

https://dspace.org/

# Molecular basis of the interplay between the Nth and Nei DNA glycosylases of the base excision repair pathway in Mycobacterium smegmatis

**Rantsi, Tebogo Christina**

**2017**
https://hdl.handle.net/10539/24787

Downloaded from a DSpace repository. https://dspace.org/

# Molecular basis of the interplay between the Nth and Nei DNA glycosylases of the base excision repair pathway in *Mycobacterium smegmatis*

**Tebogo Christina Rantsi**

**(447626)**

A thesis submitted to the Faculty of Health Science, University of the Witwatersrand, Johannesburg, in the fulfilment of the requirements of the degree of Master of Science in Medicine.

2017

**Declaration**

I declare that this dissertation is my own, unaided work. It is being submitted for the degree of Master of Science in Medicine at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

_____ _____

Tebogo C. Rantsi Date

**Presentations from this study**

Tebogo Rantsi, Bavesh Kana and Bhavna Gordhan

Molecular basis of the interplay between the Nth and Nei DNA glycosylases of the base excision repair pathway in *Mycobacterium smegmatis*

Wits Faculty of Health Sciences Research Day and PG Expo 2016

01 September 2016

University of Witwatersrand, Medical School, Johannesburg

Poster presentation

Tebogo Rantsi, Bavesh Kana and Bhavna Gordhan

Molecular basis of the interplay between the Nth and Nei DNA glycosylases of the base excision repair pathway in *Mycobacterium smegmatis*

Molecular and biosciences research thrust 2015

03 December 2015

University of Witwatersrand, West Campus, Johannesburg

Poster presentation

**Abstract**

During infection, *Mycobacterium tuberculosis* encounters hostile conditions which result in the generation of host-derived reactive oxygen (ROS) and nitrogen species (RNS) as part of the immune response to control the infection. Exposure to these reactive radicals can lead to oxidative damage of DNA, which ultimately destabilises the genome and introduces mutations. However, *M. tuberculosis* is well equipped with a number of DNA repair pathways such as the base excision repair (BER) pathway, which plays a role in maintaining genome stability and survival of the pathogen. A number of DNA glycosylases are involved in the BER pathway, including formamidopyrimidine (Fpg), endonuclease VIII (Nei) and endonuclease III (Nth), which are the initial enzymes responsible for recognition and excision of damaged DNA bases. It was previously demonstrated that combinatorial deletion of *nth* and two *nei* homologues in *Mycobacterium smegmatis* resulted in reduced survival under oxidative stress conditions with a corresponding increase in mutation rates, suggestive of interplay between these enzymes. To understand the molecular basis of this interplay, the individual effects of the Nei homologues (NeiI and NeiII), together with Nth on survival and mutagenesis under oxidative stress conditions, expected to induce DNA damage, were investigated in the current study. Two mutants lacking *nth* and either *neiI* or *neiII* were generated by homologous recombination. These double deletion mutants together with the individual deletion mutants, the parental strain and the respective complemented strains were phenotypically characterized under oxidative stress conditions and assessed for increased mutagenesis as measured by rifampicin resistance. Defects in the BER system resulted in reduced survival under oxidative stress conditions. Deletion of *nth* combined with the *neiII* homologue led to reduced survival under oxidative stress conditions and an increase in spontaneous mutagenesis to rifampicin when compared to the deletion of *nth* combined with the *neiI* homologue. Collectively these data suggest that NeiII may play an important physiological role in BER in comparison to the NeiI homologue.

**Acknowledgements**

I would like to sincerely thank God the Almighty for the strength and courage.

It has been an honour and privilege to be awarded funding from the National Research Foundation (NRF) and the DST-NRF Centre of Excellence for Biomedical TB Research. The following funding is acknowledged to Dr Bhavna Gordhan: The South African Tuberculosis and AIDS Training (SATBAT) program (1U2RTW007370/3), NHLS Research Trust Grant and the NRF incentive funding.

To my supervisor Dr Bhavna Gordhan, thank you for your constant encouragement and guidance. Your intelligence and leadership is invaluable. I appreciate your patient guidance, good judgment and useful critiques in this research project. I would like to extend my sincere gratitude to Prof Bavesh for the opportunity and unique experience. Being a part of this institution has led me on a path of self-discovery and personal growth which will hopefully put me in good stead of becoming a good scientist one day. Thank you for your valuable time and constructive suggestions.

Heartfelt thanks to all my colleagues (past and present) and support staff at the Centre of Excellence for Biomedical TB Research for your assistance and kindness. It has been a motivational and a rewarding journey.

A special thanks to my mother and my father for their support and encouragement. Lastly I would like to express my appreciation to my family and friends for their endless, support.

## Contents

Declaration ..... 2
Presentations from this study ..... 3
Abstract ..... 4
Acknowledgements ..... 5
Nomenclature ..... 9
List of figures ..... 11
List of Tables ..... XIII
1. Introduction ..... 1
1.1 Tuberculosis ..... 1
1.3 Pathogenesis ..... 8
1.4 Host cell activity in response to entry of *M. tuberculosis* ..... 11
1.5 DNA damage ..... 12
1.5.1 Endogenous Stresses ..... 12
1.5.1.1 Hydrolytic DNA damage ..... 12
1.5.1.2 Oxidative DNA damage ..... 13
1.5.2 Exogenous stress ..... 14
1.5.2.1 Alkylation damage ..... 14
1.5.2.2 Ionizing Radiation-induced Damage ..... 15
1.5.2.3 UV induced damage ..... 15
1.6 DNA repair pathways ..... 16
1.6.1 Mismatch repair ..... 16
1.6.2 Recombination repair ..... 17
1.6.3 Nucleotide excision repair ..... 18
1.7 Base excision repair pathway ..... 20
1.9 DNA glycosylases ..... 22
1.10 Preliminary studies in mycobacteria that are important for the work presented in this dissertation ..... 22
1.11 Aims and objectives of the study ..... 24
2. Materials and methods ..... 26
2.1 Bacterial strains, plasmids and culturing conditions ..... 26
2.1.1 Bioinformatics analysis ..... 28
2.1.2 *E. coli*- DH5α strains ..... 28
2.1.3 Mycobacterial strains ..... 29

2.2.1 Plasmid DNA extraction from *E. coli* ..... 29
2.2.1.1 Small scale plasmid DNA extraction ..... 29
2.2.1.2 Large scale plasmid DNA extraction ..... 30
2.2.2 Chromosomal DNA extraction ..... 30
2.2.2.1 Small scale chromosomal DNA extraction (colony boil) ..... 30
2.2.2.2 Large scale chromosomal DNA extraction ..... 30
2.3.1 Polymerase chain reaction ..... 31
2.3.1.1 Primer design ..... 31
2.3.1.2 PCR reactions ..... 31
2.3.2 Restriction enzyme digestions ..... 31
2.3 Agarose gel electrophoresis ..... 32
2.4 Bacterial Transformations ..... 32
2.4.1 Preparations of chemically competent *E coli*-DH5α cells ..... 32
2.4.2 Transformations into *E. coli*-DH5α competent cells ..... 32
2.4.3 Preparation of competent *M. smegmatis* cells ..... 33
2.4.4 Electroporation into competent *M. smegmatis* cells ..... 33
2.5 Southern Blot analysis (Southern, 1975) ..... 33
2.6 Knockout vector ..... 34
2.7 Inactivation of the Nth DNA glycosylase to generate double deletion mutants ..... 36
2.7.1 Generation of SCO and DCO transformants ..... 37
2.8 Complementation of mutant strains ..... 38
2.9 Phenotypic analysis of mutants ..... 39
2.9.1 Growth Kinetics under normal culture conditions ..... 39
2.9.2 Mutation frequency ..... 39
2.9.3 Sensitivity to hydrogen peroxide ..... 39
2.9.4 Spontaneous mutation rates ..... 39
3. Results ..... 43
3.1 Bioinformatics analysis ..... 43
3.2 Confirmation of pre-existing mutants ..... 49
3.2.1 Confirmation of the *ΔneiI* mutant strain ..... 49
3.2.2 Confirmation of the Δ*neiII* mutant strain ..... 50
3.3 Confirmation of previously constructed knockout vector ..... 50
3.4 Generation of double deletion mutants deficient in the Nth and Nei DNA glycosylases ..... 52
3.4.1 Isolation of single cross over single cross over (SCO) and double cross over (DCO) ..... 52

3.4.2 Confirmation of double deletion mutants by PCR ..... 52
3.4.3 Southern blot confirmation of double mutants ..... 54
3.4.4 Genetic complementation of double mutant strains ..... 56
3.5 Phenotypic characterization ..... 59
3.5.1 Growth kinetics ..... 59
3.5.2 Sensitivity to hydrogen peroxide ..... 60
3.5.3 Assessment of mutation rates by the fluctuation assay ..... 64
4. Discussion and Conclusion ..... 66
5. References ..... 73
6. Appendices ..... 80

## Nomenclature

| | |
|---|---|
| $^{\cdot}OH$ | Hydroxyl radicals |
| 8OxoG | 8-oxo-7,8-dihydroguanine |
| A | Adenine |
| AIDS | Acquired Immune Deficiency Syndrome |
| amp | Ampicillin |
| AP | Apurinic/Apyrimidinic |
| *aph* | Gene encoding kanamycin resistance |
| *aphC* | alkyl hydroperoxidase gene |
| BCG | Bacille Calmette - Guerin |
| BER | Base Excision Repair |
| BLAST | Basic Local Alignment Search Tool |
| bp/s | Base pair/s |
| BSA | Bovine Serum Albumin |
| C | Cytosine |
| CFU | Colony Forming Unit |
| CFZ | Clofazamine |
| CTAB | Cetyltrimethylammonium bromide |
| DCO | Double cross over |
| DIG | Dioxygen |
| DMSO | Dimethylsulphoxide |
| DNA | Deoxyribonucleic acid |
| DNase | Deoxyribonuclease |
| dNTPs | Deoxyribonucleotide triphosphates |
| DS | Downstream |
| DSBs | Double strand breaks |
| dTTP | Deoxythymine triphosphate |
| dUTP | Deoxyuridine triphosphate |
| EDTA | Ethylenediaminetetra acetic acid |
| EMB | Ethambutol |
| EP-TB | Extra-pulmonary TB |
| Fapy | Formamidopyrimidine |
| Fpg | Formamido-pyrimidine-DNA-glycosylase |
| FU | Formyluracil |
| G | Guanine |
| *g* | Gravitational force |
| G+C | Guanine plus Cytosine |
| $H_2O_2$ | Hydrogen peroxide |
| HIV | Human Immunodeficiency Virus |
| HMU | Hydroxylmethyluracil |
| hr/s | hour/s |
| IFN | Interferon |
| INH | Isoniazid |
| *inhA* | NADH dependent trans enoyl-acyl ACP |
| kanR | Kanamycin resistance |
| *katG* | Catalase-peroxidase |
| Kb | Kilo base pairs |
| LA | Luria-Bertani agar |

| | |
|---|---|
| lacZ | Gene encoding β galactosidase activity |
| *lacZ* | Gene encoding β-galactosidase |
| M | Molar |
| Mb | Mega base pairs |
| MDR-TB | Multidrug-resistant tuberculosis |
| MIC | Minimum inhibitory concentration |
| min | Minutes |
| MutY | Adenine DNA glycosylase |
| m-value | Number of mutational events |
| n | Nano |
| NaCl | Sodium chloride |
| NADPH | Nicotinamide Adenine Dinucleotide Phosphate |
| NaOH | Sodium hydroxide |
| Nei | Endonuclease VIII |
| Nth | Endonuclease III |
| NY | New York |
| $O_2$ | Oxygen |
| $OD_{600nm}$ | Optical density at 600 nanometer wavelength |
| Ori | Origin of replication |
| P | Phosphate |
| PCR | Polymerase Chain Reaction |
| PYZ | Pyrazinamide |
| RNA | Ribonucleic acid |
| RNase | Ribonuclease |
| RNI | Reactive nitrogen intermediates |
| RNS | Reactive nitrogen species |
| ROI | Reactive oxygen intermediates |
| ROS | Reactive oxygen species |
| rpm | Revolutions per minute |
| s | Seconds |
| *sacB* | Gene encoding levansucrase |
| SSBs | Single strand breaks |
| STR | Streptomycin |
| T | Thymine |
| TE | Tris EDTA |
| Tg | cis- and trans-thymine glycol |
| TLR | Toll-like receptors |
| Tween | Polyoxyethylenesorbitan monooleate |
| Ug | Uracil glycol |
| US | Upstream |
| USA | United States of America |
| UV | Ultraviolet |
| v/v | Volume/volume |
| WHO | World Health Organization |
| X-gal | 5-bromo-4-chloro-3-indolyl- β-galactosidase |
| α | Alpha |
| β | Beta |
| μ | Mutation rate |

**ist of figures**

Figure 1.1 Cellular components targeted by the first line anti-TB drug .....6

Figure 1.2 Life cycle of *M. tuberculosis* within the human host and the resultant formation of granuloma within the lung…….....…….....10

Figure 1.3 Hydrolytic DNA lesion...……..…….....13

Figure 1.4 Oxidative DNA Lesion…..……..…….14

Figure 1.5 Schematic diagram of the base excision repair pathway ……..21

Figure 2.1 Schematic representation of the construction of knockout constructs using three way cloning ……..…….....35

Figure 2.2 Schematic representation of knock out mutant generation ….…….....37

Figure 2.3 Example of the MSS-LC distribution graph/curve used to estimate the $Rif^R$ mutants per culture (*m*-value) ……...40

Figure 2.4 Schematic outline of the fluctuation assay experiment ……..42

Figure 3.1 Structural characteristics of the *M. smegmatis* Nei protein. …..……. 44-45

Figure 3.2 Alignments of the amino acid sequences of the Nei protein from various microorganisms…… 46 -47

Figure 3.3 Structural characteristics of the *M. tuberculosis* Nei protein……..... 48

Figure 3.4 Confirmation of ∆*neiI* single mutant ……..… 49

Figure 3.5 Confirmation of ∆*neiII* single mutant ……..... 50

Figure 3.6 Confirmation of the p2NIL∆*nth*:: pGOAL19 knockout construct using restriction enzyme digestion ……. 51

Figure 3.7 PCR screen for identification of DCOs using *nth* primers ……..….. 53

Figure 3.8 Southern blot upstream analysis of Δ*nth* mutants ……..….... 56

Figure 3.9 Confirmation of the pTWEETY::*nth* complementation construct using restriction enzyme digestion ……. 57

Figure 3.10 PCR confirmations of complemented strains..……...……..…… 58

Figure 3.11 Growth kinetics of the parental and mutant strains under normal culture conditions ……. 59

Figure 3.12 Assessment of washed and unwashed wild type *M. smegmatis* strain in the presence of hydrogen peroxide ……..….62

Figure 3.13 Assessment of wild type *M. smegmatis* single, double deletion mutant and their respective complemented strain with and without 2 mM $H_2O_2$……63

## List of Tables

Table 1.1 Bacterial strains……..……27

Table 1.2 Cloning vectors……..……28

Table 2.1 Spontaneous rifampicin resistance mutation rates……...66

Table 3.1 PCR Primers used in this study…..……..82

Table 3.2 Primers used for qPCR…….……83

Table 3.3 PCR reagents and cycling conditions…..……84

# 1. Introduction

## 1.1 **Tuberculosis**

Tuberculosis (TB), caused by *Mycobacterium tuberculosis*, is one of the oldest recorded human diseases, and is still one of the biggest killers among the infectious diseases, despite the worldwide use of a live attenuated vaccine and several antibiotics. *M tuberculosis* is primarily transmitted through inhalation of aerosol droplets expelled by infected hosts.

TB is spread from person-to-person through the air by droplet nuclei which are produced when persons with pulmonary or laryngeal TB cough and sneeze (Schaaf and Zumla, 2009). Droplet nuclei can also be generated during aerosol treatments, sputum induction, aerosolization during bronchoscopy, and through manipulation of lesions or processing of tissue or secretions (Mlambo, 2004). Due to their small size, droplet nuclei can remain airborne for minutes to hours after expectoration (Nardell, 2004). When inhaled, droplet nuclei are carried down the bronchial tree and implant in respiratory bronchioles or alveoli in the lungs. Whether or not an inhaled tubercle bacillus establishes an infection in the lung depends on both bacterial virulence and the inherent microbicidal ability of the alveolar macrophage that ingests it (Dannenberg, 1989, Schaaf and Zumla, 2009). The last two decades has seen a dramatic rise in drug resistant TB, with composite forms of resistance that are difficult to treat. Hence, new vaccines and drugs are needed to stem the worldwide epidemic of TB. To rationally develop new anti-tubercular agents, it is essential to study the genetics and physiology of *M. tuberculosis* and related mycobacteria. It is equally important to understand the *M. tuberculosis*-host interaction to learn how these bacteria overcome host defences and cause disease (Smith, 2003).

Worldwide almost 8.7 million cases of TB occur every year and cause about 1.4 million deaths each year (Glaziou et al., 2013). The recent data obtained in 2015 surveillance shows that India has the highest incidence rate of 2.0 million–2.5 million, followed by China with 0.9 million–1.1 million (WHO, 2016). South Africa is in third position with an incidence

rate of approximately 0.4 million–0.6 million annually and countries such as Indonesia and Pakistan account for 0.4 million–0.5 million and 0.3 million–0.5 million cases respectively (WHO, 2015). TB accounts for extraordinarily high rates of morbidity and mortality worldwide primarily because of its ability to persist for long periods of time in the face of an active immune response, and to adapt rapidly to the changing conditions inside and outside the host (Ulrichs and Kaufmann, 2002, zu Bentrup and Russell, 2001).

*M. tuberculosis* and HIV have been closely entwined since early years of the HIV/AIDS pandemic. They individually carry the risk of creating social, economic and political instability. The overwhelming burden of disease due to both HIV and TB is borne by resource –limited countries and the hardest hit among these is sub-Saharan Africa. In sub-Saharan Africa the HIV epidemic is accelerating together with the already massive TB epidemic. TB is now the leading cause of death among HIV positive patients. HIV increases the risk of reactivation of latent TB (LTBI) and progression to active TB disease more than other risk factors. In some countries, the percentages of patients with active TB who are co infected with HIV are now greater than 60% (WHO, 2015), even with appropriate management of TB. Patients with HIV co-infection have increased mortality as a consequence of HIV related complications (Corbett et al., 2003).

TB/HIV co-infection represents a novel pathogenic scenario at the global level. It constitutes a serious diagnostic and therapeutic challenge and, particularly in poor countries, where it weighs heavily on already strained health care budgets. There is growing appreciation that the epidemiology, clinical manifestations, and management of both HIV and *M. tuberculosis* infections are different and far more complex in co-infected compared to mono-infected patients. However, our knowledge about the mechanisms of interaction of the two pathogens still has many gaps that need to be filled in order to develop preventive measures against the two diseases (Pawlowski et al., 2012). Despite the epidemiological data supporting the notion that *M. tuberculosis* infection has a negative impact on the immune response to HIV and on progression to AIDS, research on possible mechanisms is scarce. The function of many immune cells, including macrophages and dendritic cells (DCs), is modulated by both HIV and *M. tuberculosis* infection (Pawlowski et al., 2012).

The primary target for *M. tuberculosis*, the alveolar macrophage, can also be infected with HIV (Padayatchi et al., 2017). Mycobacteria exacerbate HIV replication in macrophages and lung cells obtained by broncho-alveolar lavage from co-infected individuals. Also, in vitro studies have demonstrated that *M. tuberculosis* infection can up-regulate both HIV infection and replication within monocyte-derived macrophages (MDMs), increase the efficiency of virus transmission from infected MDMs to T cells, and favour replication of X4 HIV variants by up-regulation of CXCR4 (Adu, 2014). Furthermore, monocytes from $HIV^+$ patients display an impaired response to toll-like receptors (TLR) ligands, and viral proteins, which can interfere with both MDM and dendritic cell (DC) maturation and function in vitro, including their ability to phagocytose mycobacteria and kill intracellular bacteria (Pawlowski et al., 2012). SOCS1, which is stimulated by infection with *M. tuberculosis*, has been shown to facilitate the late replication pathways of HIV infection and mediate viral evasion of type I interferon (IFN) anti-viral signalling (Fenner et al., 2006). While tumour necrotic factor (TNF) production in response to *M. tuberculosis* infection is required for control of bacterial growth, TNF is known to activate HIV replication in macrophages, indicating that the host immune response initiated against one pathogen may promote the replication of another. Thus, both HIV and *M. tuberculosis* stimulate TNF release from infected cells, and TNF hampers bacterial growth while enhancing HIV replication (Pawlowski et al., 2012).

It is estimated that approximately 1.1 million incident cases, of the 8.7 million people who developed TB infection worldwide, were HIV positive (Gao et al., 2013, WHO, 2015). In Africa TB co-infection with HIV cases are high (WHO, 2015). South Africa accounts for 28% of the world's HIV-related TB cases (Corbett et al., 2003). Another serious problem associated with HIV infection is that it allows for the dissemination of bacilli, which leads to the development of extra-pulmonary TB (EP-TB) (Caws et al., 2008). The diagnosis of TB in HIV infected individuals and EP-TB are challenging considering the fact that with EP-TB, the number of bacilli are usually low in the tissues at sites of the disease and another difficulty is in obtaining clinical specimens from deep-seated organs (Prihoda et al., 2008). At least one-third of HIV patients are more prone to TB, however, antiviral agents play an important role in the effectiveness of TB treatment (Sun et al., 2016).

## 1.2 Treatment of TB infection

Infection with *M. tuberculosis* can result in active TB or more commonly, latent infection, which is defined as having evidence of *M. tuberculosis* infection by immunologic tests (tuberculin skin test [TST] or IFN-γ release assay [IGRA]) without clinical signs or symptoms of disease and a normal chest radiograph (Organization and Initiative, 2010). This represents a state of equilibrium in which the host is able to control the infection but not completely eradicate the bacteria. Patients with latent infection are the largest reservoir for potential transmission. Although most patients with latent infection will not die of TB, the greatest danger is in reactivation (active TB after Latent TB infection [LTBI]) cases and the subsequent silent spread to close contacts. The latent stage is normally tolerant to multiple anti-TB drugs therefore, rendering a need to discover and develop TB drugs and drug regimens that are well understood and superior to those available today in their efficacy, speed of action, safety and tolerability, ease of use for all patient populations, and accessibility.

The standard treatment regimen for drug-susceptible TB includes an intensive phase consisting of rifampin, isoniazid, and pyrazinamide, to which ethambutol is added as protection against unrecognized resistance to one of the three core drugs. Once susceptibility to isoniazid, rifampin, and pyrazinamide has been confirmed, ethambutol can be discontinued (Parumasivam et al., 2016). In young children, this drug is frequently omitted if the source of transmission is known to have drug-susceptible TB, because recognizing the toxic effects of ethambutol is challenging in children. The intense phase is followed by a continuation phase consisting of rifampin and isoniazid for an additional 4 months of treatment (Parumasivam et al., 2016).

The standard 6-month treatment regimen for drug-susceptible tuberculosis is an exceptionally long course of treatment as compared with the duration of treatment of other bacterial infectious diseases (Albanna et al., 2011). The prolonged regimen poses two major challenges to success: managing drug toxicity and ensuring that patients adhere to the full course of treatment. Drug toxicity is substantial; a review of retrospective studies using similar definitions estimates that 3 to 13% of patients have hepatotoxic effects (Saukkonen et al., 2012). A recent prospective cohort study of patients with drug-

susceptible disease who received standard tuberculosis therapy documented a 15% incidence of adverse drug reactions resulting in interruption or discontinuation of one or more of the drugs. Of these adverse reactions, 7.7% resulted in hospitalization, disability, or death. A wide variety of reactions were reported; the most common were hepatotoxic effects, gastrointestinal disorders, allergic reactions, and arthralgia (Lv et al., 2013).

Multidrug-resistant (MDR) tuberculosis, is defined as the disease caused by *M. tuberculosis* that is resistant to both rifampin and isoniazid (Gunther et al., 2015). Recommendations for the treatment of drug resistant tuberculosis comprise a regimen consisting of a rifamycin, ethambutol, and pyrazinamide, with or without a fluoroquinolone, for 6 to 12 months (Allan, 1977). However, it can be difficult for patients to complete 9 months of treatment with an injectable agent. Eighteen-month regimens comprising isoniazid, pyrazinamide, and ethambutol, with or without a fluoroquinolone, are also recommended, on the basis of clinical experience. If the isolate is susceptible to isoniazid and pyrazinamide, a shorter and less toxic regimen can be used. For MDR tuberculosis, all the national guidelines surveyed recommend regimens that are individualized on the basis of the results of drug-susceptibility testing (WHO, 2016). In order to control the drug resistance epidemic it is necessary to gain insight into how *M. tuberculosis* develops drug resistance. This knowledge will help to understand how to prevent the occurrence of drug resistance as well as identifying genes associated with resistance to new drugs.

The various anti-mycobacterial agents have varying mechanisms of action (see Figure 1.1) (Kimerling et al., 2008). The molecular mechanisms which are responsible for isoniazid (INH) resistance are quite complex as they have been associated with a variety of mutations which affect one or several genes involved in mycolic acid biosynthesis (Figure 1.1) (Johnson et al., 2006). Genes which are associated with INH resistance are *katG, inhA, and ahpC*; however, the more frequent mutation in INH resistant isolates is located in codon 315 of *KatG*. Mutation in the *katG* gene is responsible for 60-70% of isoniazid-resistant strains (Johnson et al., 2006, Miesel et al., 1998).

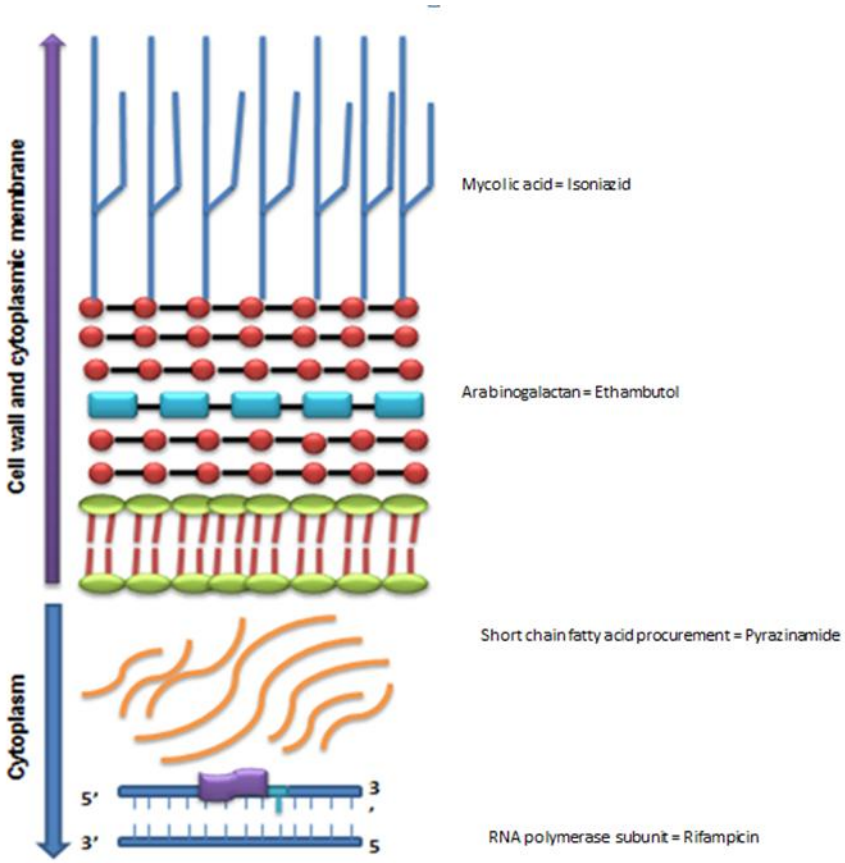



**Figure 1.1: Cellular components targeted by the first line anti-TB drugs** In this illustration, Pyrazinamide, Isoniazid, Rifampicin and Ethambutol are drugs that target different pathways that are essential for the survival of *M. tuberculosis* [adapted from (Shi et al., 2007)].

Rifampicin (RIF) is known to interfere with transcription in bacteria by binding to the β-subunit of RNA polymerase (the product of the *rpoB* gene). Mutations in codons 516, 526 and 531 of *rpoB* are most commonly associated with high-level RIF resistance, especially in the areas with high incidence of MDR-TB, but the frequency with which these mutations are observed, varies by geographic location and treatment history (Ramaswamy and Musser, 1998, Cavusoglu et al., 2002, Van Deun et al., 2009), in addition mutations in this region causes cross-resistance to rifamycins (Palomino, 2009). Of all RIF resistance mutations, 95 to 98% are found in an 81 bp 'hotspot' called the RIF resistance-determining region (RRDR), codons 507 to 533 (Telenti et al., 1993, Gillespie, 2002). The point mutation that occurs in the *rpoB* gene in codons 511-533 accounts for approximately 90% of RIF resistant

*M. tuberculosis* strains from different countries (Nunn and Tweed, 2016). Ninety-five percent of *M. tuberculosis* strains with resistance to RIF contain distinct mutations located within 81 bp (27 codon) region of the β-subunit of the RNA polymerase (*rpoB*) gene (Palomino, 2009). RIF resistance have also been reported to be caused by mutations in other parts of *rpoB* outside the hot spot region, such as codon 176 (146) in the N-terminal part and codons 541 and 553 (Telenti et al., 1993).

RIF remains a major drug used in the treatment of TB infections, and an increased RIF resistance worldwide, represents a global clinical problem. Resistance to anti-TB drugs is mostly due to inadequate chemotherapy and poor clinical management, which results in the selection of resistant isolates. RIF's mono-resistance is extremely rare and as a result, resistance to RIF has been a surrogate marker for MDR-TB (Gillespie, 2002).

The result of drug susceptibility testing (DST) is the critical component of initiating treatment when infection is detected. There has been a marked increase recently in the development of both phenotypic and molecular assays to detect *M. tuberculosis* complex (Gagneux and Small, 2007). DST for certain drugs can also be technically challenging and in resource-limited areas, cost prohibitive. The use of molecular methods to identify mutations associated with drug resistance can decrease diagnostic delay and, in some cases, may prove to be more useful than phenotypic DST.

In 2010, South African Department of Health rolled out the line probe assay (LPA) (GenoType MTBDR*plus*, Hain Life Science) nationwide for the rapid detection of drug resistance and thereby tailoring of treatment regimens. However, rapid methods are not a replacement for culture, most are not reliable when used with smear-negative specimens. Conventional DST is not sensitive enough whilst the specificity of LPA is high (Williamson et al., 2011). Discordant results with the BACTEC Mycbacteria Growth Indicator Tube (MGIT) 960 and LPA have been widely observed nationally since the rollout of the rapid LPA (Parsons et al., 2011). Approximately half of MDR suspects screened by LPA (GenoType MTBDR*plus*, Hain Life Science) that had neither wild type nor high level mutations bands were susceptible by MGIT (Parsons et al., 2011). The cause of discordance between molecular and phenotypic DST may include human error or laboratory error, unknown mechanisms of resistance, *rpoB* mutations due to fitness loss that results in false

susceptible phenotypic DST, limited genes and sites targeted in the assays, emerging resistance and/or mixed populations, not all mutations are associated with phenotypic resistance, silent mutations that leads to no change in protein, neutral polymorphisms (e.g., *gyrA* codon 95 may be Ser or Thr), and phenotypic DST mutations which may result in elevated minimal inhibitory concentrations (MIC) but susceptibility at critical concentrations for DST (e.g., Leu511Pro in *rpoB*) (Van Deun et al., 2009, Williamson et al., 2011). In DST proficiency testing among the Supranational Reference Laboratories, 10% of the strains yielded less than 80% agreement for Rif resistance, leading to exclusion of these strains from a judicial result. DNA sequencing of these strains showed the presence of particular *rpo*B gene mutations despite their variable susceptibility. Further investigation of these strains showed that 85% of the strains classified as resistant or probably resistant strains (on the basis of mutations in the *rpo*B gene and poor treatment outcome), yielded susceptible results by the rapid BACTEC MGIT 960 or radiometric method (Williamson et al., 2011).

Susceptible MGIT results from genotypically RIF-resistant strains have been reported in the literature. A recent retrospective study conducted in New Zealand showed that 4.3% of phenotypically INH-resistant, RIF-susceptible strains harboured clinically significant *rpo*B mutations and had led to treatment failure (Parsons et al., 2011). These variable strains with low-level resistance, linked to specific *rpo*B mutations, thus seem to have clinical relevance (occult MDR-TB). They may also not be so rare, representing around 15% of all *rpoB* mutations found in primo-treatment failures and relapses from Bangladesh and Kinshasa (Van Deun et al., 2013). The frequency of these mutations in different geographical settings is not known and should be assessed. There are two important issues related to these mutations, their effect on sensitivity of DST for RIF by BACTEC MGIT 960, widely used in several countries and their impact on clinical outcome of TB.

### 1.3 Pathogenesis

Infection of tuberculosis is initiated by inhalation of droplet nuclei, which are particles of 1–5 *μ*m in diameter containing *M. tuberculosis*, expectorated by patients with active pulmonary TB (open TB), typically when the patient coughs. The droplet nuclei, due to their small size, can remain suspended in the air for several minutes to hours (Figure 1.2). The risk of

infection is dependent on several factors such as the infectiousness of the source case, the closeness of contact, the bacillary load inhaled, and the immune status of the potential host response (Zhang, 2005) .

The primary route of infection involves the lungs. Inhaled droplet nuclei avoid the defences of the bronchi due to their small size and penetrate into the terminal alveoli (Figure 1.2 B) where they are engulfed by phagocytic immune cells (macrophages and dendritic cells). *M. tuberculosis* can also infect non-phagocytic cells in the alveolar space including M cells, alveolar endothelial, and type 1 and type 2 epithelial cells (pneumocytes) (Abebe et al., 2007) (Figure 1.2 C).

In the early phase of infection, *M. tuberculosis,* internalized by phagocytic immune cells, replicates intracellularly and the bacterially loaded immune cells may cross the alveolar barrier to cause systemic dissemination (Zhang, 2005). The intracellular replication and simultaneous dissemination of the pathogen to the pulmonary lymph nodes and to various other extra-pulmonary sites occurs prior to the development of the adaptive immune response. This exemplifies the extraordinary ability of *M. tuberculosis* to establish a protected niche where it can avoid elimination by the immune system and persist indefinitely (Cooper, 2009). In the vast majority of the infected individuals, an effective cell-mediated immune response develops 2–8 weeks after infection that stops further multiplication of the tubercle bacilli. The activated T lymphocytes, macrophages, and other immune cells form granulomas (Figure 1.2 D) that wall off the growing necrotic tissue limiting further replication and spread of the tubercle bacilli.

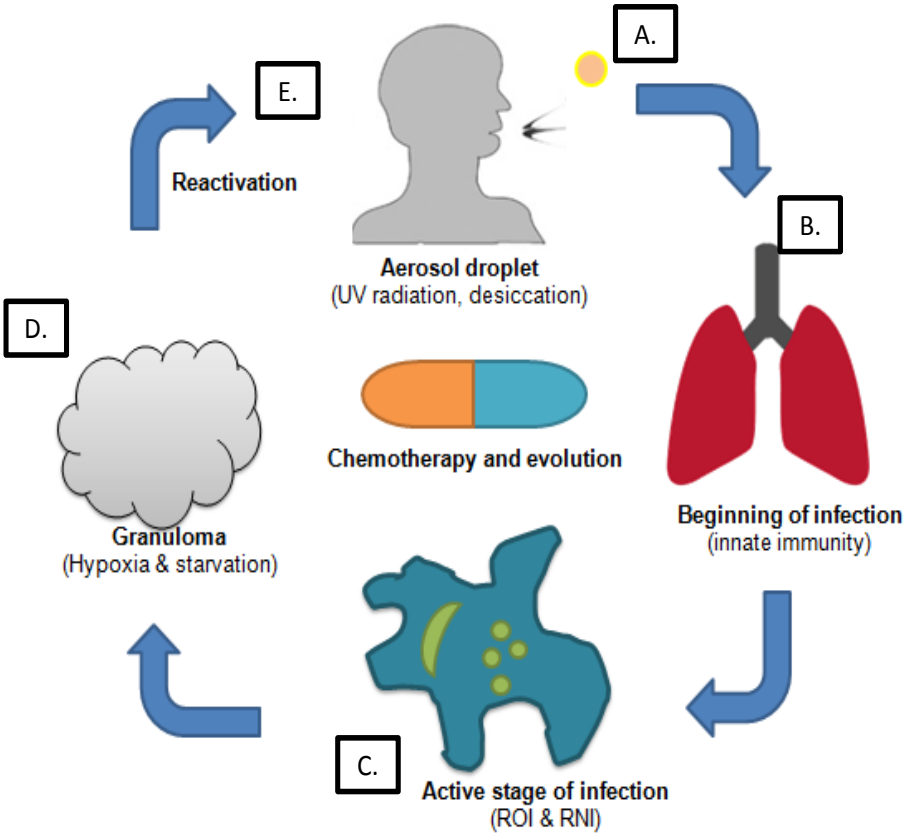



**Figure 1.2: Life cycle of *M. tuberculosis* within the human host and the resultant formation of granuloma caseate within the lung.** When *M. tuberculosis* is inhaled and gains entry into the host via phagocytosis, it persists in the macrophages of the lung tissues. A sub-population of the infected macrophages undergoes differentiation, and become filled with lipid droplets, referred to as foamy macrophages. *M. tuberculosis* eventually escapes when granuloma caveats, resulting in transmission to other hosts [Adapted from (Ahmad,

Most of the *M. tuberculosis* is killed in the caseating granulomas, and disease progression is arrested. However, the pathogen is not completely eradicated in some individuals as *M. tuberculosis* has evolved effective strategies to evade the immune response resulting in survival and persistence of some bacilli in a non-replicating state in the host (Al-Harbi, 2012). In this case, that is LTBI, for reasons such as a weakened immune system; bacilli can start replicating inside this primary lesion, allowing for active disease to follow.

## 1.4 Host cell activity in response to entry of *M. tuberculosis*

The precepts of cell mediated immunity (CMI) maintain that macrophages, activated by interferon-gamma (IFN-□), are primary effector cells responsible for controlling intracellular pathogens such as *M. tuberculosis* (Kahn et al., 1987). An impressive array of monokines and cell products are produced by activated immune macrophages, and many of these have been implicated not only in host resistance to TB, but also in the immunopathology of this disease. Two of the most powerful products of activated macrophages are toxic radicals of oxygen and nitrogen (Nathan and Ehrt, 2004).

From the point of granuloma formation, the mycobacteria are continuously exposed to both oxidative and nitrogenous stress generated by the activated macrophages that they inhabit. Macrophages produce ROIs (reactive oxygen intermediates) such as hydroxyl radical (•OH), singlet oxygen ($^1O_2$) and one-electron and reactive nitrogen species (RNS) which form part of the anti-mycobacterial mechanisms (Zahrt and Deretic, 2002, Cadet and Wagner, 2014). RNIs (reactive nitrogen intermediates) serve as cytotoxic agents in the immune response to pathogens. Production of these compounds by host cells appears to be essential in the control of mycobacterial infection. Reactive oxygen species (ROS) and reactive nitrogen species (RNS) are produced by two enzymes namely NADPH oxidase (NOX2/gp91$^{phox}$) and inducible nitric oxide synthase (iNOS) respectively, from activated macrophages (Ehrt and Schnappinger, 2009b). The superoxide anions ($O_2^{\cdot -}$) produced by NOX2 as catalysts for the production of ROS species, •OH, $^1O_2$ and hydrogen peroxide ($H_2O_2$) (Babior, 1999). The reactive oxygen species are formed when NOX2 assembles into an active enzyme complex that transports electrons across the membrane from NADPH to molecular oxygen, and results in the production of ($O_2^{\cdot -}$), that dismutase into $H_2O_2$ and consequently generates lethal hydroxyl. IFN-□ induces iNOS and make nitrite along with nitrate through nitric oxide radicals (Ehrt and Schnappinger, 2009b). The mutagenic and cytotoxic effect of ROIs and RNIs is based on several chemical reactions leading to damage of cellular macromolecules including DNA.

## 1.5 DNA damage

DNA in living cells is subjected to many chemical alterations. If the genetic information encoded in the DNA is to remain uncorrupted, any chemical changes must be corrected. A failure to repair DNA before it is replicated produces a mutation. Damage to DNA is induced by certain wavelengths of radiation, ionizing radiation and ultraviolet rays, highly-reactive oxygen radicals, chemicals in the environment, and chemicals used in chemotherapy of cancers. Various forms of DNA damage arise from exposure to endogenous and exogenous DNA damaging agents. All four of the different bases in DNA (A, T, C, and G) can be covalently modified at various positions. One of the most frequent base modifications is a deamination, i.e. a loss of an amino group. Mismatches of DNA bases can arise because of a failure of proofreading during DNA replication. DNA strand breaks from ionizing radiation or chemicals can be limited to one of the two strands giving rise to a single strand break (SSB), or on both strands causing a double strand break (DSB). Crosslinks or covalent linkages, induced by several chemotherapeutic drugs, can be formed between bases on the same DNA strand (intra-strand breaks) or on the opposite strand (inter-strand breaks).

### 1.5.1 Endogenous Stresses

#### 1.5.1.1 Hydrolytic DNA damage

Endogenous DNA lesions are caused by chemical hydrolysis of nucleoside bases or by reaction of DNA with electrophiles or reactive free radicals generated during metabolism. Uracil (U) is normally confined to RNA, but it can arise in DNA by replication error and by spontaneous deamination of cytosine (Figure 1.3 C) (Mosbaugh, 1988). Uracil in DNA is subject to removal by Uracil DNA glycosylase (UDG or Ung)-initiated base excision repair (BER). Spontaneous deamination of 5-methylcytosine (5-meC) generates thymine (T) leading to T:G mis-pairs (Figure 1.3 B), which is a major cause of mutations in eukaryotes (Ehrlich et al., 1990). Adenine (A) and guanine (G) can also deaminate albeit with a lower rate creating hypoxanthine and xanthine respectively (Figure 1.3 A & D). Hypoxanthine is potentially mutagenic, mis-pairing with C during DNA replication (Batey, 2004). Xanthine, on the other hand, is only able to pair with C or T unstably, which may arrest DNA synthesis (Greer and Zamenhof, 1962). Furthermore, the N-glycosylic bond of DNA, particularly with

purine residues, is susceptible to spontaneous hydrolysis resulting in an apurinic/apyrimidinic site (APsite) (Lindahl and Barnes, 2000).

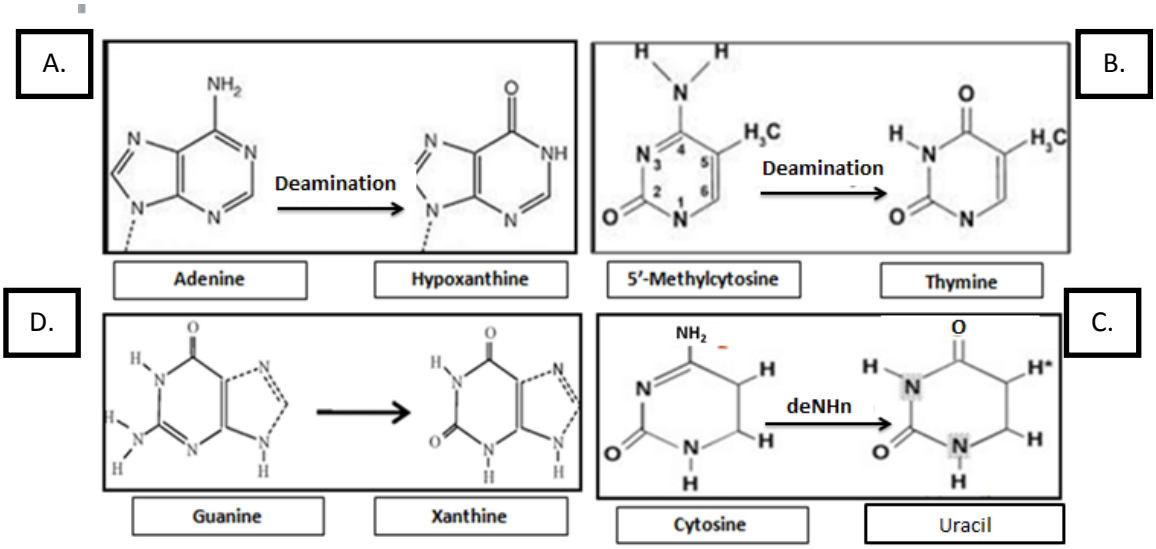



**Figure 1.3: Hydrolytic DNA lesions.**

DNA damaged lesions that arise from deamination of Adenine; Cytosine, Guanine and 5'-methylcytosine (Gates, 2009).

### 1.5.1.2 Oxidative DNA damage

Reactive oxygen species (ROS), such as superoxide, hydrogen peroxide ($H_2O_2$), and hydroxyl radicals, which are generated during normal cellular metabolism and by 3 external sources, such as UV light, ionizing radiation and some chemicals, can lead to various types of DNA damage, including oxidized or ring-opened base lesions, abasic sites and strand breaks (Cheng, 1992). Common oxidative base lesions are shown in Figure 1.4. Oxidative damage is generally repaired by the base excision repair (BER) pathway. Guanine (G) has a low redox potential thus making it particularly vulnerable to oxidation (Cooke, 2003). One of the most abundant oxidized base lesions is 7, 8-dihydro-8-oxo-2'-deoxyguanosine (8-oxoG), Figure 1.4. 8-oxoG is a major form of mutagenic base damage, which can mispair with adenine (A) during replication, giving rise to G to thymine (T) transversions (Cheng, 1992). The ring opened purine derivatives, 2,6-diamino-4-hydroxy-5-formamidopyrimidine (FapyG) and 4,6-diamino-5-formamidopyrimidine (FapyA) are generated by $H_2O_2$ treatment, as well as ionizing radiation (Kim and M Wilson III, 2012).

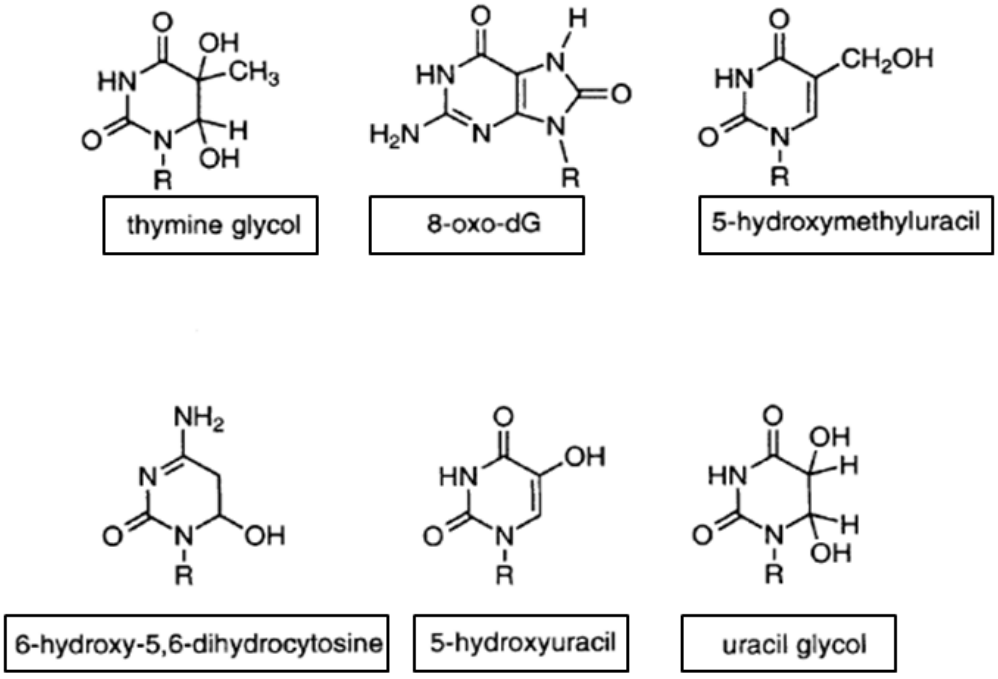



**Figure 1.4: Oxidative DNA Lesions**

DNA damaged lesions that arise from oxidation of adenine; cytosine, guanine and cytosine (Marnett, 2000).

### 1.5.2 Exogenous stress

#### 1.5.2.1 Alkylation damage

Alkylating agents are electrophilic compounds which can attack nucleophilic centres of DNA resulting in N-alkylated or O-alkylated purines and pyrimidines as well as phosphotriesters (Sinha and Hader, 2002). The most abundant alkylation products in double-stranded DNA are 7-methylguanine and 3-methyladenine, both of which block DNA replication and thus are cytotoxic. Other alkylation products, such as $O^6$-methylguanine and $O^4$-methylthymine can mispair during DNA replication resulting in mutagenesis. The $O^6$-alkylated guanine lesions can be directly reversed to guanine; the other alkylated bases are repaired by the BER pathway (Sancar et al., 2004).

#### 1.5.2.2 Ionizing Radiation-induced Damage

Ionizing radiation, including α-particles, β-particles, X rays and γ-rays, damage DNA by direct deposition of radiation energy (direct effect) and more commonly by ionization of water molecules to produce hydroxyl radicals (indirect effect), thus inducing multiple forms of DNA damage, including base lesions, AP sites, single-strand breaks (SSBs) as well as double-strand breaks (DSBs) (Ward, 1973). The damage generated by radiation-induced hydroxyl radicals overlaps considerably with those generated during normal oxidative metabolism, which is efficiently removed by BER. A special feature of ionizing radiation is the formation of clustered damage or multiple damage sites in close proximity on the DNA (Cadet, 2016). Clustered damage on the opposite strands of DNA can readily cause DSBs, which are among the most cytotoxic and mutagenic lesions. DSBs are repaired by recombinational repair (Cadet, 2016).

#### 1.5.2.3 UV induced damage

UV-induced DNA damage can result from the direct absorption of photons by DNA. The most frequent DNA lesions formed by UV radiation are cyclobutane pyrimidine dimers (CPD) and pyrimidine-pyrimidone (6-4) photoproducts [(6-4) PP] wherein two adjacent pyrimidines are covalently linked. Moreover, UV radiation results in DNA-protein cross-links and occasionally inter-chain DNA-DNA cross-links (Drabla¸s et al., 2004). These photoproducts, as well as other minor photoproducts in DNA, can interfere with DNA replication and transcription. Both CPD and 6-4 PP can be excised from DNA by the nucleotide excision repair pathway (NER). UV radiation can also form reactive oxygen species causing oxidative damage (Sinha and Hader, 2002). There are two photoproducts, both of which are di-pyrimidine structures, believed to be responsible for most of the carcinogenic effects of UV radiation. These two photoproducts, cyclobutane pyrimidine dimer (CPD) and 6-4 pyrimidine –pyrimidone, are caused by exposure to UV-B radiation. CPD is more common than 6-4 pyrimidine-pyrimidone and is thought to account for approximately 85% of primary lesions in UV-irradiated DNA. The photoproducts interfere with DNA replication if not repaired and cause specific mutations in DNA. The specific mutations induced by UV-B in DNA sequences most often include single-base substitutions

of cytosine (C) for thymine (T). Double-base changes from CC to TT also occur, but in an order of less magnitude (Cooke, 2003).

## 1.6 DNA repair pathways

DNA repair can occur by one of two mechanisms that involves either the reversal of DNA damage or the excision of the damaged bases. With regards to excision repair, three basic repair pathways have evolved to prevent the cytotoxic or mutagenic outcomes of DNA damage thus maintaining genomic integrity. BER predominantly copes with small lesions resulting from oxidation, alkylation or deamination, mismatch repair (MMR) corrects mismatches as well as insertion/deletion loops formed during replication, and nucleotide excision repair (NER) removes bulky helix-distorting lesions. In addition, recombinational repair is crucial for repairing DSBs. Moreover, in some bacteria, the physiological responses to DNA damage are under the control of an SOS regulatory network, which leads to not only DNA repair but also to damage tolerance and mutagenesis (Dos Vultos et al., 2009).

The G/C content of the *M. tuberculosis* genome is very high (65.4%), making the pathogen particularly susceptible to guanine and cytosine oxidation and cytosine deamination (O'Sullivan et al., 2005). Oxidative stress plays an important role in the host's innate immune response. Macrophages generate ROS and RNS resulting in lethal damage to the DNA of pathogenic microorganisms. However, *M. tuberculosis* manages to survive and replicate in its host's macrophages. Therefore, the ability to repair DNA damage caused by exposure of the microbe to the environment within macrophages is likely to play a particularly important role in pathogen proliferation or colonization conferring a virulence advantage for *M. tuberculosis* (Master et al., 2008).

### 1.6.1 Mismatch repair

The *M. tuberculosis* response to DNA damage has distinguishing features compared to some other bacteria. Most strikingly, the genome sequences of *M. tuberculosis, Mycobacterium leprae,* and *M. smegma*tis indicate that these microbes lack the mismatch repair (MMR) homologues (Dos Vultos et al., 2009, Mizrahi and Andersen, 1998). The deficiency of MMR possibly confers a selective advantage under certain stressful situations such as the consequent increase in mutation rate, which may serve as a starting point for

the evolution of antibiotic resistance, fitness for survival and pathogenicity, and allowing adaptation during specific periods of the life cycle (Eisen and Hanawalt, 1999). A similar mismatch repair deficiency was observed for *Helicobacter pylori (H. pylori)*, suggesting that the absence of MMR allows for increased mutation frequencies, emergence of antibiotic resistance and fitness for survival and pathogenesis under certain stressful situations (Kang and Blaser, 2006).

### 1.6.2 Recombination repair

DNA damage is a fact of life for bacteria as a consequence of endogenous sources and processes as well as exogenous stresses. Homologous recombination (HR) is a DNA metabolic process found in all forms of life that provides high-fidelity, template-dependent repair or tolerance of complex DNA damages including DNA gaps, DNA double-stranded breaks (DSBs), and DNA inter-strand crosslinks (ICLs). In addition to its role in preserving the genome, HR plays a prominent role in faithfully duplicating the genome by providing critical support for DNA replication and telomere maintenance (Boshoff et al., 2003).

Homologous recombination repair involves the exchange of two complementary strands of DNA. This repair system relies on a homologous chromosomal copy as template for DNA synthesis and hence is considered not to be an error prone system; (Warner & Mizrahi, 2011). The RecBCD complex is made up of a helicase and nuclease proteins that unwinds double stranded DNA into single strands, which are cut at a specific nucleotide sequence known as the Chi site (Kowalczykowski et al., 1994). The single stranded DNA is bound by the RecA protein that acts as a signal to induce the SOS response; in particular the RuvABC complex, which searches for homologous duplex DNA (Kowalczykowski *et al*., 1994). Upon recognition of homologous DNA, RecA coated single stranded DNA invades the duplex recipient DNA (Kowalczykowski et al., 1994, Warner and Mizrahi, 2011). One of the recipient strands forms a bulge which is cleaved by the endonuclease from the RuvABC complex (Kowalczykowski et al., 1994). The complex then exchanges the strands which form a structure known as a holiday junction, followed by the ligation of swapped strands (Kowalczykowski et al., 1994).

The two homologous recombinational (HR) pathways, RecBCD and RecFOR, are present in *M. tuberculosis*. Both provide single-stranded DNA for RecA binding. The expression of *ruvA* and *ruvC* are up regulated upon DNA damage in *M. tuberculosis* (Boshoff et al., 2003). Taken together with previous observations that RuvC is important for the survival of *Helicobacter pylori* in macrophages (Loughlin et al., 2003), the data suggest that HR involving RuvC may also play an important role in repairing damage in *M. tuberculosis* (Brooks et al., 2001). Although RecJ is missing in *M. tuberculosis*, the expression of mycobacterial *recR* was shown to be up regulated upon capreomycin treatment, and both *recR* and *recF* are up regulated upon translational inhibition (Boshoff et al., 2003) suggesting that the RecFOR pathway is active in *M. tuberculosis*. The NHEJ pathway was once thought to be restricted to eukaryotes but was also identified in prokaryotes. The process in *M. tuberculosis* is aided by the DNA end binding protein Ku homo dimer and DNA ligase D (LigD) (Shuman and Glickman, 2007). The physical interaction between mycobacterial Ku and LigD stimulates the ligase activity of LigD on DSBs (Shuman and Glickman, 2007). Deleting Ku or LigD significantly lowers the efficiency of NHEJ-directed DSBs repair in mycobacteria (Shuman and Glickman, 2007). Moreover, *M. tuberculosis* contains a LigD-independent NHEJ pathway in which LigC may function redundantly in NHEJ. The absence of NHEJ sensitizes mycobacteria to desiccation and ionizing radiation (Rassool, 2003). NHEJ may be the only pathway available to repair DSBs during LTBI, assuming that no daughter chromatid is present for HR. NHEJ may also be important for repairing DSBs induced by clustered oxidative damages within macrophages and promoting the survival of the pathogen (Dos Vultos et al., 2009, Shuman and Glickman, 2007).

### 1.6.3 Nucleotide excision repair

Nucleotide excision repair (NER) identifies and removes bulky DNA lesions generated from UV irradiation and genotoxic compounds (Friedberg and Gerlach, 1999, Kurthkoti and Varshney, 2011, Kurthkoti and Varshney, 2012,). This repair pathway executes its function using several Uvr proteins (Muir et al., 1998). The Uvr proteins have low base specificity and act in combination, usually forming a complex because they remove multiple lesions. UvrA and UvrB identify and bind to the DNA lesion forming a UvrAB complex (Kurthkoti and Varshney, 2011, Kurthkoti and Varshney, 2012). UvrA is released when UvrC binds to UvrB.

The UvrBC complex cleaves and excises the surrounding upstream and downstream regions of the DNA damage (Sancar, 1996). Thereafter, UvrD removes the UvrBC complex with the lesion and the surrounding regions including 12-13 nucleotides which creates a gap (Kurthkoti and Varshney, 2011, Kurthkoti and Varshney, 2012). The gap is filled by DNA polymerase and ligase (Kurthkoti and Varshney, 2011, Sancar, 1996). In *M. tuberculosis* the *uvrB* gene is essential *in vivo* (in mice) for infection and *in vitro* (human macrophages) for survival (Darwin and Nathan, 2005). *M. tuberculosis* possesses two homologues of UvrD proteins - UvrD1 and UvrD2 (Cole et al., 1998, Mizrahi and Andersen, 1998). UvrD1 is a Ku dependent protein and is bi-functional (Guthlein et al., 2008). It has helicase and DNA-dependent ATPase activities that allow this protein to repair double and single strand breaks (Guthlein et al., 2008, Sinha and Hader, 2002). Hence, UvrD1 can function within NER and non-homologous end joining repair (Kurthkoti and Varshney, 2011, Shuman and Glickman, 2007). UvrD2 on the other hand is a helicase specific to NER that is essential for *M. tuberculosis* survival when exposed to UV irradiation (Boshoff et al., 2003, Kurthkoti and Varshney, 2011).

*M. tuberculosis* possesses all the genes involved in the NER pathway. The expression of *uvr* genes is enhanced in intracellular *M. tuberculosis* demonstrating the importance of NER for the pathogen's survival upon infection (Thi et al., 2012). In addition, these genes are also involved in damage caused by $H_2O_2$ and are up regulated upon UV exposure (Friedberg et al., 2005). Mutants lacking *uvrB* were highly susceptible to acidified nitrite and UV light and were markedly attenuated for survival in mice and bone marrow-derived macrophages (de Laat et al., 1999). Moreover, *M. tuberculosis* has two putative genes encoding DNA helicase II: *uvrD1* and *uvrD2* (de Boer and Hoeijmakers, 2000). The Ku-dependent UvrD1 seems to be essential for mycobacterial repair of DNA damage caused by UV and ionizing radiation and is up regulated upon $H_2O_2$ treatment (Boshoff et al., 2003, Sinha and Hader, 2002). The absence of UvrD1 attenuates *M. tuberculosis* during infection (Curtiet al., 2007). On the other hand, UvrD2 is a Ku-independent helicase, which is also up regulated upon UV radiation (Boshoff et al., 2003). The potentially redundant functions of UvrD1 and UvrD2 strongly suggest the important role of NER in *M. tuberculosis*.

## 1.7 Base excision repair pathway

Base excision repair (BER) is an important multistep, multi-enzyme DNA repair system. BER of modified nucleotides is initiated by damage-specific DNA glycosylases, Figure 1.5. The repair of the resulting apurinic/apyrimidinic site (AP site) involves the replacement of either a single nucleotide (short patch BER) or of several nucleotides (long patch BER) (Kurthkoti and Varshney, 2012). The mechanism that controls the selection of either BER pathway is unknown. DNA glycosylases cleave the *N*-glycosidic bond between the target base and the deoxyribose, releasing a free base, thus creating an AP site.

In the short patch model as shown in Figure 1.5, DNA glycosylases are the initial enzymes that recognize and excise damaged bases, (I). Following this an AP endonuclease cleaves the phosphodiester bond immediately 5′ to the AP site, generating 5′-sugar phosphate and 3′-OH ends as it nicks the DNA (II). Removal of the 5′- sugar phosphate moiety by a deoxyribophosphodiesterase (dRpase) (III) results in a single nucleotide gap that is then filled by a DNA polymerase (IV) and sealed by DNA ligase (V). As in short patch BER, AP sites in long patch BER are processed by an AP endonuclease which cleaves immediately 5′ to the AP site, generating 5′- sugar phosphate and 3′-OH ends (Kurthkoti and Varshney, 2012). However, in this process, the 5′-sugar phosphate residue is not removed by a dRpase, rather a DNA polymerase adds several nucleotides to the 3′ end of the nick displacing the 5′- sugar phosphate as part of a single stranded flap structure. This flap structure is recognized and excised by flap endonuclease (FEN1) and DNA is finally ligated by DNA ligase. This pathway is dependent on proliferating cell nuclear antigen (PCNA), which presumably plays a role in loading DNA polymerase onto the DNA as well as stimulating the activity of FEN1 (Dos Vultos et al., 2009).

Many DNA glycosylases have been discovered and isolated, and their reaction mechanisms and substrate specificities have been elucidated. Most of the known products of oxidative damage to DNA (described previously) are substrates of DNA glycosylases with broad or narrow substrate specificities. Some possess cross-activity and remove both pyrimidine- and purine-derived lesions (Boshoff et al., 2003). Overlapping activities between enzymes also exist. DNA glycosylases with overlapping substrate ranges provide back-up or

redundant functions, so that the absence of one will be compensated by another during the repair of a critical mutagenic lesion. The stability and correct function of the DNA is necessary for normal cellular functions and damage to the DNA can lead to cellular dysfunction, cancer and other diseases, or to cell death. Deficiencies and polymorphisms in DNA glycosylase-encoding genes have been shown to be related to human disease susceptibility (Hazra et al., 2001).

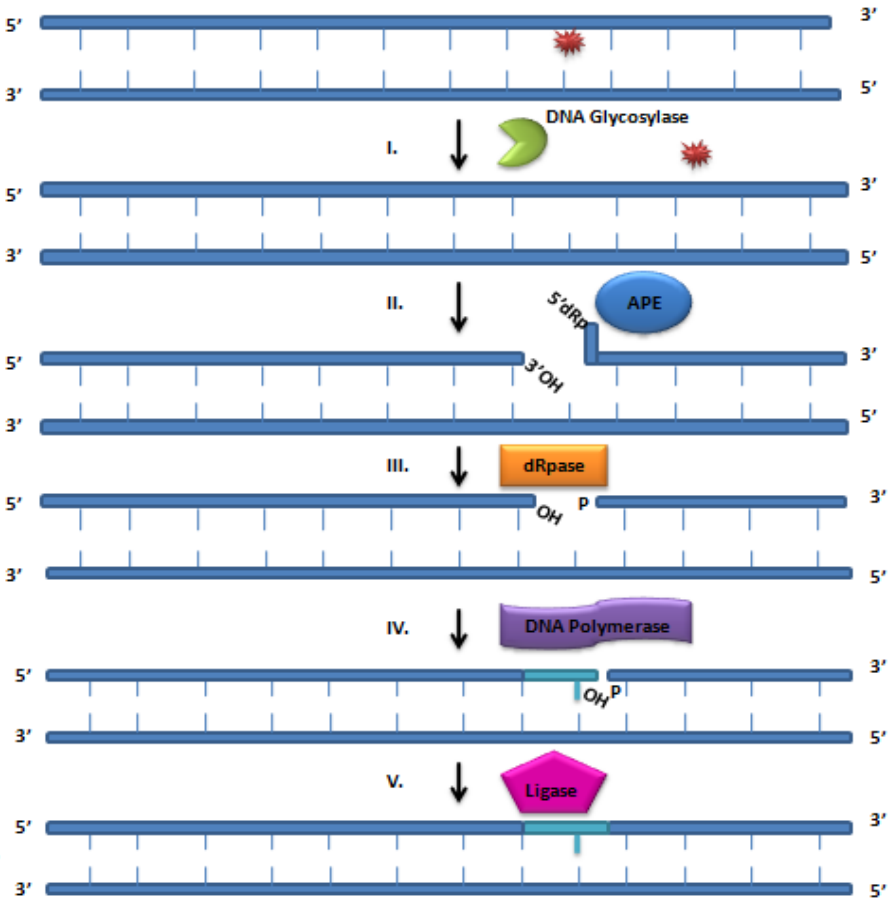



**Figure 1.5: Schematic diagram of the base excision repair pathway**
I. Oxidatively damaged DNA base is recognized and excised by DNA glycosylases (as shown by the red star) generating an AP (apurinic/apyrimidic) site. II. The AP endonuclease cleaves the sugar-phosphate backbone to generate a single strand break at the 3' hydroxyl group. III. Poly nucleotide kinase and a deoxyribophosphodiesterase respectively, excise the 5'-terminal deoxyribose and change the 3'-terminal to a hydroxyl group. IV. DNA polymerase fills in the new correct DNA base and, V. DNA ligase seals the gap (Kurthkoti and Varshney, 2011).

## 1.9 DNA glycosylases

DNA glycosylases fall into two general families based on structural and sequence homology (David and Williams, 1998). The Nth superfamily includes endonuclease III (Nth), which removes both oxidized pyrimidines, such as Tg, urea, DHU and formamidopyrimidines: FapyA and FapyG; MutY which removes A mispaired with 8-oxoG and FapyG; Ogg1 whose primary substrate is 8-oxoG and FapyG opposite C; and AlkA which removes alkylated bases. The Fpg/Nei family is defined by formamidopyrimidine DNA glycosylase (Fpg), also called MutM, and endonuclease VIII (Nei). Fpg has substrate specificity similar to Ogg and removes 8-oxoG and FapyG paired with C as well as FapyA paired with T. The hydantoin lesions, Gh and Sp, are also substrates for Fpg (Leipold et al., 2000). In addition to purine lesions, Fpg also recognizes oxidized pyrimidines such as 5-OHC, 5-hydrouracil, and 5,6-dihydroxy-5,6-dihydroxythymine but to a much lesser extent (Saparbaev et al., 2002). Nei has substrate specificity similar to that of Nth, removing oxidized pyrimidines and FapyA. Gh and Sp lesions are also removed by endonuclease VIII (Nei) (Mitra et al., 2001). MutY and AlkA do not have glycosylase activities and bind the substrate to sequester it for other DNA repair enzymes; the other glycosylases listed above are bi-functional and exhibit lyase activity in addition to glycosylase activity. These DNA glycosylases exhibit redundancy with respect to substrate specificities; therefore the loss of a single DNA glycosylase usually does not result in a complete loss of repair for a specific oxidative lesion.

## 1.10 Preliminary studies in mycobacteria that are important for the work presented in this dissertation

Several studies using gene knockout strategies were conducted to assess the various roles and functions of the base excision repair DNA glycosylases in mycobacteria. A study conducted by Goosens using bioinformatics analyses identified two copies each of Fpg and Nei-encoding genes in *M. tuberculosis* as well as in its non-pathogenic relative, *Mycobacterium smegmatis* (Goosens, 2009). To understand the role of these multiple glycosylases in the maintenance of genomic integrity and survival of mycobacteria, the genes encoding the four Fpg/Nei glycosylases were individually deleted in the *M. smegmatis* strain $mc^2155$ by homologous recombination. In addition to the four single mutants, double and triple Fpg and Nei glycosylase knockout mutants were generated by sequential gene

inactivation. Their observation was that when comparing the parental strain, the single and double mutants showed no variation in growth kinetics, no increased sensitivity to hydrogen peroxide and no increase in spontaneous mutation rates (Moolla et al., 2014). However, a slight increase in frequency of spontaneous C - T transition mutations was observed in double knockout mutants compared to the wild type and single mutant strains. The results suggested that these enzymes may be part of an extensive network of enzymes which collectively work to enhance the overall survival of *M. smegmatis* through the repair of oxidatively damaged DNA.

A study conducted by Hassim et al aimed to understand the role of MutY in combination with the Fpg/Nei family of DNA glycosylases in mycobacteria and whether these DNA glycosylases display overlapping and/or compensatory functions in dealing with oxidative damage (Hassim et al., 2015). Using homologous recombination, the *mutY* gene was deleted in the parental and in the Fpg/Nei deficient mutant strains. Deletion of *mutY* together with the Fpg/Nei family of DNA glycosylases yielded similar *in vitro* growth kinetics as the parental strain under normal culture conditions. However, under *in vitro* oxidative stress conditions, *mutY* deficiency, especially in the absence of all four Fpg/Nei DNA glycosylases, resulted in a greater reduction in survival of the mutants, with a general increase in mutation rates. Consistent with these data was the significant increase in C → A and A → C mutations with the progressive loss of the DNA glycosylases as assessed by the spectral analysis of rifampicin resistant mutants. Taken together, the authors concluded that the mycobacterial MutY DNA glycosylase has antimutator properties and possibly has a more significant role in mycobacterial genome maintenance compared to the Fpg/Nei family of DNA glycosylases.

Previously, it has been shown that the lack of Fpg/Nei DNA glycosylases in *M. smegmatis* resulted in no differences in growth and survival under normal and oxidative stress conditions, with no increase in spontaneous mutation rates as compared to the parental strain, suggesting that *nth* may also be a significant role player in mycobacterial genome maintenance (Moolla et al., 2014). These findings led to the study conducted by Moolla et al. wherein the *nth* gene was site specifically inactivated by homologous recombination in the parental *M. smegmatis* strain and in selected combinatorial mutant strains deficient in the

Fpg/Nei glycosylases (Moolla et al., 2014). Inactivation of the *nth* gene did not affect the *in vitro* growth of the mutant strains under normal culture conditions. Interestingly, UV induced DNA damage of the single *nth* deletion mutant resulted in a dramatic increase in mutation frequency that was not observed in any of the other single DNA glycosylase deletion mutants. The progressive loss of *fpg*, *nei* and *nth* genes resulted in exaggerated survival defects under oxidative stress. Furthermore, the subsequent deletion of *nth* in mutants deficient in *fpg*/*nei* resulted in a dramatic increase in spontaneous mutation rates and frequencies, implying that *nth* is integral for the repair of both spontaneous and induced DNA damage. These results indicate that the *M. smegmatis nth*-encoding DNA glycosylase is involved in DNA repair and has anti-mutator properties and that it is able to exert a combinatorial effect with the Nei homologues in mutation avoidance. As a result, this study highlighted a novel interplay between the Nth and Nei homologues in DNA damage tolerance under oxidative stress conditions and in spontaneous mutagenesis.

### 1.11 Aims and objectives of the study

As mentioned above it was previously shown in our laboratory that the Nth DNA glycosylase was a key enzyme of the BER pathway responsible for preventing and/or reducing mutations in *M. smegmatis*. The previous study also demonstrated that combinatorial deletion of Nth and Nei in *M. smegmatis* resulted in reduced survival under oxidative stress conditions with a corresponding increase in mutation rates, suggestive of interplay between these enzymes (Moolla et al., 2014). Hence, it was important to assess the combined role of these genes in maintaining mycobacterial genome integrity. In this study we aimed to highlight the hierarchical role of the two Nei homologues (NeiI and NeiII) in BER. The single deletion mutant lacking the *nth* gene, that were generated previously (Moolla et al., 2014, Hassim et al., 2015), was used to generate double mutants that lacked both the *neiI* and *nth*, as well as *neiII* and *nth* genes.

Aim: The overall aim of this study was to further investigate the interplay between the Nth and NeiI/NeiII homologues through combinatorial gene deletion and subsequent characterization of mutant strains.

Specific objectives: To achieve the above-mentioned aim the following specific objectives were set out:

1. Deletion of NeiI in the Nth-defective mutant using two-step allelic exchange mutagenesis.
2. Deletion of NeiII in the Nth-defective mutant using two-step allelic exchange mutagenesis.
3. Genomic confirmation of the resulting mutant strains using PCR and Southern blotting
4. Genetic complementation of the mutant strains using integrating vectors to deliver single gene doses
5. Confirmation of gene expression in genetically complemented strains.
6. Characterization of the mutant and genetically complemented strains in vitro under normal culture conditions and under oxidative stress.
7. Assessment of mutation rates in combinatorial mutants.

# 2. Materials and methods

## 2.1 Bacterial strains, plasmids and culturing conditions

All bacterial strains and cloning vectors used in this study are listed in Table 1.1 and Table 1.2 respectively. Glycerol stocks of bacterial strains were prepared in 66 % glycerol and stored at -80 °C.

**Table 1.1: Bacterial strains**

| Bacterial strain | Genotype | Reference |
|---|---|---|
| ***M. smegmatis* mc$^2$ 155** | *ept*-1, High frequency plasmid transformation mutant of mc$^2$6 | (Snapper et al., 1990) |
| **Δ*neiI*** | Mutant of mc$^2$155 carrying an internal 486 bp deletion in *neiI* | (Goosens, 2009) |
| **Δ*neiII*** | Mutant of mc$^2$155 carrying an internal 637 bp deletion in *neiII* | (Goosens, 2009) |
| **Δ*nth*** | Derivative of *M. smegmatis* mc$^2$155 containing an internal 570 bp deletion in *nth* | (Moolla et al., 2014) |
| **Δ*neiI* Δ*nth*** | Derivative of Δ*neiI* carrying an internal 570 bp deletion in *nth* | This study |
| **Δ*neiII* Δ*nth*** | Derivative of Δ*neiII* carrying an internal 570 bp deletion in *nth* | This study |
| **Δ*neiI* Δ*nth::nth*** | Derivative of Δ*neiI* Δ*nth* carrying pTWEETY::*nth* integrated at the *attP* phage attachment site | This study |
| **Δ*neiII* Δ*nth::nth*** | Derivative of Δ*neiII* Δ*nth* carrying pTWEETY::*nth* integrated at the *attP* phage attachment site | This study |

**Table 1.2: Cloning vectors**

| Plasmids | Characteristics | Size(bp) | Reference |
|---|---|---|---|
| **pGEM3Zf(+)** | *E. coli* cloning vector; amp$^R$; *lacZ*-alpha; *ori*E | 3199 | Promega |
| **p2NIL** | *E. coli* cloning vector and mycobacterial suicide plasmid; kan$^R$; *ori*E | 4753 | (Parish and Stoker, 2000) |
| **pGOAL19** | Plasmid carrying *lac*Z, *hyg* and *sacB*g enes as a *Pac*I cassette; amp$^R$, *ori*E | 10435 | (Parish and Stoker, 2000) |
| **p2NILΔ*nth*::pGOAL17** | p2NIL suicide vector carrying the deleted *nth* gene and the selectable and counter selectable markers from pGOAL17, kan$^R$; *lacZ*; *sacB*; *oriE* | 14508 | (Hassim et al., 2015) |
| **p2NILΔ*nth*::pGOAL19** | p2NIL suicide vector carrying the deleted *M. smegmatis nth* gene and the selectable and counter selectable markers (*hyg-lacZ-sacB)* from pGOAL19; kan$^{R,}$ hyg$^R$; *lacZ*; *sacB*; *oriE.* | 14748 | (Moolla et al., 2014) |
| **pTWEETY** | Mycobacterial integrating vector; kan$^R$, *ori*E, *int* | 5835 | (Pham et al., 2007) |
| **pTWEETY:::*nth*** | pTWEETY integrating vector carrying *nth*: kan$^R$ | 6990 | (Moolla et al., 2014) |
| **pSE100** | *E. coli*-Mycobacterial shuttle vector carrying Pmyc1*tetO*; | 5538 | (Ehrt and Schnappinger, |

| | | | |
|---|---|---|---|
| | hyg$^R$ | | 2009a) |

amp$^R$-ampicillin resistance; *ori*E–origin of replication in *E. coli*; kan$^R$-kanamycin resistance; hyg$^R$-hygromycin resistance

### 2.1.1 Bioinformatics analysis

Various bioinformatics tools were used to identify and analyse the Nei DNA glycosylases in *M. smegmatis* and *M. tuberculosis*. Genome sequence data for *M. tuberculosis* was obtained from the Tuberculosis database TubercuList (http://genolist.pasteur.fr/TubercuList/) and for *M. smegmatis* from the Smegmalist database (http://cmr.tigr.org).

### 2.1.2 *E. coli*- DH5α strains.

*E. coli* cells containing plasmids were grown overnight at 37 °C in Luria-Bertani (LB) broth containing the appropriate antibiotics with vigorous shaking (Labcon Shaking Incubator). For selection on solid media, *E. coli* containing plasmids were grown overnight in a Labotec incubator at 37 °C on Luria-Bertani agar (LA) containing the appropriate antibiotics. *E. coli* strains carrying plasmids >8000 bp was grown at 30°C for 48 hours to avoid plasmid rearrangement.

The following antibiotics were used at these concentrations: 200 µg/mL Ampicillin (amp), 200 µg/mL hygromycin (hyg) and 50 µg/mL kanamycin (kan). For counter selection of clones carrying *sac*B, 5 % w/v sucrose was used. For confirmation of disruption of the *lac*Z cassette during cloning or identification of a clone containing the *lac*Z cassette, 40 µg/mL of 5-bromo-4-chloro-3-indolyl- α -galactosidase (X-gal) was used.

### 2.1.3 Mycobacterial strains.

*M. smegmatis* strains, unless otherwise stated, were grown in Middlebrook 7H9 media supplemented with glucose salt and 0.05 % Tween or on Middlebrook 7H10 solid media. All culturing of *M. smegmatis* strains was performed in a Bio-safety Level 2 laboratory. For selection of strains containing vectors with various selection markers, 50 µg/mL hyg and 25 µg/mL kan was used, and for counter selection of clones carrying the *sacB* and *lac*Z cassettes*,* 2 % w/v sucrose and 40 µg/mL X-gal was included in 7H10 plates

## 2. 2 DNA extractions.

### 2.2.1 Plasmid DNA extraction from *E. coli* (Sambrook et al., 1989).

#### 2.2.1.1 Small scale plasmid DNA extraction.

A single colony of *E. coli* carrying the plasmid of interest was re-suspended in 2 mL of LB media supplemented with the appropriate antibiotic and grown overnight at 37 °C with shaking at 250 revolutions per minute (rpm). A one mL aliquot was transferred to a 1.5 mL micro centrifuge tube and the cells harvested at 12 470 × *g* for 5 min*.* The pellets were re-suspended in 100 µL of Solution I (50 mM glucose, 25 mM Tris-Cl pH 8.0, 10 mM EDTA). To this suspension, 200 µL of Solution II (1% SDS, 0.2 M NaOH) was added and mixed by gently inverting tubes 6-8 times. The mixture was incubated at room temperature for 5 min; lastly 150 µL of Solution III (3 M Potassium acetate, pH 5.5) was added and the tubes were incubated on ice for 5 min. Thereafter, the suspensions were centrifuged at 12 470 × *g* for 5 min. The supernatants were decanted into a fresh 1.5 mL micro centrifuge tube containing 1 µL RNaseA (10 mg/mL) and incubated at 42 ˚C for 15 min. To precipitate the plasmid DNA, 600 µL of iso-propanol was added and the iso-propanol-DNA mixture was incubated at –20 ˚C for ~ 30 min, followed by centrifugation at 12 470 × *g* for 30 min. The iso-propanol supernatant was poured off gently and the pellet was washed with 1 mL of 70% EtOH and dried at 45 °C in a Speed Vacuum concentrator, (5301 Eppendorf). The DNA was re-suspended in 20 – 30 µL sterile distilled water (sd$H_2$0).

#### 2.2.1.2 Large scale plasmid DNA extraction.

For large scale DNA extractions, *E. coli* cultures were grown in 50 mL LB, overnight with shaking at 100 rpm, and cells were harvested by centrifugation in a Beckmann J2-21 centrifuge (3901 × g for 15 min at 4 °C). The DNA extraction method was as described above except that the solution volumes were increased as follows: 800 µL of Solution I, 1600 µL of Solution II and 1200 µL of Solution III.

The small scale plasmid DNA extraction protocol was followed up until the 70 % ethanol wash step and larger volumes of RNaseA (3 µL) and iso-propanol (800 µL) were used. Following the ethanol wash, pellets were dried and re-suspended in 50-100 µL of $sdH_2O$.

### 2.2.2 Chromosomal DNA extraction.

#### 2.2.2.1 Small scale chromosomal DNA extraction (colony boil).

Half a colony (~10 mm diameter) of *M. smegmatis* was re-suspended in 50 µL of $sdH_2O$. An equal volume of chloroform was added to the suspension, mixed vigorously and boiled at 95 ˚C for 5 min. The mixture was centrifuged at 12 470 × *g* for 5 min to separate the cell debris. The aqueous phase above the cell debris interface was removed into a fresh 1.5 mL micro centrifuge tube and used as DNA template in PCR reactions (see section 2.3.1).

#### 2.2.2.2 Large scale chromosomal DNA extraction.

Mycobacterial chromosomal DNA was isolated using a modified cetyltrimethylammonium bromide (CTAB; ICN Biomedicals, Aurora, Ohio) method (Somerville et al., 2005). Briefly, mycobacterial cells were harvested and re-suspended in 500 µL TE buffer (10 mM Tris-HCl pH 8.0, and 1 mM EDTA). The cells were heat killed at 65 °C for 10 min to which 50 µL of lysozyme (10 mg/mL) was added and the suspension was incubated at 37 °C for 1 hour. A solution of 70 µL 10 % SDS and 6 µL proteinase K (10 mg/mL) was added followed by incubation at 65 °C for 1 hour. Following this, 100 µL of a 5 M solution of sodium chloride and 100 µL pre-warmed CTAB/NaCl mix (10 % CTAB made in 0.7 M NaCl) was added to the sample and incubated at 65 °C for a further 10 min. An equal volume of chloroform: isoamyl alcohol

(24:1 v/v) was added to remove proteins. Subsequent to centrifugation at 6168 × g for 30 min, the aqueous phase containing the DNA was precipitated by addition of 420 µL iso-propanol and incubation at -20 °C for 1 hour. The DNA was pelleted by centrifugation (16168 × g for 20 min), washed with ice-cold 70 % ethanol, dried in a vacuum centrifuge and re-suspended in 50 -100 µL sdH2O

## 2.3 DNA manipulations.

### 2.3.1 Polymerase chain reaction.

#### 2.3.1.1 Primer design.

Primers were designed using the online program Primer3 (http://frodo.wi.mit.edu/) which suggested the most appropriate primer sequences. Primers were purchased and synthesized from Inqaba Biotec.

#### 2.3.1.2 PCR reactions.

For PCR reactions using FastStart Taq DNA polymerase (Thermo Scientific™), the reactions were set up in 50 µL volumes with the following components: 1× recommended buffer, 1× GC Rich solution, dNTPs to a final concentration of 0.2 mM each, forward and reverse primers to a final concentration of 1 µM each, DNA template between 10-100 ng/µL and 2 U enzyme. Reactions were made up to volume with sterile distilled nuclease free water. Cycling conditions were carried out as follows: one cycle of an initial denaturation at 94 °C for 4 min; 30-35 cycles of 30 sec denaturation at 94 °C, 30 sec annealing at 55-65 °C and 30-90 sec elongation at 72 °C which was followed by a final elongation step at 72 °C for 5-7 min.

### 2.3.2 Restriction enzyme digestions.

All restriction enzymes and buffers used were purchased from New England Biolabs (NEB) or Roche Applied Sciences. Reactions were carried out according to the manufactures instructions. All digestion reactions were incubated at 37 °C with the appropriate buffer for optimal enzyme performance. For plasmid DNA 1-2 µg was digested in a 10-20 µL reaction and incubated at 37 °C for 1-2 hours. For chromosomal DNA, 2-5 µg was digested in a 50 µL reaction and incubated at 37 °C for 8-16 hours.

## 2.3 Agarose gel electrophoresis.

For the separation of high molecular weight DNA fragments, 0.8 % - 1 % agarose gels; prepared in 1 × TAE buffer (1 mM EDTA, 40 mM Tris-acetic acid pH 8.5); were used. For low molecular weight DNA fragments in the size range of 1 kb or lower, 2 % agarose gels were used. All gels contained 0.5 µg/mL ethidium bromide and the DNA samples were loaded with a tracking dye (0.025 % bromophenol blue in 30 % glycerol). Lambda DNA molecular weight markers (III, IV, and VI respectively; Roche Applied Science) were used to assess DNA fragment sizes. The agarose gels were electrophoresed at 80 –100 V in a Mini-Sub Cell GT mini gel horizontal submarine unit (BIO-RAD) and DNA visualized under UV-light using the Gel Doc 2000 system (BIORAD).

## 2.4 Bacterial Transformations.

### 2.4.1 Preparations of chemically competent *E coli*-DH5α cells.

Rubidium chloride cells were prepared as follows. Briefly, 1mL of a stationary phase overnight culture was inoculated into 100 mL LB and grown to an $OD_{600nm}$ of between 0.5- 0.8. The cells were chilled on ice for 15 min and harvested (3901 × g for 5 min at 4 °C). The pellets were re-suspended in 20 mL TfbI solution (30 mM potassium acetate, 100 mM rubidium chloride, 10 mM calcium chloride, 50 mM manganese chloride, and 15 % v/v glycerol - pH 5.8), and chilled on ice for 15 min. The cells were re-harvested (3901 × g for 5 min at 4 °C), re-suspended in 2 mL TfbII solution (10 mM MOPS [3-N-morpholino propanesulfonic acid], 75 mM calcium chloride, 10 mM rubidium chloride and 15 % v/v glycerol-pH 6.5) and 500 µL aliquots were flash-frozen in ethanol and stored at -80 °C until further use.

### 2.4.2 Transformations into *E. coli*-DH5α competent cells.

*E. coli*-DH5α competent cells stored at -80 °C were thawed on ice and a 100 µL aliquot was used per transformation. Up to 1 µg plasmid DNA was incubated with the cells on ice for 1 hour, heat-shocked for 90 sec at 42 °C and chilled on ice for 2 min. One mL of 2XTY (Tryptone Yeast) was then added to rescue the cells at 37 °C for 1hour. The cells

were then spread on LA media containing the appropriate antibiotics, and incubated for 1 - 2 days at 37 °C.

### 2.4.3 Preparation of competent *M. smegmatis* cells.

One millilitre of a stationary phase *M. smegmatis* culture was inoculated in 100 mL LB containing 0.05 % Tween80 and grown to an $OD_{600nm}$ of 0.5 - 0.8. The cells were harvested by centrifugation (2360 × g for 10 minutes at 4°C) and the pellet washed thrice by gentle re-suspension in 30 mL ice-cold 10 % glycerol and harvested by centrifugation at 2360 × g for 10 min at 4°C. The pellet was suspended in 1 mL ice-cold 10 % glycerol and these competent cells were used immediately.

### 2.4.4 Electroporation into competent *M. smegmatis* cells.

A maximum of 5 µg plasmid DNA was added to 400 µL of *M. smegmatis* competent cells and incubated on ice for 15 min. This mixture was transferred to a 0.2 cm electroporation cuvette and pulsed using the Biorad Gene PulserX cells™ electroporator at the following settings: 2.5 kV, 25 µF and 1000 Ω. The cells were rescued immediately with 800 µL of 2XTY for at least 3-16 hours at 37 °C. The cells were subsequently spread on Middlebrook 7H10 media containing the appropriate supplements and antibiotics, and incubated for 3 - 7 days at 37 °C.

## 2.5 Southern Blot analysis (Southern, 1975).

Approximately 2-3 µg of genomic DNA was digested overnight at 37 °C with the appropriate restriction enzyme and the DNA fragments were separated on a 0.8% agarose gel at 80 V for 2 hours. The DNA was initially depurinated by treating the agarose gel with a 0.25 M HCl solution for 15 min, then denatured by soaking in a 0.5 M NaOH/ 1.5 M NaCl solution for 15 min. The gel was subsequently equilibrated in 1 ×TBE buffer (Tris-Borate-EDTA pH 8.0, Sigma). The agarose gel was overlaid with a Hybond™ – N nitrocellulose membrane, sandwiched between two pre-soaked Whatman filter papers, and two pre-soaked sponges. After ensuring that no air bubbles were present, this “sandwich” was carefully placed in a TE 22 Mini Transphor cassette which

was transferred to a TE 22 Mini Transphor unit (OmniPAGEElectro-blotting Unit (Cleaver Scientific Ltd CS-300V) containing 1 × TBE buffer. The DNA was transferred to the nitrocellulose membrane (0.5 A for 2hours at 4°C) and cross-linked by irradiation at 1200 mJ/cm$^2$ in a UV Stratalinker (1800 Stratagene). Thereafter, the DNA was transferred onto a nitrocellulose membrane, the membrane was pre-hybridised in pre-hybridisation solution (0.5 % SDS, 6 ×SSC, 5 × Denhardts and 50 % deionised formamide) at 42 °C with 10 µg/mL heat denatured salmon sperm DNA (Roche Applied Science) in a Techne HybridizerHB-1 roller bottle.

The probe was hybridised to the membrane with 12 mL DIG (Dioxygen) EasyHyb solution for 1 hour at 56 °C in Hybaid HB-OV-BM roller bottles incubated in a hybridization oven (Hybaid Micro-4). Prior to hybridization, the probe was heat-denatured (95 °C for 10 min) and added to the pre-hybridisation solution and hybridization was allowed to occur overnight at 56 °C. The membrane was then washed twice with Solution I (2× SSC, 0.1 % SDS) at room temperature for 5 min and then once with Solution II (0.5 × SSC, 0.1 % SDS) at 65 °C for 30 min. Following these washes, detection of the DIG-labelled hybrids was carried out by dephosphorylation of the CSPD® substrate upon addition of a specific DIG alkaline phosphatase-conjugated antibody, which resulted in chemiluminescence at 477 nm. This enabled visualization of DNA on X-ray film after 30 min–4 hours of incubation at room temperature.

## 2.6 Knockout vectors used in this study

Plasmid constructs containing inactivated alleles of the *nth* and *nei* genes were previously generated by PCR and cloning as part of a Masters study (Goosens, 2009, Moolla et al., 2014). The strategy for creating the knockout vector is shown in Figure 2.1. Briefly, approximately 1500 bp upstream (green bar) and downstream (orange bar) (shown in Figure 2.1 below) of the genes of interest were amplified by PCR using a high fidelity polymerase and primers that had appropriate restriction sites incorporated within the sequence to facilitate cloning. The amplified upstream and downstream fragments were cloned into pGEM3Zf (+) and sequenced to ensure mutations were not introduced during the PCR cycling (II, in Figure 2.1). Subsequently, the upstream and downstream

fragments were excised from pGEM3Zf (+) and cloned by three way cloning into the suicide vector, p2NIL which carries a kan$^R$ marker (III – in Figure 2.1). The *Pac*I cassette from pGOAL19 containing marker genes *sac*B and *lac*Z and hyg$^R$ was inserted into the *Pac*I site of p2NIL to generate 2NILΔ*nth*::pGOAL19 (IV – in Figure 2.1).

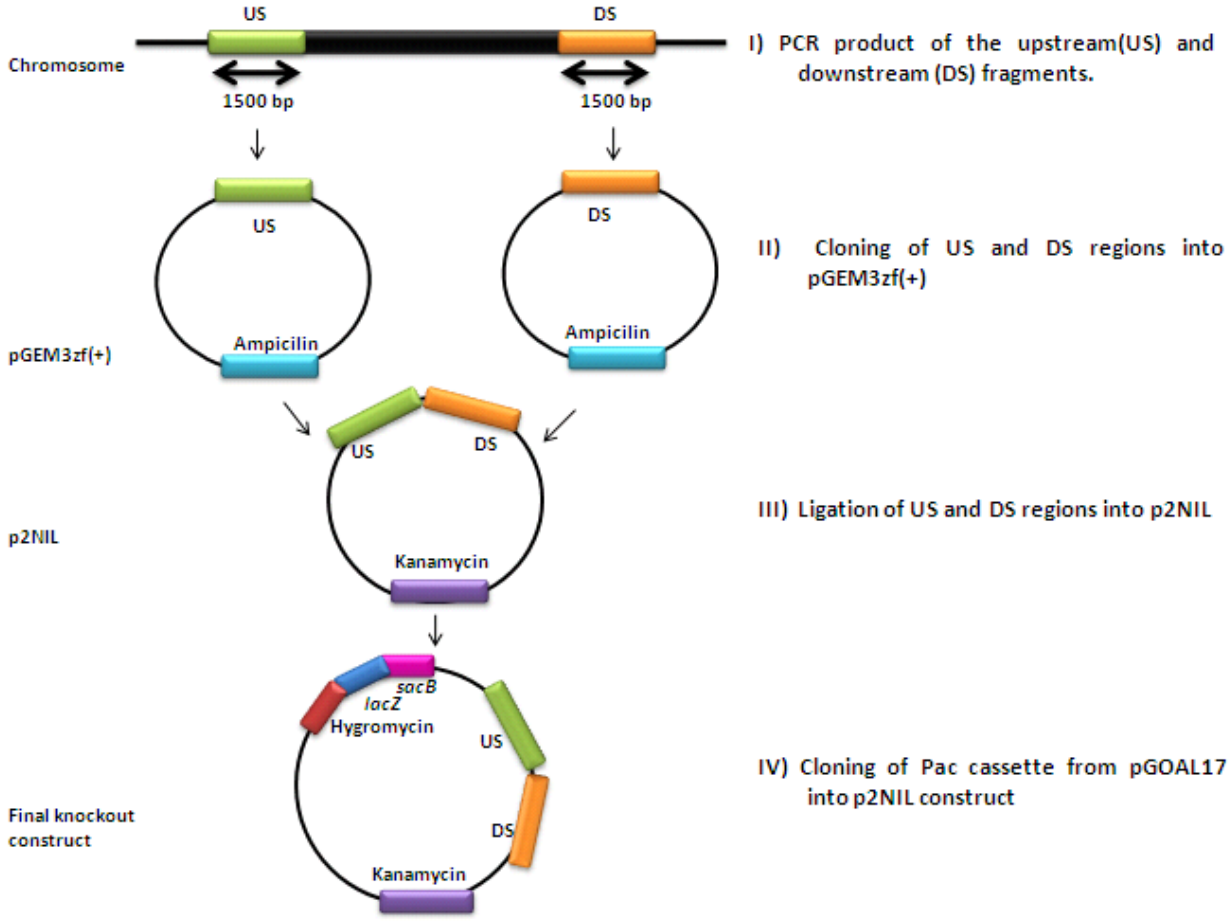



**Figure 2.1: Schematic representation of the construction of knockout constructs using three way cloning.** Initially PCR products of upstream region (US - green) and downstream (DS - orange) region are cloned separately into the pGEM vector, subsequently the US and DS regions and a Pac cassette are ligated together into p2NIL vector to make the final construct. The blue bar represents ampicillin resistance, purple bar represents kanamycin resistance and red bar represents hygromycin resistance. The black arrows represent regions amplified by the forward and reverse primers.

## 2.7 Inactivation of the Nth DNA glycosylase to generate double deletion mutants.

Two-step allelic exchange by homologous recombination, described by (Gordhan and Parish, 2001), was used to generate *M. smegmatis* knock-out mutants. A schematic representation of this technique is depicted in Figure 2.2 below. The method involves the construction of a knockout vector, shown in Figure 2.1. Once introduced into the *M. smegmatis* cell, a single cross-over (SCO) event between the chromosome and one of the regions of homology on the knockout vector results in the integration of the knockout vector into the chromosome. The knockout vector carries the selectable marker genes *aph*, *hyg* and *lacZ* which allow for the selection of SCO homologous recombinants, identified as blue colonies growing in the presence of X-gal, kan and hyg. Also carried on the knockout vector is the *sacB* gene, which encodes levansucrase. When grown in the presence of sucrose, cells carrying the *sacB* gene produce levansucrase which converts sucrose to fructose polymers. Accumulation of these polymers in the cell envelope becomes toxic, resulting in cell death. The *sacB* gene therefore serves as a counter-selectable marker, facilitating a second cross-over event which would result in the expulsion of the vector backbone to either generate a knock-out mutant or reconstitute the wild type allele if the second cross over occurs on the same side. Cells in which the second cross-over event occurs will be white in the presence of X-gal and will be sucrose resistant.

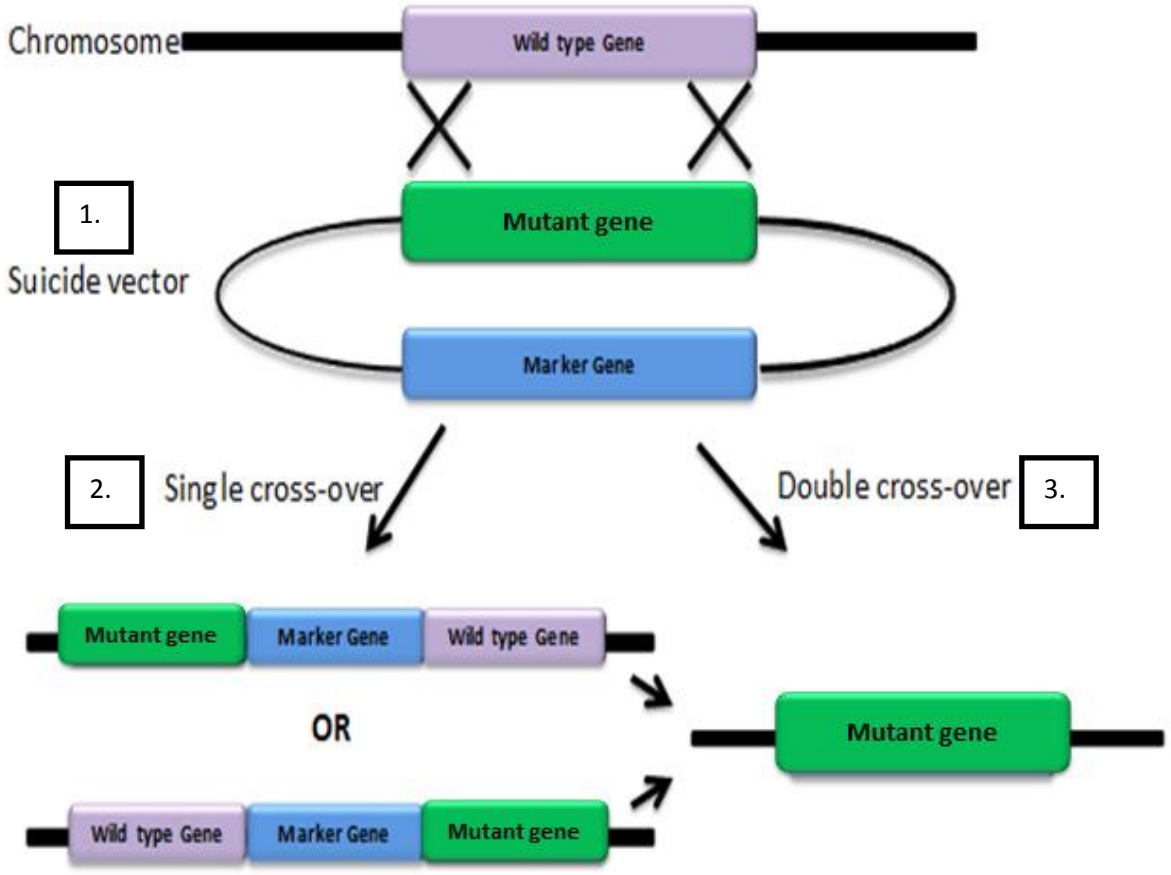



**Figure 2.2: Schematic representation of knock out mutant generation.**

A non-replicating plasmid carrying a mutated version of the gene is introduced into *M. smegmatis 1*. Homologous recombination leads to single cross-overs (SCO) 2, where a single recombination event has occurred or double cross-overs (DCO) 3 where two recombination events have occurred. Single crossovers carry both the wild-type (purple bar) and mutated alleles (green bar) and marker genes (blue bar) found on the plasmid, whereas double cross-overs will carry only the deleted copy of the gene (green bar) and no marker genes.

### 2.7.1 Generation of SCO and DCO transformants.

The p2NILΔ*nth*::pGOAL19 suicide plasmid was electroporated into mid-logarithmic phase cultures of the Δ*neiI* and Δ*neiII* deletion mutant strains respectively and selected on 7H10 plates containing kan and X-gal to identify the SCO's. Blue colonies on the plate were considered to be potential SCO's that result due to the hydrolysis of X-gal by

β-galactosidase, encoded by the *lacZ* gene in p2NIL, which is integrated into the genome. The genotype of the SCO's was confirmed by PCR and Southern blot analysis. For the PCR, chromosomal DNA was extracted from potential SCO's using the colony boil method and the DNA amplified using primers listed in Appendix B. Table 6.1 that distinguished the mutant and the wild type alleles.

To generate DCO mutants, one of the PCR confirmed SCO's was selected and grown overnight in 10 mL 7H9 with kanamycin. The overnight culture was inoculated into 20 mL of fresh 7H9 without antibiotics and grown at 37 °C overnight to allow for the second crossover to take place. The cells were serially diluted (10-fold) and 100 µL of each dilution ($10^0$-$10^8$) was spread on 7H10 media supplemented with 2 % sucrose and X-gal. Plates were incubated at 37 °C and scored for white sucrose resistant colonies after 3 days, which were potential DCO mutants. Potential DCO's were further confirmed for kanamycin sensitivity by spotting on 7H10 media with and without the antibiotic. The colonies with confirmed kanamycin sensitivity were screened initially by PCR to distinguish between wild type and mutant alleles. The PCR confirmed mutants together with the SCO and parental stain were then analysed by Southern blot to ensure that the genes upstream and downstream of the mutated allele were genotypically correct.

## 2.8 Complementation of mutant strains.

To confirm whether the observed phenotypes were due to the inactivation of the genes of interest, complementation strains were generated. The complemented constructs were generated by a previous Masters student (Moolla et al., 2014). Briefly the full length functional *nth* gene with an additional 350 bp upstream sequence carrying the promoter region was amplified and cloned into the integrating vector pTWEETY to generate pTWEETY::*nth.* In this study the integrity of the construct was re-evaluated by restriction analysis. The appropriate complementation vector was electroporated into the respective mutant strain and the resulting clones were screened by PCR. Using this strategy, two complemented strains, (Δ*neiI* Δ*nth::nth* and Δ*neiII* Δ*nth*::*nth)* were generated.

## 2.9 Phenotypic analysis of mutants.

### 2.9.1 Growth Kinetics under normal culture conditions.

To determine the growth kinetics of the various mutant strains, 1 mL freezer stocks were grown overnight at 37 °C in 5 mL 7H9 medium with shaking at 100 rpm. The overnight late log phase culture was diluted into 50 mL of fresh media to a starting $OD_{600nm}$ of 0.02 and growth of the cultures was monitored by $OD_{600nm}$ measurements every 3 hours for a period of 36 hours. Growth was represented graphically as optical density over time.

### 2.9.2 Mutation frequency.

For mutation frequency analysis, bacterial strains were grown overnight in 20 mL 7H9 at 37 °C to an $OD_{600nm}$ of 0.3-0.5. The cultures were washed twice by centrifuging at 16168 *x g* for 10 min and re-suspended in 20 mL of 7H9 broth. Aliquots (100 µL) of 10 fold serial dilutions ($10^4$-$10^7$) were spread onto 7H10 media for colony forming unit (CFU) counts. A 1 mL aliquot of the cells was spread onto 7H10 supplemented with 200 µg/mL rifampicin (rif) and incubated at 37 °C. The plates were scored after 7 days for rifampicin resistant ($rif^R$) colonies. The mutation frequency was calculated as total number of $rif^R$ colonies over the total number of colony forming units (CFU/mL) per culture.

### 2.9.3 Sensitivity to hydrogen peroxide.

To assess the sensitivity of the various mutants under oxidative stress conditions, about 20 mL of the various mutant and complemented strains were grown to an $OD_{600nm}$ of 0.3-0.4. A 5 mL aliquot of each culture was dispensed into a flask and treated with 2 mM of hydrogen peroxide. The treated cultures were incubated at 37 °C for three hours in the dark. After the three hours incubation, 100 µl aliquots of 10 fold serial dilutions ($10^4$-$10^7$) were spread onto 7H10 media for CFU counts.

### 2.9.4 Spontaneous mutation rates.

A mutation rate measures the probability of a cell acquiring a mutation during its lifetime and takes into account the possibility of mutations arising early or late during the growth phase. The Luria-Delbrück fluctuation analysis was used to measure spontaneous

mutation rates for the various mutants (Rosche and Foster, 2000). The numbers of mutational events that occur are accurately calculated by using parallel cultures, thereby minimizing variability. The mutation rate was calculated using a previously created excel spreadsheet (E. Machowski, CBTBR). Briefly, mutation rate is taken as the number of observed mutations per culture (*m*) within the total population (*Nt*). There are two methods suggested for estimating the overall mutation rate of the population: the p0 method, which is based on the proportion of cultures in which there are no mutants observed, and the method of the mean (Lea-Coulson method), which relies on the determination of the mean number of mutants. Both methods assume a Poisson distribution, shown in Figure 2.3, with a mean and variance equal to the product of the probability of a mutation and the number of bacteria. This method uses an estimate of the number of mutational events (not the number of mutants), m (Purple arrow), to determine the mutation rate that have a Luria–Delbrück distribution. The parameter m will be influenced by the amount of growth and the mutation rate (μ). The estimated value of m can be divided by the total number of cells to give the mutation rate. The Lea-Coulson Maximum Likelihood method (MSS-LC) was used when a high number of $rif^R$ mutants (*m*-value) were obtained on selective plates. Mutation rate was calculated using the following formula:

$$\mu = \frac{\text{m - Value X In2}}{\text{Nt - value}}$$

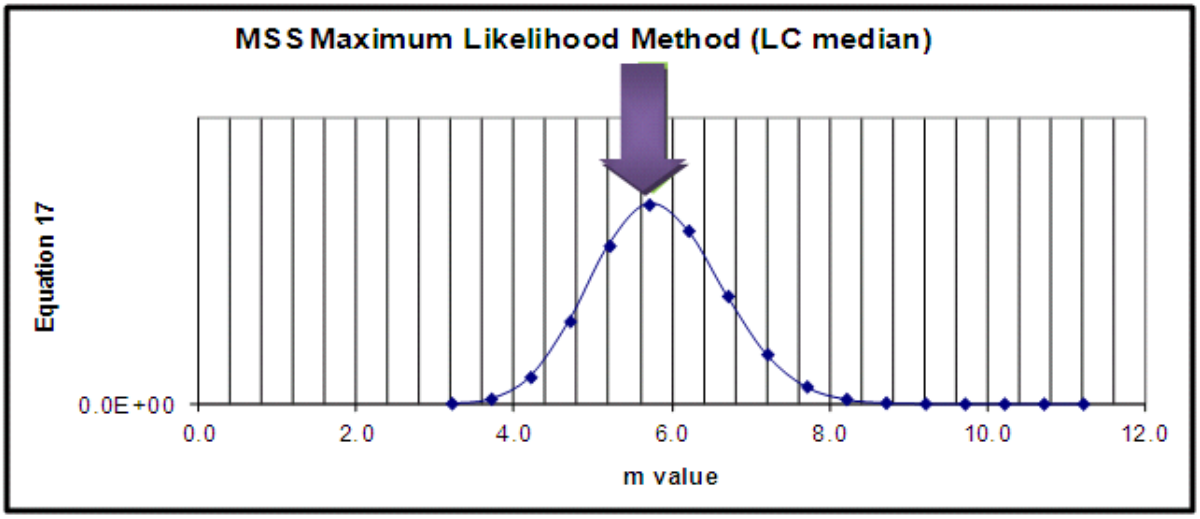



**Figure 2.3: Example of the MSS-LC distribution graph/curve used to estimate the $Rif^R$ mutants per culture (*m*-value).** The purple arrow shows the peak on the graph at which the *m*-value is estimated.

Freezer stocks of the parental, mutant and complemented strains were streaked on 7H10 and incubated at 37 °C until single colonies emerged. Single colonies from the respective strains were grown in 7H9 media to late log phase ($10^8$cells/mL). A dilution series from $10^8$-$10^0$ of the culture was performed and spread on 7H10 media to measure the initial inoculum ($N_0$). Also, 1 mL of the culture was spread onto 7H10 plates containing 200 µg/mL of rif to ensure no pre-existing mutants were present. The cell density was diluted from $10^8$cells/mL to $10^2$ cells/mL and 2 mL aliquots of the $10^2$ dilution were dispensed into 30 culture tubes and incubated at 37 °C with constant shaking (100 rpm) for 7 days until the cultures were in late log phase. The total population size (CFU/mL) was determined by randomly selecting 5 culture tubes out of the 30 and spreading a serial dilution of the cells from $10^8$-$10^0$ on 7H10 plates (*Nt*).The culture from the remaining 25 tubes was plated individually on 7H10 rif (200 µg/mL) to score for spontaneous $rif^R$ mutants (*m*-value) (refer to Figure 2.4). The 7H10 plates were incubated at 37 °C for 3 days for total population size count and 7 days for scoring $rif^R$ colonies.

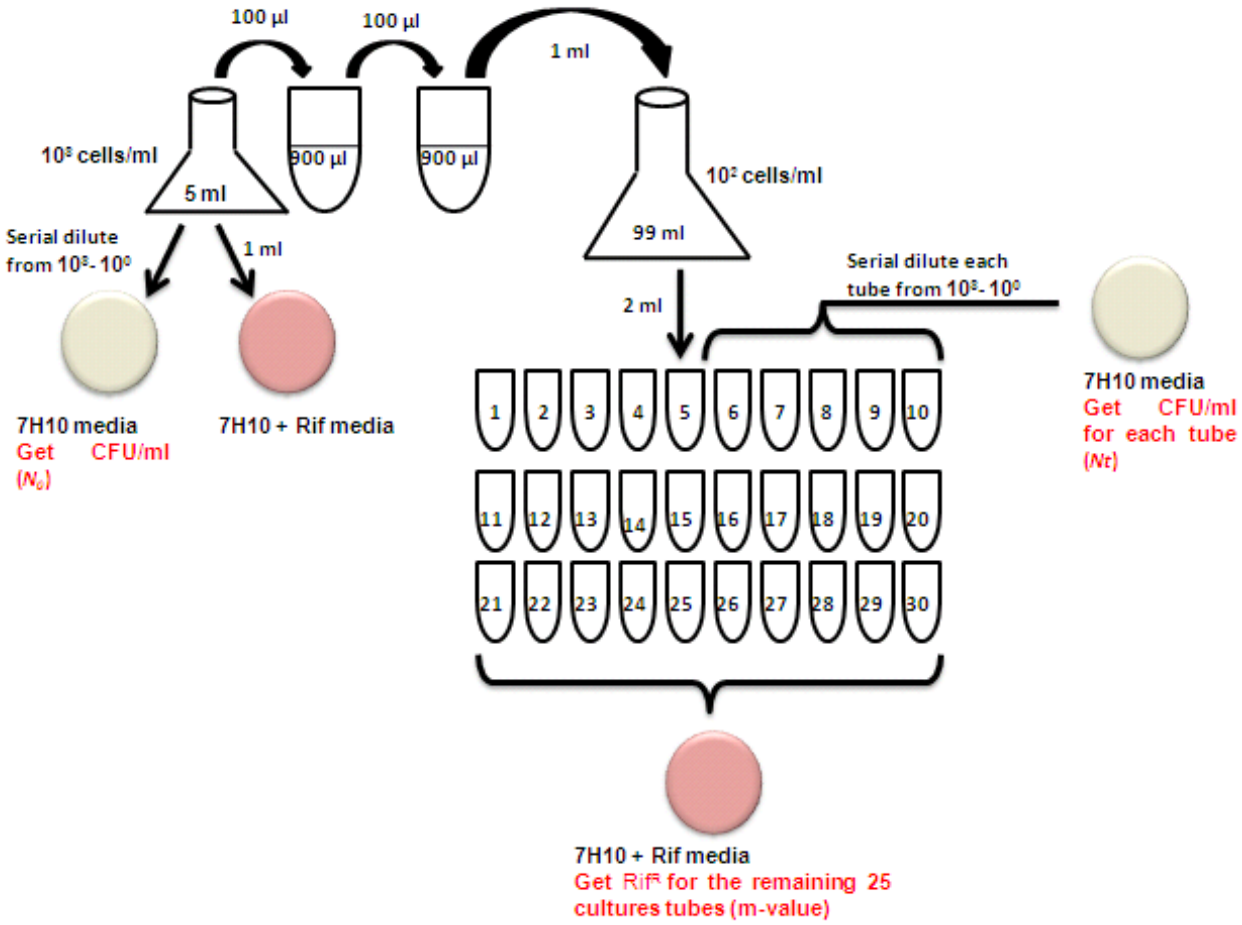



**Figure 2.4: Schematic outline of the fluctuation assay experiment**.

The fluctuation analysis is based on the probability of an organism developing a mutation per cell division in the lifetime of the bacterial cell. The fluctuation assay was set up with a number of parallel cultures (30 tubes), each containing the same number of cells, which were allowed to grow and 5 out of the 30 tubes were selected to determine the total number of cells. The remaining culture was plated on 7H10 media containing rifampicin. The mutation rates were calculated for each strain using the computer software program developed in the laboratory.

# 3. Results

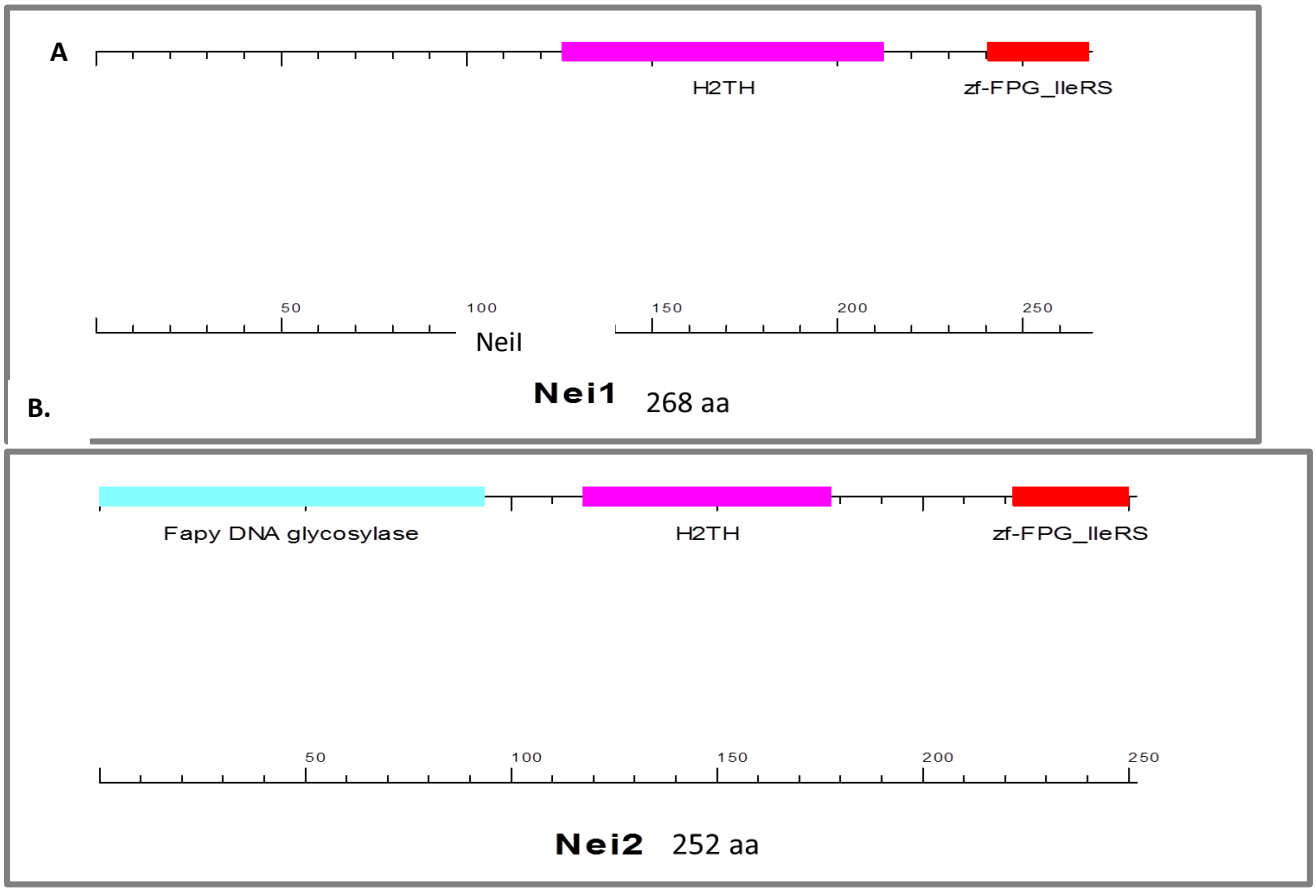
## 3.1 Bioinformatics analysis

The bioinformatics analysis was aimed at highlighting the differences and the similarities between the two Nei homologues, and then comparing them in *M. smegmatis* and *M. tuberculosis*.

The general structure of the Nei DNA glycosylases consists of three distinct domains including the zinc finger and the helix-2-turn-helix (H2TH) domains, which are involved in DNA binding, and a Fapy DNA glycosylase domain that has hydrolytic activity and hydrolyses the N-glycosidic bond in the DNA. The Pfam database (http://pfam.xfam.org/) was used to confirm the structural features of the Nei protein in *M. smegmatis*: As illustrated in Figure 3.1 A, NeiI does show the presence of the H2TH and the zinc finger motifs. However, interestingly the NeiII homologue has an additional Fapy DNA glycosylase, see Figure 3.1 B.

Three dimensional protein structures of the Nei homologues were also modelled to further examine the positioning of the motifs within the two Nei homologues. Nei consists of two domains and possesses several defined structural motifs whose functional importance is evident from the structure. In particular, the enzyme flips the damaged nucleotide from DNA to gain access to C1', a target for nucleophilic attack by the N-terminal proline moiety. This flipping, referred to as "base flipping", is accompanied by the insertion of three amino acid residues (Gln-69, Leu-70 and Tyr-71, referred to here as the QLY loop or the intercalation loop) into the void left in the helix after base extrusion. The amino acids forming this short loop also participate in un-stacking of the base, opposite the lesion, introducing a sharp kink into the DNA, and forming hydrogen bonds with the base opposite the lesion, thereby contributing to the observed opposite-base specificity. NeiI uses the GLY-loop and NeiII uses a KME-loop for intercalation (Figure 3.1 C and D). The overall crystal structure of *M. smegmatis* NeiI and NeiII comprises two domains connected by a hinge polypeptide (Fig. 3.1 C and D) and resembles that of *E. coli* Nei (Zharkov et al., 2002). The N-terminal domain contains strands, forming a sandwich with two helices parallel to its edges shown in red in

Figure. 3.1 C and D. The C-terminal domain includes four -helices, two of which (shown in yellow, orange, light green and blue in Fig. 3.1 C and D) form the helix-two turn-helix (H2TH) motif, and two strands that form a hairpin zinc finger. There is also a $Cys_4$ type zinc finger (indicated on the yellow strand with sticks) (Arg-171) which is involved in tight network hydrogen bonds surrounding the lesion that hold DNA in a highly distorted conformation. Amino acids Arg-171 and Gln-252 are involved in the positioning of the zinc finger (Fromme and Verdine, 2002).



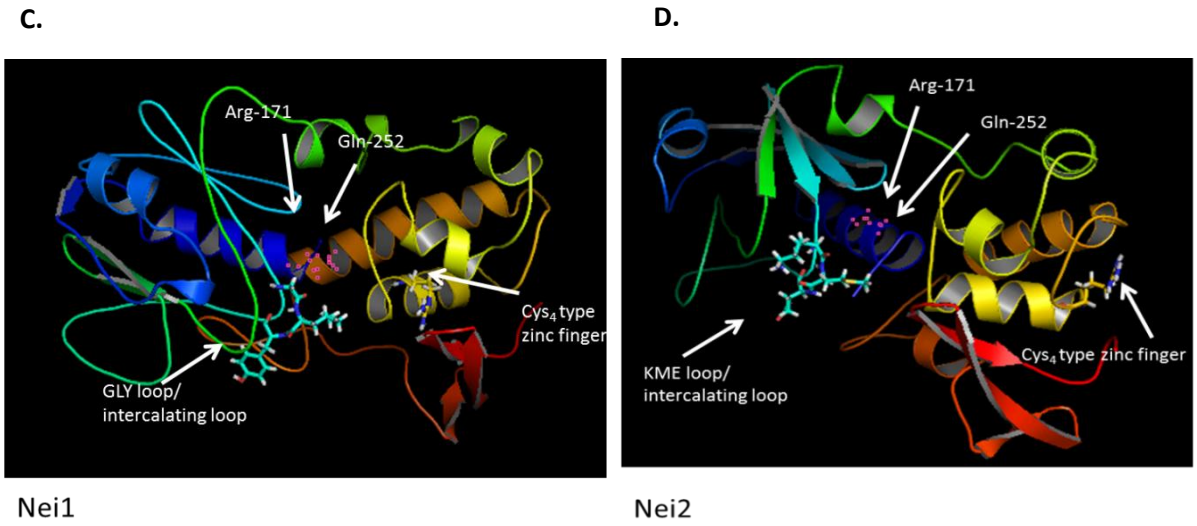



**Figure 3.1: Structural characteristics of the *M. smegmatis* Nei protein.** A and B *pfam* results of the *M. smegmatis* NeiI and NeiII homologues respectively depicting the different motifs. The pink bar indicates the H2TH, red bar shows the zinc finger and the turquoise bar depicts the Fapy DNA glycosylase motif. C and D shows three-dimensional structures of the *M. smegmatis* Nei showing H2TH (orange cluster), Fapy DNA glycosyse (blue cluster) and the zinc finger (red cluster) motifs. The white arrows indicate the important amino acids residues for DNA binding and excision.

The protein database of Clusters of Orthologous Groups (http://www.ebi.ac.uk/Tools/msa/clustalo/) was used to identify the Nei glycosylase of several microorganisms and the amino acid composition of the Nei protein in these organisms was analysed by Clustal Omega alignments (Fig. 3.2). Sequences of the Nei homologues from the following microorganisms were assessed: *Escherichia coli (E. coli), Salmonella typhimurium (S), Xanthomonas campestris (X), Streptomyces (Strep), Corynobacterium (Cory), M. tuberculosis (Tb-neiII and Tb-neiI)* and compared to the *M. smegmatis* counterparts *(Smeg-neiI and Smeg-neiII)* The Nei glycosylases motifs in the various organisms showed significant similarity except for the difference in the intercalating loop between the *M. smegmatis* and *M. tuberculosis* NeiII DNA glycosylases.

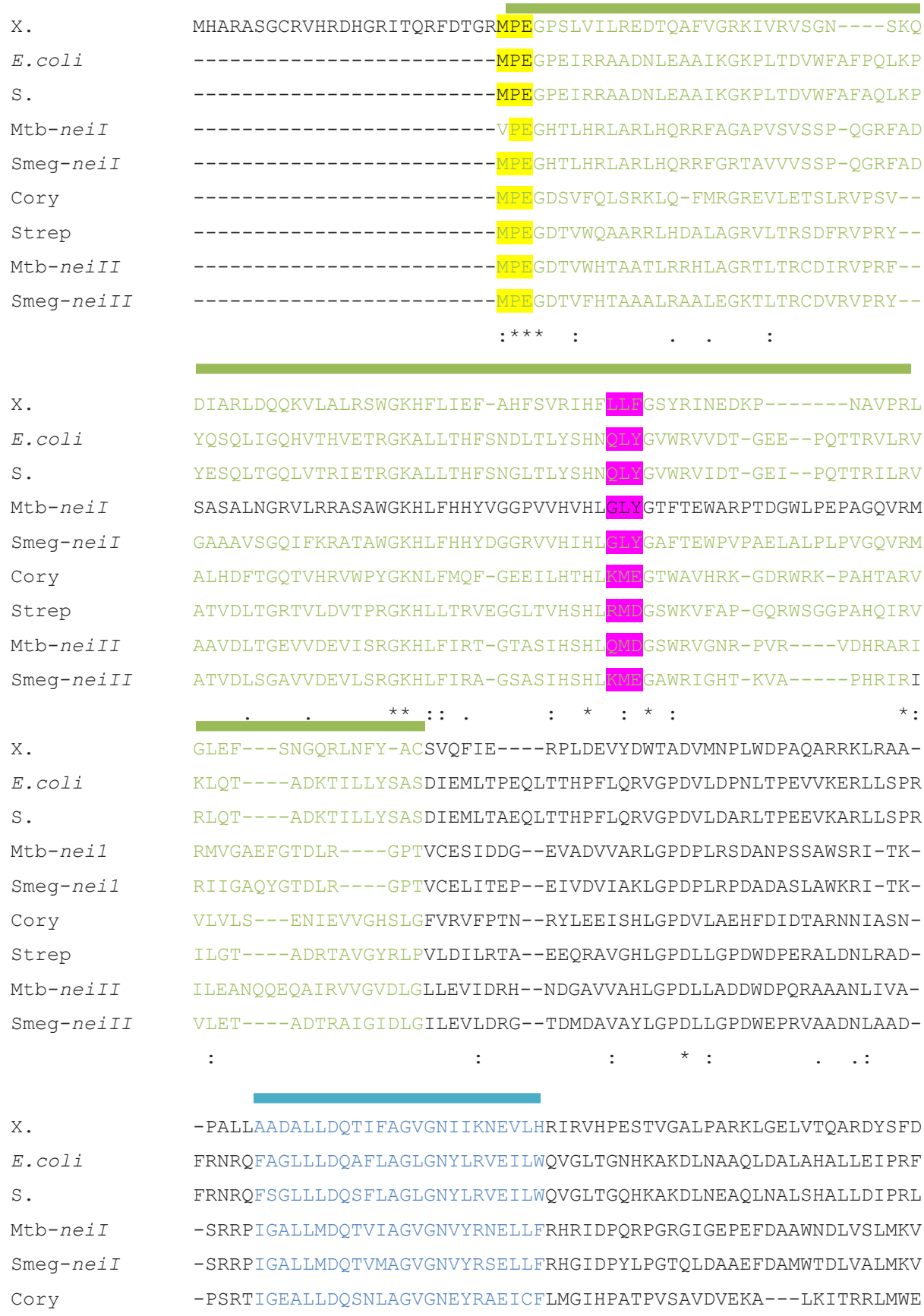

X. MHARASGCRVHRDHGRITQRFDTGRMPEGPSLVILREDTQAFVGRKIVRVSGN----SKQ
E.coli ----------MPEGPEIRRAADNLEAAIKGKPLTDVWFAFPQLKP
S. ----------MPEGPEIRRAADNLEAAIKGKPLTDVWFAFAQLKP
Mtb-neiI ----------VPEGHTLHRLARLHQRRFAGAPVSVSSP-QGRFAD
Smeg-neiI ----------MPEGHTLHRLARLHQRRFGRTAVVVSSP-QGRFAD
Cory ----------MPEGDSVFQLSRKLQ-FMRGREVLETSLRVPSV--
Strep ----------MPEGDTVWQAARRLHDALAGRVLTRSDFRVPRY--
Mtb-neiII ----------MPEGDTVWHTAATLRRHLAGRTLTRCDIRVPRF--
Smeg-neiII ----------MPEGDTVFHTAAALRAALEGKTLTRCDVRVPRY--
X. DIARLDQQKVLALRSWGKHFLIEF-AHFSVRIHFLLFGSYRINEDKP-------NAVPRL
E.coli YQSQLIGQHVTHVETRGKALLTHFSNDLTLYSHNQLYGVWRVVDT-GEE--PQTTRVLRV
S. YESQLTGQLVTRIETRGKALLTHFSNGLTLYSHNQLYGVWRVIDT-GEI--PQTTRILRV
Mtb-neiI SASALNGRVLRRASAWGKHLFHHYVGGPVVHVHLGLYGTFTEWARPTDGWLPEPAGQVRM
Smeg-neiI GAAAVSGQIFKRATAWGKHLFHHYDGGRVVHIHLGLYGAFTEWPVPAELALPLPVGQVRM
Cory ALHDFTGQTVHRVWPYGKNLFMQF-GEEILHTHLKMEGTWAVHRK-GDRWRK-PAHTARV
Strep ATVDLTGRTVLDVTPRGKHLLTRVEGGLTVHSHLRMFGSWKVFAP-GQRWSGGPAHQIRV
Mtb-neiII AAVDLTGEVVDEVISRGKHLFIRT-GTASIHSHLQMDGSWRVGNR-PVR----VDHRARI
Smeg-neiII ATVDLSGAVVDEVLSRGKHLFIRA-GSASIHSHLKMEGAWRIGHT-KVA-----PHRIRI
X. GLEF---SNGQRLNFY-ACSVQFIE----RPLDEVYDWTADVMNPLWDPAQARRKLRAA-
E.coli KLQT----ADKTILLYSASDIEMLTPEQLTTHPFLQRVGPDVLDPNLTPEVVKERLLSPR
S. RLQT----ADKTILLYSASDIEMLTAEQLTTHPFLQRVGPDVLDARLTPEEVKARLLSPR
Mtb-nei1 RMVGAEFGTDLR----GPTVCESIDDG--EVADVVARLGPDPLRSDANPSSAWSRI-TK-
Smeg-nei1 RIIGAQYGTDLR----GPTVCELITEP--EIVDVIAKLGPDPLRPDADASLAWKRI-TK-
Cory VLVLS---ENIEVVGHSLGFVRVFPTN--RYLEEISHLGPDVLAEHFDIDTARNNIASN-
Strep ILGT----ADRTAVGYRLPVLDILRTA--EEQRAVGHLGPDLLGPDWDPERALDNLRAD-
Mtb-neiII ILEANQQEQAIRVVGVDLGLLEVIDRH--NDGAVVAHLGPDLLADDWDPQRAAANLIVA-
Smeg-neiII VLET----ADTRAIGIDLGILEVLDRG--TDMDAVAYLGPDLLGPDWEPRVAADNLAAD-
X. -PALLAADALLDQTIFAGVGNIIKNEVLHRIRVHPESTVGALPARKLGELVTQARDYSFD
E.coli FRNRQFAGLLLDQAFLAGLGNYLRVEILWQVGLTGNHKAKDLNAAQLDALAHALLEIPRF
S. FRNRQFSGLLLDQSFLAGLGNYLRVEILWQVGLTGQHKAKDLNEAQLNALSHALLDIPRL
Mtb-neiI -SRRPIGALLMDQTVIAGVGNVYRNELLFRHRIDPQRPGRGIGEPEFDAAWNDLVSLMKV
Smeg-neiI -SRRPIGALLMDQTVMAGVGNVYRSELLFRHGIDPYLPGTQLDAAEFDAMWTDLVALMKV
Cory -PSRTIGEALLDQSNLAGVGNEYRAEICFLMGIHPATPVSAVDVEKA---LKITRRLMWE




**Fig. 3.2: Alignments of the amino acid sequences of the Nei protein from various microorganisms.** Sequences were downloaded from CMR for *Escherichia coli, Salmonella typhimurium (S.), Xanthomonas campestris (X), Streptomyces (Strep), Corynobacterium (Cory), Mycobacterium tuberculosis (Mtb-neiII and Mtb-nei1) and Mycobacterium smegmatis (Smeg-neiI and Smeg-neiII).* The H2TH motif is outlined by the blue line and elements of the Fapy DNA glycosylase motif is shown by the green line. The zinc-finger domain is showed by red line. Important motifs are shown in coloured highlights.

Analysis of the *M. tuberculosis* Nei homologues as described for *M. smegmatis* also showed conservation of the intercalating loop in both the homologues, Figure 3.3. Eukaryotic homologues of Nei proteins have been identified and exhibit some structural variations. The structure of the human NEIL1 is similar to that of *E. coli* Nei, it lacks the cysteine residues coordinating $Zn^{2+}$ but shows an unusual "zinc-less finger" motif that functions as a zinc finger (Brieba, 2004). According to sequence alignments and pfam analysis this "zinc-less finger" motif was also observed in the *M. tuberculosis* NeiII homologue (see Figure 3.3 A and B). The functional consequences of this difference remain unknown.

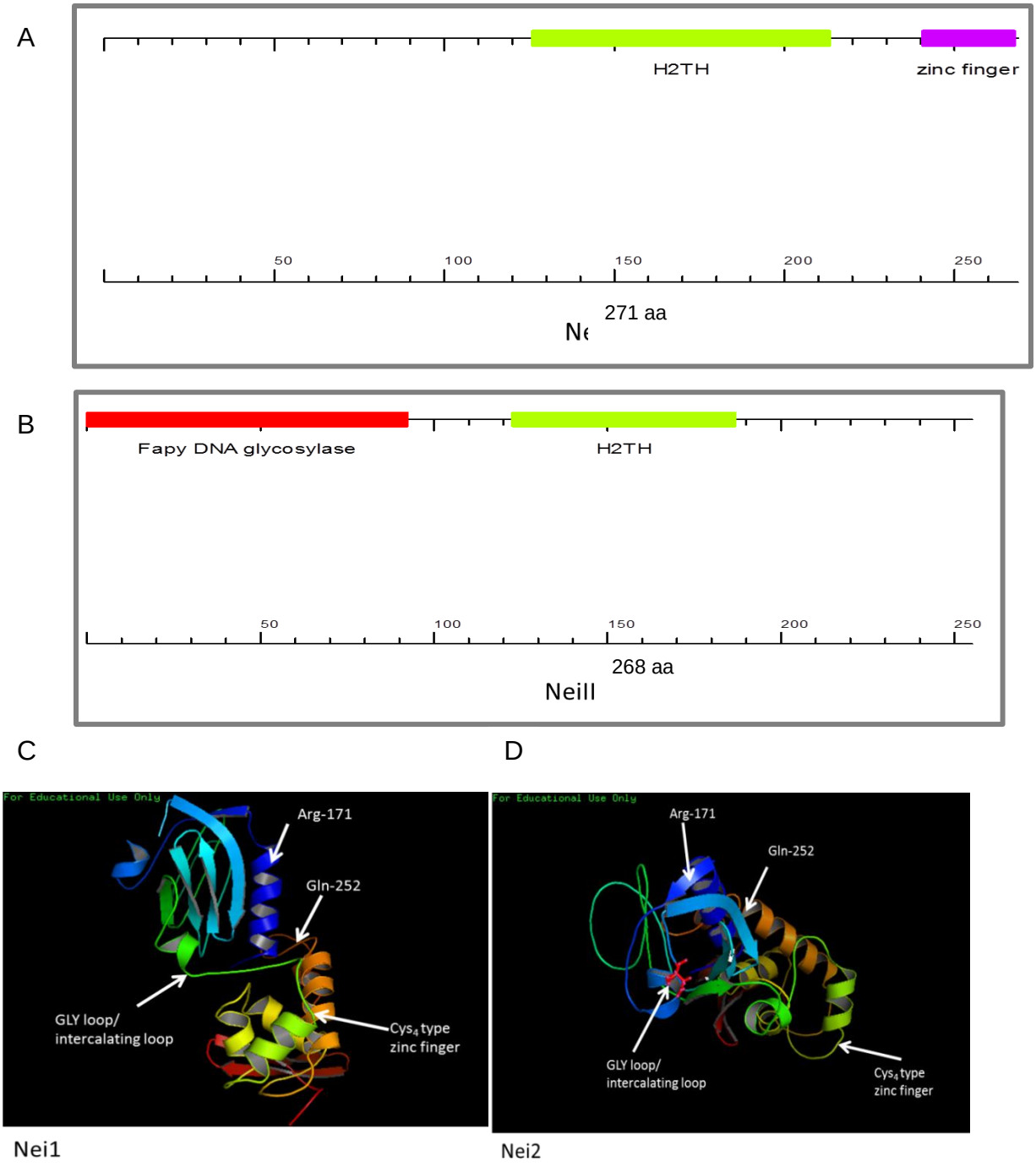


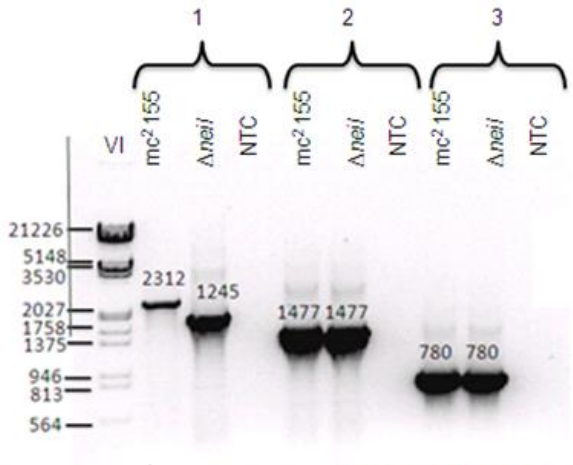

**Figure 3.3: Structural characteristics of the *M. tuberculosis* Nei protein.** A & B *pfam* results of the *M. tuberculosis* NeiI and NeiII homologues respectively depicting the different motifs. The green bar indicates the H2TH, pink bar shows the zinc finger and the red bar depicts the Fapy DNA glycosylase motif. C and D shows three-dimensional structures of the *M. tuberculosis* Nei showing H2TH (orange cluster), Fapy DNA glycosyse (blue cluster) and zinc finger (red cluster) motifs. The white arrows indicate the important amino acids residues for DNA binding and excision.

### 3.2 Confirmation of pre-existing mutants.

Single mutants lacking either *neiI* or *neiII* (Δ*neiI* and Δ*neiII*) that were previously constructed in our laboratory (Goosens, 2009, Moolla et al., 2014) were used to generate double deletion mutants lacking the *nth* DNA glycosylase in each of these backgrounds to generate Δ*neiI* Δ*nth* and Δ*neiII* Δ*nth*. Prior to using these single mutants as recipient strains, PCR confirmation was carried out to ensure that the mutant strains were correct and did indeed lack the respective Nei homologue (see Figure 3.4 and 3.5). The expected band sizes were observed for all the primer sets used (listed in Table 6.1, Appendix D) with genomic DNA from the respective mutants and the parental strains.

#### 3.2.1 Confirmation of the Δ*neiI* mutant strain.

A.



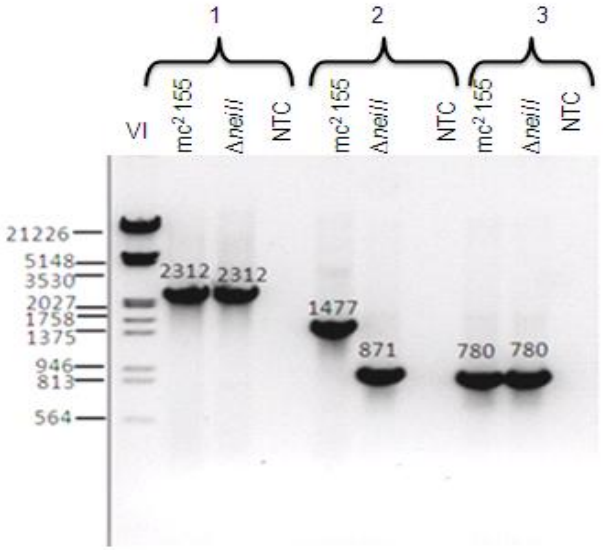

B.

| | | Amplicon size (bp) | |
|---|---|---|---|
| Primers | | Wild type (WT mc² 155) | Mutant |
| 1 | *neiI* | 2312 | 1245 |
| 2 | *neiII* | 1477 | 1477 |
| 3 | *nth* | 780 | 780 |

**Figure 3.4: Confirmation of Δ*neiI* single mutant.** A. Gel electrophoresis image showing the Δ*neiI* mutants, VI depicts Molecular Weight Marker VI, NTC depicts no template control, 1 represents amplification of the DNA with *neiI* specific primers, 2 represent amplification of the DNA with *neiII* specific primers and 3 represent amplification of the DNA with *nth* specific primers. B. Table listing the expected band sizes from each primer set with DNA from the wild type and mutant strains.

### 3.2.2 Confirmation of the Δ*neiII* mutant strain.

A.




B.

| | | Amplicon size (bp) | |
|---|---|---|---|
| Primers | | Wild type (WT mc² 155) | Mutant |
| 1 | *neil* | 2312 | 2312 |
| 2 | *neiII* | 1477 | 871 |
| 3 | *nth* | 780 | 780 |

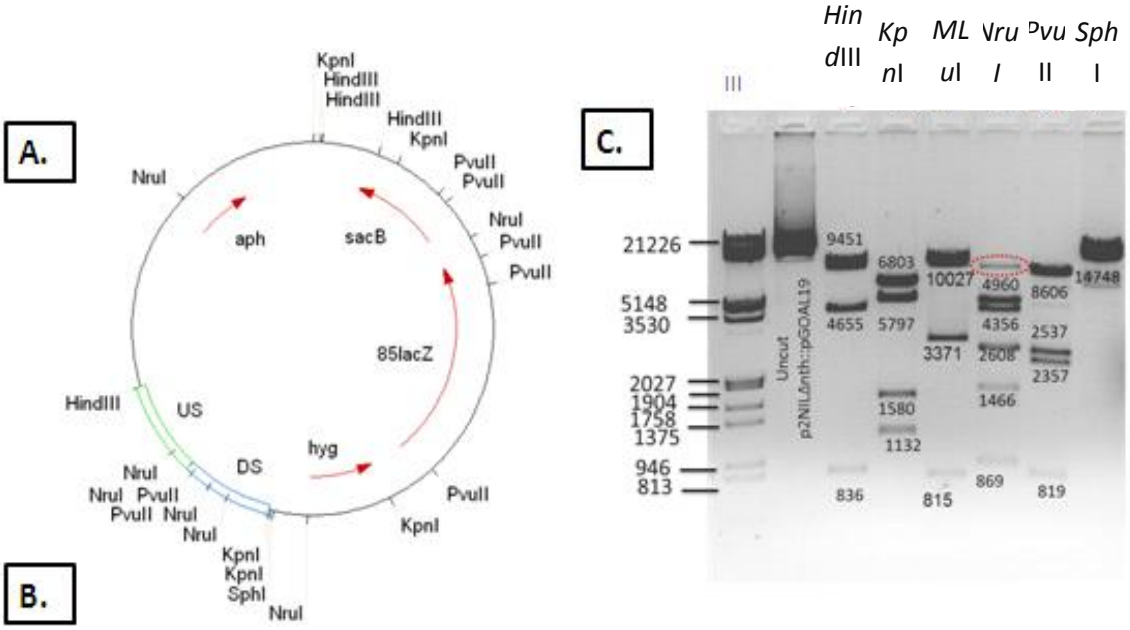
**Figure 3.5: Confirmation of Δ*neiII* single mutant.** A. Gel electrophoresis image showing the Δ*neiII* mutants, VI depicts Molecular Weight Marker VI, NTC depicts no template control, 1 represents amplification of the DNA with *neil* specific primers, 2 represent amplification of the DNA with *neiII* specific primers and 3 represent amplification of the DNA with *nth* specific primers. B. Table listing the expected band sizes from each primer set with DNA from the wild type and mutant strains.

All previously generated strains amplified the expected band sizes (Figure 3.4 and 3.5) confirming that the strains were correct. The strains were taken further for gene knockout using homologous recombination to generate double mutant strains.

### 3.3 Confirmation of previously constructed knockout vector.

In order to ensure that the suicide vector construct carrying the deleted allele for *nth* generated in a previous study (Moolla et al., 2014) was still intact, DNA from the *E. coli* strains carrying the p2NILΔ*nth*::pGOAL19 knockout vector, was isolated and profiled with various restriction enzymes. The restriction profile of the knockout construct yielded

the expected band sizes thus confirming that the knockout vector was correct (see Figure 3.6). The extra band circled in red in Fig 3.6 C is the result of partial digestion of plasmid DNA. This could be due to contaminants or inhibitors that affect the activity of the restriction enzyme.




| Enzymes used | Size of fragments (bp) |
|---|---|
| *Hind*III | 9451, 4655, 836 |
| *Kpn*I | 6803, 5797, 1580, 1132, 30 |
| *MLu*I | 10027, 3371, 815, 490 |
| *Nru*I | 4960, 4356, 2608, 1466, 869, 270 |
| *Pvu*II | 8606, 2537, 2357, 819, 114, 102 |
| *Sph*I | 14784 |

**Figure 3.6: Confirmation of the p2NILΔ*nth*:: pGOAL19 knockout construct using restriction enzyme digestion.** A. Restriction map of the p2NILΔ*nth*::pGOAL19 vector. B. Expected sizes of fragments for a selection of restriction enzymes used for profiling. C. Restriction analysis of p2NILΔ*nth*::pGOAL19 in which the restriction digests all corresponded to the expected sizes.

## 3.4 Generation of double deletion mutants deficient in the Nth and Nei DNA glycosylases

### 3.4.1 Isolation of single cross over single cross over (SCO) and double cross over (DCO).

The knockout vector p2NIL∆*nth*::pGOAL19 was electroporated into the Nei deficient mutant strains (∆*neiI* and ∆*neiII*) that were previously generated using homologous recombination in the laboratory (Goosens, 2009). Blue transformants/colonies represented single cross over (SCO) mutants containing the functional copy and a deleted allele of *nth* together with the selectable and counter-selectable vector marker genes (Figure 2.2). The genotype of the SCO was confirmed by PCR before the SCO mutants were subjected to undergo a second cross over event in the presence of sucrose to generate a double cross over (DCO) mutant. This two-step selection process ensures that only colonies that have lost the *sacB* cassette will be able to survive in the presence of sucrose as the expression of the *sacB* gene produces levansucrase which is toxic to the cell. Hence, colonies that retain the *sacB* cassette will not be viable in the presence of sucrose. Loss of the *sacB* gene results in the simultaneous loss of the *lac*Z gene therefore, the resulting colonies which are possible DCOs (∆*nei* ∆*nth* double deletion mutants) will be white and sucrose resistant. The white colonies could also be wild type revertants if the second crossover occurs on the same side as the SCO (depicted in Figure 2.2). This process yielded several white colonies which were assessed further by PCR.

### 3.4.2 Confirmation of double deletion mutants by PCR.

Primer set NthF1 and NthR1 listed in Table 6.1 (Appendix D) were used to discriminate between clones containing an intact copy of the *nth* allele and mutants that had a disrupted copy of the *nth* allele. Chromosomal DNA from 10 possible DCO isolates from the two mutant strains (∆*neiI* ∆*nth* and ∆*neiII* ∆*nth*) was amplified and the expected size for the intact *nth* allele (780 bp) was obtained, (see Figure 3.7 A and B) in the parental strain (WT) a smaller 270 bp fragment was observed when the functional copy was replaced with a disrupted copy of the *nth* gene in the Nei deficient mutant strains (∆*neiI* and ∆*neiII*).

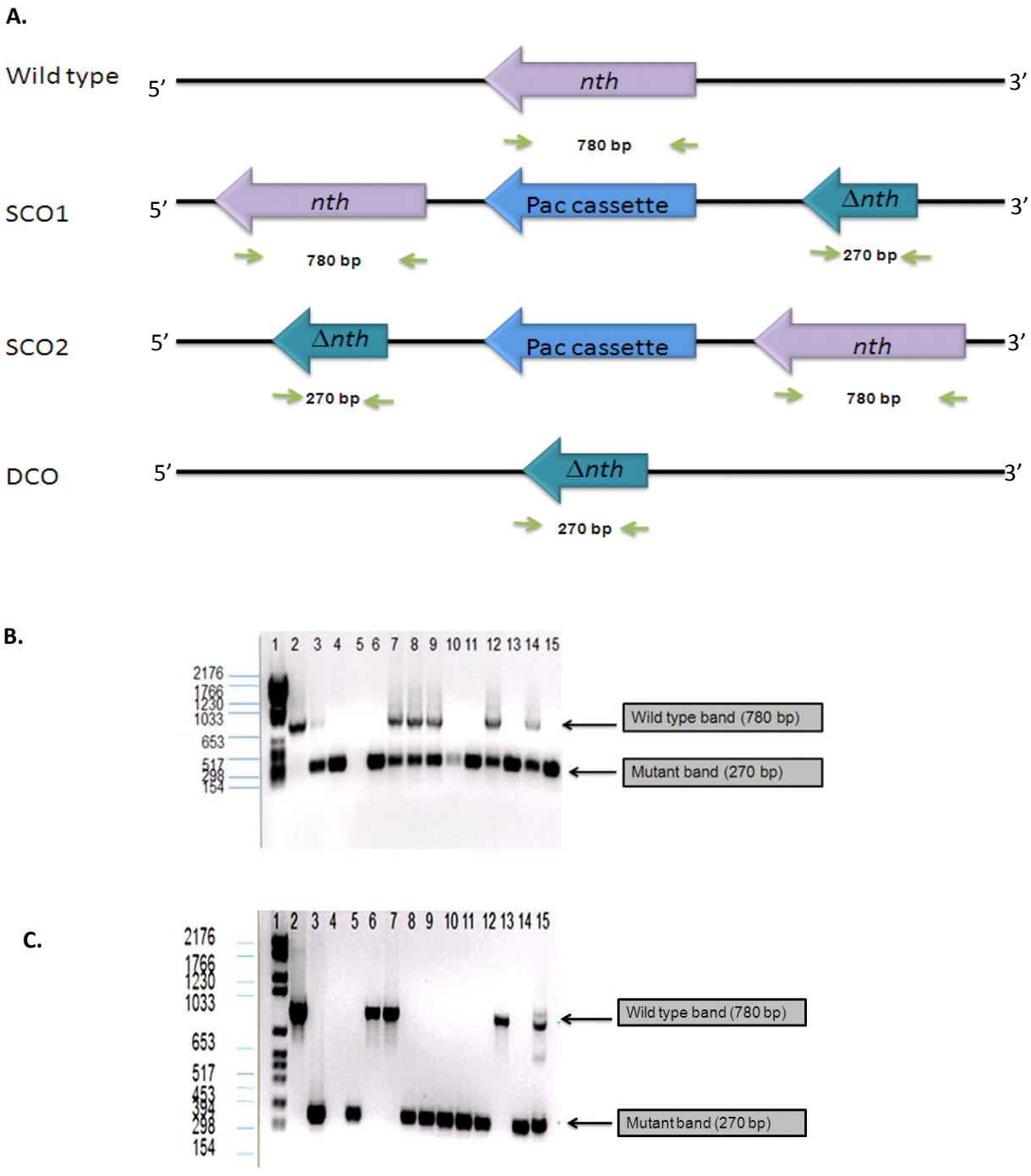



**Figure 3.7: PCR screen for identification of DCOs using *nth* primers.** A. Schematic diagram showing screening primers (green) with expected band sizes. B. and C. Agarose gel showing the PCR screen for deletion of *nth* in the Δ*neil* and Δ*neilI* deletion mutants respectively. [Lane 1] Marker VI, [Lane 2] $mc^2155$, [Lane 3] SCO strain, [Lane 4] p2NILΔ*nth*:pGOAL19, [Lane 5 to 15] Possible DCO Δ*neil* Δ*nth* (B) or Δ*neilI* Δ*nth* clones (C). Expected sizes, for wild type ($mc^2$ 155) is 780 bp and for the Δ*nth is* 270 bp.

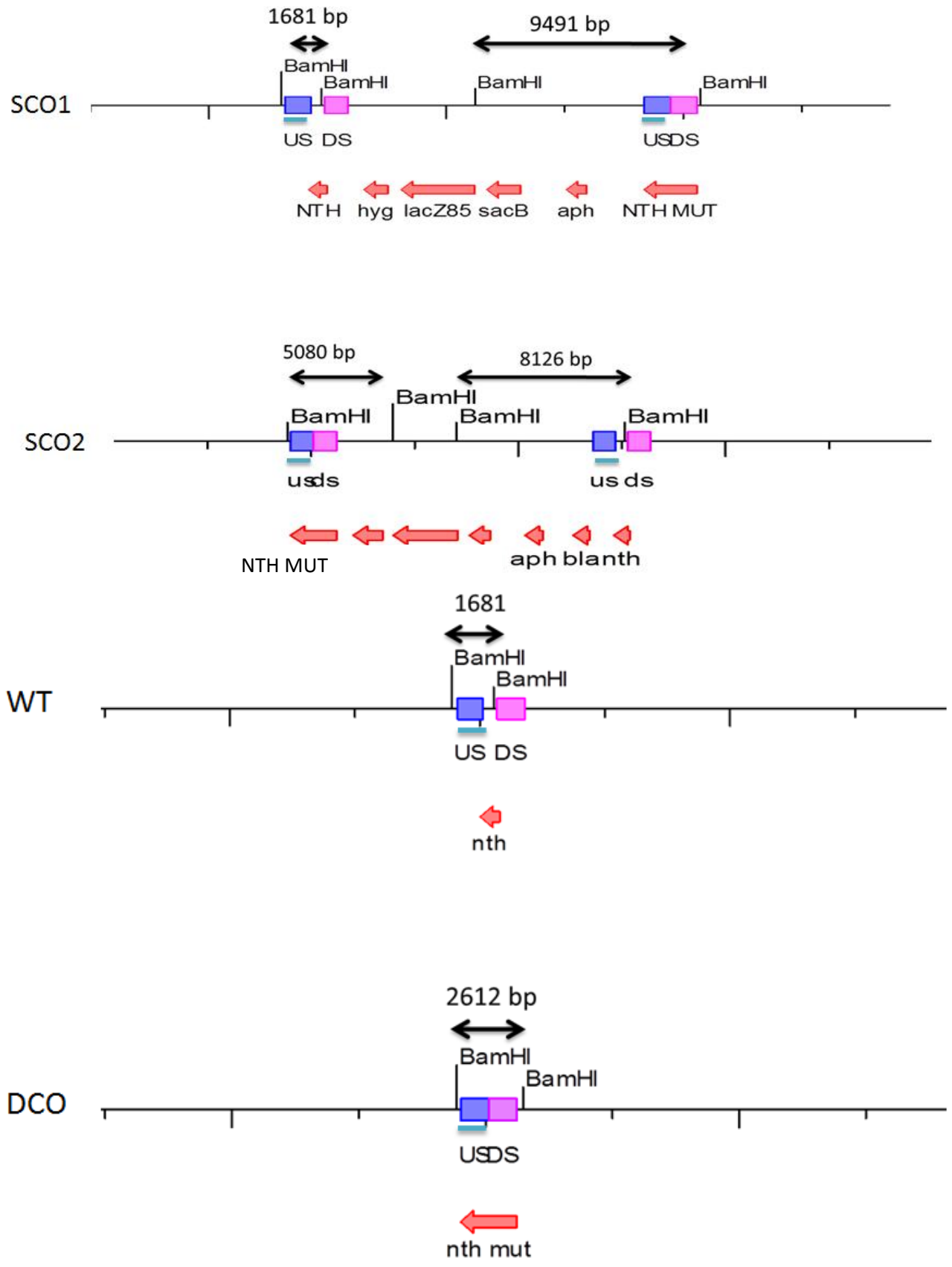
From the 10 colonies that were picked for the Δ*neiI* background, PCR confirmed that 5 clones were the Δ*nth ΔneiI* mutants (Fig. 3.7 B). SCOs showed the expected two amplicons with a preference for the deleted allele as it is smaller. From the 10 colonies that were picked for the Δ*neiII* background, PCR confirmed that 7 clones were the Δ*nth* Δ*neiII* mutants (Fig. 3.7 C). Five PCR verified DCOs from each strain were analysed further by Southern blot analyses to ensure that the mutants were genotypically intact.

### 3.4.3 Southern blot confirmation of double mutants.

Southern blot analysis was used to ascertain the integrity of the upstream (US) and downstream (DS) sequences of the deletion region around the *nth* gene. Genomic DNA from the DCOs that were confirmed by PCR was digested with *Bam*HI and the agarose gel-separated DNA was transferred onto a nitrocellulose membrane (outlined in section 2.7) and probed with a chemiluminescently labelled US probe to identify the respective predicted sizes. The fragment sizes identified by the US probe were in accordance with the expected fragments as predicted by the sequence map (Figure 3.8 A). The WT yielded a 1681 bp fragment and all the DCOs (*nth* mutants) produced a bigger fragment of 2612 bp, thus confirming the replacement of the intact *nth* allele with a non-functional allele (Figure 3.8 B and C). For the SCOs, the expected fragment sizes were obtained only for the ∆*neiI* ∆*nth* (SCO1 9491 bp and 1861 bp) while for the ∆*neiII* ∆*nth SCO* the bigger expected fragment size (9491 bp) was not observed (Figure 3.6 C) This could be due to DNA degradation leading to loss of the larger molecular weight bands.

Since the SCO mutants were correct by PCR analysis (data not shown) and were required only for the generation of the DCO in the two step selection process during homologous recombination, the Southern blot analysis of the ∆*neiII* ∆*nth* SCO was not pursued further as the resulting DCOs were all correct with the US and DS regions in the correct genotypic context.

**A.**


1681 bp
9491 bp
SCO1
BamHI
BamHI
BamHI
BamHI
US DS
USDS
NTH
hyg
lacZ85
sacB
aph
NTH MUT
5080 bp
8126 bp
SCO2
BamHI
BamHI
BamHI
BamHI
usds
us ds
NTH MUT
aph
blanth
1681
WT
BamHI
BamHI
US DS
nth
2612 bp
DCO
BamHI
BamHI
USDS
nth mut

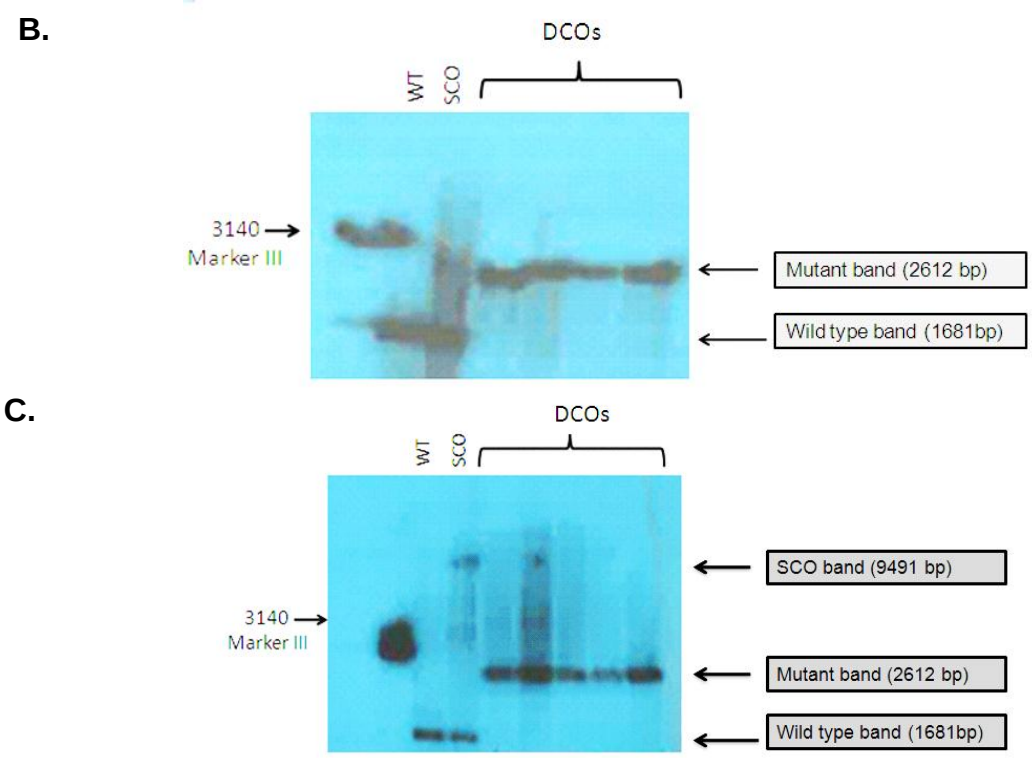
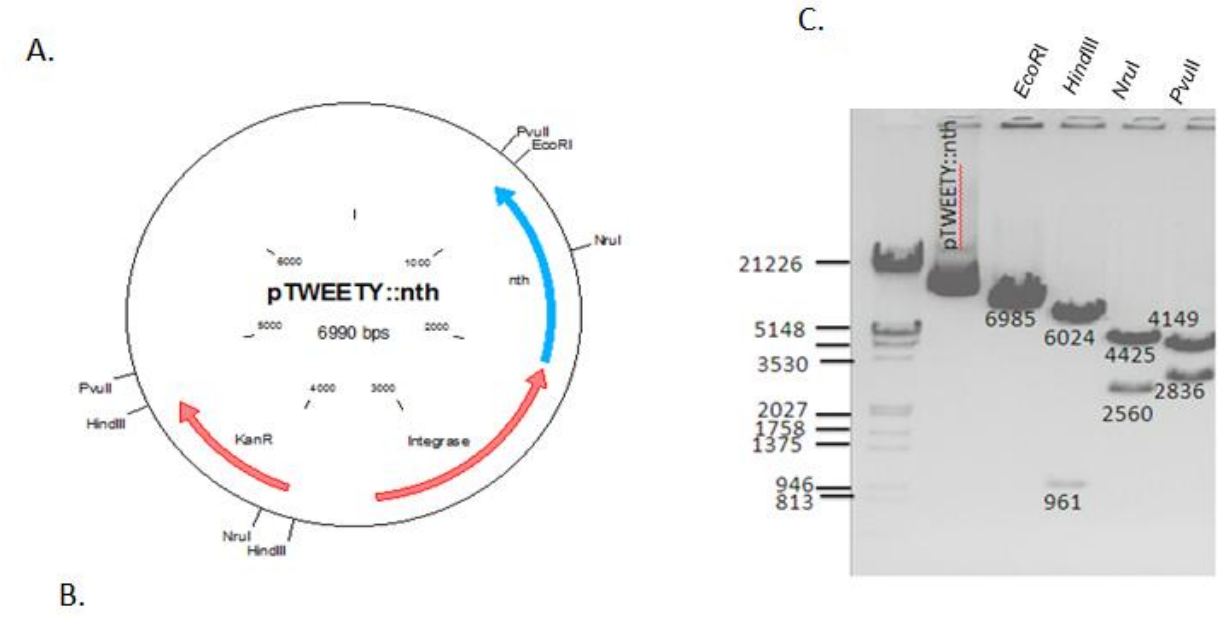



**Figure 3.8: Southern blot upstream analysis of Δ*nth* mutants.** A. Southern blot strategy/Map to distinguish the wild type, the SCO and DCO. The chromosomal DNA of the parental, SCO and DCO was digested with *Bam*HI (probed with the upstream fragment indicated by the blue line – US probe) for confirmation of the upstream region. B. Radiograph showing the Southern blot analysis of the Δ*neiI* Δ*nth* mutants and C. Radiograph showing the Southern blot analysis of the Δ*neiII* Δ*nth* mutants.

### 3.4.4 Genetic complementation of double mutant strains.

The double mutants generated in this study were complemented with an intact copy of the *nth* gene to confirm that any observed phenotype is due to the deleted *nth* allele and not due to downstream polar effects that may have arisen during mutant construction. Complementation of Δ*neiI* Δ*nth* and Δ*neiII* Δ*nth* double deletion mutants with the *nth* gene was accomplished by utilizing a previously constructed integrating vector containing the intact gene (Moolla et al., 2014). The integrating vector, pTWEETY::*nth* (Table 1.2) has a functional copy of the *nth* gene with approximately 350 bp additional upstream sequence that contains the promoter sequence. To reconfirm genetic integrity

of the integrating vector, restriction enzyme mapping was carried out and the expected fragment sizes were observed, thus confirming the integrity of the complementation vector (Figure 3.9). The complementation vector, pTWEETY::*nth* was electroporated into the double mutant strains, Δ*neiI* Δ*nth* and Δ*neiII* Δ*nth* to genetically revert the strain to their Δ*neiI and* Δ*neiII* parent single deletion mutant backgrounds. The resulting complemented strains were annotated as Δ*neiI* Δ*nth*::*nth* and Δ*neiII* Δ*nth*::*nth* (Table 1.1).




| Enzymes used | Expected Fragment size in bp |
|---|---|
| *EcoR*I | 6985 |
| *Hind*III | 6024, 961 |
| *Nru*I | 4425, 2560 |
| *Pvu*II | 4149, 2836 |

**Figure 3.9: Confirmation of the pTWEETY::*nth* complementation construct using restriction enzyme digestion.** A. Restriction map of the pTWEETY::*nth* integrating vector. B. Expected sizes of fragments for a selection of restriction enzymes used for profiling. C. Agarose gel showing the restriction analysis of pTWEETY::*nth* in which the restriction digests all corresponded to the expected sizes.

The pTWEETY::*nth* vector was electroporated into the Δ*neiI* Δ*nth* and Δ*neiII* Δ*nth* strains and the integrants were selected on kanamycin. The integration of the vector in the complemented strains was confirmed by PCR using the same conditions as described in section 2.3.1 in which the wild type and the mutants *nth* alleles were distinguished based on size difference in the complemented strains (Figure 3.10). The re-appearance of the wild type band in the genetically completed strains provided proof that the complementation vector integrated into the chromosome.

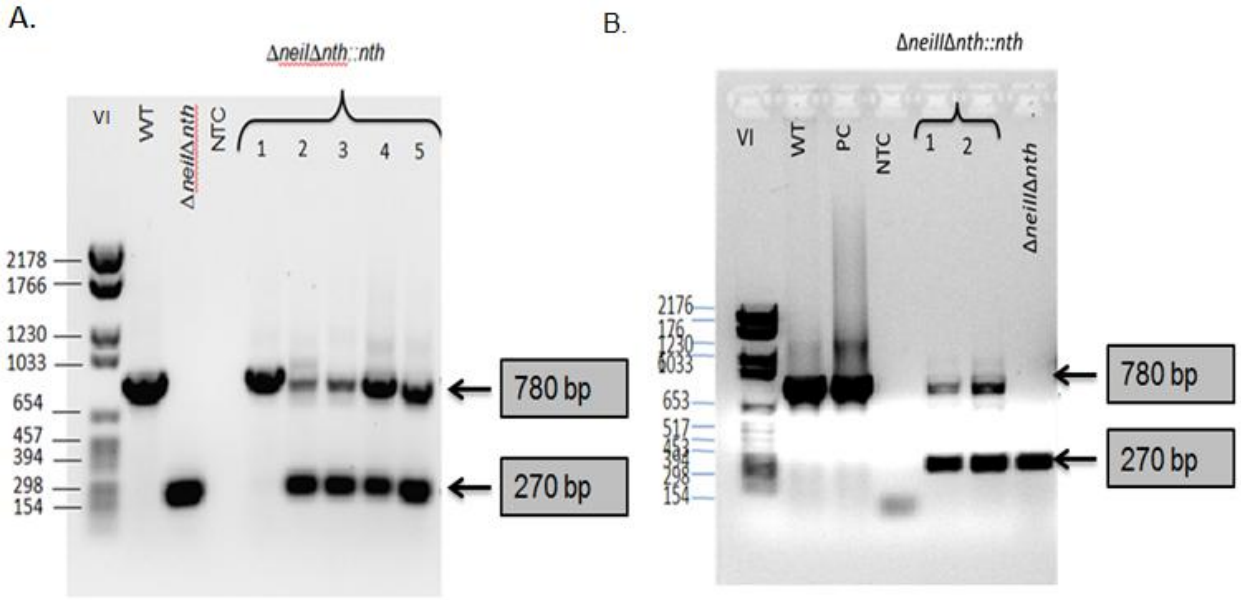



**Figure 3.10: PCR confirmation of complemented strains.** A. and B. represent complementation strains shown as the intact copy of the *nth* gene (780 bp) and the disrupted copy of the gene (270 bp). VI -marker VI, WT-wild type, NTC- no template control and 1 to 5 represent the complemented strains.

### 3.5 Phenotypic characterization.

#### 3.5.1 Growth kinetics.

Growth kinetics of mc$^2$155, the single mutants (Δ*nth*, Δ*neiI* and Δ*neiII*), the double mutants (Δ*neiI* Δ*nth* and Δ*neiII* Δ*nth*) and the complemented strains (Δ*neiI* Δ*nth*::*nth* and ΔneiII Δnth::*nth*) was assessed under normal culture conditions to evaluate the effect of the combinatorial loss of *nth* and *nei* DNA glycosylases. Log phase cultures of the various strains were inoculated in 7H9 media and bacterial growth monitored by measuring optical density ($OD_{600}$nm) at 3 hour intervals over a period of 36 hours. Both the double deletion mutant strains showed comparable growth to the parental strain and the respective single deletion mutants suggesting that the combinatorial loss of Nth and Nei do not affect growth of *M. smegmatis* under normal culture conditions (Figure 3.11). Furthermore, both the complemented strains also displayed the same growth kinetics as the respective single gene deletion mutants indicating that integration of the functional *nth* gene at the *attB* site does not affect the growth kinetics of these strains.

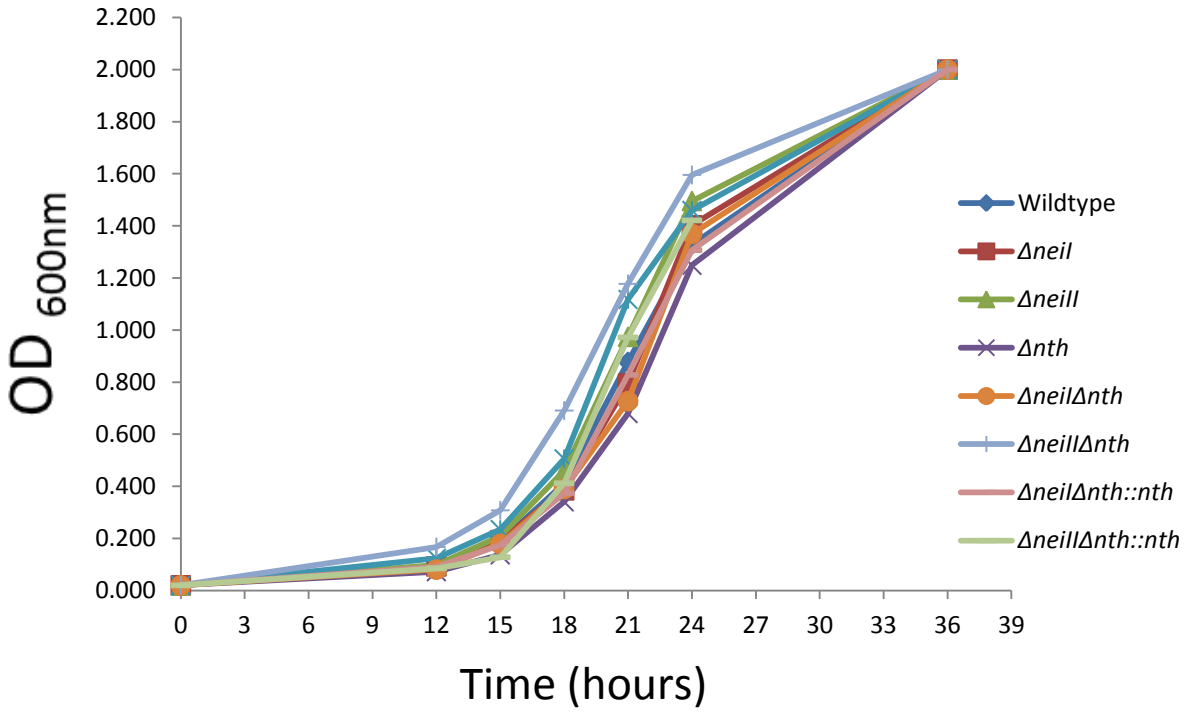
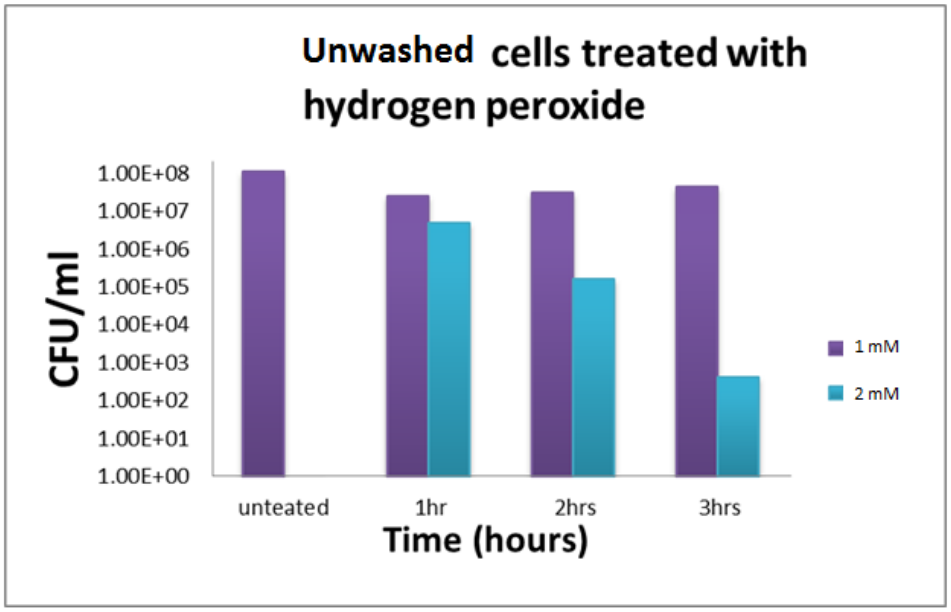



**Figure 3.11: Growth kinetics of the parental and mutant strains under normal culture conditions.** Growth of the wild type (WT), Nei and Nth deficient mutants and their respective complemented strains was monitored every 3 hours over a 36 hour period. The data shows the mean of two independent experiments.

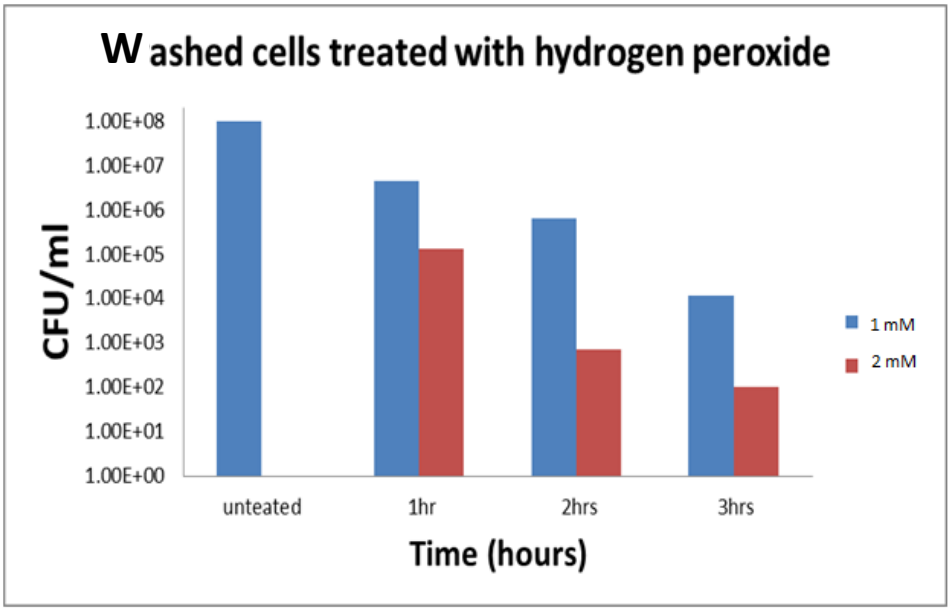
### 3.5.2 Sensitivity to hydrogen peroxide.

Early log phase cultures ($OD_{600nm}$ of 0.3-0.4) of $mc^2155$, Δ*nth*, Δ*neiI*, Δ*neiII*, Δ*neiI* Δ*nth*, Δ*neiII* Δ*nth,* Δ*neiI* Δ*nth*::*nth* and Δnei*II* Δnth::*nth* were exposed to 2.0 mM hydrogen peroxide as described in section 2.11.3. The treated cells were monitored for survival by assessing cell viable counts (log CFU/mL) over a 3 hour period. This experiment was carried out 3 times however, the results were not reproducible for all 3 replicates; the different strains gave varying results in each experiment (data not shown). This could be due to the effect of the "catalase" enzyme which is found in mycobacterial cells, which is secreted into the culture medium. When the enzyme is exposed to oxygen it catalyzes the decomposition of hydrogen peroxide to water and oxygen. It is a very important enzyme in protecting the cell from oxidative damage by ROS. The varying results obtained in this experiment could have been due to the neutralization of hydrogen peroxide by the catalase within the mycobacterial cells or within the culture medium. To eliminate this possibility, wild type cells from early log phase ($OD_{600nm}$ of 0.3-0.4) cultures were pelleted and washed twice with 7H9 broth to remove cellular biomolecules in the media that might react with the hydrogen peroxide before treating with 1 mM and 2 mM concentrations of hydrogen peroxide. The treated cells were monitored for survival by assessing cell viable counts (log CFU/mL) hourly over a period of 3 hours (Figure 3.12). The washed cells showed notable sensitivity to both 1 and 2 mM hydrogen peroxide, with a 2 and 3 log reduction after 3 hours, respectively of viable cells after 3 hours (Figure 3.12 B). As expected, for the unwashed cells 1 mM of hydrogen peroxide was insufficient to reduce the survival of the wild type strain (Figure 3.2 A) but at 2 mM hydrogen peroxide there was a 5- log reduction in cfu/mL after 3 hours of exposure. Consolidation of the data for the washed and unwashed cells as shown in Figure C, clearly show that washing the cells before treatment with hydrogen peroxide does have a greater impact on survival of cells at both the concentrations of hydrogen peroxide tested, suggesting that the washing of the cells does remove cellular biomolecules that might react with the hydrogen peroxide.

A.


Unwashed cells treated with hydrogen peroxide
1.00E+08
1.00E+07
1.00E+06
1.00E+05
1.00E+04
1.00E+03
1.00E+02
1.00E+01
1.00E+00
CFU/ml
unteated
1hr
2hrs
3hrs
Time (hours)
1 mM
2 mM

B.


Washed cells treated with hydrogen peroxide
1.00E+08
1.00E+07
1.00E+06
1.00E+05
1.00E+04
1.00E+03
1.00E+02
1.00E+01
1.00E+00
CFU/ml
unteated
1hr
2hrs
3hrs
Time (hours)
1 mM
2 mM

C.

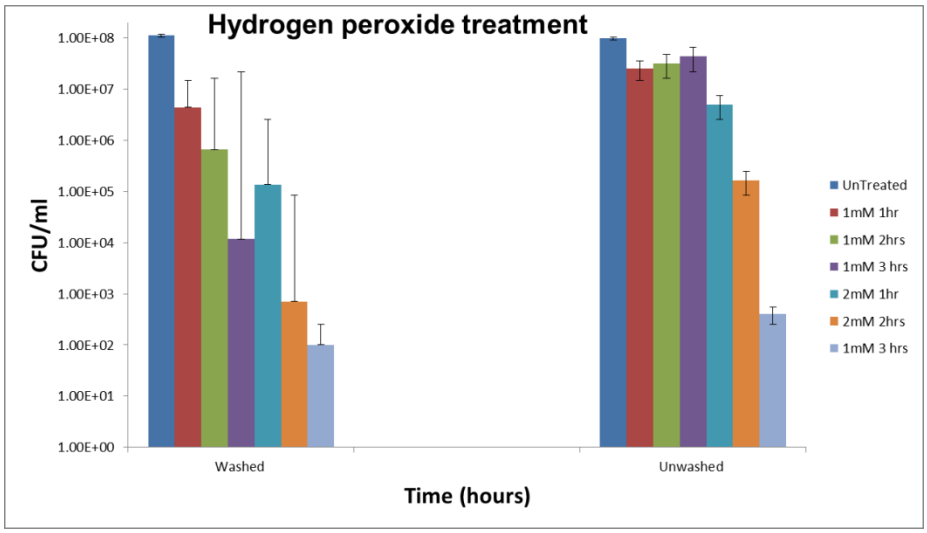



**Figure 3.12: Assessment of washed and unwashed wild type *M. smegmatis* strain in the presence of hydrogen peroxide.** A. Graph showing the unwashed cells when treated with 1 mM (Purple bars) and 2 mM (light blue bars) hydrogen peroxide. B. Graph showing the washed cells when treated with 1 mM (blue bars) and 2 mM (red bars) hydrogen peroxide. C. Graph showing the combined data from graphs A and B. The treated strains were monitored hourly for CFU/mL over a period of 3 hours; the experiment was repeated 3 times. Error bars represent standard deviations between experiments.

Hence, based on this observation, the parental, mutant and complemented strains (mc$^2$155, Δ*nth,* Δ*neiI*, Δ*neiII*, Δ*neiI* Δ*nth*, Δ*neiII* Δ*nth,* Δ*neiI* Δ*nth*::*nth* and Δnei*II* Δnth::*nth)* were grown to early log phase ($OD_{600nm}$ 0.3-0.4), washed twice with 7H9 broth and treated with 2 mM hydrogen peroxide. The data from these experiments are represented in Figure 3.11

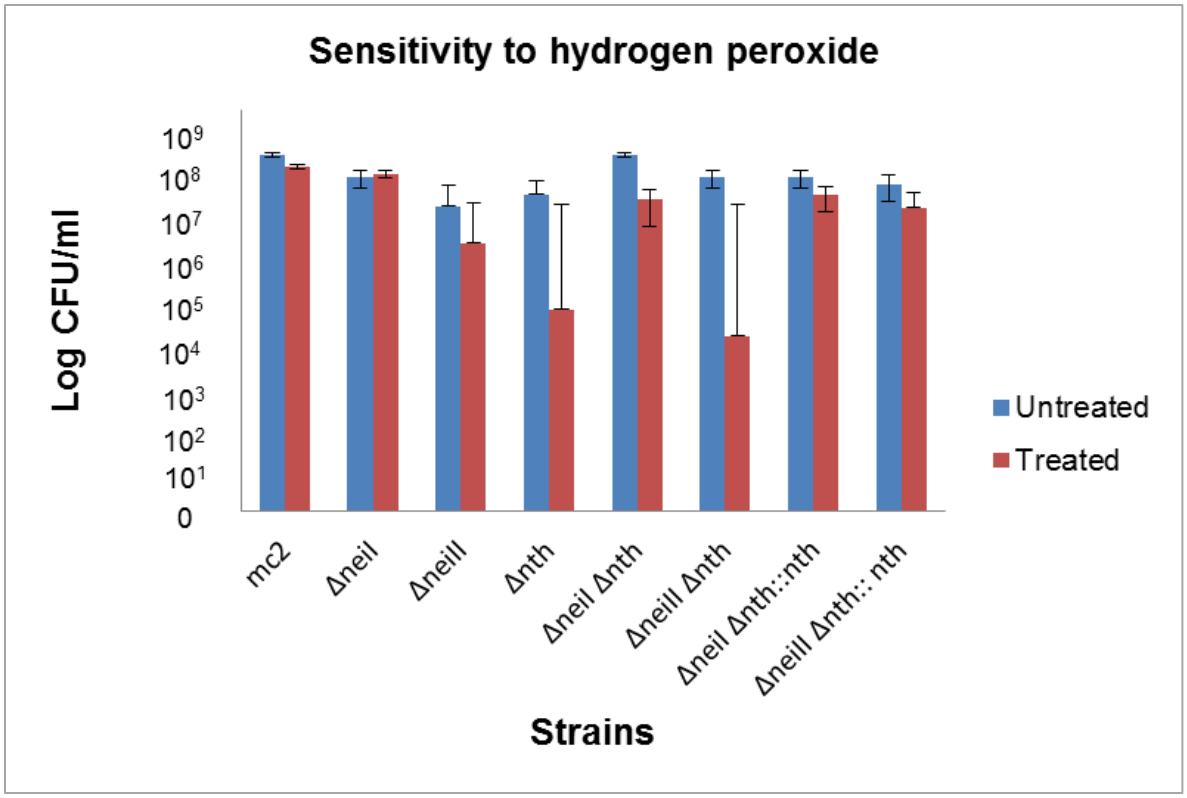



**Figure 3.13: Assessment of wild type *M. smegmatis* single, double deletion mutants and their respective complemented strain with and without 2 mM $H_2O_2$.** The blue bars shows cells before treatment and red bars indicate the viable cells after treatment with 2 mM $H_2O_2$, growth was monitored as cfu/mL after 3 hour. Error bars represent standard deviations from three experiments.

The two single Nei deletion mutants clearly show differential survival under oxidative stress conditions with the NeiI-defective mutant having similar tolerance to the wild type strain. Deletion of the *neiII* allele results in a 1 log reduction in CFU/mL compared to the untreated cells. The *nth* single deletion mutant showed a 3 log reduction in viability in the presence of hydrogen peroxide, which was marginally reduced when both the *nth* and *neiII* genes were inactivated suggesting that in this double mutant, the killing effect was due to the loss of the *nth* DNA glycosylase only. Interestingly deletion of *nth* together with the *neiI* DNA glycosylase only displayed a one log reduction in viable counts suggesting that the remaining intact *neiII* homologue is able to sufficiently protect the DNA under oxidative stress conditions. The observed phenotypes were fully

complemented to the levels observed for the single deletion mutants with the intact *nth* gene.

### 3.5.3 Assessment of mutation rates by the fluctuation assay.

The mutation rates were calculated using the Luria-Delbrück fluctuation analysis as described in section 2.11.4. The mutation rates were calculated for each strain using the computer software program developed in the laboratory by Machowski *et al.*, 2007. The Lea-Coulson method was used to create the Luria-Delbrück distribution to estimate the *m* value required for the mutation rate calculation. The mutation rates (μ) are dependent on the distribution of mutations (*m*) in the entire population (*Nt*) (Rosche and Foster, 2000).

To investigate the role of the two Nei homologues during mutagenesis the spontaneous mutation rates to rifampicin of the various *nth* and *nei* deficient mutants were assessed and compared to the parental strain as indicated in Table 3.1. The single Δ*nth,* Δ*neiI* and Δ*neiII* deletion mutants displayed no increase in mutation rates (Table 3.1 – outlined with the blue line) which clearly suggested that the Nei homologs are able to compensate for Nth and vice versa. In contrast, deletion of *nth* together with either *neiI or neiII* homologues results in a 2 and 6-8 fold increase in mutation rate respectively (Table 3.1– outlined in the grey box). Both the complemented strains restored the functionality of the *nth* in the double mutants. Loss of the NeiII homologue in combination with Nth again showed an increased mutation rate compared to when the NeiI homologue was deleted in combination with the Nth DNA glycosylase. Taken together these data clearly suggested that NeiII has a greater role in maintaining genome integrity under oxidative stress conditions compared to the NeiI homologue. These findings indicate a hierarchal role for the various DNA glycosylases in the BER pathway for genome maintenance during stressful conditions.

**Table 3.1:** Spontaneous rifampicin resistance mutation rates of the various *nth* knockout double mutant strains compared to *M. smegmatis* and single mutant parental strains as determined by fluctuation analysis. Various *nth* deleted mutants are outlined in the grey box for comparison. The experiment was repeated three times.

| Mutation rate | | | | Fold change | | |
|---|---|---|---|---|---|---|
| **Strains** | **Experiment 1** | **Experiment 2** | **Experiment 3** | **Experiment 1** | **Experiment 2** | **Experiment 3** |
| $mc^2$ 155 | 2.06E-09 | 2.09E-09 | 1.99E-09 | 1.0 | 1.0 | 1.0 |
| Δ*neiI* | 3.94E-09 | 3.11E-09 | 3.75E-09 | 1.9 | 1.5 | 1.9 |
| Δ*neiII* | 2.214E-09 | 3.47E-09 | 2.87E-09 | 1.1 | 1.7 | 1.4 |
| Δ*nth* | 2.66E-09 | 2.41E-09 | 3.00E-09 | 1.3 | 1.2 | 1.5 |
| Δ*neiI* Δ*nth* | 5.64E-09 | 4.52E-09 | 4.78E-09 | 2.7 | 2.2 | 2.4 |
| Δ*neiII* Δ*nth* | 1.42E-08 | 1.74E-08 | 1.21E-09 | 6.9 | 8.3 | 6.1 |
| Δ*neiI* Δ*nth::nth* | 7.02E-09 | 6.33E-09 | 5.67E-09 | 3.4 | 3.0 | 2.8 |
| Δ*neiII* Δ*nth::nth* | 8.43E-09 | 7.87E-09 | 7.32E-09 | 4.1 | 3.8 | 3.7 |

# 4. Discussion and Conclusion

The ability to repair damaged DNA forms a crucial component of the cellular machinery in *M. tuberculosis* as the pathogen is constantly exposed to DNA damaging conditions. As such, successful repair of damaged DNA is a key feature of virulence. *M. tuberculosis* is a pathogen that is constantly exposed to a myriad of DNA damaging stresses within the host macrophage rendering it important to understand the mechanisms involved in repairing DNA damage for the survival of the organism under these conditions.

The importance of DNA repair proteins in survival of *M. smegmatis* under oxidative stress conditions, exposure to UV and during mutagenesis has been previously studied (Moolla et al., 2014). These authors assessed the contribution of Nei and Nth DNA repair enzymes in *M. smegmatis*. *M. smegmatis* mutants lacking *neiI*, *neiII* and *nth* individually or in combination did not display aberrant growth in broth culture. The ∆*neiII* ∆*neiI* ∆*nth* triple mutant showed elevated spontaneous mutation frequencies compared to the wild type, the Fpg/Nei deficient mutants and their respective genetically complemented strains. These data confirmed that the Nth DNA glycosylase has an antimutator role in mycobacteria, similar to its demonstrated role in *E. coli* and *Salmonella typhimurium*, wherein the Nth DNA glycosylase also exerted a combinatorial effect with the Nei homologues in mutation avoidance (Altuvia et al., 1997). In comparison the study by Moolla et al also clearly demonstrated a novel interplay between the Nth and Nei homologues in DNA damage tolerance (with subsequent survival under oxidative stress conditions) and maintenance of spontaneous mutagenesis in *M. smegmatis* (Moolla et al., 2014). Hence, in the present study, we aimed to further interrogate the molecular basis of this interplay between the Nth and Nei homologues, with the particular aim to understand the individual contribution of the Nei homologues in DNA damage tolerance (and subsequent survival) under oxidative stress conditions.

***M. smegmatis* NeiI protein lacks a Fapy DNA glycosylase motif.**

Nei DNA glycosylases are members of the Fpg/Nei family of DNA glycosylases involved in the identification, excision and repair of oxidatively damage DNA. Genome analysis performed in this study identified two putative *nei* DNA glycosylase-encoding genes in *M. tuberculosis* and *M. smegmatis.* Bioinformatic analysis confirmed that the *M. tuberculosis* and *M. smegmatis* genes were DNA glycosylase homologues. Furthermore, it was observed that the *M. smegmatis* and *M. tuberculosis neiII*-encoded proteins had maintained the domains and motifs for DNA glycosylase activity, which include the N-terminal region that is predominantly β-sheet rich and is composed of a β-sandwich flanked by α-helices. In addition, these proteins retained the C-terminal domain, comprising α-helices, two of which form a conserved H2TH motif, as well as two anti-parallel β-strands that fold into a zinc finger motif necessary for protein function. Overall, the NeiII homologue showed a strong resemblance to previously characterised Nei DNA glycosylases in *Escherichia coli*, *Thermus thermophilus*, *Geobacillus stearothermophilus* and *Lactococcus lactis* (Zharkov et al., 2003). In addition, the zinc finger and H2TH motifs have been shown to be essential for Nei to bind to DNA. Both the *M. smegmatis* and *M. tuberculosis* NeiI proteins have maintained the domains and motifs required for function, however the *M. smegmatis* NeiI glycosylase lacks the Fapy DNA glycosylase motif. Therefore, from these observations, it is possible that the NeiII DNA glycosylase is the more dominant DNA glycosylase during BER as it possesses an additional Fapy DNA glycosylase motif. Alternatively, the absence/presence of the Fapy DNA glycosylate motif may confer some level of substrate specificity to these enzymes. It has been previously shown that most organisms have maintained one member of the Nei family of DNA glycosylases but rarely both, suggesting that the retention and duplication of these BER enzymes in intracellular organisms, such as *M. tuberculosis* and *S. typhimurium*, might be a deliberate adaptation that has allowed increased survival of these pathogens when exposed to oxidative stress (Suvarnapunya et al., 2003).

In order to identify the individual and/or combined role(s) of the Nei DNA glycosylase homologues in mycobacterial DNA repair, double knockout mutant strains that lacked the Nth DNA glycosylase in combination with either the NeiI homologue or NeiII homologue were generated in *M. smegmatis* by homologous recombination and phenotypically assessed. The resulting phenotypes were compared to the wild type strain as well as the single deletion mutant strains that lacked Nth or the respective individual Nei homologues.

**Nei DNA glycosylases are non-essential for growth under normal culture conditions.**

Comparison of growth kinetics for the parental, single *nei* and *nth* deletion mutants and double *nei*/*nth* deletion mutant strains showed no differences. These results compared well to previously published data, where a single Δ*neiI,* Δ*neiII* and Δ*nth* deletion mutants of *M. smegmatis* showed no growth defects under normal culture conditions (Goosens, 2009, Moolla et al., 2014). This observation clearly indicates that these glycosylases are not essential for survival. Interestingly, some bacterial species such as *H. pylori* and *Haemophilus influenzae* do not possess genes that encode for the Nei DNA glycosylases (Tomb et al., 1997) suggesting that in *M. smegmatis* not only the retention but the duplication of the Nei DNA glycosylases may be a deliberate evolutionary event for the survival of the organism during stressful conditions.

***M. smegmatis* strain lacking both Nth and NeiII DNA glycosylases is sensitive to hydrogen peroxide treatment.**

Under oxidative stress conditions as generated by 2.0 mM hydrogen peroxide ($H_2O_2$) the wild type strain showed no significant difference in killing rates compared to the untreated cells as measured by CFU/mL. However, variations in killing rates between biological replicates was noted which could be due to the decomposition of $H_2O_2$, since $H_2O_2$ spontaneously decomposes into water and oxygen, in a manner dependent on temperature, initial concentration and pH (Wilson et al., 2000). Mycobacterial cells also contain the catalase enzyme which as mentioned previously, is able to inactivate the $H_2O_2$ thus providing variable concentrations from one experiment to the next. Although

attempts were made to ensure conditions between experiments remained the same, there might have been differences in the effective concentration of $H_2O_2$ that the cells were exposed to between biological replicates.

To minimise the variation in the effective concentration of $H_2O_2$ and to determine the appropriate concentration of $H_2O_2$ required for killing the *M. smegmatis* wild type, the cells were washes twice as explained in section 3.5.2 and treated with 1 mM and 2 mM concentration of $H_2O_2$. The unwashed cells that were treated with 1 mM $H_2O_2$ showed no difference in killing compared to the washed cells. However, when the same unwashed cells were treated with 2 mM $H_2O_2$ a slight reduction in survival was observed over time. In contrast the cells that were washed showed enhanced killing at both concentrations of hydrogen peroxide, with a greater reduction in survival at the 2 mM concentration which conferred with previous observations by (Moolla et al., 2014). These data indicate that there clearly are components in the spent media that inactivate the $H_2O_2$, thus reducing its oxidative capacity and resulting in variable outcomes from one experiment to the next.

The single Δ*nth, ΔneiI and ΔneiII* deletion mutants showed no differences in survival under when exposed to $H_2O_2$ compared to the parental and the respective genetically complemented strains. This observation compared well to that observed for the *E. coli nth* and the *M. smegmatis neiI* and *neiII* deficient mutants (Cunningham and Weiss, 1985, Zheng and Zhu, 2003, Goosens, 2009). It is highly likely that survival of the *E. coli* and *M. smegmatis* strains in the absence of a single DNA glycosylase can be attributed to the ability of the remaining DNA glycosylases to compensate for the defect and maintain genome integrity under genotoxic conditions. However, in this study when both the *neiII* and *nth* genes were deleted, the combinatorial mutant displayed an exaggerated 3 log reduction in survival under oxidative stress conditions, compared to the single *nth* and *neiII* mutants suggesting that these two DNA glycosylases in combination are important during BER. However, when *neiI* was inactivated together with the *nth* glycosylase the reduction in growth was marginal (1 log) under oxidative stress conditions compared to the single *nth* and *neiI* mutants suggesting that the NeiII DNA glycosylase is a more dominant player in BER and that NeiI may be serving a

backup role. Similarly, *E. coli* (Zheng and Zhu, 2003) and *S. typhimurium* (Suvarnapunya et al., 2003) mutants deficient in both *nei* and *nth* genes displayed increased hypersensitivity which was absent in the respective single mutants. These data clearly indicate that *neiII* and *nth* DNA glycosylases have a major role in BER for the repair of damaged DNA under oxidative stress, and in mycobacteria, the NeiII DNA glycosylase seems to have a more significant role in genome maintenance compared to the NeiI DNA glycosylase.

The first reported study in mycobacteria that shows functional overlap between two DNA repair glycosylases (Nth and Nei) under oxidative stress was by Moolla et al (Moolla et al., 2014). In this study, these observations were further interrogated to establish if there was any functional hierarchy between the two Nei homologues. Since Δ*neiI* and Δ*neiII* single mutants did not display any survival defects under oxidative stress, it clearly indicates that in these mutant strains of *M. smegmatis* the remaining *nei* homologue and more significantly the Nth DNA glycosylase which has been shown to have anitmutator properties (Moolla et al., 2014) can compensate for the loss of the single Nei glycosylase. It has been biochemically proven that the Nth endonuclease and Nei DNA glycosylases have similar substrate specificity (Guo et al., 2010), they mainly excise pyrimidine lesions hence, it is highly possible that they have compensatory roles (Dizdaroglu, 2003) as generated by $H_2O_2$.

**Combinatorial loss of the *nth* and *neiII* genes yielded increased mutation rate to rifampicin.**

To further investigate the role of the individual Nei homologues in combination with the Nth DNA glycosylase, the number of spontaneous mutational events for the various single and double mutant strains was measured using the Luria-Delbrück fluctuation analysis (mutation rate), which measures the mutational events that occur within a cell during its lifetime and excludes pre-existing mutants that result in “jackpots” thus over estimating mutations (Rosche and Foster, 2000).

The single mutants (Δ*nth,* Δ*neiI* and Δ*neiII*) displayed only a slight elevation in mutation rates compared to the wild type strain. This observation was in contrast to the

previously observed mutation phenotypes wherein Moolla et al observed a significant increase in mutation rate in the absence of Nth thus identifying it to have antimutator properties in the BER pathway (Moolla et al., 2014, Goosens, 2009). It is possible that as these differences could be due to technical variations such as inoculum size, growth phase of cells which may have resulted in differences in the total final number of viable cells that were plated on rifampicin between the current and previously reported studies.

However, combinatorial deletion of the *nth* and *nei* DNA glycosylases resulted in the expected increase in spontaneous mutations with NeiII showing a 6-8 fold increase in mutation rates compared to the 2 fold increase observed with the simultaneous deletion of NeiI with Nth. This observation also correlated with the earlier observed phenotype of the Δ*neiII* Δ*nth* mutant being more sensitive to hydrogen peroxide than the Δ*nth* Δ*neiI* mutant. Thus, these data clearly confirm that NeiII is the major role player during DNA repair.

In conclusion this study provided insights on the relationship between the Nth and NeiI/NeiII homologues the base excision repair pathway during repair of damaged lesions. Based on the data obtained, it is clear that the loss of Nth and NeiII has a greater impact on the survival of *M. smegmatis* under oxidative stress conditions, compared to the combinatorial loss of Nth and NeiI. Therefore, a greater role can be assigned to NeiII compared to NeiI in the maintenance of mycobacterial genome integrity.

**Limitations of this study**

As hydrogen peroxide is unstable and results in variable experimental outcomes, novel robust assay systems are required to assess for DNA damage in mycobacteria. The current assays are based on the genetic assessment of the various mutants which, provides a crude measurement for genome maintenance based on survival and mutation rates. More sophisticated tools need to be developed that can perhaps measure specific lesions associated with specific DNA glycosylases in order to assign accurate function to them. Development in these areas will certainly intensify our understanding of the BER pathway in mycobacterial genome maintenance.

## Future studies

A measure of the expression levels of the remaining DNA glycosylases in the various combinatorial deletion mutants under normal culture conditions as well as under oxidative stress conditions will confirm and provide new information on the back up function of these various DNA glycosylases for genome maintenance. Isolation of mycobacterial proteins also presents a hurdle and therefore restricts the biochemical assessment of these various DNA glycosylases and their interaction with each other. Such confirmatory experiments will allow for better understanding of the role of each of the DNA glycosylases in BER and ultimately help delineate the complexities involved in the maintenance of the mycobacterial genome integrity under oxidative stress conditions.

# 5. References


ABEBE, F., HOLM-HANSEN, C., WIKER, H. G. & BJUNE, G. (2007) Progress in serodiagnosis of *Mycobacterium tuberculosis* infection. *Scand J Immunol,* 66**,** 176-91.

ADU, P. (2014) Role of chemokine receptors in inflammation-induced haematopoietic progenitor cell mobilisation. University of Glasgow.

AHMAD, S. (2010) Pathogenesis, immunology, and diagnosis of latent *Mycobacterium tuberculosis* infection. *Clinical and Developmental Immunology,* 2011.

AL-HARBI, A. M. (2012) Challenges of tuberculosis prevention through early detection of latent tuberculosis infection in new immigrants to the State of Kuwait.

ALBANNA, A. S., REED, M. B., KOTAR, K. V., FALLOW, A., MCINTOSH, F. A., BEHR, M. A. & MENZIES, D. (2011) Reduced transmissibility of East African Indian strains of *Mycobacterium tuberculosis*. *PloS one,* 6**,** e25075.

ALLAN, W. (1977) Controlled trial of 6-month and 9-month regimens of daily and intermittent streptomycin plus isoniazid plus pyrazinamide for pulmonary tuberculosis in hong-kong. *American Review of Respiratory Disease,* 115**,** 727-735.

ALTUVIA, S., WEINSTEIN-FISCHER, D., ZHANG, A., POSTOW, L. & STORZ, G. (1997) A small, stable RNA induced by oxidative stress: role as a pleiotropic regulator and antimutator. *Cell,* 90**,** 43-53.

BABIOR, B. M. (1999) NADPH oxidase: an update. *Blood,* 93**,** 1464-1476.

BATEY, R. T. G., SUNNY D MONTANGE, REBECCA K (2004) Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine. *Nature,* 432**,** 411-415.

BOSHOFF, H. I., 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**,** 183-193.

BRIEBA, L. G. E., BRANDT F KOKOSKA, ROBERT J DOUBLIÃ©, SYLVIE KUNKEL, TOM A ELLENBERGER, TOM (2004) Structural basis for the dual coding potential of 8â€oxoguanosine by a highâ€fidelity DNA polymerase. *The EMBO journal,* 23**,** 3452-3461.

BROOKS, P. C., MOVAHEDZADEH, F. & DAVIS, E. O. (2001) Identification of some DNA damage-inducible genes of *Mycobacterium tuberculosis*: apparent lack of correlation with LexA binding. *Journal of bacteriology,* 183**,** 4459-4467.

CADET, J. & WAGNER, J. R. (2014) Oxidatively generated base damage to cellular DNA by hydroxyl radical and one-electron oxidants: Similarities and differences. *Archives of biochemistry and biophysics,* 557**,** 47-54.

CADET, J. W., J RICHARD (2016) Radiation-induced damage to cellular DNA: Chemical nature and mechanisms of lesion formation. *Radiation Physics and Chemistry*.

CAVUSOGLU, C., HILMIOGLU, S., GUNERI, S. & BILGIC, A. (2002) Characterization of rpoB mutations in rifampin-resistant clinical isolates of *Mycobacterium tuberculosis from* Turkey by DNA sequencing and line probe assay. *Journal of clinical microbiology,* 40**,** 4435-4438.

CAWS, M., THWAITES, G., DUNSTAN, S., HAWN, T. R., LAN, N. T. N., THUONG, N. T. T., STEPNIEWSKA, K., HUYEN, M. N. T., BANG, N. D. & LOC, T. H. (2008) The influence of host and bacterial genotype on the development of disseminated disease with *Mycobacterium tuberculosis*. *PLoS Pathog,* 4**,** e1000034.

CHENG, K. C. C., DAVID S KASAI, HIROSHI NISHIMURA, SUSUMU LOEB, LAWRENCE A (1992) 8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G----T and A----C substitutions. *Journal of Biological Chemistry,* 267**,** 166-172.

COLE, S., BROSCH, R., PARKHILL, J., GARNIER, T., CHURCHER, C., HARRIS, D., GORDON, S., EIGLMEIER, K., GAS, S. & II III, C. B. (1998) Erratum: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature,* 396**,** 190.

COOKE, M. S. E., MARK D DIZDAROGLU, MIRAL LUNEC, JOSEPH (2003) Oxidative DNA damage: mechanisms, mutation, and disease. *The FASEB Journal,* 17**,** 1195-1214.

COOPER, A. M. (2009) Cell mediated immune responses in tuberculosis. *Annual review of immunology,* 27**,** 393.

CORBETT, E. L., WATT, C. J., WALKER, N., MAHER, D., WILLIAMS, B. G., RAVIGLIONE, M. C. & DYE, C. (2003) The growing burden of tuberculosis: global trends and interactions with the HIV epidemic. *Archives of internal medicine,* 163**,** 1009-1021.

CUNNINGHAM, R. P. & WEISS, B. (1985) Endonuclease III (nth) mutants of *Escherichia coli*. *Proceedings of the National Academy of Sciences,* 82**,** 474-478.

DANNENBERG, A. M. (1989) Immune mechanisms in the pathogenesis of pulmonary tuberculosis. *Review of Infectious Diseases,* 11**,** S369-S378.

DARWIN, K. H. & NATHAN, C. F. (2005) Role for nucleotide excision repair in virulence of *Mycobacterium tuberculosis*. *Infection and immunity,* 73**,** 4581-4587.

DAVID, S. S. & WILLIAMS, S. D. (1998) Chemistry of glycosylases and endonucleases involved in base-excision repair. *Chemical reviews,* 98**,** 1221-1262.

DE BOER, J. & HOEIJMAKERS, J. H. (2000) Nucleotide excision repair and human syndromes. *Carcinogenesis,* 21**,** 453-460.

DE LAAT, W. L., JASPERS, N. G. & HOEIJMAKERS, J. H. (1999) Molecular mechanism of nucleotide excision repair. *Genes & development,* 13**,** 768-785.

DIZDAROGLU, M. (2003) Substrate specificities and excision kinetics of DNA glycosylases involved in base-excision repair of oxidative DNA damage. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis,* 531**,** 109-126.

DOS VULTOS, T., MESTRE, O., TONJUM, T. & GICQUEL, B. (2009) DNA repair in *Mycobacterium tuberculosis* revisited. *FEMS microbiology reviews,* 33**,** 471-487.

DRABLA¸S, F., FEYZI, E., AAS, P. A., VAAGBA¸, C. B., KAVLI, B., BRATLIE, M. S., PEA-DIAZ, J., OTTERLEI, M., SLUPPHAUG, G. & KROKAN, H. E. (2004) Alkylation damage in DNA and RNA repair mechanisms and medical significance. *DNA repair,* 3**,** 1389-1407.

EHRLICH, M., ZHANG, X.-Y. & INAMDAR, N. M. (1990) Spontaneous deamination of cytosine and 5-methylcytosine residues in DNA and replacement of 5-methylcytosine residues with cytosine residues. *Mutation Research/Reviews in Genetic Toxicology,* 238**,** 277-286.

EHRT, S. & SCHNAPPINGER, D. (2009a) Mycobacterial survival strategies in the phagosome: defence against host stresses. *Cell Microbiol,* 11**,** 1170-8.

EHRT, S. & SCHNAPPINGER, D. (2009b) Mycobacterial survival strategies in the phagosome: defence against host stresses. *Cellular microbiology,* 11**,** 1170-1178.

EISEN, J. A. & HANAWALT, P. C. (1999) A phylogenomic study of DNA repair genes, proteins, and processes. *Mutation Research/DNA Repair,* 435**,** 171-213.

FENNER, J. E., STARR, R., CORNISH, A. L., ZHANG, J.-G., METCALF, D., SCHREIBER, R. D., SHEEHAN, K., HILTON, D. J., ALEXANDER, W. S. & HERTZOG, P. J. (2006) Suppressor of cytokine signaling 1 regulates the immune response to infection by a unique inhibition of type I interferon activity. *Nature immunology,* 7**,** 33-39.

FRIEDBERG, E. C. & GERLACH, V. L. (1999) Novel DNA polymerases offer clues to the molecular basis of mutagenesis. *Cell,* 98**,** 413-416.

FRIEDBERG, E. C., WALKER, G. C., SIEDE, W. & WOOD, R. D. (2005) *DNA repair and mutagenesis*, American Society for Microbiology Press.

FROMME, J. C. & VERDINE, G. L. (2002) Structural insights into lesion recognition and repair by the bacterial 8-oxoguanine DNA glycosylase MutM. *Nature Structural & Molecular Biology,* 9**,** 544-552.

GAGNEUX, S. & SMALL, P. M. (2007) Global phylogeography of *Mycobacterium tuberculosis* and implications for tuberculosis product development. *The Lancet infectious diseases,* 7**,** 328-337.

GAO, L., ZHOU, F., LI, X. & JIN, Q. (2013) HIV/TB co-infection in mainland China: a meta-analysis. *PloS one,* 5**,** e10736.

GATES, K. S. (2009) An overview of chemical processes that damage cellular DNA: spontaneous hydrolysis, alkylation, and reactions with radicals. *Chemical research in toxicology,* 22**,** 1747-1760.

GILLESPIE, S. H. (2002) Evolution of drug resistance in *Mycobacterium tuberculosis*: clinical and molecular perspective. *Antimicrobial agents and chemotherapy,* 46**,** 267-274.

GLAZIOU, P., FALZON, D., FLOYD, K. & RAVIGLIONE, M. (2013) Global epidemiology of tuberculosis. *Seminars in respiratory and critical care medicine.* Thieme Medical Publishers.

GOOSENS, V. J. (2009) The construction and phenotypic characterization of mycobacterial mutants deficient in DNA glycosylases.

GORDHAN, B. G. & PARISH, T. (2001) Gene Replacement using Pretreated DNA. *Methods Mol Med,* 54**,** 77-92.

GREER, S. & ZAMENHOF, S. (1962) Studies on depurination of DNA by heat. *Journal of molecular biology,* 4**,** 123-141.

GUNTHER, G., LETH, F. V., ALEXANDRU, S., ALTET, N., AVSAR, K., BANG, D., BARBUTA, R., BOTHAMLEY, G., CIOBANU, A. & CRUDU, V. (2015) Multidrug-resistant tuberculosis in Europe, 2010-2011.

GUO, Y., BANDARU, V., JARUGA, P., ZHAO, X., BURROWS, C. J., IWAI, S., DIZDAROGLU, M., BOND, J. P. & WALLACE, S. S. (2010) The oxidative DNA glycosylases of *Mycobacterium tuberculosis* exhibit different substrate preferences from their *Escherichia coli* counterparts. *DNA repair,* 9**,** 177-190.

GUTHLEIN, C., WANNER, R. M., SANDER, P., BOTTGER, E. C. & SPRINGER, B. (2008) A mycobacterial smc null mutant is proficient in DNA repair and long-term survival. *Journal of bacteriology,* 190**,** 452-456.

HASSIM, F., PAPADOPOULOS, A. O., KANA, B. D. & GORDHAN, B. G. (2015) A combinatorial role for MutY and Fpg DNA glycosylases in mutation avoidance in *Mycobacterium smegmatis*. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis,* 779**,** 24-32.

HAZRA, T. K., MULLER, J. G., MANUEL, R. C., BURROWS, C. J., LLOYD, R. S. & MITRA, S. (2001) Repair of hydantoins, one electron oxidation product of 8-oxoguanine, by DNA glycosylases of *Escherichia coli*. *Nucleic acids research,* 29**,** 1967-1974.

JOHNSON, J., HADAD, D., BOOM, W., DALEY, C., PELOQUIN, C., EISENACH, K., JANKUS, D., DEBANNE, S., CHARLEBOIS, E. & MACIEL, E. (2006) Early and extended early bactericidal activity of levofloxacin, gatifloxacin and moxifloxacin in pulmonary tuberculosis. *The International Journal of Tuberculosis and Lung Disease,* 10**,** 605-612.

KAHN, J., GULLY, R., COOPER, T., PERUMAL, A., SMITH, T. & KLEIN, D. (1987) Correlation of type A behaviour with adrenergic receptor density: implications for coronary artery disease pathogenesis. *The Lancet,* 330**,** 937-939.

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

KIM, Y.-J. & M WILSON III, D. (2012) Overview of base excision repair biochemistry. *Current molecular pharmacology,* 5**,** 3-13.

KIMERLING, M. E., LAMBREGTS-VAN WEEZENBEEK, K. & JARAMILLO, E. (2008) Programmatic Control of Multidrug-Resistant Tuberculosis. *Tuberculosis: The Essentials,* 237**,** 249.

KOWALCZYKOWSKI, S. C., DIXON, D. A., EGGLESTON, A. K., LAUDER, S. D. & REHRAUER, W. M. (1994) Biochemistry of homologous recombination in *Escherichia coli*. *Microbiological reviews,* 58**,** 401-465.

KURTHKOTI, K. & VARSHNEY, U. (2011) Base excision and nucleotide excision repair pathways in mycobacteria. *Tuberculosis,* 91**,** 533-543.

KURTHKOTI, K. & VARSHNEY, U. (2012) Distinct mechanisms of DNA repair in mycobacteria and their implications in attenuation of the pathogen growth. *Mechanisms of ageing and development,* 133**,** 138-146.

LEIPOLD, M. D., MULLER, J. G., BURROWS, C. J. & DAVID, S. S. (2000) Removal of hydantoin products of 8-oxoguanine oxidation by the *Escherichia coli* DNA repair enzyme, FPG. *Biochemistry,* 39**,** 14984-14992.

LINDAHL, T. & BARNES, D. (2000) Repair of endogenous DNA damage. *Cold Spring Harbor symposia on quantitative biology.* Cold Spring Harbor Laboratory Press.

LOUGHLIN, M. F., BARNARD, F. M., JENKINS, D., SHARPLES, G. J. & JENKS, P. J. (2003) *Helicobacter pylori* mutants defective in RuvC Holliday junction resolvase display reduced macrophage survival and spontaneous clearance from the murine gastric mucosa. *Infection and immunity,* 71**,** 2022-2031.

LV, X., TANG, S., XIA, Y., WANG, X., YUAN, Y., HU, D., LIU, F., WU, S., ZHANG, Y. & YANG, Z. (2013) Adverse reactions due to directly observed treatment strategy therapy in Chinese tuberculosis patients: a prospective study. *PloS one,* 8**,** e65037.

MARNETT, L. J. (2000) Oxyradicals and DNA damage. *Carcinogenesis,* 21**,** 361-370.

MASTER, S. S., RAMPINI, S. K., DAVIS, A. S., KELLER, C., EHLERS, S., SPRINGER, B., TIMMINS, G. S., SANDER, P. & DERETIC, V. (2008) *Mycobacterium tuberculosis* prevents inflammasome activation. *Cell host & microbe,* 3**,** 224-232.

MIESEL, L., WEISBROD, T. R., MARCINKEVICIENE, J. A., BITTMAN, R. & JACOBS, W. R. (1998) NADH dehydrogenase defects confer isoniazid resistance and conditional lethality in *Mycobacterium smegmatis*. *Journal of bacteriology,* 180**,** 2459-2467.

MITRA, S., BOLDOGH, I., IZUMI, T. & HAZRA, T. K. (2001) Complexities of the DNA base excision repair pathway for repair of oxidative DNA damage. *Environmental and molecular mutagenesis,* 38**,** 180-190.

MIZRAHI, V. & ANDERSEN, S. J. (1998) DNA repair in *Mycobacterium tuberculosis*. What have we learnt from the genome sequence? *Molecular microbiology,* 29**,** 1331-1339.

MLAMBO, C. K. (2004) Molecular and epidemiological characterization of multidrug-resistant *Mycobacterium tuberculosis* isolates in Johannesburg, South Africa. Faculty of Health Sciences, University of the Witwatersrand, Johannesburg.

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**,** 32-41.

MOSBAUGH, D. W. (1988) Purification and characterization of porcine liver DNA polymerase Î³: utilitzation of dUTP and dTTP during in vitro DNA synthesis. *Nucleic acids research,* 16**,** 5645-5659.

MUIR, T. W., SONDHI, D. & COLE, P. A. (1998) Expressed protein ligation: a general method for protein engineering. *Proceedings of the National Academy of Sciences,* 95**,** 6705-6710.

NARDELL, E. A. (2004) Catching droplet nuclei: toward a better understanding of tuberculosis transmission. Am Thoracic Soc.

NATHAN, C. & EHRT, S. (2004) Nitric oxide in tuberculosis. *Tuberculosis***,** 215-235.

NUNN, A. & TWEED, C. (2016) Treatment regimens for rifampicin-resistant tuberculosis: highlighting a research gap. *The international journal of tuberculosis and lung disease: the official journal of the International Union against Tuberculosis and Lung Disease,* 20**,** 1418.

O'SULLIVAN, D. M., MCHUGH, T. D. & GILLESPIE, S. H. (2005) Analysis of rpoB and pncA mutations in the published literature: an insight into the role of oxidative stress in *Mycobacterium tuberculosis* evolution? *Journal of Antimicrobial Chemotherapy,* 55**,** 674-679.

ORGANIZATION, W. H. & INITIATIVE, S. T. (2010) *Treatment of tuberculosis: guidelines*, World Health Organization.

PADAYATCHI, N., NAIDU, N., FRIEDLAND, G., NAIDOO, K., CONRADIE, F., NAIDOO, K. & OÂ€™DONNELL, M. R. (2017) Turning the tide against tuberculosis. *International Journal of Infectious Diseases,* 56**,** 6-9.

PALOMINO, J. C. (2009) Molecular detection, identification and drug resistance detection in *Mycobacterium tuberculosis*. *FEMS Immunology & Medical Microbiology,* 56**,** 103-111.

PARISH, T. & 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**,** 1969-1975.

PARSONS, L. M., SOMOSKÃ¶VI, Ã. K., GUTIERREZ, C., LEE, E., PARAMASIVAN, C., ABIMIKU, A. L., SPECTOR, S., ROSCIGNO, G. & NKENGASONG, J. (2011) Laboratory diagnosis of tuberculosis in resource-poor countries: challenges and opportunities. *Clinical microbiology reviews,* 24**,** 314-350.

PARUMASIVAM, T., LEUNG, S. S., QUAN, D. H., TRICCAS, J. A., BRITTON, W. J. & CHAN, H.-K. (2016) Rifapentine-loaded PLGA microparticles for tuberculosis inhaled therapy: Preparation and in vitro aerosol characterization. *European Journal of Pharmaceutical Sciences,* 88**,** 1-11.

PAWLOWSKI, A., JANSSON, M., SKALD, M., ROTTENBERG, M. E. & KALLENIUS, G. (2012) Tuberculosis and HIV co-infection. *PLoS Pathog,* 8**,** e1002464.

PHAM, T. T., JACOBS-SERA, D., PEDULLA, M. L., HENDRIX, R. W. & HATFULL, G. F. (2007) Comparative genomic analysis of mycobacteriophage Tweety: evolutionary insights and construction of compatible site-specific integration vectors for mycobacteria. *Microbiology,* 153**,** 2711-2723.

PRIHODA, N. D., ARJANOVA, O. V., YURCHENKO, L. V., SOKOLENKO, N. I., VIHROVA, L. A., PYLYPCHUK, V. S. & KUTSYNA, G. A. (2008) Adjunct immunotherapy of tuberculosis in drug-resistant TB and TB/HIV co-infected patients. *Int J Biomed Pharmaceut Sci,* 2**,** 59-64.

RAMASWAMY, S. & MUSSER, J. M. (1998) Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*: 1998 update. *Tubercle and Lung disease,* 79**,** 3-29.

RASSOOL, F. V. (2003) DNA double strand breaks (DSB) and non-homologous end joining (NHEJ) pathways in human leukemia. *Cancer letters,* 193**,** 1-9.

ROSCHE, W. A. & FOSTER, P. L. (2000) Determining mutation rates in bacterial populations. *Methods,* 20**,** 4-17.

SAMBROOK, J., FRITSCH, E. F. & MANIATIS, T. (1989) Molecular Cloning. A laboratory manual. *Cold Spring Harbour Laboratory Press.* Second Edition. ed. New York, Cold Spring Harbour.

SANCAR, A. (1996) DNA excision repair. *Annual review of biochemistry,* 65**,** 43-81.

SANCAR, A., LINDSEY-BOLTZ, L. A., ÃŒNSAL-KAÃ§MAZ, K. & LINN, S. (2004) Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annual review of biochemistry,* 73**,** 39-85.

SAPARBAEV, M., SIDORKINA, O. M., JURADO, J., PRIVEZENTZEV, C. V., GREENBERG, M. M. & LAVAL, J. (2002) Repair of oxidized purines and damaged pyrimidines by *E. coli* Fpg protein: different roles of proline 2 and lysine 57 residues. *Environmental and molecular mutagenesis,* 39**,** 10-17.

SAUKKONEN, J. J., POWELL, K. & JEREB, J. A. (2012) Monitoring for tuberculosis drug hepatotoxicity: moving from opinion to evidence. Am Thoracic Soc.

SCHAAF, H. S. & ZUMLA, A. (2009) *Tuberculosis: a comprehensive clinical reference*, Elsevier Health Sciences.

SHI, R., ITAGAKI, N. & SUGAWARA, I. (2007) Overview of anti-tuberculosis (TB) drugs and their resistance mechanisms. *Mini reviews in medicinal chemistry,* 7**,** 1177-1185.

SHUMAN, S. & GLICKMAN, M. S. (2007) Bacterial DNA repair by non-homologous end joining. *Nature Reviews Microbiology,* 5**,** 852-861.

SINHA, R. P. & HADER, D.-P. (2002) UV-induced DNA damage and repair: a review. *Photochemical & Photobiological Sciences,* 1**,** 225-236.

SMITH, I. (2003) *Mycobacterium tuberculosis* pathogenesis and molecular determinants of virulence. *Clin Microbiol Rev,* 16**,** 463-96.

SNAPPER, S., MELTON, R., MUSTAFA, S., KIESER, T. & WR JR, J. (1990) Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Molecular Microbiology,* 4**,** 1911-1919.

SOMERVILLE, W., THIBERT, L., SCHWARTZMAN, K. & BEHR, M. A. (2005) Extraction of *Mycobacterium tuberculosis DNA*: a question of containment. *J Clin Microbiol,* 43**,** 2996-7.

SOUTHERN, E. M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol,* 98**,** 503-17.

SUN, J., SCHAAF, K., DUVERGER, A., WOLSCHENDORF, F., SPEER, A., WAGNER, F., NIEDERWEIS, M. & KUTSCH, O. (2016) Protein phosphatase, Mg2+/Mn2+-dependent 1A controls the innate antiviral and antibacterial response of macrophages during HIV-1 and *Mycobacterium tuberculosis* infection. *Oncotarget,* 7**,** 15394-15409.

SUVARNAPUNYA, A. E., LAGASSE, H. & STEIN, M. A. (2003) The role of DNA base excision repair in the pathogenesis of *Salmonella enterica serovar Typhimurium*. *Molecular microbiology,* 48**,** 549-559.

TELENTI, A., IMBODEN, P., MARCHESI, F., MATTER, L., SCHOPFER, K., BODMER, T., LOWRIE, D., COLSTON, M. & COLE, S. (1993) Detection of rifampicin-resistance mutations in *Mycobacterium tuberculosis*. *The Lancet,* 341**,** 647-651.

THI, E. P., LAMBERTZ, U. & REINER, N. E. (2012) Sleeping with the enemy: how intracellular pathogens cope with a macrophage lifestyle. *PLoS Pathog,* 8**,** e1002551.

TOMB, J.-F., WHITE, O., KERLAVAGE, A. R., CLAYTON, R. A., SUTTON, G. G., FLEISCHMANN, R. D., KETCHUM, K. A., KLENK, H. P., GILL, S. & DOUGHERTY, B. A. (1997) The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *nature,* 388**,** 539-547.

ULRICHS, T. & KAUFMANN, S. (2002) Mycobacterial persistence and immunity. *Frontiers in bioscience: a journal and virtual library,* 7**,** d458-69.

VAN DEUN, A., BARRERA, L., BASTIAN, I., FATTORINI, L., HOFFMANN, H., KAM, K., RIGOUTS, L., RASCH-GERDES, S. & WRIGHT, A. (2009) *Mycobacterium tuberculosis* strains with highly discordant rifampin susceptibility test results. *Journal of clinical microbiology,* 47**,** 3501-3506.

VAN DEUN, A., MAUG, A. K., BOLA, V., LEBEKE, R., HOSSAIN, M. A., DE RIJK, W. B., RIGOUTS, L., GUMUSBOGA, A., TORREA, G. & DE JONG, B. C. (2013) Rifampicin drug resistance tests for tuberculosis: challenging the gold standard. *Journal of clinical microbiology***,** JCM. 00553-13.

WARD, J. F. (1973) Molecular mechanisms of radiation induced damage to nucleic acids. California Univ., Los Angeles (USA). Lab. of Nuclear Medicine and Radiation Biology.

WARNER, D. F. & MIZRAHI, V. (2011) Making ends meet in mycobacteria. *Molecular microbiology,* 79**,** 283-287.

WHO (2015) Global Tuberculosis report. Available at https://www.health-e.org.za/wp-content/uploads/2015/10/Global-TB-Report-2015-FINAL-2.pdf. *World Health Organization (WHO)*.

WHO (2016) Global tuberculosis report. Available at http://www.who.int/tb/publications/global_report/en/. In World Health Organization (WHO).

WILLIAMSON, D. A., BASU, I., BOWER, J., FREEMAN, J. T., HENDERSON, G. & ROBERTS, S. A. (2011) An evaluation of the Xpert MTB/RIF assay and detection of false-positive rifampicin resistance in *Mycobacterium tuberculosis*. *Diagnostic microbiology and infectious disease,* 74**,** 207-209.

WILSON, C. L., HINMAN, N. W., COOPER, W. J. & BROWN, C. F. (2000) Hydrogen peroxide cycling in surface geothermal waters of Yellowstone National Park. *Environmental science & technology,* 34**,** 2655-2662.

ZAHRT, T. C. & DERETIC, V. (2002) Reactive nitrogen and oxygen intermediates and bacterial defenses: unusual adaptations in *Mycobacterium tuberculosis*. *Antioxidants and Redox Signaling,* 4**,** 141-159.

ZHANG, Y. (2005) The magic bullets and tuberculosis drug targets. *Annu Rev Pharmacol Toxicol,* 45**,** 529-64.

ZHARKOV, D. O., GOLAN, G., GILBOA, R., FERNANDES, A. S., GERCHMAN, S. E., KYCIA, J. H., RIEGER, R. A., GROLLMAN, A. P. & SHOHAM, G. (2002) Structural analysis of an *Escherichia coli* endonuclease VIII covalent reaction intermediate. *The EMBO journal,* 21**,** 789-800.

ZHARKOV, D. O., SHOHAM, G. & GROLLMAN, A. P. (2003) Structural characterization of the Fpg family of DNA glycosylases. *DNA repair,* 2**,** 839-862.

ZHENG, L.-Y. & ZHU, J.-F. (2003) Study on antimicrobial activity of chitosan with different molecular weights. *Carbohydrate Polymers,* 54**,** 527-530.

ZU BENTRUP, K. H. N. & RUSSELL, D. G. (2001) Mycobacterial persistence: adaptation to a changing environment. *Trends in microbiology,* 9**,** 597-605.

# 6. Appendices

## Appendix A

### 6.1 Media preparation and solutions

### 6.1.1 *Mycobacterium smegmatis*

Middlebrook 7H9

4.9 g 7H9 powder, 0.2% glycerol, 0.05% Tween 80 (filter sterilized) and 0.2% glucose

Middlebrook 7H10

19 g 7H10 powder, 0.085 % NaCl, 0.2 % glucose, 0.2 % glycerol

### 6.2 *Escherichia coli*

2XTY

5 g sodium chloride, 10 g yeast extract, 16 g tryptone

Luria-Bertani Broth (LB)

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

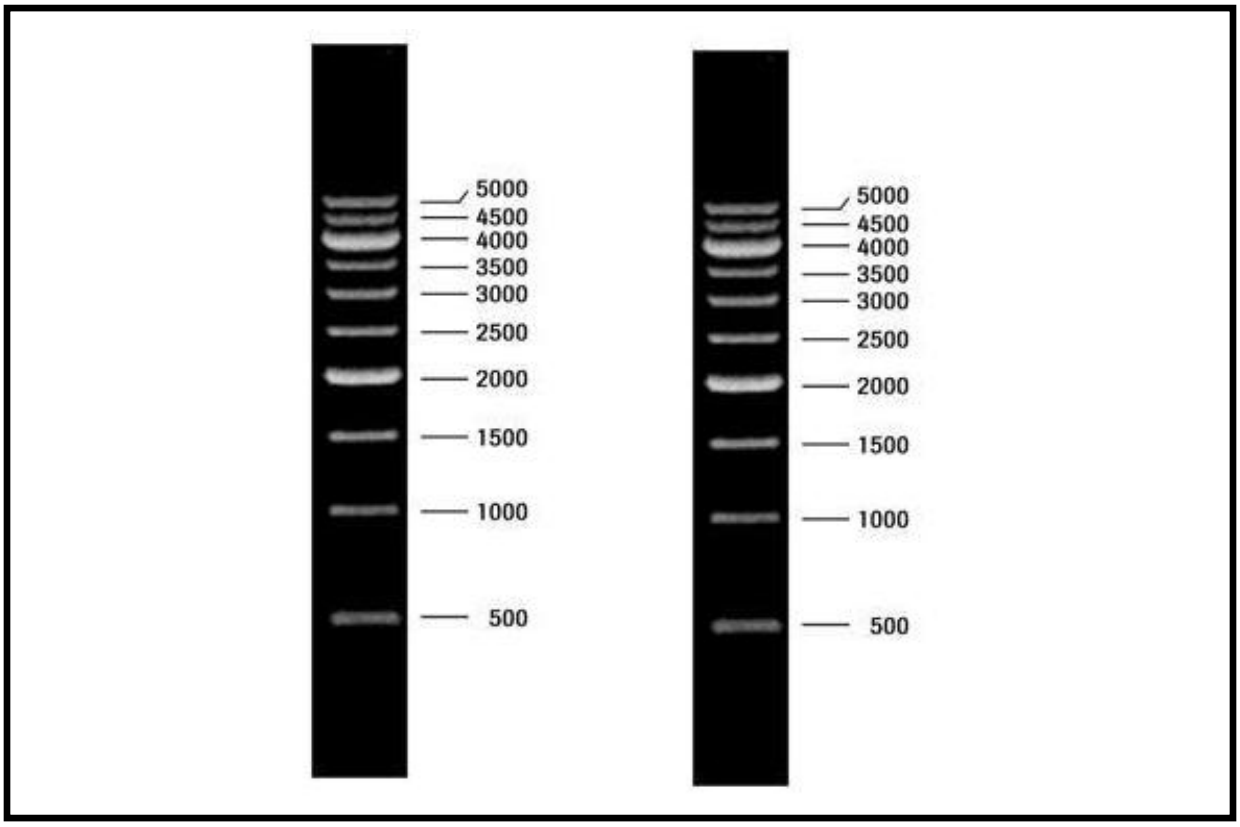
Luria-Bertani Agar (LA)

5 g yeast, 10 g tryptone, 10 g sodium chloride, 15 g agar

All media was made up to 1000 mL and autoclaved at 121 ˚C for 10 min.

## Appendix B

Lambda DNA molecular weight markers




The above DNA molecular weight markers III and VI used in this study were supplied by Roche Applied Science.

## Appendix C

## Maps of vectors used in this study

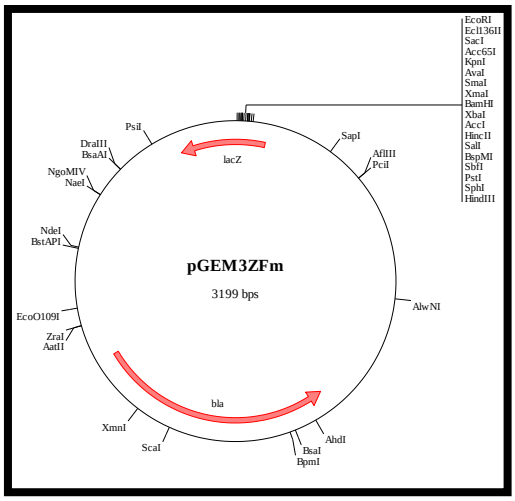


**Figure 6.1: pGEM3Zf(+).**Vector used for cloning upstream and downstream fragments.

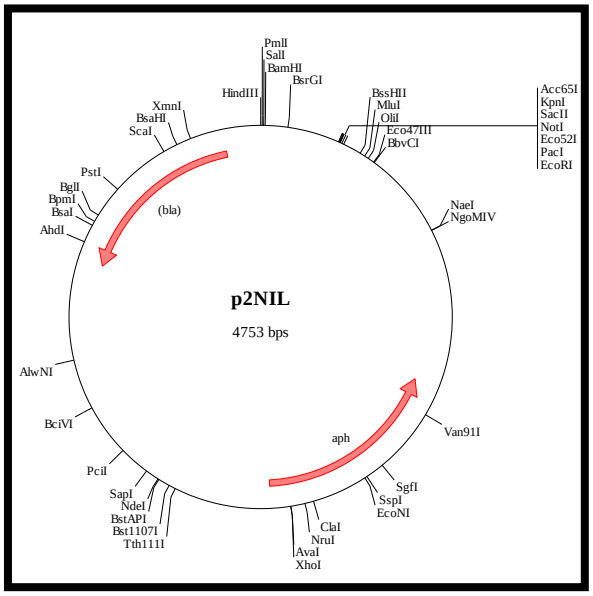



**Figure 6.2: p2NIL suicide vector**. Vector used in 3 way cloning, where the upstream and downstream fragments were simultaneously ligated into p2NIL.

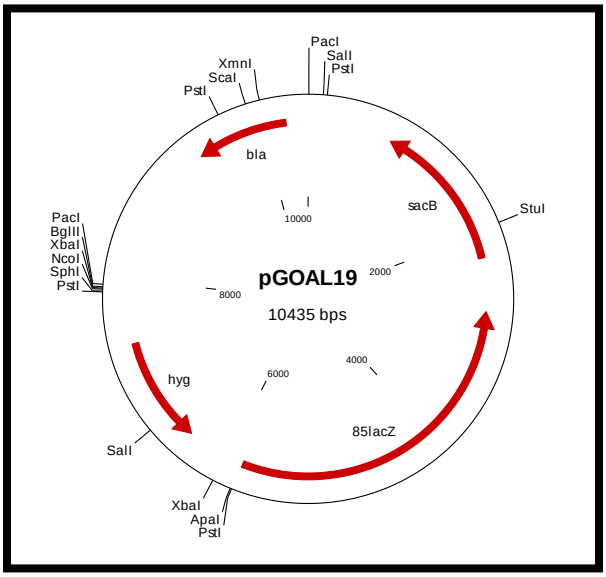



**Figure 6.3: pGOAL19**. Vector with the Pac cassette ( *sacB*, *lacZ*, *bla* and hyg) which was inserted into p2NIL to generate p2NILΔ*nth*::pGOAL19 ( Moolla et al., 2014).

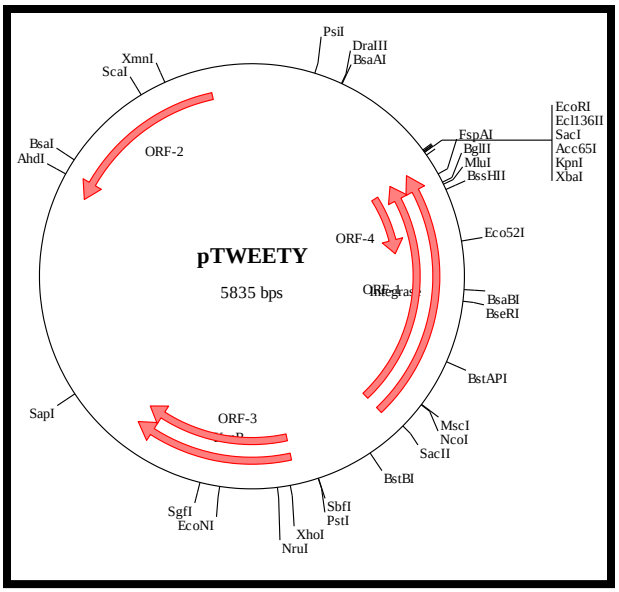



**Figure 6.4: pTWEETY.** Vector used for constructing the complementation vector.

## Appendix D

**Table 6.1:** PCR primers used in this study

| Name of fragment | Name of primers | Oligonucleotide sequence (5'- 3') | Annealing temperature and DNA polymerase | Amplicon properties |
|---|---|---|---|---|
| *M. smegmatis nth* | NthCF | AGCTTGACGGTC TGAAGGACCAGT | Taq/Phusion = 63.0 $^{o}$C | 1150 bp *nth* amplicon generated from wild type *M. smegmatis* genomic DNA including 350 bp upstream of the gene. |
| | NthCR | GGTACCCGAGAC GCTCATCGTGTG | | |
| Us | NthUsF1 | *AAGCTT* GTGTGCGTGCTG GCCAAG | Taq/Phusion = 63.5 $^{o}$C | 1185 bp amplicon generated from wild type genomic *M. smegmatis* DNA retaining 100 bp upstream of the *nth* gene |
| | NthUsR1 | *AGATCT* CACACCGGTGAT CGTCTT | | |
| Ds | NthDsF2 | *AGATCT* ACGACGCACGAG ACCCAG | Taq/Phusion = 63.5 oC | 1170 bp amplicon generated from the wild type *M. smegmatis nth* allele retaining 110 bp downstream of the gene. |
| | NthDsR2 | *GGTACC* AGCTGGGCGATC TCTTCC | | |

**Table 6.2:** Primers used for qPCR

| Name of primers | Oligonucleotide sequence (5'- 3') | Annealing temperature and DNA polymerase | Amplicon properties |
|---|---|---|---|
| RTF1 | TTCCTGGCCGAG CTTGAT | Taq = 56.5 ˚C | Amplification of fragment from position 135-360 in *nth* |
| RTR1 | ACAAGCGGGTCA ACCTGA | | |
| SigAF1 | GGGCGTGATGTC CATCTCCT | Taq = 56.0 ˚C | Amplification of fragment from position 367-488 in *sigA* |
| SigAR1 | GTATCCCGGTGC ATGGTC | | |
| MutYRT F1 | TGGCGCTCGGCC GGATTCC | Taq = 60 ˚C | Amplification of fragment from position 102-232 in *muty* |
| MutYRT R1 | GTCCACCAGCAG CGAGTC | | |

**Table 6.3:** PCR reagents and cycling conditions

| Reagents | 1X reaction |
|---|---|
| $dH_20$ | 16.6 µl |
| dNTPs | 8 µl |
| 5X G-C rich Sol. | 10 µl |
| Genomic DNA | 2 µl |
| Rev and fwd primers | 2.5 µl |
| Taq polymerase | 0.2 µl |
| Green buffer | 5 µl |
| Pcr dig probe | 2.5 µl |

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| 95°C | 95°C | 61°C | 72°C | 72°C | 16°C |
| 3 min | 0.30 | 2 min | 1 min | 10 min | ∞ |

34 X