rv2517c	Unknown protein	tcgCTcaaatAttcga	-92
Rv2578c	Conserved hypothetical protein	tAgaacggttgttcga	-24
Rv2579	DhaA	tcgaacaaccgttcTa	-99
Rv2592c/Rv2593c	RuvB/RuvA	GAgaacacccgttcga	-364
Rv2594c	RuvC	tcgaacgattgttcgG	-37
Rv2719c	Possible conserved membrane protein	tcAaacatgtgttcga	-218
Rv2720	LexA	tcgaacacatgttTga	-105
Rv2737c	RecA	tcgaacaggtgttcgG	-123

Rv3074	Conserved hypothetical protein	tcgaacacatgttcga	-8
Rv3260c	WhiB2	tcgaTcacttgttcga	-186
Rv3370c	DnaE2	tcgaacaattgttcga	-65
Rv3394c/Rv3395c	ImuB/ImuA'	tcgaacatatTttcga	-30
Rv3776	Conserved hypothetical protein	tcgaacgtatgttcga	-8

[a] Genes identified to have a LexA binding site using ChIP-seq in Smollet *et al*., 2012. Genes indicated in **bold** are co-regulated by both pathways.
[b] The LexA binding SOS-box sequence identified by Davis *et al.*, 2002 where capital letters indicate deviations from the consensus TCGAAC(N_4)GTTCGA.
[c] Location of the SOS-box with regards to the first nucleotide of the predicted initiation codon as reported in Davis *et al.*, 2002.

1.3.2 RecA-ND response in mycobacteria

Unlike *E. coli*, Mtb possesses two DNA damage response mechanisms. A non-classical mechanism, independent of RecA and LexA, was first described in *Acinetobacter calcoaceticus* and has since been referred to as the RecA-ND response (Rauch *et al.*, 1996). The majority of Mtb DNA damage response genes were found to be induced even in the absence of RecA (Gamulin *et al*., 2004; Rand *et al*., 2003) and therefore regulated by the RecA-ND mechanism. An unidentified σ factor (Gamulin *et al*., 2004) is thought to bind the consensus sequence, Rec-NDp (TTGTC(G/A)GTGN_8TAN_3T), upstream of RecA-ND regulated genes. It was later shown that the transcriptional regulator, ClpR (Rv2745c), binds the RecA-NDp and has an essential role in the induction of DNA repair genes (Wang *et al*., 2011). Genes with an identifiable RecA-NDp (Table 1.2) can be classified into four classes. Firstly, DNA repair, which includes genes contributing to BER, NER and recombination. Secondly, like the SOS response, a number of mobile elements also have a RecA-NDp. Thirdly, genes contributing to transcriptional regulation such as SigH and SigG might indirectly regulate additional genes which could explain the absence of a RecA-NDp in some genes which show RecA-independent induction. The fourth class consists of genes that are grouped as "other", again with many contributors having unknown functions (Gamulin *et al*., 2004; Rand *et al*., 2003). One such gene is predicted to encode the putative SOS-

response associated peptidase (SRAP) (Aravind *et al.*, 2013; Gamulin *et al.*, 2004), which is investigated in detail in Chapter 3. Notably, the two damage response pathways show partial overlap with some genes regulated by both mechanisms: e.g., *recA, ruvC* and *dnaB* (Gamulin *et al.*, 2004; Rand *et al.*, 2003; Smollett *et al.*, 2012).

Table 1.2 Genes with a RecA-NDp recognition sequence in Mtb

Gene[a]	**Product**	**RecA-NDp[b]**	**Location[c]**	**Experimental evidence (Rand *et al.* 2003)[d]**
Rv0054	SSB	Ctgtcggtgcggttacgtaggct	-77 (45)	Δ*recA* < WT
Rv0058	DnaB	Gtgtcggtgcggcgacttacgct	-79 (49)	Δ*recA* < WT
Rv0095c	Conserved hypothetical protein	ttgtcAgtgcccgacgttatgat	+85	Δ*recA* < WT
Rv0181c	SigG	GtgtcggAgactctgcgtaggct	+85 (109)	Δ*recA*=WT
Rv0184	Conserved hypothetical protein	CtgtcgCAgccctccactagttt	-30	Δ*recA* < WT
Rv0427c	XthA	CAgtcggtCggctgggatacctt	-29	Δ*recA*=WT
Rv0605	Possible resolvase	ttgtcggtggcgtgctgtaggaA	-6	Δ*recA* < WT
Rv1128c	Conserved hypothetical protein	Ttgtcggtgcctgcacatagcat	-29	Δ*recA*=WT
Rv1148c	Conserved hypothetical protein	ttgtcAgtggtgcttcgtagaat	+54	Δ*recA* < WT
Rv1277	Conserved hypothetical protein	Ttgtcggtgtgcagacctacact	+42 (29)	Δ*recA*=WT
Rv1317c	AlkA	Atgtcggtgggcgggtgtaattg	-30	None

Rv1406	Probable methionyl-tRNA formyltransferase Fmt	ACgtcggCggcgcttcttagact	-29	Δ*recA*=WT
Rv1547	DnaE1	GtgtcggtCctcggtcatagact	-32	None
Rv1588c	Partial REP13E12 repeat protein	ttgtcAgtgcccgacgttatgat	-29	Δ*recA* < WT
Rv1633	UvrB	GtgtcCgCgggtggctctaggct	-93 (61)	Δ*recA*=WT
Rv1638	UvrA	ttgtcggtacgggcacatagcat	-71 (39)	Δ*recA*=WT
Rv1765c	Conserved hypothetical protein	ttgtcAgACcctgctgcGatgat	-29	Δ*recA*=WT
Rv1833c	Possible haloalkane dehalogenase	AtgtcACCCttcgggcataccat	-47 (15)	Δ*recA*=WT
Rv1907c	Hypothetical protein	ttgtcAgACcccgcgtgCatggt	+92 (29)	Δ*recA* > WT
Rv1945	Conserved hypothetical protein	ttgtcAgtggtgcttcgtagaat	-30	Δ*recA* < WT
Rv1955	Possible toxin HigB	ttgtcgCtgccctccggtaccgt	-78 (46)	Δ*recA* > WT
Rv2015c	Conserved hypothetical protein	tCgtcAgGCaacgcccaCacgat	+210	Δ*recA*=WT
Rv2016	Conserved hypothetical protein	ttgtcAgCggataccggtatatt	-66 (34)	Δ*recA* > WT
Rv2024c	Conserved hypothetical protein	GtgtcggtTgccgccagtatggt	-32	Δ*recA*=WT

Rv2100	Conserved hypothetical protein	Ctgtcggtggccgctactagtat	+84	Δ*recA* < WT
Rv2119	Conserved hypothetical protein	ttgtcggCgcccaaccctaagct	-28	Δ*recA* > WT
Rv2428	AhpC	ttgGcggCgatgccgatAaatat	-116 (84)	Δ*recA* > WT
Rv2594c	RuvC	AtgtcggGAcgatccgctagcgt	-61 (29)	Δ*recA* < WT
Rv2719c	Possible conserved membrane protein	Gtgtcggtggtgtggtctaccgt	-127 (95)	Δ*recA* < WT
Rv2734	Conserved hypothetical protein	ttgtcggAgccgctgtctagcat	-71 (39)	Δ*recA* > WT
Rv2737c	RecA	ttgtcAgtggctgtctctagtgt	-77 (45)	N/A
Rv2792c	Possible resolvase for IS 1602	ttgtcAgtggcgtgctgtaggaA	-30	Δ*recA*=WT
Rv2975c	Conserved hypothetical protein	CtgtcACCAcgtctggttaggct	-74 (51)	Δ*recA*=WT
Rv2979c	Possible resolvase	ttgtcAgtggcgtgcggtaggaA	-30	Δ*recA*<WT
Rv3048c	NrdF2	ttgtcggACcccaggggtagtgt	-70 (38)	Δ*recA* < WT
Rv3198c	UvrD2	CtgtcATCTaccgctgcCaagat	-30 (125)	Δ*recA*=WT
Rv3223c	SigH	GtgtcggtCgcggcctctacatt	-157	None
Rv3226c	Conserved hypothetical protein	GtgtcgTGggtaccaggCaccat	-30	Δ*recA*=WT
Rv3263	Probable DNA methylase	AtgtcggtAgctcccttGatact	+43	Δ*recA* > WT
Rv3296	Lhr	ttgtcggtgcctgttggCacgat	-29	Δ*recA*=WT

Rv3466	Conserved hypothetical protein	ttgtcggtggcccctcgtagaat	-30	Δ*recA*<WT
Rv3517	Conserved hypothetical protein	Gtgtcggtgcacgacggcacact	-29	Δ*recA*=WT
Rv3555c	Conserved hypothetical protein	tGtcggtgggtgatgcCaggat	-28	Δ*recA*=WT
Rv3585	RadA	ttgtcACACccccccgatacggt	-29	Δ*recA* < WT
Rv3644c	Possible DNA polymerase	CtgtcggCgcggatagctactgt	-31	Δ*recA*=WT
Rv3714c	Conserved hypothetical protein	tGtcggtgcggagtgcCagcat	+16	Δ*recA*=WT
Rv3828c	Possible resolvase	ttgtcggtggtgtgctgtaggaA	0	Δ*recA*<WT

[a] Genes predicted to be regulated by the RecA-ND pathway by identification of RecA-NDp recognition sites. Genes indicated in **bold** are co-regulated by both pathways.
[b] The RecA-NDp recognition sequence where capital letters indicate the recognition sequence deviating from the consensus sequence TTGTCRGTG(N_8)TA(N_3)T.
[c] Location of the RecA-NDp sequence in relation to the first nucleotide in the predicted start codon. Nucleotides before initiation codon as recorded by Gamulin *et al*., 2004 is indicated in brackets.
[d] Microarray data from Rand *et al.*, 2003 studying gene expression in wildtype and Δ*recA*. Δ*recA*<WT indicates higher expression in wildtype than Δ*recA*, Δ*recA*=WT indicates the same levels of expression in the two strains and, lastly, Δ*recA*>WT indicates genes expressed independently of RecA and showing higher levels of expression in Δ*recA* than wildtype.

1.4 Adaptive mutagenesis

Adaptive mutagenesis has been the topic of great debate for many years (Foster, 2007; Maisnier-Patin & Roth, 2015; Rosenberg, 2001; Roth *et al.*, 2006). It is described as the acquisition of mutations upon encountering non-lethal stresses, such as nutrient deprivation, allowing cell survival and replication (Foster, 2000; Foster, 1999; Rosenberg, 2001). The most widely used system for the study of adaptive mutagenesis is the *E. coli* FC40 strain due to the high rate of mutation acquisition (Foster, 2000). Many years of studying FC40 has brought forth two major models attempting to explain adaptive mutation (Figure 1.3) (Foster, 2000; Foster, 1997; MacLean *et al.*, 2013; Maisnier-Patin & Roth, 2015; Rosenberg & Hastings, 2004; Rosenberg, 2001; Roth *et al.*, 2006; Stumpf *et al.*, 2007). In the first model, nutrient deprivation causes an evolved stress response pathway to create new point mutations in non-growing cells. Here, the cells do not grow under selective pressure but survive and induce a set of genes, including the TLS polymerases, that introduces mutations (Foster, 2007; Rosenberg, 2001). The second model relies on mutations forming in cells with a gene duplication before application of the selective pressure rather than in response to the pressure. The growth rate of cells with duplications is improved and thus, upon encountering stress, amplification is selected for (Roth *et al.*, 2006). This is referred to as "amplification under selection" where selection appears like mutagenesis, driving evolution and adaptation without altering the mutation rate (Maisnier-Patin & Roth, 2015; Roth *et al*., 2006). A major limitation of most systems used to study adaptive mutation – in some ways, quite similar to studying Mtb mutation rate modulation during infection – is the inability to discern between the contribution of selection and mutation (McGrath *et al*., 2014; Roth, 2010; Yosef *et al*., 2016). In other words, to determine the extent to which a state of hypermutation is induced as opposed to the effect that selection has on shaping the high *apparent* mutation rate. Recently, a novel system capable of separating the effects of mutation and growth limitation was used to show that growth limitation selects for mutation but does not induce it (Yosef *et al.*, 2016). However, even though this model brings new perspective in separating selection from DNA metabolism and possible mutagenesis, it might not be representative of the more complex systems.

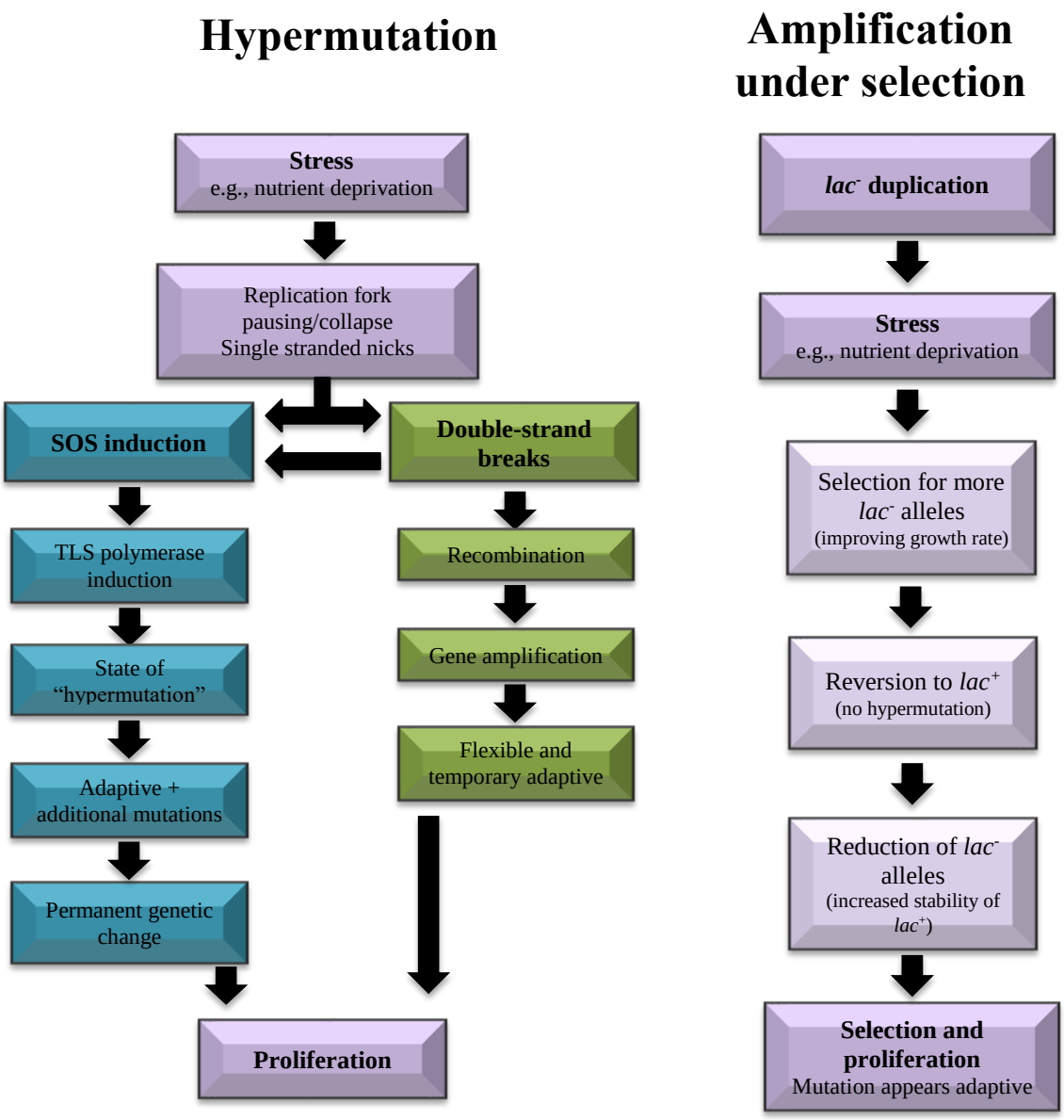


Figure 1.3 Models for adaptive genetic change. The two major models representing contrasting views of how adaptive mutagenesis is brought about. The figure is adapted from Rosenberg *et al*., 2001 and Roth *et al*., 2006.

There are numerous systems that support the hypermutation model, but do not exclude amplification. One such example is accumulation of mutations in aging colonies of *E. coli*, also known as *m*utagenesis in *a*ging *c*olonies (MAC) (Bjedov *et al.*, 2003). Rif^R was used as a marker to measure mutagenesis after a day and also a week of growth. A genome-wide increase in mutagenesis was observed due to the stress conditions encountered in a week old colony (Bjedov *et al.*, 2003). The authors concluded that MAC is genetically encoded and subject to environmental dependent diverse selective

pressures which impact the adaptive evolution of bacteria. The importance of the SOS response was elucidated when 95% of the observed mutations were found to be RecA dependent and 73% dependent on the error-prone Pol II (Bjedov *et al.*, 2003). This was consistent with the hypothesis for hypermutation as previously discussed. Contradictory results were obtained in a later study (Wrande *et al.*, 2008), suggesting selection instead of mutagenesis caused an increase in Rif^R frequency. Experiments by Bjedov *et al*. and Wrande *et al*. were repeated by studying nalidixic $acid^R$ in addition to Rif^R (Katz & Hershberg, 2013); using WGS, it was suggested that increased mutagenesis could not explain the high frequency resistance observed across the population. In agreement with Wrande *et al.*, it was found that certain mutations confer a growth advantage, and it was proposed that mutations occurring early on in the experiment, natural selection and the growth advantage of some mutations could contribute to the high frequency of Rif^R and nalidixic $acid^R$ (Katz & Hershberg, 2013).

MAC, also known as stationary phase mutation, has also been studied extensively in *Bacillus* (Robleto *et al.*, 2007) and *Pseudomonas* (Kivisaar, 2010) species. In *B. subtilis*, stationary phase mutation was studied under amino acid starvation and was found to be quite different to that of *E. coli* (Sung & Yasbin, 2002). Two genes, encoding transcription factors regulating competence, were found to be involved in regulating some aspects of stationary phase mutation. Furthermore, unlike *E. coli*, RecA was not required although mutation was TLS polymerase dependent (Robleto *et al.*, 2007; Sung & Yasbin, 2002). In this sense, the polymerase could be regulated in a RecA-independent manner. It was proposed that only a small sub-population contributed to the state of hypermutation that was strictly controlled by means of communication between cells similar to competence development (Sung & Yasbin, 2002). Although this was an important contributor, additional factors such as oxidative stress, a major stress encountered during stationary phase, was also found to contribute (Gomez-Marroquin *et al*., 2015; Vidales *et al.*, 2009). Furthermore, stationary phase mutation was associated with certain aspects of transcription. Transcription allows coupling of physiology and genetic variation, allowing mutation accumulation in highly transcribed genes without increasing the mutation rate across the whole genome which could cause undesired mutations (Ross *et al*., 2006). It was shown that a deletion mutant lacking Mfd, the transcription repair coupling factor, was associated with a decreased mutation frequency (Gomez-Marroquin *et al*., 2016; Ross *et al*., 2006); the deficiency

in transcription termination and elongation factor increased mutations (Ross *et al.*, 2006).

In a *Pseudomonas* model relying on the restoration of a promoter allowing phenol metabolism, an increase in mutations was also observed during stationary phase (Kasak *et al.*, 1997; Kivisaar, 2010). Like in *Bacillus*, stationary phase mutation was RecA-independent even though Pol IV was shown to contribute significantly to simple deletions, again pointing to a RecA-independent manner of regulation of Pol IV (Tegova *et al.*, 2004). The introduction of Pol V also increased the frequency of substitutions (Saumaa *et al.*, 2007). Surprisingly, ImuB and DnaE2 were shown to have opposite roles during stationary phase mutation, with ImuB inducing mutation and DnaE2 acting as an anti-mutator (Koorits *et al.*, 2007). As in *Bacillus*, oxidative stress greatly exacerbated stationary phase mutation (Saumaa *et al.*, 2007). Lastly, the NHEJ components, Ku and LigD, were found to contribute to MAC, with LigD seemingly functioning independently of Ku. This could be because the LigD polymerase domain, previously implicated in lesion bypass, also contributes to TLS without requiring Ku. (Paris *et al.*, 2015). In a recent study, a set of known and novel genes impacting mutation rate during stationary phase was obtained by combining a papillation assay with transposon mutagenesis (Tagel *et al.*, 2016). This assay is a promising tool to help unravel the genetic adaptation network in *Pseudomonas* (Tagel *et al*., 2016).

Like *Bacillus* and *Pseudomonas*, a state of hypermutation was also identified in stationary phase Msm (Karunakaran & Davies, 2000). Furthermore, translational misreading was also found to increase in stationary phase Msm (Leng *et al.*, 2015). Mistranslation, although not mutagenic, can be a major contributor to the adaptive response of Mtb (Javid *et al.*, 2014). It is apparent that very little is known about mycobacterial adaptive evolution and, given the complex nature of adaptive responses of *Bacillus* and *Pseudomonas*, it highlights the need for research into stationary phase mycobacteria. The development of tools, similar to the system that was developed to identify genes contributing in stationary phase mutation in *Pseudomonas*, will be of great value to the TB field. Furthermore, there is currently no other system that can be used to identify and study adaptive mutation in Mtb in addition to MAC. Studying multiple systems will help in addressing the different aspects of adaptive mutation

under different conditions to inform a comprehensive understanding of the underlying process.

Despite the apparent state of hypermutation in stationary phase Msm (Karunakaran & Davies, 2000), it is important not to disregard amplification since it can have far reaching consequences. Although quite difficult to detect, gene duplications are thought to be a common event (Sandhu & Akhter, 2015). These duplications can act as a reservoir for variation and can be stabilised if they confer an adaptive advantage when the cells are exposed to selective pressures (Domenech *et al*., 2014; Sandhu & Akhter, 2015). A major 350 kb genome duplication including more than 300 genes was identified in the "modern" sub-lineages of lineage 2 (W/Beijing) (Domenech *et al.*, 2010); however, even though it was speculated that this duplication arose only under host applied pressures, it was shown to occur more readily *in vitro* than *in vivo*. Regardless, this points to the ability of Mtb to evolve rapidly (Domenech *et al*., 2014). Amplification was directly implicated in the evolution of MmpL (*M*ycobacterial *M*embrane *P*rotein *L*arge) which is involved in lipid transfer and efflux (Sandhu & Akhter, 2015) and has been shown to contribute to Mtb viability and virulence (Chim *et al*., 2015; Domenech *et al.*, 2005). MmpL has also been linked to drug efflux-dependent azole resistance (Chim *et al*., 2015; Milano *et al*., 2009). With regards to genetic resistance, griselimycin resistance was mapped to duplication of the *dnaN* region (Kling *et al*., 2015). Furthermore, resistance to a potent drug that causes slow lysis of Mtb also mapped to a duplication of 40 genes (Park *et al*., 2016). In these examples, amplification might be the only viable option for Mtb when a targeted mutation cannot be readily accommodated to confer the resistance phenotype (Park *et al*., 2016). This highlights the potentially significant threat amplification poses to treatment regimens and the major gap in our understanding of amplification and how it can be exploited to prevent adaptation and improve treatment strategies (Kling *et al*., 2015).

1.5 General remarks

A high level of drug resistance, caused by chromosomal alterations such as mutations (Almeida Da Silva & Palomino, 2011; Koch *et al*., 2014; Warner *et al.*, 2015) is observed in Mtb, even though the organism is not considered to be a natural mutator (Ford *et al*., 2011). It remains unclear how resistance is obtained at such a high level and how the mutation rates differ between different stages of disease (McGrath *et al*., 2014). A system that might aid Mtb adaptation is the SOS response. Although fairly well studied, the functions of many genes that form part of the SOS response and RecA-ND response remain unknown (Davis *et al*., 2002; Gamulin *et al*., 2004; Rand *et al*., 2003; Smollett *et al*., 2012). Determining the functions of these genes should be pursued; for example, the putative *srap* is linked to mutagenesis cassettes in many bacteria, yet its molecular function remains unknown. Genes with important contributions might be overlooked because their function is redundant. Studies looking at the DNA damage response as a whole, as have been conducted in *E. coli* (Kumar *et al*., 2016), will aid in characterising contributors, not only at a gene or protein level, but in providing insight into how the proteins fit into the networks which underlie acquisition of mutations. They will also aid in identifying functional redundancy, thus shedding some light on how to target pathways that lead to mutation and subsequent adaptation.

There is a major gap in what is known about the factors and processes driving Mtb evolution. Apart from studying antibiotic resistance, limited systems are available in mycobacteria to study adaptive mutation to other stresses, such as starvation, *in vitro*, highlighting the need for different models to enable a better understanding of the pathways and stresses that may contribute to adaptive mutation in Mtb under different conditions. It is known that the DinB TLS polymerase, forming part of the SOS response significantly contributes to adaptive mutation in *E.coli*, (McKenzie *et al.*, 2001). Similarly, it is known in *Bacillus* and *Pseudomonas* that Pol IV contributes to adaptive mutation (Tegova *et al*., 2004) although DnaE2 seems to play the role of an anti-mutator in *Pseudomonas* (Koorits *et al*., 2007). The potential role of the Mtb mutagenic cassette (DnaE2, ImuA' and ImuB) in adaptive mutation remains to be explored, together with the possibility that DinB2 might contribute to mutagenesis on some level (Ordonez & Shuman, 2014; Ordonez *et al*., 2014).

The general hypothesis investigated here is that mechanisms modulating mutation rate require an in depth understanding in order to elucidate and comprehend pathways leading to drug resistance. The main aims were to investigate a new proposed component of the mutagenesis machinery and to elucidate the role of the known translesion synthesis in adaptive mutation during starvation. In the third chapter we test the hypothesis that SRAP plays a role in the DNA damage response in mycobacteria by means of its peptidase function. The main aims were to evaluate the DNA damage sensitivity of *srap* deletion mutants, visualise and localise Msm SRAP and express and purify Msm SRAP to test its peptidase function to elucidate the role of SRAP in the mycobacterial DNA damage response. In the fourth chapter of this study we test the hypothesis that DnaE2 contributes to adaptive mutation upon nutrient deprivation by exploiting the vitamin B_{12} sensitive phenotype of a methionine synthase (*metH*) deletion mutant. By identifying contributors to adaptive mutation we can identify new targets that can be exploited for combination drug therapy and decrease the level of resistance acquisition.

Chapter 2: Materials and Methods

2.1 General materials and methods

2.1.1 Strains and growth conditions

General bacterial strains used in this study are listed in Table 2.1. All strains used were stored in 30% (v/v) glycerol at -80°C. Msm was used as a surrogate model for Mtb to overcome the difficulties encountered working with a slow growing organism and to limit the utilisation of the biosafety level III resources.

Table 2.1. Strains used in this study.

Strain	Description/Genotype	Reference
Escherichia coli		
DH5α	*supE44*Δ*lacU169* (F80*LacZ*Δ*M15*) *hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	Promega
BL21	*fhuA2 [lon] ompT gal [dcm] ΔhsdS*	New England BioLabs
BL21-SRAP	BL21 with pCOLDI::msms*rap*; Amp^R	This study
Mycobacterium smegmatis **(Msm)**		
mc^2155 (wildtype)	High frequency transformation mutant of Msm ATCC 706	Snapper *et al.*, 1990
ΔY	Triple *dinB1*, *dinB2* and *dinB3* deletion mutant of mc^2155; Hyg^R	D.F. Warner, MMRU
Δ*dnaE2*	*dnaE2* deletion mutant of mc^2155	E. Machowski, MMRU
Δmsms*rap*	*MSMEG_1891* deletion mutant of mc^2155	This study
ΔYΔmsms*rap*	*MSMEG_1891* deletion mutant of ΔY	This study

SSE	pMCAINT::*srapegfp* integrated at *attB* of mc^2155	This study
mc^2155::pSOS-*egfp*	pMCAINT::pSOS-*egfp* integrated at *attB* of mc^2155	M.A. Reiche, MMRU
Mycobacterium tuberculosis **(Mtb)**		
H37Rv (Ma)	Virulent laboratory reference strain of Mtb ATCC27294	Ioerger *et al*., 2010
Δ*dnaE2*	*dnaE2* deletion mutant of H37Rv (Ma); Hyg^R	Boshoff *et al.*, 2003
Δmtb*srap*	*Rv3226c* deletion mutant of H37Rv	This study
Δ*metH*	*metH* deletion mutant of H37Rv, vitamin B_{12} and Cbi sensitive	Gopinath *et al*., 2013
Δ*metHdnaE2*^{AIA} SCO	Homologous recombination leaves two incomplete copies of *dnaE2* in Δ*metH* background, effectively a deletion mutant, Kan^R, Hyg^R.	This study
Δ*metHdnaE2*^{AIA} DCO3	*dnaE2* site specific mutant D439A and D441A in Δ*metH* background	This study
Δ*metHdnaE2*^{AIA} DCO5	*dnaE2* site specific mutant D439A and D441A in Δ*metH* background, biological replicate of Δ*metHdnaE2*^{AIA} DCO3	This study

All ***E. coli*** strains were grown at 37 °C overnight in Luria-Bertani (LB) broth (IncoShake Incubator, Labotec) or on LB agar (IncoTherm Incubator, Labotec). Strains harbouring plasmids larger than 8 kb were cultured at 30°C for 48 h (IncoCool Incubator, Labotec) to minimise plasmid rearrangement.

All **Msm** strains, unless otherwise stated, were grown in Middlebrook 7H9 (Difco™) supplemented with 0.2% (v/v) glycerol, 10% (v/v) Middlebrook Oleic Acid Dextrose Catalase (OADC) enrichment (Difco™) and 0.05% (v/v) Tween80 or grown on Middlebrook 7H10 supplemented with 0.5% (v/v) glycerol and 10% (v/v) Middlebrook OADC. All strains were grown at 37°C with liquid cultures grown in Erlenmeyer flasks at 37°C, shaking (IncoShake Incubator, Labotec).

All **Mtb** strains were cultured similarly to Msm strains. Liquid cultures were grown in tissue culture flasks placed flat at 37°C. All culturing and manipulations of Mtb was done in a Class II flow cabinet at negative pressure (160 – 170 kPa) within Biosafety Level 3 laboratory.

2.1.2 Selection of transformants

Selection of *E. coli* strains was done on media supplemented with antibiotics at concentrations: 50 µg/ml kanamycin (Kan), 200 µg/ml hygromycin (Hyg) and 200 µg/ml ampicillin (Amp). For mycobacterial strains the following concentrations were used: 25 µg/ml Kan, 50 µg/ml Hyg and 100 µg/ml Amp. All antibiotics were stored at 4°C. For positive selection X-gal (5-bromo-4-chloro-3-indolyl-β-D-galacto-pyranoside) was used at 4 µg/ml for both *E. coli* and mycobacteria. The presence of *sacB* was utilised for negative selection by plating cultures on 5% (w/v) sucrose for *E. coli* and 2% (w/v) sucrose for mycobacteria.

Table 2.2 List of plasmids used in this study

Plasmid	Description	Reference
p2NIL	Cloning vector; Kan^R	Parish & Stoker, 2000
pGOAL17	Plasmid with *lacZ* and *sacB* genes in a *Pac*I cassette; Amp^R	Parish & Stoker, 2000
pGOAL19	Plasmid with *hyg*, *lacZ* and *sacB* genes in a *Pac*I cassette, Amp^R	Parish and Stoker, 2000
pCOLD1	Cold shock expression system for high yield,	Takara Bio

	high purity protein production with N-terminus His-tag; Amp^R	
pAINT	*E. coli – Mycobacterium* integrating shuttle vector with *aph*; Kan^R, Amp^R	Boshoff & Mizrahi, 2000
pMCAINT	Derivative of pAINT with additional multiple cloning site, Amp^R	D.F. Warner, MMRU
pMCAINT::*srapegfp*	pMCAINT carrying a 1933bp fragment consisting of *MSMEG_1891* with 400bp of upstream regulatory sequence, a 24bp linker region and *egfp*; Kan^R	This study
pUC57::msms*rap*	pUC57 carrying a Δ*MSMEG_1891* allele; Kan^R	This study
p2NIL::mtb*srap*	p2NIL carrying a Δ*Rv3226c* allele; Kan^R	This study
pUC57::msms*rap*-P17	p2NIL::msms*rap* containing the *Pac*I cassette from pGOAL17; Kan^R	This study
p2NIL::mtb*srap*-P17	p2NIL::mtb*srap* containing the *Pac*I cassette from pGOAL17; Kan^R	This study
pCOLDI::msms*rap*	pCOLDI carrying the *MSMEG_1891* gene with	This study

	400 bp of upstream regulatory region; Amp^R	
p2NIL::mtb*dnaE2*^{AIA}	p2NIL containing *Rv3370c* with site specific mutations modifying 439DID441 to 439AIA441, Kan^R	This study
p2NIL::mtb*dnaE2*^{AIA}-P19	p2NIL::mtb*dnaE2*^{AIA} vector containing the *Pac*I cassette from pGOAL19, Kan^R, Hyg^R	This study

2.1.3 DNA extraction

2.1.3.1 Plasmid DNA extraction and purification

2.1.3.1.1 Small scale extraction

E. coli cultures were grown at 37°C overnight or at 30°C for 48 h for strains harbouring plasmids larger than 8 kb. Cells were harvested by centrifugation at 15 000 × *g* for 1 min (Eppendorf 5415D), supernatant was discarded and the pellet was resuspended in 100 µl lysis solution I (0.5 M glucose, 50 mM Tris-HCl pH 8 and 10 mM EDTA) after which 200 µl Solution II (0.2 M NaOH, 1% SDS) was added and gently mixed. The suspension was incubated at room temperature for 5 min after which 150 µl of neutralising Solution III (3 M potassium acetate, pH 5.5) was added. The supernatant and pellet was separated through centrifugation at 15 000 x *g* for 5 min at room temperature, followed by RNase treatment (1 µl of 10 µg/ml stock, Sigma Aldrich) for 10 min at room temperature. Plasmid DNA was precipitated through the addition of 350 µl isopropanol followed by centrifugation at 15 000 x *g* for 10 min after 10 min incubation at room temperature. The DNA pellet was washed with 70% ice cold ethanol and dried at 35°C in vacuum centrifuge (MiVac DNA concentrator, GeneVac). Plasmid

DNA was resuspended in 20 µl ddH_20 and DNA concentration was estimated using NanoDrop ND-1000 Spectrophotometer (Thermo Scientific).

2.1.3.1.2 Large scale extraction

50 ml *E. coli* cultures were cultured overnight at 37ºC or at 30ºC for 48 h for strains harbouring plasmids larger than 8 kbp. Cells were harvested by centrifugation at 3901 x *g* for 10 min (Beckmann Allegra X-22R). The same protocol was used downstream as for the small scale extraction, except the solution volumes were increased ten times. The plasmid DNA was resuspended in 500 µl ddH_20 and DNA was purified through the addition of equal volume phenol:chloroform (1:1; v/v) followed by vigorous vortexing and centrifugation at 15 000 × *g* for 10 min at room temperature (Eppendorf 5415D Centrifuge). The aqueous phase was collected and plasmid DNA was precipitated by adding 2.5 × volumes ice cold 100% ethanol and 0.1 × 3 M NaOAc. The solution was incubated at -20ºC for at least 45 min and precipitated DNA was collected by centrifugation at 15 000 × *g* for 20 min at room temperature. The pellet was washed with ice cold 70% ethanol, dried using a vacuum centrifuge and resuspended in 50 µl ddH_20.

2.1.3.2 Chromosomal DNA extraction

2.1.3.2.1 Small scale DNA extraction

The colony boil method was used to obtain DNA for PCR screening. A section of a colony was picked off of agar plates, resuspended in 100 µl ddH_2O and heat killed at 95º for 20 min. An equal amount of chloroform was added, mixed well and centrifuged at 15 000 × *g* for 5 min. The aqueous phase was recovered and 2 µl was used for PCR reactions.

2.1.3.2.2 CTAB DNA extraction

Msm: A modified version of the cetyltrimethylammonium bromide (CTAB) method was used for mycobacterial DNA extraction (Larsen, 2000). Briefly, 20 ml cells were grown to mid-log and harvested by centrifugation 3901 × *g* (Beckmann Allegra X-22R) for 10 min. The pellet was resuspended in 1 ml ddH_2O and heat killed at 65ºC for 20 min. The cells were harvested again using centrifugation at 15 000 × *g* and resuspended in 500 µl fresh ddH_2O after which 50 µl lysozyme (10 mg/ml) was added and incubated overnight at 37ºC. The following morning 70 µl 10% SDS and 6 µl proteinase K (10 mg/ml) was added and incubated in the Thermomixer Compact (Eppendorf) at 65 ºC

for 1 h at 400 rpm. Thereafter, 100 µl 10% CTAB (10% CTAB prepared in 0.7 M NaCl) and 100 µl 5 M NaCl were added and incubated at 65°C for 15 min at 400 rpm. An equal volume chloroform:isoamyl alcohol (24:1) was added, gently mixed and centrifuged at 15 000 × *g* for 5 min. The aqueous phase was recovered and equal volume ice cold isopropanol was added and incubated at -20°C for at least 30 min. The DNA was recovered by centrifugation, 15 000 × *g* for 20 min at room temperature, the pellet was washed with 70% ethanol, dried in a vacuum centrifuge and resuspended in 50 µl ddH_2O.

Mtb: The method used for Mtb was very similar to that used for Msm except all steps up to the overnight incubation with lysozyme was done in a Class II flow cabinet at negative pressure (160-170 kPa) in a Biosafety Level 3 laboratory. Minor modifications were made. Briefly, 20 ml Mtb cells were harvested by centrifugation 3901 × *g* for 10 min and the pellet was resuspended in 1 ml ddH2O and split into two Eppendorf tubes with 500 µl each which was heat killed at 80°C for 1 h. Similar to before 50 µl lysozyme (10 mg/ml) was added and incubated overnight at 37°C. The next morning, 70 µl 10% SDS and 50 µl proteinase K (10mg/ml) were added and incubated at 65°C for 1 h in a Thermomixer Compact (Eppendorf) at 400 rpm. Following incubation, 100 µl pre-warmed 10% CTAB and 100 µl 5 M NaCl were added and incubated for another 15 min at 65°C. The tubes were then incubated at -80°C for 15 min, removed and allowed to thaw followed by another 15 min incubation at 65°C in the thermomixer. Equal volume chloroform:isoamylalcohol (24:1) was added to the mixture, mixed gently and centrifuged at 15 000 × *g* for 10 min. The aqueous phase was recovered and equal volume ice cold isopropanol was added and incubated at -20°C for minimum 1 h followed by centrifugation at 15 000 × *g* for 20 min at room temperature. The pellet was washed with 70% ethanol, dried in the vacuum centrifuge and resuspended in 50 µl ddH_2O.

2.1.4 DNA manipulations

Al DNA manipulations and molecular biology techniques were performed according to standard protocols (Sambrook *et al.*, 1989; Sambrook &Russell 2001).

2.1.4.1 Agarose gel electrophoresis

Standard electrophoretic techniques were used (Sambrook *et al.*, 1989). The percentage agarose gel used was altered according to the size of the DNA fragments. High

molecular weight fragments (>500 bp) were separated on a 1% agarose gel, whereas low molecular weight fragments (<500 bp) were separated on a 2-4% gel. Gels were prepared by dissolving agarose powder (Sigma-Aldrich) in 1 × TAE (40 mM Tris-acetic acid, 1 mM Na_2EDTA pH 8), adding 0.5 µg/ml ethidium bromide. DNA samples were loaded with tracking dye (0.025% bromophenol blue in 30% glycerol). Lambda DNA molecular weight markers (III-VI; Roche Applied Science, Germany) were run to estimate DNA fragment sizes. A Mini-Sub® Cell GT was used for electrophoresis with a mini gel horizontal submarine unit (Bio-Rad) at 80-100 V. Gels were visualised under UV-light using the Gel Doc (WealTeach Ketal Imaging System). Fragments used for cloning were visualised using the blue-light Dark Reader DR88M transilluminator (Inqaba Biotec).

2.1.4.2 Gel extraction and quantification

DNA fragments of the desired size were cut from the agarose gel using a scalpel blade and purified using the Nucleospin Extract II Kit (Macherey-Nagel) as per the manufacturer's instructions. Briefly, the desired band was cut from the gel and incubated with DNA-binding buffer at 60°C until completely melted. The solution was then loaded onto a Nucleospin column. For purification of PCR reactions, 1 volume × PCR reaction was mixed with 2 × volume DNA-binding buffer and loaded onto the column. The column was washed twice with wash buffer to and DNA was eluted in 30 µl ddH2O and quantified by NanoDrop ND-1000 Spectrophotometer (Thermo Scientific) and gel electrophoresis, comparing to concentration of molecular weight marker.

2.1.4.3 Restriction digests

Restriction enzymes from New England Biolabs Inc. (Inqaba Biotech, South Africa) or Fermentas (Thermo Scientific) were used. Digests were performed at optimal temperatures of the respective enzymes, mostly 37ºC. Plasmid DNA (1 µg) or mycobacterial chromosomal DNA (5 µg) was digested for at least 3 h and overnight, respectively in the appropriate buffer. For PCR-based screening, PCR products were digested for 2-3 h to identify presence of point mutations with appropriate enzymes in their optimal buffer. Fragments from the digestions were analysed using gel electrophoresis.

2.1.4.4 Dephosphorylation

To minimise vector re-ligation during ligations, linearized plasmid DNA was treated with Antarctic Alkaline Phosphatase (New England Biolabs Inc) to remove the 5' phosphate group. Dephosphorylation was performed for 1 h at 37°C, followed by heat inactivation of the phosphatase at 65°C for 20 min. Dephosphorylation reactions were purified using the Nucleospin Extract II Kit (as per manufacturer's instructions).

2.1.4.5 Ligation

The Fast-link™ ligation kit (Epicentre® Biotechnologies) was used for ligation reactions as per manufacturer's instructions. A minimum of 10 ng dephosphorylated vector was added to the appropriate amount of insert, 1 µl buffer, 1 µl ATP and 1 µl Fast-link ligase. Ligations were performed at room temperature for 2 h followed by heat inactivation of the ligase at 70°C for 15 min.

2.1.4.6 Polymerase Chain Reaction (PCR)

Fragments amplified for subsequent cloning purposes were amplified using Phusion High-Fidelity DNA polymerase (Finnzymes) due to the high fidelity of the enzyme (4.4×10^{-7} error rate). As per manufacturer's instructions, 50 µl reactions contained: 1 × HF reaction buffer, 200 µM of each dNTP, 0.5 µM of each primer, 3% DMSO, 0.02 U/µl DNA polymerase and 50 – 100 ng of DNA. The following cycling conditions were used: initial denaturation at 98°C for 3 min, followed by 30 cycles of denaturation (98°C for 10 s), annealing (appropriate temperature for 30 s), extension (72°C for 30s/kb) and final extension at 72°C for 7 min. The annealing temperature was calculated according to the melting temperatures of the primers, as per manufacturer's instructions. For the purposes of screening, FastStart Taq (Roche) was used as per manufacturer's specifications. Again, 50 µl reactions contained 50 - 100 ng template DNA, 1 × PCR reaction buffer (including $MgCl_2$), 200 µM of each dNTP, 0.5 µM of each primer and 2U/50µl DNA polymerase. The following cycling conditions were used: initial denaturation at 95°C for 4min, followed by 30 cycles of denaturation (95°C for 30 s), annealing (appropriate temperature for 30 s), extension (72°C for 60s/kb) and final extension at 72°C for 7 min. All PCR reactions of were performed in the MyCycler™ thermal cycler (Bio-Rad) with oligonucleotide primers purchased from UCT (Molecular and Cell Biology, University of Cape Town).

Table 2.3 List of primers used in this study

Primer	Sequence (5'-3')	Application	Comments
msms*rap*SFe	ACGGCTATATCCTGCCTCAT	Screening	Forward and reverse primers used for PCR based screening
msms*rap*SRe	TGGTCGTGGCGCAACGGT	Screening	
msms*rap*SFi	CAGCGGTTGGCCACCGAA	Screening	Forward primer and internal reverse primer used for PCR based screening to exclude merodiploids
msms*rap*SRi	GGGCGAGGTGGTCACCTT	Screening	
mtbs*rap*SFe	CCATGTCTTCACCGCTTCAT	Screening	Forward and reverse primers used for PCR based screening
mtbs*rap*SRe	GTCGGACTCACACTGGCTT	Screening	
mtbs*rap*SFi	CGACTCCGGCTTGCCGCA	Screening	Forward primer and internal reverse primer used for PCR based screening to exclude merodiploids
mtbs*rap*SRi	GCGGTGGGCCTTTGGCAT	Screening	
msms*rap*expF	ACGTACTCTCGAGTCTGCAACCATGGGTCTCAT	Expression	Gene amplification for cloning into an expression vector
msms*rap*expR	GCATCGGAAAGCTTCGTCGATCTCGCCGCTCA	Expression	

msms*rap*SBF	GCGCGGTCATCGATGCGA	Southern blot probe	Amplification of downstream homologous region for southern blot probe
msms*rap*SBR	GCGCGGTCATCGATGCGA	Southern blot probe	
SSEF	GCGATCGACGAGACCGCA	Screening and sequencing	Forward and reverse primers used for screening of fusion protein also included for sequencing
SSER	GGTGCTCAGGTAGTGGTTG	Screening and sequencing	
SSEsF1	CCTCGATCATGAGGTCGGT	Sequencing	Forward and reverse primers used for sequencing of SSE fusion protein
SSEsF2	CGACGGTGAGCCGCTGTT	Sequencing	
SSEsF3	GCAAGGGCGAGGAGCTGT	Sequencing	
SSEsF4	CGTCCAGGAGCGCACCAT	Sequencing	
SSEsF5	CACATGGTCCTGCTGGAGT	Sequencing	
SSEsR1	CCAGTCGTCCCGGCTGAT	Sequencing	
SSEsR2	GTGGACTCGTCCTCCGGT	Sequencing	
SSEsR3	GCAGGGCACCTCCACGAT	Sequencing	
SSEsR4	GCAGGGCACCTCCACGAT	Sequencing	
pCOLDF	ACGCCATATCGCCGAAAGG	Screening and sequencing	Forward and reverse primers used for PCR based screening
pCOLDR	TGGCAGGGATCTTAGATTCTG	Screening and sequencing	
pCOLDsF1	GATCCCGCCGTGGGTCAA	Sequencing	Forward and reverse primers used for sequence confirmation
pCOLDsF2	CCCAAGGATCACGCACCGTT	Sequencing	
pCOLDsR1	CCGCGTCGGTGGTGATGA	Sequencing	
pCOLDsR2	CCTTCTCGGCACGCGCAT	Sequencing	

dnaE2scF	GTGCAGCGCCGAGAGCAT	Screening	Forward and reverse primers used for PCR based screening
dnaE2scR	CACCCGCTCGACGACGAA	Screening	
dnaE2sF1	CTCCGGATCGCGGGACTT	Sequencing	Forward and reverse primers to amplify *Rv3370c* for sequencing using all 19 indicated primers
dnaE2sF2	GAACAACTGCGCTGTATCGA	Sequencing	
dnaE2sF3	CGCATTCGGCGTACAGCTT	Sequencing	
dnaE2sF4	CGGGCACTGGCACATCCT	Sequencing	
dnaE2sF5	GGCGAGCGGTGCGCCTT	Sequencing	
dnaE2sF6	CGAGTCGGATCAGCGCGA	Sequencing	
dnaE2sF7	GGTGGAATGGGCGCGCAT	Sequencing	
dnaE2sF8	CGACCTGGTGGTGGAGGT	Sequencing	
dnaE2sF9	ACGAGGTGATCGACCGGAT	Sequencing	
dnaE2sF10	CTTGGGCGCCGTCCGCT	Sequencing	
dnaE2sF11	GCCGGCGCGGTGACTCA	Sequencing	
dnaE2sR1	CCGCTCGCCGCCAGAGT	Sequencing	
dnaE2sR2	GTGCAGAGCACGTTGACCA	Sequencing	
dnaE2sR3	CCGCCGAGACATGCCGAA	Sequencing	
dnaE2sR4	CAGCAGCGCCGCACAGAA	Sequencing	
dnaE2sR5	GACGGGTGCTCGTAGATGA	Sequencing	
dnaE2sR6	GTAGGTGATGACGTTGGCG	Sequencing	
dnaE2sR7	CAACGACCGCAGCCAGCT	Sequencing	
dnaE2sR8	CCGTGCGCACGTCGAGTT	Sequencing	
bacAF1	CGTGACCACAAATGACAT	Sequencing	bacAF1 and bacAR5 were used to amplify 2263bp of genomic DNA containing *Rv1819c* flanked by 150bp upstream and 190bp
bacAF2	TGGCTGATGCTCGGCGTG	Sequencing	
bacAF3	ATCGGGCGGCCCCTGATC	Sequencing	
bacAF4	ATCATCCGATTGCATGGG	Sequencing	
bacAF5	TGACCGGCTGGACGAGGA	Sequencing	
bacAR1	AGATGGAAAACAGGTGGC	Sequencing	
bacAR2	ATCCTCGCACGCCTGAGC	Sequencing	
bacAR3	ACGGTCCAGAACATTGCG	Sequencing	

bacAR4	GAGTCGTGAATGTTGCCG	Sequencing	downstream for sequencing using all ten indicated primers from Gopinath *et al.*, 2013
bacAR5	CGCCACCTTGGTCAGCGT	Sequencing	
Rv1314cF1	CCATGAGCTTGAGCACGCT	Sequencing	*Rv1314c*F1 and *Rv1314c*R1 were used to amplify 855bp of genomic DNA containing *Rv1314c* flanked by 160bp upstream and 110bp downstream for sequencing using all four sequencing primers
Rv1314cF2	CACGGATGTGCTGCGGCA	Sequencing	
Rv1314cR1	CCACAACGACGTGCGCCT	Sequencing	
Rv1314cR2	CACTATCGGAGTCGACAGAT	Sequencing	
$metB_{12}$seqF1	GGCCCAGTAGCCTTCGGT	Sequencing	metB12seqF1 and metB12seqR2 were used to amplify a 500bp of genomic DNA containing the vitamin B_{12} riboswitch upstream of *metE* for sequencing using all four indicated primers by Gopinath *et al.*, 2013
$metB_{12}$seqF2	GGCTGGCAGGTCTTCGGA	Sequencing	
$metB_{12}$seqR1	GGCCAGTAGGAGCACCCA	Sequencing	
$metB_{12}$seqR2	GGCCGATGTCACCGGAGT	Sequencing	

2.1.5 Bacterial transformation

2.1.5.1 *E. coli*

2.1.5.1.1 Preparation of competent cells

A rubidium chloride method was used to prepare competent *E. coli* cells. Briefly, 1 ml of and overnight culture of *E. coli* DH5α was inoculated into 100 ml fresh LB and cultured to an OD_{600} of 0.5. Cells were kept on ice for 15 min and then harvested by centrifugation at 3901 × *g* for 5 min at 4°C. The pellet was resuspended in 20 ml transformation buffer I (30 mM potassium acetate, 100 mM rubidium chloride, 50 mM manganese chloride and 15% (v/v) glycerol, pH 5.8). Again, the mixture was kept on ice for 15 min and the cells were recovered by centrifugation as before. The pellet was resuspended in 2 ml transformation buffer II (10 mM MOPS, 75 mM calcium chloride, 10 mM rubidium chloride and 15% (v/v) glycerol, pH 6.5). Aliquots of 500 µl were flash-frozen and stored at -80°C.

2.1.5.1.2 Transformation plasmid DNA

Competent *E. coli* DH5α cells were thawed on ice and a 100 µl aliquot of cells were added to up to 1 µg of DNA and kept on ice for 20 min. The cells were heat-shocked at 42°C for 90 s, immediately placed back on ice for 1 min and rescued through the addition of 500 µl of 2TY (Tryptone Yeast broth) for 1 h at 37° C, shaking. Cells were then plated on LA plates containing the appropriate antibiotics for screening, and incubated overnight at 37°C, or for 48 h at 30°C, for plasmids larger than 8 kbp.

2.1.5.2 Mycobacteria

Mycobacteria were transformed by electroporation as described by Larsen 2000 and Gordhan & Parish, 2001.

2.1.5.2.1 Msm

Briefly, 100 ml cells were cultured in standard 7H9/OADC to an OD_{600} of 0.8 at 37°C. Cells were recovered by centrifugation at 3901 × *g* for 10 min at 4°C and wash three times in pre-chilled 10% (v/v) glycerol. Following the final wash the cell pellet was resuspended in 2 ml 10% (v/v) glycerol. Aliquots of 400 µl were added to pre-chilled electroporation cuvettes (0.2 cm electrode gap, Bio-Rad) containing 1-5 µg of plasmid DNA and pulsed once in a GenePulser™ (Bio-Rad) set at 2.5 kV, 1000 Ω resistance and capacitance of 25 µF. Cells were rescued after pulsing by adding 800 µl LB to the cuvette and then transferring it to a fresh microcentrifuge tube. Tubes were incubated

at 37°C for 3 h after which the cells were plated on 7H10/OADC with the appropriate antibiotics and incubated at 37°C for 3-5 days.

2.1.5.2.2 Mtb

Electroporation was done similar to that previously described for Msm except that all manipulations were performed at room temperature (Wards & Collins, 1996). Again, 100 ml cells were grown to an OD_{600} of 0.8. Cells were harvested by centrifugation at 3901 × *g* for 10 min and washed twice using 10% (v/v) glycerol. Following the final wash the cell pellet was resuspended in 2 ml 10% (v/v) glycerol. An aliquot of 400 µl cells was added to the electroporation cuvettes (0.2 cm electrode gaps, Bio-Rad) with 1-5 µg plasmid DNA. The cuvettes were pulsed in a GenePulser™ (Bio-Rad) set at 2.5 kV, resistance 1000 Ω and capacitance 25 µF. Cells were rescued by the addition of 800 µl 7H9/OADC and incubated overnight at 37°C. The next morning the cells were plated on 7H10/OADC plates containing the appropriate antibiotic and incubated for 21-28 days at 37°C.

2.1.6 DNA sequencing

All sequencing of constructs, plasmids or PCR products were performed by the Central Analytical Sequencing Facility (CAF) at the University of Stellenbosch. Sequencing reads were visualised using CLC DNA Workbench 7.5 software (http://www.clcbio.com/blog/clc-genomics-workbench-7-5/). WGS of isolates was done by Professors Tom R. Ioerger and James C. Sacchettini (Texas A&M University, Texas, USA).

2.1.7 Southern Blotting

2.1.7.1 Electroblotting

Genomic DNA (5 µg) was digested with 2.5 units of the two appropriate enzymes in the appropriate buffer at 37°C. Digested DNA was separated on a 1% agarose gel using gel electrophoresis at 80 V, viewed and photographed with a fluorescent marker with the Gel Doc (WealTeach Keta Imaging System). The gel was then submerged and shaken gently in depurination buffer (0.25 M HCL) for 15 min at room temperature, followed by treatment with denaturation buffer (1.5 M NaCl, 0.5 M NaOH) for 25 min and neutralisation buffer (1.5 M NaCl, 5 M Tris-HCl, pH 7.5) for 30 min. The DNA was transferred to a nitrocellulose membrane (Hybond™-N^+ membrane, Amersham)

using a vacuum blotter for 120 min (Model 785, Bio-Rad). Once transferred, the DNA was fixed using a UV Stratalinker 1800 (Stratagene) at 1200 mJ/cm^2.

2.1.7.2 Probe preparation

Probes were synthesised using PCR with the indicated primers. The ECL Direct Nucleic Acid Labelling and Detection System (Amersham) was used to label probes. Briefly, 100 ng PCR product in 10 µl was denatured by heating the probe to 95^{0}C for 5 min in a water bath and immediately chilled on ice for another 5 min. Equal volumes of DNA labelling agent and glutaraldehyde reagents (Amersham) were added to the denatured probe, mixed gently and then incubated at 37^{0}C for 15 min.

2.1.7.3 Hybridisation

The hybridisation buffer was used as per the ECL Direct Nucleic Acid Labelling and Detection System (Amersham) protocol. Briefly, the membrane (10 cm^2), 20 ml hybridisation buffer, 5% (w/v) blocking agent and 0.5 M NaCl were stirred at room temperature for 1 h and then incubated at 42°C for another hour. After cross-linking the membrane it was pre-hybridised in roller bottles in the Hybridisation oven/shaker SI30H (Stuart) for 1 h at 42°C followed by the addition of the labelled probe and hybridisation at 42°C overnight.

2.1.7.4 Detection

Following hybridisation, the membrane was washed twice in primary was buffer (6 M urea, 0.4 % (w/v) SDS, 0.5 × SSC) for 20 min at 42°C followed by two washes in secondary wash buffer (0.5 × SSC) for 5 min at room temperature. Detection reagents 1 and 2 (Amersham) were mixed in equal volumes and transferred to the membrane followed by a 1-min incubation at room temperature. The excess detection reagents were drained off, the membrane was wrapped in saran wrap, and exposed to X-ray film (High performance chemiluminescence film, Amersham Hyperfilm™ECL) in a cassette for ~ 10 min at room temperature. Development of the X-ray film was done in a dark room with red-light facilities. Following exposure the X-ray film was submerged in developer solution (GBX Developer, Carestream®Kodak®, Sigma-Aldrich) for 1 min, rinsed in water, and submerged in fixer solution (GBX Fixer, Carestream®Kodak®, Sigma-Aldrich) for 1 min. X-ray films were air-dried and photographed.

2.2 Mutant generation

2.2.1 Construction of Δmsm*srap*

We searched for homologs of SRAP in Msm and Mtb. The *Bordetella bronchiseptica* protein sequence (PDB: 1ZN6) was used to conduct a BLASTp search on the TubercuList website (EPFL). A 400 bp fragment was deleted in the middle of the 780 bp *MSMEG_1891*, leaving a 380 bp fragment of Msm *srap*. To allow for homologous recombination, 706 bp of both upstream and downstream regions were included, flanked by *Bam*HI and *Hin*dIII restriction sites. The resulting construct of 1819 bp was synthesised commercially (Genewiz, Sigma). In an attempt to clone the construct into p2NIL, it was established that the *Hin*dIII and *Bam*HI restriction sites on p2NIL are located too close to each other to allow complete double digestion. This eliminated the possibility of cloning the construct into p2NIL as intended. Instead, the *Pac*I cassette of pGOAL17 was cloned into pUC57::msm*srap* as received from Sigma to give pUC57::msm*srap*-P17. Due to the absence of a *Pac*I cut site in pUC57::msm*srap* the *Pac*I cassette was excised out of pGOAL17 and the 3' overhangs were removed using DNA Pol I as per manufacturer's instructions. Briefly, one unit of DNA polymerase I was added per μg of DNA in the presence of 1 × NEB buffer and 33 μM of each dNTP. The reaction was incubated at 25°C for 15 min and then inactivated by adding EDTA to a final concentration of 10 mM and heating it to 75°C for 20 min. The blunt-end cutter *Dra*I was used to cut pUC57::msm*srap* and clone the blunt-ended *Pac*I cassette. This method for cloning was not optimal and thus the Mtb construct was designed differently. Clones were screened as Kan^R, blue and sucrose sensitive colonies and confirmed using PCR and restriction mapping. pUC57::msm*srap*-P17 was electroporated (Larsen *et al.*, 2007) into mc^2155 and ΔY, a previously constructed triple *din* deletion mutant, to generate *srap* deletion mutants. The presence of the desired deletion was confirmed in all constructs and after homologous recombination in single cross overs (SCOs) and double cross overs (DCOs) using restriction digests, PCR (msm*srap*SF + msm*srap*SR + msm*srap*SFi + msm*srap*SRi) and southern blot.

2.2.2 Construction of Δmtb*srap*

A 399 bp fragment was deleted in the middle of the 759 bp *Rv3226c*, leaving a 360 bp fragment of Mtb *srap*. To allow for homologous recombination, 520 bp of both upstream and downstream regions was included, flanked by two *Hin*dIII restriction

sites. As for Msm *srap* the resulting construct of 1412 bp was synthesised (Sigma). The deletion construct was cloned into p2NIL using *Hin*dIII to give p2NIL::mtb*srap*. The *Pac*I cassette of pGOAL17 was subsequently cloned into p2NIL::mtb*srap* to give p2NIL::mtb*srap*-P17. Clones were screened as Kan^R, blue and sucrose sensitive after which it was confirmed using PCR and restriction mapping. This construct was electroporated into Mtb H37Rv. The desired deletion was confirmed in the constructs using restriction mapping and PCR (mtb*srap*SFe, mtb*srap*SRe, mtb*srap*SFi and mtb*srap*SRi). The final DCOs were PCR screened and WGS was used to confirm the deletion (Thomas R. Ioerger and James C. Sacchettini, Texas A&M University, Texas, USA).

2.2.3 Construction of suicide vector containing *dnaE2*^{AIA}

For the second part of the study a construct containing *dnaE2*^{AIA} was synthesised (Sigma) harbouring 1.4 kbp of the *Rv3370c* gene where the polymerase active site residues were altered as follows: D439A and D441A. The *dnaE2*^{AIA} construct was sub-cloned into p2NIL using *Hin*dIII restriction cut sites. Transformants were selected on Kan containing LB agar plates. Resistant colonies were confirmed using restriction mapping to be p2NIL::mtb*dnaE2*^{AIA}. The *Pac*I-cassette of pGOAL19 was cloned into p2NIL::mtb*dnaE2*^{AIA} generating p2NIL::mtb*dnaE2*^{AIA}-P19. Positive transformants were screened as blue, Kan resistant, Hyg resistant and sucrose sensitive colonies followed by restriction mapping. To simplify screening *Nru*I and *Pvu*I cut sites were introduced by mutating ^{439}DID^{441}to ^{439}AIA^{441} (Figure 4.1A).

Mtb was electroporated with p2NIL::mtb*dnaE2*^{AIA}-P19. SCOs were confirmed using restriction digests (*Nru*I and *Pvu*I) of PCR products (dnaE2scF and dnaE2scR, Table 2.3). If the construct harboured the desired mutations *Nru*I rendered fragments of 449 bp and 661 bp as opposed to uncut wildtype. For the mutant fragment *Pvu*I gave 135 bp, 312 bp and 663 bp as opposed to 974 bp and 135 bp for wildtype. Positive transformants were confirmed by Sanger sequencing of PCR products (Central Analytical Sequencing Facility, Stellenbosch University). DCOs were generated by culturing a SCO in the absence of selective pressures (Kan and Hyg) for 7 days and plated on 2% sucrose (w/v) 7H10/OADC plates and incubated for 21 days. Sucrose resistant potential DCOs were selected and screened in the same way as the SCOs. To

confirm that the DnaE2 polymerase activity was indeed inactivated, UV mutagenesis and MMC survival assays were performed (Section 2.8).

2.3 DNA damage sensitivity and UV mutagenesis assays

A range of different DNA damaging agents/treatments was implemented to test the sensitivity of the deletion mutants in the different strains.

2.3.1 Msm

2.3.1.1 UV mutagenesis assays:

For UV mutagenesis assays, 40 ml cultures were grown to an OD_{600} of 0.6-0.8. At this point colony forming units (CFU) counts were obtained by plating a log-fold dilution series onto 7H10/OADC and incubating the plates at 37°C for 3 days. Cells were then harvested using centrifugation at 3901 × *g* (Beckmann Allegra X-22R) for 10 min. Cells were resuspended in 5 ml fresh 7H9/OADC and irradiated with 20 mJ/cm^2 (UV Stratalinker 1800, Stratagene) in open petri dishes. The cells were rescued in a total volume of 40 ml 7H9/OADC for 3 h at 37°C after which 1 ml was plated on 7H10/OADC containing 200 µg/ml Rif to calculate Rif^R frequencies. Plates were incubated at 37°C for 5-6 days. Similar to before UV exposure, a log-fold dilutions were plated for CFU determination to study UV survival, plates were also incubated for 3 days at 37°C.

2.3.1.2 Damage survival assays:

Cultures were grown to an OD_{600} of 0.6-0.8 after which 5 µl of log-fold dilutions were spotted onto 7H10/OADC and 7H10/OADC supplemented with MMC at 60 nM and 120 nM concentrations. Msm strains were also spotted on 7H10/GN as an untreated control and 7H10/GN containing H_2O_2 (0.05 mM and 0.1 mM). Plates were incubated for 3-4 days at 37°C.

2.3.2 Mtb

2.3.2.1 UV mutagenesis assays:

UV mutagenesis assays were conducted in a similar manner to that described for Msm. However, cells were left to recover for 24 h at 37°C following UV treatment after which they were plated on 7H10 plates containing 2 ug/ml Rif. Standard 7H10 plates for CFU

determination were incubated for 21 days at 37°C and Rif containing plates were incubated 28-35 days at 37°C for Rif^R determination.

2.3.2.2 Damage survival assays:

These were executed in the same manner to Msm, with 5 µl of 10-fold serially diluted cells spotted on 7H10/OADC and 7H10/OADC containing: MMC at 60 nM and 120 nM, ofloxacin (Ofx) at 0.83 µM and 1.66 µM, plumbagin (Pmb) at 13.3 µM and 26.6 µM, nitrofurazone (Nfz) at 80.8 µM and 161.6 µM. Plates were incubated at 37°C for 10-12 days.

2.4 MIC_{90} analysis

The broth micro-dilution method was used to determine the minimum inhibitory concentration (MIC) of wildtype and Δmsm*srap* to H_2O_2 (Collins & Franzblau, 1997). This method allows for the testing of strain sensitivity to a range of drug concentrations in a single 96-well microtiter plate. Briefly, 50 µl media was added to all rows but row one and 100 µl 7 mM H_2O_2 was added to row one and serially diluted, two-fold, in a U-bottom 96-well microtiter plate (PGRE650180, Lasec, SA). Media controls were included to allow maximum growth. Strains were grown in LB broth to an OD_{600} of 0.5 and diluted 1:1000 LB, 50 µl of the cell dilution was added to each well to give a final volume of 100 µl. The inoculated microtiter plate was incubated at 37^0C for 3 days after which 20 µl alamarBlue® was added at the end of day three to each well and incubated at 37^oC for another 6 h. Fluorescence readings of wells were obtained using the spectrophotometer (FLUOstar Optima, BMG Labtech). The lowest concentration of the drug that prevented a colour change from blue to pink was reported as the MIC that inhibited at least 90% bacterial growth (MIC_{90}).

2.5 Msm *srap* induction upon DNA damage

Cultures were grown to an OD_{600} of 0.4 and split, half was treated with 0.6 µM MMC for 1 h and the other half was left untreated. RNA was extracted using FastRNA ProBlue Kit (MP Biomedicals) as per manufacturer's instructions. Cells were harvested by centrifugation at 15 000 × *g* for 15 min at 4^0C. The pellet was resuspended in 1 ml RNApro™ solution and transferred to a blue capped tube with Lysing MatrixB. These were then processed for 40 sec at 6.0 setting of the FastPrep® Instrument and centrifuged at 12 000 × *g* for 5 min at 4^0C. The supernatant was transferred to a new

microcentrifuge tube and incubated at room temperature for 5 min after which 300 µl chloroform was added and vortexed for 10 sec. Again, the sample was incubated at room temperature for 5 min and centrifuged for 12 000 × *g* for 5 min at 4^0C. The supernatant was transferred to a new tube and 500 µl cold 100% ethanol was added, tubes were inverted 5 times and incubated at -20^0C for a minimum of 30 min. Samples were centrifuged at 12 000 × *g* for 15 min at 4^0C and the supernatant was discarded. The pellet was washed with 300 µl 70% ethanol after which the ethanol was removed and the pellets were air dried for 5 min. Pellets were resuspended in 100 µl DEPC-treated H_2O and incubated at room temperature for 5 min after which the RNA concentration was measured using a NanoDrop ND-1000 Spectrophotometer (Thermo Scientific). RNA was stored at -80^0C.

DNase treatments were carried out using Turbo DNase (Ambion). In each round, 2 µg RNA was added to 2 µl DNase, 2 µl 10 × buffer and DEPC-treated ddH_20 to a total volume of 10 µl. The reaction was incubated at 37^0C for 30 min after which the RNA was purified. The reaction volume was made up to a total volume of 100 µl using DEPC-treated H_2O after which 100 µl acid phenol: chloroform: isoamylalcohol (25:24:1) was added. The samples were vortexed briefly and centrifuged for 30 sec at 12 000 × *g*. The aqueous phase was collected (~75 µl), 0.1 volumes 3 M CH_3COONa and 2.5 volumes cold 100% ethanol were used to precipitate nucleic acids at -80^0C for 45 min. Samples were centrifuged at 12 000 × *g* for 5 min, supernatant was discarded and pellets were washed using 800 µl 70% ethanol (prepared with DEPC-treated H_2O) and dried in a vacuum centrifuge (MiVac DNA concentrator, GeneVac) for 3 min at 30^oC. The purification steps were repeated and the pellet was ultimately resuspended in 10 µl DEPC-treated H_2O and incubated at 4^0C after which a PCR was done to confirm the absence of any contaminating DNA.

To generate cDNA from the purified, DNA-free RNA, the High Capacity RNA to cDNA kit (Thermo Fisher Scientific) was used. A total of 100 ng RNA was added to a reaction mixture consisting of 10 µl, 2 × RT buffer and 1 µl RT enzyme mix. The reaction volume was made up to a total volume of 20 µl with ddH_2O. Reactions without reverse transcriptase (RT^-) were included as no template controls in addition to water

controls. Reverse transcription was performed in a T100 Thermal Cycler (Bio-Rad) at 37^{0}C for 1 h followed by 5 min at 95^{0}C. The cDNA was stored at 4^{0}C.

Droplet Digital™ PCR (ddPCR) was used to study gene expression in MMC treated and untreated bacilli. Primers and probes were designed using Primer3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) and are listed in Table 2.4. Prior to ddPCR, primer amplification was optimised using a gradient PCR with all reactions performed optimally at 62^{0}C. Reaction mixtures (20 µl) contained 2 µl cDNA, 10 µl 2 × ddPCR Supermix for probes (Bio-Rad), 500 nM of the forward and reverse primers and 250 nM of TaqMan probe. This reaction mixture was emulsified with droplet generation oil using a QX200 Droplet Generator (Bio-Rad). Amplification was performed using a protocol of 95^{0}C for 10 min followed by 40 cycles of 94^{0}C for 30 sec, 60^{0}C for 60 sec and ended with 98^{0}C for 10 min in a T100 Thermal Cycler (Bio-Rad). Droplets were analysed using the QX200 Droplet Reader (Bio-Rad). All samples, including the water controls and no template controls were done in duplicate. The QuantaSoft Software version 1.4.0.99 (Bio-Rad) provided absolute quantification of the respective genes which was then normalised to *sigA*.

Table 2.4 Primers and probes used for ddPCR reactions to study SRAP expression levels

Gene	Primer	Primer sequence (5'- 3')	Probe (5' – 3')*
srap	*srap*F	CAGGCGCCAATTACAACGT	CCGACAACGACCATC
	*srap*R	GTTCGGTGTGGCGTTTGAC	(VIC)
recA	*reca*F	TGACCGGCGCGTTGA	AACTCGGGCACCACC
	*reca*R	CGGAGCTGGTTGATGAAGATC	(VIC)
sigA	*siga*F	TCGGTTCGCGCCTACCT	ATCGGCAAGGTGGC
	*siga*R	ACCTCTTCTTCGGCGTTGAG	(FAM)

*TaqMan minor groove binder (MGB) probes were either labelled with 4,7,2'-trichloro-7'-phenyl-6- carboxyfluorescein (VIC) or with 6- carboxyfluorescein (FAM).

2.6 Fusion protein construction and visualisation

The construct was designed by taking Msm *srap* with 400 bp upstream of the gene including the native promoter and cloning it in frame with *egfp* using a 24 bp linker at the C-terminus of SRAP. This fusion was cloned into pMCAINT using *Hin*dIII restriction sites to give pMCAINT::msm*srap-egfp*. The construct was confirmed using restriction mapping, PCR and Sanger sequencing (Central Analytical Sequencing Facility, Stellenbosch University). The resulting plasmid was subsequently used to transform mc^2155 by electroporation to allow integration at the *attB* site. A previously constructed reporter strain, mc^2155::pSOS-*egfp* (Reiche, unpublished), was used as a positive control for MMC DNA damage detection. This strain has *egfp* under the control of a SOS inducible promoter, thus expressing eGFP under DNA damaging conditions. Wildtype, the SRAP-eGFP (SSE) fusion protein and reporter strain were grown to an OD_{600} of 0.4 in Middlebrook 7H9 supplemented with OADC (Middlebrook) and treated with 0.6 mM MMC for 3 h at 37°C, shaking. An aliquot of both treated and untreated samples were taken and viewed using fluorescence microscopy (Zeiss, Axio Scope A1). In addition to MMC, wildtype and SSE were also treated with 7 mM H_2O_2 for 1 h and 3 h at 37^0C, shaking in Middlebrook 7H9 supplemented with 0.11 M glucose, 0.75 M bovine serum albumin (BSA), 0.14 M NaCl and 3% (v/v) Na-oleate solution. Again, aliquots of both treated and untreated samples were viewed using fluorescence microscopy. Images were analysed and fluorescence was measured using Fiji software (Schindelin *et al*., 2012). Wilcoxon signed rank tests were used to detect significant increase fluorescence levels.

2.7 Protein modelling and expression

2.7.1 Bioinformatic modelling

Dr. Natasha Wood and Dr. Česlovas Venclovas generated models of the Msm and Mtb SRAP proteins, respectively. These models were analysed and compared to available crystal structures of homologs using Chimera 1.10.2 (UCFS). Crystal structures were obtained from PDB (http://www.rcsb.org/pdb/home/home.do).

2.7.2 Cloning and expression of Msm *srap*

The predicted Msm *srap* gene and upstream regulatory region was PCR amplified using msm*srap*expF and msm*srap*expR (Table 2.3), which have engineered *Xho*I and *Hin*dIII restriction sites, respectively. The resulting fragment (840 bp, ~ 30 kDa protein) was

cloned into pCOLDI using *Xho*I and *Hin*dIII to add an N terminal His-tag to Msm SRAP. pCOLDI::msm*srap* was used to transform BL21 *E. coli*, transformants were selected for using Amp^R and confirmed using PCR (pCOLDF + pCOLDR) and sequencing (pCOLDF + pCOLDR + pCOLDsF1 + pCOLDsF2 + pCOLDsR1 + pCOLDsR2). Flasks with 200 ml BL21-SRAP were cultured in LB with Amp at 37°C, shaking, until an OD_{600} of 0.4-0.5. The culture was cooled to 15°C for 30 min. Gene expression was induced by adding IPTG to a final concentration of 0.75 mM and the cultures were incubated at 15°C for 24 h. The cells were harvested by centrifugation, 3901 × *g* for 10 min. The pellet was resuspended in 7.5 ml lysis buffer (50 mM Tris-HCl, 300 mM NaCl and 20 mM imidazole, pH 8). The cells were sonicated (Misonix Sonicator 3000) at an amplitude of 20, 8 × 15 sec with 15 sec breaks in between pulses and finally centrifuged at 27 000 × *g* for 30 min at 4°C. The supernatant was recovered and the pellet was resuspended in the same volume as supernatant recovered. Aliquots of the uninduced cell lysate, induced whole cell lysate and supernatant were run on a 12.5% reducing SDS-PAGE stained with AquaStain (Bulldog Bio) to confirm soluble expression.

2.7.3 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Protein size and purity was confirmed using 12.5% SDS-Poly Acrylamide Gel Electrophoresis and stained by submerging the gel in AquaStain (Bulldog Bio) for 10 min, with shaking. Fermentas Pageruler™ prestained protein ladder and prestained protein ladder plus (5 µl) were used as molecular markers. Samples were prepared by mixing 5 µl of 4 × sample application buffer (SAB) (8 % SDS, 40 % glycerol, 4 % β-mercaptoethanol, 50 mM EDTA, 0.08 % bromophenol blue and 200 mM Tris-HCl, pH 6.8) with 15 µl sample and boiled for 5 min at 95°C to denature proteins. Depending on the protein concentration, 5 – 20 µl was loaded on the gel and ran for 90 min at 20 mA in running buffer (0.025 M Tris-HCl, 0.192 M glycine, 0.1 % SDS, pH 8.3).

2.7.4 Protein concentration determination

The Bradford protein assay was used to estimate protein concentration. Bovine serum albumin was used as standard in increments of 0.1 mg/ml from 0.1 mg/ml to 0.5 mg/ml. The dye was prepared by adding 1 part dye to 4 parts water. The sample (10 µl) was added to 200 µl dye dilution, incubated at room temperature before the absorbance was measured using the spectrophotometer (FLUOstar Optima, BMG Labtech). Protein

samples of unknown concentration were processed similarly to obtain absorbance values, the standard curve was then used to calculate protein concentration.

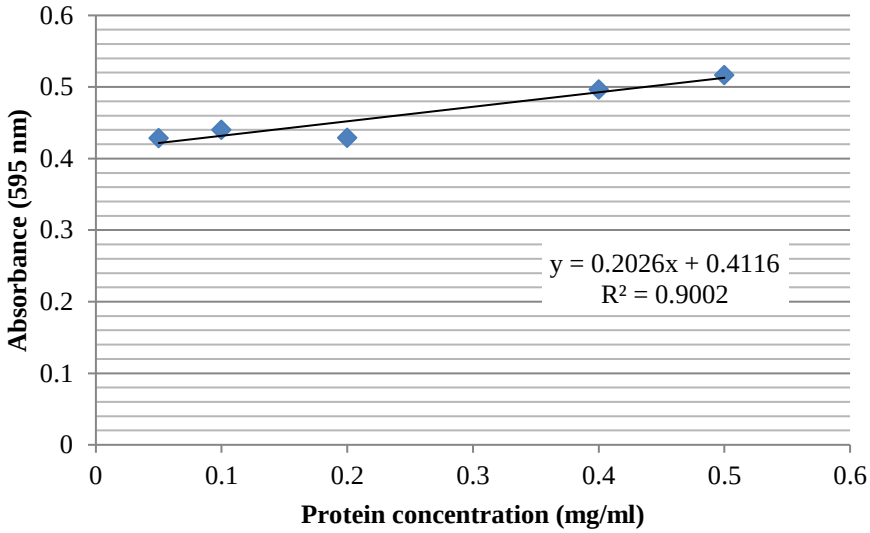


Figure 2.1 Bradford standard curve. Standard curve and equation used for calculating protein concentration.

2.7.5 Protein purification

The supernatants were loaded on a nickel-charged column on the AKTA system, which was pre-equilibrated with the lysis buffer used before. Ten column volume washes were done to get to baseline following loading and flow through, after which elution was facilitated over 20 column volumes on an imidazole gradient from 20 mM to 180 mM at 2.5 ml fraction volume. Lastly, the gradient was stepped up to 100% for five column volumes. Fractions were run on a reducing SDS-PAGE gel as before. To confirm that the band at the expected size represented SRAP it was sent for peptide mass fingerprinting (PMF) which confirmed the presence of the desired protein (CSIR, Pretoria, South Africa). Fractions with the highest concentrations of pure protein were selected, pooled and concentrated down to 5 ml using an Amicon® stirred cells device (Merck Millipore). The concentrate was loaded onto an in house packed column with GE Healthcare Sephacryl™ S200 high resolution beads for size exclusion chromatography (SEC) to study the native conformation. The column was used with a Gilson 305 pump (flow rate of 0.5 ml/min), Gilson 806 manometric module, Gilson uv/vis-151 absorbance detector (set at 280 nm) and a Gilson FC 203B fraction collector. Data transfer was accomplished using a Gilson 506C system interface and data

interpretation was done using Gilson Unipoint software, version 3.3. The column was pre-equilibrated with two column volumes of the following buffer: 50 mM Tris-HCL, 300 mM NaCl, pH 8. Sizes were estimated based on a previously constructed calibration curve that was generated using dilutions of proteins standards with known sizes recorded in Table 2.5 (Biorad). Fractions were checked by SDS-PAGE as before. To test whether the size distribution was in equilibrium, fractions from the large peak (280 kDa) were left in the fridge overnight and the SEC run was repeated the next morning. A dilution series of fractions representing the largest peak (280 kDa) was prepared and used to make transmission electron microscopy (TEM) carbon grids for TEM (Electron Microscopy Unit, UCT).

Table 2.5. Protein standards used to construct calibration curve for protein size determination

Protein	**MW**
Tobacco Mosaic Virus	50000 kDa
Thyroglobulin	670 kDa
Gammaglobulin	158 kDa
Ovalbumin	44 kDa
Myoglobulin	17.5 kDa
Vitamin B-12	1.35 kDa
Acetone	0.58kDa

2.7.6 Crystal trials

Fresh, purified protein was concentrated down using Amicon® Ultra-15 centrifugal filter units (Merck Millipore) to ~7.8 mg/ml. Initial crystal trials were set up using the PEG-Ion Screen® (Hampton Research). TTP Labtech's mosquito® was used to generate three hanging droplets of 200 µl for each precipitant consisting of (i) 150 nl protein + 50 nl precipitant (ii) 100 nl protein + 100 nl precipitant (iii) 50 nl protein + 150 nl precipitant. Plates were incubated at room temperature for 21 days and screened on a regular basis to note a change in precipitation. Precipitants with precipitate in the

drops with 150 nl and 100 nl protein but lacked precipitate in the drop with 50 nl protein were recorded. Seven precipitants were identified using these criteria, of which only KCl (pH 7), NH_4Cl (pH 6.3) and CH_3CO_2K (pH 8.1) were used in downstream screens. All three were used at 0.2 M with 20% w/v polyethylene glycol 3350. Plates were set up with 0.15 M, 0.165 M, 0.175 M, 0.2 M, 0.225 M and 0.25 M salt in combination with 18%, 20%, 22% and 24% polyethylene glycol 3350 (PEG) and protein at ~7.8 mg/µl. Three 30 µl hanging drops were prepared for each salt and PEG 3350 combination (i) 15 µl protein + 15 µl precipitant (ii) 12 µl protein and 18 µl precipitant (iii) 7.5 µl protein and 22.5 µl precipitant. Plates were incubated at room temperature and carefully monitored for change in precipitate over 56 days.

2.7.7 Functional analyses

2.7.7.1 Zymography

Fresh, purified protein was quantified by Bradford protein assay (Biorad). Gelatin (0.5 mg/ml) and casein (0.2 mg/ml and 0.5 mg/ml) gels were used as substrates for zymography. Human matrix metalloproteinase 1 (MMP1) (5 ng) was run as a positive control with 10 ng, 20 ng and 100 ng of Msm SRAP on a non-reducing SDS-PAGE gel for 50 min at 40 mA. The gel was then washed twice using 2.5% TritonX for 15 min after which it was incubated at 37°C in incubation buffer (50 mM Tris-HCL, 5 mM $CaCl_2$, 0.2 M NaCl, 0.5% Brij-35, pH7.6) for either 20 h or 40 h. The gel was stained (0.1% Coomassie Brilliant Blue, 40% methanol, 10% acetic acid) for 1 h and subsequently destained (40% methanol, 10% acetic acid) twice for 30 min, shaking. A non-reducing PAGE gel was run to confirm the expected sizes of native SRAP on the gel.

2.7.7.2 N-terminus cleavage

Fresh, purified protein was divided in three and incubated at room temperature, 37°C or 42°C, respectively. A sample of protein at each temperature was collected at 0 minutes (T0), after 24 h, 48 h and 72 h. These samples were run on a reducing SDS-PAGE gel and stained with the InVision™ His-Tag In-Gel Staining kit (Invitrogen) followed by normal AquaStaining (Bulldog-Bio) to confirm the presence of equal amounts of protein. The InVision staining was used as per manufacturer's protocol, after electrophoresis the gel was fixed using Fixing Solution (2% v/v acetic acid, 90% v/v methanol) for 1 h. The gel was washed twice for 10 min with ddH20, followed by

staining with the InVision His-tag In-gel Stain (Invitrogen) for 1 h at room temperature, followed by two 10 min washes with 20 mM phosphate buffer (pH 7.8). The gel was then visualised on the UV transilluminator with the maximum excitation wavelength being 560 nm and maximum emission wavelength 590 nm.

2.8 VitaminB$_{12}$ sensitivity analyses of *dnaE2*^{AIA}Δ*metH*

2.8.1 Determining vitamin B$_{12}$ sensitivity

To determine the strains' sensitivity to B_{12} and Cbi the cultures were grown to an OD_{600} of 0.6-0.8 after which log-fold dilutions were plated on 7H10/OADC plates and 7H10/OADC containing 7.38 μM B_{12} and 10 μM Cbi, respectively. Plates were incubated for 21 days at 37°C and CFUs were calculated.

2.8.2 Confirmation of hereditary nature and identifying resistance causing mutations

A random subset of B_{12} resistant colonies was picked, cultured in 7H9/OADC (in the absence of B_{12} and Cbi) for 7 days and log-fold dilutions were plated on 7H10/OADC supplemented with B_{12} and Cbi. Three regions that were previously implicated in the appearance of the suppressor phenotype, were sequenced (Table 2.3) to identify mutations that could possibly account for the suppressor phenotype. These regions were *Rv1819c* (*bacA*), *Rv1314c* and the riboswitch upstream of *metE* (*Rv1133c*). Furthermore, genomic DNA of the parental Δ*metH* strain as well as the two DCOs were sent for whole genome sequencing (Thomas R. Ioerger and James C Sacchettini, Texas A&M University, Texas, USA.

Chapter 3: The role of SRAP in the mycobacterial DNA damage response

3.1 Introduction

The DNA damage response pathways of mycobacteria are widely studied and some of the major elements in these pathways have been characterised in great detail for example RecA and LexA, but the functions of numerous damage inducible genes remain to be elucidated. A bioinformatics study using comparative genomics and gene neighbourhood analyses identified two new domain superfamilies that strongly co-occur with major SOS response operons or genes, including *imuAB-dnaE2*, *lexA*, *dinB* and *ku*/*ligD* (Aravind *et al.*, 2013). The first superfamily comprises a protein with a globular domain of unknown function (DUF159), typified by *Escherichia coli* YedK. Both Msm and Mtb possess homologs of the domain encoding gene, *MSMEG_1891* and *Rv3226c*. The second identified domain showed significant homology to the C-terminus of Y-family polymerases in the ImuB-family, and was therefore named ImuB-C. The prototype for this protein is *Geobacter sulfurreducens* GSU0042, no homologs of which are found in Msm or Mtb.

Further investigation into the DUF159 protein revealed the conservation of cysteine (second from the N-terminus), glutamate (middle of the chain) and histidine (C-terminus) residues. These three conserved residues form a binding pocket that resembles the catalytic triad of structurally unrelated cysteine proteases. Autocatalytic cleavage at the conserved cysteine is possibly required for activation of the peptidase. This prediction was based on the absence of sequences N-terminal to the cysteine in the crystal structures of homologs, and further supported by the loss of an N-terminal tag during purification of YedK (Aravind *et al.*, 2013). Based on the structures, cleavage of the N-terminus causes hydrogen bond formation between the cysteine and a highly conserved glutamine in the centre of the protein, inactivating the peptidase function. This protein was subsequently referred to as the SOS response-associated peptidase (SRAP). Furthermore, the presence of helix-hairpin-helix motifs, commonly associated with DNA binding, together with two conserved arginines, could facilitate contact with DNA (Aravind *et al.*, 2013). Together with the predicted manner of regulation, it suggests a role as a functional switch, possibly through DNA binding following DNA damage.

Autocatalytic cleavage plays a critical role in the DNA damage response. As noted above, repression by LexA is relieved by autocatalytic cleavage, which itself is triggered by the co-protease activity of activated RecA. This cleavage event separates the N-terminal DNA binding domain from the C-terminal dimerization domain (Neher *et al.*, 2003). Similarly, *E. coli* UmuD, a specialist TLS repair polymerase, DNA Pol V, is also regulated by RecA, mediating the cleavage of 3 kDa from the N-terminus (Nohmi *et al.*, 1990; Shinagawa *et al.*, 1988) to generate an active UmuD'. The predicted SRAP cleavage is therefore reminiscent of LexA and UmuD cleavage, although structurally and mechanistically unrelated (Aravind *et al*., 2013).

Cysteine proteases are abundant in all forms of life, performing a variety of functions (Verma *et al.*, 2016). These proteases all hydrolyse amide bonds by a shared catalytic mechanism which requires a catalytic triad with a nucleophilic cysteine. The first step involves the deprotonation of the cysteine by a neighbouring amino acid with a basic side chain. This is followed by a nucleophilic attack on the carbonyl group of the substrate. Unwanted hydrolysis by the newly synthesised proteases is controlled by the production of long inactive zymogens which are activated by cleavage of the pro-peptide, commonly by means of autocatalytic cleavage (Chapman *et al.*, 1997).

Members of the SRAP family are widely found throughout bacteria and eukaryotes (Aravind *et al*., 2013). Notably, they are commonly located in the same operon as the organism's specific induced mutagenesis cassette: *umuD-umuC* in *E.coli*, and *imuA'-imuB-dnaE2* in *Agrobacterium* and *Burkholderia* (Aravind *et al.*, 2013). These cassettes are commonly associated with drug resistance acquisition (Bos *et al.*, 2015; Boshoff *et al*., 2003; Zhang *et al*., 2011). However, the function of the bacterial SRAP homologs - for example, YedK in *E. coli* (EcoGene) and YoqW in *Bacillus subtilis* (SubtiList) - remains unknown in these model organisms. Interestingly, the mammalian homolog (C3Orf37) was found to bind oxidised derivatives of 5-methyl cytosine (Spruijt *et al*., 2013). In that system, however, it is thought unlikely that C3Orf37 contributes to adaptive mutagenesis but rather plays a pivotal role in the recognition of oxidised methylated cytosines and recruitment of repair machinery. The putative autopeptidase activity of SRAP has been proposed to act as a functional switch between the damaged DNA and recruiting repair enzymes (Aravind *et al*., 2013).

The potential hierarchies of gene and protein regulation and the signalling networks mediating the DNA damage response remain poorly understood. The inferred functioning of the mammalian homolog suggests a potential regulatory role for SRAP, possibly influencing other factors within the DNA damage response network. Since *srap* occurs within the same operon as the mutagenesis cassette *imuA'-imuB-dnaE2* in many bacteria, it could be an important component in the DNA damage response and even contribute to induced mutagenesis. To date, however, no role has been assigned to SRAP in the DNA damage response in bacteria. A key focus in our laboratory is to identify alternative factors that may interact with the mutasome (ImuA', ImuB and DnaE2); therefore, in this study, we tested the hypothesis that SRAP plays a role in the DNA damage response in Msm and Mtb. To this end, the sensitivities of Msm and Mtb deletion mutants to a range of DNA damaging agents were investigated. Owing to the profound sensitivity phenotype observed when any of *imuA', imuB* or *dnaE2* deletion mutants are exposed to the genotoxin, MMC, we focused specifically on SRAP in response to MMC treatment. We also addressed a possible role for SRAP during oxidative stress as suggested by the mammalian counterpart. As detailed below, the absence of a clear phenotype of any of the *srap* knock-out mutants, however, means that the potential contribution of SRAP to the mycobacterial DNA damage response remains cryptic.

3.2 Aims and objectives

The main aim of this study was to determine whether SRAP contributes to the DNA damage response in mycobacteria. The specific objectives were to:

I. evaluate if SRAP contributes to induced mutagenesis and DNA damage sensitivity in Msm and Mtb;
II. determine the subcellular location of Msm SRAP;
III. express and purify recombinant Msm SRAP for functional studies
IV. test for predicted peptidase activity and N-terminal cleavage of Msm SRAP

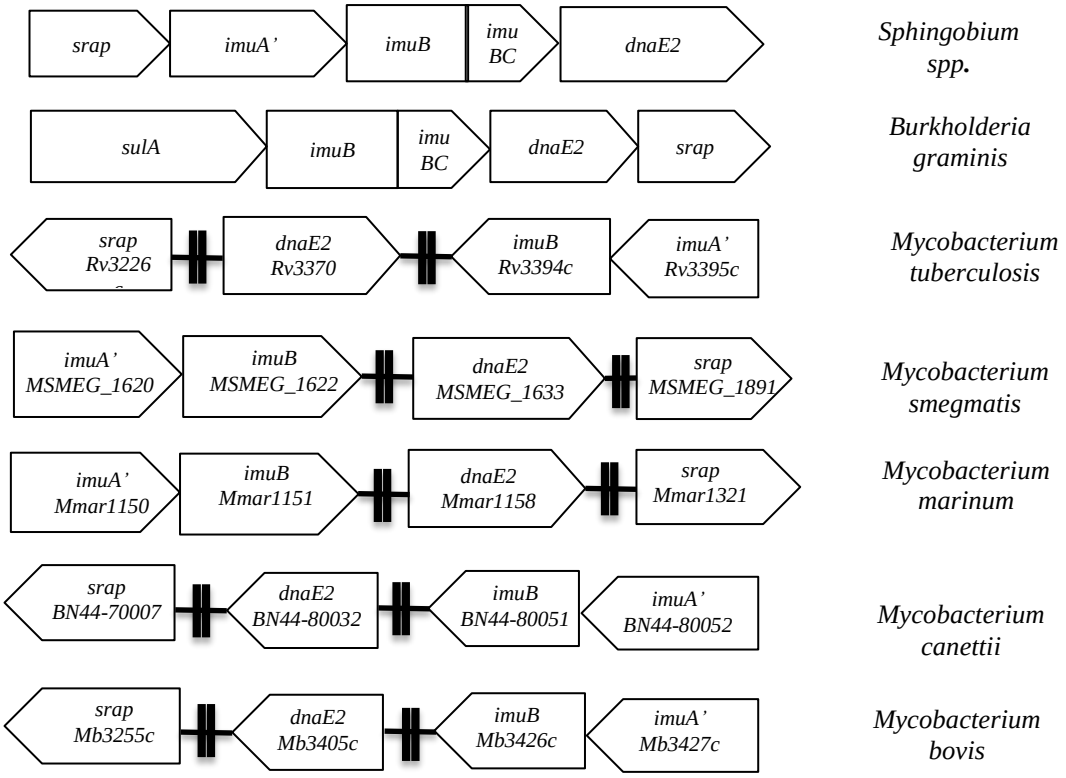
3.3 Results

3.3.1 SRAP in Mycobacteria

Homologs of *E. coli yedK*, which encodes the SRAP protein in that organism, were identified in mycobacterial species including Msm (*MSMEG_1891*) and Mtb (*Rv3226c*). A protein BLAST search using the Mtb SRAP showed it to be most similar to the DUF159 family protein of *Saccharopolyspora erythraea* with 48% identity and 61% similarity. Furthermore, Mtb SRAP exhibited 49.4% similarity and 32.5% identity to *E. coli* YedK. Unlike *Sphingobium spp.*, *Burkholderia graminis* and several other species, the mycobacterial SRAP-encoding gene does not form part of an operon containing other DNA damage response genes (Figure 3.1A); instead, SRAP is surrounded by a variety of hypothetical proteins and transferases. The genomic loci of Msm *srap* and Mtb *srap* are illustrated in Figure 3.1B. Both Msm and Mtb *srap* genes are located close to *aroA* which contributes to the biosynthesis of aromatic amino acids. The Mtb *srap* gene is also in close proximity to *sigH*, which encodes the alternative σ factor forming part of the oxidative stress response. However, *srap* does not form part of the σ^H regulon in Mtb (Manganelli *et al*., 2002).

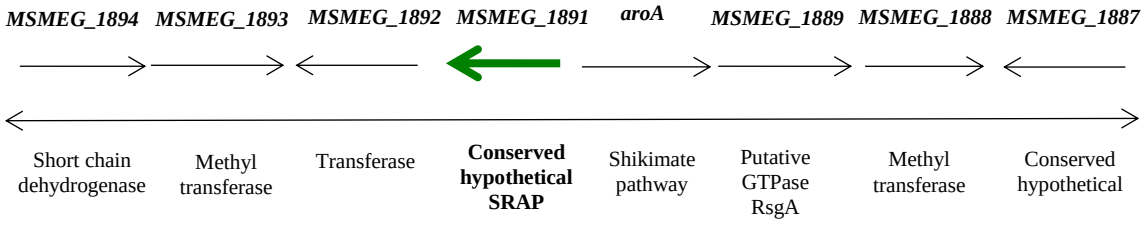
In support of the idea that SRAP could play a role in DNA damage response in mycobacteria, a RecA-NDp motif was identified upstream of *srap* genes in both species (Figure 3.1C). Both RecA-NDp sequences are found 30 bp upstream of the predicted *srap* start codons. The Mtb RecA-NDp differs from the consensus sequence by two nucleotides, whereas Msm RecA-NDp differs by only one. Both Msm and Mtb RecA-NDp differ from the consensus sequence at four positions (Figure 3.1C). These differences could contribute to temporal regulation, with sequences identical to the consensus sequence being more tightly bound to the repressor than sequences deviating from the consensus (Qin *et al*., 2015; Simmons *et al*., 2008).

A

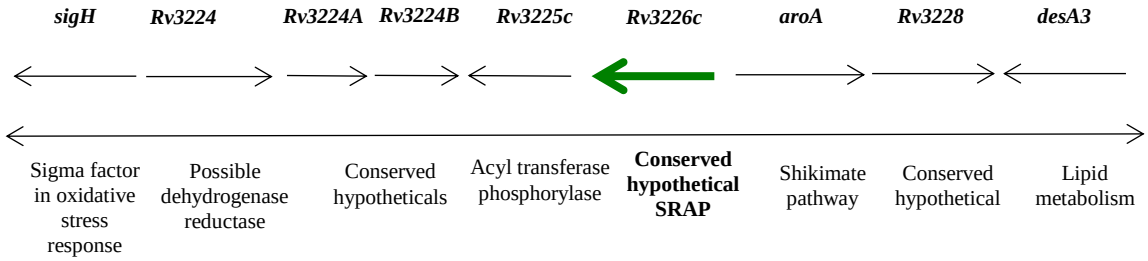

srap
imuA'
imuB
imu BC
dnaE2
Sphingobium spp.
sulA
imuB
imu BC
dnaE2
srap
Burkholderia graminis
srap Rv3226
dnaE2 Rv3370
imuB Rv3394c
imuA' Rv3395c
Mycobacterium tuberculosis
imuA' MSMEG_1620
imuB MSMEG_1622
dnaE2 MSMEG_1633
srap MSMEG_1891
Mycobacterium smegmatis
imuA' Mmar1150
imuB Mmar1151
dnaE2 Mmar1158
srap Mmar1321
Mycobacterium marinum
srap BN44-70007
dnaE2 BN44-80032
imuB BN44-80051
imuA' BN44-80052
Mycobacterium canettii
srap Mb3255c
dnaE2 Mb3405c
imuB Mb3426c
imuA' Mb3427c
Mycobacterium bovis

B

Msm:


MSMEG_1894
MSMEG_1893
MSMEG_1892
MSMEG_1891
aroA
MSMEG_1889
MSMEG_1888
MSMEG_1887
Short chain dehydrogenase
Methyl transferase
Transferase
Conserved hypothetical SRAP
Shikimate pathway
Putative GTPase RsgA
Methyl transferase
Conserved hypothetical

Mtb:


sigH
Rv3224
Rv3224A
Rv3224B
Rv3225c
Rv3226c
aroA
Rv3228
desA3
Sigma factor in oxidative stress response
Possible dehydrogenase reductase
Conserved hypotheticals
Acyl transferase phosphorylase
Conserved hypothetical SRAP
Shikimate pathway
Conserved hypothetical
Lipid metabolism

C

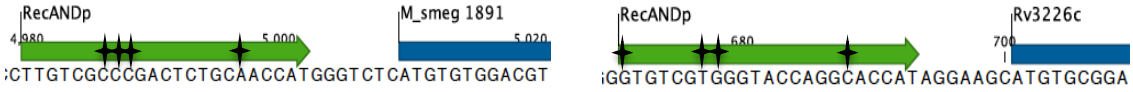

RecANDp
4,980
5,000
M_smeg 1891
5,020
CTTGTCGCCCGACTCTGCAACCATGGGTCTCATGTGTGGACGT
RecANDp
680
Rv3226c
700
GGTGTCGTGGGTACCAGGCACCATAGGAAGCATGTGCGGA

Figure 3.1 Genomic contexts of *srap* homologues in mycobacteria and other organisms. A. Schematic illustration of the association between SRAP and SOS-inducible operons in *Sphingobium* spp. and *Burkholderia graminis* but lack thereof in mycobacterial species. The three mutasome genes, *imuA'*, *imuB* and *dnaE2*, are all pseudogenes in *M. leprae* and are therefore not included in the comparison. Vertical double lines indicate that genes are not adjacent to each other. **B.** Gene environments of Msm and Mtb *srap*. **C.** The RecA-NDp sequences found upstream of Msm and Mtb *srap* suggesting a possible role in DNA damage response. The black stars indicate deviations from the consensus sequence (TTGTCRGTG(N8)TA(N3)T).

3.3.2 Generation and validation of *srap* knock-out mutants

Deletion mutants were successfully generated in both Msm (Δ*MSMEG_1891*) and Mtb (Δ*Rv3226c*), eliminating ~400 bp of the internal region of *srap*. Initial confirmation was done using PCR by designing primers flanking the gene, and an internal set with a reverse primer within the deleted region (Table 2.3, Figure 3.2A). The Mtb mutant contained an in-frame deletion of *srap* whereas *srap* was not deleted in-frame in the Msm mutant. However, this was not expected to make a difference since *srap* was not predicted to form part of an operon (http://tuberculosis.bu.edu/tbdb_sysbio/operon/Rv3226c.html). Deletion mutants were constructed in both wildtype and ΔY backgrounds in Msm. The ΔY strain was previously constructed in our research group, and is a four-gene knockout lacking all Msm Pol IV homologs (*dinP* (*MSMEG_1002*) [which is duplicated in Msm as an unannotated ORF between *MSMEG_2295* and *MSMEG_2296*], *dinX*, (*MSMEG_3178*) and *dinP3* (*MSMEG_6405*); D.F. Warner, unpublished). The ΔY strain background was selected in addition to wildtype in order to eliminate the potential contribution of the mycobacterial Pol IV homologs to the DNA damage response, thereby unmasking any redundancy in function with SRAP. The presence of the mutant allele was confirmed in both the single cross over (SCO) and double cross overs (DCOs) using PCR. When using the external primers, the presence of a single band ~400 bp smaller than the wildtype suggested deletion of the gene (Figure 3.2D). For Msm, the wildtype band was 1261 bp and the mutant band 862 bp, and for Mtb the wildtype band was 939 bp with the mutant band being 540 bp. The possibility of merodiploidy was eliminated by confirming the absence of the deleted region using the reverse primer that binds within the deleted region (internal). Msm Δmsm*srap* and ΔYΔmsm*srap* were further confirmed by Southern blot; the strategy used is set apart in Figure 3.2E. In Mtb, the

deletion mutant, Δmtbs*rap*, was confirmed by WGS to have the 399 bp deletion and no other mutations.

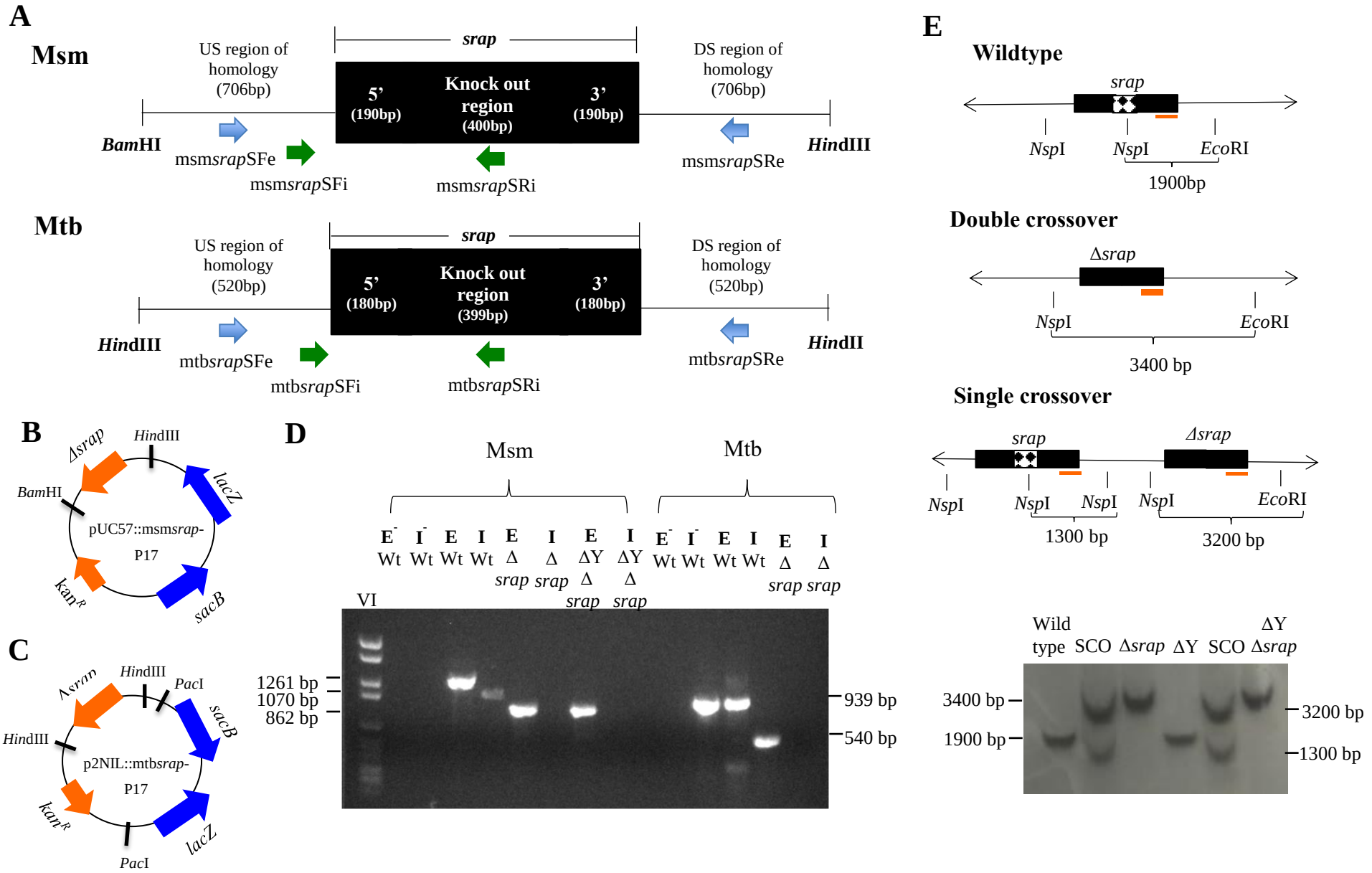

A
Msm
srap
US region of homology (706bp)
5' (190bp)
Knock out region (400bp)
3' (190bp)
DS region of homology (706bp)
BamHI
msmsrapSFe
msmsrapSFi
msmsrapSRi
msmsrapSRe
HindIII
Mtb
srap
US region of homology (520bp)
5' (180bp)
Knock out region (399bp)
3' (180bp)
DS region of homology (520bp)
HindIII
mtbsrapSFe
mtbsrapSFi
mtbsrapSRi
mtbsrapSRe
HindII
B
Δsrap
HindIII
BamHI
lacZ
pUC57::msmsrap-P17
kanR
sacB
C
Δsrap
HindIII
PacI
sacB
HindIII
p2NIL::mtbsrap-P17
kanR
lacZ
PacI
D
Msm
Mtb
E- Wt
I- Wt
E Wt
I Wt
E Δsrap
I Δsrap
E ΔY Δsrap
I ΔY Δsrap
VI
1261 bp
1070 bp
862 bp
939 bp
540 bp
E
Wildtype
srap
NspI
NspI
EcoRI
1900bp
Double crossover
Δsrap
NspI
EcoRI
3400 bp
Single crossover
srap
Δsrap
NspI
NspI
NspI
NspI
EcoRI
1300 bp
3200 bp
Wild type
SCO
Δsrap
ΔY
SCO
ΔY Δsrap
3400 bp
1900 bp
3200 bp
1300 bp

Figure 3.2 Construction and confirmation of *srap* knock-out mutants. A. Design of *srap* constructs with restriction sites used to clone the deletion alleles. The deleted regions as well as the upstream (US) and downstream (DS) regions of homology included to allow homologous recombination are shown. Primers used to screen constructs, SCOs and DCOs are indicated with SFe and SRe flanking the deleted region and SFi and SRi used to confirm the absence of the deleted region elsewhere in the genome. Plasmids used to transform mycobacterial strains of Msm (**B**) and Mtb (**C**) respectively. **D**. PCR confirmation of the deletion mutants in the different genetic backgrounds. Lane one has molecular weight marker VI followed by the Msm and Mtb strains with E and I indicating external and internal primer sets, respectively, with E^- and I^- being non-template controls. **E**. Southern blot strategy used to confirm Msm deletion mutants and resulting southern blot. The wildtype contains the extra ~400 bp internal, indicated by white blocks and the probe is indicated by the orange line.

3.3.3 Lack of DNA damage sensitivity in knock-out mutants

The presence of a RecA-NDp (Figure 3.1C), coupled with the previously reported upregulation of *srap* following different DNA damaging conditions (MMC and H_2O_2) (Namouchi *et al*., 2016; Rand *et al*., 2003; Voskuil *et al*., 2011), supported the idea that SRAP might form part of the DNA damage response in mycobacteria. A suite of well-established assays was employed to evaluate UV mutagenesis and DNA damage sensitivity of Δmsm*srap* and Δmtb*srap*. Firstly, UV mutagenesis assays study induced mutagenesis by measuring the frequency of Rif^R conferring *rpoB* mutations upon DNA damage. DnaE2 has been implicated in induced mutagenesis, and previous work in our laboratory showed that induced mutagenesis was effectively eliminated by functional inactivation of *dnaE2* (Boshoff *et al.,* 2003). Therefore, corresponding Δ*dnaE2* mutants were used as positive controls in both Msm and Mtb. As expected, Δ*dnaE2* in both species showed a mutation frequency two $\log_{10}$ lower than wildtype, whereas no difference could be detected between the SRAP deletion mutants and the parental strain. Mutation frequencies were calculated in Msm wildtype (2.3×10^{-6}), Δ*dnaE2* (9.2×10^{-8}), Δmsm*srap* (1.3×10^{-6}), ΔY (2×10^{-6}) and ΔYΔmsm*srap* (1.7×10^{-6}) (Figure 3A). In Mtb, they were recorded as H37Rv (4.4×10^{-7}), Δ*dnaE2* (2.35×10^{-9}) and Δmtb*srap* (4.7×10^{-7}) (Figure 3.3B). Therefore, there appeared to be no role for SRAP in induced mutagenesis.

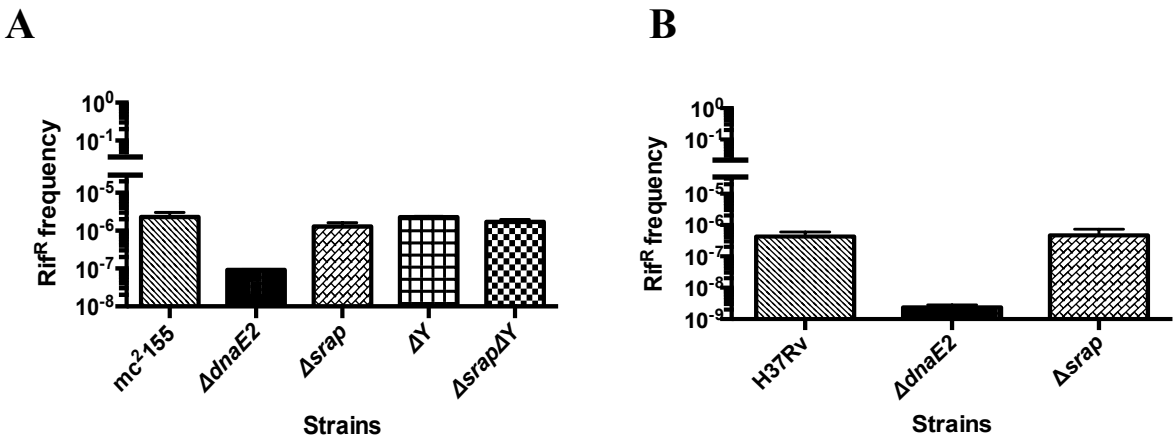


Figure 3.3 Mutation frequencies measured in Msm and Mtb strains. Rif^R frequencies were recorded after UV exposure by plating the cells on Rif containing plates. A. All Msm strains and B. Mtb strains tested. In both cases, Δ*dnaE2* was used as a positive control with a two log-fold reduction in mutation frequency.

Mtb *srap* was found to be upregulated upon MMC treatment in a previous study (Namouchi *et al*., 2016; Rand *et al*., 2003). Here, we assessed the upregulation of Msm *srap* upon DNA damage by means of droplet digital PCR (ddPCR), using *sigA* as reference (Singh *et al*., 2016). The cells were exposed to MMC for 1 h since *recA* was previously determined to reach its upper limit of expression at this time point (Boshoff *et al*. 2003). Readings in the "water only" and "no-template" controls were negligible. The *water only* wells enabled us to identify water contamination whereas *no-template* controls were used to identify contaminated reagents. Msm *srap* was upregulated 6 fold after 1 h of 1 × MIC MMC treatment, which was comparable to *recA* - which was found to be significantly upregulated (13 fold) (Figure 3.4).

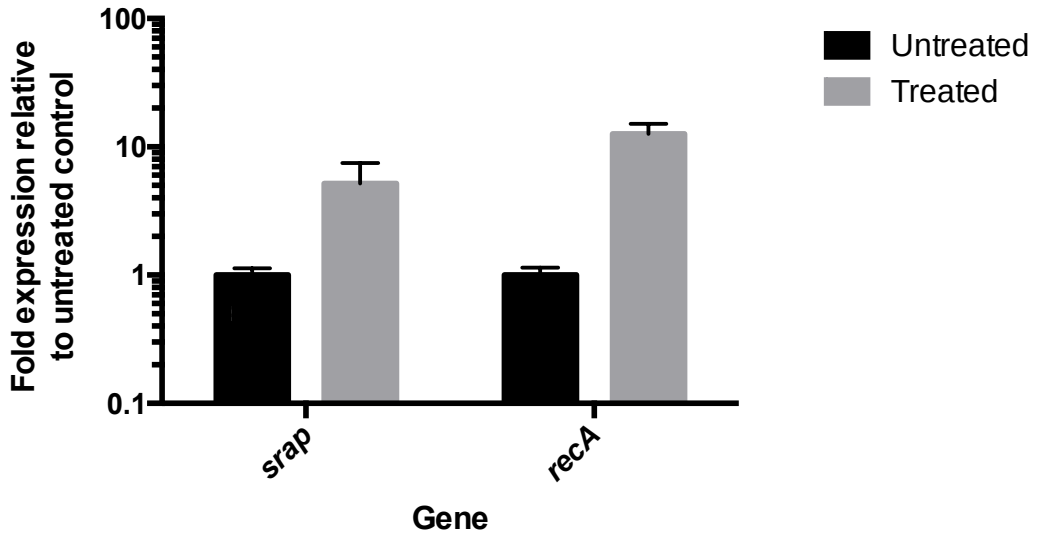
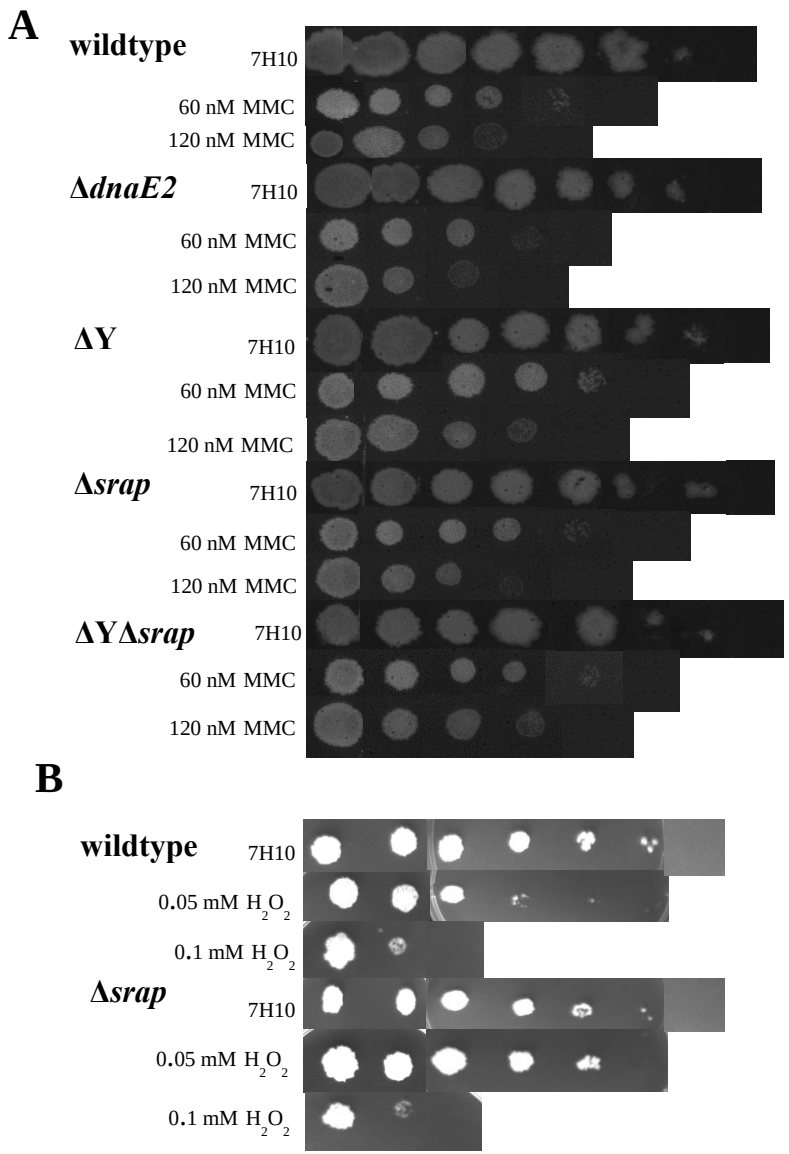


Figure 3.4 Relative expression of *recA* and *srap* in Msm after 1 h of MMC treatment. The expression of *srap* and *recA* in treated samples are represented relative to its abundance in untreated samples, indicating the fold increase in treated samples on a log scale. Gene expression of *srap* and *recA* were normalised to *sigA* expression levels. Represented results are from three biological replicates with error bars representing standard deviation.

The confirmed induction of *srap* following MMC exposure prompted us to test the DNA damage sensitivity of the Δmsm*srap* strain. Again, Δ*dnaE2* was used as a positive control. The SRAP deletion strains phenocopied the wildtype strains whereas the Δ*dnaE2* mutant was 1 $\log_{10}$ more sensitive than wildtype at both MMC concentrations tested (Figure 3.5A). Upregulation of Msm *srap* has been reported previously following H_2O_2 treatment (Li *et al*., 2015), therefore sensitivity of the *srap* mutants to H_2O_2 was also tested (Figure 3.5B). Again, similar sensitivities to H_2O_2 were observed for both Δmsm*srap* and wildtype. This was confirmed by obtaining similar MIC_{90} values for wildtype and Δmsm*srap* in broth microdilution assays (Collins & Franzblau, 1997), both being 0.25 – 0.49 mM (Figure 3.5C).


A
wildtype
7H10
60 nM MMC
120 nM MMC
ΔdnaE2
7H10
60 nM MMC
120 nM MMC
ΔY
7H10
60 nM MMC
120 nM MMC
Δsrap
7H10
60 nM MMC
120 nM MMC
ΔYΔsrap
7H10
60 nM MMC
120 nM MMC
B
wildtype
7H10
0.05 mM H2O2
0.1 mM H2O2
Δsrap
7H10
0.05 mM H2O2
0.1 mM H2O2

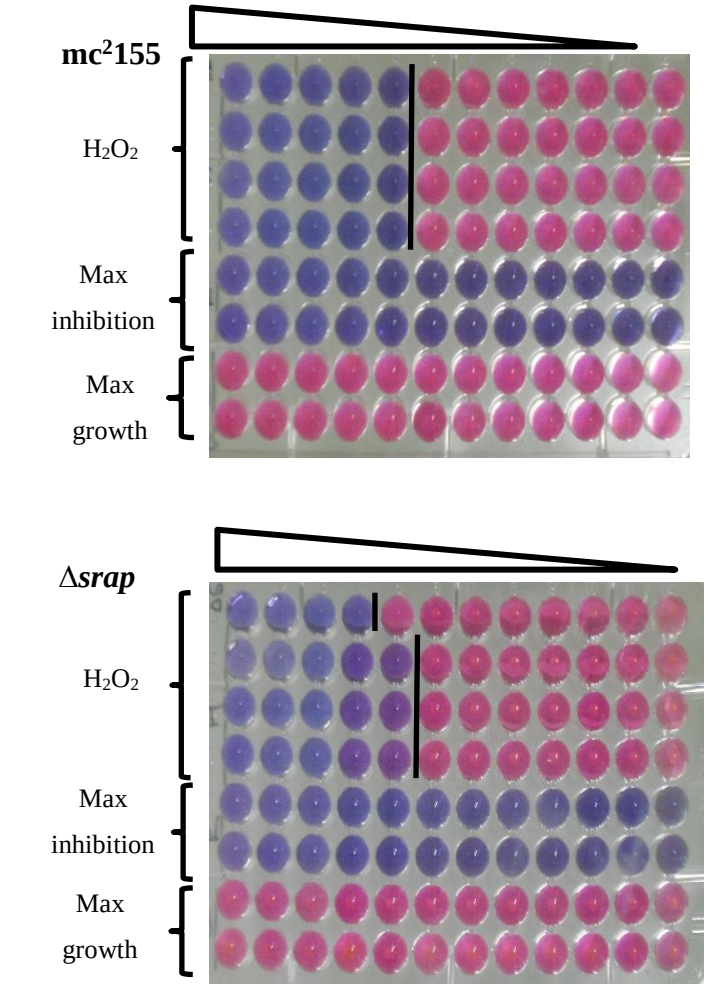

C
mc²155
H2O2
Max inhibition
Max growth
Δsrap
H2O2
Max inhibition
Max growth

Figure 3.5 DNA damage sensitivity of Msm *srap* deletion mutants. A. Sensitivities to MMC of the panel of Msm strains. **B.** Sensitivities of wildtype and the *srap* deletion mutant to H_2O_2. For both MMC and H_2O_2, log-fold dilutions were spotted on media without and with antibiotic at specified concentrations. **C.** H_2O_2 MIC plates of wildtype and Δmsms*rap* starting at 3.9 mM H_2O_2. Data are from a single representative experiment performed in triplicate.

A wider range of DNA damaging agents was tested against Δmtbs*rap*: MMC, ofloxacin (Ofx), plumbagin (Pmb) and nitrofurazone (Nfz) (Figure 3.6). The agents chosen have different mechanisms of action, causing DNA damage in different ways. Pmb induces oxidative DNA damage by generating superoxide (Mosel *et al*., 2003), MMC generates DNA crosslinks (Matsumoto *et al*., 1989), Ofx inhibits DNA gyrase (Drlica & Zhao, 1997) and Nfz causes N^2-dG adducts (Jarosz *et al*., 2006). For Pmb, both Δ*dnaE2* and Δmtbs*rap* phenocopied H37Rv, whereas the Δ*dnaE2* strain was $1\log_{10}$-fold more sensitive than H37Rv and Δmtbs*rap* at both concentrations of Ofx. The high concentration of Nfz completely abolished growth at 161.6 μM and Δ*dnaE2* was $1\log_{10}$-fold more sensitive to 80.0 μM Nfz than H37Rv and Δmtbs*rap*. In summary, none of the Msm or Mtb deletion mutants showed any hypersensitivity to the DNA damaging agents tested. .

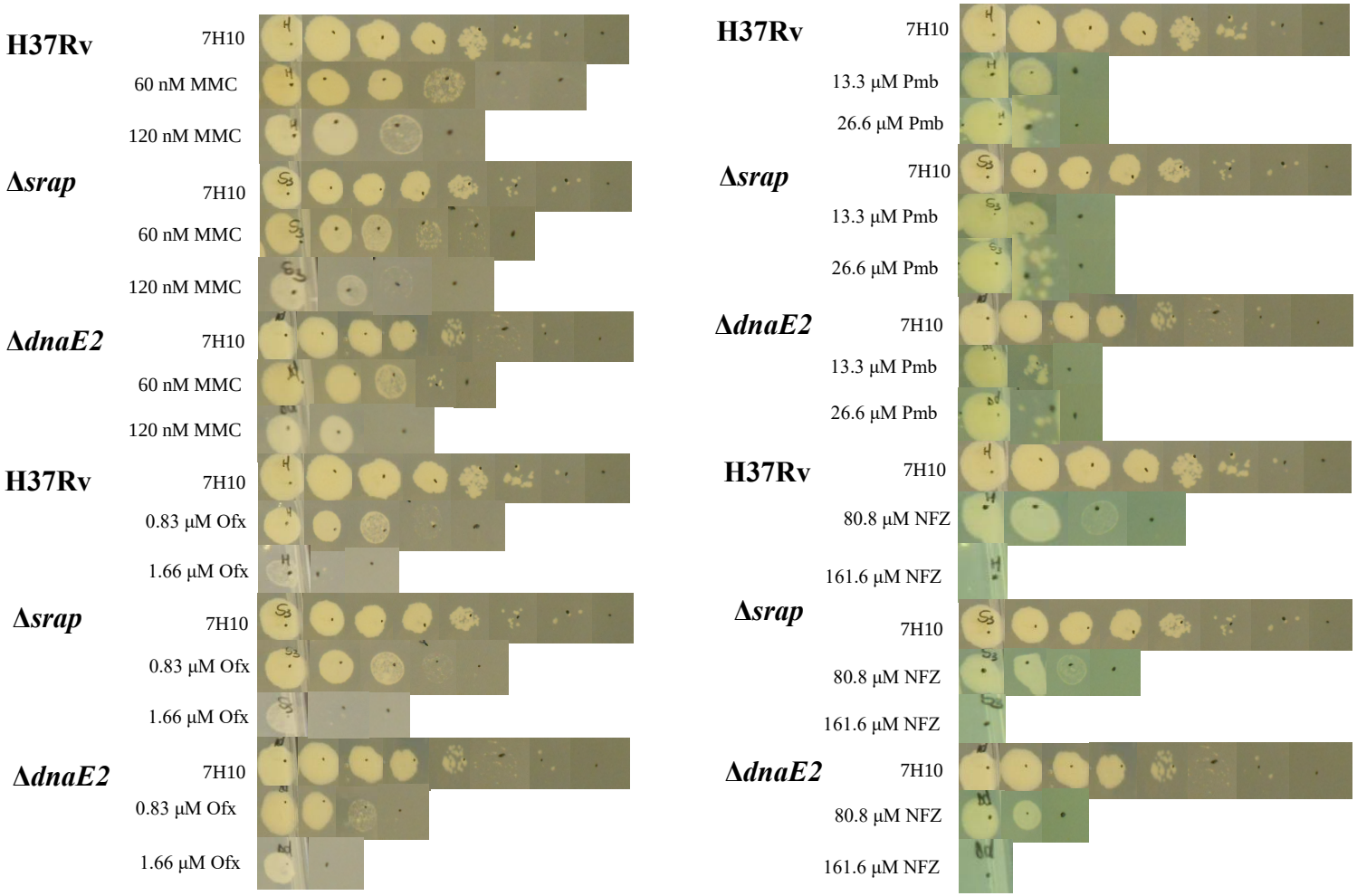


Figure 3.6 DNA damage sensitivity of Mtb strains. Log-fold dilutions were spotted on media without antibiotic and media with antibiotic at specified concentrations. Data are from a single representative experiment performed in triplicate.

Lack of detectable sensitivity suggested that SRAP might not have a unique role in the mycobacterial DNA damage response. However, an alternate explanation is that these assays were not sufficiently sensitive to detect low-level hypersensitivity or functional redundancy. Owing to the upregulation of Msm *srap* upon MMC treatment, a reporter mutant expressing a translational fusion linking Msm SRAP with enhanced green fluorescent protein (eGFP) was constructed to enable visualisation and localisation of *Msm* SRAP. Visualising and ultimately localising SRAP production could give insight into its function, for example if it co-localises with the mutasome it could suggest a role during DNA damage response. The Msm strain containing the fusion protein was referred to as SSE. As a first step to test if any expression could be detected in SSE a positive control where the eGFP gene is under the control of the damage inducible *imuA'* promoter was included in the experiment (pSOS-*egfp*). The pSOS-*egfp* strain was previously constructed in our laboratory (M Reiche, unpublished), it is routinely used and known to fluoresce upon DNA damage. eGFP expression was detected in the positive control, as was expected. However, no clear difference in eGFP expression could be detected between SSE following treatment with 1 × MIC MMC for 3 h and the untreated sample. Low level expression was observed, even in the untreated samples of wildtype and SSE, and slight variation between samples could be due to auto-fluorescence. Variable levels of auto-fluorescence were observed throughout the fluorescent microscopy experiments (Figure 3.7). The lack of significant fluorescence in SSE could be because the expression level of *srap* is not high enough to detect or another possibility is that the eGFP protein does not fold correctly in the fusion protein.

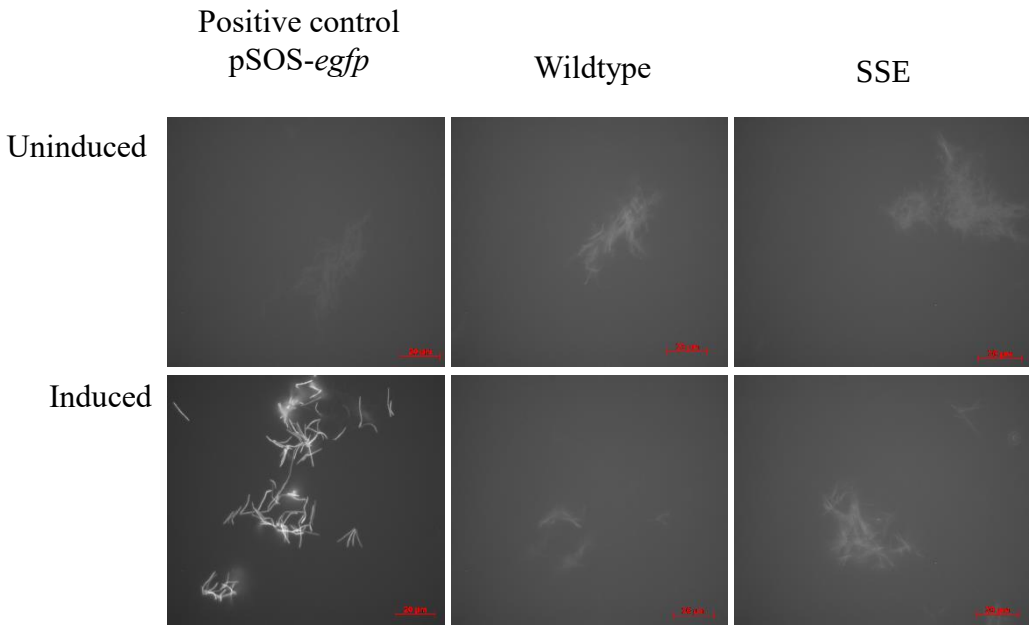


Figure 3.7 MMC treatment of pSOS-*egfp*, wildtype and SSE strains. Strains were treated with 1 × MIC MMC for 3 h and visualised using fluorescent microscopy. Induction of eGFP and subsequent fluorescence is only visible in pSOS-*egfp*.

3.3.4 Msm SRAP and oxidative stress

There are several indicators that point to a potential role for SRAP in the response of mycobacterial to oxidative stress: both Msm (Li *et al*., 2015) and Mtb *srap* were reportedly induced upon H_2O_2 treatment (Voskuil *et al*., 2011), and Mtb *srap* was upregulated in mouse macrophages (Schnappinger *et al*., 2003). To attempt visualisation of SRAP expression in response to oxidative DNA damage, we cultured and exposed wildtype and SSE mutant to 7 mM H_2O_2 (where maximum upregulation was observed by Li *et al*., 2015) in the absence of extracellular catalase and visualised expression using fluorescent microscopy (Figure 3.8). Catalase was omitted to maximise the effect of H_2O_2 treatment. As is evident in Figure 3.8, fluorescence increases after H_2O_2 treatment in both strains after 1 h and 3 h; however, fluorescence intensity seems to decrease from 1 h to 3 h of treatment. There is no apparent difference between wildtype and the SSE strain; however, due to the slight variation in fluorescence between cells, fluorescence was quantified per cell (ImageJ-Fiji).

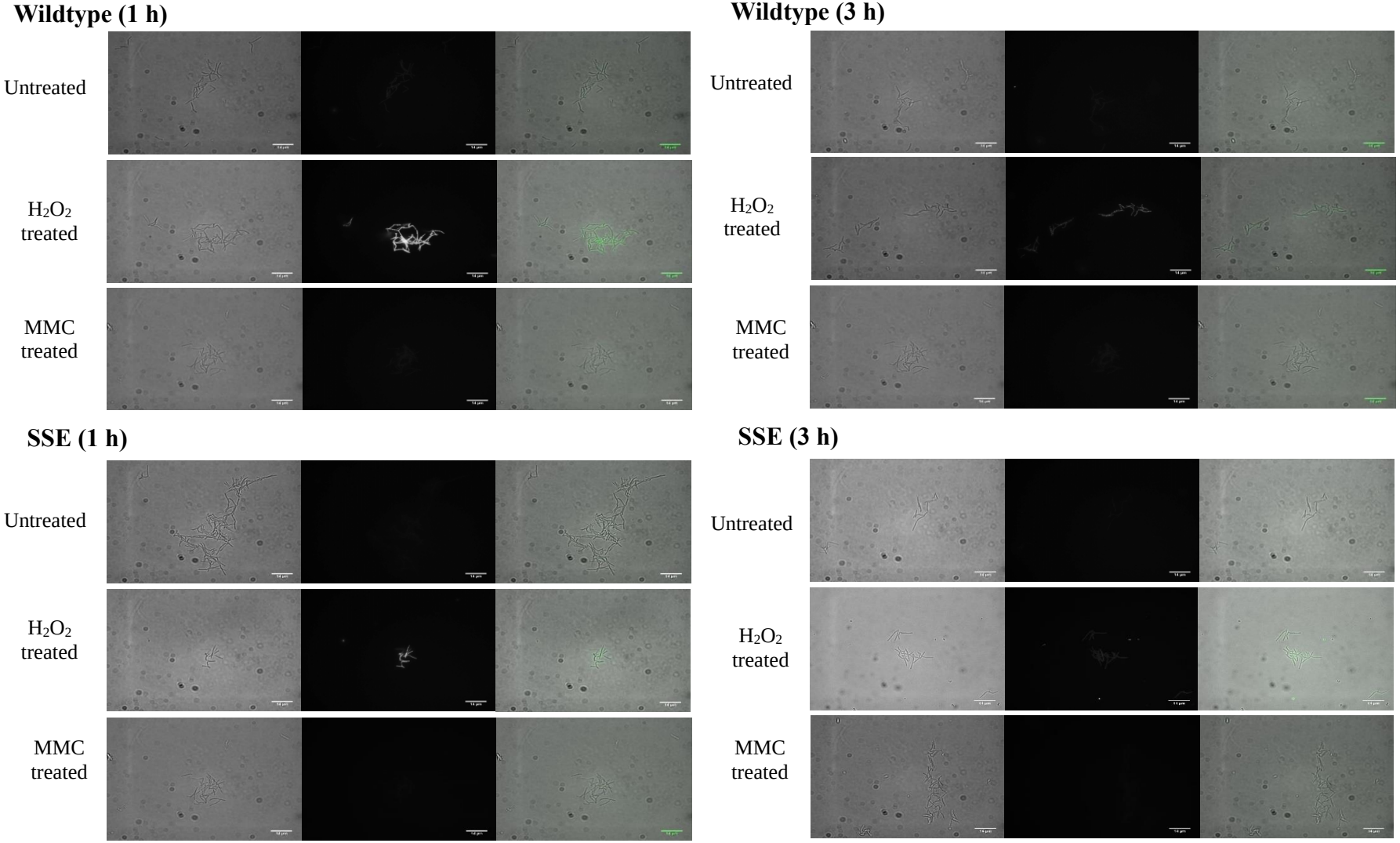

Wildtype (1 h)
Untreated
H_2O_2 treated
MMC treated
Wildtype (3 h)
Untreated
H_2O_2 treated
MMC treated
SSE (1 h)
Untreated
H_2O_2 treated
MMC treated
SSE (3 h)
Untreated
H_2O_2 treated
MMC treated

Figure 3.8 H_2O_2 treatment of wildtype and SSE strains. Wildtype and SSE were treated with 7 mM H_2O_2 for 1 and 3 hours and 1 × MIC MMC for 1 and 3 h in the absence of external catalase. Samples were viewed using a fluorescent microscope and fluorescence were detected in all strains treated with H_2O_2, however no fluorescence in the MMC treated samples. Data are from a single representative experiment done in triplicate.

Unexpectedly, all strains, including wildtype which does not carry an eGFP construct, showed a significant increase in fluorescence ($p < 0.05$) upon H_2O_2 treatment (Figure 3.9). Maximum fluorescence was observed after 1 h of H_2O_2 treatment in both wildtype and SSE strains; around 3 h, the total fluorescence decreased, presumably as cells recovered or died, although the difference between untreated and treated remained significantly different. There was, however, no significant difference between fluorescence intensity of wildtype and SSE strains. To test whether wildtype fluorescence in the absence of external catalase was unique to H_2O_2 exposure, we also treated the strains with 1 × MIC MMC for 1 h and 3 h without adding catalase (Figure 3.8). Unlike H_2O_2, no fluorescence was observed in wildtype or SSE at either time point, establishing that this phenomenon is not a general consequence of genotoxic stress. No fluorescence was visible in the red spectrum, and variable auto-fluorescence was observed in the cyan spectrum.

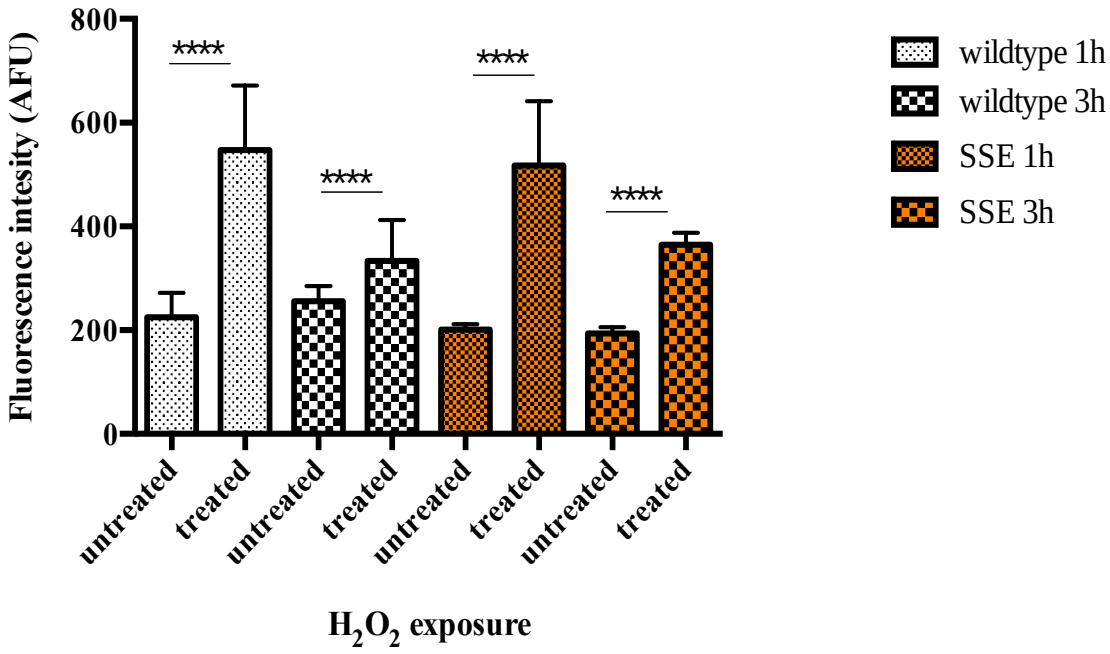


Figure 3.9 Fluorescence intensity of wildtype and SSE strains following exposure to oxidative stress. The fluorescence intensity of the two strains were quantified in both H_2O_2 treated and untreated samples after 1 h and 3 h. Wilcoxon signed-rank tests were used to determine if fluorescence were significantly different. Data are from a single representative experiment performed in triplicate.

3.3.5 SRAP modelling and expression

To gain better insight into SRAP function, we set out to express a recombinant form of Msm SRAP with the aim of solving its crystal structure and establishing whether this protein has peptidase activity, as has been predicted by Aravind *et al.* (2013). As a starting point, homology models were constructed for both Msm and Mtb SRAP. Six of the seven conserved residues reported by Aravind *et al.* (2013) are present in both proteins, with the catalytic triad consisting of a cysteine (C2), glutamic acid (E119) and histidine (H183). The putative catalytic triad of Mtb SRAP is represented in Figure 3.10C. Furthermore, two conserved arginines (R102 and R185) are present that could potentially facilitate DNA binding, as well as the conserved glutamine (N89). However, in place of conserved glutamic acid, an aspartic acid is present in the mycobacterial SRAP proteins (D93) (Figure 3.10D).

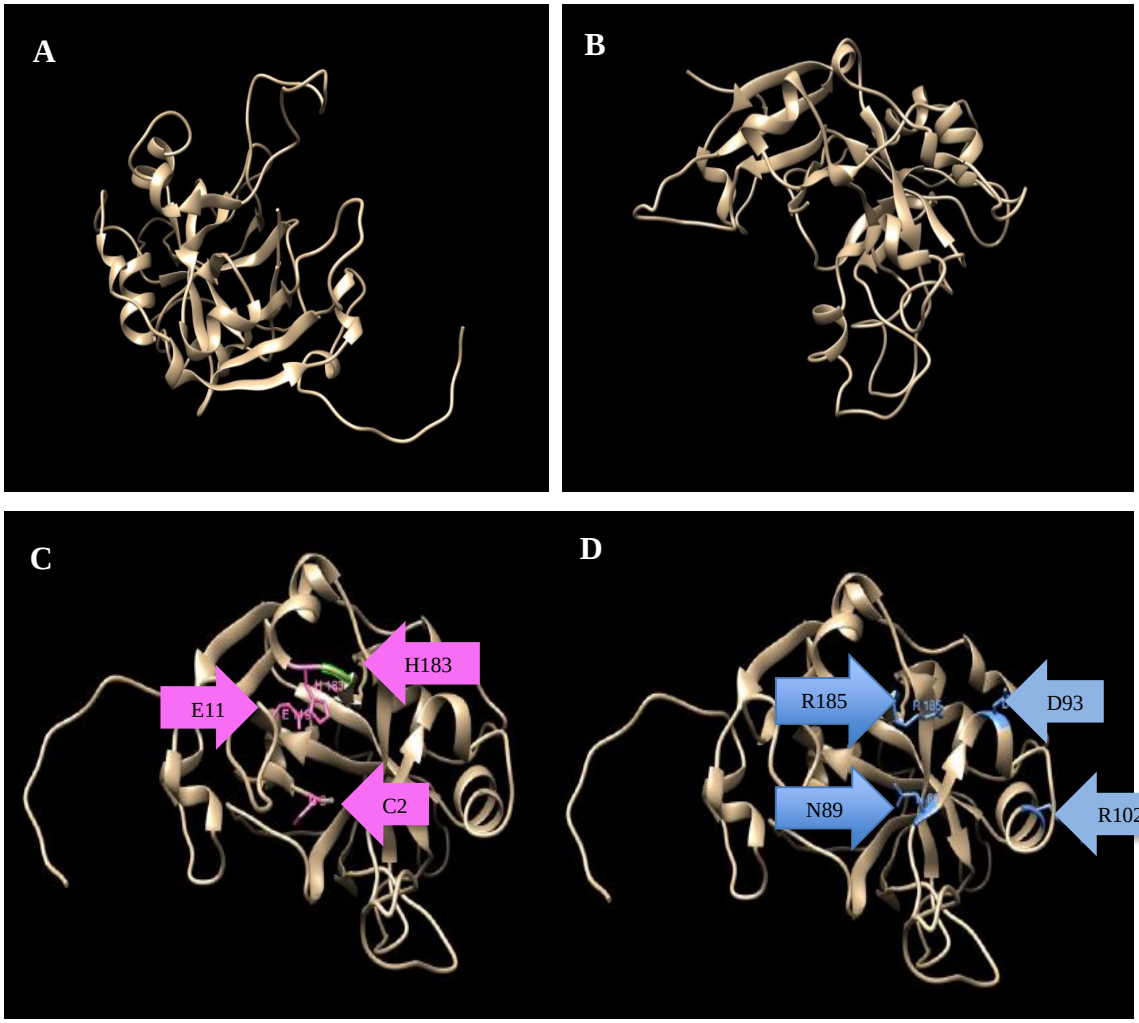


Figure 3.10 Predicted SRAP structure and conserved residues. Homology models of **A.** Mtb and **B.** Msm SRAP proteins. Important conserved residues are shown on Mtb SRAP. **C.** Representation of the catalytic triad. **D.** Other conserved residues as predicted by Aravind *et al.*, 2013. Figures courtesy of Dr Natasha Wood (Msm SRAP) and Dr. Česlovas Venclovas (Mtb SRAP).

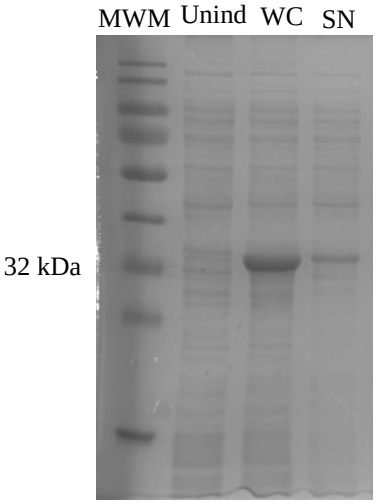
Both models are consistent with the prediction that SRAP is a cysteine peptidase, and locate the putative catalytic residues in an active site pocket. Furthermore, the two conserved arginines and the conserved glutamine are also present in close proximity to the catalytic triad (Figure 3.10D). However, the presence of these conserved residues is only a starting point to determine if mycobacterial SRAP indeed has peptidase activity. To test this experimentally, we expressed the Msm SRAP by cloning it in frame with an N-terminal hexa-histidine (His)-tag in the pCOLD1 vector. This was

expressed in *E. coli* BL21, lacking proteases Lon and OmpT to optimise expression of the 287 amino acid protein, estimated to be ~31.57 kDa.

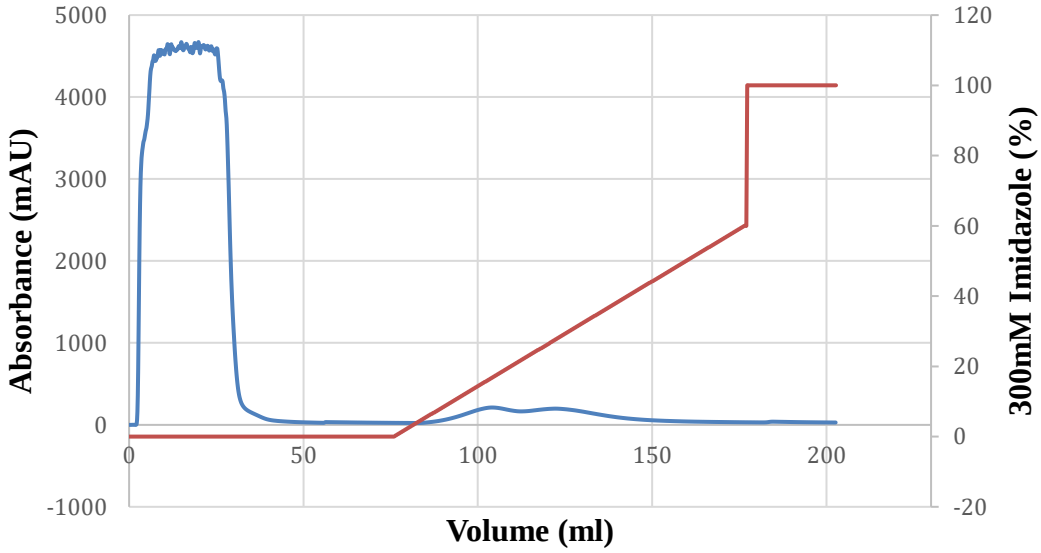
3.3.5.1 Soluble expression and purification

Following 24 hours of IPTG induction, the cells were recovered and sonicated after which the supernatant, with the soluble protein fraction, was separated from the pellet with the insoluble fraction. Aliquots of the supernatant, together with the uninduced and induced whole cell lysates, were run on an SDS-PAGE gel (Figure 3.11A). The presence in the supernatant fraction of the induced protein, at the expected size, was indicative of soluble expression; this was encouraging, and suggested proper protein folding, allowing downstream analyses. Some of the induced protein was observed in the insoluble fraction (data not shown) but, enough protein (~20 mg) was obtained from the soluble fraction to continue to purification (Figure 3.14).

A


MWM
Unind
WC
SN
32 kDa

B


Absorbance (mAU)
5000
4000
3000
2000
1000
0
-1000
0
50
100
150
200
Volume (ml)
120
100
80
60
40
20
0
-20
300mM Imidazole (%)

C

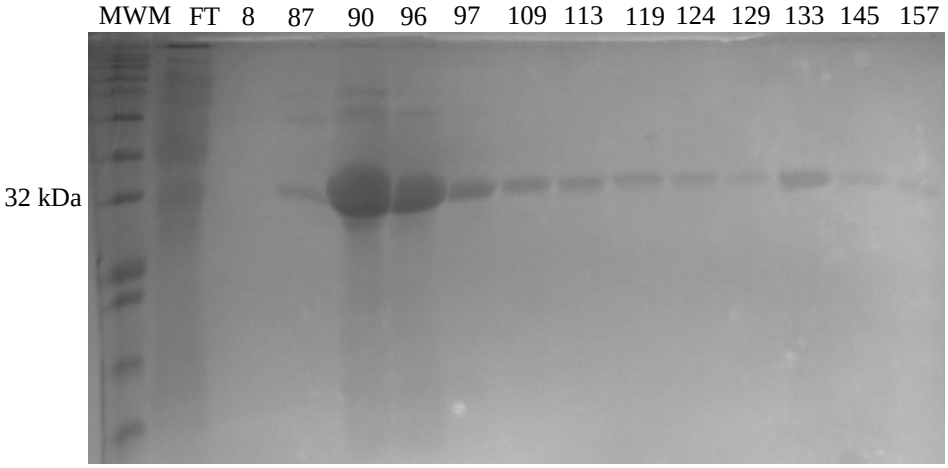

MWM
FT
8
87
90
96
97
109
113
119
124
129
133
145
157
32 kDa

Figure 3.11 Msm SRAP expression and purification. A. The protein sample was extracted from uninduced cells (Unin) and following IPTG induction (WC). The WC sample was then separated by centrifugation into pellet and supernatant (SN) and soluble expression was confirmed by the presence of the expected band size in the SN fraction. **B.** Nickel affinity chromatography was used to purify Msm SRAP. The graph depicts the elution, and increase in absorbance, of the proteins over a gradient increase in imidazole concentration. The blue line represents protein elution and the red line the imidazole gradient. **C.** Samples from fragments representing the two elution peaks were run to evaluate purity. All samples eluting between 97 ml and 157 ml were considered >90% pure.

The addition of an N-terminal His-tag to Msm SRAP allowed for purification using nickel affinity chromatography. Purification was accomplished using an imidazole gradient (0% to 60% of 300 mM imidazole) over 20 column volumes (Figure 3.11B). The initial spike in absorbance was due to the flow through of all unbound proteins. Most of the bound proteins seemed to elute between 6% and 46% imidazole, forming two peaks. Samples representative of fractions containing the shoulders and apexes of both peaks were run on a SDS–PAGE gel (Figure 3.11C). The two peaks represented only the target protein and could indicate two general species of protein with one binding the column more strongly possibly because the His-tag is more accessible to chelate nickel. The samples of the left shoulder of the first peak contained contaminating proteins and were excluded from downstream analyses. A subsample of pure protein was analysed by peptide mass fingerprinting (PMF) at the CSIR (Pretoria, South Africa) and was confirmed to be Msm SRAP. Samples between volumes 97 ml and 157 ml were pooled (60 ml × 0.3 mg/ml) and concentrated for size exclusion chromatography (SEC). The protein concentration of the sample loaded for SEC was 6.4 mg/ml (Figure 3.14).

3.3.5.2 Native multimeric conformation of Msm SRAP

SEC is a chromatographic method used to separate molecules in solution based on size, and is non-reducing and allows identification of multimer formation. A Superdex-200 column was used with a range of 10 000 – 60 000 Da. We consistently observed three peaks, indicating three different conformations (Figure 3.12A). SDS-PAGE confirmed that the peaks were only represented by the target protein (Figure 3.12B). The nonameric peak was detected by samples eluting at 55 min, the trimer eluted at 70 min and the monomer at 85 min.

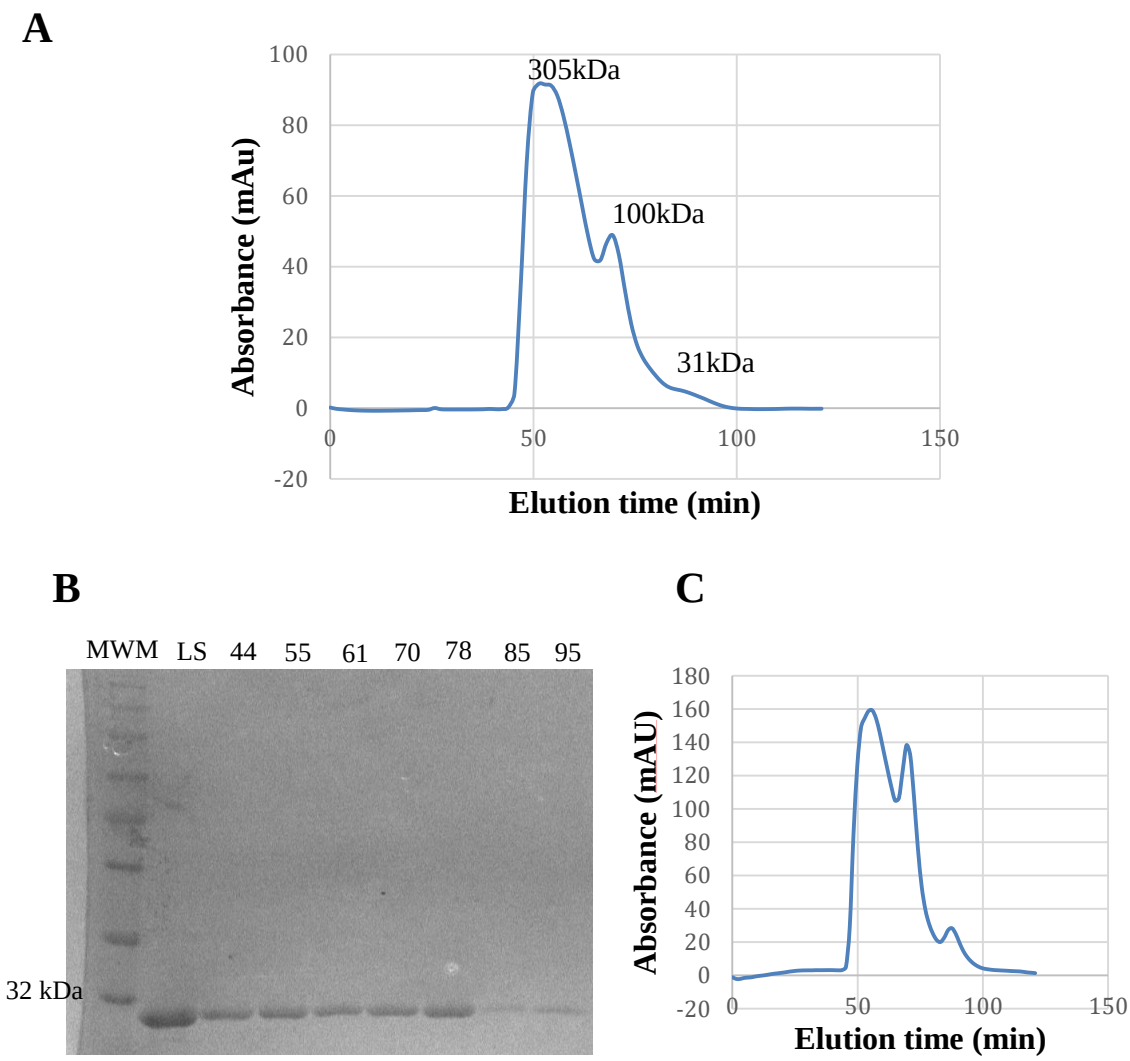


Figure 3.12 Size exclusion chromatography of Msm SRAP. A. Three peaks were consistently observed corresponding to the approximate size of a monomer (31 kDa), trimer (100 kDa) and nonamer (305 kDa). **B.** The loaded sample (LS) and fractions representative of the three peaks were run on a SDS-PAGE gel to confirm purity (>95% pure). The concentration of protein in the fractions representing the monomeric peak is lower than those representing the other two peaks. **C.** The fractions representing the nonamer peak were pooled, incubated at 4°C overnight and re-run on the S200. The three peaks were observed again indicating protein equilibrium.

The sizes of these peaks were calculated to be ~31 kDa, ~100 kDa and ~305 kDa using the standard curve (y = -0.4827x + 2.734) and V_o = 48,67 and V_t = 114,9 (Figure 3.13). These peaks correspond to size estimates of a monomer, trimer and nonamer. To determine whether this system was in equilibrium, the samples from the nonamer were collected, pooled and stored at 4°C overnight. The S200 run was repeated and the previously observed three peaks were detected again, illustrating break down of the nonamer to trimers and monomers to re-establish equilibrium (Figure 3.12 C).

Kav=(Ve-Vo)/(Vt-Vo)

Plot Kav on Y-axis vs LogMW on X-axis

Vo			48,67
Vt			114,9
Elution time	**Kav**	**LogMW**	**MW (Da)**
Standards			
158000	0,22	5,19	63,48
44000	0,49	4,64	81,45
17000	0,69	4,23	94,41
Test			
54,39	0,08	5,48	305528,26
69,88	0,32	5,00	100120,46
86,11	0,56	4,49	31106,20

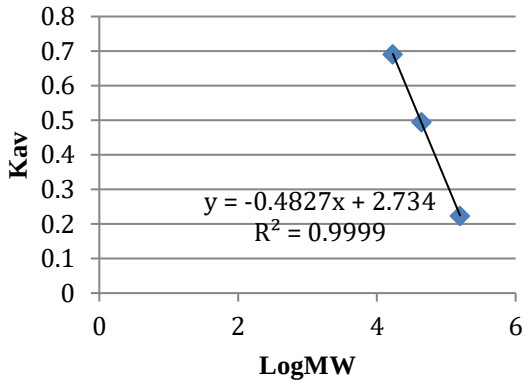


Figure 3.13 Molecular weight estimations of SEC peaks. Standard curve and calculations used to calculate the size of the three peaks from the S200 size exclusion chromatography.

3.3.5.3 Crystal trials and TEM

In an attempt to solve the crystal structure of this protein, samples of the trimeric peak were pooled and concentrated to 70 µl at 7.8 mg/ml (Figure 3.14). The Hampton PEG-Ion Screen™ was used as a starting point for screening. Plates were incubated at room temperature for 21 days and regularly screened for changes in precipitation. Conditions that yielded precipitation in the two droplets with highest protein concentration but remained clear at the low protein concentration were identified to set up fine grid screens to try reach the supersaturation zone. The supersaturation zone is between the precipitation zone (too high protein concentration) and under-saturated zone (too low protein concentration) and is where nucleation and crystal growth occurs. A total of seven precipitants were identified, however only three (KCl, NH_4Cl and CH_3CO_2K) were chosen based on availability to set up fine grid screens. Three hanging drops with different protein concentrations were used in each condition. No crystallisation was observed within 51 days and, owing to time constraints, further optimisation was not pursued.

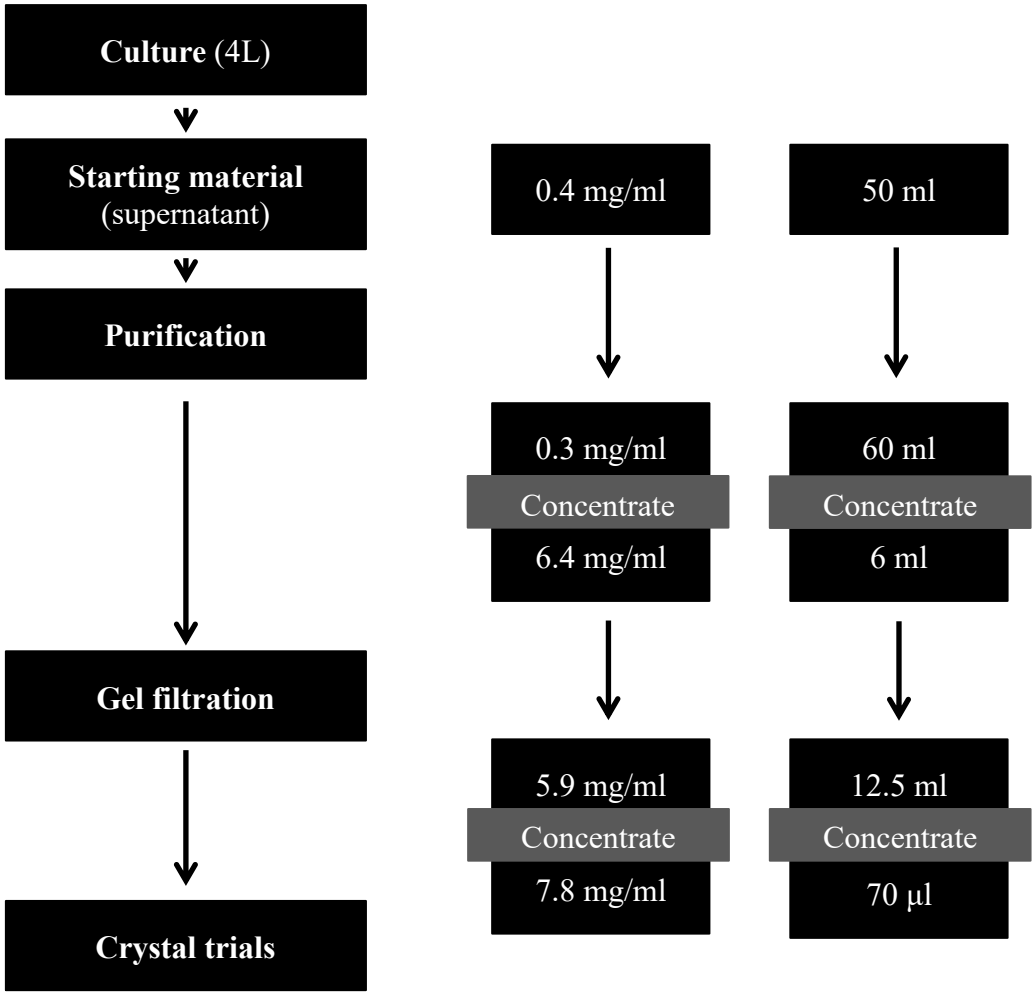


Figure 3.14 Schematic representation of the steps obtaining pure protein for crystal trials. The protein concentrations and volumes obtained throughout the various steps leading up to crystal trials are indicated at each step.

As an alternative to crystallisation, transmission electron microscopy (TEM) at cryogenic temperatures, cryo-electron microscopy (cryo-EM), was explored (Kennaway *et al*., 2005; Wang *et al*., 2015). The homogeneity of the protein mixture was examined using TEM. The nonamer was the only molecular conformation large enough to view using EM, and TEM grids were prepared using dilutions of the fractions representing the nonamer. The protein seemed to be very homogenous (Figure 3.15) and a good candidate for cryo-EM. However, time constraints prevented any further work, and it is recommended that this aspect is pursued later in our laboratory.

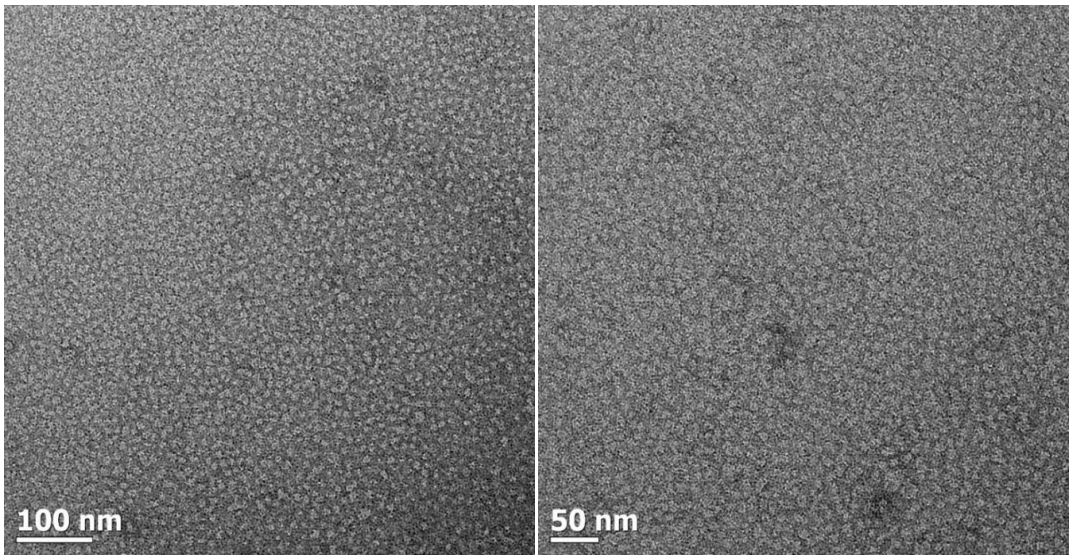


Figure 3.15 Transmission electron microscopy grids of pure SRAP. Dilutions of the samples from the nonameric peak were used to make TEM grids to assess the potential for cryo-EM.

3.3.6 Msm SRAP lacks detectable protease activity

The predicted proteolytic activity of SRAP was evaluated using zymography. Zymography is a functional assay that separates hydrolytic enzymes based on their molecular weights by means of non-reducing SDS-PAGE. Hydrolysis is detected by the ability of the protein to degrade a specific substrate embedded in the gel. Coomassie blue staining reveals sites of proteolysis as white bands on a dark blue background. Human matrix metalloproteinase 1 (MMP1) was used as a positive control. MMP1 is an interstitial collagenase that is synthesised as an inactive zymogen and only gets activated upon removal of the pro-peptide. In this study, hydrolysis of two general substrates was tested, casein and gelatin. Different concentrations of the purified Msm SRAP were tested, but no activity could be detected at any of the concentrations even though MMP1 (~45 kDa) showed proteolysis of both substrates (Figure 3.16). The incubation period allowed for renaturation of the proteins after the SDS was removed by incubating the gel in Triton-X. Since no activity could be detected after 20 h incubation with casein as the substrate, a 40 h incubation period was also tested. Despite the extended duration, no proteolytic activity could be detected for either casein or gelatin. A non-reducing SDS-PAGE was run to confirm the expected sizes of the protein multimers on the zymography gels (data not shown). No proteolysis could be

seen corresponding to any of the bands observed in the non-reducing SDS-PAGE. The lack of proteolysis could indicate the lack of peptidase activity but it could also point towards substrate specificity and/or specific conditions required for activation of SRAP.

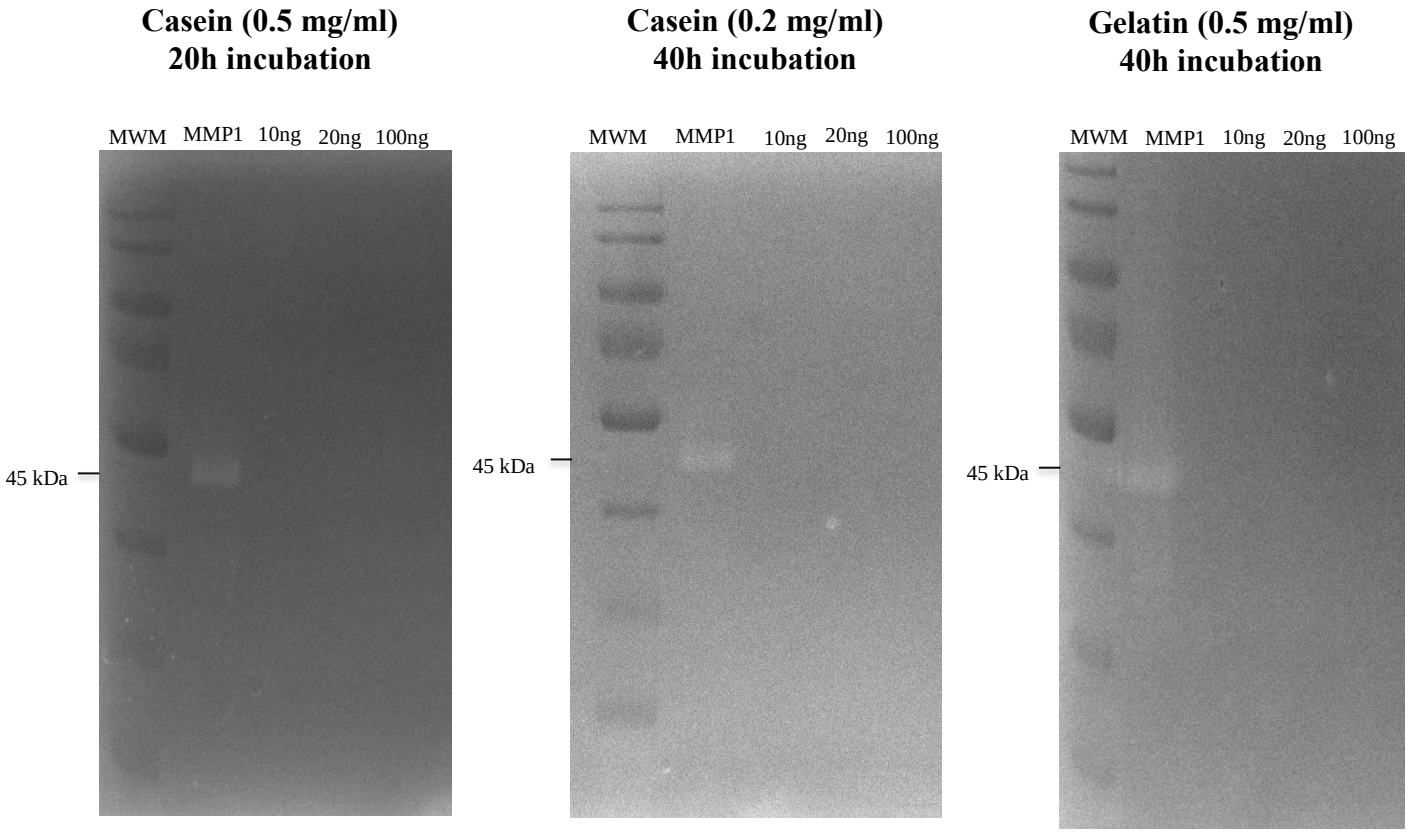


Figure 3.16 Zymography gels to test proteolytic function of Msm SRAP. Zymography gels showing proteolysis of gelatin and casein by MMP1 but lack of proteolysis by different concentrations of Msm SRAP.

3.3.7 Msm SRAP shows no N-terminal cleavage

The lack of activity seen in the zymography runs could be due to the lack of enzymatic activation. Activation of this protein is predicted to require cleavage of the N-terminus at C2. This prediction was based on the absence of the N-terminal tag used for purification in the crystal structure of YedK, the *E. coli* homolog (PDB: 2icu) (Aravind *et al*., 2013). We tested N-terminal cleavage of Msm SRAP using the His-tag. Pure protein was incubated at room temperature, 37°C and 42°C for 24 h, 48 h and 72 h and the presence of the N-terminal His-tag was visualised using a His-tag stain. Equal amounts of protein were run in each lane (Figure 3.17) and this was verified by staining the gel with AquaStain. The His-tag stain intensity seemed to be similar at all temperatures at all times, indicating lack of cleavage. The possibility remains that ideal conditions for cleavage were not met; e.g., pH or presence of specific metabolites or substrate. However, under standard conditions used for zymography, no visible N-terminus cleavage occurs which could explain lack of peptidase activity.

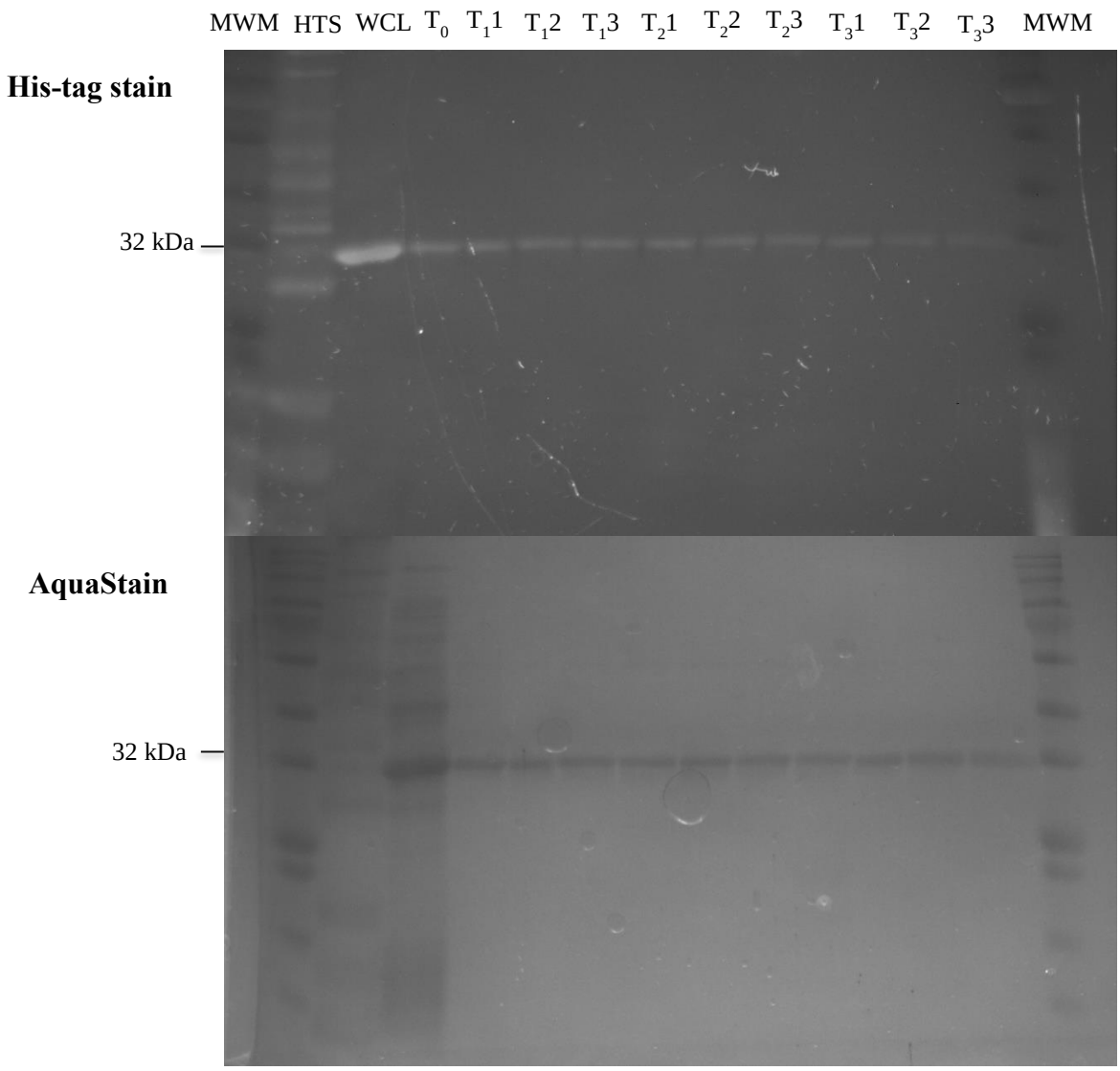


Figure 3.17 His-tag stain of SDS-PAGE gel to detect the N-terminus hexa His-tag. A molecular weight marker was run followed by a His-tag standard (HTS) and whole cell lysate (WCL). Samples were incubated at the three different temperatures for three different time periods were evaluated. T_1 represents 24 h, T_2 is 48 h and T_3 is 72 h, the 1, 2 and 3 following the time represents room temperature, 37 ^{0}C and 42 ^{0}C, respectively. The N-terminal His-tag was observed in all the samples. AquaStaining revealed similar protein concentrations in all samples.

3.4 Discussion

The bacterial DNA damage response is a versatile, dynamic and complicated interactive network. The major DNA damage response pathway in Mtb is the RecA-ND mechanism (Rand *et al.*, 2003). As part of this pathway, a whole suite of genes responsible for DNA repair is induced to facilitate segregation of intact, repaired genomes. This includes the classical repair genes required for NER and BER. The identification of a RecA-NDp consensus sequence allowed prediction of a group of genes that is regulated in a RecA-ND fashion (Gamulin *et al.*, 2004). One of the genes that seems to be part of this pathway is *srap*, which encodes a non-essential conserved hypothetical protein of unknown function. A study by Aravind *et al*., (2013) predicted SRAP to be part of the SOS response, the classical DNA damage pathway, and functions by binding DNA, potentially facilitated by autocatalytic cleavage. Unlike the prediction by Aravind *et al.*, there is no identifiable SOS-box, however the RecA-NDp upstream of Mtb *srap* differs by two nucleotides from the conserved consensus sequence (Gamulin *et al.*, 2004). We were also able to identify a sequence resembling RecA-NDp upstream of Msm *srap* which differs from the consensus by only one nucleotide. It is expected that these deviations from the consensus sequence does not affect normal recognition and binding since DNA damage induction of *srap* is observed. Further studies with regards to the importance of the different residues in the RecA-NDp recognition and binding is required to determine functionality of *srap* sites as was done for LexA recognition and binding of the SOS-box (Davis *et al.*, 2002). In addition to *srap* not having a SOS-box, Mtb and Msm *srap* also do not seem to be associated with any damage inducible operon, in fact it appears not to be operonic (Operon correlation browser: http://tuberculosis.bu.edu/tbdb_sysbio/operon/Rv3226c.html). In both Msm and Mtb, *srap* is located adjacent to *aroA* which forms part of the shikimate pathway and biosynthesis of aromatic amino acids. In Msm, there are three transferase encoding genes in the immediate environment. Mtb *srap* is located adjacent to an acyl transferase encoding gene and is four genes upstream of *sigH* which encodes the alternative σ factor for the oxidative stress response. Further upstream is *desA3*, which encodes a possible linoleoyl-CoA desaturase, involved in lipid metabolism. The genomic context of *srap* seems to differ significantly from the other bacterial species studied in Aravind *et al.*

The upregulation of *srap* following MMC treatment (Namouchi *et al*., 2016; Rand *et al*., 2003) in Mtb prompted us to confirm upregulation of Msm *srap* upon MMC induced damage. Absolute quantification of RNA expression levels was determined using ddPCR which allowed calculation of expression levels by determining the ratio of region of interest to reference gene concentrations (Mazaika *et al.*, 2014), calculated by simplex or duplex reactions (Morisset *et al*., 2013). Cells were treated for 1 h since maximum *recA* expression was observed after 1 h of MMC treatment in Msm (Boshoff *et al*., 2003). The induction level of *recA* in our study, ~13 fold, was similar although slightly lower than was previously determined using qRT-PCR, ~18-20 fold (Boshoff *et al.*, 2003). This difference could be due to the different methods of expression quantification. However, an induction level similar to what was previously reported for Mtb *srap* (~6.1 fold) as well as RecA (~11.4 fold) using a microarray was seen in this study (Rand *et al*., 2003).

With SRAP predicted to be a DNA response protein, the first step was to conduct analyses that are routinely applied to test for a role in DNA damage response. From the genomic context study, it was evident that *srap* co-localised with DinB, the major TLS polymerase in *E. coli*, and DnaE2, the TLS polymerase in *Sphingobium* spp. and *Nitrobacter* spp. (Aravind *et al*., 2013). Even though *srap* does not occur in an operon with damage response genes in mycobacteria, it remains possible that it can be co-regulated with other damage inducible genes. With DnaE2 being the major error-prone polymerase in Msm and Mtb, even though *srap* did not co-localise with DnaE2 it raised the question whether SRAP could play a role in induced mutagenesis. UV mutagenesis analyses were conducted to address this question. A deletion strain of DnaE2 was used as a positive control, previously reported to eliminate induced mutagenesis and show greatly reduced levels of Rif^R acquisition (Boshoff *et al.*, 2003). However, both Δmsm*srap* and Δmtb*srap* phenocopied wildtype. Even when *srap* was deleted in the ΔY background, where the Pol IV homologs *dinB1*, *dinB2* and *dinB3* were eliminated in Msm mc^2155, there was no notable change in the Rif^R frequency. Therefore, it does not seem that SRAP plays a non-redundant role (if any) in induced mutagenesis.

The deletion mutants showed no sensitivity to any of the DNA damaging agents tested. A suite of genotoxins, causing a range of different lesions, were tested. Pmb is a metabolic generator of superoxide, causing oxidative DNA damage (Mosel *et al.*,

2013), MMC forms DNA crosslinks (Matsumoto *et al.*, 1989), Ofx is a second generation fluoroquinolone and DNA gyrase inhibitor (Drlica & Zhao, 1997) and lastly, Nfz is a nitrofuran that forms N^2-dG adducts (Jarosz *et al.*, 2006). The results of this study argued against a major role for SRAP in the DNA damage response of either Msm or Mtb. However, the assays applied here might not be sensitive or specific enough to pick up any subtle sensitivities of the deletion strains. Another limitation of this type of assay is that it is limited to detecting only non-redundant, unique functions. It is therefore possible that functional redundancy may have masked any effect of loss of SRAP function that these assays were designed to evaluate. .

Ongoing work in our laboratory has given key insights into mutasome recruitment and localisation using fluorescent microscopy. In an attempt to visualise SRAP expression and localisation, we used a fusion protein, linking *egfp* to *srap* at the C-terminus via a short linker sequence, maintaining the native *srap* promoter. The *egfp* gene was cloned at the C-terminus of SRAP to avoid loss during possible N-terminal cleavage required for enzymatic activation as discussed later (Aravind *et al.*, 2013). As a positive control, pSOS-*egfp* was included, where *egfp* is expressed under control of the damage inducible promoter of *imuA'*. LexA-regulated ImuA' is responsible for DNA damage induced mutagenesis and damage tolerance in *Caulobacter crescentus* (Erill *et al.*, 2006). This was later also confirmed to be true in mycobacteria (Warner *et al*., 2010). After MMC exposure, eGFP is clearly visible in the pSOS-*egfp* strain, however, no fluorescence could be consistently detected in the SSE strain: in some instances there was a slightly higher intensity in MMC-treated SSE cells, but in other instances, the converse was observed with the untreated SSE control showing higher fluorescence. It is most likely that this inconsistent fluorescence is the result of auto-fluorescence linked to the metabolic and redox state of the cell. Auto-fluorescence is intrinsic fluorescence, visible without staining. Mtb is known to have a high level of auto-fluorescence, especially in the cyan range of the visible spectrum possibly due to the fluorescent properties of the F_{420} coenzyme, further increased by an increase in temperature and basic pH (Patiño *et al*., 2008). Low level auto-fluorescence in the green spectrum is observed similar to the cyan spectrum. Alternatively, the lack of consistent fluorescence could also be due to incorrect folding of the fusion protein disrupting eGFP fluorescence and possible SRAP function. However, the inability to produce a

phenotype in the DNA damage sensitivity assays made it impossible to confirm functionality of the fusion protein.

Experimental evidence suggests a role for SRAP in the oxidative stress response of mycobacteria, possibly through oxidation damage of DNA. Upregulation of *srap* has been observed upon H_2O_2 treatment in both Msm (Li *et al*., 2015) and Mtb (Voskuil *et al.*, 2011) and modest induction of srap was also seen in Mtb following macrophages infection (Schnappinger *et al*., 2003). In further support of a role for SRAP in oxidative stress, the human SRAP homolog, C3Orf37, was found to bind oxidised derivatives of 5-methyl cytosine in embryonic cells (Spruijt *et al.*, 2013). However, the Msm deletion mutant showed no sensitivity to H_2O_2 when spotted on standard media (no external catalase) containing H_2O_2. In addition to this, the MIC analysis also revealed no difference in sensitivity between wildtype and Δmsm*srap* with both strains having an MIC_{90} of 0.245 mM – 0.49 mM in LB. The MIC observed in this study is higher than previously reported (0.039 mM) (Li *et al*., 2014). A possible reason for this difference could be the difference in growth media used for the MIC analyses: LB was used in this study whereas 7H9 supplemented with 10% ADS (Albumin-Dextrose-Saline) in the previous study. It is also not clear if the MIC reported in the previous study is indeed MIC_{90}, as reported in this study.

In an attempt to investigate expression of SRAP under oxidative stress, the SSE strain was used. SSE and the parental wildtype control were treated with H_2O_2 to visualise expression. Unexpectedly, fluorescence was not only observed for the strain expression eGFP: instead, both strains showed a significant difference in fluorescence between treated and untreated samples after 1 h and 3 h treatments. Again, catalase was omitted to only allow neutralisation of H_2O_2 by natural occurring catalases, maximising the effect of H_2O_2 treatment. The high level of auto-fluorescence could possibly be due to oxidation of flavins which emit fluorescence at 515-520 nm, overlapping with GFP (Barile *et al*., 2013). Flavin plays a pivotal role in cellular processes. Many flavin containing enzymes have been identified in Mtb for example, 20 genes were identified to encode cytochrome P450 enzymes, responsible for metabolising xenobiotics (Ouellet *et al*., 2010). However, it has been reported that cytochrome P450 emits auto-fluorescence in the cyan (Wong *et al*., 2016) and not the green spectrum. Another system, complementary to cytochrome P450, is flavin containing monooxygenases

(FMO), *MSMEG_1682* and *MSMEG_5022* (Roxas & Li, 2009). An important difference between FMO and cytochrome P450 is the manner in which their substrates get oxidised. Unlike cytochrome P450 which uses an iron containing cofactor, FMO uses flavin adenine dinucleotide (FAD) to oxidise its substrate (Cashman, 2005). The location of FMO, with oxidised flavins, could be the source of foci formation upon oxidative stress. Perhaps the bacterial flavo-proteins like eukaryotic peroxisomal oxidases co-localise in microcompartments similar to peroxisomes (Chen & Silver, 2012). This was confirmed to be a H_2O_2 specific phenomenon since no such auto-fluorescence was detected in cells treated with MMC. Also, no auto-fluorescence was observed in the red range. Auto-fluorescence in the cyan range was very inconsistent and generally did not seem to differ between treated and untreated samples. More research will have to be conducted to determine the source of fluorescence foci formation in the green spectrum following oxidative stress.

The absence of any microbiological phenotype of the deletion mutants under tested conditions directed us to attempt to understand the structure of the protein, which could give some insight into the function. Protein modelling is not an exact reflection of protein folding but the accuracy of this process has been improving (Zhang *et al*., 2010). For this reason protein models were generated for both Msm and Mtb SRAP as a point of entry to attempt solving the structure and possibly gain some insight into the function. The two mycobacterial proteins were modelled based on the crystal structure of the *E. coli* homolog, YedK (PDB: 2icu). From these models, a putative catalytic triad consisting of C2, E119 and H183 in Mtb and C2, E124 and H190 in Msm was evident as previously predicted. Furthermore, the two conserved arginines, R107 and R192 in Msm and R103 and R186 in Mtb were located close to the triad, predicted to facilitate contact with DNA. However, these are only predictions and might not be accurate and require further investigation. A conserved asparagine is present in both homologs however, Mtb and Msm homologs have an aspartic acid in place of the second conserved glutamic acid. Nonetheless, the close structural similarity between aspartic acid and glutamic acid might still allow for normal folding and functioning.

As the first step to attempt to solve the structure we managed to express soluble Msm SRAP. Using the hexa His-tag we were able to obtain protein of high purity using nickel affinity chromatography. The native conformation of the pure protein was studied using

SEC, which is non-reducing. Three peaks were consistently observed which correspond to the estimated sizes of a monomer, trimer and nonamer, respectively. The reproducibility of the three peaks could be indicative of a state of equilibrium which was confirmed by seeing break down of the large nonameric peak to monomers and trimers after overnight incubation. The oligomer assembly seems to form transient protein complexes which could depend on different factors such as pH, temperature and salt concentration (Park & Raines, 2000). Regardless, under the tested conditions the putative nonamer seems to be the preferred state. Such large complexes have not been reported for any of the homologs but the *Agrobacterium tumefaciens* homolog (PDB: 2aeg) reportedly forms trimers. Trimers might be preferred in a physiological system but under lab conditions three trimers associate to form a nonamer. However, the nonamer is the only complex large enough to view using TEM and cryo-EM in future.

The publically available crystal structures of SRAP homologs (PDB: 1zn6, 2f2o, 2aeg and 2icu) were solved as part of structural genomics efforts with no crystallisation conditions published. The samples representing the trimeric peak were used to set up crystal trials since the concentration was higher than the monomer and SRAP trimers have previously been observed in *A. tumefaciens*. Crystallography trials can be a lengthy process and to try speed up the process the Hampton screen was used to allow for rapid screening of crystallisation conditions. Vapour diffusion is the most commonly used method for crystallisation and allows subtle changes in protein and precipitant concentration which allows growth of well-organised crystals. A variety of factors influence the formation of protein crystals including protein purity, pH, concentration and precipitant (Chayen & Saridakis, 2008). Crystal growth occurs in the nucleation zone that lies between the clear and the precipitation zones. Protein was diluted into different volumes of water for drop formation which allowed us to identify concentrations at which the protein precipitates out and remains soluble for each precipitant. Even after an attempt to optimise crystallisation conditions for precipitants that showed precipitate at the highest protein concentration and remained clear at the lowest concentration no crystals formed after two months and was subsequently abandoned for the purposes of this study. If the trimer is not the natural conformation then regular packing of proteins, required for crystal formation, might not be established. Optimisation of purification and SEC to allow for higher concentrations of

monomer should be pursued to also attempt crystal trials with the monomer. The TEM images showed a high level of homogeneity and could be possible candidate for cryo-EM but was also not further pursued in this study due to time constraints. However, crystallography and cryo-EM could be pursued together to solve the structure in future.

The predicted thiol-peptidase activity of SRAP could not be confirmed by means of zymography. The putative peptidase activity is based on the presence of three highly conserved residues that are positioned in such a way as to resemble the catalytic triad of structurally unrelated cysteine proteases (Aravind *et al*., 2013). The absence of visible proteolysis in the zymography gels indicated a lack of peptidase activity under the conditions tested. However, only two general substrates were tested which might not be optimal substrates for SRAP. An alternative explanation for the lack of proteolysis is that SRAP is produced in an inactive form and requires activation. Since it is expected that N-terminal cleavage is required for activation the lack of N-terminal cleavage at a range of different temperatures is likely to be the reason we could not detect any proteolysis. However, MMP1, used as a positive control, is synthesised as an inactive pro-peptide and is activated by a "cysteine switch" caused by SDS (Springman *et al.*, 1990). It is possible that SRAP requires a specific set of conditions for N-terminal cleavage and activation, more complicated than is required for MMP1 activation. More research is required to establish whether this protein is indeed a peptidase and which specific conditions are required for activation, a pull down assay could aid in identifying possible substrates.

The DNA damage response is of central importance in preserving genome integrity. The coordination of DNA repair, replication and recombination is required which in turn is linked to pathways directing cell division and proliferation. These systems are highly integrated, allowing one enzyme/pathway to compensate for the loss of another. A recent study looked at the interaction between functional pathways instead of studying contributions of individual genes in isolation which showed a highly adaptive response (Kumar *et al*., 2016). This stresses the importance of large scale interaction network studies and looking at the response as a whole rather than a gene in isolation to capture functional dependencies and redundancies during DNA damage response. This could explain the absence of a DNA damage sensitivity phenotype, not because SRAP plays no role in DNA damage response but rather because of functional

redundancy where another protein can compensate for the loss of SRAP. Future studies will have to be aimed at studying networks as a whole instead of studying SRAP in isolation. This will also aid in addressing the roles of numerous protein with unknown function forming part of the DNA damage response. Furthermore, the availability of protein structure could give great insight into the function of the protein and must be further pursued.

Chapter 4: DnaE2 is not required for stress-induced adaptive mutagenesis in *Mycobacterium tuberculosis*

4.1 Introduction

Adaptive mutation is described as an increased mutation rate in a non-proliferating bacterial population upon encountering certain environmental stresses; for example, nutritional deprivation (Maisnier-Patin and Roth, 2015; Rosenberg, 2001). In general, the concept denotes the acquisition of advantageous mutations that allow the cell to overcome the selective pressure. Adaptive mutation has been extensively studied and characterised in *E.coli*, most notably in studies which looked at the rate at which a *lacZ* +1 frameshift mutant, incapable of utilising lactose as sole carbon source, reverted to wildtype due to a genetic change (Cairns & Foster, 1991). One common source of genetic change is through induction of the DNA damage - or SOS - response which includes upregulation of specialist TLS polymerases. The activity of these polymerases, which have also been described as "error prone", can result in hypermutation (Rosenberg, 2001). Although adaptive, these mutations are not targeted at the gene conferring a selective advantage since unselected mutations also occur randomly throughout the genome (Cairns & Foster, 1991; Rosenberg, 2001).

The major SOS-inducible TLS polymerase in mycobacteria is DnaE2. Unlike the error-prone polymerases in *E. coli* which belong to the Y-family polymerases, DnaE2 belongs to the C-family. Most C-family polymerases are essential for DNA replication but, in Mtb it was shown that DnaE1 was the major replicative polymerase whereas DnaE2 was non-essential (Boshoff *et al*., 2003). A deletion mutant of Mtb DnaE2 was further linked to hypersensitivity to DNA damaging agents as well as reduced drug resistance acquisition (Boshoff *et al.*, 2003). It was later shown that altering the highly conserved active site residues ^{441}DID^{443} to ^{441}AIA^{443} eliminated UV-induced mutagenesis in Msm similar to the *dnaE2* deletion mutant (Warner *et al*., 2010). Studies of DnaE2 function in mycobacteria have been mostly limited to the acquisition of *rpoB* mutations upon UV exposure. In these UV mutagenesis assays, the UV-treated cells are exposed to Rif, a lethal pressure, allowing only those cells with *rpoB* mutations to survive and proliferate - and to be quantified for mutation frequency calculations. In order to study the role of DnaE2 in adaptive mutagenesis in bacilli exposed to a nonlethal stress (here, we focus on nutrient deprivation), we aimed to exploit the vitamin-B_{12}-sensitive phenotype of a Δ*metH* mutant (Warner *et al*., 2007).

The genome of Mtb encodes two methionine synthases, MetH and MetE (Cole *et al.*, 1998), which catalyse the final methyl transfer in *de novo* methionine synthesis, converting homocysteine to methionine. MetH is dependent on a B_{12}-derived cofactor for catalytic function, whereas MetE is B_{12}-independent (Warner *et al.*, 2007). It is common in organisms with B_{12}-dependent and B_{12}-independent enzymes that the activity of the B_{12}-independent enzyme is regulated by a riboswitch (Vitreschak *et al.*, 2003). A riboswitch is a highly structured region in the mRNA that changes in response to ligand binding, regulating gene expression (Blount & Breaker, 2006). This is true also for *metE*: as a result, expression of *metE* is repressed in the presence of B_{12}, in turn causing effective methionine starvation in a Δ*metH* mutant exposed to a high-dose B_{12} (Warner *et al*., 2007).

A previous study in my host laboratory identified a major deletion (1196 bp) in the *metH* locus of the clinical strain, CDC1551 (Warner *et al*., 2007). This deletion, which disrupts the cobalamin (B_{12}) - and *S*-adenosyl-L-methionine- binding domains of MetH, was subsequently shown to eliminate methionine synthase function. When plated on solid media containing B_{12}, a corresponding Δ*metH*(B) mutant of Mtb H37Rv showed a $3\log_{10}$ reduction in colony forming units (CFU). This phenotype – profound B_{12}-sensitivity - appeared to result from methionine starvation: B_{12}-mediated repression of *metE* in a mutant lacking functional MetH effectively eliminated all methionine synthesis. Isolates that grew on the B_{12}-containing plates were referred to as "B_{12}-resistant suppressor mutants" and were observed at a frequency of $\sim10^{-3}$. The genetic heritability of the suppressor phenotype led to the hypothesis that mutations in the riboswitch might prevent B_{12}-mediated repression, allowing growth. This was confirmed when mutations in the riboswitch were identified in ~20% of suppressor mutants. Furthermore, this study illustrated transport of exogenous B_{12}, despite the absence of a B_{12} specific transporter, as well as the ability of Mtb to convert exogenous B_{12} to the cofactor required for MetH function (Warner *et al.*, 2007). This emphasised the possibility that mutations in genes contributing to B_{12} transport and assimilation could also contribute to the suppressor phenotype.

In a subsequent study, random mutagenesis was used to identify genes required for exogenous B_{12} utilisation. The B_{12} sensitivity phenotype of an unmarked Δ*metH* mutant, analogous to Δ*metH*(B), was exploited to identify mutations allowing the

suppressor phenotype (Gopinath *et al*., 2013). Again, mutations in the riboswitch region upstream of *metE* were observed in ~20% of the suppressor mutants. In addition to these, 40% of the suppressors had mutations in *Rv1819c*, encoding a predicted ABC-transporter and homolog of BacA. However, unlike BacA, Rv1819c combines a predicted transmembrane domain and nucleotide-binding domain (NBD) in a single polypeptide, and so appears to be the only corrinoid transporter during *in vitro* growth of Mtb. No substrate binding proteins (SBP), usually associated with ABC transporters, could be identified which could indicate that transport relies solely on the transporter, or that multiple different SBPs could play a role during transport. In addition to *Rv1819c*, a few mutations were also observed in *Rv1314c* encoding the sole active adenosyltransferase required for B_{12} assimilation during *in vitro* growth of Mtb (Gopinath *et al.*, 2013).

In the study by Warner *et al*. (2007), B_{12} suppressor mutants were observed at a frequency markedly higher (10^{-3}) than can be explained by spontaneous mutagenesis (10^{-8}) alone. To investigate this phenomenon further, we aimed to determine the contribution of DnaE2 to the appearance of these suppressor mutants by eliminating the polymerase activity of the DnaE2 in the Δ*metH* background. We hypothesised that DnaE2 contributed to the high frequency of suppressor mutant appearance in cells under nutrient (methionine) stress.

4.2 Aim and objectives

The roles of various TLS polymerases in adaptive mutagenesis in *E. coli* have been studied extensively using a *lacZ* reversion assay. Critically, this system imposes a non-lethal stress in which cells incapable of utilising lactose are grown on media with lactose as sole carbon source, starving cells that have not acquired an adaptive mutation. In a similar way, our system causes methionine deprivation and starvation when a MetH deletion strain is exposed to high concentrations of B_{12}. The main aim of this study, therefore, was to determine the contribution of DnaE2 to the high frequency appearance of suppressor mutants. The specific objectives include:

I. evaluate the B_{12} sensitivity of a catalytically dead *dnaE2*^{AIA} mutant in the Δ*metH* background;
II. identify mutations in three target genes of suppressor mutants that could account for the suppressor phenotype.

4.3 Results

4.3.1 Inactivating the DNA-damage induced TLS polymerase, DnaE2

The high degree of conservation between Msm and Mtb DnaE2 allowed us to follow a strategy similar to that used previously for Msm (Warner *et al*., 2010), to design and synthesise a construct to generate Mtb DnaE2^{AIA} by mutating ^{339}DID^{341} to ^{339}AIA^{341} (Figure 4.1A and B). The presence of the mutant allele was confirmed in both the SCO and DCOs by restriction digests of PCR products using dnaE2scF and dnaE2scR (Table 2.3, Figure 4.1C). The PCR products were also sequenced to confirm the presence of the desired mutations.

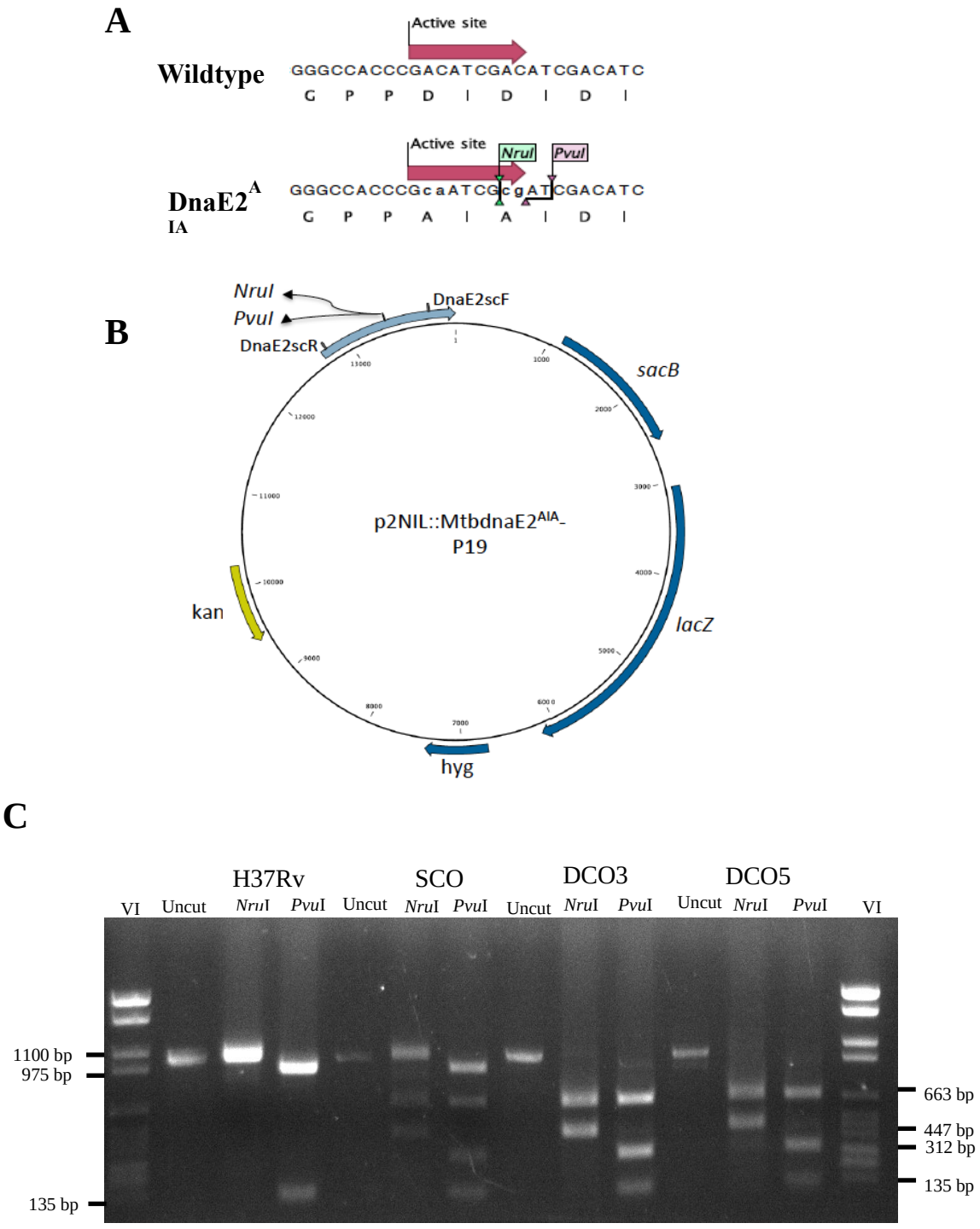


Figure 4.1. Construction and genetic confirmation of the *dnaE2*^{AIA}Δ*metH* strains. **A.** DnaE2 active site of wildtype (H37Rv) and mutations introduced to inactivate polymerase activity in DnaE2^{AIA} as well as introduce *Nru*I and *Pvu*I cut sites. **B.** Schematic representation of the p2NIL::mtb*dnaE2*^{AIA}-P19 suicide vector carrying the *dnaE2*^{AIA} allele. The primer binding sites of dnaE2scF and dnaE2R, used to amplify the active site with the *Nru*I and *Pvu*I sites for screening, are also indicated. **C.** Restriction digest screening of PCR products. Wildtype contains no *Nru*I cut site, and so remains uncut (1110 bp), whereas the mutant allele contains an *Nru*I site, resulting in two fragments (447 bp and 663 bp). *Pvu*I cuts the wildtype product once (974 bp and 135 bp) and the mutant allele twice (663 bp, 312 bp and 135 bp).

The loss of DnaE2 activity was confirmed phenotypically, first by means of the established induced mutagenesis assay in which UV-exposed bacilli were plated on Rif containing media in order to score the number of Rif^R mutants. The Rif^R mutation frequency was calculated based on the number of Rif^R colonies that formed post-UV treatment. The constructed SCO is effectively a *dnaE2* knock out because only part of the gene was included in the design of the mutant allele and therefore we expected loss of DnaE2 function in all the generated strains, including the partial merodiploid intermediate. As expected, the *dnaE2* point mutants phenocopied the previously characterised *dnaE2* deletion strain in which the frequency of induced mutagenesis to Rif^R was $2\log_{10}$-fold lower than that of the parental wildtype strain (Figure 4.2A). Secondly, MMC damage survival assays were conducted by spotting serial 10-fold dilutions of log-phase cultures on MMC containing plates. Again, the SCO and DCOs phenocopied Δ*dnaE2*, showing a log-fold increase in sensitivity compared to wildtype and Δ*metH* at both concentrations of MMC (Figure 4.2B).

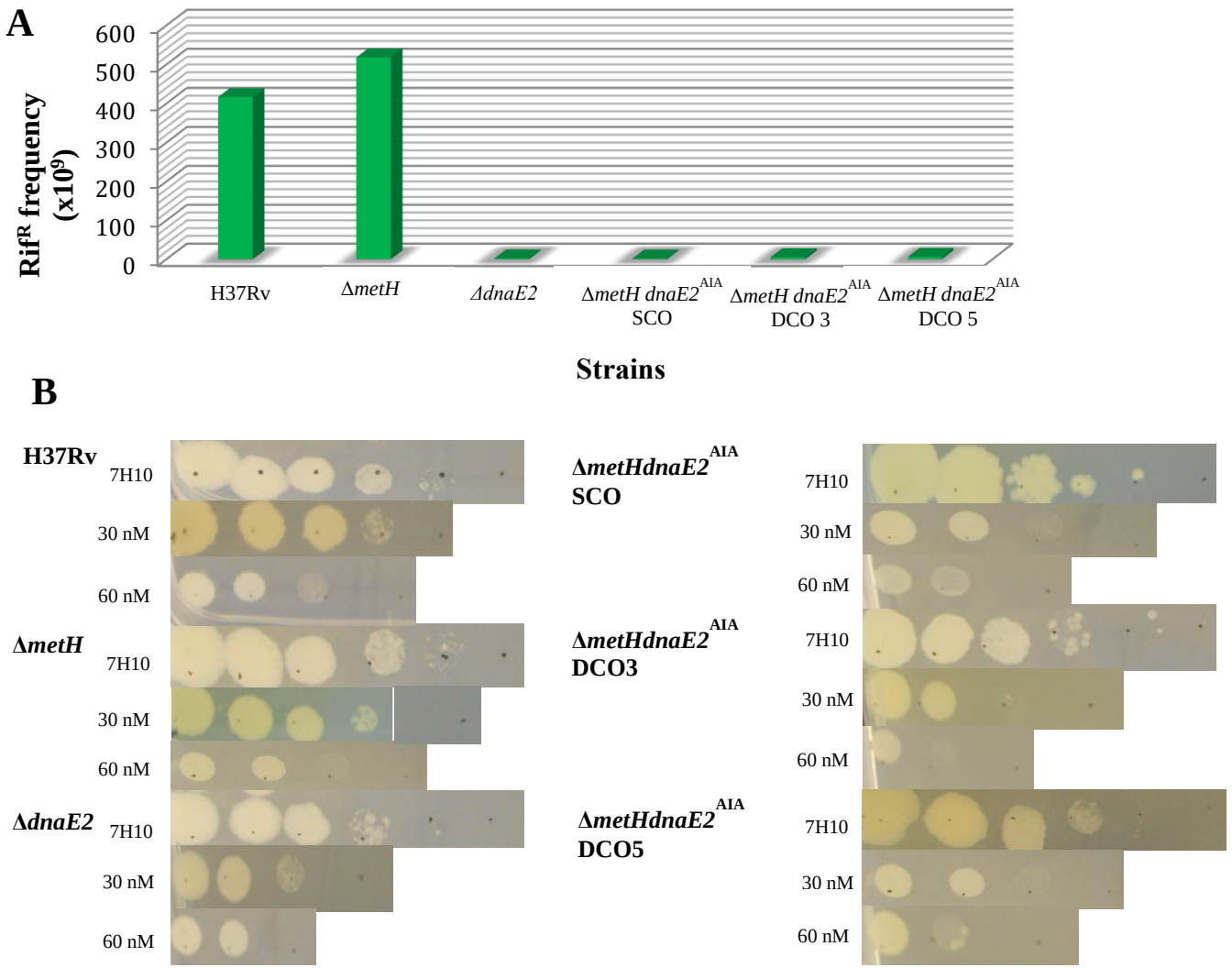

A
600
500
400
300
200
100
0
Rif^R frequency (x10^9)
H37Rv
ΔmetH
ΔdnaE2
ΔmetH dnaE2^AIA SCO
ΔmetH dnaE2^AIA DCO 3
ΔmetH dnaE2^AIA DCO 5
Strains
B
H37Rv
7H10
30 nM
60 nM
ΔmetH
7H10
30 nM
60 nM
ΔdnaE2
7H10
30 nM
60 nM
ΔmetHdnaE2^AIA SCO
7H10
30 nM
60 nM
ΔmetHdnaE2^AIA DCO3
7H10
30 nM
60 nM
ΔmetHdnaE2^AIA DCO5
7H10
30 nM
60 nM

Figure 4.2. Phenotypic confirmation of loss of DnaE2 activity in double mutants. A. UV mutagenesis confirmed a two log-fold reduction in Rif^R acquisition in Δ*metHdnaE2*^{AIA}, as previously reported for Δ*dnaE2* (Boshoff *et al*., 2003). **B.** MMC survival assays showed an increase in sensitivity to MMC similar to Δ*dnaE2*.

4.1.1 B_{12} and Cbi sensitivity of Δ*metHdnaE2*^{AIA}

CFU counts of H37Rv, Δ*metH* and Δ*metHdnaE2*^{AIA} SCO and DCOs were obtained by plating serial log-fold dilutions of cultures on standard 7H10 medium supplemented with 7.38 µM B_{12} and 10 µM Cbi, respectively (Figure 4.3). As expected, the parental H37Rv showed no sensitivity to either of the corrinoid compounds whereas Δ*metH* exhibited a 4log_{10} reduction in CFU on both B_{12} and Cbi. However, inactivation of DnaE2 had no effect on the B12 or Cbi sensitivity of Mtb in a metH mutant background as evidenced by the fact that the SCO and DCO mutants were indistinguishable from the parental Δ*metH* mutant.

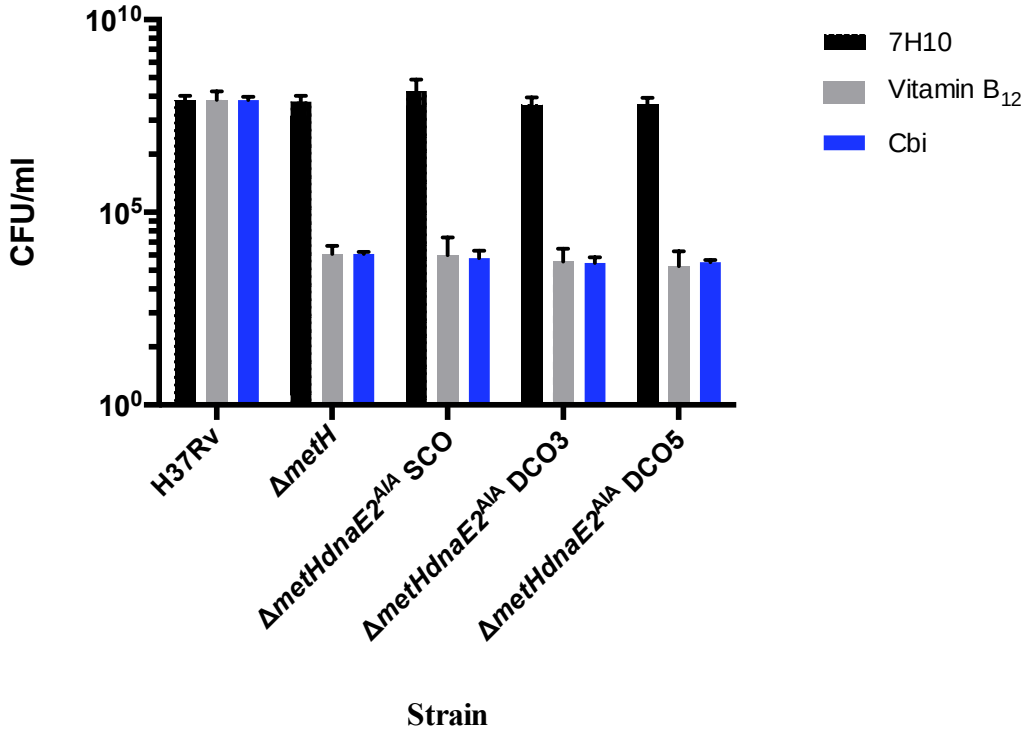


Figure 4.3. B_{12} and Cbi sensitivity of wildtype and *dnaE2*^{AIA} strains. The sensitivity of wildtype (H37Rv), parental Δ*metH*, and Δ*metHdnaE2*^{AIA} SCO and DCO mutants were evaluated by plating log-fold dilutions of exponential-phased cultures on 7H10 agar (black bars) or 7H10 containing 7.38 µM B_{12} (grey bars) or 10 µM Cbi (blue bars), respectively.

4.1.2 Identification of mutations responsible for resistance acquisition

The genetic heritability of the B_{12} suppressor phenotype was confirmed in a subset of isolates by growing them in the absence of B_{12} and then re-plating on media containing vitamin B12 or Cbi (Figure 4.4).

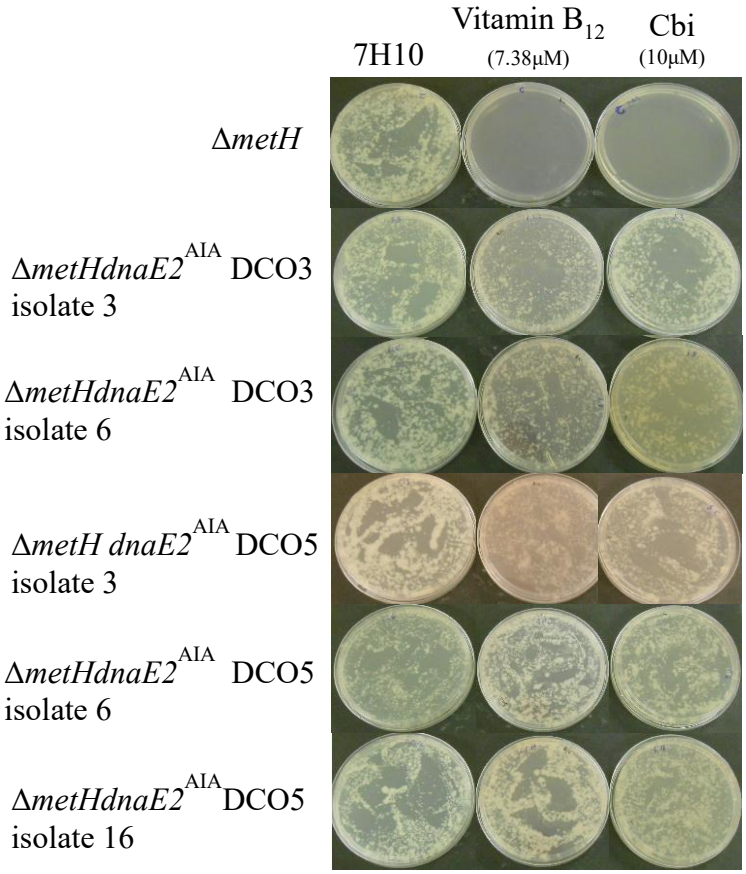


Figure 4.4. Heritability of B_{12} and Cbi insensitivity. Randomly selected suppressor mutants were grown in standard 7H9, and log-fold dilutions were then plated on 7H10 supplemented with B_{12} or Cbi, respectively.

Three Mtb genomic regions have previously been associated with the acquisition of the B_{12} suppressor phenotype: the predicted riboswitch region upstream of *metE*; the adenosyltransferase, *Rv1314c*; and the B_{12} transporter, *Rv1819c* (Gopinath *et al*., 2013; Warner *et al*., 2007). A subset of suppressor colonies on B_{12} containing plates from each strain was randomly selected, and each of the three genomic regions PCR amplified and sequenced. Ten suppressor isolates from one technical repeat of Δ*metH* and the SCO were included, as well as 30 suppressor isolates from DCO3 (20 from repeat one and 10 from a second repeat) and 20 suppressor isolates from DCO5 (10 from repeat one and 10 from repeat two). Suppressor mutants from the DCOs were labelled based on their parental strain (D3 or D5), followed by the technical repeat (one or two) and, lastly, its isolate number. Suppressors from the Δ*metH* and the SCO were named either M (Δ*metH*) or S (SCO) followed by the isolate number. The number of isolates sequenced and the frequency and nature of identified mutations is indicated in Table 4.1. Mutations included SNPs, single nucleotide insertions and deletions (indels), and also complex nucleotide deletions. There were only two suppressor isolates (D3.1.10 and D3.1.16) for which no mutations could be identified. Some isolates harboured more than one mutation; therefore the total number of mutations was higher than the number of isolates sequenced. For example, M10 and D5.2.3 both contained two mutations, an insertion and a SNP, in the riboswitch region.

Table 4.1 Summary of mutations identified in the B_{12}-resistant mutants.

Gene	Strain	Number of isolates sequenced	Number of isolates with mutations	Number of mutations[a]	Type of mutation		
					SNPs	**Insertions**	**Deletions**
riboswitch upstream of *metE*	Δ*metH*	10	7	8	6	1	1
	Δ*metHdnaE2*^{AIA} SCO	10	8	8	7	0	1
	Δ*metHdnaE2*^{AIA}DCO3	30	19	19	9	9	1
	Δ*metHdnaE2*^{AIA} DCO5	20	9	10	9	1	0
Total		**70**	**43**	**45**	**31**	**11**	**3**
Rv1314c	Δ*metH*	10	0	0	0	0	0
	Δ*metHdnaE2*^{AIA} SCO	10	0	0	0	0	0
	Δ*metHdnaE2*^{AIA} DCO3	30	2	2	2	0	0
	Δ*metHdnaE2*^{AIA} DCO5	20	2	2	2	0	0
Total		**70**	**4**	**4**	**4**	**0**	**0**
Rv1819c	Δ*metH*	10	1	1	1	0	0
	Δ*metHdnaE2*^{AIA} SCO	10	1	1	0	0	1
	Δ*metHdnaE2*^{AIA} DCO3	29	7	7	4	2	1
	Δ*metHdnaE2*^{AIA} DCO5	19	8	8	3	1	4
Total		**68**	**17**	**17**	**8**	**3**	**6**

[a]The total number of mutations across all isolates combined

The majority of mutations were found in the riboswitch region upstream of *metE*: 45 mutations, comprising 31 substitutions, 11 insertions, and three deletions (Figure 4.5). Only three mutations occurred in the conserved B_{12}-box, with the vast majority located outside the B_{12}-box. Two mutations occurred in multiple isolates in the same technical repeat; that is, T -> C at -203 in four isolates from the same DCO as well as one isolate from another DCO (Table 4.2). The T-insertion at -206 was another example seen in multiple isolates from a single technical repeat. Most of the mutations in the riboswitch region only affected a single nucleotide; however, one isolate, S8, had a complex deletion eliminating 53 bp of the riboswitch region upstream of the B_{12}-box. The scale of this deletion was likely to cause a major change in the structure and folding of the riboswitch, eliminating B_{12}-mediated regulation.

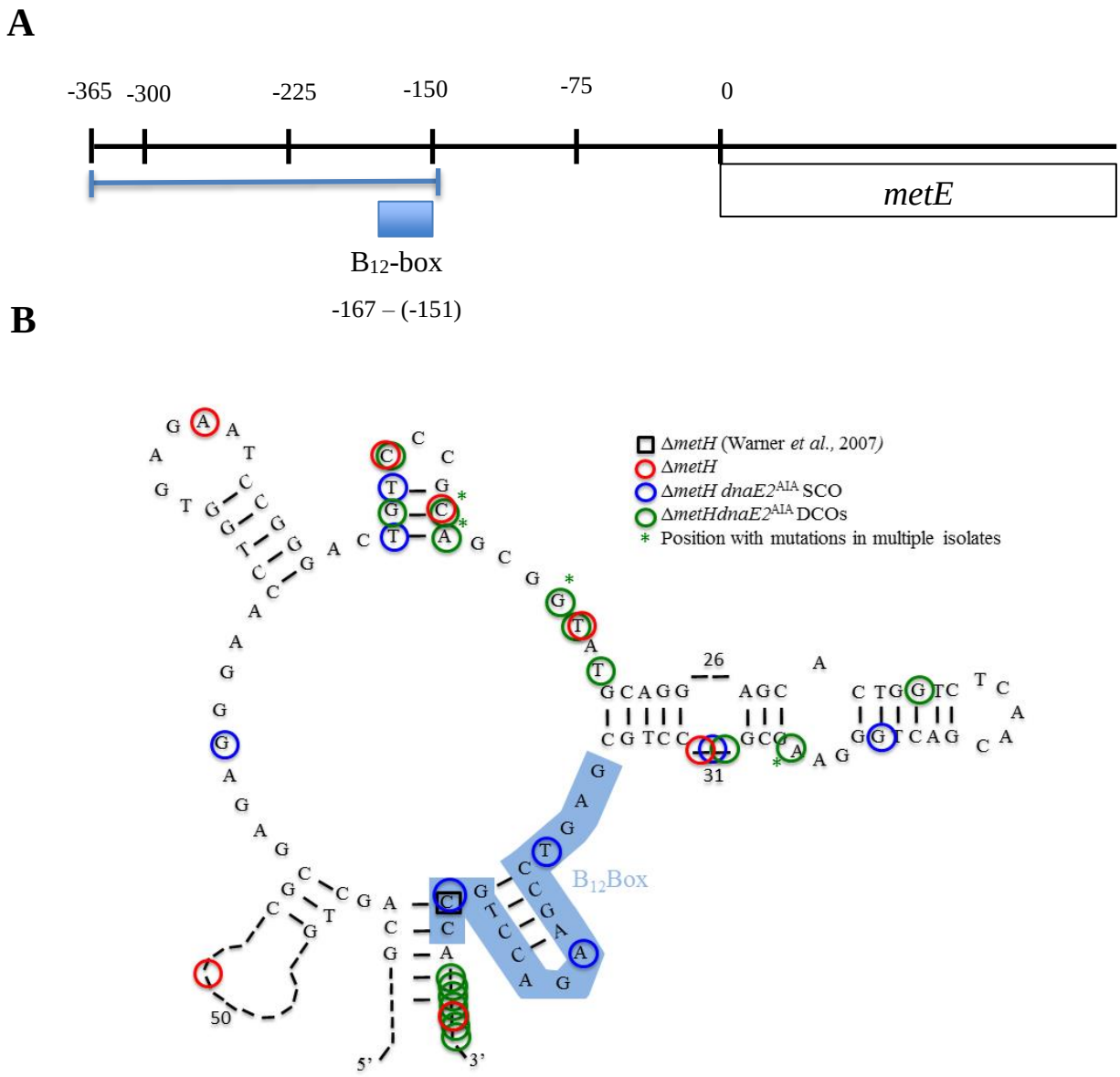


Figure 4.5. Schematic representation of the riboswitch upstream of *metE* with the conserved B_{12}-box. A. Illustration of the region upstream of *metE* and the position of the conserved B_{12}-box (blue box) in the riboswitch (blue line) in relation to *metE*. **B.** The Mtb riboswitch region upstream of *metE* as adapted from Warner *et al.*, 2007. Mutations identified in the previous study are indicated with black circles. The red (Δ*metH*), blue (Δ*metHdnaE2*^{AIA} SCO) and green (Δ*metHdnaE2*^{AIA} DCOs) circles indicate mutations identified in this study. Green circles with asterisks indicate positions where mutations were identified in more than one isolate.

Two mutations were identified in the riboswitch regions of isolates M10 and D5.2.3. Interestingly, the same two mutations – a G insertion (-123) and C -> T substitution (-124) – occurred in both isolates. This makes it tempting to speculate that both mutations might be required to acquire the suppressor phenotype or that both bases together acts as a target for mutagenic machinery. Furthermore, multiple independent isolates had mutations at -274 (two isolates), -270 (three isolates) and -269 (two isolates).

Table 4.2. Mutations identified in the upstream region of *metE* containing the riboswitch.

Isolate[a]	**Mutation position**	**Base mutated**
D5.2.4	-110	C -> T
D5.2.9	-115	C -> A
D5.2.10	-122	C -> T
M10	-122	G insertion
	-123	C -> T
D5.2.3	-122	G insertion
	-123	C -> T
D3.1.12	-125	C -> T
D3.1.14	-130	C ->A
D3.1.19	-146	C -> T
S2	-152	G -> A
S7	-159	T -> C
S4	-164	A -> C
M7	-173	C -> A
S8	-227 – (-175)	53 bp deletion
D5.1.7	-203	T -> C
D3.1.1	-203	T -> C
D3.1.2	-203	T -> C
D3.1.3	-203	T -> C
D3.1.4	-203	T -> C
D3.1.5	-206	T insertion
D3.1.7	-206	T insertion

D3.1.8	-206	T insertion
D3.1.9	-206	T insertion
D3.1.11	-206	T insertion
S9	-211	C -> G
D3.1.17	-223	C -> T
D3.1.13	-262	A -> G
D5.1.10	-264	A -> G
M6	-264	A -> G
D3.1.18	-265	C -> T
D5.2.5	-265	C -> T
D3.1.20	-269	T deletion
D5.2.7	-269	T -> C
M4	-270	G -> C
D3.1.15	-270	G -> A
D3.2.7	-270	G -> A
M3	-274	G deletion
D5.1.5	-274	G -> A
S5	-275	A -> G
D3.2.1	-276	C -> T
S1	-277	A-> C
M1	-287	T -> G
S6	-301	C -> T
M8	-326	T -> C

[a] Isolates are named as follows: First letter indicate strain with M = Δ*metH*, S = Δ*metHdnaE2*^{AIA} SCO, D3 = Δ*metHdnaE2*^{AIA} DCO3 and D5 = Δ*metHdnaE2*^{AIA} DCO5. The second letter following M and S indicates isolate number whereas the second letter following strains 3 and 5 refers to the technical repeat. The third number following the technical repeat represents the isolate number e.g., 3.2.20 is isolate 20 from technical repeat of Δ*metHdnaE2*^{AIA} DCO3.

The other two genes harboured fewer mutations than the riboswitch. Four isolates were found to have non-synonymous SNPs in *Rv1314c* and a total of 17 mutations (eight non-synonymous SNPs, three insertions and six deletions) were identified in *Rv1819c*. All the mutations in these two genes cause a change in predicted protein structure either by changing the amino acid, causing a frameshift or by introducing a stop codon. Tables 4.3 and 4.4 provide detailed descriptions of the mutations identified including the position and type of mutation as well as the isolate in which it was identified.

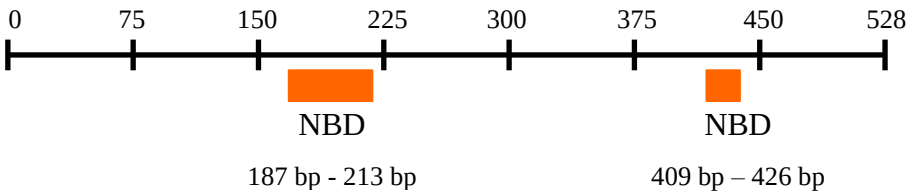


Figure 4.6. *Rv1314c* with two putative nucleotide-binding domains (NBD). A total of four mutations were identified of which two were identical. All three different mutations are non-synonymous which could alter the function of the protein and cause the suppressor phenotype.

The Rv1314c transferase has two putative NBDs, indicated in Figure 4.6. No mutations were identified in the first NBD, three mutations were identified in the region between the two NBDs and one mutation was found in the second NBD. All the mutations were non-synonymous SNPs, which could disrupt protein function (Table 4.3). Two isolates from the same repeat harboured the same mutation, and were probably siblings.

Table 4.3. Mutations identified in the *Rv1314c*, transferase encoding gene in this study.

Isolate[a]	Mutation position	Base mutated	Type of mutation
D5.2.2	233	C -> T	Non-synonymous (S -> L)
D5.2.6	233	C -> T	Non-synonymous (S -> L)
D3.2.6	389	A -> G	Non-synonymous (H -> R)
D3.2.3	419	A -> G	Non-synonymous (E -> G)

[a] Isolates are named as follows: First letter indicate strain with M = Δ*metH*, S = Δ*metHdnaE2*^{AIA} SCO, D3 = Δ*metHdnaE2*^{AIA} DCO3 and D5 = Δ*metHdnaE2*^{AIA} DCO5. The second letter following M and S indicates isolate number whereas the second letter following strains 3 and 5 refers to the technical repeat. The third number following the technical repeat represents the isolate number e.g., 3.2.20 is isolate 20 from technical repeat of Δ*metHdnaE2*^{AIA} DCO3.

Rv1819c is a single polypeptide with both transmembrane and nucleotide binding domains (Figure 4.7). All but one of the identified mutations were found in either of the two functional domains (Table 4.4). Seven of the different mutations identified were found in the transmembrane domain: three mutations were non-synonymous SNPs and four mutations were predicted to cause premature termination either by introducing a stop codon or by causing a frameshift mutation. Furthermore, all three mutations that occurred in the NBD were indels that were expected to result in a frameshift, again causing premature termination. The one mutation that occurred between the two domains was also a deletion, which was predicted to introduce a premature stop codon. All these mutations most likely lead to the synthesis of a dysfunctional protein.

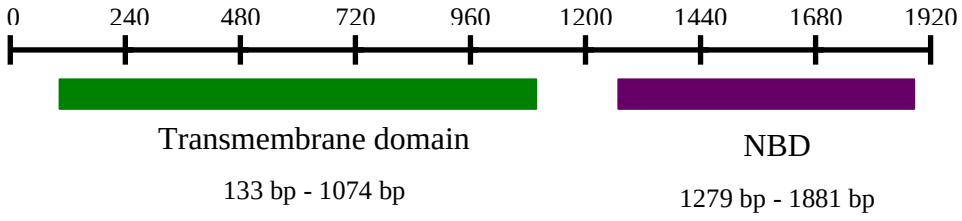


Figure 4.7. Schematic representation of *Rv1819c*. A total of 17 mutations were identified consisting of 11 different mutations. The vast majority of these mutations cause a reading frame shift which could lead to the production of a dysfunctional protein.

Table 4.4. Mutations identified in the ABC-transporter encoding gene, *Rv1819c* in the B_{12}-resistant suppressors.

Isolate[a]	Mutation position	Base mutated	Type of mutation
D5.2.1	230	T insertion	Non-synonymous (N -> I)
D5.1.3	270-272	CCA deletion	Premature termination
M2	270	C -> A	Nonsense mutation
D3.2.4	892	G -> A	Non-synonymous (E -> K)
D3.2.5	892	G -> A	Non-synonymous (E -> K)
D3.2.8	892	G -> A	Non-synonymous (E -> K)
D3.2.9	892	G -> A	Non-synonymous (E -> K)
D3.1.6	906	T insertion	Frame-shift and premature termination
D5.1.8	1013	A -> T	Non-synonymous (I -> N)
D5.1.9	1013	A -> T	Non-synonymous (I -> N)
D5.2.8	1057	C -> T	Nonsense mutation
D3.2.10	1235	G deletion	Frame-shift and premature termination
D3.2.2	1356	G insertion	Frame-shift and premature termination
S3	1453	A deletion	Frame-shift and premature termination
D5.1.2	1796	A deletion	Frame-shift and premature termination
D5.1.4	1796	A deletion	Frame-shift and premature termination
D5.1.6	1796	A deletion	Frame-shift and premature termination

[a] Isolates are named as follows: First letter indicate strain with M = Δ*metH*, S = Δ*metHdnaE2*^{AIA} SCO, D3 = Δ*metHdnaE2*^{AIA} DCO3 and D5 = Δ*metHdnaE2*^{AIA} DCO5. The second letter following M and S indicates isolate number whereas the second letter following strains 3 and 5 refers to the technical repeat. The third number following the technical repeat represents the isolate number e.g., 3.2.20 is isolate 20 from technical repeat of Δ*metHdnaE2*^{AIA} DCO3.

4.2 Discussion

The components which contribute to adaptive mutagenesis in *E. coli* are well defined. Both point mutations (substitutions and indels) and gene amplification (large scale genomic rearrangements) have been found to contribute to adaptive genetic changes using the *lac*-frameshift assay (Rosenberg, 2001). Adaptive mutagenesis is initiated when stress conditions that cause DNA damage are encountered. SOS induction, in combination with limiting mismatch repair activity, leads to a state of hypermutation, which can introduce permanent genetic changes that allow the cell to survive and proliferate. Polymerase errors are often found in simple repeat sequences due to strand slippage, making these sequences unstable and a major contributor to lac^+ reversion. Furthermore, overexpression of pol IV in *E. coli* led to an increased mutation rate, especially -1 frameshift mutations (Jarosz *et al.*, 2007; Kim *et al.*, 1997). The importance of pol IV in the acquisition of adaptive mutation was confirmed (Mckenzie & Rosenberg, 2001) despite not contributing to growth dependent mutations (Kuban *et al*., 2004) and survival of UV and oxidative damage (Kim *et al.*, 1997).

The genome of Mtb revealed two homologs belonging to the DinB subfamily of Y-family polymerase (Kana *et al*., 2010). Based on sequence similarity to the *E. coli* and *Pseudomonas aeruginosa* homologs, these were thought to be functional polymerases. However, when deletion mutants failed to show a phenotype in a suite of routinely used DNA damage sensitivity assays, it emphasised the difference in behaviour between the Mtb homologs and their *E. coli* and *P. aeruginosa* counterparts (Kana *et al*., 2010). Initially, it was thought that Mtb was incapable of damage-inducible mutagenesis since the Y-family polymerases were not SOS-inducible. This was until a polymerase belonging to the C family of polymerases, DnaE2, was found to be SOS inducible and implicated in induced mutagenesis (Boshoff *et al.*, 2003). Sensitivity to both UV and MMC was associated with loss of DnaE2 function. Although DnaE2 overexpression did not lead to an increase in spontaneous mutagenesis, the loss of DnaE2 function led to a reduction in drug resistance emergence upon DNA damage *in vitro*. In the same study, it was shown that the deletion mutant was also associated with reduced drug resistance acquisition *in vivo* in a mouse model of TB. This identified DnaE2 as a potentially important contributor to the evolution of Mtb since drug resistance in this organism appears only to be acquired through chromosomal mutations. Furthermore,

the attenuation the *dnaE2* knockout in the same model emphasised the importance of the enzyme in the survival and pathogenesis of Mtb (Boshoff *et al.*, 2003).

Studies of DnaE2-dependent mutation acquisition are mostly limited to exposure to lethal stresses such as Rif. Here, we tested the role of DnaE2 in adaptive mutagenesis in response to methionine deprivation, a non-lethal stress. No difference in B_{12} sensitivity was observed between isolates with active DnaE2 and inactive DnaE2^{AIA}, suggesting that DnaE2 is not a major factor in the high frequency occurrence of the B_{12} suppressor phenotype. One possibility is that the extent of DNA damage produced during nutrient deprivation and starvation is not as severe as the damage caused by genotoxic agents, which have been shown to induce DnaE2 mutagenesis. Furthermore, we cannot eliminate the possibility that DnaE2 contributes to the increase in mutagenesis observed to some degree with an additional factor compensating for the loss of DnaE2 function. Functional redundancy is very likely, especially due to the complex nature of the networks that underlie DNA repair (Kumar *et al.*, 2016).

A model that dismisses a role for hypermutation in adaptive mutagenesis was proposed to explain the *lacZ* frameshift reversion system in *E. coli.* This is based on the frequent duplication of the *lacZ* region that allows the acquisition of a mutation in at least one of the copies of *lacZ* prior to encountering the selective pressure. Selection then makes the mutations appear adaptive, even though the mutations were pre-existing (Maisnier-Patin & Roth, 2015). However, *lacZ* studied in this system is located on a F' plasmid (Cairns & Foster, 1991); it seems unlikely, therefore, that the genome of Mtb would accumulate enough copies of the *metE* region to allow a mutation to occur in one of the copies at the normal rate of spontaneous mutagenesis. Instead, it is possible that an alternative element(s) contributes to a state of hypermutation, for example the two homologs belonging to the DinB subfamily of polymerases. Notably, even in the alternative model which argues against a role for DinB, the overexpression of *dinB* was shown to increase adaptive mutations (Maisnier-Patin & Roth, 2015). We failed to show any decrease in adaptive mutation acquisition by eliminating DnaE2 function, which supports the notion that another element contributes to hypermutation or compensates for the loss of DnaE2. Future work should include eliminating *metH* in Δ*dinB1*Δ*dinB2* (Kana *et al*., 2010) and a triple Δ*dinB1*Δ*dinB2dnaE2*^{AIA} mutant and evaluating B_{12} sensitivity. This process can be fast tracked by using the recently

developed CRISPRi technology for conditional gene silencing in mycobacteria (Rock *et al.*, 2017).

The mutations identified in this study are similar to those identified previously (Gopinath *et al*., 2013; Warner *et al*., 2007). However, in both previous studies only two out of ten suppressor isolates were found to harbour mutations in the riboswitch region, whereas more than 60% of the suppressor isolates in the present study were found to carry mutations within the riboswitch region. This high frequency of mutation detection in the riboswitch could be due to mutations arising early on to give rise to multiple isolates with the same mutation; for example, nine isolates from the same technical repeat have the same T insertion, suggesting they are siblings. There are several isolates from independent experiments with mutations at the same position which could indicate that the region is easily accessible to the mutagenesis machinery or hard to reach for repair machinery making it more prone to errors. All the previously identified mutations in the riboswitch region were C -> T substitutions in the conserved B_{12}-box; in contrast, the vast majority of mutations identified in the current work occurred elsewhere in the riboswitch. Only three out of the 45 mutations found in the riboswitch in this study occur in the conserved B_{12}-box. It is difficult to predict the exact impact of the mutation on the folding of the riboswitch. However, mutations, although not all in the conserved B_{12} box region, could disrupt the functioning of the riboswitch, allowing expression of *metE* even in the presence of high concentrations B_{12}.

Rv1314c encodes a PduO like adenosyltransferase which was identified as the sole active adenosyltransferase during growth *in vitro,* converting cyanocobalamin to the active co-factor, adenosylcobalamin (Gopinath *et al.*, 2013). Inactivation of PduO disrupts the assimilation of B_{12}, giving rise to the suppressor phenotype. None of the unique mutations identified in this gene mapped to the two putative NBDs (Figure 4.6 and Table 4.3) but, since all of the mutations are non-synonymous, they could alter the structure of the protein. For example, a mutation at position 233 in the gene causes a change from a polar, uncharged amino acid to a hydrophobic amino acid side chain. *Rv1819c* encodes BacA, an ABC transport protein required for the uptake of exogenous B_{12}. The vast majority of the mutations in *Rv1819c* were found in the conserved domains of BacA, seven in the transmembrane domain and three in the NBD with one

mutation located in between the two domains. Disruption of this protein could lead to the inability of the cell to take up exogenous B_{12} allowing *metE* expression and growth in the presence of high concentrations B_{12} giving rise to the suppressor phenotype.

In contrast to what we hypothesised, it appears that DnaE2 plays no role in the acquisition of the B_{12} suppressor phenotype in a B_{12}-sensitive Δ*metH* background when Mtb is exposed to high concentrations of B_{12} and Cbi on agar-solidified media. We therefore conclude that the major TLS polymerase in Mtb does not contribute to adaptive mutagenesis during nutrient deprivation. The question arises as to whether one of the Y-family polymerases, whose physiological function has yet to be determined, is contributing to the elevated level of mutagenesis. A major limitation of such a targeted approach is that one is limited to the studying a selected gene(s) whereas WGS of the suppressor isolates will allow identification of random, non-selected mutations as well. Out of the isolates sequenced, only two had no mutations in any of the three genes sequenced. It will be interesting to investigate the source of B_{12} resistance in these isolates, which could also contribute to a better understanding of B_{12} transport and assimilation. Future work will be aimed at eliminating the function of these Y-family polymerases in addition to DnaE2. Moreover, elucidating the mechanism enabling rapid emergence of B_{12} suppressors remains a key goal given the potential link to adaptive mutagenesis of drug resistance in clinical TB.

Appendices

Appendix 1: List of abbreviations

ADS	Albumin, dextrose and saline
Amp	Ampicillin
Amp^R	Ampicillin resistance
aph	Gene encoding aminoglycoside phosphotransferase
ATP	Adenosine triphosphate
attB	bacterial tRNAGly attachment site
ART	Anti-retroviral treatment
B_{12}	Vitamin B_{12}
BCG	Bacille Calmette-Guérin
BER	Base excision repair
BLAST	Basic Local Alignment Search Tool
bp	Base pair(s)
Cbi	Dicyanocobinamide
cDNA	Complementary DNA
CFU	Colony forming units
CH_3CO_2K	Potassium acetate
Cryo-EM	Cryo Electron Microscopy
CTAB	Cetyltrimethylammonium bromide
Da	Dalton
DCO	Double cross over
ddH_2O	Double distilled DNA
ddPCR	Digital droplet PCR
DEPC	Diethylpyrocarbonate
dG	Deoxyguanosine
DMSO	Dimethylsulphoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleic acid
DOTS	Directly Observed Therapy – Short Course

DS	Downstream
DSBs	Double stranded breaks
dsDNA	Double stranded DNA
DUF	Domain of Unknown Function
EDTA	Ethylenediaminetetraacetic acid
eGFP	Enhanced Green Fluorescent Protein
FAM	6-Carboxyfluorescein
g	Gravitational force
GN	Glucose salt
h	Hour
H_2O_2	Hydrogen peroxide
HCl	Hydrochloric acid
HDT	Host directed therapy
His	Histidine
HIV	Human Immunodeficiency virus
HR	Homologous recombination
Hyg	Hygromycin
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IRIS	Immune reconstitution inflammatory syndrome
Kan	Kanamycin
Kan^R	Kanamycin resistance
kbp	Kilo base pairs
KCl	Potassium chloride
LA	Luria-Bertani agar
lacZ	Gene encoding β-galactosidase
LB	Luria-Bertani broth
LTBI	Latent tuberculosis infection
MAC	Mutagenesis in aging colonies
MDR	Multidrug resistant
MIC_{90}	Minimum Inhibitory Concentration
MMC	Mitomycin C
MMP1	Matrix-metalloproteinase 1
MmpL	Mycobacterial membrane protein large

MMR	Mismatch repair
MOPS	3–(N-morpholino) propanesulfonic acid
mRNA	Messenger ribonucleic acid
Msm	*Mycobacterium smegmatis*
Mtb	*Mycobacterium tuberculosis*
NaCl	Sodium chloride
NaOAc	Sodium acetate
NaOH	Sodium hydroxide
NBD	Nucleotide-binding domain
NER	Nucleotide excision repair
Nfz	Nitrofurazone
NHEJ	Non-homologous end joining
NH_4Cl	Ammonium chloride
OADC	Oleic acid-albumin-dextrose-catalase
OD	Optical density
Ofx	Ofloxacin
ORF	Open reading frame
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PEG	Polyethylene glycol
PHP	Polymerase and histidinol phosphatase
Pmb	Plumbagin
PMF	Peptide Mass Fingerprinting
qRT-PCR	Quantitative Real Time PCR
RecA-ND	RecA/LexA-independent mechanism
RecA-NDp	RecA/LexA-independent promoter
Rif	Rifampicin
Rif^R	Rifampicin resistance
RNA	Ribonucleic acid
rNTPs	Ribonucleotides
rpm	Revolutions per minute
RT	Reverse transcriptase
PAGE	Polyacrylamide Gel Electrophoresis

SAB	Sample application buffer
sacB	Gene encoding levansucrase
SBP	Substrate binding proteins
SDS	Sodium Dodecyl Sulfate
SEC	Size Exclusion Chromatography
SCO	Single cross over
SNP	Single nucleotide polymorphism
SRAP	SOS Response Associated Peptidase
SSA	Single strand annealing
SSBP	Single stranded binding protein
SSC	Saline-sodium citrate
ssDNA	Single stranded DNA
TB	Tuberculosis
TE	Tris-EDTA
TEM	Transmission Electron Microscopy
TLS	Translesion synthesis
Tris	Tris(hydroxymethyl)aminomethane Polyoxyethylene
Tween	Polyoxyethylene sorbitan monooleate
TY	Tryptone Yeast
UCT	University of Cape Town
US	Upstream
UV	Ultraviolet
VIC	4,7,2'-trichloro-7'-phenyl-6-carboxyfluorescein
v/v	Volume/volume
WCV	Whole-cell vaccine
WGS	Whole genome sequencing
WHO	World Health Organization
w/v	Weight/volume
X-gal	5-bromo-4-chloro-3-indolyl-beta-D-galacto-pyranoside
XDR	Extensively drug-resistance

Appendix 2: Culture media

All media were made up to a final volume of 1 L with dH_2O and autoclaved at 121°C for 20 min.

Luria-Bertani Broth

10 g tryptone, 10 g NaCl, 5 g yeast extract

Luria-Bertani Agar

10 g tryptone, 10 g NaCl, 5 g yeast extract, 15 g bacteriological agar

2TY

4 g NaCl, 10 g yeast extract, 16 g tryptone

7H9/OADC

4.7 g Middlebrook 7H9 broth powder (DifcoTM, USA), 2 ml glycerol (Merck, Germany), 2 ml 25% Tween80

100 ml OADC Middlebrook OADC Enrichment (BD Microbiology Systems, USA) was added after autoclaving

7H10/OADC

19 g Middlebrook 7H10 agar powder (DifcoTM, USA), 5 ml glycerol (Merck, Germany)

100 ml OADC Middlebrook OADC Enrichment (BD Microbiology Systems, USA) was added after autoclaving

References

Abubakar, I., Zignol, M., Falzon, D., Raviglione, M., Ditiu, L., Masham, *et al*. (2013). Drug-resistant tuberculosis: time for visionary political leadership. *The Lancet Infectious Diseases*, *13*(6), 529–539.

Achtman, M. (2012). Insights from genomic comparisons of genetically monomorphic bacterial pathogens. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*(1590), 860–867.

Almeida Da Silva, P. E., & Palomino, J. C. (2011). Molecular basis and mechanisms of drug resistance in Mycobacterium tuberculosis: classical and new drugs. *Journal of Antimicrobial Chemotherapy*, *66*(7), 1417–1430.

Alvarez, B., & Radi, R. (2003). Peroxynitrite reactivity with amino acids and proteins. *Amino Acids*, *25*(3–4), 295–311.

Andersson, D. I., & Hughes, D. (2009). Gene amplification and adaptive evolution in bacteria. *Annual Review of Genetics*, *43*, 167–195.

Andersson, D. I., & Hughes, D. (2012). Evolution of antibiotic resistance at non-lethal drug concentrations. *Drug Resistance Updates*, *15*(3), 162–172.

Andersson, D. I., & Hughes, D. (2014). Microbiological effects of sublethal levels of antibiotics. *Nature Reviews. Microbiology*, *12*(7), 465–478.

Aniukwu, J., Glickman, M. S., & Shuman, S. (2008). The pathways and outcomes of mycobacterial NHEJ depend on the structure of the broken DNA ends. *Genes & Development*, *22*(4), 512–527.

Anthony, R. M., Schuitema, A. R., Bergval, I. L., Brown, T. J., Oskam, L., & Klatser, P. R. (2005). Acquisition of rifabutin resistance by a rifampicin resistant mutant of Mycobacterium tuberculosis involves an unusual spectrum of mutations and elevated frequency. *Annals of Clinical Microbiology and Antimicrobials*, *4*(1), 9.

Arana, M. E., & Kunkel, T. A. (2010). Mutator phenotypes due to DNA replication infidelity. *Seminars in Cancer Biology*, *20*(5), 304–311.

Aravind, L., Anand, S., & Iyer, L. M. (2013). Novel autoproteolytic and DNA-damage sensing components in the bacterial SOS response and oxidized methylcytosine-induced eukaryotic DNA demethylation systems. *Biology Direct*, *8*(1), 20.

Baharoglu, Z., & Mazel, D. (2014). SOS, the formidable strategy of bacteria against aggressions. *FEMS Microbiology Reviews*, *38*(6), 1126–1145.

Balu, S., Reljic, R., Lewis, M. J., Pleass, R. J., McIntosh, R., van Kooten, C., *et al.* (2011). A Novel Human IgA Monoclonal Antibody Protects against Tuberculosis. *J Immunol*, *186*(5), 3113–3119.

Barile, M., Giancaspero, T. A., Brizio, C., Panebianco, C., Indiveri, C., Galluccio, M., Gianazza, E. (2013). Biosynthesis of flavin cofactors in man: implications in health and disease. *Current Pharmaceutical Design*, *19*(14), 2649–2675.

Bebenek, K., Tissier, A., Frank, E. G., McDonald, J. P., Prasad, R., Wilson, S. H., *et al*. (2001). 5'-Deoxyribose phosphate lyase activity of human DNA polymerase iota in vitro. *Science*, *291*(5511), 2156–2159.

Bell, J. C., & Kowalczykowski, S. C. (2016). Mechanics and Single-Molecule Interrogation of DNA Recombination. *Annual Review of Biochemistry*, *85*(1), 193–226.

Bhaskar, A., Chawla, M., Mehta, M., Parikh, P., Chandra, P., Bhave, D., *et al.* (2014). Reengineering Redox Sensitive GFP to Measure Mycothiol Redox Potential of *Mycobacterium tuberculosis* during Infection. *PLoS Pathogens*, *10*(1), e1003902.

Bjedov, I., Tenaillon, O., Souza, V., Denamur, E., & Radman, M. (2003). Stress-induced mutagenesis in bacteria. *Science*, *300*(MAY), 1404–1409.

Blount, K. F., & Breaker, R. R. (2006). Riboswitches as antibacterial drug targets. *Nature Biotechnology*, *24*(12), 1558–1564.

Borsari, C., Ferrari, S., Venturelli, A., & Costi, M. P. (2016). Target-based approaches for the discovery of new antimycobacterial drugs. *Drug Discovery Today*.

Bos, J., Zhang, Q., Vyawahare, S., Rogers, E., Rosenberg, S. M., & Austin, R. H. (2015). Emergence of antibiotic resistance from multinucleated bacterial filaments. *Proceedings of the National Academy of Sciences*, *112*(1), 178–183.

Boshoff, H. I. M., Reed, M. B., Barry, C. E., & Mizrahi, V. (2003). DnaE2 polymerase contributes to in vivo survival and the emergence of drug resistance in *Mycobacterium tuberculosis*. *Cell*, *113*(2), 183–193.

Boshoff, H. I., & Mizrahi, V. (2000). Expression of *Mycobacterium smegmatis* pyrazinamidase in *Mycobacterium tuberculosis* confers hypersensitivity to pyrazinamide and related amides. *Journal of Bacteriology*, *182*(19), 5479–5485.

Brauner, A., Fridman, O., Gefen, O., & Balaban, N. Q. (2016). Distinguishing between resistance, tolerance and persistence to antibiotic treatment. *Nature Reviews. Microbiology*, *14*(5), 320–330.

Briken, V., & Miller, J. L. (2008). Living on the edge: inhibition of host cell apoptosis

by. *Future Medicine*, *3*(4), 415–422.

Bruchfeld, J., Correia-Neves, M., & Kallenius, G. (2015). Tuberculosis and HIV coinfection. *Cold Spring Harbor Perspectives in Medicine*, *5*(7), 1–16.

Bryant, J. M., Harris, S. R., Parkhill, J., Dawson, R., Diacon, A. H., van Helden, P., *et al*. (2013). Whole-genome sequencing to establish relapse or re-infection with *Mycobacterium tuberculosis*: A retrospective observational study. *The Lancet Respiratory Medicine*, *1*(10), 786–792.

Butala, M., Žgur-Bertok, D., & Busby, S. J. W. (2009). The bacterial LexA transcriptional repressor. *Cellular and Molecular Life Sciences*, *66*(1), 82-93.

Cabrera, S. E., Santos, D., Valverde, M. P., Domínguez-Gil, A., González, F., Luna, G., & García, M. J. (2009). Influence of the cytochrome P450 2B6 genotype on population pharmacokinetics of efavirenz in human immunodeficiency virus patients. *Antimicrobial Agents and Chemotherapy*, *53*(7), 2791–2798.

Cairns, J., & Foster, P. L. (1991). Adaptive reversion of a frameshift mutation in Escherichia coli. *Genetics*, *128*(4), 695–701.

Cardona, P. J. (2006). RUTI: A new chance to shorten the treatment of latent tuberculosis infection. *Tuberculosis*, *86*, 273–289.

Casali, N., Nikolayevskyy, V., Balabanova, Y., Harris, S. R., Kontsevaya, I., Corander, J., *et al*. (2014). Evolution and transmission of drug resistant tuberculosis in a Russian population. *Nature Genetics*, *46*(3), 279–286.

Cashman, J. R. (2005). Some distinctions between flavin-containing and cytochrome P450 monooxygenases. *Biochemical and Biophysical Research Communications*, *338*(1), 599–604.

Castañeda-García, A., Prieto, A. I., Rodríguez-Beltrán, J., Alonso, N., Cantillon, D., Costas, C., *et al*. (2017). A non-canonical mismatch repair pathway in prokaryotes. *Nature Communications*, *8*, 14246.

Chapman, H. A., Riese, R. J., & Shi, G. P. (1997). Emerging roles for cysteine proteases in human biology. *Annual Review of Physiology*, *59*(1), 63–88.

Chayen, N. E., & Saridakis, E. (2008). Protein crystallization: from purified protein to diffraction-quality crystal. *Nature Methods*, *5*(2), 147–153.

Chen, A. H., & Silver, P. A. (2012). Designing biological compartmentalization. *Trends in Cell Biology*, *22*(12), 662–670.

Cheo, D. L., Bayles, K. W., & Yasbin, R. E. (1991). Cloning and characterization of DNA damage-inducible promoter regions from *Bacillus subtilis*. *Journal of*

Bacteriology, *173*(5), 1696–1703.

Chim, N., Torres, R., Liu, Y., Capri, J., Batot, G., Whitelegge, J. P., & Goulding, C. W. (2015). The Structure and Interactions of Periplasmic Domains of Crucial MmpL Membrane Proteins from *Mycobacterium tuberculosis*. *Chemistry & Biology*, *22*(8), 1098–1107.

Cohen, J. (2013). Approval of Novel TB Drug Celebrated-With Restraint. *Science*, *339*(6116), 130.

Cohen, N. R., Lobritz, M. A., & Collins, J. J. (2013). Microbial persistence and the road to drug resistance. *Cell Host and Microbe*, *13*(6), 632–642.

Colangeli, R., Arcus, V. L., Cursons, R. T., Ruthe, A., Karalus, N., Coley, K., *et al*. (2014). Whole genome sequencing of *Mycobacterium tuberculosis* reveals slow growth and low mutation rates during latent infections in humans. *PLoS ONE*, *9*(3), e91024.

Cole, S., Brosch, R., Parkhill, J., & Garnier, T. (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature*, *393*, 537–544.

Collins, L., & Franzblau, S. G. (1997). Microplate alamar blue assay versus BACTEC 460 system for high-throughput screening of compounds against *Mycobacterium tuberculosis* and *Mycobacterium avium*. *Antimicrobial Agents and Chemotherapy*, *41*(5), 1004–1009.

Courcelle, J., Khodursky, A., Peter, B., Brown, P. O., Hanawalt, P. C., Arthur, H. M., *et al*. (2001). Comparative gene expression profiles following UV exposure in wild-type and SOS-deficient *Escherichia coli*. *Genetics*, *158*(1), 41–64.

Cox, H. S., Morrow, M., & Deutschmann, P. W. (2008). Long term efficacy of DOTS regimens for tuberculosis: systematic review. *BMJ (Clinical Research Ed.)*, *336*(7642), 484–487.

Dabbs, E. R., Yazawa, K., Mikami, Y., Miyaji, M., Morisaki, N., Iwasaki, S., & Furihata, K. (1995). Ribosylation by mycobacterial strains as a new mechanism of rifampin inactivation. *Antimicrobial Agents and Chemotherapy*, *39*(4), 1007–9.

Darwin, K. H., Ehrt, S., Gutierrez-Ramos, J. C., Weich, N., & Nathan, C. F. (2003). The proteasome of *Mycobacterium tuberculosis* is required for resistance to nitric oxide. *Science*, *302*(5652), 1963–1966.

Darwin, K. H., & Nathan, C. F. (2005). Role for Nucleotide Excision Repair in Virulence of *Mycobacterium tuberculosis*. *Infection and Immunity*, *73*(8), 4581–

4587.

David, H. L. (1970). Probability distribution of drug-resistant mutants in unselected populations of *Mycobacterium tuberculosis*. *Applied Microbiology*, *20*(5), 810–814.

Davies, J. (2013). Specialized microbial metabolites: functions and origins. *The Journal of Antibiotics*, *66*(7), 361–364.

Davies, J., Gilbert, W., & Gorini, L. (1964). Streptomycin, Suppression, and the Code. *Proceedings of the National Academy of Sciences*, *51*(1960), 883–890.

Davis, E., Dullaghan, E., & Rand, L. (2002). Definition of the mycobacterial SOS box and use to identify LexA-regulated genes in *Mycobacterium tuberculosis*. *Journal of Bacteriology*, *184*(12), 3287–3295.

Davis, E., & Forse, L. (2009). DNA repair: key to survival? *In Mycobacterium: Genomics and Molecular Biology* (Parish, T. and Brown, A., eds.), pp. 79–119, Caister Academic Press.

Davis, E. O., Jenner, P. J., Brooks, P. C., Colston, M. J., & Sedgwick, S. G. (1992). Protein splicing in the maturation of *M. tuberculosis* RecA protein: A mechanism for tolerating a novel class of intervening sequence. *Cell*, *71*(2), 201–210.

Davis, E. O., Sedgwick, S. G., & Colston, M. J. (1991). Novel structure of the *recA* locus of *Mycobacterium tuberculosis* implies processing of the gene product. *Journal of Bacteriology*, *173*(18), 5653–5662.

Davis, E. O., Springer, B., Gopaul, K. K., Papavinasasundaram, K. G., Sander, P., & Böttger, E. C. (2002). DNA damage induction of recA in *Mycobacterium tuberculosis* independently of RecA and LexA. *Molecular Microbiology*, *46*(3), 791–800.

Deriano, L., & Roth, D. B. (2013). Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. *Annual Review of Genetics*, *47*, 433–455.

Dheda, K., Limberis, J. D., Pietersen, E., Phelan, J., Esmail, A., Lesosky, M., *et al.* (2017). Outcomes, infectiousness, and transmission dynamics of patients with extensively drug-resistant tuberculosis and home-discharged patients with programmatically incurable tuberculosis: a prospective cohort study. *The Lancet Respiratory Medicine*, (January).

Dietrich, L. E. P., Teal, T. K., Price-Whelan, A., & Newman, D. K. (2008). Redox-Active Antibiotics Control Gene Expression and Community Behavior in

Divergent Bacteria. *Science*, *321*(5893), 1203–1206.

Ditse, Z. (2015). *Replication fidelity in the microevolution of mycobacteria*. Ph.D thesis. University of Cape Town, South Africa

Domenech, P., Kolly, G. S., Leon-Solis, L., Fallow, A., & Reed, M. B. (2010). Massive gene duplication event among clinical isolates of the *Mycobacterium tuberculosis* W/Beijing family. *Journal of Bacteriology*, *192*(18), 4562–4570.

Domenech, P., Reed, M. B., & Iii, C. E. B. (2005). Contribution of the *Mycobacterium tuberculosis* MmpL Protein Family to Virulence and Drug Resistance Contribution of the *Mycobacterium tuberculosis* MmpL Protein Family to Virulence and Drug Resistance. *Infection and Immunity*, *73*(6), 3492–3501.

Domenech, P., Rog, A., Moolji, J. ud din, Radomski, N., Fallow, A., Leon-Solis, *et al.* (2014). Origins of a 350-kilobase genomic duplication in *Mycobacterium tuberculosis* and its impact on virulence. *Infection and Immunity*, *82*(7), 2902–2912.

Donigan, K. A., McLenigan, M. P., Yang, W., Goodman, M. F., & Woodgate, R. (2014). The steric gate of dna polymerase iota regulates ribonucleotide incorporation and deoxyribonucleotide fidelity. *Journal of Biological Chemistry*, *289*(13), 9136–9145.

Dörr, T., Lewis, K., & Vulić, M. (2009). SOS Response Induces Persistence to Fluoroquinolones in *Escherichia coli*. *PLoS Genetics*, *5*(12), e1000760.

Dos Vultos, T., Mestre, O., Tonjum, T., & Gicquel, B. (2009). DNA repair in *Mycobacterium tuberculosis* revisited. *FEMS Microbiology Reviews*, *33*(3), 471–487.

Drlica, K., & Zhao, X. (1997). DNA gyrase, topoisomerase IV, and the 4-quinolones. *Microbiology and Molecular Biology Reviews* , *61*(3), 377–392.

Durbach, S. I., Andersen, S. J., & Mizrahi, V. (1997). SOS induction in mycobacteria: analysis of the DNA-binding activity of a LexA-like repressor and its role in DNA damage induction of the *recA* gene from *Mycobacterium smegmatis*. *Molecular Microbiology*, *26*(4), 643–653.

Durbach, S. I., Springer, B., Machowski, E. E., North, R. J., Papavinasasundaram, K. G., Colston, M. J., *et al*. (2003). DNA Alkylation Damage as a Sensor of Nitrosative Stress in *Mycobacterium tuberculosis*. *Infection and Immunity*, *71*(2), 997–1000.

Ebrahimi-Rad, M., Bifani, P., Martin, C., Kremer, K., Samper, S., Rauzier, J., *et al.*

(2003). Mutations in putative mutator genes of *Mycobacterium tuberculosis* strains of the W-Beijing family. *Emerging Infectious Diseases*, *9*(7), 838–845.

Ehrt, S., & Schnappinger, D. (2009). Mycobacterial survival strategies in the phagosome: defence against host stresses. *Cellular Microbiology*, *11*(8), 1170–1178.

Eldholm, V., & Balloux, F. (2016). Antimicrobial Resistance in *Mycobacterium tuberculosis*: The Odd One Out. *Trends in Microbiology*, *24*(8), 637–648.

Eldholm, V., Rieux, A., Monteserin, J., Lopez, J. M., Palmero, D., Lopez, B., *et al*. (2016). Impact of HIV co-infection on the evolution and transmission of multidrug- resistant tuberculosis. *eLIFE*, *5*, e16644.

Erill, I., Campoy, S., Barbé, J., Abella, M., Erill, I., Jara, M., *et al*. (2007). Aeons of distress: an evolutionary perspective on the bacterial SOS response. *FEMS Microbiology Reviews*, *31*(6), 637–656.

Erill, I., Campoy, S., Mazon, G., & Barbé, J. (2006). Dispersal and regulation of an adaptive mutagenesis cassette in the bacteria domain. *Nucleic Acids Research*, *34*(1), 66–77.

Evans, T. G., Schrager, L., & Thole, J. (2016). Status of vaccine research and development of vaccines for tuberculosis. *Vaccine*, *34*(26), 2911–2914.

Farhat, M. R., Jacobson, K. R., Franke, M. F., Kaur, D., Sloutsky, A., Mitnick, C. D., & Murray, M. (2016). Gyrase Mutations Are Associated with Variable Levels of Fluoroquinolone Resistance in *Mycobacterium tuberculosis*. *Journal of Clinical Microbiology*, *54*(3), 727–733.

Fijalkowska, I. J., Schaaper, R. M., & Jonczyk, P. (2012). DNA replication fidelity in *Escherichia coli*: a multi-DNA polymerase affair. *FEMS Microbiology Reviews*, *36*(6), 1105–1121.

Ford, C. B., Lin, P. L., Chase, M. R., Shah, R. R., Iartchouk, O., Galagan, J., *et al*. (2011). Use of whole genome sequencing to estimate the mutation rate of *Mycobacterium tuberculosis* during latent infection. *Nature Genetics*, *43*(5), 482–486.

Ford, C. B., Shah, R. R., Maeda, M. K., Gagneux, S., Murray, M. B., Cohen, T., *et al*. (2013). *Mycobacterium tuberculosis* mutation rate estimates from different lineages predict substantial differences in the emergence of drug-resistant tuberculosis. *Nature Genetics*, *45*(7), 784–790.

Foster, P. L. (1997). Nonadaptive mutations occur on the F' episome during adaptive

mutation conditions in *Escherichia coli*. *Journal of Bacteriology*, *179*(5), 1550–1554.

Foster, P. L. (1999). Mechanisms of stationary phase mutation: a decade of adaptive mutation. *Annual Review of Genetics*, *33*, 57–88.

Foster, P. L. (2000). Adaptive mutation: implications for evolution. *BioEssays : News and Reviews in Molecular, Cellular and Developmental Biology*, *22*(12), 1067–1074.

Foster, P. L. (2007). Stress-induced mutagenesis in bacteria. *Critical Reviews in Biochemistry and Molecular Biology*, *42*(5), 373–397.

Friedberg, E. C., Walker, G. C., Siede, W., Wood, R., Schultz, R., & Ellenberger, T. (1995). *DNA Repair and Mutagenesis* (2nd ed.). Washington, DC: ASM Press.

Gagneux, S. (2012). Host-pathogen coevolution in human tuberculosis. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*(1590), 850–859.

Gagneux, S. (2013). Genetic diversity in *Mycobacterium tuberculosis*. *Current Topics in Microbiology and Immunology*, *374*, 1–25.

Galagan, J. E. (2014). Genomic insights into tuberculosis. *Nature Reviews Genetics*, *15*(5), 307–320.

Galan, J. C., Gonzalez-Candelas, F., Rolain, J. M., & Canton, R. (2013). Antibiotics as selectors and accelerators of diversity in the mechanisms of resistance: From the resistome to genetic plasticity in the β-lactamases world. *Frontiers in Microbiology*, *4*, article 9.

Galhardo, R. S., Rocha, R. P., Marques, M. V, & Menck, C. F. M. (2005). An SOS-regulated operon involved in damage-inducible mutagenesis in *Caulobacter crescentus*. *Nucleic Acids Research*, *33*(8), 2603–2614.

Gamulin, V., Cetkovic, H., & Ahel, I. (2004). Identification of a promoter motif regulating the major DNA damage response mechanism of *Mycobacterium tuberculosis*. *FEMS Microbiology Letters*, *238*(1), 57–63.

Gengenbacher, M., & Kaufmann, S. H. E. (2012). *Mycobacterium tuberculosis*: success through dormancy. *FEMS Microbiology Reviews*, *36*(3), 514–532.

Georgescu, R. E., Yurieva, O., Kim, S., Kuriyan, J., Kong, X., O'Donnell, M. (2008) Structure of a small-molecule inhibitor of a DNA polymerase sliding clamp. *Proceedings of the National Academy of Sciences, 105*(32), 11116-11121.

Gler, M. T., Skripconoka, V., Sanchez-Garavito, E., Xiao, H., Cabrera-Rivero, J. L.,

Vargas-Vasquez, D. E., *et al*. (2012). Delamanid for Multidrug-Resistant Pulmonary Tuberculosis. *New England Journal of Medicine*, *366*(23), 2151–2160.

Glickman, M. S. (2014). Double-Strand DNA Break Repair in Mycobacteria. *Microbiology Spectrum*, *2*(5), 657–666.

Goh, E.-B., Yim, G., Tsui, W., McClure, J., Surette, M. G., & Davies, J. (2002). Transcriptional modulation of bacterial gene expression by subinhibitory concentrations of antibiotics. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(26), 17025–17030.

Gomez-Marroquin, M., Martin, H. A., Pepper, A., Girard, M. E., Kidman, A. A., Vallin, C., *et al*. (2016). Stationary-phase mutagenesis in stressed *Bacillus subtilis* cells operates by Mfd-dependent mutagenic pathways. *Genes*, *7*(7), 1–13.

Gomez-Marroquin, M., Vidales, L. E., Debora, B. N., Santos-Escobar, F., Obregon-Herrera, A., Robleto, E. A., & Pedraza-Reyes, M. (2015). Role of *Bacillus subtilis* DNA glycosylase MutM in counteracting oxidatively induced DNA damage and in stationary-phase-associated mutagenesis. *Journal of Bacteriology*, *197*(11), 1963–1971.

Gong, C., Bongiorno, P., Martins, A., Stephanou, N. C., Zhu, H., Shuman, S., & Glickman, M. S. (2005). Mechanism of nonhomologous end-joining in mycobacteria: a low-fidelity repair system driven by Ku, ligase D and ligase C. *Nature Structural & Molecular Biology*, *12*(4), 304–312.

Gopaul, K. K., Brooks, P. C., Prost, J.-F., & Davis, E. O. (2003). Characterization of the Two *Mycobacterium tuberculosis recA* Promoters. *Journal of Bacteriology*, *185*(20), 6005–6015.

Gopinath, K., Venclovas, C., Ioerger, T. R., Sacchettini, J. C., McKinney, J. D., Mizrahi, V., & Warner, D. F. (2013). A vitamin B_{12} transporter in *Mycobacterium tuberculosis*. *Open Biology*, *3*(2), 120–175.

Gordhan, B. G. and Parish, T. (2001). Gene Replacement using Pretreated DNA. *Mycobacterium tuberculosis* Protocols. T. Parish and N. G. Stoker. Otowa, New Jersey, Humana Press: 59-92.

Gorna, A. E., Bowater, R. P., Dziadek, J., Cooper, A. M., Friedberg, E. C., Warner, D. F., *et al*. (2010). DNA repair systems and the pathogenesis of *Mycobacterium tuberculosis*: varying activities at different stages of infection. *Clinical Science*, *119*(5), 187–202.

Gu, S., Li, W., Zhang, H., Fleming, J., Yang, W., Wang, S., *et al*. (2016). The β2 clamp

in the *Mycobacterium tuberculosis* DNA polymerase III αβ2ε replicase promotes polymerization and reduces exonuclease activity. *Scientific Reports*, *6*, 18418.

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 Pathogens*, *7*(7), e1002158.

Guo, Y., Bandaru, V., Jaruga, P., Zhao, X., Burrows, C. J., Iwai, S., *et al*. (2010). The oxidative DNA glycosylases of *Mycobacterium tuberculosis* exhibit different substrate preferences from their *Escherichia coli* counterparts. *DNA Repair*, *9*(2), 177–190.

Gupta, R., Barkan, D., Redelman-Sidi, G., Shuman, S., & Glickman, M. S. (2011). Mycobacteria exploit three genetically distinct DNA double-strand break repair pathways. *Molecular Microbiology*, *79*(2), 316–330.

Gupta, R., Ryzhikov, M., Koroleva, O., Unciuleac, M., Shuman, S., Korolev, S., & Glickman, M. S. (2013). A dual role for mycobacterial RecO in RecA-dependent homologous recombination and RecA-independent single-strand annealing. *Nucleic Acids Research*, *41*(4), 2284–2295.

Gupta, R., Shuman, S., & Glickman, M. S. (2015). RecF and RecR play critical roles in the homologous recombination and single-strand annealing pathways of mycobacteria. *Journal of Bacteriology*, *197*(19), 3121–3132.

Gupta, R., Unciuleac, M., Shuman, S., & Glickman, M. S. (2016). Homologous recombination mediated by the mycobacterial AdnAB helicase without end resection by the AdnAB nucleases. *Nucleic Acids Research*, *1*, 1–13.

Harding, C. V., & Boom, W. H. (2010). Regulation of antigen presentation by *Mycobacterium tuberculosis*: a role for Toll-like receptors. *Nature Reviews Microbiology*, *8*(4), 296–307.

Helaine, S., Cheverton, A. M., Watson, K. G., Faure, L. M., Matthews, S. A., & Holden, D. W. (2014). Internalization of Salmonella by macrophages induces formation of nonreplicating persisters. *Science*, *343*(6167), 204–208.

Hershberg, R., Lipatov, M., Small, P. M., Sheffer, H., Niemann, S., Homolka, S., *et al*. (2008). High Functional Diversity in *Mycobacterium tuberculosis* Driven by Genetic Drift and Human Demography. *PLoS Biology*, *6*(12), e311.

Hmama, Z., Peña-Díaz, S., Joseph, S., & Av-Gay, Y. (2015). Immunoevasion and immunosuppression of the macrophage by *Mycobacterium tuberculosis*. *Immunological Reviews*, *264*(1), 220–232.

Hoagland, D. T., Liu, J., Lee, R. B., & Lee, R. E. (2016). New agents for the treatment of drug-resistant *Mycobacterium tuberculosis*. *Advanced Drug Delivery Reviews*, *102*, 55–72.

Ioerger, T. R., Feng, Y., Ganesula, K., Chen, X., Dobos, K. M., Fortune, S., *et al.* (2010). Variation among genome sequences of H37Rv strains of *Mycobacterium tuberculosis* from multiple laboratories. *Journal of Bacteriology*, *192*(14), 3645–3653.

Ischiropoulos, H., Zhu, L., Chen, J., Tsai, M., Martin, J. C., Smith, C. D., & Beckman, J. S. (1992). Peroxynitrite-mediated tyrosine nitration catalyzed by superoxide dismutase. *Archives of Biochemistry and Biophysics*, *298*(2), 431–437.

Ishino, S., Nishi, Y., Oda, S., Uemori, T., Sagara, T., Takatsu, N., *et al*. (2016). Identification of a mismatch-specific endonuclease in hyperthermophilic Archaea. *Nucleic Acids Research*, *44*(7), 2977–2986.

Jain, R., Kumar, P., & Varshney, U. (2007). A distinct role of formamidopyrimidine DNA glycosylase (MutM) in down-regulation of accumulation of G , C mutations and protection against oxidative stress in mycobacteria. *DNA Repair*, *6*, 1774–1785.

Janion, C. (2008). Inducible SOS response system of DNA repair and mutagenesis in *Escherichia coli*. *International Journal of Biological Sciences*, *4*(6), 338-344.

Jarosz, D. F., Beuning, P. J., Cohen, S. E., & Walker, G. C. (2007). Y-family DNA polymerases in Escherichia coli. *Trends in Microbiology*, *15*(2), 70–77.

Jarosz, D. F., Godoy, V. G., Delaney, J. C., Essigmann, J. M., & Walker, G. C. (2006). A single amino acid governs enhanced activity of DinB DNA polymerases on damaged templates. *Nature*, *439*(7073), 225–228.

Javid, B., Sorrentino, F., Toosky, M., Zheng, W., Pinkham, J. T., Jain, N., *et al*. (2014). Mycobacterial mistranslation is necessary and sufficient for rifampicin phenotypic resistance. *Proceedings of the National Academy of Sciences*, *111*(3), 1132–1137.

Johnson, P. J. T., & Levin, B. R. (2013). Pharmacodynamics, Population Dynamics, and the Evolution of Persistence in *Staphylococcus aureus*. *PLoS Genetics*, *9*(1), e1003123.

Kamarthapu, V., & Nudler, E. (2015). Rethinking transcription coupled DNA repair. *Current Opinion in Microbiology*, *24*, 15–20.

Kana, B. D., Abrahams, G. L., Sung, N., Warner, D. F., Gordhan, B. G., Machowski, E. E., *et al*. (2010). Role of the DinB homologs Rv1537 and Rv3056 in

Mycobacterium tuberculosis. *Journal of Bacteriology*, *192*(8), 2220–2227.

Kang, J., & Blaser, M. J. (2006). Bacterial populations as perfect gases: genomic integrity and diversification tensions in *Helicobacter pylori*. *Nature Reviews. Microbiology*, *4*(11), 826–836.

Kaplan, G., Post, F. A., Moreira, A. L., Wainwright, H., Kreiswirth, B. N., Tanverdi, M., *et al*. (2003). *Mycobacterium tuberculosis* Growth at the Cavity Surface: a Microenvironment with Failed Immunity. *Infection and Immunity*, *71*(12), 7099–7108.

Karunakaran, P., & Davies, J. (2000). Genetic antagonism and hypermutability in Mycobacterium smegmatis. *Journal of Bacteriology*, *182*(12), 3331–3335.

Kasak, L., Horak, R., & Kivisaar, M. (1997). Promoter-creating mutations in *Pseudomonas putida* : A model system for the study of mutation in starving bacteria. *Proceedings of the National Academy of Science*, *94*, 3134–3139.

Katz, S., & Hershberg, R. (2013). Elevated Mutagenesis Does Not Explain the Increased Frequency of Antibiotic Resistant Mutants in Starved Aging Colonies. *PLoS Genetics*, *9*(11).

Kaufmann, S. H. E., Lange, C., Rao, M., Balaji, K. N., Lotze, M., Schito, M., *et al*. (2014). Progress in tuberculosis vaccine development and host-directed therapies-a state of the art review. *The Lancet Respiratory Medicine*, *2*(4), 301–320.

Kennaway, C. K., Benesch, J. L. P., Gohlke, U., Wang, L., Robinson, C. V., Orlova, E. V., *et al*. (2005). Dodecameric structure of the small heat shock protein Acr1 from *Mycobacterium tuberculosis*. *Journal of Biological Chemistry*, *280*(39), 33419–33425.

Kesavan, A. K., Brooks, M., Tufariello, J., Chan, J., & Manabe, Y. C. (2009). Tuberculosis genes expressed during persistence and reactivation in the resistant rabbit model. *Tuberculosis*, *89*(1), 17–21.

Kim, S. R., Maenhaut-Michel, G., Yamada, M., Yamamoto, Y., Matsui, K., Sofuni, T., *et al*. (1997). Multiple pathways for SOS-induced mutagenesis in *Escherichia coli*: an overexpression of dinB/dinP results in strongly enhancing mutagenesis in the absence of any exogenous treatment to damage DNA. *Proceedings of the National Academy of Sciences*, *94*(25), 13792–13797.

Kiran, D., Podell, B. K., Chambers, M., & Basaraba, R. J. (2016). Host-directed therapy targeting the *Mycobacterium tuberculosis* granuloma: a review. *Seminars in Immunopathology*, *38*(2), 167–183.

Kisker, C., Kuper, J., & Houten, B. Van. (2013). Prokaryotic Nucleotide Excision Repair. *Cold Spring Harbor Perspectives in Biology*, *5*, a012591.

Kivisaar, M. (2010). Mechanisms of stationary-phase mutagenesis in bacteria: Mutational processes in pseudomonads. *FEMS Microbiology Letters*, *312*(1), 1–14.

Kling, A., Lukat, P., Almeida, D. V., Bauer, A., Fontaine, E., Sordello, S., *et al*. (2015). Targeting DnaN for tuberculosis therapy using novel griselimycins. *Science*, *348*(6239), 1106–1112.

Koch, A., Mizrahi, V., & Warner, D. F. (2014). The impact of drug resistance on *Mycobacterium tuberculosis* physiology: what can we learn from rifampicin? *Emerging Microbes & Infections*, *3*(3), e17.

Koorits, L., Tegova, R., Tark, M., Tarassova, K., Tover, A., & Kivisaar, M. (2007). Study of involvement of ImuB and DnaE2 in stationary-phase mutagenesis in *Pseudomonas putida*. *DNA Repair*, *6*(6), 863–868.

Kowalczykowski, S. C. (2000). Initiation of genetic recombination and recombination-dependent replication. *Trends in Biochemical Sciences*, *25*(4), 156–165.

Krokan, H. E., & Bjøra, M. (2013). Base Excision Repair. *Cold Spring Harbor Perspectives in Biology*, *5*, a012583.

Kuban, W., Jonczyk, P., Gawel, D., Malanowska, K., Schaaper, R. M., & Fijalkowska, I. J. (2004). Role of *Escherichia coli* DNA Polymerase IV in In Vivo Replication Fidelity. *Journal of Bacteriology*, *186*(14), 4802–4807.

Kumar, A., Beloglazova, N., Bundalovic-Torma, C., Phanse, S., Deineko, V., Gagarinova, A., *et al*. (2016). Conditional Epistatic Interaction Maps Reveal Global Functional Rewiring of Genome Integrity Pathways in *Escherichia coli*. *Cell Reports*, *14*(3), 648–661.

Kumar, R. A., Vaze, M. B., Chandra, N. R., Vijayan, M., & Muniyappa, K. (1996). Functional characterization of the precursor and spliced forms of RecA protein of *Mycobacterium tuberculosis*. *Biochemistry*, *35*(6), 1793–802.

Kurthkoti, K., & Varshney, U. (2010). Detrimental effects of hypoxia-specific expression of uracil DNA glycosylase (Ung) in *Mycobacterium smegmatis*. *Journal of Bacteriology*, *192*(24), 6439–6446.

Kurthkoti, K., & Varshney, U. (2011). Base excision and nucleotide excision repair pathways in mycobacteria. *Tuberculosis*, *91*(6), 533–543.

Larsen, M. H. (2000). Appendix 1. Molecular Genetics of Mycobacteria. G. F. Hatfull

and W. R. Jacobs, Jr. Washington, D.C., ASM Press, 313-320.

Larsen, M. H., Biermann, K., Tandberg, S., Hsu, T., & Jacobs, W. R. (2007). Genetic Manipulation of *Mycobacterium tuberculosis*. *Current Protocols in Microbiology*, *Chapter 10* (August), Unit 10A.2.

Lawn, S. D., Meintjes, G., McIlleron, H., Harries, A. D., & Wood, R. (2013). Management of HIV-associated tuberculosis in resource-limited settings: a state-of-the-art review. *BMC Medicine*, *11*, 253.

Lee, H. H., Molla, M. N., Cantor, C. R., & Collins, J. J. (2010). Bacterial charity work leads to population-wide resistance. *Nature*, *467*(7311), 82–85.

Leng, T., Pan, M., Xu, X., & Javid, B. (2015). Translational misreading in *Mycobacterium smegmatis* increases in stationary phase. *Tuberculosis*, *95*(6), 678–681.

Levin-reisman, I., Ronin, I., Gefen, O., Braniss, I., Shoresh, N., & Balaban, N. Q. (2017). Antibiotic tolerance facilitates the evolution of resistance. *Science*, 10.1126/science.aaj2191, 1–10.

Li, X., Tao, J., Han, J., Hu, X., Chen, Y., Deng, H., Mi, K. (2014). The gain of hydrogen peroxide resistance benefits growth fitness in mycobacteria under stress. *Protein and Cell*. Springer.

Li, X., Wu, J., Han, J., Hu, Y., Mi, K., Zahrt, T., Clark-Curtiss, J. (2015). Distinct Responses of *Mycobacterium smegmatis* to Exposure to Low and High Levels of Hydrogen Peroxide. *PLOS ONE*, *10*(7), e0134595.

Lieberman, T. D., Wilson, D., Misra, R., Xiong, L. L., Moodley, P., Cohen, T., & Kishony, R. (2016). Genomic diversity in autopsy samples reveals within-host dissemination of HIV-associated *Mycobacterium tuberculosis*. *Nature Medicine*, (October), 1–7.

MacLean, R. C., Torres-Barceló, C., & Moxon, R. (2013). Evaluating evolutionary models of stress-induced mutagenesis in bacteria. *Nature Reviews Genetics*, *14*(3), 221–227.

Maisnier-Patin, S., & Roth, J. R. (2015). The Origin of Mutants Under Selection: How Natural Selection Mimics Mutagenesis (Adaptive Mutation). *Cold Spring Harbor Perspectives in Biology*, *7*(7), a018176.

Manganelli, R., Voskuil, M. I., Schoolnik, G. K., Dubnau, E., Gomez, M., & Smith, I. (2002). Role of the extracytoplasmic-function sigma factor sigma(H) in *Mycobacterium tuberculosis* global gene expression. *Molecular Microbiology*,

45(2), 365–374.

Mangtani, P., Abubakar, I., Ariti, C., Beynon, R., Pimpin, L., Fine, P. E. M., *et al.* (2014). Protection by BCG vaccine against tuberculosis: A systematic review of randomized controlled trials. *Clinical Infectious Diseases*, *58*(4), 470–480.

Martin, C. J., Peters, K. N., & Behar, S. M. (2014). Macrophages clean up: Efferocytosis and microbial control. *Current Opinion in Microbiology*, *17*(1), 17–23.

Matsumoto, A., Vos, J. M. H., & Hanawalt, P. C. (1989). Repair analysis of mitomycin C-induced DNA crosslinking in ribosomal RNA genes in lymphoblastoid cells from Fanconi's anemia patients. *Mutation Research-DNA Repair*, *217*(3), 185–192.

Mayanja-Kizza, H., Jones-Lopez, E., Okwera, A., Wallis, R. S., Ellner, J. J., Mugerwa, R. D., & Whalen, C. C. (2005). Immunoadjuvant prednisolone therapy for HIV-associated tuberculosis: a phase 2 clinical trial in Uganda. *The Journal of Infectious Diseases*, *191*(6), 856–865.

Mazaika, E., Homsy, J., Mazaika, E., & Homsy, J. (2014). Digital Droplet PCR: CNV Analysis and Other Applications. In *Current Protocols in Human Genetics* (p. 7.24.1-7.24.13).

McGlynn, P., & Lloyd, R. G. (2002). Genome stability and the processing of damaged replication forks by RecG. *Trends in Genetics*, *18*(8), 413-419.

McGrath, M., Gey van pittius, N. C., Van helden, P. D., Warren, R. M., & Warner, D. F. (2014). Mutation rate and the emergence of drug resistance in *Mycobacterium tuberculosis*. *Journal of Antimicrobial Chemotherapy 69*(2), 292-302.

McHenry, C. S. (2011). DNA replicases from a bacterial perspective. *Annual Review of Biochemistry*, *80*(1), 403–436.

McInerney, P., Johnson, A., Katz, F., & O'Donnell, M. (2007). Characterization of a Triple DNA Polymerase Replisome. *Molecular Cell*, *27*(4), 527–538.

McKenzie, G. J., Lee, P. L., Lombardo, M. J., Hastings, P. J., & Rosenberg, S. M. (2001). SOS mutator DNA polymerase IV functions in adaptive mutation and not adaptive amplification. *Molecular Cell*, *7*(3), 571–579.

Mckenzie, G. J., & Rosenberg, S. M. (2001). Adaptive mutations , mutator DNA polymerases and genetic change strategies of pathogens. *Current Opinion in Microbiology*, 586–594.

McVicker, G., Prajsnar, T. K., Williams, A., Wagner, N. L., Boots, M., Renshaw, S.

A., *et al*. (2014). Clonal Expansion during *Staphylococcus aureus* Infection Dynamics Reveals the Effect of Antibiotic Intervention. *PLoS Pathogens*, *10*(2), e1003959.

Mdluli, K., Kaneko, T., & Upton, A. (2015). The Tuberculosis Drug Discovery and Development Pipeline and Emerging Drug Targets. *Cold Spring Harbor Perspectives in Medicine*, *5*, a021154.

Mestre, O., Luo, T., Dos Vultos, T., Kremer, K., Murray, A., Namouchi, A., *et al*. (2011). Phylogeny of *Mycobacterium tuberculosis* Beijing strains constructed from Polymorphisms in genes involved in DNA replication, recombination and repair. *PLoS ONE*, *6*(1), 1–7.

Mikušová, K., & Ekins, S. (2016). Learning from the past for TB drug discovery in the future. *Drug Discovery Today*, *0*(0).

Milano, A., Pasca, M. R., Provvedi, R., Lucarelli, A. P., Manina, G., Luisa de Jesus Lopes Ribeiro, A., *et al*. (2009). Azole resistance in *Mycobacterium tuberculosis* is mediated by the MmpS5-MmpL5 efflux system. *Tuberculosis*, *89*(1), 84–90.

Mizrahi, V., & Andersen, S. J. (1998). DNA repair in *Mycobacterium tuberculosis*. What have we learnt from the genome sequence? *Molecular Microbiology*, *29*(6), 1331–1339.

Mokrousov, I. (2014). Widely-used laboratory and clinical *Mycobacterium tuberculosis* strains: To what extent they are representative of their phylogenetic lineages? *Tuberculosis*, *94*(3), 355–356.

Moolla, N., Goosens, V. J., Kana, B. D., & Gordhan, B. G. (2014). The contribution of Nth and nei DNA glycosylases to mutagenesis in *Mycobacterium smegmatis*. *DNA Repair*, *13*(1), 32–41.

Morisset, D., Štebih, D., Milavec, M., Gruden, K., Žel, J., Holst-Jensen, A., *et al*. (2013). Quantitative Analysis of Food and Feed Samples with Droplet Digital PCR. *PLoS ONE*, *8*(5), e62583.

Mosel, M., Li, L., Drlica, K., & Zhao, X. (2013). Superoxide-mediated protection of Escherichia coli from antimicrobials. *Antimicrobial Agents and Chemotherapy*, *57*(11), 5755–5759.

Mouton, J. W., Brown, D. F. J., Apfalter, P., Cantón, R., Giske, C. G., Ivanova, M., *et al*. (2012). The role of pharmacokinetics/pharmacodynamics in setting clinical MIC breakpoints: The EUCAST approach. *Clinical Microbiology and Infection*, *18*(3), E37–E45.

Movahedzadeh, F., Colston, M. J., & Davis, E. O. (1997). Determination of DNA Sequences required for regulated *Mycobacterium tuberculosis* RecA expression in response to DNA-Damaging agents suggests that two modes of regulation exist. *Journal of Bacteriology*, *179*(11), 3509–3518.

Namouchi, A., Gómez-Muñoz, M., Frye, S. A., Moen, L. V., Rognes, T., Tønjum, T., & Balasingham, S. V. (2016). The *Mycobacterium tuberculosis* transcriptional landscape under genotoxic stress. *BMC Genomics*, *17*(1), 791.

Nautiyal, A., Rani, P. S., Sharples, G. J., & Muniyappa, K. (2016). *Mycobacterium tuberculosis* RuvX is a Holliday junction resolvase formed by dimerisation of the monomeric YqgF nuclease domain. *Molecular Microbiology*, *100*(4), 656–674.

Neher, S. B., Flynn, J. M., Sauer, R. T., & Baker, T. A. (2003). Latent ClpX-recognition signals ensure LexA destruction after DNA damage. *Genes and Development*, *17*(9), 1084–1089.

Newton, G. L., Buchmeier, N., & Fahey, R. C. (2008). Biosynthesis and functions of mycothiol, the unique protective thiol of Actinobacteria. *Microbiology and Molecular Biology Reviews* , *72*(3), 471–494.

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

Nohmi, T., Battista, J. R., Dodson, L. A., & Walker, G. C. (1990). Dominant negative umuD mutations decreasing RecA-mediated cleavage suggest roles for intact UmuD in modulation of SOS mutagenesis. *Proceedings of the National Academy of Sciences*, *87*(18), 7190–7194.

O'Brien, P. J., & Ellenberger, T. (2004). The *Escherichia coli* 3-Methyladenine DNA Glycosylase AlkA Has a Remarkably Versatile Active Site. *Journal of Biological Chemistry*, *279*(26), 26876–26884.

O'Donnell, M. (2006). Replisome architecture and dynamics in *Escherichia coli*. *Journal of Biological Chemistry*, *281*(16), 10653–10656.

Ordonez, H., & Shuman, S. (2014). *Mycobacterium smegmatis* DinB2 misincorporates deoxyribonucleotides and ribonucleotides during templated synthesis and lesion bypass, *42*(20), 12722–12734.

Ordonez, H., Uson, M. L., & Shuman, S. (2014). Characterization of three mycobacterial DinB (DNA polymerase IV) paralogs highlights DinB2 as naturally adept at ribonucleotide incorporation, *42*(17), 11056–11070.

Ouellet, H., Johnston, J. B., & Ortiz de Montellano, P. R. (2010). The Mycobacterium tuberculosis cytochrome P450 system. *Archives of Biochemistry and Biophysics*, *493*(1), 82–95.

Pagès, V. (2016). Single-strand gap repair involves both RecF and RecBCD pathways. *Current Genetics*, *62*(3), 519–521.

Paris, Ü., Mikkel, K., Tavita, K., Saumaa, S., Teras, R., & Kivisaar, M. (2015). NHEJ enzymes LigD and Ku participate in stationary-phase mutagenesis in Pseudomonas putida. *DNA Repair*, *31*, 11–18.

Parish, T. and Stoker, N. G. (2000). Use of a flexible cassette method to generate a double unmarked *Mycobacterium tuberculosis* tlyA plcABC mutant by gene replacement. *Microbiology 146*(8), 1969- 1975.

Park, Y., Pacitto, A., Bayliss, T., Cleghorn, L. A. T., Wang, Z., Hartman, T., *et al*. (2016). Essential but Not Vulnerable: Indazole Sulfonamides Targeting Inosine Monophosphate Dehydrogenase as Potential Leads against *Mycobacterium tuberculosis*. *ACS Infectious Diseases*, *3*, 18-33.

Park, C., & Raines, R. T. (2000). Dimer formation by a "monomeric" protein. *Protein Science*, *9*(10), 2026–2033.

Pasipanodya, J. G., McNabb, S. J., Hilsenrath, P., Bae, S., Lykens, K., Vecino, E., *et al*. (2010). Pulmonary impairment after tuberculosis and its contribution to TB burden. *BMC Public Health*, *10*, 259.

Patiño, S., Alamo, L., Cimino, M., Casart, Y., Bartoli, F., García, M. J., & Salazar, L. (2008). Autofluorescence of mycobacteria as a tool for detection of *Mycobacterium tuberculosis*. *Journal of Clinical Microbiology*, *46*(10), 3296–302.

Pepper, D. J., Meintjes, G. A., McIlleron, H., & Wilkinson, R. J. (2007). Combined therapy for tuberculosis and HIV-1: the challenge for drug discovery. *Drug Discovery Today*, *12*(21–22), 980–989.

Pérez-Lago, L., Comas, I., Navarro, Y., González-Candelas, F., Herranz, M., Bouza, E., & García-De-Viedma, D. (2014). Whole Genome Sequencing Analysis of Intrapatient Microevolution in *Mycobacterium tuberculosis*: Potential Impact on the Inference of Tuberculosis Transmission. *Journal of Infectious Diseases*, *209*(1), 98–108.

Philips, J. A., & Ernst, J. D. (2012). Tuberculosis Pathogenesis and Immunity. *Annual Review of Pathology: Mechanisms of Disease*, *7*(1), 353–384.

Prabha, S., Rao, D. N., & Nagaraja, V. (2011). Distinct Properties of Hexameric but Functionally Conserved *Mycobacterium tuberculosis* Transcription- Repair Coupling Factor. *PLoS ONE 6*(4), e19131.

Prabowo, S. A., Groschel, M. I., Schmidt, E. D. L., Skrahina, A., Mihaescu, T., Hasturk, S., *et al*. (2013). Targeting multidrug-resistant tuberculosis (MDR-TB) by therapeutic vaccines. *Medical Microbiology and Immunology*, *202*(2), 95–104.

Puri, R. V., Singh, N., Gupta, R. K., & Tyagi, A. K. (2013). Endonuclease IV Is the Major Apurinic / Apyrimidinic Endonuclease in *Mycobacterium tuberculosis* and Is Important for Protection against Oxidative Damage. *PLoS ONE*, *8*(8), e71535.

Qin, T., Kang, H., Ma, P., Li, P., Huang, L., & Gu, B. (2015). SOS response and its regulation on the fluoroquinolone resistance. *Annals of Translational Medicine*, *3*(22), 358.

Rachman, H., Strong, M., Ulrichs, T., Grode, L., Schuchhardt, J., Mollenkopf, H., *et al*. (2006). Unique Transcriptome Signature of *Mycobacterium tuberculosis* in Pulmonary Tuberculosis. *Infection and Immunity*, *74*(2), 1233–1242.

Radman, M. (1975). SOS repair hypothesis: phenomenology of an inducible DNA repair which is accompanied by mutagenesis. *Basic Life Sciences*, *5A*, 355–367.

Rand, L., Hinds, J., Springer, B., Sander, P., Buxton, R. S., & Davis, E. O. (2003). The majority of inducible DNA repair genes in *Mycobacterium tuberculosis* are induced independently of RecA. *Molecular Microbiology*, *50*(3), 1031–1042.

Rangarajan, S., Woodgate, R., & Goodman, M. F. (2002). Replication restart in UV-irradiated *Escherichia coli* involving pols II, III, V, PriA, RecA and RecFOR proteins. *Molecular Microbiology*, *43*(3), 617–628.

Rauch, P. J. G., Palmen, R., Burds, A. A., Gregg-Jolly, L. A., Van Der Zee, J. R., & Hellingwerf, K. J. (1996). The expression of the *Acinetobacter calcoaceticus recA* gene increases in response to DNA damage independently of RecA and of development of competence for natural transformation. *Microbiology*, *142*(4), 1025–1032.

Ravikumar, B., Vacher, C., Berger, Z., Davies, J. E., Luo, S., Oroz, L. G., *et al*. (2004). Inhibition of mTOR induces autophagy and reduces toxicity of polyglutamine expansions in fly and mouse models of Huntington disease. *Nature Genetics*, *36*(6), 585–595.

Reardon, J. T., & Sancar, A. (2005). Nucleotide Excision Repair. *Progress in Nucleic Acid Research and Molecular Biology*, *79*, 183-235.

Reha-Krantz, L. J. (2010). DNA polymerase proofreading: Multiple roles maintain genome stability. *Biochimica et Biophysica Acta - Proteins and Proteomics*, *1804*(5), 1049-1063.

Reynolds, N. M., Lazazzera, B. A., & Ibba, M. (2010). Cellular mechanisms that control mistranslation. *Nature Reviews. Microbiology*, *8*(12), 849–56.

Rivas-Santiago, C. E., Hernández-Pando, R., & Rivas-Santiago, B. (2013). Immunotherapy for pulmonary TB: antimicrobial peptides and their inducers. *Immunotherapy*, *5*(10), 1117–1126.

Roberts, J., & Park, J.-S. (2004). Mfd, the bacterial transcription repair coupling factor: translocation, repair and termination. *Current Opinion in Microbiology*, *7*(2), 120–125.

Robleto, E. A., Yasbin, R., Ross, C., & Pedraza-Reyes, M. (2007). Stationary Phase Mutagenesis in *B. subtilis* : A Paradigm to Study Genetic Diversity Programs in Cells Under Stress. *Critical Reviews in Biochemistry and Molecular Biology*, *42*(5), 327–339.

Rock, J. M., Hopkins, F. F., Chavez, A., Diallo, M., Chase, M. R., Gerrick, E. R., *et al*. (2017). Programmable transcriptional repression in mycobacteria using an orthogonal CRISPR interference platform. *Nature Microbiology*, *2*, 16274.

Rock, J. M., Lang, U. F., Chase, M. R., Ford, C. B., Gerrick, E. R., Gawande, R., *et al*. (2015). DNA replication fidelity in *Mycobacterium tuberculosis* is mediated by an ancestral prokaryotic proofreader. *Nature Genetics*, *47*(6), 677–681.

Rodgers, K., & Mcvey, M. (2016). Error-Prone Repair of DNA Double-Strand Breaks. *Journal of Cellular Physiology*, *231*(1), 15–24.

Rosenberg, S., & Hastings, P. (2004). Adaptive point mutation and adaptive amplification pathways in the *Escherichia coli* Lac system: stress responses producing genetic change. *Journal of Bacteriology*, *186*(15), 4838–4843.

Rosenberg, S. M. (2001). Evolving responsively: adaptive mutation. *Nature Reviews. Genetics*, *2*(7), 504–515.

Ross, C., Pybus, C., Pedraza-Reyes, M., Sung, H. M., Yasbin, R. E., & Robleto, E. (2006). Novel role of mfd: Effects on stationary-phase mutagenesis in *Bacillus subtilis*. *Journal of Bacteriology*, *188*(21), 7512–7520.

Roth, J. R., Kugelberg, E., Reams, A. B., Kofoid, E., & Andersson, D. I. (2006). Origin of mutations under selection: the adaptive mutation controversy. *Annual Review of Microbiology*, *60*, 477–501.

Roth, J. R. (2010). Genetic Adaptation: A New Piece for a Very Old Puzzle. *Current Biology*, *20*(1), R15–R17.

Roxas, B. A. P., & Li, Q. (2009). Acid stress response of a mycobacterial proteome: Insight from a gene ontology analysis. *International Journal of Clinical and Experimental Medicine*, *2*(4), 309–328.

Rubbo, H., Radi, R., Trujillo, M., Telleri, R., Kalyanaraman, B., Barnes, S., *et al*. (1994). Nitric oxide regulation of superoxide and peroxynitrite-dependent lipid peroxidation. Formation of novel nitrogen-containing oxidized lipid derivatives. *The Journal of Biological Chemistry*, *269*(42), 26066–75.

Russell, D. G. (2001). *Mycobacterium tuberculosis*: here today, and here tomorrow. *Nature reviews Molecular Cell Biology*, *2*(8), 569–577.

Sambrook, J., Fritsch, E. and Maniatis, T. (1989). Molecular cloning: a laboratory manual, Cold Spring Laboratory. Cold Spring Harbor, NY. VIII. Appendix A. pBIND Vector Sequence (continued) A. pBIND Vector Sequence (continued) B. pBIND Vector Restriction Sites Enzyme# of Sites Location Dra I 4(1857): 4877.

Sambrook, J. and Russell, D. (2001). Molecular Cloning: A Laboratory Manual 2001 3rd edition Cold Spring Harbor. NY Cold Spring Harbor Laboratory.

Sandgren, A., Strong, M., Muthukrishnan, P., Weiner, B. K., Church, G. M., & Murray, M. B. (2009). Tuberculosis drug resistance mutation database. *PLoS Medicine*, *6*(2), 0132–0136.

Sandhu, P., & Akhter, Y. (2015). The internal gene duplication and interrupted coding sequences in the MmpL genes of *Mycobacterium tuberculosis*: Towards understanding the multidrug transport in an evolutionary perspective. *International Journal of Medical Microbiology*, *305*(3), 413–423.

Saumaa, S., Tover, A., Tark, M., Tegova, R., & Kivisaar, M. (2007). Oxidative DNA damage defense systems in avoidance of stationary-phase mutagenesis in *Pseudomonas putida*. *Journal of Bacteriology*, *189*(15), 5504–5514.

Saunders, N. J., Trivedi, U. H., Thomson, M. L., Doig, C., Laurenson, I. F., & Blaxter, M. L. (2011). Deep resequencing of serial sputum isolates of *Mycobacterium tuberculosis* during therapeutic failure due to poor compliance reveals stepwise mutation of key resistance genes on an otherwise stable genetic background. *Journal of Infection*, *62*(3), 212–217.

Saxowsky, T. T., & Doetsch, P. W. (2006). RNA polymerase encounters with DNA

damage: Transcription-coupled repair or transcriptional mutagenesis? *Chemical Reviews*, *106*(2), 474–488.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods*, *9*(7), 676–682.

Schnappinger, D., Ehrt, S., Voskuil, M. I., Liu, Y., Mangan, J. A., Monahan, I. M.,*et al*. (2003). Transcriptional adaptation of *Mycobacterium tuberculosis* within macrophages : Insights into the phagosomal environment. *Journal of Experimental Medicine*, *198*(5), 693–704.

Scriba, T. J., Kaufmann, S. H. E., Lambert, P. H., Sanicas, M., Martin, C., & Neyrolles, O. (2016). Vaccination against tuberculosis with whole-cell mycobacterial vaccines. *Journal of Infectious Diseases*, *214*(5), 659–664.

Sherman, D., & Gagneux, S. (2011). Estimating the mutation rate of *Mycobacterium tuberculosis* during infection. *Nature Genetics*, *43*(5), 1–3.

Shinagawa, H., Iwasaki, H., Kato, T., & Nakata, A. (1988). RecA protein-dependent cleavage of UmuD protein and SOS mutagenesis. *Proceedings of the National Academy of Sciences of the United States of America*, *85*(6), 1806–1810.

Simmons, L. A., Foti, J. J., Cohen, S. E., & Walker, G. C. (2008). The SOS Regulatory Network. *EcoSal Plus – Escherichia coli* and *Salmonella*: Cellular and Molecular Biology, A. Bock III, R. C. Kaper, J. B. Karp *et al*., Eds., chapter 5.4.3., ASM Press, Washington, DC, USA, 2008.

Singh, A., Bhagavat, R., Vijayan, M., & Chandra, N. (2016). A comparative analysis of the DNA recombination repair pathway in mycobacterial genomes. *Tuberculosis*, *99*, 109–119.

Singh, V., Donini, S., Pacitto, A., Sala, C., Hartkoorn, R. C., Dhar, N., *et al*. (2016). The Inosine Monophosphate Dehydrogenase, GuaB2, Is a Vulnerable New Bactericidal Drug Target for Tuberculosis. *ACS Infectious Diseases*, *2007*(3), 5-17.

Sinha, K. M., Unciuleac, M., Glickman, M. S., Niu, H., Raynard, S., Sung, P., & Shuman, S. (2009). AdnAB : a new DSB-resecting motor − nuclease from mycobacteria. *Genes & Development*, *23*, 1423–1437.

Smith, K. C. (2004). Recombinational DNA repair: the ignored repair systems. *BioEssays*, *26*(12), 1322–1326.

Smith, P. A., Romesberg, F. E. (2007). Combating bacteria and drug resistance by

inhibiting mechanisms of persistence and adaptation. *Nature Chemical Biology*, *3*(9), 549-556.

Smollett, K. L., Smith, K. M., Kahramanoglou, C., Arnvig, K. B., Buxton, R. S., & Davis, E. O. (2012). Global analysis of the regulon of the transcriptional repressor LexA , a key component of SOS response in *Mycobacterium tuberculosis*. *The Journal of Biological Chemistry*, *287*(26), 22004–22014.

Snapper, S. B., Melton, R. E., Mustafa, S., Kieser, T., & Jacobs, W. R. (1990). Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Molecular Microbiology*, *4*(11), 1911–1919.

Sonnenberg, P., Glynn, J. R., Fielding, K., Murray, J., Godfrey-Faussett, P., & Shearer, S. (2005). How soon after infection with HIV does the risk of tuberculosis start to increase? A retrospective cohort study in South African gold miners. *Journal of Infectious Diseases*, *191*(2), 150–158.

Springer, B., Sander, P., Sedlacek, L., Hardt, W.-D., Mizrahi, V., Schär, P., & Böttger, E. C. (2004). Lack of mismatch correction facilitates genome evolution in mycobacteria. *Molecular Microbiology*, *53*(6), 1601–1609.

Springman, E. B., Angleton, E. L., Birkedal-Hansen, H., & Van Wart, H. E. (1990). Multiple modes of activation of latent human fibroblast collagenase: evidence for the role of a Cys73 active-site zinc complex in latency and a “cysteine switch” mechanism for activation. *Proceedings of the National Academy of Sciences*, *87*(1), 364–368.

Spruijt, C. G., Gnerlich, F., Smits, A. H., Pfaffeneder, T., Jansen, P. W. T. C., Bauer, C., M. (2013). Dynamic Readers for 5-(Hydroxy)Methylcytosine and Its Oxidized Derivatives. *Cell*, *152*, 1146–1159.

Stallings, C. L., & Glickman, M. S. (2010). Is *Mycobacterium tuberculosis* stressed out? A critical assessment of the genetic evidence. *Microbes and Infection*, *12*(14), 1091–1101.

Stephanou, N. C., Gao, F., Bongiorno, P., Ehrt, S., Schnappinger, D., Shuman, S., & Glickman, M. S. (2007). Mycobacterial nonhomologous end joining mediates mutagenic repair of chromosomal double-strand DNA breaks. *Journal of Bacteriology*, *189*(14), 5237–5246.

Stucki, D., & Gagneux, S. (2013). Single nucleotide polymorphisms in *Mycobacterium tuberculosis* and the need for a curated database. *Tuberculosis*, *93*(1), 30–39.

Stumpf, J. D., Poteete, A. R., & Foster, P. L. (2007). Amplification of lac cannot

account for adaptive mutation to Lac+ in *Escherichia coli*. *Journal of Bacteriology*, *189*(6), 2291–2299.

Sun, G., Luo, T., Yang, C., Dong, X., Li, J., Zhu, Y., *et al*. (2012). Dynamic population changes in *Mycobacterium tuberculosis* during acquisition and fixation of drug resistance in patients. *The Journal of Infectious Diseases*, *206*(11), 1724–1733.

Sung, H. M., & Yasbin, R. E. (2002). Adaptive, or stationary-phase, mutagenesis, a component of bacterial differentiation in *Bacillus subtilis*. *Journal of Bacteriology*, *184*(20), 5641–5653.

Szabó, C. (2003). Multiple pathways of peroxynitrite cytotoxicity. *Toxicology Letters*, *140*, 105–112.

Szumowski, J. D., Adams, K. N., Edelstein, P. H., & Ramakrishnan, L. (2013). Antimicrobial efflux pumps and *Mycobacterium tuberculosis* drug tolerance: Evolutionary considerations. *Current Topics in Microbiology and Immunology*, *374*, 81–108.

Tagel, M., Tavita, K., Horak, R., Kivisaar, M., & Ilves, H. (2016). A novel papillation assay for the identification of genes affecting mutation rate in *Pseudomonas putida* and other pseudomonads. *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*, *790*, 41–55.

Talaat, A. M., Lyons, R., Howard, S. T., & Johnston, S. A. (2004). The temporal expression profile of *Mycobacterium tuberculosis* infection in mice. *Proceedings of the National Academy of Sciences*, *101*(13), 4602–4607.

Tanaka, Y., Yazawa, K., Dabbs, E. R., Nishikawa, K., Komaki, H., Mikami, Y.,*et al*. (1996). Different Rifampicin Inactivation Mechanisms in Nocardia and Related Taxa. *Microbiology and Immunology*, *40*(1), 1–4.

Taylor, N. G. H., Verner-Jeffreys, D. W., & Baker-Austin, C. (2011). Aquatic systems: Maintaining, mixing and mobilising antimicrobial resistance? *Trends in Ecology and Evolution*, *26*(6), 278–284.

Tegova, R., Tover, A., Tarassova, K., Tark, M., & Kivisaar, M. (2004). Involvement of error-prone DNA polymerase IV in stationary-phase mutagenesis in *Pseudomonas putida*. *Journal of Bacteriology*, *186*(9), 2735–2744.

Tham, K. C., Kanaar, R., & Lebbink, J. H. G. (2016). Mismatch repair and homeologous recombination. *DNA Repair*, *38*, 75–83.

Truglio, J. J., Croteau, D. L., Van Houten, B., & Kisker, C. (2006). Prokaryotic Nucleotide Excision Repair: The UvrABC System. *Chemical Reviews*, *106*(2),

233–252.

Trunz, B. B., Fine, P., & Dye, C. (2006). Effect of BCG vaccination on childhood tuberculous meningitis and miliary tuberculosis worldwide: a meta-analysis and assessment of cost-effectiveness. *Lancet*, *367*(9517), 1173–1180.

van der Veen, S., & Tang, C. M. (2015). The BER necessities: the repair of DNA damage in human-adapted bacterial pathogens. *Nature Reviews Microbiology*, *13*(2), 83–94.

Venkatesh, J., Kumar, P., Sai, P., Krishna, M., Manjunath, R., & Varshney, U. (2003). Importance of Uracil DNA Glycosylase in *Pseudomonas aeruginosa* and *Mycobacterium smegmatis* , G + C-rich Bacteria , in Mutation Prevention , Tolerance to Acidified Nitrite, and Endurance in Mouse Macrophages. *The Journal of Biological Chemistry*, *278*(27), 24350–24358.

Verma, S., Dixit, R., & Pandey, K. C. (2016). Cysteine proteases: Modes of activation and future prospects as pharmacological targets. *Frontiers in Pharmacology*, *7*, article 107.

Vidales, L. E., Cárdenas, L. C., Robleto, E., Yasbin, R. E., & Pedraza-Reyes, M. (2009). Defects in the error prevention oxidized guanine system potentiate stationary-phase mutagenesis in *Bacillus subtilis*. *Journal of Bacteriology*, *191*(2), 506–513.

Vitreschak, A. G., Rodionov, D. A, Mironov, A. A, & Gelfand, M. S. (2003). Regulation of the vitamin B12 metabolism and transport in bacteria by a conserved RNA structural element. *RNA*, *9*(9), 1084–1097.

Voskuil, M., Lamichhane, G., & Hopkins, J. (2011). The response of *Mycobacterium tuberculosis* to reactive oxygen and nitrogen species. *Frontiers in Microbiology*, *2*(105), 1–12.

Wakamoto, Y., Dhar, N., Chait, R., Schneider, K., Signorino-Gelo, F., Leibler, S., & McKinney, J. D. (2013). Dynamic persistence of antibiotic-stressed mycobacteria. *Science*, *339*, 91–95.

Walker, T. M., Ip, C. L. C., Harrell, R. H., Evans, J. T., Kapatai, G., Dedicoat, M. J., *et al*. (2013). Whole-genome sequencing to delineate *Mycobacterium tuberculosis* outbreaks: A retrospective observational study. *The Lancet Infectious Diseases*, *13*(2), 137–146.

Wallis, R. S., & Hafner, R. (2015). Advancing host-directed therapy for tuberculosis. *Nature Reviews in Immunology*, *15*(4), 255–263.

Wang, R. Y., R., Kudryashev, M., Li, X., Egelman, E. H., Basler, M., Cheng, Y., *et al*.

(2015). De novo protein structure determination from near-atomic-resolution cryo-EM maps. *Nature Methods*, *12*(4), 335–338.

Wang, Y., Huang, Y., Xue, C., He, Y., & He, Z. G. (2011). ClpR protein-like regulator specifically recognizes RecA protein-independent promoter motif and broadly regulates expression of DNA damage-inducible genes in mycobacteria. *Journal of Biological Chemistry*, *286*(36), 31159–31167.

Wards, B. J. and Collins, D. M. (1996). Electroporation at elevated temperatures substantially improves transformation efficiency of slow-growing mycobacteria. F*EMS Microbiology Letters, 145*(1), 101-105.

Wanner, R. M., Guethlein, C., Springer, B., Boettger, E. C., Ackermann, M., Cole, S., *et al*. (2008). Stabilization of the genome of the mismatch repair deficient *Mycobacterium tuberculosis* by context-dependent codon choice. *BMC Genomics*, *9*(1), 249.

Warner, D. F. (2010). The role of DNA repair in *M. tuberculosis* pathogenesis. *Drug Discovery Today: Disease Mechanisms*, *7*(1), e5–e11.

Warner, D. F., Evans, J. C., & Mizrahi, V. (2014). Nucleotide Metabolism and DNA Replication. *Microbiology Spectrum*, *2*(5), MGM2- 0001-2013.

Warner, D. F., Koch, A., & Mizrahi, V. (2015). Diversity and disease pathogenesis in *Mycobacterium tuberculosis*. *Trends in Microbiology*, *23*(1), 14-21.

Warner, D. F., & Mizrahi, V. (2006). Tuberculosis chemotherapy: The influence of bacillary stress and damage response pathways on drug efficacy. *Clinical Microbiology Reviews*, *19*(3), 558-570.

Warner, D. F., & Mizrahi, V. (2011). Making ends meet in mycobacteria. *Molecular Microbiology*, *79*(2), 283–287.

Warner, D. F., Ndwandwe, D. E., Abrahams, G. L., Kana, B. D., Machowski, E. E., Venclovas, C., & Mizrahi, V. (2010). Essential roles for *imuA*'- and *imuB*-encoded accessory factors in DnaE2-dependent mutagenesis in *Mycobacterium tuberculosis*. *Proceedings of the National Academy of Sciences*, *107*(29), 13093–13098.

Warner, D. F., Savvi, S., Mizrahi, V., & Dawes, S. S. (2007) A Riboswitch Regulates Expression of the Coenzyme B 12 -Independent Methionine Synthase in *Mycobacterium tuberculosis* : Implications for Differential Methionine Synthase Function in Strains H37Rv andn CDC1551. *Journal of Bacteriology*, *189*(9), 3655–3659.

Warner, D., Tønjum, T., & Mizrahi, V. (2013). DNA Metabolism in Mycobacterial Pathogenesis. *Current Topics in Microbiology and Immunology*, 374, 27-51.

Warrier, T., Kapilashrami, K., Argyrou, A., Ioerger, T. R., Little, D., Murphy, K. C., *et al*. (2016). N-methylation of a bactericidal compound as a resistance mechanism in *Mycobacterium tuberculosis*. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(31), E4523-30.

Weller, G. R., Kysela, B., Roy, R., Tonkin, L. M., Scanlan, E., Della, M., *et al*. (2002). Identification of a DNA nonhomologous end-joining complex in bacteria. *Science*, *297*(5587), 1686–1689.

Werngren, J., Werngren, J., Hoffner, S. E., & Hoffner, S. E. (2003). *Mycobacterium tuberculosis*. *Journal of Clinical Microbiology*, *41*(4), 1520–1524.

West, S. C. (1995). Formation, translocation and resolution of Holliday junctions during homologous genetic recombination. *Philosophical Transactions of the Royal Society of London*, *347*(1319), 21–25.

Wigle, T. J., Sexton, J. Z., Gromova, A. V., Hadimani, M. B., Hughes, M. A., Smith, G. R., Yeh, L., Singleton, S. F. (2009). Inhibitors of RecA activity discovered by high throughput screening: Cell permeable small molecules attenuate the SOS response in *Escherichia coli*. *Journal of Biomolecular Screening*, *7*(1), 1092-1101.

Wigley, D. B. (2013). Bacterial DNA repair: recent insights into the mechanism of RecBCD, AddAB and AdnAB. *Nature Review. Microbiology*, *11*(1), 9–13.

Wilkinson, M., Chaban, Y., & Wigley, D. B. (2016). Mechanism for nuclease regulation in RecBCD. *eLife*, *5*, 1–13.

Williams, A. B., & Schumacher, B. (2016). DNA damage responses and stress resistance: Concepts from bacterial SOS to metazoan immunity. *Mechanisms of Ageing and Development*, 1–6.

Witney, A. A., Cosgrove, C. A., Arnold, A., Hinds, J., Stoker, N. G., & Butcher, P. D. (2016). Clinical use of whole genome sequencing for *Mycobacterium tuberculosis*. *BMC Medicine*, 1–7.

Wong, C., Ha, N. P., Pawlowski, M. E., Graviss, E. A., & Tkaczyk, T. S. (2016). Differentiating between live and dead *Mycobacterium smegmatis* using autofluorescence. *Tuberculosis*, *101*, 1–5.

Wrande, M., Roth, J. R., & Hughes, D. (2008). Accumulation of mutants in "aging" bacterial colonies is due to growth under selection, not stress-induced mutagenesis. *Proceedings of the National Academy of Sciences*, *105*(33), 11863–

11868.

Yang, W. (2003). Damage repair DNA polymerases Y. *Current Opinion in Structural Biology*, *13*(1), 23–30.

Yang, W., & Woodgate, R. (2007). What a difference a decade makes: insights into translesion DNA synthesis. *Proceedings of the National Academy of Sciences*, *104*(40), 15591–15598.

Yermilov, V., Rubio, J., Becchi, M., Friesen, M. D., Pignatelli, B., & Ohshima, H. (1995). Formation of 8-nitroguanine by the reaction of guanine with peroxynitrite *in vitro*. *Carcinogenesis*, *16*(9), 2045–2050.

Yi, C., & He, C. (2013). DNA Repair by Reversal of DNA Damage. *Cold Spring Harbor Perspectives in Biology*, *5*(1), a012575–a012575.

Yin, J., Yuan, J., Hu, Y., & Wei, X. (2016). Association between directly observed therapy and treatment outcomes in multidrug-resistant tuberculosis: A systematic review and meta-analysis. *PLoS ONE*, *11*(3), 1–14.

Yosef, I., Edgar, R., Levy, A., Amitai, G., Sorek, R., Munitz, A., & Qimron, U. (2016). Natural selection underlies apparent stress-induced mutagenesis in a bacteriophage infection model. *Nature Microbiology Letters*, (April), 16047.

Zhang, Q., Lambert, G., Liao, D., Kim, H., Robin, K., Tung, C., *et al*. (2011). Acceleration of Emergence of Bacterial Antibiotic Resistance in Connected Microenvironments. *Science*, *333*, 1764–1767.

Zhang, J., Wang, Q., Barz, B., He, Z., Kosztin, L., Shang, Y., & Xu, D. (2010). MUFOLD: A new solution for protein 3D structure prediction. *Proteins: Structure, Function and Bioinformatics*, *78*(5), 1137–1152.

Zhu, H., Bhattarai, H., Yan, H.-G., Shuman, S., & Glickman, M. S. (2012). Characterization of *Mycobacterium smegmatis* PolD2 and PolD1 as RNA/DNA polymerases homologous to the POL domain of bacterial DNA ligase D. *Biochemistry*, *51*(51), 10147–10158.

Zumla, A., Chakaya, J., Hoelscher, M., Ntoumi, F., Rustomjee, R., Vilaplana, C., *et al*. (2015). Towards host-directed therapies for tuberculosis. *Nature Reviews. Drug Discovery*, *14*(8), 511–512.