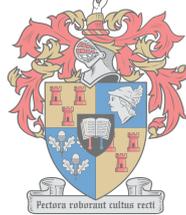


Functional characterisation
of a SufT homologue
in *Mycobacterium smegmatis*

by

Tsaone Tamuhla



Thesis presented in partial fulfilment of the requirements for the degree of Master of
Science in Molecular Biology in the Faculty of Medicine and Health Sciences at
Stellenbosch University



Supervisor: Dr Monique Joy Williams

Co-supervisor: Dr Danicke Willemse

December 2018

Author's declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third-party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Tsaone Tamuhla

Date: December 2018

Copyright © 2018 Stellenbosch University
All rights reserved

Abstract

Mycobacterium tuberculosis is one of the leading causes of death globally and with drug resistant tuberculosis (TB) on the rise, there is an urgent need to find new anti-TB drugs and drug targets. Increasing our understanding of the physiology of *M. tuberculosis* can aid in elucidating novel essential pathways which can be used as new drug targets. One such pathway is the iron-sulphur (Fe-S) biogenesis pathway, which is encoded by the *sufR-sufB-sufD-sufC-csd-sufU-sufT* operon (*suf* operon) in mycobacteria. Fe-S biogenesis is a vital process in cellular physiology yet the functioning of the Fe-S biogenesis machinery in mycobacteria is not fully understood. The last gene in the *suf* operon, *sufT*, encodes the only protein in the genome that contains a DUF59 domain. This study used targeted gene deletion and phenotypic characterisation of the resultant mutant to investigate the role of SufT in the physiology of mycobacteria, using *Mycobacterium smegmatis* as a model organism. An *M. smegmatis* Δ *sufT* knockout mutant harbouring an unmarked deletion in *sufT* was generated using allelic exchange mutagenesis. SufT was confirmed to be dispensable for growth in standard aerobic culture. Loss of SufT significantly decreased the activity of the Fe-S containing enzyme succinate dehydrogenase (SDH) and is therefore proposed to be a putative Fe-S maturation protein. No decrease in aconitase (ACN) activity was observed, suggesting that its role could be limited to certain Fe-S cluster proteins. Loss of SufT did not impact the survival of *M. smegmatis* after exposure to oxidative stress induced by the redox cycler 2,3-dimethoxy-1,4-naphthoquinone (DMNQ), or the sensitivity of *M. smegmatis* to the anti-TB drugs isoniazid, clofazimine or rifampicin. The *M. smegmatis* Δ *sufT* mutant displayed a growth defect during planktonic growth under iron limiting conditions. This defect was characterised by an extended lag phase, which was observed for all iron concentrations below 2 μ M. This suggests that SufT is needed for adaptation to growth under iron limitation. The exponential growth and final cell density achieved by the *M. smegmatis* Δ *sufT* mutant under iron limiting conditions was comparable to wild-type, suggesting that induction of a protein that compensates for the loss of *sufT* occurs. The study also confirmed that the cellular demand for iron during biofilm formation far exceeds that for planktonic growth, particularly during the maturation of the biofilms to form an extracellular matrix. This is the first study to functionally characterise SufT in mycobacteria providing a basis for further mechanistic studies.

Opsomming

Mycobacterium tuberculosis is een van die grootste oorsake van sterftes wêreldwyd en met antibiotika weerstandige tuberkulose wat aanhou toeneem, is dit noodsaaklik om nuwe antibiotika teen TB en antibiotika teikens te identifiseer. Deur ons kennis van die fisiologie van *M. tuberculosis* te verskerp kan nuwe noodsaaklike prosesse wat vir anti-TB antibiotika ontwikkeling gebruik kan word, gevind word. Een proses wat as potensieële teiken kan dien is die yster swawel (Fe-S) ko-faktor sintese proses, waarvoor die gene wat betrokke is in die proses in die *sufR-sufB-sufD-sufC-csd-sufU-sufT* (*suf*) operon enkodeer word in mycobacteria. Ten spyte van die belangrike rol wat Fe-S sintese in sellulêre fisiologie speel, word die proses nie ten volle verstaan in mycobacteria. Die laaste geen in die *suf* operon, *sufT*, enkodeer die enigste proteïen in die genoom wat 'n DUF59 domein bevat. In hierdie studie word geteikende geen verwydering en fenotipiese karakterisering van die mutant gebruik om die rol van SufT in die fisiologie van mycobacteria te ondersoek. *Mycobacterium smegmatis* is as 'n model organisme gebruik. 'n *M. smegmatis* Δ *sufT* mutant waarvan die *sufT* geen verwyder is deur alleliese uitruil mutagenese is gegeneer. Die verlies van SufT het geen effek op die groei van die bakterie onder standaard aerobiese kondisies gehad nie. Die aktiwiteit van die Fe-S ko-faktor bevattende ensiem succinate dehydrogenase (SDH) is wel verlaag, wat aandui dat SufT 'n Fe-S ko-faktor vorming proteïen is. Geen afname in aktiwiteit is waargeneem vir aconitase (ACN), wat aandui dat SufT slegs 'n rol speel in die vorming van Fe-S ko-faktore vir sekere proteïene. Die verlies van SufT het geen invloed gehad op die oorlewing van *M. smegmatis* na blootstelling aan die redoks genereerder 2,3-dimethoxy-1,4-naphthoquinone (DMNQ) of die sensitiwiteit van *M. smegmatis* tot anti-TB antibiotikas isoniasied, rifampisien en clofazemien. Die Δ *sufT* mutant het 'n groei defek gedurende planktoniese groei in yster beperkende media gehad. Hierdie defek is gekarakteriseer deur 'n toename in die lag-groEIFase en is in al die yster konsentrasies onder 2 μ M waargeneem. Dit stel voor dat SufT benodig word vir die aanpassing tot groei in yster beperkende kondisies. Die eksponensieële groei van die bakterie en finale sel digtheid wat op die ou einde bereik is, was vergelykbaar tussen die mutant en die ongemuteerde bakterie. Dit kan aandui dat 'n ander proteïen, met 'n soortgelyke funksie as SufT moontlik kan intree en die funksie daarvan oorneem. In hierdie studie is daar ook gewys dat die sellulêre vraag na yster gedurende biofilm groei baie hoër is as die van planktoniese groei, veral wanneer dit kom by die vorming van die biofilm se ekstrasellulêre matriks. Hierdie studie is die eerste om SufT in mycobacteria te karakteriseer en bied die basis vir addisionele meganistiese studies.

Acknowledgements

I would like to extend a heartfelt thank you to my supervisor, **Dr Monique Williams**, for giving me the opportunity to work on this project. You taught me a lot and it was a great pleasure working with you and learning from you. I will be forever grateful for having been given the opportunity to become part of the iron sulphur clusters cluster.

To my co-supervisor, **Dr Danicke Willemse**, thank you for your patience and support especially when imparting new technical skills. Thank you for the laughs, you made this MSc journey a fun one.

To the iron sulphur clusters cluster, thank you for the impromptu trouble shooting sessions and the many laughs we shared. It has been a pleasure sharing my MSc journey with you.

To **Dr Liezel Smith**, thank you for your support and encouragement, I am truly grateful.

Thank you to the **Harry Crossley Fund** for financial support on this research project.

Thank you to the **DST-NRF Centre of Excellence for Biomedical Tuberculosis Research**, through **Prof Gerhard Walzl** for the financial support in the form of a student bursary.

Finally, to my family, Neo, Rorisang, Ludo and Isabella, I am truly blessed to have you in my life. Thank you for supporting me in this journey.

Table of Contents

Author's declaration	ii
Abstract.....	iii
Opsomming	iv
Acknowledgements.....	v
Table of Contents.....	vi
List of Figures	xi
List of Tables	xiii
List of Abbreviations.....	xiv
Chapter 1	1
1. Literature review: Functional characterisation of protein domains.....	1
1.1 Introduction.....	1
1.2 Domains of Unknown Function (DUFs).....	1
1.3 Functional characterisation of DUFs	2
1.3.1 Sequence homology-based characterisation.....	3
1.3.1.1 Detection of remote homology	3
1.3.2 Structural homology-based characterisation.....	4
1.3.2 Biological and biochemical characterisation of proteins.....	5
1.4 Domain of unknown function 59 (DUF59)	6
1.4.1. Structural characterisation of DUF59 containing proteins	8
1.4.2 Functional characterisation of DUF59 containing proteins in eukaryotes	10
1.4.2.1 Mammalian DUF59 containing proteins	10
1.4.2.2 Yeast DUF59 containing proteins	10
1.4.2.3 Plant DUF59 containing proteins	11
1.4.3 Functional characterisation of bacterial DUF59 containing proteins.....	12
1.4.3.1 Bioinformatic analysis	12
1.4.3.2 Biochemical and biological characterisation of bacterial SufT proteins	13
1.4.3.2.1 <i>Staphylococcus aureus</i> (Gram-positive bacterium)	13
1.4.3.2.2 <i>Sinorhizobium meliloti</i> (Gram-negative bacterium)	14
1.4.3.3 DUF59 (SufT) containing proteins in mycobacteria.....	15
1.5 Conclusion.....	17

1.6 Study rationale.....	17
1.7 Approach.....	18
1.8 Hypothesis.....	18
1.9 Aim.....	19
1.9.1 Objectives	19
Chapter 2: Materials and Methods	20
2.1 Bacterial strains and culture conditions.....	20
2.1.1 Standard culture.....	20
2.1.2. Iron limitation.....	20
2.1.3 Assessment of bacterial growth.....	21
2.1.3.1 Preparation of whole cell lysates.....	22
2.1.3.1.1 Standard culture biofilms.....	22
2.1.3.1.2 Iron limitation biofilms.....	22
2.1.3.2 Bradford Assay	22
2.2 DNA isolation.....	23
2.2.1 Plasmid DNA isolation and purification from <i>E. coli</i>	23
2.2.2 Genomic DNA extraction from <i>M. smegmatis</i>	24
2.2.2.1 Large scale genomic DNA extraction	24
2.2.2.2 Small scale crude genomic DNA extraction	25
2.3 Cloning	25
2.3.1 Polymerase Chain Reaction (PCR)	25
2.3.2 DNA sequencing	27
2.3.3 Agarose gel electrophoresis.....	28
2.3.4 DNA fragment extraction and purification from agarose gels.....	28
2.3.5 Restriction endonuclease digestion	28
2.3.6 Ligation reactions	28
2.4 Transformation of <i>E. coli</i> XL1 Blue cells	29
2.4.1 Preparation of <i>E. coli</i> XL1 Blue chemically competent cells.....	29
2.4.2 Transformations	29

2.5 Transformation of <i>M. smegmatis</i> by electroporation	29
2.5.1 Preparation of electrocompetent <i>M. smegmatis</i> cells	29
2.5.2 Electroporation of <i>M. smegmatis</i>	30
2.6 Construction of the Δ <i>sufT</i> knockout mutant	30
2.6.1 Construction of Δ <i>sufT</i> allelic exchange vector	30
2.6.1.1 Generation of upstream and downstream regions for allelic exchange	30
2.6.1.2 Three-way cloning into p2NIL	31
2.6.1.3 Selectable marker cloning.....	31
2.6.2 Generation of single cross over (SCOs)	31
2.6.3 Generation of double cross over mutants (DCOs)	32
2.7 Southern Blot analyses.....	33
2.7.1 Southern transfer	33
2.7.2. Preparation of the DNA probes by PCR	33
2.7.3 Pre-hybridisation	33
2.7.4 ECL labelling of probe	34
2.7.5 Hybridisation	34
2.7.6 Detection of hybridisation	34
2.8 Genetic complementation	34
2.8.1 Generation of the <i>sufT</i> complementation vector	34
2.8.2 Sub-cloning into the pSE100 expression vector	35
2.8.3 Electroporation into competent <i>M. smegmatis</i> Δ <i>sufT</i>	35
2.9 Phenotypic characterisation.....	35
2.9.1 Growth kinetics	35
2.9.1.1 Standard culture media (7H9 GST) growth kinetics	35
2.9.2 Enzyme kinetics	36
2.9.2.1 Succinate dehydrogenase activity assay	36
2.9.2.2 Aconitase activity	37
2.9.3 Survival under oxidative stress conditions.....	37
2.9.4 Drug sensitivity testing	38

2.9.5 Iron limitation growth kinetics	39
2.9.6 Biofilm formation	39
2.9.6.1 Standard culture media (Sauton's) biofilm formation.....	39
2.9.6.2 Iron limitation pellicle biofilm formation.....	40
2.9.7 Statistical analyses.....	40
Chapter 3: Results	41
3.1 Construction of a <i>M. smegmatis</i> Δ <i>sufT</i> knockout mutant strain.....	41
3.1.1 Construction of Δ <i>sufT</i> allelic exchange vector	41
3.1.2 Two-step allelic exchange (Homologous recombination).....	44
3.1.2.1 Identification of SCOs and DCOs.....	44
3.1.3 Southern blot analysis	45
3.2 Genetic complementation	47
3.3 Phenotype characterisation	49
3.3.1 Growth in standard aerobic culture.....	49
3.3.2 Enzyme activity assays	49
3.3.2.1 Succinate dehydrogenase activity.....	50
3.3.2.2. Aconitase activity	51
3.3.3 Survival under oxidative stress.....	52
3.3.4 Drug sensitivity testing	53
3.3.5 Growth in iron limiting conditions.....	54
3.3.6 Standard culture (Sauton's media) pellicle biofilm formation	59
3.3.6.1 Method optimisation.....	60
3.3.6.2 Biofilm formation under standard culture conditions (Sauton's media)	61
3.3.7 Iron limitation biofilm formation.....	62
3.3.7.1 Impact of media components on pellicle biofilm formation	62
3.3.7.2 Biofilm formation under iron limiting conditions	63
3.3.4.3 Growth of pellicle biofilms under iron limitation in a mycosin 3 mutant	67
Chapter 4: Discussion and Conclusion.....	70
4.1 Discussion.....	70
4.2 Conclusion.....	75

References 76

List of Figures

Figure 1. 1. Iron-sulphur (Fe-S) cluster biogenesis systems in bacteria and eukaryotes	8
Figure 1. 2. Cartoon images of the predicted structure of DUF59 containing proteins in bacteria and eukaryotes.....	9
Figure 1. 3. Clustal Omega generated multiple sequence alignment of DUF59 domains in bacteria.....	12
Figure 1. 4. Proposed model of differential utilisation of Fe-S maturation proteins in <i>S. aureus</i>	14
Figure 1. 5. Schematic of the gene order in the <i>suf</i> operon of mycobacteria.....	15
Figure 2. 1. Construction of a <i>sufT</i> knockout mutant by two-step allelic exchange.....	32
Figure 3. 1. Agarose gel electrophoresis of 3127US (1953bp) and 3127DS (2029bp) allelic exchange PCR amplicons.....	41
Figure 3. 2. Restriction digest of pJET3127US and pJET3127DS.....	42
Figure 3. 3. Restriction map analyses of p2NIL Δ <i>sufT</i>	43
Figure 3.4. Restriction map analyses of p2NIL Δ <i>sufT</i> pGOAL17.....	44
Figure 3. 5. Genotype confirmation of <i>M. smegmatis</i> Δ <i>sufT</i> mutant strains by colony PCR.....	45
Figure 3. 6. Genotype characterisation of <i>M. smegmatis</i> Δ <i>sufT</i> mutant strains by Southern blot analysis.....	46
Figure 3. 7. Restriction digest of pJET3127.....	48
Figure 3. 8. Restriction map analyses of pSE3127.....	48
Figure 3. 9. Growth curve of <i>M. smegmatis</i> mc ² 155-1 (blue circles), Δ <i>sufT</i> (red squares) and Δ <i>sufT</i> (pSE3127) (green triangles) growing under standard aerobic culture conditions.....	49
Figure 3. 10. Activity of succinate dehydrogenase in <i>M. smegmatis</i> mc ² 155-1, Δ <i>sufT</i> and Δ <i>sufT</i> (pSE3127) relative to the activity in mc ² 155-1.....	51
Figure 3. 11. Activity of aconitase in <i>M. smegmatis</i> mc ² 155-1, Δ <i>sufT</i> and Δ <i>sufT</i> (pSE3127) relative to the activity in mc ² 155-1.....	52
Figure 3. 12. Effect on the survival of <i>M. smegmatis</i> mc ² 155, Δ <i>sufT</i> and Δ <i>sufT</i> (pSE3127) after exposure to 30 μ M DMNQ.....	53
Figure 3. 13. Growth curves of <i>M. smegmatis</i> mc ² 155-1 (blue circles), Δ <i>sufT</i> (red squares) and Δ <i>sufT</i> (pSE3127) (green triangles) in varying concentrations (A) 0 μ M, (B) 0.1 μ M, (C) 0.5 μ M and (D) 2 μ M of supplemental iron.....	55
Figure 3. 14. Sigmoidal four-parameter logistic regression (4PL) models of the growth curves of <i>M. smegmatis</i> mc ² 155-1 (blue circles), Δ <i>sufT</i> (red squares) and Δ <i>sufT</i>	

(pSE3127) (green triangles) in varying concentrations (A) 0 μM , (B) 0.1 μM , (C) 0.5 μM and (D) 2 μM of supplemental iron.	56
Figure 3. 15. Box and whisker plots of the logIC50 of <i>M. smegmatis</i> mc ² 155-1, ΔsufT and ΔsufT (pSE3127) under iron limitation at varying concentrations of iron.	57
Figure 3. 16. Box and whisker plot showing the increase in maximum cell density by <i>M. smegmatis</i> (A) mc ² 155, (B) ΔsufT and (C) ΔsufT (pSE3127) as iron concentration is increased.	58
Figure 3. 17. Growth curves of <i>M. smegmatis</i> mc ² 155 (blue circles), ΔsufT (red squares) and ΔsufT (pSE3127) (green triangles) after, (A) one growth cycle and (B) two growth cycles in MM with 0 μM supplemental iron.	59
Figure 3. 18. Optimised growth of <i>M. smegmatis</i> mc ² 155-1 pellicle biofilms in 24 well plates.	60
Figure 3. 19. Growth of pellicle biofilms in <i>M. smegmatis</i> mc ² 155-1, ΔsufT and ΔsufT (pSE3127) under standard culture conditions in Sauton's media.	61
Figure 3. 20. Impact of media components on pellicle biofilm formation in <i>M. smegmatis</i> mc ² 155-1.	63
Figure 3. 21. Pellicle biofilm formation of <i>M. smegmatis</i> mc ² 155-1 in increasing iron (Fe^{3+}) concentrations.	64
Figure 3. 22. Pellicle biofilm formation of <i>M. smegmatis</i> ΔsufT in increasing iron (Fe^{3+}) concentrations.	65
Figure 3. 23. Pellicle biofilm formation of <i>M. smegmatis</i> ΔsufT (pSE3127) in increasing iron (Fe^{3+}) concentrations.	66
Figure 3. 24. Quantification of the biofilm biomass in <i>M. smegmatis</i> mc ² 155-1, ΔsufT and ΔsufT (pSE3127) under iron limitation with varying concentrations of supplemental iron.	67
Figure 3. 25. Growth curve of <i>M. smegmatis</i> mc ² 155-2 (blue circles), $\Delta\text{MycP3ms}$ (red squares) and $\Delta\text{MycP3ms}::\text{Pr1MycP3ms}$ (green triangles) under iron limitation without the addition of supplemental iron.	68
Figure 3. 26. Pellicle biofilm formation of <i>M. smegmatis</i> (A) $\Delta\text{MycP3ms}$ (B) mc ² 155-2 in increasing iron (Fe^{3+}) concentrations.	69

List of Tables

Table 2. 1. List of bacterial strains used and generated in this study.....	21
Table 2. 1. List of bacterial plasmids used and generated in this study.	23
Table 2. 3. List of primers used in this study	26
Table 3. 1. Minimum inhibitory concentration for the drugs isoniazid, clofazimine and rifampicin.	54

List of Abbreviations

Amp	Ampicillin
<i>A. thaliana</i>	<i>Arabidopsis thaliana</i>
<i>B. anthracis</i>	<i>Bacillus anthracis</i>
bp	Base pairs
BSA	Bovine serum albumin
CFUs	Colony forming units
°C	Degrees Celsius
DCOs	Double cross overs
DNA	Deoxyribonucleic acid
DUF	Domain of Unknown Function
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetic acid
Fe-S	Iron-sulphur
Fe ³⁺	Ferric iron
xg	Centrifugal force
g	Gram
hrs	Hours
Hyg	Hygromycin
INH	Isoniazid
<i>Isc</i>	Iron sulphur cluster system
KatG	Catalase-peroxidase
km	Kanamycin
l	Litres
<i>lacZ</i>	β-galactosidase encoding gene
LA	Lysogeny broth agar
LB	Lysogeny broth
M	Molar
μ	Micro
mins	Minutes
ml	Milliliters
mM	Millimolar
MM	Chelex resin 100 treated mineral defined media
mV	Millivolt
<i>M. tuberculosis</i>	<i>Mycobacterium tuberculosis</i>
<i>M. smegmatis</i>	<i>Mycobacterium smegmatis</i>

nm	Nanometres
<i>nif</i>	Nitrogen fixation-specific system
OD	Optical density
PCR	Polymerase chain reaction
RE	Restriction enzyme
ROS	Reactive oxygen species
RIF	Rifampicin
s	Seconds
<i>S. meliloti</i>	<i>Sinorhizobium meliloti</i>
<i>S. aureus</i>	<i>Staphylococcus aureus</i>
SCO	Single cross over
<i>suf</i>	Sulphur assimilation
TB	Tuberculosis
Tet	Tetracycline
<i>T. moritima</i>	<i>Thermotoga moritima</i>
Tris	Tris(hydroxymethyl)aminomethane
Tween 80	Polyoxyethylene (20) sorbitan monooleate
UV	Ultraviolet
WHO	World Health Organization
X-gal	5-Bromo-4 Chloro-3 Indolyl β -D-galactosidase

Chapter 1

1. Literature review: Functional characterisation of protein domains

1.1 Introduction

Tuberculosis (TB) is a treatable and curable disease (WHO, 2017), but for centuries, *Mycobacterium tuberculosis*, the bacteria that causes TB has evaded eradication with devastating consequences to human lives (Daniel, 2006; Donoghue, 2009). In 2016 alone, 1.7 million people died from TB and 6.3 million new TB cases were reported (WHO, 2017). The emergence of drug resistant strains of *M. tuberculosis* is a growing public health concern (WHO, 2017) and has highlighted the urgent need to increase our understanding of the pathogenesis of *M. tuberculosis*.

Understanding the mechanisms that drive bacterial pathogenesis relies on the identification of the proteins that aid their survival and proliferation in host cells (Hensel & Holden, 1996; Welch 2015). In this regard, TB research was greatly accelerated by the publication of the genome of the laboratory strain of *M. tuberculosis* H37Rv (Cole et al., 1998) which contributed significantly to research efforts aimed at improving the understanding of the pathogenesis and general physiology of the bacteria (Cole et al., 1998; Sasseti et al., 2003; Mao et al., 2013; Ramakrishnan et al., 2015). However, our understanding of the physiology of *M. tuberculosis* is still limited because there are over 1000 *M. tuberculosis* proteins, which still need functional annotation (Mao et al., 2013). This is concerning because proteins drive cellular processes and each protein has (a) specific role(s) to play in an organisms' metabolism (Bateman et al., 2010, Goodacre et al., 2013). There is therefore a need to prioritise the functional characterisation of proteins in pathogens like *M. tuberculosis* because without complete knowledge of their functions, our understanding of the mechanisms that drive their pathogenesis will remain limited (Hensel & Holden, 1996; Prakash et al., 2011; Goodacre et al., 2013)

The aim of this review is to highlight the strategies that have been used to determine the function of protein domains that lack functional annotation. The application of this approach to determine the function of Domain of Unknown Function/DUF 59 will also be discussed.

1.2 Domains of Unknown Function (DUFs)

Protein sequence data is curated online in the Universal Protein Resource (UniProt) which is the most comprehensive protein sequence database (Mulder et al., 2007). The sequences in UniProt are further curated as families based on statistically significant sequence similarity in databases such as Interpro (Apweiler et al., 2001), Cluster of Orthologous genes (COG)

(Tatusov et al., 2000), PROSITE (Sigrist et al., 2009) and Pfam (Finn et al., 2010). Each protein family database classifies families on different themes (Xu, 2014). In COG, families are based on evolutionary relatedness (Tatusov et al., 2000) and in Interpro, PROSITE and Pfam, families are based on functional relatedness (Apweiler et al., 2001, Sigrist et al., 2009; Bateman et al., 2010). Even though Interpro, PROSITE and Pfam serve a similar purpose, the protein families in Pfam are discussed in this review because the database specialises in the curation of protein domains of unknown function (DUFs) (Bateman et al., 2010; Punta et al., 2012). DUFs are families of conserved protein domains that have not been assigned a function and the overarching aim of Pfam is to functionally annotate all these domains (Bateman et al., 2010; Punta et al., 2012). Protein domains are the functional and/or structural unit of a protein, and a protein can be made up of one or more domains (Bateman et al., 2010, Goodacre et al., 2013).

The number of DUF's has been increasing with each version of Pfam that is released (Punta et al., 2012). Between 2010 and 2015, the number of DUFs increased by 45% with more than 1000 new DUF families being added to the database in a five-year period (Bateman et al., 2010; Mudgal et al., 2015). In addition, DUFs currently represent a quarter of all the protein families in Pfam (Mudgal et al., 2015). This rapid increase in the number of DUFs has been attributed to numerous factors, but the most cited is the prolific use of whole genome sequencing, which has resulted in an increase in the number of sequenced genomes available for annotation (Galperin and Koonin, 2004; Jaroszewski et al., 2009; Bateman et al., 2010; Buttigieg et al., 2013; Goodacre et al., 2014; Fang and Gough, 2013). While newly sequenced genomes are the biggest contributors to this constant increase (Galperin and Koonin, 2004), even well studied organisms such as *Saccharomyces cerevisiae* (Fidler et al., 2015) and *Arabidopsis thaliana* (Neihaus et al., 2015) have a considerable proportion of their genome lacking functional annotation.

DUFs are ubiquitously distributed in all life forms (Jaroszewski et al., 2009) and concerted efforts have been made to characterise DUFs in eukaryotes (Hausmann et al., 2005; Horan et al., 2008; Schwenkert et al., 2010 Luo et al., 2012; Stehling et al., 2013), while bacterial DUFs have been neglected. The need to assign functions to DUFs therefore represents a major research challenge.

1.3 Functional characterisation of DUFs

The functional characterisation of DUFs is a multi-step process, which is anchored in the ability to accurately predict or infer a hypothetical function for a domain by establishing homology with proteins of known function (Jaroszewski et al., 2009; Bateman et al., 2010; Punta et al., 2011; Fidler et al., 2016). The exponential increase in the number of sequenced genomes

available for annotation has increased the reliance on the use of computational methods as search tools for finding biologically relevant data (Galperin and Koonin, 2004; Mulder et al., 2007; Atkinson et al., 2009; Jaroszewski et al., 2009; Bateman et al., 2010; Buttigieg et al., 2013; Goodacre et al., 2014; Fang and Gough, 2013).

1.3.1 Sequence homology-based characterisation

Sequence homology-based characterisation has become a fundamental starting point in the functional characterisation of DUFs based on the principle that proteins with homologous sequences have similar functions (Mulder et al., 2007; Atkinson et al., 2009; Ramakrishnan et al., 2015). However, the usefulness of sequence-based homology relies on the accurate detection of homology between sequences (Eddy, 1996; Dunbrack, 2006; Mulder et al., 2007; Mudgal et al., 2015; Ramakrishnan et al., 2015). Traditional homology detection methods use pairwise alignment algorithms such as Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1990) to compare the nucleotides or amino acids in protein sequences and determine the probability of two sequences being true homologues (Altschul et al., 1990; Eddy, 1996). While useful in identifying homologues with a high sequence similarity, pairwise alignment tools like BLAST lack the sensitivity to detect remote homologies (Dunbrack, 2006; Mulder et al., 2007; Ramakrishnan et al., 2015).

1.3.1.1 Detection of remote homology

Improved computing power has seen the development of newer, more sensitive homology detection tools such as the widely used profile hidden Markov models (HMMs) (Eddy, 1996; Eddy, 1998) which have helped to increase the utility of sequence-based homology (Park et al., 2005; Pearson & Siek, 2005). Unlike pairwise alignments, which compare two sequences, HMMs compare protein profiles which have been generated from multiple protein sequence alignments (Altschul et al., 1997; Eddy, 1996; Eddy, 1998; Mulder et al., 2007). HMMs have increased sensitivity and can detect remote homologies missed by BLAST searches (Park et al., 2005; Prakash et al., 2011; Mudgal et al., 2015; Fidler et al., 2016). When used retrospectively to search the Protein Data Bank (PDB), for PH-like proteins in *S. cerevisiae*, HH_{search} (HMMs based search tool) identified 1200 positive matches, while BLAST could only match 350 sequences (Fidler et al., 2011). Although the function predictions inferred through sequence homology need to be validated experimentally, confidence in the reliability of the results produced using these techniques should increase as computational biology tools continue to improve (Pearson & Sierk, 2005; Fidler et al., 2016).

1.3.2 Structural homology-based characterisation

Structural homology has also enjoyed wide usage in the characterisation of DUFs because structural homology is a good predictor of function based on the principle that conserved structure is an indication of conserved function (Jaroszewski et al., 2009). As with sequence-based homology, the use of HMMs based tools to determine structural homology has improved structure homology-based function predictions (Jaroszewski et al., 2009; Mudgal et al., 2015). In addition, when used synergistically, sequence and structural homology have proven to be powerful methods in predicting the functions of DUFs (Fang & Gough, 2011; Fidler et al., 2015; Mudgal et al., 2015; Ramakrishnan et al., 2015). This coupled approach to functional characterisation was successfully used to assign putative functions to 26 DUFs in *M. tuberculosis* (Ramakrishnan et al., 2015) and to reassign 614 DUFs to protein families with a known function (Mudgal et al., 2015). The function of DUF3233 in *Vibrio cholerae* as an autotransporter was also successfully predicted using this approach (Prakash et al., 2011).

While these results demonstrate the value of combining sequence and structural homology, the biggest limitation of *in silico* prediction of homology is that sometimes the predicted functions are discordant with the physiologic relevance of those functions in the organism (Mao et al., 2013; Ramakrishnan et al., 2015). This was illustrated in *M. tuberculosis* where the protein encoded by the gene *Rv0141* was predicted with 99.9% confidence to share structural homology with enzymes involved in the degradation of cyclic terpenes and biosynthesis of nitrogen containing polycyclic phenazines (Mao et al., 2013). However, those specific biosynthesis and degradation pathways have not been identified in *M. tuberculosis* casting doubt on the reliability of the function prediction on the premise that encoding for an enzyme used in a non-existing metabolic pathway does not make physiological sense (Mao et al., 2013). Therefore, the function of *Rv0141* could not be inferred based on those results alone and will still need to be confirmed experimentally (Mao et al. 2013). Similarly, in *M. tuberculosis*, two members of the DUF4185 family, *Rv1754* and *Rv3707*, were also predicted with 95-100% confidence to be structurally homologous to a family of enzymes called sialidases (Ramakrishnan et al., 2015). However, because there are no sialidases in mycobacteria, this assignment is questionable and will have to be confirmed experimentally (Ramakrishnan et al., 2015). The results from these studies highlight the importance of factoring in physiological relevance when interpreting *in silico* function predictions (Mao et al., 2013; Ramakrishnan et al., 2015).

Genomic enzymology is the term used to describe the systematic process of predicting the functions and associated metabolic pathways of enzymes based on shared sequence homology and genomic co-location (Gerlt, 2016; Zhang et al., 2016). It has been proposed as

an accurate and robust way of making large-scale function predictions (Zhang et al., 2016) using sequence similarity networks (SSNs) (Atkinson et al., 2009) and genome neighbourhood networks (GNNs) (Zhao et al., 2014). SSNs are based on multiple comparisons of pairwise protein alignments, which are presented as clusters based on homology (Atkinson et al., 2009). Each protein in the network is represented by a “node” and is joined to other proteins in the cluster by lines, which represent the similarity scores for each alignment (Atkinson et al., 2009; Gerlt, 2016). GNNs search for gene order conservation (synteny) and assign functions to genes on the principle that genes that are organised as clusters or operons in the genome are likely to be involved in the same metabolic pathway (Guerrero et al., 2005; Zhao et al., 2014). The utility of genomic enzymology in assigning functions to DUFs was illustrated in the functional characterisation of DUF1537 (Zhang et al., 2016). The researchers used SSNs and GNNs synergistically to form an accurate function hypothesis about the potential *in vitro* characteristics and *in vivo* functions of DUF1537 containing proteins. These hypotheses were then used to inform the design of suitable biological and biochemical assays to test the predicted functions and the precise function of DUF1537 containing proteins as novel ATP-dependent four-carbon acid sugar kinases was subsequently confirmed (Zhang et al., 2016).

1.3.2 Biological and biochemical characterisation of proteins

Biological and biochemical characterisation remains the gold standard for confirming *in silico* protein function predictions. A good illustration of this was in the functional characterisation of DUF1792 (Zhang et al., 2014). In this study, analysis of sequence and structural homology data for DUF1792 was used to accurately hypothesise that the domain had a similar fold to glycotransferases. This hypothesis was then used to inform the experimental design of *in vitro* glycosylation assays and mass spectrophotometry which confirmed that *in vivo*, DUF1792 was a novel glycotransferase, which was needed for the third step of the glycosylation of the Fimbriae-associated protein (Fap1) (Zhang et al., 2015). The use of this approach has however proven to be challenging (Goodacre et al., 2013) because, these methods are resource intensive and limited by a need to have an accurate functional hypothesis to guide the experimental design (Jaroszewski et al., 2009; Goodacre et al., 2013; Neihaus et al., 2015).

Traditional biological and biochemical functional characterisation strategies lack robustness in that usually only one protein is studied at a time. In this regard, *in silico* tools like genetic enzymology provide a more robust strategy where multiple proteins can be studied simultaneously, and this can help to speed up the functional characterisation of genes in clusters or operons. This approach will have great applicability in the functional characterisation of bacterial DUFs because bacterial genomes display a high level of synteny

and many of the enzymes involved in fundamental metabolic pathways in bacteria are encoded by operons (Guerrero et al., 2005; Zhao et al., 2014).

1.4 Domain of unknown function 59 (DUF59)

There are 3710 species of archaea, bacteria and eukaryotes that have a protein containing a DUF59 domain, but the domain is more prevalent in bacteria (Finn et al., 2016). DUF59 containing proteins are arranged in 39 different domain architectures, however 64% of all the DUF59 domain containing proteins in Pfam are made up entirely of the DUF59 domain, making it the most prevalent domain arrangement (Finn et al., 2015; Mashruwala et al., 2016). The second most common architecture (32%) consists of a DUF59 domain at the N-terminus and a P-Loop NTPase at the C-terminus of the gene-coding region (Finn et al., 2015; Mashruwala et al., 2016).

Proteins containing the DUF59 domain are hypothesized to be involved in various vital cellular activities such as phenyl acetate catabolism (Grishin and Cygler, 2015), DNA repair (Luo et al., 2012) and iron sulphur (Fe-S) cluster biogenesis (Almeida et al., 2005; Chen et al., 2012; Luo et al., 2012; Mashruwala et al., 2016a; Mashruwala et al., 2016b; Sasaki et al., 2016; Stehling et al., 2013). However, in Pfam version 31 the DUF59 family was assigned a putative function as iron sulphur (Fe-S) assembly proteins (Pfam ID: PF01883) (Finn et al., 2016).

Some proteins require co-factors to become functionally active (Lill & Mühlenhoff, 2006). One such group of co-factors is Fe-S clusters, which are found in all living organisms (Lill & Mühlenhoff, 2006). Fe-S clusters are co-factors to proteins involved in an array of fundamental cellular processes such as respiration, DNA repair, photosynthesis, and nitrogen assimilation, so they play a vital role in cellular physiology (Castro et al., 2008; Py et al., 2011). Fe-S clusters are unstable when they are not bound to protein and are synthesised *de novo* in the cell in a tightly regulated process which requires complex assembly systems (Lill & Mühlenhoff, 2006; Fontcave & Ollagnier-de-Choudens, 2007; Blanc et al., 2014).

There are three Fe-S biogenesis pathways in bacteria (Fig 1.1B), namely the nitrogen-fixation-specific (*nif*), iron sulphur cluster (*isc*) and sulphur assimilation (*suf*) system (Johnson et al., 2005; Ayala-Castro et al., 2008). These systems are encoded as operons (Fontcave & Ollagnier-de-Choudens, 2007) and the number and type of system varies among bacterial species. The *nif* system is specific to nitrogen fixing bacteria, while in gram-negative bacteria, the *isc* system generally serves as the house-keeping system and the *suf* system is used under conditions of cell stress (Castro et al., 2005), except in *Sinorhizobium melliloti*, which contains only the *suf* system (Sasaki et al., 2016). Unlike Gram-negative bacteria which contain multiple Fe-S biogenesis systems, Gram-positive bacteria such as *Bacillus subtilis*

(Selbach et al., 2013) and *Staphylococcus aureus* (Mashruwala et al., 2016) generally contain only the *suf* system for Fe-S biogenesis (Santos et al., 2014). Mycobacteria also contain the *suf* system and while it is the primary Fe-S biogenesis system (Huet et al., 2005), they also contain a single component of the *isc* system, IscS, which is also involved in Fe-S biogenesis (Rybniker et al., 2014). It has been suggested that the type of system present in a bacterium is highly dependent on the conditions to which it is exposed (Py and Barras, 2010; Jang and Imlay, 2010).

The first step in Fe-S cluster biogenesis requires the mobilisation of sulphur by a cysteine desulphurase and iron from an unknown source (Fontcave & Ollagnier-de-Choudens, 2007). The cluster is then assembled directly on a scaffold protein or may be transferred via transfer proteins to a scaffold protein (Fontcave & Ollagnier-de-Choudens, 2007). Next, the Fe-S cluster is transferred either via transfer proteins or directly to the apoproteins (Fig. 1.1A) (Fontcave & Ollagnier-de-Choudens, 2007). Transfer of assembled clusters to apoproteins is necessary since Fe-S clusters are sensitive to oxidation and in their free form they fuel intracellular oxidative stress by Fenton chemistry (Wardman et al., 1996).

The steps involved in Fe-S biogenesis in eukaryotes are conserved and many of the bacterial proteins have eukaryote homologues (Lill & Mühlenhoff, 2006). However, the major difference is that Fe-S biogenesis in eukaryotes is compartmentalised (Fig. 1.1C) (Hausmann et al., 2005). In the mitochondria, an *isc* like system is used and in the plastids of photosynthetic eukaryotes, Fe-S clusters are assembled using a *suf*-like system (Balk & Pilon, 2011; Lill & Mühlenhoff, 2006 (Fig. 1.1). In the cytosol, the cytosolic Fe-S cluster assembly (CIA) pathway is used to assemble Fe-S clusters some of which are transported to the nucleus, however, the CIA system is not independent and some of its' components require prior synthesis by the *isc*-like system (Fig. 1.1C) (Balk & Pilon, 2011; Lill & Mühlenhoff, 2006).

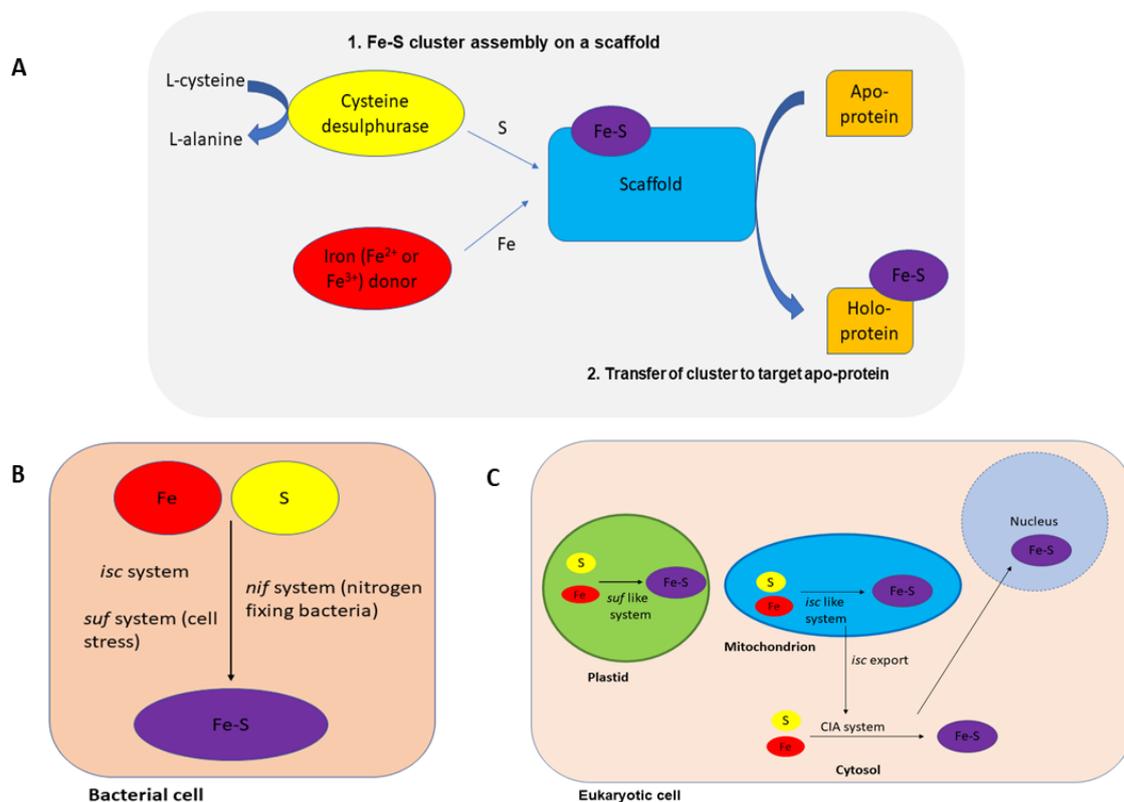


Figure 1. 1. Iron-sulphur (Fe-S) cluster biogenesis systems in bacteria and eukaryotes. (A). Schematic showing the steps of Fe-S cluster biogenesis, which are conserved for all the Fe-S biogenesis systems. (B). Bacteria use three Fe-S biogenesis systems, namely the *isc*, *nif* and *suf*. (C). In eukaryotes Fe-S Biogenesis is compartmentalised, in the plastids the *suf* like system is used, in the cytosol the CIA system and in the mitochondria the *isc* like system. This figure is adapted from Fontcave & Ollagnier-de-Choudens, 2007 and Lill & Mühlhoff, 2006.

1.4.1. Structural characterisation of DUF59 containing proteins

The three published structures of eukaryotic DUF59 containing proteins are of the mammalian protein Fam96a (Chen et al., 2012; Ouyang et al., 2013). The solution nuclear magnetic resonance (NMR) structure of Fam96a (PDB ID:2M5H) showed a secondary structure that consisted of a combination of five α -helices and three-strand mixed β -sheets (Fig. 1.2) (Ouyang et al., 2012). In addition, the solution NMR structure revealed a closed monomeric conformation, which is typical of a domain swapping protein (Ouyang et al., 2013). Domain swapping is when monomeric proteins exchange identical structural elements such as α -helices or β -sheets to form dimers or oligomers (Liu & Eisenberg, 2002). The crystal structure of Fam96a showed that it formed two distinct types of domain-swapped dimers (Chen et al., 2012). These were designated Fam96a major dimer (PDB ID:3UX2) and Fam96b minor dimer (PDB ID:3UX3) (Fig. 1.2) (Chen et al., 2012). A comparison of the two dimers showed that they had the same subunit arrangement and that in both dimers it is the α -helix that is swapped

(Chen et al., 2012). However, the subunits were hinged differently in the two dimeric states and the formation of the minor dimer required zinc (Fig. 1.2) (Chen et al., 2012). In addition, they also found that the regions involved in domain swapping are highly conserved in eukaryotic but not bacterial DUF59 containing proteins, suggesting that the latter do not dimerise (Chen et al., 2012).

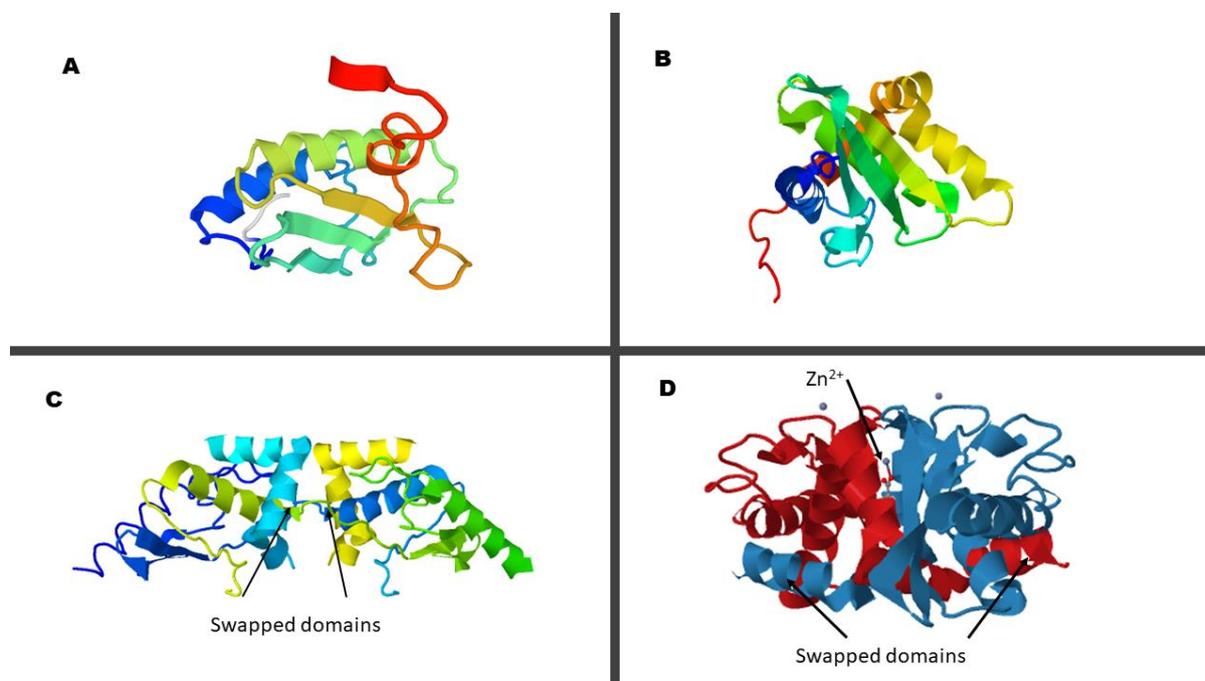


Figure 1. 2. Cartoon images of the predicted structure of DUF59 containing proteins in bacteria and eukaryotes. A solution nuclear magnetic resonance (NMR) structure of *Thermotoga maritima* TM0478 monomer (PDB ID:1UWD). B. Solution nuclear magnetic resonance (NMR) structure of *Homo sapiens* Fam96a monomer (PDB ID:2M5H) C. Crystal structure of domain-swapped *Homo sapiens* Fam96a major dimer (PDB ID: 3UX2). The black arrows indicate the swapped domains. D. Crystal structure of domain-swapped *Homo sapiens* Fam95a minor dimer (PDB ID: 3UX3). The zinc binding and swapped domains are indicated by the black arrows.

The only published structure of a DUF59 containing protein in bacteria is the solution nuclear magnetic resonance (NMR) structure of TM0487 (Protein Data Bank (PDB) ID :1UWD) from *Thermotoga maritima*, which showed a secondary structure characterised by an α/β topology (Fig. 1.2). Unlike the eukaryotic DUF59 containing proteins, which can exist as both monomers and dimers (Chen et al., 2013; Ouyang et al., 2013), DUF59 in bacteria exist only as monomers (Almeida et al., 2005). TM0487 has 20 structural homologues, however all these homologues display low sequence similarity with TM0487 and have divergent functions, precluding functional prediction using only this structural homology data. However, when limiting the search to structural homologues of the putative active site, a partial positive match with an Fe-S cluster containing protein from *Desulfovibrio desulfuricans* was identified

(Almeida et al., 2015). This made the case for the involvement of TM0487 and by inferred homology, DUF59 proteins in Fe-S cluster biogenesis, although the structural information alone was not enough to make a definitive conclusion (Almeida *et al.*, 2005).

1.4.2 Functional characterisation of DUF59 containing proteins in eukaryotes

1.4.2.1 Mammalian DUF59 containing proteins

Fam96a and Fam96b are the only two mammalian DUF59 containing proteins (Chen et al., 2012). They are both involved in iron homeostasis and Fe-S cluster assembly in the cytosol of mammalian cells (Chen et al., 2012; Stehling et al., 2013). The role of Fam96a in Fe-S cluster biogenesis was first investigated using yeast two hybrid assays and co-immunoprecipitation assays, which showed that *in vitro*, Fam96a can bind the CIA assembly protein (Cia1). This suggested that Fam96a could be part of the CIA system (Chen et al., 2012). Another study using co-immunoprecipitation and enzyme assays, demonstrated that Fam96a was directly involved in the maturation of Fe-S clusters for iron regulatory protein 1 (IRP1) (Stehling et al., 2013). The study also showed that Fam96b is also an Fe-S cluster maturation protein and that unlike Fam96a which has only one target (IRP1), Fam96b matured the Fe-S cluster containing proteins dihydropyrimidine dehydrogenase (DYPD), glutamyl amidotransferase (GPAT) and DNA polymerase delta (POLD) (Stehling et al., 2013). DYPD is involved in pyrimidine degradation, GPAT in the synthesis of purines and POLD in DNA synthesis (Stehling et al., 2013). These results illustrated that Fe-S clusters are required for proteins in multiple pathways, which could possibly explain why DUF59 containing proteins are associated with multiple cellular pathways.

1.4.2.2 Yeast DUF59 containing proteins

In yeast, DUF59 containing proteins have been studied in *Saccharomyces cerevisiae*. Unlike the Fam96a and Fam96b proteins which are made up entirely of the DUF59 domain (Chen et al., 2012), the *S. cerevisiae* DUF59 containing protein Nbp35 only carries the domain at its' N-terminus (Hausmann et al., 2005). The N-terminus of Nbp35 has four conserved cysteine residues suggesting a role in metal binding (Hausmann et al., 2005). Additionally, phylogenetic data also showed that the closest homologs of Nbp35 are the yeast protein Cfdp1 and the *Salmonella enterica* protein ApbC, which are both part of the CIA system. (Tsaousis et al., 2014) Nbp35 was therefore hypothesized to be involved in Fe-S cluster assembly in yeast (Hausmann et al., 2005).

Co-immunoprecipitation assays showed protein-protein interactions between Nbp35 and proteins involved in the CIA system (Hausmann et al., 2005). Iron (^{55}Fe) radio labelling experiments showed that Nbp35 binds Fe-S clusters *in vivo*, and that cells lacking Nbp35 were impaired for growth (Hausmann et al., 2005). Collectively these results implicated Nbp35 in *de novo* assembly of Fe-S clusters however, Nbp35 was only needed for the maturation of cytosolic Fe-S clusters and a lack of Nbp35 had no effect on mitochondrial Fe-S cluster assembly (Hausmann et al., 2005). These results had significant implications for the mechanisms of Fe-S cluster biogenesis in higher organisms as they introduced the idea of localised function by showing that *S. cerevisiae* had a cytosolic-system separate from the mitochondrial-iron-sulphur-assembly system (Hausmann et al., 2005).

1.4.2.3 Plant DUF59 containing proteins

In plants, two DUF59 containing proteins, HCF101 (Schwenkert et al., 2010) and AE7, (Luo et al., 2012) have been extensively studied. The AE7 protein is made up entirely of a DUF59 domain (Luo et al., 2012), while, like Nbp35, (Hausmann et al., 2005) HCF101 only carries the DUF59 domain at its N-terminus (Schwenkert et al., 2010). HCF101 was identified as an essential protein in Fe-S cluster assembly, but initially its mechanism of action was not determined (Lezhneva et al., 2004). A combination of spectrophotometry and co-immunoprecipitation experiments in a yeast model showed that HCF101 was able to bind 4Fe-4S clusters *in vivo*, suggesting that it could function as a scaffold protein on which clusters are assembled (Schwenkert et al., 2010). As a scaffold protein, HCF101 should be able to perform the dual function of binding clusters and then releasing them to target apoproteins (Schwenkert et al., 2010). *In vitro* enzyme activity assays showed that HCF101 was indeed able to transfer the bound cluster to a target apoprotein (Schwenkert et al., 2010). In combination, these results characterised HCF101 as a scaffold protein for the assembly of 4Fe-4S clusters in photosystem 1 in the plastids of *A. thaliana* (Schwenkert et al., 2010).

AE7 was also implicated in Fe-S cluster assembly in *A. thaliana* (Luo et al., 2012). Quantification of the activity of the Fe-S cluster containing enzyme aconitase in cells lacking AE7 showed that it was reduced by 75% and 55% in the cytosol and mitochondria respectively, indicating that the protein was directly involved in *de novo* Fe-S biogenesis. Additional, co-immunoprecipitation and bimolecular fluorescence assays showed that AE7 interacts with the CIA pathway proteins CIA1, NAR1 and MET18 (Luo et al., 2012). These four proteins form an AE7-CIA1-NAR1-MET18 complex and this complex facilitates the transfer of Fe-S clusters to target apoproteins in the cytosol or nucleus of the plant cells (Luo et al., 2012). The target apoproteins of AE7 are involved in DNA repair, making the protein indispensable for growth of *A. thaliana* (Luo et al., 2012).

1.4.3 Functional characterisation of bacterial DUF59 containing proteins

1.4.3.1 Bioinformatic analysis

Sequence and structural analyses of TM0487 from *T. maritima* revealed that the protein had 216 homologues, and that 96 of these homologues all had six conserved residues (D20, E22, L23, T51, T/S52 and C55) with respect to the sequence of TM0487 (Fig. 1.3) (Almeida *et al.*, 2005). Furthermore, all six residues were positioned in close proximity to each other in the folded structure at the C-terminus of the domain, suggesting they might form part of the active site (Almeida *et al.*, 2005).

<i>T. maritima</i>	-----PMSKKVTKEDVLNALKNVIDFELGLDVVSLGLVYDIQI	38
<i>S. meliloti</i>	MGLEQTQEKVNVREGIVHSAIPPEELARLSDDIIAALKTVYDPEIPADIFELGLIYKIDI	60
<i>M. tuberculosis</i>	-----MSETSAPEELLADVEEAMRDVVDPELGINVVDLGLVYGLDV	42
<i>B. anthracis</i>	-----MSQEAFENKLYANLEAVIDPELGVDIVNLGLVYDVTA	37
<i>S. aureus</i>	-----MVIDPELGDIVNLGLVYKVVV	22
	* * * : : : . . * * * : * :	
<i>T. maritima</i>	DD---QNNVKVLMTMTTPMOPLAGMILSDAEEAI-KKIEGVNNVEVELTFDPPWTPERMS	94
<i>S. meliloti</i>	ED---DRMVKIDMTLTAPGQPVAGEMPGWVENAV-GTVEGVSGVEVSMTFDPPWTPDRMS	116
<i>M. tuberculosis</i>	QDGDEGTVALIDMTLTSAAOPLTDVIEDQSRSAI-VGSGLVDDIRINVWVNPWPWPKIT	101
<i>B. anthracis</i>	DE---NNNAVITMTTSIGCPMAGQIVSDVKKVLSTNVPEVNEIEVNVVWVNPWWSKERMS	94
<i>S. aureus</i>	DD---EGVCTVDMTTLTSMGCPMGPQIIDQVKTVL-AEIPETIQDTEVNIWVSPWTKDMMS	78
	: : : * * * : * * : . . . : : * * * : : :	
<i>T. maritima</i>	PELREKFGV-----	103
<i>S. meliloti</i>	EEAQVALGWY----	126
<i>M. tuberculosis</i>	EDGREQLRALGFTV	115
<i>B. anthracis</i>	RMAKIALGIRD---	105
<i>S. aureus</i>	RYAKIALGVS----	88
	: :	

Figure 1. 3. Clustal Omega generated multiple sequence alignment of DUF59 domains in bacteria. The six conserved residues (D20, E22, L23, T51, T/S52 and C55) highlighted in red. An asterisk (*) indicates positions which have a single, fully conserved residue, a colon (:) indicates conservation between groups of strongly similar properties and a period (.) indicates conservation between groups of weakly similar properties.

Recent bioinformatic analysis of 1092 published genomes encoding *sufBC*, revealed that 761 of them also encoded a protein containing a DUF59 domain (Mashruwala *et al.*, 2016). In addition, these proteins were associated with the *suf* operon in 49% of the genomes encoding both *sufBC* and a DUF59 domain (Mashruwala *et al.*, 2016). Although these sequence homology searches were done using BLASTp (pairwise alignments), which has limited sensitivity in detecting remote homologies (Dunbrack, 2006; Mulder *et al.*, 2007; Ramakrishnan *et al.*, 2015), these findings support the predicted function of DUF59 containing proteins as putative Fe-S biogenesis proteins (Almeida *et al.*, 2005).

1.4.3.2 Biochemical and biological characterisation of bacterial SufT proteins

1.4.3.2.1 *Staphylococcus aureus* (Gram-positive bacterium)

The first DUF59 containing protein in bacteria to undergo biological and biochemical characterisation was SAUSA300_0875 from the Gram-positive facultative anaerobe *Staphylococcus aureus*. It was subsequently re-named SufT, because of its proximity to the *sufCDSUB* operon (Tsaousis et al., 2014; Mashruwala et al., 2016). To determine the function of SufT in *S. aureus*, an *S. aureus* Δ *sufT* mutant strain was generated and phenotypically characterised (Mashruwala et al., 2016). *In vitro* enzyme activity assays showed that the activity of aconitase was decreased in cells lacking SufT, and that the biggest decrease was observed under conditions of oxidative stress (Mashruwala et al., 2016). Fe-S clusters are susceptible to damage by reactive oxygen species (ROS) (Djaman et al, 2004; Imlay, 2006). Mashruwala et al. (2016a) observed that the activity of aconitase was impacted by oxidative stress so, they reasoned that the observed differences could be due to reduced activity of ROS scavenging enzymes in the mutant, or that SufT was needed either for physically shielding clusters from ROS, repairing clusters damaged by ROS or *de novo* synthesis of Fe-S clusters under oxidative stress. Using a series of enzyme activity assays, they eliminated the different options and concluded from the data generated the most probable explanation for the observed phenotype was that SufT was needed for *de novo* synthesis of Fe-S clusters during oxidative stress.

In addition, the activity of aconitase in the Δ *sufT* mutant was comparable to that of an *S. aureus* mutant strain lacking the Nfu protein (Mashruwala et al., 2016a). Nfu is an Fe-S biogenesis protein that is directly involved in the maturation of aconitase (Mashruwala et al., 2015) and taken together, these results suggested that SufT may be functioning as a maturation protein like Nfu (Mashruwala et al., 2016a). Based on previous results (Mashruwala et al., 2016a), the researchers hypothesized that *in vivo*, there are multiple maturation proteins performing the same function and that the choice of protein is determined by demand for the target apoprotein (Fig. 1.4) (Mashruwala et al., 2016b). Maturation proteins transfer assembled Fe-S clusters from the scaffold proteins to target apoproteins and *S. aureus* has three maturation proteins Nfu (Mashruwala et al., 2015), SufA and SufT (Mashruwala et al., 2016a). They subsequently attempted to prove the model of the differential utilisation of Fe-S maturation proteins in *S. aureus* using phenotype characterisation in a Δ *sufT* mutant experiencing different cellular demands for lipoic acid (Mashruwala et al., 2016b). Lipoic acid is synthesized by LipA, an Fe-S containing protein They hypothesized that when the cellular demand for lipoic acid is low, either of the three proteins can be used for LipA maturation but,

under conditions of high lipoxic acid demand, SufT is indispensable for growth because it is preferentially used for LipA maturation (Fig. 1.4) (Mashruwala et al., 2016b).

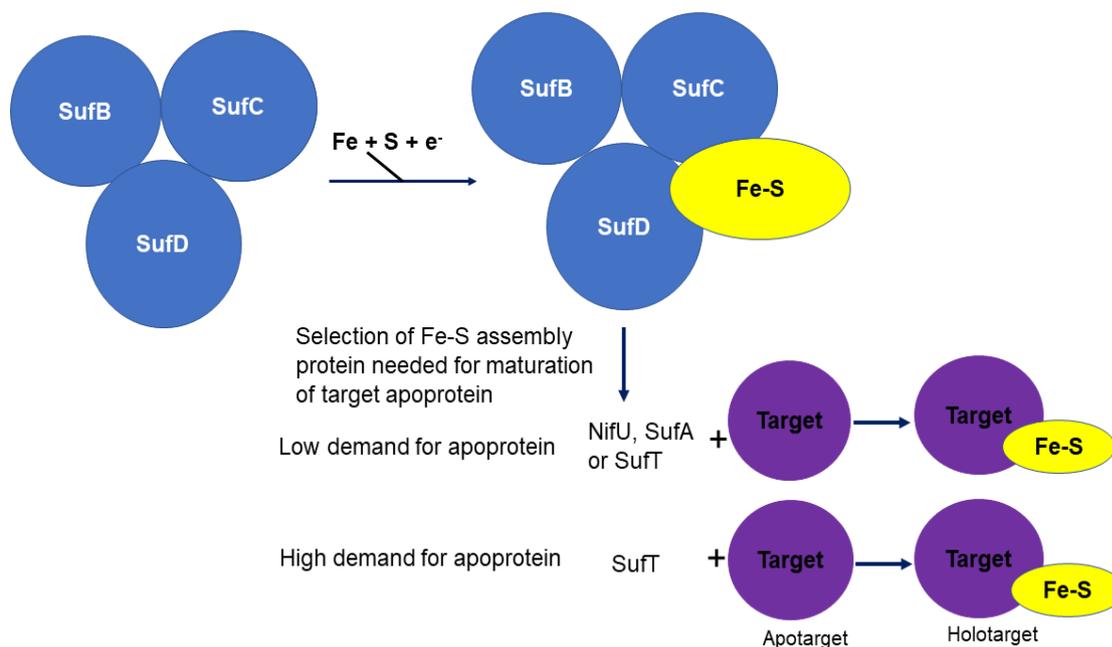


Figure 1. 4. Proposed model of differential utilisation of Fe-S maturation proteins in *S. aureus*. This figure is adapted from Mashruwala et al., (2016b).

1.4.3.2.2 *Sinorhizobium meliloti* (Gram-negative bacterium)

A mutant generation and phenotypic characterisation strategy was also used to determine the function of the DUF59 containing protein SMc00302 in *S. meliloti* (Sasaki et al., 2016). *S. meliloti* is a Gram-negative bacterium that is involved in symbiotic nitrogen-fixation with some leguminous plants (Sasaki et al., 2016). Like the *S. aureus* *sufT*, SMc00302 is part of an operon encoding the *suf* system, *sufBCDS*- SMc00302-*sufA*, and was therefore re-named SufT. Phenotypic studies showed that the Δ *sufT* mutants grew slower than the wild-type, and that they were more sensitive to environmental changes such as iron depletion, high temperatures and an acidic pH (Sasaki et al., 2016). Quantification of enzyme activity showed that loss of SufT resulted in a statistically significant loss of activity in some Fe-S cluster dependent enzymes in *S. meliloti* suggesting a role for SufT in Fe-S biogenesis in the bacterium.

The researchers also hypothesized that during symbiosis, the action of the iron uptake repressor RirA creates an intracellular iron-limiting environment in *S. meliloti*. Since symbiosis increases the cellular demand for Fe-S proteins, SufT may be required for the formation of Fe-S clusters under these intracellular iron-limiting conditions created by RirA during symbiosis (Sasaki et al., 2016). To test this hypothesis, a Δ *sufT* Δ *rirA* double mutant was generated and

the double mutants displayed wild-type growth phenotypes during symbiosis while the Δ *sufT* single mutant had impaired growth during symbiosis (Sakaki et al., 2016). Loss of *RirA* increases iron uptake in the cell, thereby creating an iron replete intracellular environment. *SufT* is dispensable for growth these conditions, suggesting a putative role for *SufT* in *de novo* synthesis of Fe-S clusters under iron-limiting conditions (Sasaki et al., 2016). However, the mechanism of action of *SufT* in *S. meliloti* is still unknown.

1.4.3.3 DUF59 (*SufT*) containing proteins in mycobacteria

In mycobacteria, the DUF59 containing protein, encoded by *Rv1466*, has also been re-named *SufT* because it is also part of the *suf* operon, (Fig. 1.5) and its ability to complement loss of *sufT* in *S. aureus* (Mashruwala et al., 2016a). The *suf* operon in mycobacteria is highly conserved across the species (Fig. 1.5) and this synteny is a good predictor of protein function on the principle that genes in an operon have evolved together and are kept together in the genome because they perform similar functions (Tamames, 2001; Guerrero et al., 2005)

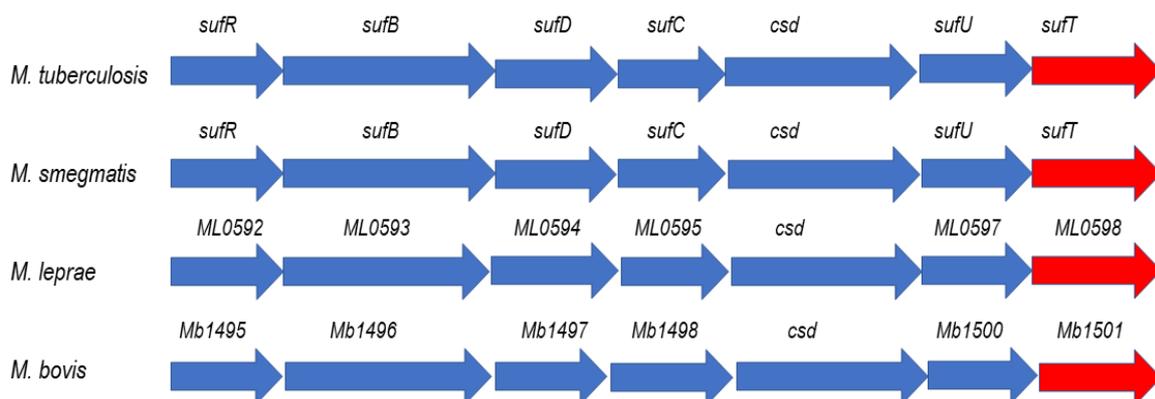


Figure 1. 5. Schematic of the gene order in the *suf* operon of mycobacteria. The genes in red represent putative *sufT* proteins.

Identifying and functionally characterising genes that are essential for growth is key to understanding the physiology of an organism because essential genes are part of fundamental cellular processes (Cole et al., 1998; Sassetti et al., 2003; Goodacre et al., 2013). Transposon mutagenesis is the most widely applied strategy for large-scale essentiality screening (Griffin et al., 2011). An insertion in an essential gene is lethal to the bacteria and therefore these mutants will not grow in subsequent growth cycles. (Craig, 1997; Sassetti et al., 2001). Transposon site hybridisation (TraSH) enables the genome-wide identification of essential genes by comparing the abundance of a mutant before (input pool) and after (output pool) selection using a microarray (Sassetti et al., 2001; Sassetti et al., 2003) The resolution of the approach was later improved by identification of transposon insertions by next generation

sequencing (Griffin et al., 2011). The initial TraSH screen identified *sufT* as an essential gene (Sasseti et al., 2003) while the subsequent screen using next generation sequencing identified it as non-essential gene (Griffin et al., 2011). Screening by next generation sequencing is more accurate than TraSH because TraSH is more susceptible to false positive results caused by downstream polar effects and lacks sensitivity to detect low mutant abundance (Sasseti et al., 2003; de Wet et al., 2018).

In a recent study, a new large-scale essentiality screening method using CRISPR interference (CRISPRi) was used to screen for essentiality of *M. tuberculosis* homologues in *M. smegmatis* (de Wet et al., 2018). CRISPRi builds on transposon mutagenesis and has great applicability in studying the essentiality of genes in operons (de Wet et al., 2018). There are however significant differences between the two methods. Unlike transposon mutagenesis which inserts a transposon into a gene to disrupt it, CRISPRi uses single guide RNAs (sgRNAs) that allows an inactive cas9 endonuclease to bind a specific transcript and stop transcription of the gene. A limitation of the technique is that it also prevents transcription of downstream operonic genes.

The genes in the *suf* operon of *M. tuberculosis* are transcribed (*sufR-sufT*) (Fig. 1.5) as an operon (Huet et al., 2005). Additionally, a *M. smegmatis* mutant in which *sufB* was interrupted with a hygromycin cassette could not be generated suggesting that this gene is essential and supporting the essentiality predictions for the operon in the TraSH screen (Huet et al., 2005). A recent study showed that the first gene in the operon, *sufR*, is a transcriptional repressor of the operon and that the *M. tuberculosis* Δ *sufR* mutant displays a growth defect under standard culture conditions. The mutant had no growth defect under iron limitation (Willemse et al., 2018), and this is agreement with an earlier study, which showed that all the genes of the operon except *sufR*, were upregulated as a response to iron starvation (Rodriguez et al., 2002). Iron starvation in *M. tuberculosis* leads to extensive repression of gene expression, with the notable exception of the iron-acquisition pathways, selected Fe-S cluster containing enzymes and the *suf* operon. Since *M. tuberculosis* is hypothesized to experience iron limitation in the host, this selective gene induction points to the importance of Fe-S cluster biogenesis during infection (Rodriguez et al., 2002).

Transcriptional studies have also shown that all the genes in the *suf* operon of *M. tuberculosis* are immediately upregulated in response to diethylenetriamine/nitric oxide adduct (DETA/NO) induced nitric oxide (NO) stress (Cortes et al., 2017; Voskuil et al., 2011) and hydrogen peroxide (H₂O₂) induced oxidative stress (Voskuil et al., 2011). Transcriptional upregulation of the operon was shown to be induced by as little as 0.05 mM H₂O₂ or DETA/NO and that *sufT* expression was upregulated 7 and 5-fold in 10mM H₂O₂ and DETA/NO respectively (Voskuil

et al., 2011). This result is not surprising because oxidative stress damages Fe-S clusters, so early transcription of genes involved in Fe-S biogenesis is likely part of the cells' adaptive response (Lamichhane, 2011; Voskuil et al., 2011). In support of this hypothesis, there was also upregulation of iron acquisition genes and repression of iron storage proteins suggesting that there is an intracellular need to acquire and use iron, presumably for the repair of damaged Fe-S clusters (Voskuil et al., 2011). This response seems counter intuitive because iron has been shown to increase oxidative stress through the creation of super hydroxide radicals in the Fenton reaction (Winterbourne, 1995), but in *M. tuberculosis* that seems to be superseded by the need to repair damaged Fe-S clusters (Voskuil et al., 2011).

Proteomic analysis under NO stress revealed that the immediate transcriptional upregulation did not result in immediate translation into proteins, and a delayed induction of all the *suf* operon encoded proteins was observed (Cortes et al., 2017). Instead, the targeted degradation of proteins containing Fe-S clusters occurred as an immediate response to NO stress (Cortes et al., 2017). Taken together these results suggest that in *M. tuberculosis* proteins of the *suf* operon, including SufT, are needed during NO stress, but that they are probably not involved in the initial response. The delayed increase in these proteins may represent a recovery strategy that involves either *de novo* synthesis or repair of damaged clusters (Cortes et al., 2017).

1.5 Conclusion

In both bacteria and eukaryotes, the involvement of DUF59 containing proteins in Fe-S cluster biogenesis has been repeatedly shown. In the eukaryote models, the functions of the DUF59 containing proteins in *de novo* Fe-S biogenesis has been confirmed experimentally both *in vitro* and *in vivo* in mammals, yeast and plants (Hausmann et al., 2005; Schwenkert et al., 2010; Luo et al., 2012; Stehling et al., 2012). In contrast, in bacteria there is still a dearth of knowledge regarding the role of *sufT* in Fe-S biogenesis. While these studies have postulated a role for *sufT* in *de novo* synthesis of Fe-S clusters (Almeida et al., 2005; Mashruwala et al., 2016; Sasaki et al., 2016; Mashruwala et al., 2017), the mechanism of action is still unknown but, the studies have laid a foundation for examining the function of *sufT* in other bacteria, including mycobacteria.

1.6 Study rationale

M. tuberculosis is one of the leading causes of death globally and with drug resistant TB on the rise (WHO, 2017) there is an urgent need to find new anti-TB drugs and drug targets. Increasing our understanding of the physiology of *M. tuberculosis* can aid in elucidating novel essential pathways, which can be used as new drug targets (Hensel & Holden, 1996; Welch,

2015). One such pathway is the Fe-S cluster biogenesis pathway, which is required for the synthesis of Fe-S cluster *in vitro* (Huet et al., 2005). The genome of *M. tuberculosis* encodes a single SufT protein and it is highly conserved in all mycobacteria (Mashruwala et al. 2016). To date, there are only three published studies on the functional characterisation of SufT in bacteria (Mashruwala et al., 2017; Mashruwala et al., 2016; Sasaki et al., 2016) but none in mycobacteria.

Evolutionary data suggests that *sufT* was recruited to the operon suggesting that it functions as part of the operon. Studies in other bacteria (Mashruwala et al., 2016a; Mashruwala et al., 2016b; Sasaki et al., 2016) have already alluded to the involvement of *sufT* in Fe-S biogenesis during conditions of low iron and/ or high Fe-S demand. During infection, the host uses iron starvation and oxidative stress to kill *M. tuberculosis* but in turn, the bacteria upregulates transcription of the *suf* operon as part of its' adaptive response to the oxidative stress (Voskuil et al, 2011) and iron limitation (Rodriguez et al., 2002) enabling it to survive and proliferate.

Mycobacteria use the *suf* operon as the primary Fe-S biogenesis pathway and the *suf* operon is essential to the survival of *M. tuberculosis* (Huet et al., 2005), therefore understanding the role of each protein in the operon is crucial to fully understanding how it contributes to pathogenesis. In addition, because the host cells do not use the *suf* system for Fe-S biogenesis (Lill & Mühlenhoff, 2006), the *suf* operon of *M. tuberculosis* presents a potential novel drug target.

1.7 Approach

This study utilised targeted gene deletion and phenotypic characterisation of the resulting mutant to investigate the role of SufT in the physiology of mycobacteria. *Mycobacterium smegmatis*, a non-pathogenic, fast-growing mycobacteria, which is closely related to *M. tuberculosis*, was used as a model organism. An *M. smegmatis* Δ *sufT* mutant strain harbouring an in-frame, unmarked deletion in the *sufT* gene was generated by homologous recombination. The wild-type, mutant and genetically complement mutant were subsequently subjected to phenotypic studies to evaluate effect of the loss of SufT on the bacteria.

1.8 Hypothesis

The SufT protein in mycobacteria is involved in *de novo* synthesis of Fe-S clusters under conditions of cell stress and/or high Fe-S cluster demand.

1.9 Aim

This study aimed to investigate the role of SufT in Fe-S cluster synthesis and physiology of mycobacteria, using *M. smegmatis* as a model organism.

1.9.1 Objectives:

1. Generate an *M. smegmatis* Δ sufT mutant and genetically complemented strain.
2. Evaluate the planktonic growth of the *M. smegmatis* Δ sufT mutant under standard culture conditions.
3. Evaluate the activity of Fe-S cluster containing enzymes in the *M. smegmatis* Δ sufT mutant.
4. Evaluate the impact of oxidative stress on the *M. smegmatis* Δ sufT mutant.
5. Evaluate the drug sensitivity of the *M. smegmatis* Δ sufT mutant.
6. Evaluate the planktonic growth of the *M. smegmatis* Δ sufT mutant under iron limiting conditions
7. Evaluate the ability of the *M. smegmatis* Δ sufT mutant to form biofilms under standard culture and iron limiting conditions.

Chapter 2: Materials and Methods

2.1 Bacterial strains and culture conditions

2.1.1 Standard culture

Escherichia coli strains were cultured aerobically in lysogeny broth (LB) at 37°C in a shaking incubator (180rpm) or on lysogeny broth agar (LA). When necessary, ampicillin (Amp) 100 µg/ml, kanamycin (Km) 50 µg/ml, tetracycline (Tet) 10 µg/ml, hygromycin (Hyg) 150 µg/ml, sucrose (5%) and 5-bromo-4 chloro-3 indolyl β-D-galactosidase (X-gal) 40 µg/ml were added.

Mycobacterium smegmatis strains were cultured aerobically in Middlebrook 7H9 media (Difco) supplemented with 0.085% NaCl, 0.2% glucose, 0.2% glycerol and 0.05% Tween 80 (7H9 GST) at 37°C in a shaking incubator (180rpm) or on Middlebrook 7H10 media supplemented with 0.085% NaCl, 0.2% glucose and 0.5% glycerol (7H10 GS). Where necessary, Amp (100 µg/ml), Hyg (150 µg/ml), Km (50 µg/ml), sucrose (2%) and X-gal (40 µg/ml) were added.

M. smegmatis strains were cultured for biofilm formation in Sauton's media (pH 7.4) containing 0.05% KH₂PO₄, 0.05% MgSO₄, 0.4% L-asparagine, 0.2% citric acid, 0.005% ferric ammonium citrate, 6% glycerol and 0.1% zinc sulphate (ZnSO₄). Where necessary, Hyg (150 µg/ml) was added.

2.1.2. Iron limitation

M. smegmatis strains were cultured in a Chelex® 100 resin treated mineral defined media (MM) prepared as follows: A solution (pH 6.8) of 0.5% L-asparagine, 0.5% potassium dihydrogen phosphate (KH₂PO₄), 0.5% bovine serum albumin (BSA) and 0.2% glucose was treated with 5% (w/v) Chelex® 100 resin (Bio-Rad) for 24 hours at 4°C, and the solution filter sterilised and supplemented with metals 0.00001% manganese sulphate (MnSO₄), 0.0001% zinc sulphate (ZnSO₄) and 0.005% magnesium sulphate (MgSO₄) to generate mineral defined media (MM). Where necessary, Hyg (150 µg/ml) was added. The MM was stored at 4°C.

Bacterial strain stocks were frozen at -80°C for long-term storage. *E. coli* strains were stored in 66% glycerol and *M. smegmatis* strains were stored in the culture media. Bacterial strains used in this study are listed in Table 2.1.

Table 2. 1. List of bacterial strains used and generated in this study.

Strain	Strain description	Source/Reference
<i>Escherichia coli</i>		
XL1 Blue	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' proABlacI ^q ZΔM15 Tn10]; Tet ^r	Stratagene
<i>Mycobacterium smegmatis</i>		
mc ² 155-1	<i>ept-1</i> (efficient plasmid transformation) mutant of mc ² 6.	Snapper et al., 1990
Δ <i>sufT</i>	Derivative of mc ² 155 with a deletion in <i>sufT</i> from codons 12 to 93.	This study
Δ <i>sufT</i> (pSE3127)	Derivative of Δ <i>sufT</i> carrying a copy of the <i>sufT</i> gene on the vector pSE3127 (Table 2.2), Hyg ^r .	This study
mc ² 155-2	<i>ept-1</i> (efficient plasmid transformation) mutant of mc ² 6 (a different laboratory stock to mc ² 155-1). Used by Fang <i>et al.</i> , 2017.	Snapper et al., 1990
Δ <i>MycP</i> _{3ms}	Derivative of mc ² 155-2 with a deletion in <i>mycP</i> ₃ (encoding <i>mycosin 3</i>)	Fang et al., 2017
Δ <i>MycP</i> _{3ms} :: <i>pr1mycP</i> _{3ms}	Derivative of Δ <i>MycP</i> _{3ms} carrying a pMV306- <i>pr1-mycP</i> ₃ vector stably integrated into the mycobacterial <i>attB</i> site, Hyg ^r	Fang et al., 2017

2.1.3 Assessment of bacterial growth

Cell density was assessed by measuring the optical density at 600 nm (OD₆₀₀) using the Ultrospec 4051 (LKB Biochrom) UV-visible spectrophotometer in liquid cultures and viability assessed by enumerating colony forming units (CFUs) on solid media. Biofilm formation was assessed by macroscopic observation, and the biofilm biomass was quantified by calculating total protein in whole cell lysates (WCLs) using the Bradford assay (Bradford, 1976).

2.1.3.1 Preparation of whole cell lysates

2.1.3.1.1 Standard culture biofilms

WCLs were prepared from the biofilm biomass as previously described (Lundblad & Struhl, 2008) with some modifications. Briefly, the liquid below the biofilm biomass was aspirated and the biomass resuspended in the well in 1 ml 0.01% Triton X-100, mixed using a Pasteur pipette and incubated at room temperature for 30 mins. After 30 mins the suspension was mixed again and then transferred to a 2 ml screw cap tube with 0.1 mm glass beads. Samples were lysed by bead beating 4 times for 20s at 4.5 m/s with 2-minute rest periods on ice between the lysis steps in a Thermo Savant FastPrep FP120 Cell Homogenizer. The cells were centrifuged at 24 100 x g for 15 mins at 4°C to remove the cellular debris. The total protein in the supernatant was quantified using a Bradford assay.

2.1.3.1.2 Iron limitation biofilms

The WCLs were prepared as previously described for standard culture biofilms with one modification. The liquid below the biofilm biomass was not aspirated instead, Triton X-100 was added directly to each well at a final concentration of 0,01% to solubilise the biofilm.

2.1.3.2 Bradford Assay

A standard curve was generated using bovine serum albumin (BSA) which was diluted in 0.01% Triton X-100 to give final concentrations of 0, 0.05, 0.1, 0.2, 0.3 and 0.4 mg/ml. The whole cell lysates were also diluted 1:5, 1:10 and 1:20 to determine the dilution that would fall within the linear range of the standard curve, and 1:10 chosen as the optimal dilution. The Bradford assays were then done as previously described (Bradford, 1976) with some modifications. The assay was done in a 96 well plate. Bradford reagent (Bio-Rad) (200 µl) was added to whole cell lysates or BSA standards (10 µl) in triplicate and incubated in the dark for 15 mins. The absorbance was read at 595 nm using the KC4 program on a Bio-Tek Synergy HT plate reader. The gradient of the standard curve was used to determine the protein concentration(mg/ml) in each sample using the following equation:

$$\frac{\text{Average Absorbance (595 nm)} - \text{Average Blank Absorbance}}{\text{Gradient of standard curve}} \times \text{Dilution factor}$$

2.2 DNA isolation

2.2.1 Plasmid DNA isolation and purification from *E. coli*

The Wizard® Plus SV Minipreps DNA Purification System (Promega) and the NucleoBond® PC100 Plasmid DNA Purification kit (Macherey-Nagel) were used according to the manufacturer's instructions for small-scale and large-scale DNA isolation and purification respectively. Plasmid DNA was quantified using the NanoDrop ND-1000 Spectrophotometer and stored at -20°C. Bacterial plasmids used and generated in this study are listed in Table 2.2.

Table 2. 2. List of bacterial plasmids used and generated in this study.

Name	Description	Source
pJET1.2	Positive selection linearized blunt end cloning vector, with lethal restriction enzyme insert that is disrupted by ligation of a DNA insert into the cloning site, Amp ^r .	CloneJet
pJET3127US	Derivative of pJET1.2 containing the 3127US DNA fragment (1953 bp product including 1920 bp upstream of <i>sufT</i> and 33 bp of the 5' end of <i>sufT</i>) cloned into the blunt end cloning position.	This study
pJET3127DS	Derivative of pJET1.2 containing the 3127DS DNA fragment (2029 bp product including 64 bp of the 3' end of <i>sufT</i> and 1965 bp downstream of <i>sufT</i>) cloned into the blunt end cloning position.	This study
pJET3127	Derivative of pJET1.2 containing the functional <i>sufT</i> gene including 26 bp upstream and 17bp downstream of <i>sufT</i> cloned in the blunt end cloning position	This study

Table 2.2 continued on the next page

Table 2.2 continued

Name	Description	Source
p2NIL	Vector which contains origin of replication for <i>E. coli</i> but not mycobacteria that is used for homologous recombination, Km ^r .	Parish & Stoker, 2000
p2NIL Δ sufT	Derivative of p2NIL recombination vector containing the 3127US and 3127DS DNA fragments directionally cloned between KpnI and HindIII, Km ^r .	This study
pGOAL17	Plasmid with the <i>lacZ</i> and <i>sacB</i> genes as a <i>PacI</i> selection cassette, Amp ^r .	Parish & Stoker, 2000
p2NIL Δ sufTpGOAL17	Derivative of P2NIL Δ sufT suicide delivery vector containing <i>lacZ</i> and <i>sacB</i> genes from pGOAL17 allowing selection and counter selection of mutants, Km ^r .	This study
pSE100	Episomal vector for the expression of genes from a PmycTetO promoter, high copy number in <i>E. coli</i> and low in mycobacteria, Hyg ^r (Addgene: Plasmid #17972).	Alland et al., 2000
pSE3127	Derivative of pSE100 with the intact <i>M. smegmatis</i> functional <i>sufT</i> gene including 26 bp upstream of <i>sufT</i> and 17 bp downstream of <i>sufT</i> , directionally cloned between BamHI and HindIII and expressed from the PmycTetO promoter, Hyg ^r .	This study

2.2.2 Genomic DNA extraction from *M. smegmatis*

2.2.2.1 Large scale genomic DNA extraction

Genomic DNA from *M. smegmatis* was extracted as previously described (Wilson, 2001) with some modifications. Bacteria were scraped off a 7H10 GS plate, dispensed into 1 ml of 7H9 GST media and vortexed to obtain a homogenous suspension. The suspension was then centrifuged at 13 900 x *g* for 2 minutes (mins) to pellet the cells. The supernatant was discarded, and cells resuspended in 500 μ l of Tris(hydroxymethyl)aminomethane-Ethylenediamine tetra-acetic acid (TE) buffer pH 8.0 (10mM Tris-Cl, 1mM EDTA) and vortexed to obtain a homogenous suspension. Lysozyme (10 mg/ml) was added to a final concentration

of 130.4 µg/ml, the suspension vortexed gently and incubated at 37°C overnight. RNase A (10 mg/ml) was added to a final concentration of 80 µg/ml and the suspension incubated at 37°C for 2 hours (hrs). Sodium dodecyl sulphate (SDS) (10%) and proteinase K (10 mg/ml) were added to a final concentration of 71.3 µg/ml proteinase K in 1% SDS and incubated at 65°C for 10 mins. Sodium chloride (5M) and cetyltrimethylammonium bromide (CTAB)/sodium chloride solution (NaCl) (10% CTAB in 0.7M NaCl) were added to a final concentration of 1% CTAB in 0.7M NaCl and the suspension vortexed and incubated at 65° for 10 mins. After incubation, 750µl chloroform-isoamyl alcohol (24:1 v/v) was added and mixed gently by inversion. The biphasic solution was centrifuged at 13 500 x g for 5 mins at 4°C to separate the aqueous and organic phases. The upper aqueous phase containing the DNA was carefully collected into clean Eppendorf tubes and the DNA precipitated by adding 450 µl of isopropanol and incubating at -20° for 1 hour. The DNA was pelleted by centrifugation at 13 500 x g for 15 mins. The DNA was washed in 70% ethanol, air dried at 42°C and resuspended in 100 µl of TE buffer at 42°C overnight. The DNA was quantified using the NanoDrop ND-1000 Spectrophotometer and the quality was assessed on a 1% agarose gel to confirm that the DNA was not degraded.

2.2.2.2 Small scale crude genomic DNA extraction

Single *M. smegmatis* colonies were resuspended in 50 µl phosphate buffered saline in 0.05% Tween 80 (PBST) and boiled at 95°C for 15 mins. Following boiling, the solution was centrifuged at 13 900 x g for 2 mins to pellet the cell debris. The DNA containing supernatant was collected in sterile Eppendorf tubes and used for PCR amplification.

2.3 Cloning

E. coli XL1 Blue was used for all cloning experiments. DNA manipulations were performed according to standard protocols (Sambrook et al., 1989; Sambrook and Russel, 2001).

2.3.1 Polymerase Chain Reaction (PCR)

Primers were designed using Primer3 (<http://bioinfo.ut.ee/primer3-0.4.0/>) software. All the primers used in this study are listed in Table 2.3. Optimal PCR parameters were determined using Faststart™ Taq DNA polymerase (Roche Diagnostics) in a 25 µl PCR mix containing; 1x PCR buffer with 2 mM Magnesium chloride (MgCl₂), 1x GC rich solution, 200 µM deoxyribonucleotide triphosphates (dNTPs), 0.5 µM primers (Table 2.3), 1 U Faststart™ Taq DNA polymerase and 50 ng template DNA. The following cycling conditions were used: denaturation (94°C) for 5 mins, 35 cycles of denaturation (94°C) for 30 seconds (s), annealing (60°C) for 30 s and elongation (72°C) for 2 mins, with a final elongation step (72°C) for 7 mins.

Gene regions were amplified from genomic DNA using the Phusion® High-Fidelity DNA polymerase (Thermo Fisher Scientific). The Phusion® polymerase PCR reactions were set up in a 50 µl PCR mix containing; 1x Phusion® High-Fidelity buffer, 1x GC rich solution, 200 µM dNTPs, 0.5 µM primers (Table 2.3), 1 U Phusion® High-Fidelity DNA polymerase and 50 ng template DNA. The following cycling conditions were used: denaturation (98°C) for 1 min, 35 cycles of denaturation (98°C) for 15 s, annealing (60°C) for 30 s and elongation (72°C) for 2 mins, with a final elongation step (72°C) for 7 mins.

Table 2. 3. List of primers used in this study

Gene/ Region	Primer name	Sequence (5'-3')	Description	Mutation
<i>sufT</i> upstream deletion	MSMEG_3127 USF	<u>AAGCTT</u> CTGCT GATCACGCACT ACAC(KpnI)	1953 bp product including 1920 bp upstream of <i>sufT</i> and 33 bp of the 5' end of <i>sufT</i> (Fig. 2.1)	The <i>sufT</i> upstream region deletion and downstream region deletions are used in combination to delete 82 of the 115 amino acids of <i>sufT</i> (codon 12 to codon 93)
	MSMEG_3127 USR	<u>GAATTC</u> GCATG GCCTCTTCGAT GTC(EcoRI)		
<i>sufT</i> downstream deletion	MSMEG_3127 DSF	<u>GAATTC</u> CGGAC AAGATCACCGA CGAC(EcoRI)	2029 bp product including 64 bp of the 3' end of <i>sufT</i> and 1965 bp downstream of <i>sufT</i> (Fig. 2.1)	
	MSMEG_3127 DSR	<u>GGTACC</u> TTTAC GTGGCGAACAT CAGA (HindIII)		

Table 2.3 continued on the next page

Table 2.3 continued

Gene/ Region	Primer name	Sequence (5'-3')	Description	Mutation
<i>sufT</i>	3127 Comp FWD	<u>GGATCC</u> GGTGAGAGG GAGGACAACC(BamHI)	386 bp product including 26 bp upstream of <i>sufT</i> 17 bp down stream of <i>sufT</i>	N/A
	3127 Comp RVS	<u>AAGCTT</u> CCGCCCGACT GATCAGAC(HindIII)		
Sequencing primers				
N/A	pJET1.2 FWD	CGACTCACTATAGGGA GAGCGGC	pJET1.2 forward sequencing primer	N/A
N/A	pJET1.2 RVS	AAGAACATCGATTTTC CATGGCAG	pJET1.2 reverse sequencing primer	N/A
<i>sufT</i>	3127 Internal RVS	CGAGGTCAACGTCAT GTCG	<i>sufT</i> reverse internal sequencing primer	N/A
<i>sufT</i>	USintFWD	CGTTCTCGGGGCACA AGA	<i>sufT</i> upstream region forward internal sequencing primer	N/A
<i>sufT</i>	USintRVS	GGAAAACCGCGGATC CCC	<i>sufT</i> upstream region reverse internal sequencing primer	N/A
<i>sufT</i>	DSintFWD	TCGAACAGGGCGGTG TAG	<i>sufT</i> downstream region forward internal sequencing primer	N/A
<i>sufT</i>	DSintRVS	ATCAAGCACCCAGATCG ACAC	<i>sufT</i> downstream region reverse internal sequencing primer	N/A

2.3.2 DNA sequencing

Sanger sequencing of all PCR products was done at the Central Analytical Facilities-DNA Sequencing Unit at Stellenbosch University. Sequencing data was analysed using Sequencher DNA Sequence Analysis Software (Gene Codes). Sequencing primers used in this study are listed in Table 2.3.

2.3.3 Agarose gel electrophoresis

Agarose gel electrophoresis was used to separate and visualise DNA fragments. 1-2% agarose gels were prepared in 1x Tris-acetate-EDTA (TAE) buffer (40mM Tris, 20 mM acetic acid, 1 mM EDTA) containing 0.5 µg/ml ethidium bromide. 1% gels were used to separate intermediate to high molecular weight fragments and 2% gels were used for low molecular weight fragments (< 1kb). Electrophoresis was done at 80-100 volts and the DNA visualised using a gel fluorescence (UV-light) imaging system (Bio-Rad). Appropriate DNA molecular weight markers (Roche Biochemicals) were used to estimate DNA fragment sizes.

2.3.4 DNA fragment extraction and purification from agarose gels

When required, DNA fragments were excised from the agarose gel after electrophoresis under blue light using the Dark Reader™ transilluminator DR-88M (Clare Chemical Research) and purified using the Wizard® SV gel and PCR Clean-Up system (Promega) following the manufacturer's instructions. The DNA was quantified using the NanoDrop ND-1000 Spectrophotometer. All DNA fragments that were excised from agarose gels were not visualised under UV light as this could potentially damage the DNA and introduce spontaneous mutations.

2.3.5 Restriction endonuclease digestion

Restriction endonucleases and their buffers were obtained from Roche Diagnostics and New England Biolabs and the restriction enzyme digestions were set up according to the manufacturer's instructions. Plasmid DNA (0.5-1 µg) was digested in a total volume of 10-20 µl for 1 hour and genomic DNA (2-4 µg) was digested overnight in a total volume of 25-50 µl. All digests were incubated at 37°C and heat inactivated at 65°C for 15 mins. The DNA fragments were separated by gel electrophoresis (section 2.3.4) and purified by gel extraction (section 2.3.4).

2.3.6 Ligation reactions

Digested purified plasmid and insert DNA were ligated using T4 DNA ligase (Roche Diagnostics) according to the manufacturer's instructions. Ligation reactions were incubated at room temperature for 3 hrs or at 4°C overnight and inactivated at 65°C for 10 mins. For optimal transformation efficiency, a 1:3 vector to insert ratio was used and calculated as follows:

$$\text{Concentration of insert DNA (ng)} = \frac{\text{size of insert (bp)} \times \text{concentration of vector DNA}}{\text{size of vector (bp)}}$$

The ligation reactions were subsequently transformed into chemically competent *E. coli* XL1 Blue cells.

2.4 Transformation of *E. coli* XL1 Blue cells

2.4.1 Preparation of *E. coli* XL1 Blue chemically competent cells

A 500 ml culture of *E. coli* XL1 Blue was grown to an OD₆₀₀ of approximately 0.5 in LB with Tet (10 µg/ml). The culture was cooled on ice for 15 mins then harvested by centrifugation at 3 220 x *g* for 10 mins at 4°C. The supernatant was poured off and the bacterial pellet resuspended in 60 ml TfbI (100mM rubidium chloride (RbCl), 50 mM manganese chloride, 30 mM potassium acetate, 10 mM calcium chloride (CaCl₂) and 15% glycerol) then chilled on ice for 15 mins. The cells were harvested by centrifuging at 3 220 x *g* for 5 mins at 4°C. The supernatant was poured off and the cells were resuspended in 6 ml TfbII (10 mM 3-(N-morpholino) propanesulfonic acid (MOPS), 10 mM RbCl, 75 mM CaCl₂ and 15% glycerol). Aliquots (150 µl) were made, flash frozen in liquid nitrogen and stored at -80°C.

2.4.2 Transformations

Chemically competent *E. coli* XL1 Blue cells were thawed on ice and 50 µl of cells aliquoted per reaction. A volume of 2 µl of a ligation reaction was added to the cells and incubated on ice for 5 mins. A “no-DNA” control with *E. coli* XL1 Blue cells was included to assess spontaneous drug resistance of the competent cells and a positive control using an uncut vector to assess transformation efficiency. After incubation, the transformation mix was heat shocked for 90 s at 45°C to allow DNA uptake. The mixture was immediately placed on ice for 5 mins. LB (800 µl) was added to each transformation mix and incubated at 37°C for 1 hour to facilitate cell recovery. Following recovery, the cells were spread on LA plates containing the appropriate antibiotics, incubated at 37°C overnight and plates were assessed for growth of bacterial colonies.

2.5 Transformation of *M. smegmatis* by electroporation

2.5.1 Preparation of electrocompetent *M. smegmatis* cells

Electrocompetent *M. smegmatis* cells were prepared as previously described (Goude & Parish, 2008) with some modifications. Briefly, a preculture of *M. smegmatis* was grown in 7H9 GST media from a freezer stock. The preculture was used to inoculate 100 ml 7H9 GST (to a starting OD₆₀₀ of 0.005) and the cells were grown at 37°C with shaking to an OD₆₀₀ of

approximately 1. The culture was chilled on ice for 15 mins and harvested by centrifugation at $3\ 220 \times g$ for 10 mins at 4°C. Cells were washed and resuspended in 100 ml 10% glycerol containing 0.05% Tween 80. The wash step was repeated by resuspending the cells in decreasing volumes of 10 % glycerol containing 0.05% Tween 80: 50 ml (2nd wash), 20 ml (3rd wash) and 10 ml (final wash). After the final wash, the cells were resuspended in the minimum volume of 10% glycerol (containing 0.05% Tween 80) required to do electroporation.

2.5.2 Electroporation of *M. smegmatis*

Chilled electrocompetent *M. smegmatis* cells (400 µl) were placed in a 0.2 cm Gene Pulser electroporation cuvette (Bio-Rad) and mixed with 5 µg of purified plasmid DNA (knockout constructs) or 0.5-1 µg purified plasmid DNA (replicating or integrating plasmids). The electroporation was done in the Gene Pulser Xcell™ (Bio-Rad) with the following settings: 2.5 kV, 25 µF and 1000 Ω. Following each electroporation, 7H9 GST media (800 µl) was added to each cuvette and incubated for 3 hours at 37°C to allow the cells to recover. Following recovery, the cells were spread on 7H10 GS plates containing the appropriate antibiotics and incubated at 37°C for 4-5 days. After incubation, the plates were assessed for growth of bacterial colonies. A no-DNA negative control with only *M. smegmatis* cells was included to assess spontaneous drug resistance of the competent cells. A positive control using pSE100 (Table 2.2) was also included to verify transformation efficiency and cell competency.

2.6 Construction of the Δ *sufT* knockout mutant

The *M. smegmatis* Δ *sufT* knockout mutant was constructed by two-step allelic exchange as previously described (Parish and Stoker, 2000) and illustrated in Figure 2.1.

2.6.1 Construction of Δ *sufT* allelic exchange vector

2.6.1.1 Generation of upstream and downstream regions for allelic exchange

The allelic exchange homologous upstream (3127US) and downstream (3127DS) DNA fragments were generated by PCR using Phusion® High-Fidelity DNA polymerase (section 2.3.1) and primers listed in Table 2.3. The 3127US and 3127DS PCR products were separated on agarose gels (section 2.3.3), gel purified (section 2.3.4) and ligated (section 2.3.6) into the sequencing vector pJET1.2 (Table 2.2). The ligation reaction was transformed into competent *E. coli* XL1 Blue cells (section 2.4.2). Colonies from the 3127US and 3127DS transformation plates were selected and cultured for plasmid DNA amplification (pJET3127US and pJET3127DS) (Table 2.2) and extraction (section 2.2.1). Restriction endonuclease digests (section 2.3.5) were performed to verify the presence of the insert DNA. Plasmids containing

the correct insert sizes were Sanger sequenced (section 2.3.2) to confirm that no mutations were introduced during PCR amplifications.

2.6.1.2 Three-way cloning into p2NIL

Following sequencing, the 3127US and 3127DS DNA fragments, were excised from pJET3127US using HindIII and EcoRI and from pJET3127DS using KpnI and EcoRI, respectively. The suicide vector p2NIL (Table 2.2) was digested with HindIII and KpnI and gel purified. A three-way ligation with purified 3127US, 3127DS and p2NIL fragments was done, and the reaction transformed into competent *E. coli* XL1 Blue cells. Potential clones were selected and screened by PCR. Restriction endonuclease digests were performed to verify the three-way ligation and construction of the p2NIL Δ sufT vector (Table 2.2).

2.6.1.3 Selectable marker cloning

The Pacl gene cassette containing *lacZ* and *sacB* genes from pGOAL17 (Parish & Stoker, 2000) was cloned into the Pacl site of p2NIL Δ sufT to construct the p2NIL Δ sufTpGOAL17 vector (Table 2.2) as illustrated in Figure 2.1 and transformed into competent *E. coli* XL1 Blue cells. Cells were plated on media containing X-gal and blue colonies containing potential clones were selected and their plasmid DNA isolated and purified. The insertion of the Pacl cassette was confirmed by restriction endonuclease digests. Functionality of the *sacB* gene was confirmed by plating on LA containing 5% sucrose.

2.6.2 Generation of single cross over (SCOs)

Large scale plasmid isolation of p2NIL Δ sufTpGOAL17 was done and the purified plasmid DNA was used to transform competent *M. smegmatis* mc²155 cells (section 2.5). SCOs were selected on 7H10 plates containing Km and X-gal. SCOs were subsequently resuspended in PBST and spotted on 7H10 plates containing 2% sucrose to confirm the presence of the *sacB* gene.

SCOs were blue, kanamycin resistant and sucrose sensitive. This is because they contain the *lacZ* gene, which expresses β -galactosidase that degrades X-gal and forms a blue product in transformed colonies, and a *sacB* gene, which expresses levansucrase which generates a toxic by-product when degrading sucrose therefore, only transformants that have gained this marker will die when grown in 5% sucrose. To confirm their genotype, chromosomal DNA was extracted from the SCOs (section 2.2.2.2) and a colony PCR was done using the primers 3127 comp fwd and 3127 comp rvs (Table 2.3) and the following PCR parameters: denaturation (95°C) for 4 mins, 35 cycles of denaturation (95°C) for 30s, annealing (60°C) for 30s and elongation (72°C) for 1 min, with a final elongation step at 72°C for 7 mins. Two bands of 386

bp and 196 bp are expected because they represent the wild-type and mutant alleles respectively, both of which are present in the SCOs (Fig. 2.1). The SCOs identified by PCR were further confirmed by Southern blot analysis.

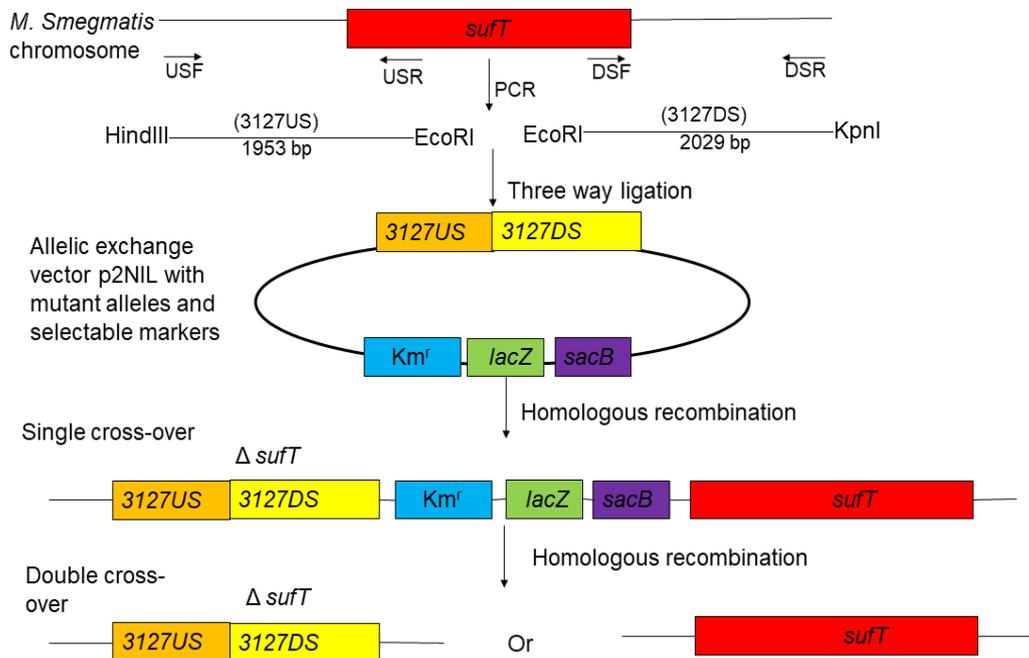


Figure 2. 1. Construction of a *sufT* knockout mutant by two-step allelic exchange. This figure was adapted from Parish & Stoker, 2000.

2.6.3 Generation of double cross over mutants (DCOs)

Genotypically confirmed SCOs were selected and counter-selected for the second homologous recombination event performed by culturing in 7H9 GST lacking *Km*, and selection on 7H10 GS containing 5% sucrose. X-gal was also added to aid selection of DCOs.

Double cross over mutants (DCOs) lose the vector carrying the selectable markers and counter selectable markers during the second homologous recombination event (Fig. 2.1). They were therefore white colonies that are *Km* sensitive and sucrose resistant. DCOs can either revert to the wild-type genotype or to the mutant genotype. To distinguish between the two, chromosomal DNA was extracted from DCOs (Section 2.2.2.2) and screened by PCR using primers 3127 comp fwd and 3127 comp rvs (Table 2.3) and the following PCR parameters: denaturation at 95°C for 4 mins, 35 cycles of denaturation at 95°C for 30s, annealing at 60°C for 30s and elongation at 72°C for 1 min, with a final elongation step at 72°C for 7 minutes. The expected sizes were 380 bp for the wild-type and 196 bp for the mutant

genotype. DCOs identified by PCR were then assessed by Southern blot analysis to confirm their genotype.

2.7 Southern Blot analyses

Southern blot analysis was done as previously described (Warren et al., 2009) with some modifications.

2.7.1 Southern transfer

Approximately 2-4 µg of genomic DNA (section 2.2.2.1) from the wild-type, mutants and SCO strains was digested overnight at 37°C with either SmaI or EcoRI. The genomic DNA fragments were separated by gel electrophoresis at 80 V in a 1% agarose gel for approximately 3 hours (hrs). The gels were removed from the tray, incubated in 0.01% (v/v) (10 mg/ml) ethidium bromide with gentle shaking for 20 mins and visualized (section 2.3.3) to confirm complete enzyme digestion. The gels were placed inverted into two separate labelled containers (SmaI and EcoRI) and washed in denaturing buffer (1.5 M sodium chloride (NaCl) and 0.5 M sodium hydroxide (NaOH)) for 30 mins at room temperature with gentle shaking. The denaturing buffer was removed, and gels washed in neutralising buffer (1.5M NaCl and 0.5 M Tris(hydroxymethyl)aminomethane (Tris-HCl)) for 30 mins at room temperature with gentle shaking, then equilibrated in transfer buffer (pH 7.4) (3 M NaCl, 0.2 M sodium dihydrogen phosphate hydrous (NaH₂PO₄·H₂O) and 20 mM EDTA). Transfer buffer was poured into the reaction tray and the DNA was transferred from the gel on to a nitrocellulose membrane (Hybond) by capillary action for 16 hrs. The membranes were washed in 2X transfer buffer for 10 mins at room temperature, placed between a cover of Whatman paper and baked at 80°C for 2 hrs to fix the DNA onto the membrane.

2.7.2. Preparation of the DNA probes by PCR

The 3127US and 3127DS DNA fragments were used as probes for the SmaI and EcoRI digested chromosomal DNA respectively. They were generated as previously described (section 2.3.1).

2.7.3 Pre-hybridisation

The membranes were rehydrated in water and partially sealed in separate plastic bags. ECL™ (25 ml) gold buffer was added to each plastic bag, the bags were sealed and incubated at 42°C for at least 1 hr in a shaking water bath.

2.7.4 ECL labelling of probe

Each probe (200 ng) (section 2.7.2) was made up in a total volume of 15 µl, incubated at 100°C for 5 mins to denature the probe, then placed on ice for 5 mins. The Amersham™ ECL™ Direct Nucleic Acid Labelling System (GE Healthcare) labelling solution (15 µl) and glutaraldehyde mix (15 µl) were added to each probe, mixed well and incubated at 37°C for 10 mins to allow labelling of the probes.

2.7.5 Hybridisation

The labelled probes were added to ECL gold buffer of the corresponding membrane, air bubbles removed, the bags re-sealed and incubated for 16 hrs at 42°C in a shaking water bath. The membrane was washed twice in primary wash buffer (pre-warmed to 42°C) (6 M urea, 0.01M SDS, 37.5 mM NaCl and 3.75 mM tri-sodium citrate ($\text{Na}_3\text{C}_6\text{H}_5\text{O}_7$)) in a water bath, with shaking at 42°C. This was followed by two wash steps at room temperature for 5 mins in hybridisation buffer (300 mM NaCl and 30 mM $\text{Na}_3\text{C}_6\text{H}_5\text{O}_7$) with gentle shaking.

2.7.6 Detection of hybridisation

The membranes were placed in clean plastic sleeves and all excess fluid was removed. A detection reagent mix was made by mixing 4 ml of each of the Clarity™ Western ECL substrate peroxide and luminol/enhancer solutions (Bio-Rad), then 4 ml of the detection reagent mix was added to each membrane. The plastic sleeves were sealed, and the detection reagent mix spread over the membranes for 90 seconds. All the excess reagents were removed, the plastic sleeves wiped dry and placed in an X-ray cassette. In the dark room, the membranes were exposed to X-ray film for an optimal time (1 min to 1 hr).

2.8 Genetic complementation

2.8.1 Generation of the *sufT* complementation vector

A functional copy of the *sufT* gene including 26 bp upstream of the start codon and 17 bp downstream of the stop codon, was generated by PCR using the Phusion® DNA polymerase and the primers listed in Table 2.3 and the following PCR parameters: denaturation (98°C) for 1 min, 35 cycles of denaturation at 98°C for 15s, annealing at 62°C for 30s and elongation at 72°C for 30s, with a final elongation step at 72°C for 7 mins. The PCR products were separated on agarose gels (section 2.3.3), gel purified (section 2.3.4) and ligated (section 2.3.6) into the sequencing vector pJET1.2. Competent *E. coli* XL1 Blue cells were transformed with the ligation mix (section 2.4.2). Transformants were selected and plasmid DNA (pJET3127) (Table 2.2) isolated (section 2.2.1). Restriction endonuclease digests (section 2.3.5) were done to

verify the presence of the insert. Plasmids containing the insert were sent for Sanger sequencing (section 2.3.2) to confirm that no mutations were introduced during PCR amplification.

2.8.2 Sub-cloning into the pSE100 expression vector

Following sequencing, the *sufT* insert was excised from pJET3127 using BamHI and HindIII. The expression vector pSE100 was digested with the same enzymes and gel purified and the *sufT* insert ligated into the BamHI and HindIII site of pSE100. *E. coli* XL1 Blue competent cells were transformed with the ligation mix and transformants selected on LA plates containing Hyg (150 µg/ml). Potential clones were selected, and restriction endonuclease digests used to verify the presence of the insert and construction of the vector pSE3127 (Table 2.2).

2.8.3 Electroporation into competent *M. smegmatis* Δ *sufT*

Large scale plasmid isolation of pSE3127 was done and competent *M. smegmatis* Δ *sufT* cells were transformed with the purified plasmid DNA. Transformants were selected on 7H10 GS containing Hyg (150 µg/ml). DNA was extracted from potential transformants (Section 2.2.2.2) and screened by PCR using primers 3127 comp fwd and 3127 comp rvs (Table 2.3) and the following PCR parameters: denaturation at 95°C for 4 mins, 35 cycles of denaturation at 95°C for 30s, annealing at 60°C for 30s and elongation at 72°C for 1 min, with a final elongation step at 72°C for 7 minutes.

2.9 Phenotypic characterisation

The *M. smegmatis* strains mc²155-1, Δ *sufT*, Δ *sufT* (pSE3127), mc²155-2, Δ *MycP3ms* and Δ *MycP3ms::pr1MycP3ms* (Table 2.1) were assessed for differences in growth kinetics, enzyme kinetics, sensitivity to oxidative stress, drug sensitivity and biofilm formation.

2.9.1 Growth kinetics

The *M. smegmatis* strains were assessed for differences in growth kinetics under standard culture conditions.

2.9.1.1 Standard culture media (7H9 GST) growth kinetics

The *M. smegmatis* strains mc²155, Δ *sufT* and Δ *sufT* (pSE3127) were pre-cultured overnight in 7H9 GST from frozen stocks. The pre-cultures were used to inoculate 50 ml cultures with appropriate antibiotics at a starting optical density measured at 600nm (OD₆₀₀) of 0.005. The cultures were incubated at 37°C in a shaking incubator (180 rpm) and growth monitored at 3-

hr intervals for 33 hrs by measuring the OD₆₀₀ using the Ultrospec 4051 UV-visible spectrophotometer (LKB Biochrom). The experiments were done in biological triplicate.

2.9.2 Enzyme kinetics

The activity of the Fe-S cluster containing enzymes aconitase and succinate dehydrogenase were evaluated in the *M. smegmatis* strains mc²155-1, Δ sufT and Δ sufT (pSE3127). The strains were pre-cultured overnight in 7H9 GST from frozen stocks. The pre-cultures were used to inoculate 50 ml of 7H9 GST and grown to mid-log (OD₆₀₀ of 0.7-1). Each sample was split into 2 x 10 ml samples, centrifuged at 1811 x g for 10 minutes, the supernatants discarded, and cells were washed by resuspending the pellets in 10 ml PBST. The cells were harvested at 1811 x g for 10 mins, the supernatants discarded, and the pellets stored at -80°C.

The *M. smegmatis* cell pellets were thawed on ice, resuspended in 20 mM Tris-Cl (pH 8.0) and transferred into 2 ml screw cap tubes containing 0.1 mm glass beads. WCLs were prepared as previously describe (section 2.1.3.1) with minor modification. Briefly, cell lysis was done 3 times for 20s at 4.5 m/s with 2-minute rest periods on ice between the lysis steps. The lysate was transferred to new tubes and centrifuged at 2 700 x g for 10 mins at 4°C and 500 μ l of the supernatant was transferred to new tubes to be used in the succinate dehydrogenase assay. The rest of the sample was then centrifuged again at 17 950 x g for another 10 mins and the supernatant transferred to new tubes to be used in the aconitase assay.

2.9.2.1 Succinate dehydrogenase activity assay

The succinate dehydrogenase activity was measured using a calorimetric continuous assay in 96 well plate format as previously described (Munujos et al., 2003). Each sample had a baseline activity control in which the substrate was omitted. All samples were done in technical duplicate. A master mix of 750 mM Tris (pH 8.0), 10 mM EDTA, 50 mM potassium cyanide (KCN), 2 mM idonitrotetrazolium chloride (INT), 1.2 mg/ml kolliphor and water was prepared and mixed well. Then, 66 μ l master mix and 30 μ l whole cell lysate were then dispensed into the wells on the plate. The assay was initiated by adding 4 μ l succinate (0.5 M) to the test wells and water (to make equal reaction volumes) to the baseline activity wells. The absorbance was read at 550 nm every minute for 30 minutes on a Bio-Tek Synergy HT plate reader and analysed using the Gen3 software. The enzyme activity was calculated using the gradient of the linear part of the graphs generated using the equation:

$$[(\text{slope}_{\text{replicate 1}} + \text{slope}_{\text{replicate 2}})/2] - [(\text{slope}_{\text{baseline replicate 1}} + \text{slope}_{\text{baseline replicate 2}})/2]$$

=Succinate dehydrogenase activity/ min

The protein concentration in the WCLs was calculated using the Bradford assay (section 2.1.3.2) and succinate dehydrogenase activity was expressed relative to protein concentration in each sample. The final succinate dehydrogenase activity was reported as enzyme activity per mg protein per min (activity/ mg/ minute). The assay was done in biological triplicate.

2.9.2.2 Aconitase activity

The aconitase assay was done in a 96 well plate format. Each sample had a background control in which the substrate was omitted, and all samples were done in technical duplicate. The aconitase assay was done using the aconitase activity assay kit (Sigma Aldrich) with some modifications. Briefly, an isocitrate standard curve (0- 20 nanomole) was generated following the manufacturer's instructions. A master mix for the test samples was also prepared containing 1x assay buffer (23 μ l), enzyme mix (1 μ l) and substrate (1 μ l) per sample, and for the background control, the substrate was replaced with water (1 μ l) in the master mix. Master mix (25 μ l) and whole cell lysate (25 μ l) were added to the wells of the plate and incubated at room temperature in the dark for 30 minutes. Detection reagent (5 μ l) was added to each well and incubated in the dark for 10 minutes. The absorbance was read at 450 nm using a Bio-Tek Synergy HT plate reader and analysed using the Gen3 software. The standard curve was used to determine the aconitase activity in each sample (mU/ ml) using the following equation:

$$\text{Aconitase activity} = [B / (TxV)] \times \text{sample dilution factor}$$

$$= \text{nmol/minute/ml}$$

$$= \text{mU/ml}$$

$$B = \frac{\text{Average Sample Absorbance (450 nm)} - \text{Average Background Absorbance (450 nm)}}{\text{Gradient of standard curve}}$$

Gradient of standard curve

T= Incubation time (mins)

V= Sample volume (ml)

The protein concentration in the WCLs was calculated using the Bradford assay (section 2.1.3.2) and aconitase activity was expressed relative to protein concentration for each sample. The final aconitase activity was reported mU/mg/ml. The assay was done in biological triplicate.

2.9.3 Survival under oxidative stress conditions

The *M. smegmatis* strains mc²155-1, Δ sufT and Δ sufT (pSE3127) were assessed for differences in survival following exposure to the redox cyler 2,3-dimethoxy-1,4-

naphthoquinone (DMNQ). The strains were pre-cultured overnight in 7H9 from frozen stocks. The pre-cultures were then used to inoculate 50 ml of 7H9 GST and grown to mid-log phase (OD_{600} of 0.7-1). The mid-log culture (20 ml) was then sonicated in a water bath for 12 minutes to separate clumping bacteria, filtered using a BD Falcon™ 40 μ M cell strainer and then diluted to OD_{600} of 0.2 in 7H9 GST. Each test had a culture exposed to DMNQ (30 μ M) and an unexposed control culture (exposed to the diluent of DMNQ, DMSO). After addition of DMNQ, time 0 hours of the exposed and unexposed cultures were sampled and a 10-fold serial dilution (10^{-1} - 10^{-7}) performed. The appropriate dilutions were plated on solid media in technical triplicate for each sample and incubated at 37°C for 3 days. Sampling and plating were done every 3 hours for 12 hours. The experiment was done in biological triplicate.

Bacterial growth was assessed by calculating colony forming units per ml (CFU/ml) using the formula:

$$\text{CFU/ml} = \frac{\text{number of colonies}}{\text{Dilution factor}} \times 10$$

Percentage survival was then calculated using the equation:

$$\% \text{ Survival} = \frac{\text{CFU/ml (DMNQ exposed)}}{\text{CFU/ml (unexposed)}} \times 100$$

2.9.4 Drug sensitivity testing

The *M. smegmatis* strains mc²155-1, Δ *sufT* and Δ *sufT* (pSE3127) were assessed for differences in sensitivity to the anti-tuberculosis (TB) drugs clofazimine, isoniazid and rifampicin using the broth micro dilution method to determine the minimum inhibitory concentration (MIC). The strains were pre-cultured overnight in 7H9 GST from frozen stocks. The pre-cultures were used to inoculate 50 ml 7H9 GST and grown to a final OD_{600} of approximately 0.2- 0.3. The culture was then diluted 1:100 in 7H9 GST. The tests are done in the rows of sterile 96 well tissue culture plates. A volume of 50 μ l of 7H9 GST was aliquoted into rows 2 to 12 of the plate. In the first row, 100 μ l of either the antibiotic, antibiotic diluent or media was aliquoted in technical duplicate. A 1:2 serial dilution of the drugs, diluent and media was done by transferring 50 μ l from row 1 into row 2 and mixing well. This process was then repeated from row 2 to row 12 to create a 2-fold dilution series down the plate. The last 50 μ l was discarded to bring the volume in each well to 50 μ l. A volume of 50 μ l of the diluted *M. smegmatis* culture were then added to all the wells of the plates. The plates were then incubated statically at 37°C for 4 days. After 4 days, resazurin (final concentration 0.0046%) was added to all the wells in the plate and incubated overnight at 37°C. Resazurin changes

from blue to pink in the presence of bacteria in respiratory growth. The plates were assessed visually and the drug concentration where the colour change from blue to pink was observed was reported as the breaking point for growth versus no growth. The MIC range reported as the highest drug concentration at which growth is observed and the lowest drug concentration at which growth is inhibited.

2.9.5 Iron limitation growth kinetics

The *M. smegmatis* strains mc²155-1, Δ *sufT*, Δ *sufT* (pSE3127), mc²155-2, Δ *MycP*_{3ms} and Δ *MycP*_{3ms::pr1MycP}_{3ms} were pre-cultured overnight in 7H9 GST from frozen stocks. The pre-cultures were washed twice harvesting by centrifugation at 1 811 x g and resuspending the pellet in 10 ml PBST. The washed culture was then used to inoculate 50 ml of MM with 0.05% Tween 80 and appropriate antibiotics at a starting OD₆₀₀ of 0.005 to starve the bacteria of iron. Cultures were incubated at 37°C in a shaking incubator (180 rpm) for 27 hours. After 27 hours the cultures were again washed in PBST as described above, then used to inoculate 50 ml cultures in MM with 0.05% Tween 80 at a starting OD₆₀₀ of 0.005. The impact of various iron concentrations on growth was evaluated by adding various concentrations (0- 2 µM) of iron(III) chloride (FeCl₃). Growth was monitored at 3-hour intervals for 56 hours by measuring OD₆₀₀ using the Ultrospec 4051 UV-visible spectrophotometer (LKB Biochrom). The experiments were done in plastic containers because glass leaches metals. The experiments were done in biological triplicate.

2.9.6 Biofilm formation

The *M. smegmatis* strains were assessed for differences in pellicle biofilm formation under standard culture conditions and under iron limitation.

2.9.6.1 Standard culture media (Sauton's) biofilm formation

M. smegmatis pellicle biofilms were cultured in tissue culture plates as previously described (Kulka et al., 2012) with some modifications. Briefly, the *M. smegmatis* strains mc²155-1, Δ *sufT* mutant and Δ *sufT* (pSE3127) were pre-cultured overnight in 7H9 GST from frozen stocks. The pre-cultures were used to inoculate 50 ml cultures and grown to mid-log phase (OD₆₀₀ of 0.7-1). The bacteria were then diluted 1:100 in Sauton's media and 2.4 ml of each culture was aliquoted in technical triplicate in a 24 well plate. The plate was sealed with several layers of parafilm to prevent media evaporation which can retard bacterial growth and incubated statically at 37°C for 5 days.

Pellicle biofilm formation was assessed by macroscopic observation and the biofilm biomass was quantified by calculating total protein using the Bradford assay (Bradford, 1976) (section 2.1.3).

2.9.6.2 Iron limitation pellicle biofilm formation

The iron limitation pellicle biofilms were cultured like the standard culture pellicle biofilms (section 2.9.2.1) with some modifications. Briefly, the *M. smegmatis* strains mc²155-1, Δ *suft*, Δ *suft* (pSE3127), mc²155-2, Δ *MycP*_{3ms}, Δ *MycP*_{3ms}::*pr1MycP*_{3ms} were pre-cultured overnight in 7H9 GST from frozen stocks. The pre-cultures were centrifuged at 1 811 x *g*, washed twice in 10 ml PBST. The washed culture was then used to inoculate 50 ml of MM with 0.05% Tween 80 to a starting OD₆₀₀ of 0.005 and incubated at 37°C in a shaking incubator (180 rpm) for 27 hours to starve the bacteria of iron. After 27 hours, the cultures were washed in PBST as described above, diluted to an OD₆₀₀ of 0.01 in MM containing a range of FeCl₃ concentrations (0- 50 µM). Each culture was aliquoted (2.4 ml) in technical triplicate into a 24 well plate. The plate was sealed with several layers of parafilm and incubated statically at 37°C for five days. Pellicle biofilm formation was assessed as previously described (section 2.1.3).

2.9.7 Statistical analyses

All the statistical analyses were done using GraphPad Prism® version 7.03 software package. Comparisons between the wild-type and mutant phenotype characterisation data was analysed by unpaired (two-sample) t-tests and each row was analysed individually, without assuming a consistent standard deviation. P-values were determined using the two-stage linear step-up procedure for controlling the false discovery rate (FDR) (Benjamini et al., 2006). FDR is the expected proportion of null hypotheses that were true but mistakenly rejected. FDR was set at 1% and a p-value greater than 0.05 was considered statistically significant.

To perform regression analysis on the growth curve data sets, a model of best fit for the data first had to be determined. A comparison between a linear or non-linear model of best fit was made using the Akaike information criterion (AIC) (Aikake, 2011). This AIC determines how well the data supports each model, accounting for both the goodness-of-fit (sum-of-squares) and the number of parameters in the model (Aikake, 2011). The results are then expressed as the probability that each model is correct, with the probabilities summing to 100%. After choosing a model, the results generated from the regression analysis will then be used to make comparison between wild-type and mutant strains using unpaired (two sample) t-tests as described above.

Chapter 3: Results

3.1 Construction of a *M. smegmatis* Δ *sufT* knockout mutant strain

To study the role of *sufT* in the physiology of mycobacteria, an allelic exchange mutagenesis strategy was used to generate an unmarked in-frame deletion in *sufT*. The principle of allelic exchange mutagenesis is based on homologous recombination (HR). HR is a process of repairing double strand DNA breaks in cells using a homologous DNA sequence as a template to repair the break (Cromie et al., 2001). HR is an essential part of every actively replicating cell and in prokaryotes it is catalysed by the recombination enzymes RecBCD complex, RecA, RuvA and RuvB (Cromie et al., 2001). When used for site directed mutagenesis, the mutant allele flanked by DNA regions that are homologous to the DNA regions flanking the wild-type allele serves as the template for DNA repair with a possibility of being incorporated into the host chromosome (Parish & Stoker, 2000; Faulds-Pain & Wren, 2013). Allelic exchange can be a one-step or two-step process. Two-step allelic exchange has been shown to be an efficient and reliable system for generating unmarked knockout mutants in mycobacteria (Parish & Stoker, 2000) and was therefore used to generate the Δ *sufT* mutant. Two-step allelic exchange requires an allelic exchange vector which contains the mutant allele and selection markers (Parish & Stoker, 2000).

3.1.1 Construction of Δ *sufT* allelic exchange vector

To construct a Δ *sufT* allelic exchange vector, allelic exchange regions flanking the upstream (3127US) and downstream (3127DS) regions of *sufT* were amplified by PCR using Phusion DNA polymerase (Fig. 3.1) and cloned into the sequencing vector pJET1.2 (Table 2.2) to generate the vectors pJET3127US and pJET3127DS (Fig. 3.2).

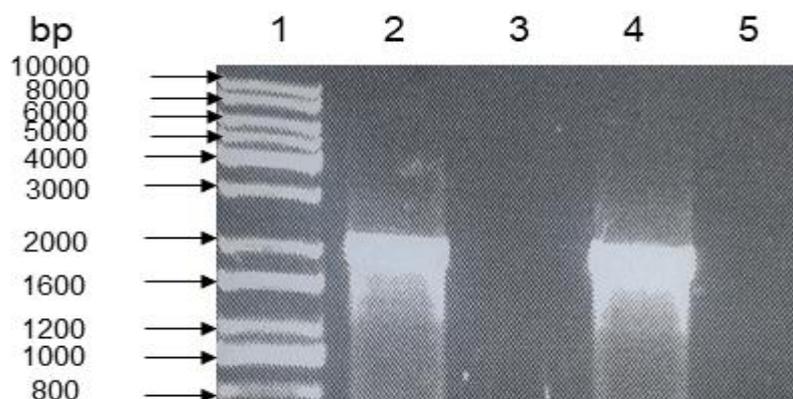


Figure 3. 1. Agarose gel electrophoresis of 3127US (1953bp) and 3127DS (2029bp) allelic exchange PCR amplicons. Lane (1) Kappa Universal DNA ladder, (2) 3127US PCR product, (3) no template DNA control, (4) 3127DS PCR product and (5) no template DNA control.

To verify the cloning, the inserts were excised from pJET1.2 using BglIII. The BglIII digests of pJET3127US (Fig. 3.2B) and pJET3127DS (Fig. 3.2D) generated the expected band sizes confirming that the vectors contained the insert DNA fragments. The inserts (3127US and 3127DS) were Sanger sequenced to check for any mutations that could have been introduced during PCR amplification. The sequences were confirmed to be correct (data not shown).

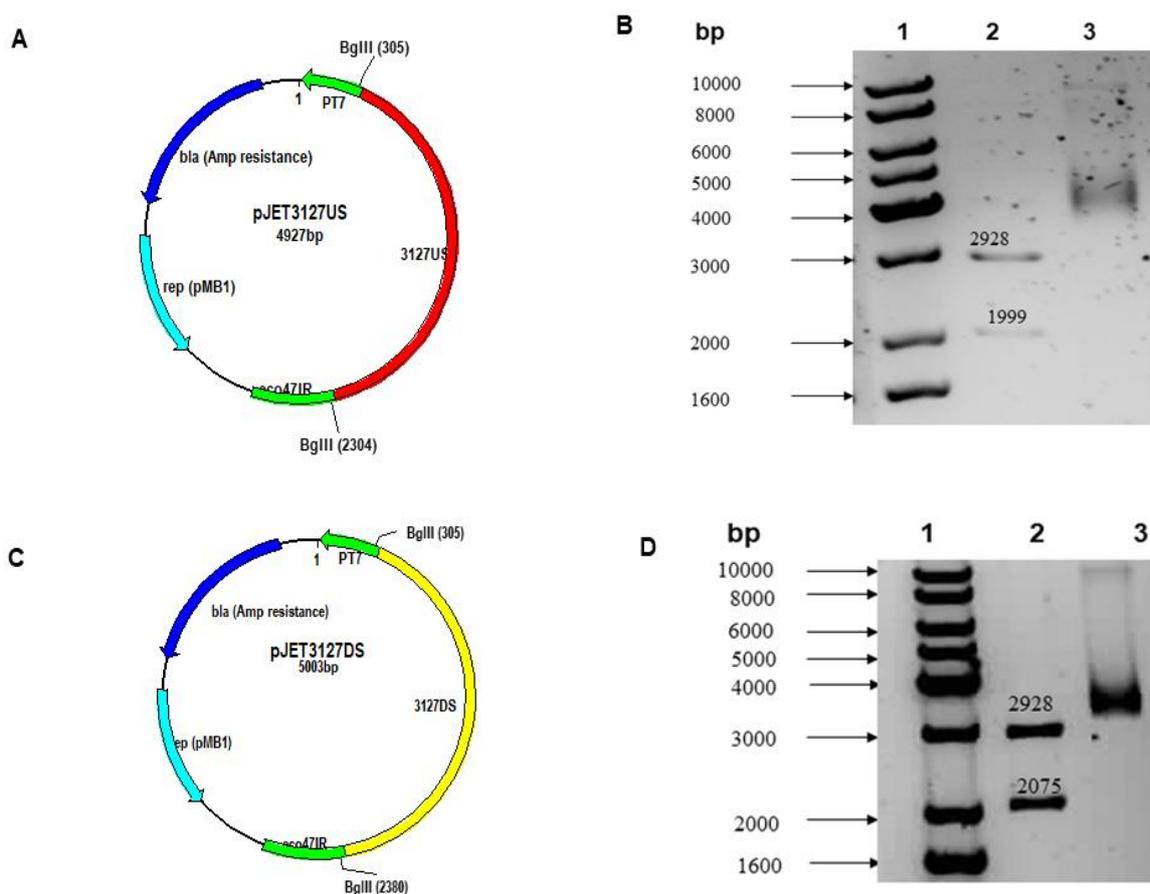


Figure 3. 2. Restriction digest of pJET3127US and pJET3127DS. (A). Vector map of pJET3127US showing the RE cut sites (B). Agarose electrophoresis of RE digest of pJET3127US. Lane (1) Kappa Universal DNA ladder, (2) pJET3127US digested with BglIII and (3) undigested pJET3127US. (C). Vector map of pJET3127DS showing the RE cut sites. (D). Agarose electrophoresis of RE digest of pJET3127DS. Lane (1) Kappa Universal DNA ladder, (2) pJET3127DS digested with BglIII and (3) undigested pJET3127DS.

The inserts (3127US and 3127DS) were subsequently directionally cloned into the HindIII and KpnI site of the non-replicating recombination vector p2NIL (Table 2.2) to generate the vector p2NIL Δ sufT which contains the mutant alleles and a gene conferring Km resistance (Fig. 3.3A). RE mapping with BamHI, EcoRI, HindIII, and PstI generated the expected band sizes (Fig 3.3B) verifying that the mutant allele was cloned correctly in p2NIL.

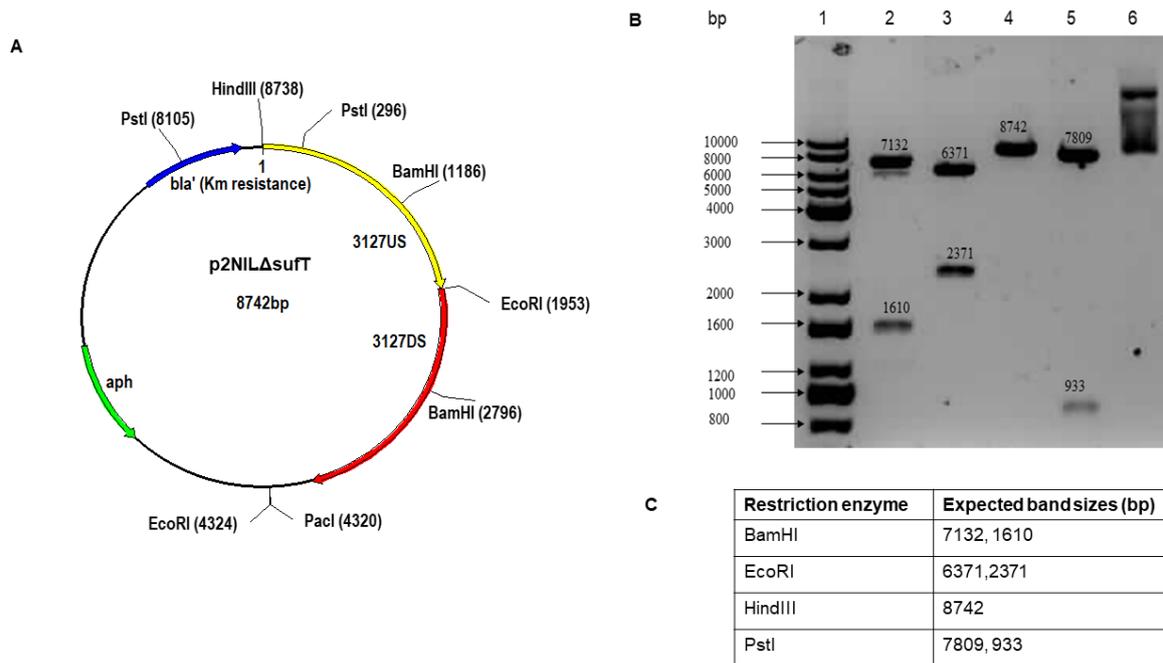


Figure 3. 3. Restriction map analyses of p2NILΔsufT. (A). Vector map of p2NILΔsufT showing the RE cut sites. (B). Agarose electrophoresis of RE mapping of p2NILΔsufT. Lane (1) Kappa Universal DNA ladder, p2NILΔsufT digested with (2) BamHI, (3) EcoRI, (4) HindIII, (5) PstI and (6) undigested vector. (C). Table showing the expected band sizes (bp).

To facilitate the identification of SCO and DCO mutants during the allelic exchange steps, the *lacZ* and *sacB* selection markers were added to p2NILΔsufT. The *lacZ* gene is a positive selection marker and it expresses β -galactosidase which degrades X-gal and forms a blue product in transformed colonies. The *sacB* gene is a counter selection marker and it expresses levansucrase which generates a toxic by-product when mycobacteria are grown in 5% sucrose, killing transformants. The markers were excised as a Pacl cassette from pGOAL17 (Table 2.2) and cloned into the Pacl site of p2NILΔsufT to generate the final allelic exchange vector p2NILΔsufTpGOAL17 (Fig. 3.4A). RE mapping with BamHI, EcoRI, HindIII and PstI generated bands of the expected sizes (Fig. 3.4B) verifying the correct insertion of the selection marker cassette in p2NILΔsufT. However, before the vector could be used for the two-step allelic exchange, the functionality of the SacB had to be verified. *E. coli* XL-1 Blue transformants were spotted on LA containing sucrose (5%) and bacteria that did not grow were confirmed as having a functional *sacB* (results not shown).

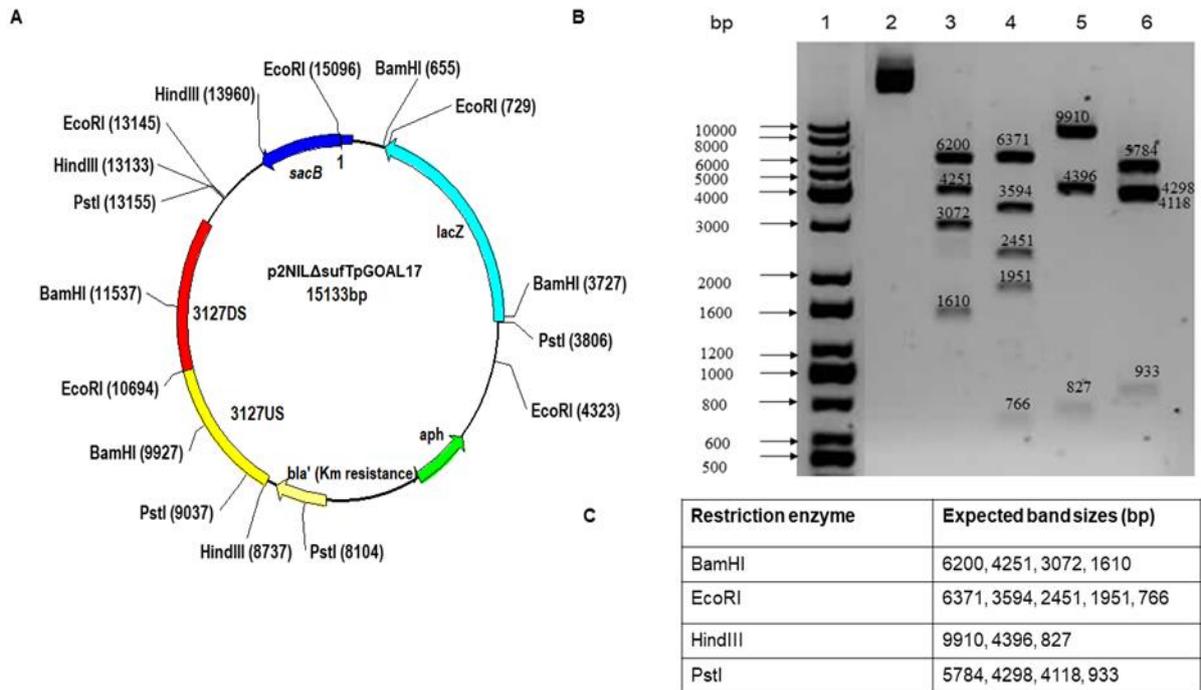


Figure 3.4. Restriction map analyses of p2NILΔsufTpGOAL17. (A). Vector map of p2NILΔsufTpGOAL17 showing the RE cut sites. (B). Agarose electrophoresis of RE mapping of p2NILΔsufTpGOAL17. Lane (1) Kappa Universal DNA ladder, (2) undigested vector, p2NILΔsufTpGOAL17 digested with (3) BamHI, (4) EcoRI, (5) HindIII and (6) PstI. (C). Table showing the expected band sizes (bp).

3.1.2 Two-step allelic exchange (Homologous recombination)

Electroporation is a method used to increase cell permeability by applying high voltage electric shocks to cells to generate temporary pores in the cell membrane (and cell wall) (Neumann et al., 1982; Potter & Heller, 2003). In bacteria, electroporation has been shown to be an efficient and reliable method of introducing vector DNA into the cell, resulting in efficient and irreversible transformations (Wu et al., 2010). Therefore, electroporation was used to transform electrocompetent *M. smegmatis* cells with the final allelic exchange vector p2NILΔsufTpGOAL17.

3.1.2.1 Identification of SCOs and DCOs

Following electroporation, the transformants were blue, Km resistant and sucrose sensitive SCOs. This indicated that the first homologous recombination event had occurred and the entire p2NILΔsufTpGOAL17 vector with all the selectable and counter selectable markers was inserted into the genome of *M. smegmatis*. SCOs were then screened by colony PCR to confirm their genotype. The PCR gave the two expected bands of 380 bp (*sufT*) and 196 bp (*ΔsufT*) (Fig. 3.5) confirming the presence of both *sufT* and *ΔsufT* alleles in the SCOs. The

verified SCO were then cultured in the absence of kanamycin to allow a second homologous recombination event to occur. This resulted in the loss of the selectable and counter selectable markers generating white, Km sensitive and sucrose resistant DCO colonies. The DCOs were either *suft* revertants or $\Delta suft$ mutants however, *suft* and $\Delta suft$ DCOs were phenotypically indistinguishable. To identify *M. smegmatis* cells with the $\Delta suft$ mutant allele, DCOs were screened by colony PCR. The colony PCR strategy employed generated products of different sizes which were used to differentiate the wild-type *suft* allele (380 bp) from the mutant $\Delta suft$ allele (196 bp) (Fig. 3.5).

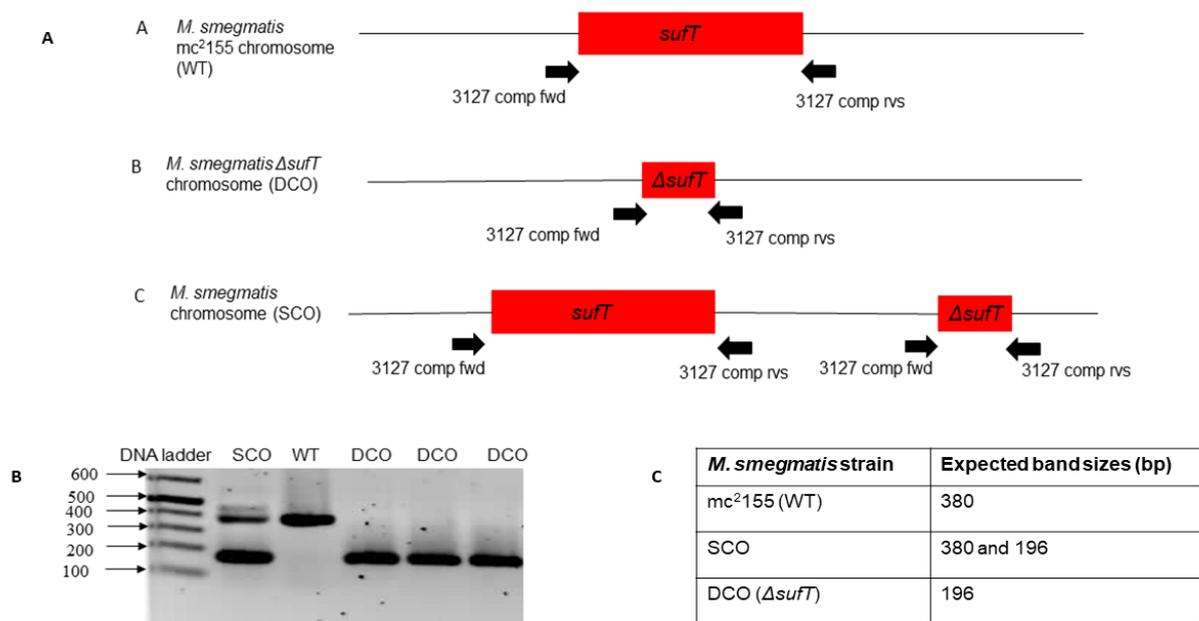


Figure 3. 5. Genotype confirmation of *M. smegmatis* $\Delta suft$ mutant strains by colony PCR. (A). PCR screening strategy to distinguish the $\Delta suft$ DCO mutants from wild-type revertants. The primers (3127 comp fwd and 3127 comp rvs) allowed for the distinction between the intact *suft* and the disrupted $\Delta suft$ gene. (B). Agarose gel electrophoresis of the PCR amplicons from the WT, SCO and DCO strains. (C). Table showing the expected band sizes.

3.1.3 Southern blot analysis

Following confirmation by colony PCR, genomic DNA was extracted from the wild-type, SCO and DCO mutant strains and Southern blot analysis was done as an additional confirmation of their genotype. Southern blotting determines if the recombination was homologous or illegitimate i.e. that it took place in the correct position in the genome. In the Southern blot strategy employed, (Fig. 3.6A) restriction enzyme cut sites outside of the upstream and downstream regions were chosen to confirm if the mutant allele had gone into the correct site

in the genome. For the genes upstream of *sufT*, the genomic DNA was digested with *Sma*I and probed with the 3127US PCR amplicon (Fig. 3.1). The two *Sma*I cut sites on either side of *sufT* generated a 2757 bp fragment in the wild-type and a 2532 bp fragment in the mutant (Fig. 3.6B). This difference in the two fragments corresponded with the approximate size of the deletion made in *sufT*. For the genes downstream of *sufT*, the genomic DNA was digested with *Eco*RI and probed with the 3127DS PCR amplicon (Fig. 3.1). The two *Eco*RI cut sites on either side of *sufT* generated an 8447 bp fragment in the wild-type however, in the mutant a 2797 bp fragment was generated because an *Eco*RI cut site was introduced when making the deletion in *sufT* (Fig. 3.6C).

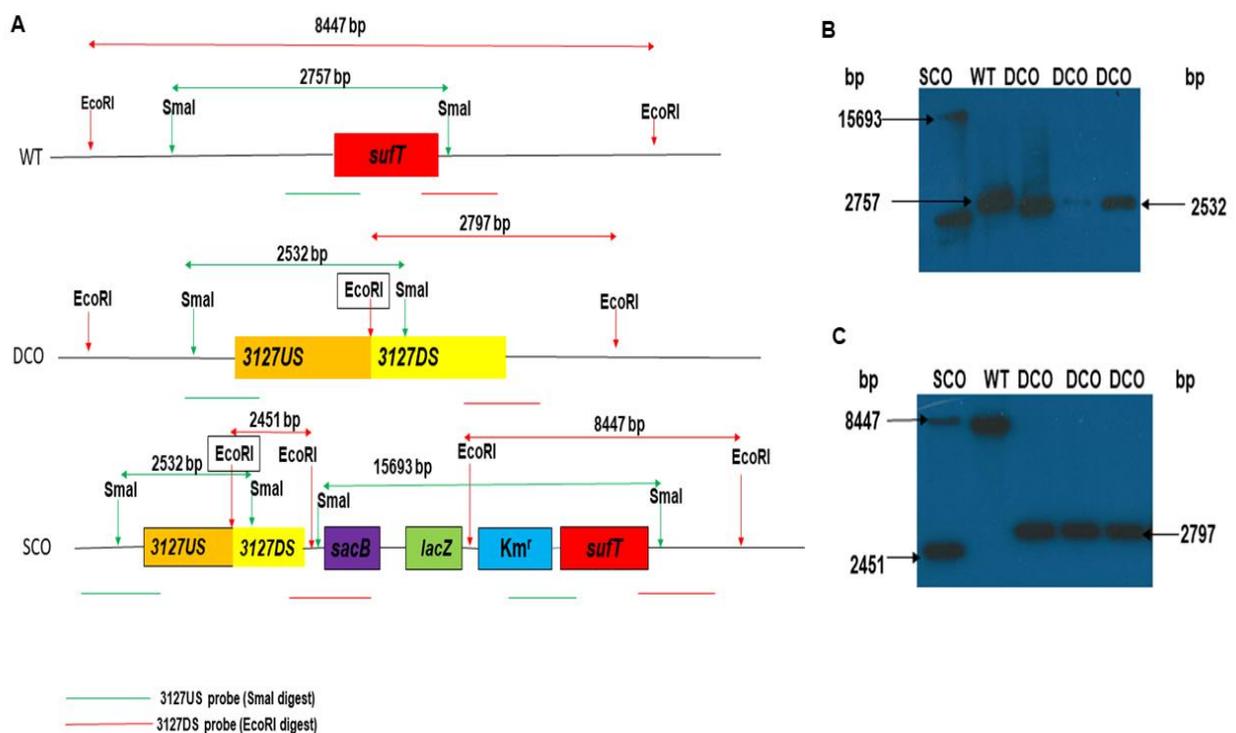


Figure 3. 6. Genotype characterisation of *M. smegmatis* Δ *sufT* mutant strains by Southern blot analysis. (A). Southern blot strategy used to distinguish the *M. smegmatis* *mc*²¹⁵⁵⁻¹ (WT), DCO and SCO mutants. Chromosomal DNA from the WT, SCO and DCO strains was digested with either *Sma*I (genes upstream of Δ *sufT*) and probed with 3127US PCR amplicon or with *Eco*RI (genes downstream of Δ *sufT*) and probed with 3127DS PCR amplicon. Green and red double-sided arrows show the expected band sizes from the *Sma*I and *Eco*RI digests respectively. (B). Southern blot analysis of the genes upstream of Δ *sufT* with generated band sizes shown in bp. (C). Southern blot analysis of the genes downstream of Δ *sufT* with generated band sizes shown in bp.

One SCO was used for both the *Sma*I and *Eco*RI digests. When digested with *Sma*I, the SCO gave two bands of 2532 bp and 15693 bp (Fig. 3.6B). When digested with *Eco*RI the SCO gave two bands of 8447 bp and 2451 bp (Fig. 3.7C). The mutant was generated from a single

crossover in which the first homologous recombination happened in the region upstream of *sufT*. The single crossover could have also occurred downstream of *sufT* but because only the single crossover that generated the mutant was analysed by Southern blot, only that result is shown. One of the DCO mutants confirmed by colony PCR and Southern blotting was then designated as the Δ *sufT* mutant strain and re-named *M. smegmatis* Δ *sufT*. The fact that a Δ *sufT* knock out mutant could be generated suggested that *sufT* is not essential for survival under standard aerobic culture.

3.2 Genetic complementation

A limitation of knockout mutagenesis is that any observed phenotypes cannot be automatically attributed to loss of the inactivated gene. This is because point mutations can arise during the culturing and selection process and there could also be downstream polar effects influencing the observed phenotype. Since the *M. smegmatis* Δ *sufT* mutant strain will be subjected to phenotypic evaluation, to control for this limitation, a complementation strain, in which *sufT* is re-introduced into the Δ *sufT* mutant on an expression vector was generated. If the complementation strain has the same phenotype as the wild-type, any change in the phenotype of the mutant can be attributed to the loss of *sufT*.

To generate a complementation vector the entire *sufT* gene including 26 bp upstream and 17 bp downstream was amplified using Phusion DNA polymerase and cloned into the sequencing vector pJET1.2 to generate the vector pJET3127 (Fig. 3.7). RE digestion with BglIII was used to verify the presence of the DNA insert (Fig. 3.7) and the insert was Sanger sequenced to confirm that no mutations were introduced during PCR amplification.

Subsequently, *sufT* was cloned into the BamHI and HindIII sites of the expression vector pSE100 to generate the vector pSE3127 (Fig. 3.8). Since, *sufT* is the last gene in the operon and presumably does not have its own promoter, it was cloned into pSE100 so that it could be expressed under control of the PmycTetO promoter. RE mapping with BamHI, HindIII and XhoI verified that the insert was cloned correctly in pSE100 (Fig. 3.8). The expression vector pSE3127 was electroporated into electrocompetent *M. smegmatis* Δ *sufT* to generate the complemented strain *M. smegmatis* Δ *sufT* (pSE3127).

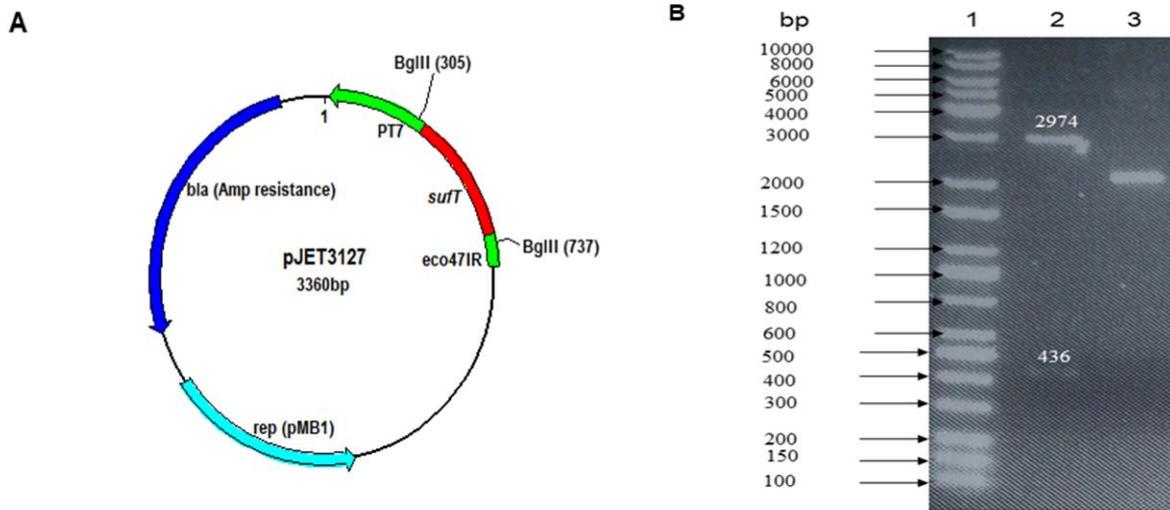


Figure 3. 7. Restriction digest of pJET3127. (A). Vector map of pJET3127 showing the RE cut sites. (B) Agarose electrophoresis of restriction digest of pJET3127. Lane (1) Kappa Universal DNA ladder, (2) pJET3127 digested with BglIII and (3) undigested pJET3127US.

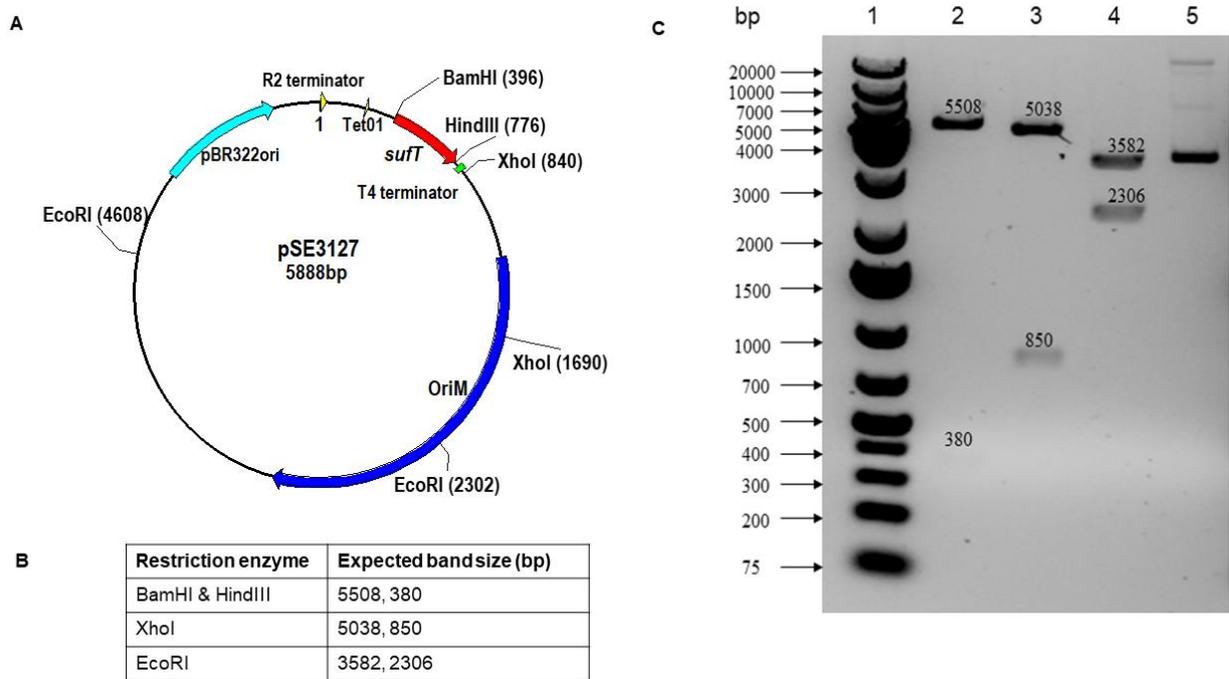


Figure 3. 8. Restriction map analyses of pSE3127. (A). Vector map of pSE3127 showing the RE cut sites. (B). Table showing the expected band sizes (bp). (C) Agarose electrophoresis of RE mapping of pSE3127. Lane (1) GeneRuler™ 1 kb Plus DNA ladder, (2) pSE3127 digested with BamHI & HindIII, (3) XhoI, (4) EcoRI and (5) undigested vector.

3.3 Phenotype characterisation

Following the successful generation of a $\Delta sufT$ mutant, the phenotype of the mutant was compared to that of the wild-type (*mc*²¹⁵⁵⁻¹) under various growth conditions to investigate the effect of loss of *sufT* on the physiology of *M. smegmatis*.

3.3.1 Growth in standard aerobic culture

The generation of a $\Delta sufT$ knockout mutant suggested that *sufT* was not essential for growth in standard aerobic culture. To investigate if loss of *sufT* has any impact on growth, the planktonic growth of *M. smegmatis* *mc*²¹⁵⁵⁻¹, $\Delta sufT$ and $\Delta sufT$ (pSE3127) was evaluated under standard aerobic culture. No difference in growth was observed between the *mc*²¹⁵⁵⁻¹ and $\Delta sufT$ confirming that SufT is dispensable for growth under standard culture conditions (Fig. 3.9).

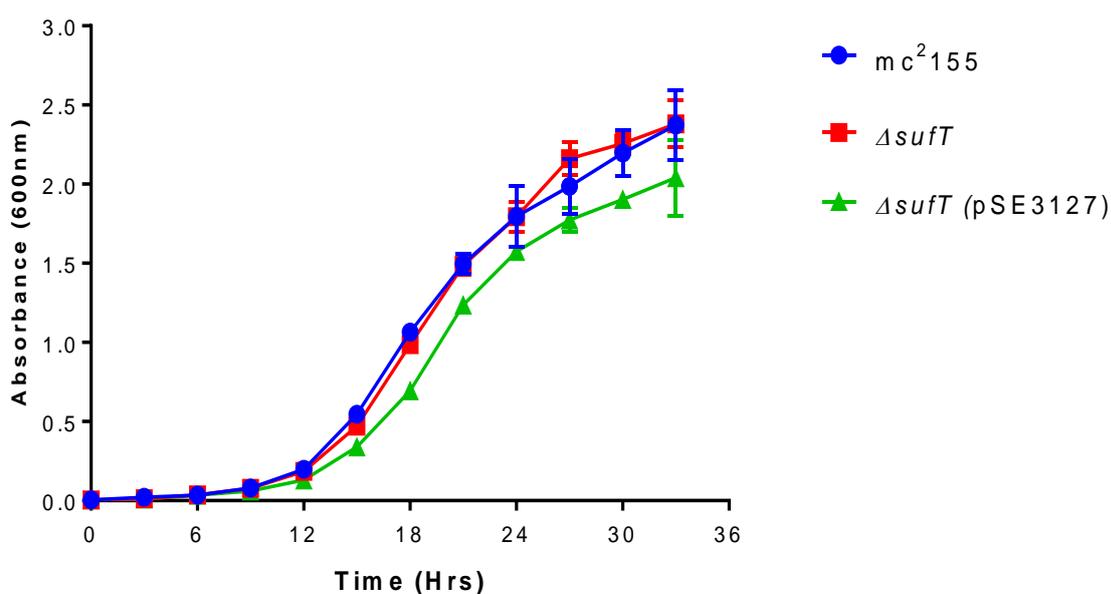


Figure 3. 9. Growth curve of *M. smegmatis* *mc*²¹⁵⁵⁻¹ (blue circles), $\Delta sufT$ (red squares) and $\Delta sufT$ (pSE3127) (green triangles) growing under standard aerobic culture conditions. The data represents the average and standard deviation (shown by error bars) of three biological replicates.

3.3.2 Enzyme activity assays

In bacteria, SufT has been showed to be involved in Fe-S cluster biogenesis (Mashruwala et al, 2016a; Mashruwala et al., 2016b, Sasaki et al., 2016) and the *sufT* from *M. tuberculosis* was able to genetically complement the loss of *sufT* in *S. aureus* (Mashruwala et al, 2016a), suggesting that SufT is potentially involved in Fe-S cluster biogenesis in mycobacteria. To

investigate the effect of loss of SufT on the ability of *M. smegmatis* to form Fe-S clusters, the formation of Fe-S clusters was measured in *mc*²¹⁵⁵⁻¹, Δ *sufT* and Δ *sufT* (pSE3127). Fe-S cluster formation was measured indirectly by measuring the activity of the Fe-S cluster dependent enzymes aconitase (ACN) and succinate dehydrogenase (SDH). This is because Fe-S clusters are unstable in their free form and can only be quantified when they are bound to their target proteins (Lill & Mühlenhoff, 2006).

3.3.2.1 Succinate dehydrogenase activity

SDH is a flavin adenine dinucleotide (FAD) containing protein with three Fe-S clusters (Cheng et al., 2006; Schirawski & Uden, 1998). It is the only enzyme to take part in both the tricarboxylic acid (TCA) cycle and the electron transport chain (Cheng et al., 2006; Schirawski & Uden, 1998). In the TCA cycle, SDH, catalyses the oxidation of succinate to fumarate and simultaneously reduces FAD to FADH₂. The FADH₂ is then oxidised back to FAD releasing 2 electrons which are then passed through a series of Fe-S clusters to menaquinone which is then reduced to menaquinol in the electron transport chain (Schirawski & Uden, 1998).

The continuous SDH activity assay uses the same principle but iodinitrotetrazolium chloride (INT) is the final electron acceptor. Briefly, succinate is oxidised to fumarate which in turns donates electrons to iodinitrotetrazolium chloride (INT) reducing it to formazan. The formation of formazan over time is therefore directly proportional to SDH activity. SDH activity was measured in the whole cell lysates of *M. smegmatis* *mc*²¹⁵⁵⁻¹, Δ *sufT* and Δ *sufT* (pSE3127) strains grown to mid-log. SDH activity in Δ *sufT* was significantly reduced, to about 10% that of *mc*²¹⁵⁵⁻¹ ($p=0.0006$) and completely restored in the Δ *sufT* (pSE3127) strain (Fig. 3.10) confirming that the observed phenotype was due to loss of *sufT*. Since SDH activity is an indirect measure of Fe-S cluster formation, these results suggest that loss of *sufT* impacts the formation of SDH associated Fe-S clusters when *M. smegmatis* is in mid-log growth in standard aerobic culture.

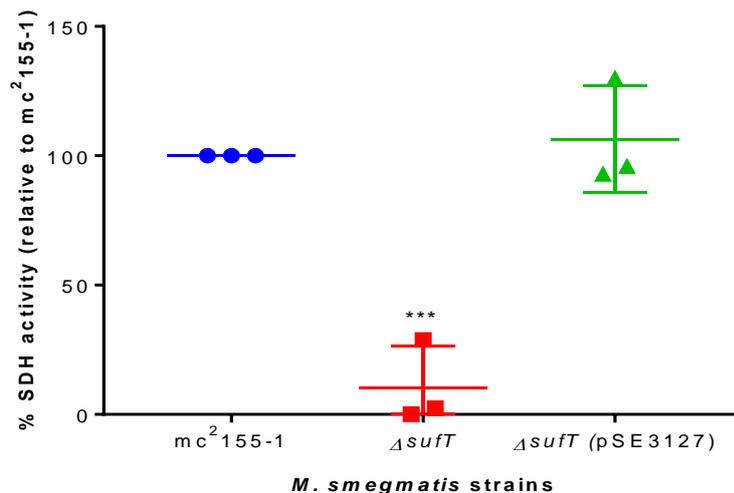


Figure 3. 10. Activity of succinate dehydrogenase in *M. smegmatis* mc²155-1, ΔsufT and ΔsufT (pSE3127) relative to the activity in mc²155-1. The activity of mc²155 was set to 100%. The data represents the average and standard deviation (shown by error bars) of three biological replicates. P-values were determined by unpaired t-tests using GraphPad Prism 7.03 software with statistically significant results indicated by p<0.05 (*), p<0.01 (**) and p<0.001 (***).

3.3.2.2. Aconitase activity

ACN is an Fe-S protein and contains a 4Fe-4S cluster (holo-form) in its active site (Robbins & Stout,1989). ACN is an important enzyme in aerobic respiration because it catalyses the isomerisation of citrate to isocitrate via a cis-aconitase intermediate in the second step of the TCA cycle. The aconitase activity assay is a coupled reaction where citrate is converted to isocitrate, then the isocitrate is processed to produces a colorimetric (450 nm) product which is used as a measure of ACN activity. ACN activity was measured in the whole cell lysates of *M. smegmatis* mc²155, ΔsufT and ΔsufT (pSE3127) strains grown to mid-log phase. There was no significant difference in ACN activity between the ΔsufT and the mc²155-1 strain (p=0.4) (Fig 3.11). A high level of variability was observed with the mutant and complemented strains in this assay, which may limit the detection of small differences between the strains. Nevertheless, these results suggest that *sufT* is dispensable for the formation of ACN associated Fe-S clusters when *M. smegmatis* is in mid-log growth in standard aerobic culture.

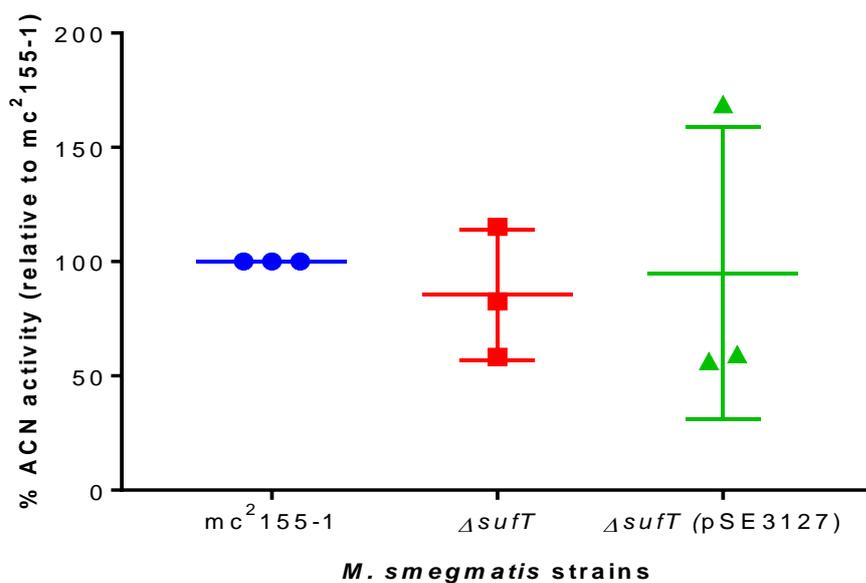


Figure 3. 11. Activity of aconitase in *M. smegmatis* mc²155-1, ΔsufT and ΔsufT (pSE3127) relative to the activity in mc²155-1. The activity of mc²155-1 was set to 100%. The data represents the average and standard deviation (shown by error bars) of three biological replicates. P-values were determined by unpaired t-tests using GraphPad Prism software with statistically significant results indicated by p<0.05 (*), p<0.01 (**) and p<0.001 (***).

3.3.3 Survival under oxidative stress

Fe-S clusters are susceptible to degradation during oxidative stress because the superoxide produced can oxidise Fe-S clusters causing them to lose their iron and degrade (Varghese et al., 2003; Imlay, 2006). The *M. smegmatis* strains mc²155-1, ΔsufT and ΔsufT (pSE3127) were assessed for differences in survival following exposure to the redox cycler 2,3-dimethoxy-1,4-naphthoquinone (DMNQ). DMNQ generates intracellular ROS using the one-electron based redox cycling mechanism. Briefly, the reduction of DMNQ by one electron produces a semiquinone product which then reduces oxygen to super oxide anions which subsequently form hydrogen peroxide (H₂O₂) (Watanabe & Forman, 2003). In the presence of iron (Fe³⁺), H₂O₂ together with superoxide can also form super hydroxyl radicals in the Fenton reaction (Winterbourne, 1995) resulting in oxidative stress.

There was no significant difference in the ability of the ΔsufT strain to survive oxidative stress when compared to mc²155-1 (Fig 3.12). It was however observed, that the ΔsufT mutant started forming clumps in the media after 3 hours of exposure to DMNQ, whereas the mc²155 strain started clumping after 9 hrs of exposure. This could suggest an increased sensitivity to DMNQ in the ΔsufT mutant, but this hypothesis is not supported by the quantitative data (Fig

3.12). It should be noted that the clumping of the cells made it challenging to accurately enumerate the colony forming units (CFUs) which were used to determine cell viability and subsequently survival.

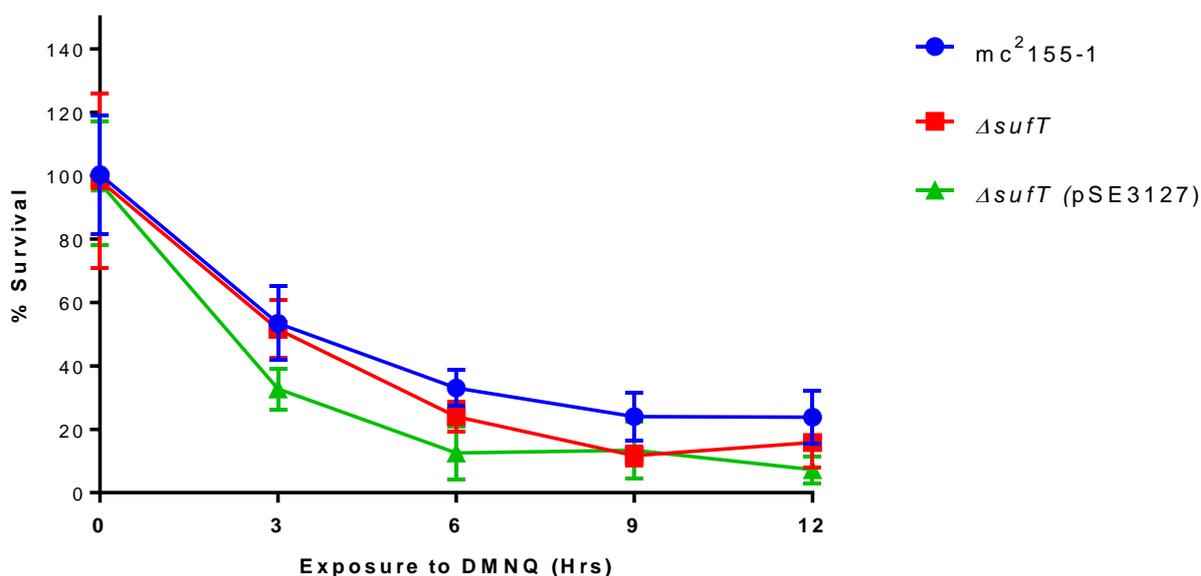


Figure 3. 12. Effect on the survival of *M. smegmatis mc*²155, Δ *sufT* and Δ *sufT* (pSE3127) after exposure to 30 μ M DMNQ. The data represents the average and standard deviation (shown by error bars) of three biological replicates. P-values were determined by unpaired t-tests using GraphPad Prism software with statistically significant results indicated by $p < 0.05$ (*), $p < 0.01$ (**) and $p < 0.001$ (***).

3.3.4 Drug sensitivity testing

KatG has a dual function as a catalase-peroxidase, and the peroxidase activity of KatG in *M. tuberculosis* has been shown to be responsible for activating the drug isoniazid (INH) (Sritharan et al., 2006). Activation of INH by katG requires iron rich conditions since the peroxidase activity of KatG was shown to be reduced in iron limiting conditions suggesting that activation of INH by katG requires iron rich conditions (Yeruva et al, 2005; Sritharan et al., 2006). In addition, the activity of INH has been shown to be reduced in mycobacteria in iron deficient conditions (Yeruva et al, 2005; Sritharan et al., 2006). Since SufT is proposed to play a role in Fe-S cluster biogenesis and changes in Fe-S cluster biogenesis can affect iron metabolism, the effect of the loss of *sufT* on the activity of INH was investigated. Broth microdilution was used to evaluate the drug sensitivity of *mc*²155-1, Δ *sufT* and Δ *sufT* (pSE3127) to INH by determining the minimum inhibitory concentration (MIC). No difference in MIC was observed between Δ *sufT* and *mc*²155-1 suggesting that the activity of INH is not affected by loss of *sufT* (Table 3.1)

Clofazimine (CFZ) has redox properties and it can harvest electrons from a normally functioning electron transport chain, to form reduced CFZ which is unstable and reacts spontaneously with oxygen to form reactive oxygen species (ROS) (Yano et al., 2010). Accumulation of CFZ in the membrane of *M. smegmatis* has been shown to produce bactericidal levels of oxidative stress (Yano et al., 2010). Since the *suf* operon is induced under conditions of oxidative stress, the effect of loss of *sufT* on the activity of clofazimine was investigated. Broth microdilution was used to evaluate the drug sensitivity of mc²155-1, Δ *sufT* and Δ *sufT* (pSE3127) to CFZ by determining the MIC. No difference in MIC was observed between Δ *sufT* and mc²155-1 suggesting that loss of *sufT* does not impact the sensitivity of *M. smegmatis* to CFZ in standard aerobic culture (Table 3.1).

Rifampicin (RIF) is an important anti-tuberculosis drug and resistance to RIF is considered a marker for multi-drug resistant TB. The effect of the loss of *sufT* on sensitivity to RIF was therefore investigated. Broth microdilution was used to evaluate the drug sensitivity of mc²155-1, Δ *sufT* and Δ *sufT* (pSE3127) to RIF by determining the MIC. No difference in MIC was observed between Δ *sufT* and mc²155-1 suggesting that loss of *sufT* does not impact the sensitivity of *M. smegmatis* to RIF in standard aerobic culture (Table 3.1).

Table 3. 1. Minimum inhibitory concentration for the drugs isoniazid, clofazimine and rifampicin.

<i>M. smegmatis</i> strain	Minimum inhibitory concentration range (µg/ml)*		
	Isoniazid	Clofazimine	Rifampicin
mc ² 155	32-64	1-2	1-4
Δ <i>sufT</i>	32-64	1-2	1-4
Δ <i>sufT</i> (pSE3127)	32-64	1-2	1-4

*The range was determined from three biological replicates

3.3.5 Growth in iron limiting conditions

The *sufR-sufB-sufD-sufC-csd-sufU-sufT* gene cluster in mycobacteria is transcribed as an operon (*suf* operon) (Huet et al., 2005) and the expression of all the genes in the *suf* operon is upregulated during iron starvation (Rodriguez et al., 2002). To investigate the role of SufT in iron metabolism, we evaluated the impact of iron limitation on the planktonic growth of *M. smegmatis* mc²155-1, Δ *sufT* and Δ *sufT* (pSE3127). The three strains were first starved of iron by growing them in Chelex® 100 resin treated mineral defined media (MM) without supplemental iron for 27 hrs, followed by growth in MM with increasing concentrations of supplemental iron up to 2 µM. An extended lag phase (~ 24 hrs) was observed for the Δ *sufT*

mutant strain, and the length of the lag phase remained unchanged even when the iron concentration in the media was increased (Fig 3.13A-D).

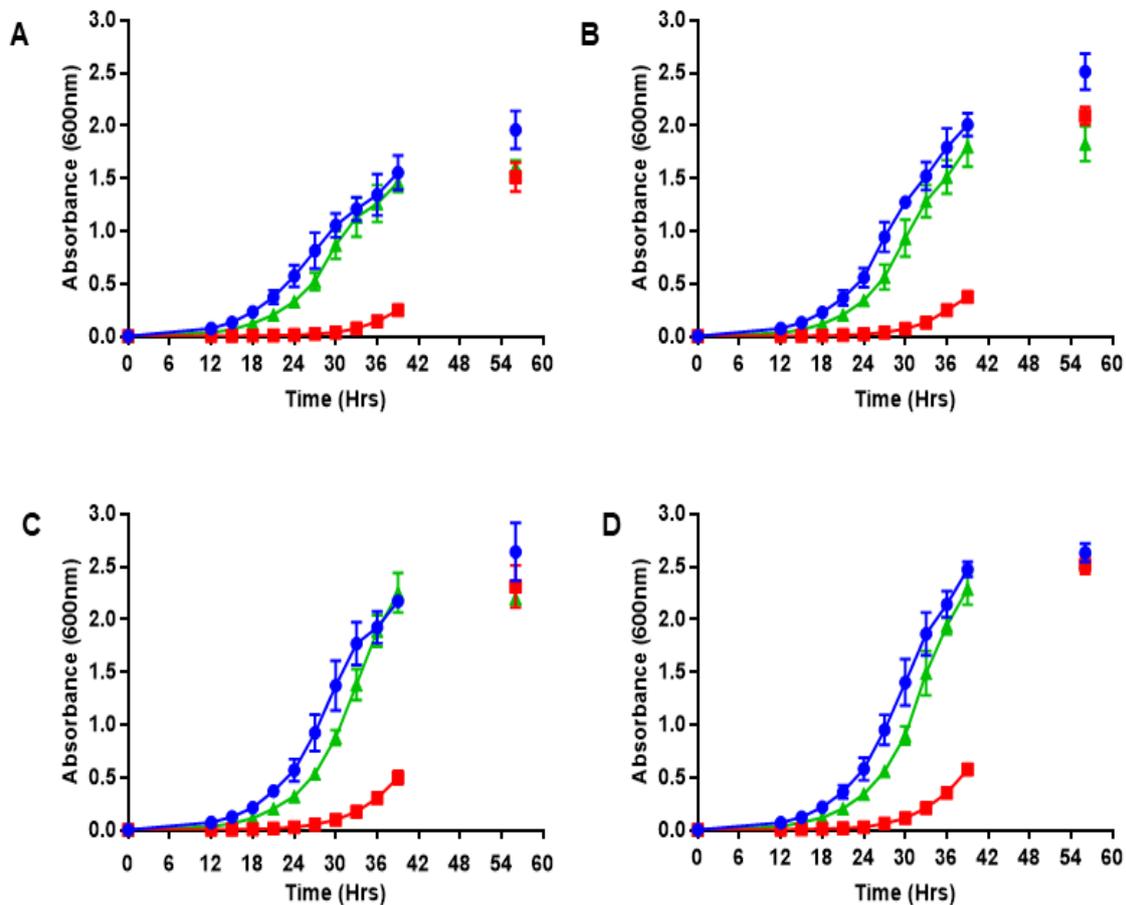


Figure 3. 13. Growth curves of *M. smegmatis* mc²155-1 (blue circles), Δ sufT (red squares) and Δ sufT (pSE3127) (green triangles) in varying concentrations (A) 0 μ M, (B) 0.1 μ M, (C) 0.5 μ M and (D) 2 μ M of supplemental iron. The unconnected points have missing time points. The data represents the average and standard deviation (shown by error bars) of three biological replicates.

To evaluate whether this extended lag phase represented a statistically significant difference in the growth cycle of the bacteria, regression analysis on the growth curve data was performed. All the curves fit the sigmoidal four-parameter logistic regression (4PL) model (Fig. 3.14). The data generated by the regression analysis was then used to compare the logIC50 of Δ sufT with that of mc²155-1 using the unpaired t-test. There was a statistically significant difference in logIC50 between the two strains for all the concentrations of supplemental iron (Fig. 3.15). The logIC50 in our data, represents the time to reach 50% maximum cell density for each strain. The Δ sufT strain took on average 1.5 times longer than the mc²155-1 strain to reach 50% maximum cell density across the range of iron concentrations (Fig. 3.14- 3.15). These results showed that loss of *sufT* had a significant impact on the time to complete a full

growth cycle in *M. smegmatis*, increasing it from 33 hrs in *mc*²155-1 in to 56 hrs in Δ *sufT*. The results of the complemented strain Δ *sufT* (pSE3127) were all comparable to *mc*²155-1 (Fig. 3.13- 3.15) confirming that the observed growth phenotype was due to loss of *sufT*.

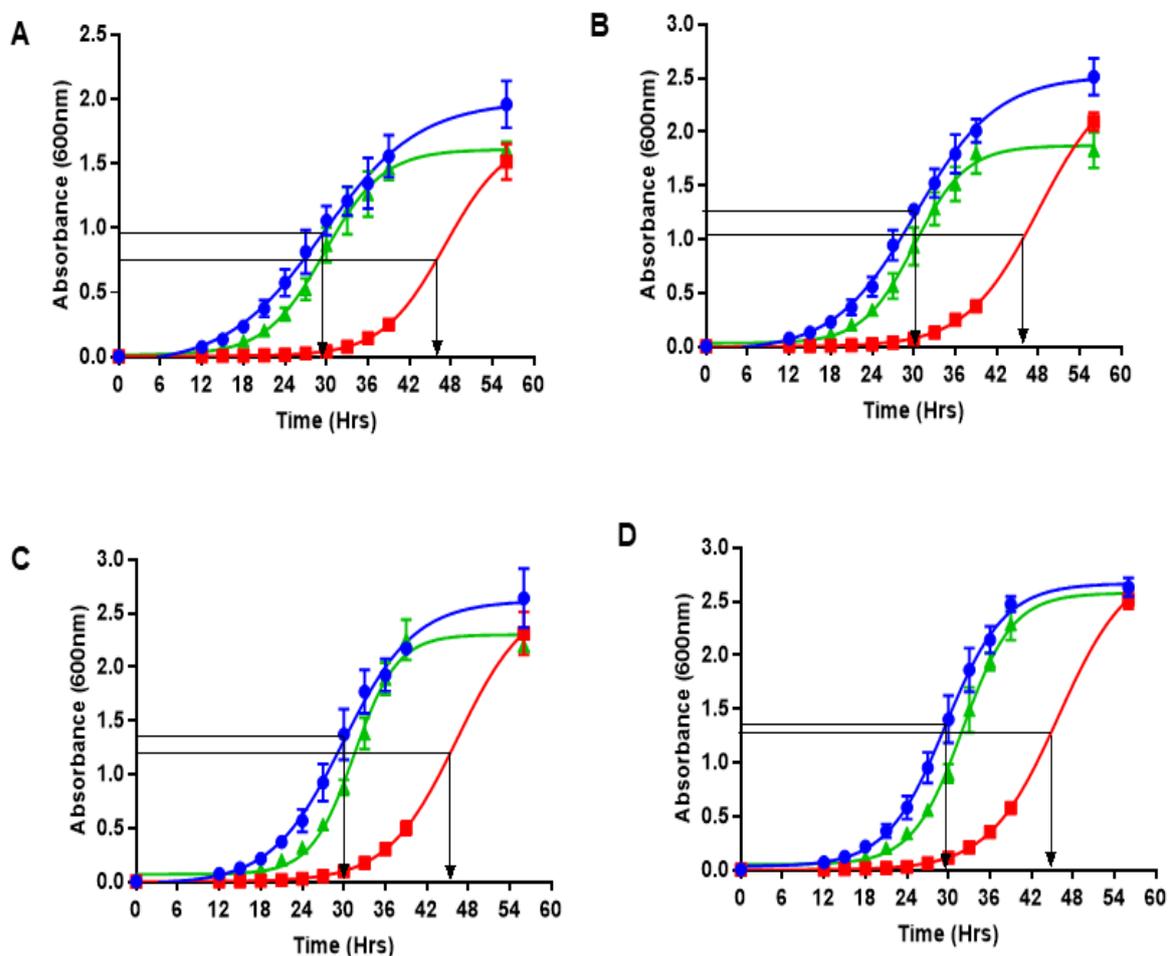


Figure 3. 14. Sigmoidal four-parameter logistic regression (4PL) models of the growth curves of *M. smegmatis* *mc*²155-1 (blue circles), Δ *sufT* (red squares) and Δ *sufT* (pSE3127) (green triangles) in varying concentrations (A) 0 μ M, (B) 0.1 μ M, (C) 0.5 μ M and (D) 2 μ M of supplemental iron. The regression model that best fit the data was generated using GraphPad Prism 7.03. The black arrows show the time (logIC₅₀) that *mc*²155-1 and Δ *sufT* took to reach 50% maximum cell density at the various concentrations of iron. The data represents the average and standard deviation (shown by error bars) of three biological replicates.

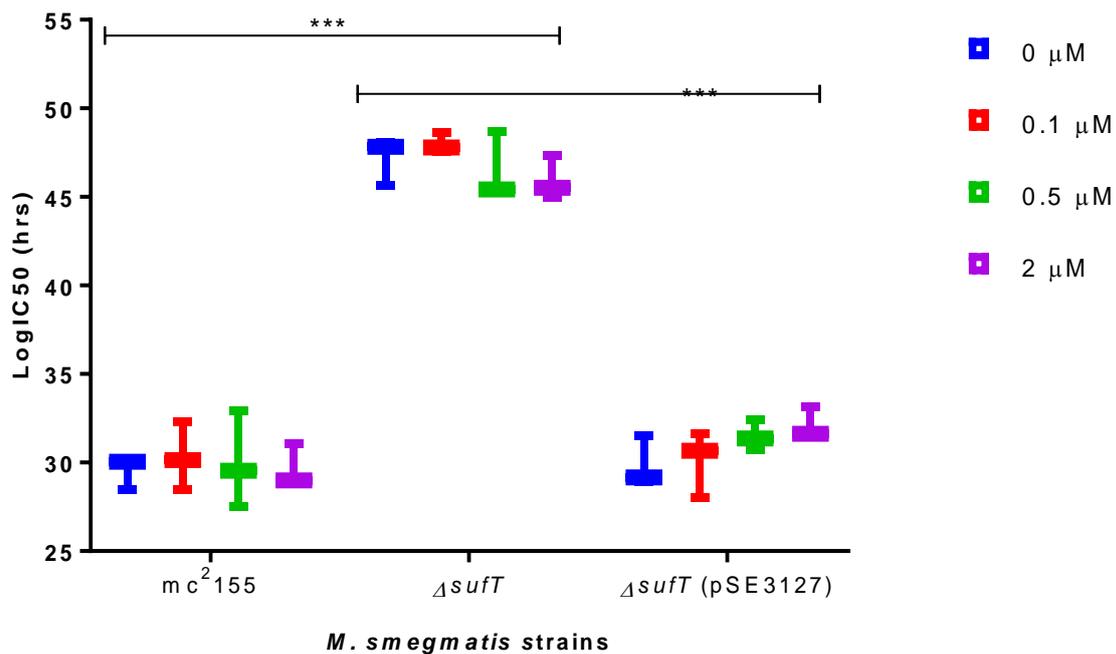


Figure 3. 15. Box and whisker plots of the logIC₅₀ of *M. smegmatis* mc²155-1, ΔsufT and ΔsufT (pSE3127) under iron limitation at varying concentrations of iron. The data represents the average and standard deviation (shown by error bars) of three biological replicates. ΔsufT and mc²155-1 took significantly different times to reach 50% maximum cell density in all the iron concentrations. The corresponding p-values were; 0 μM (p=0.00005), 0.1 μM (p=0.0001), 0.5 μM (p=0.001) and 2 μM (p=0.00009). P-values were determined by unpaired t-tests using GraphPad Prism software with statistical significance indicated by p<0.05 (*), p<0.01 (**) and p<0.001 (***) on the graph.

Interestingly, a positive correlation was observed between iron concentration and the maximum cell density of each strain at stationary phase. The 0 μM and 2 μM supplemental iron cultures had the lowest and highest maximum cell densities respectively for all three strains (Fig. 3.16). However, the addition of supplemental iron had a greater impact on the increase in maximum cell densities for the ΔsufT strain, where a significant increase was seen in all the cultures with supplemental iron when compared to the culture without any iron (Fig. 3.16B).

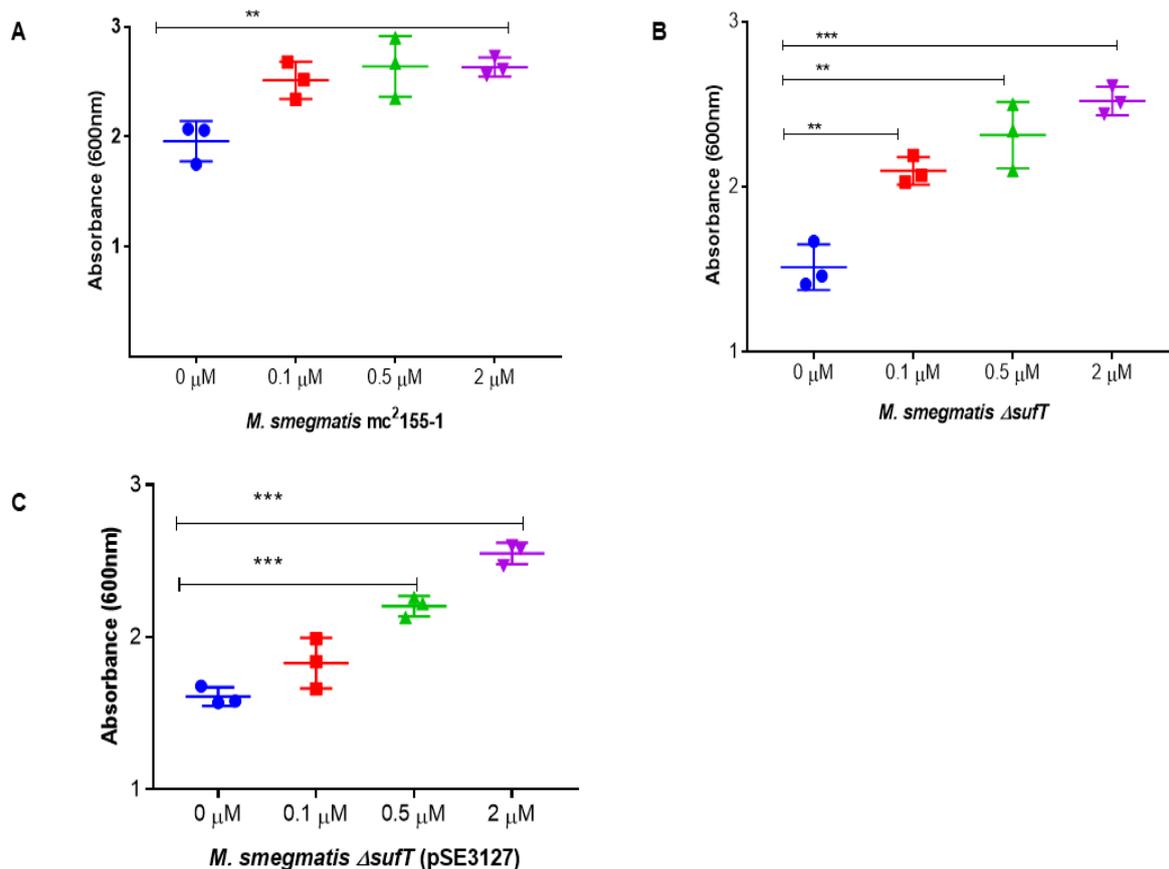


Figure 3. 16. Box and whisker plot showing the increase in maximum cell density by *M. smegmatis* (A) mc²155, (B) Δ*sufT* and (C) Δ*sufT* (pSE3127) as iron concentration is increased. The data represents the average and standard deviation (shown by error bars) of three biological replicates. (A). In mc²155-1, the maximum cell density increased significantly only when 2 μM iron was added (p=0.004). (B) In Δ*sufT* the maximum cell density increased significantly in all the cultures with supplemental iron when compared to the culture without iron with the following p-values; 0.1 μM (p=0.004), 0.5 μM (p=0.004) and 2 μM (p=0.0004). (C) In Δ*sufT* (pSE3127) the maximum cell density increased significantly only when 0.5 μM (p=0.0003) and 2 μM (p=0.00006) iron was added. P-values were determined by unpaired t-tests using GraphPad Prism software with statistical significance indicated by p<0.05 (*), p<0.01 (**) and p<0.001 (***) on the graph.

Even though the Δ*sufT* strain had an extended lag phase, it was still able to grow to comparable maximum cell densities as mc²155-1 at stationary phase. To investigate if this was a true phenotype or whether the strain had spontaneously mutated during the extended lag phase enabling it to enter log phase growth and grow to wild-type levels, the stationary phase cultures were used as inoculum and growth in MM without any supplemental iron was monitored. It was observed that during the second round of growth, the Δ*sufT* strain still exhibited the same growth phenotype in MM media without supplemental iron as had previously been observed (Fig. 3.17) confirming that the observed phenotype was the true

phenotype of the $\Delta sufT$ mutant strain in iron limiting conditions and not because of extra genomic mutations acquired because of culturing under iron limiting conditions.

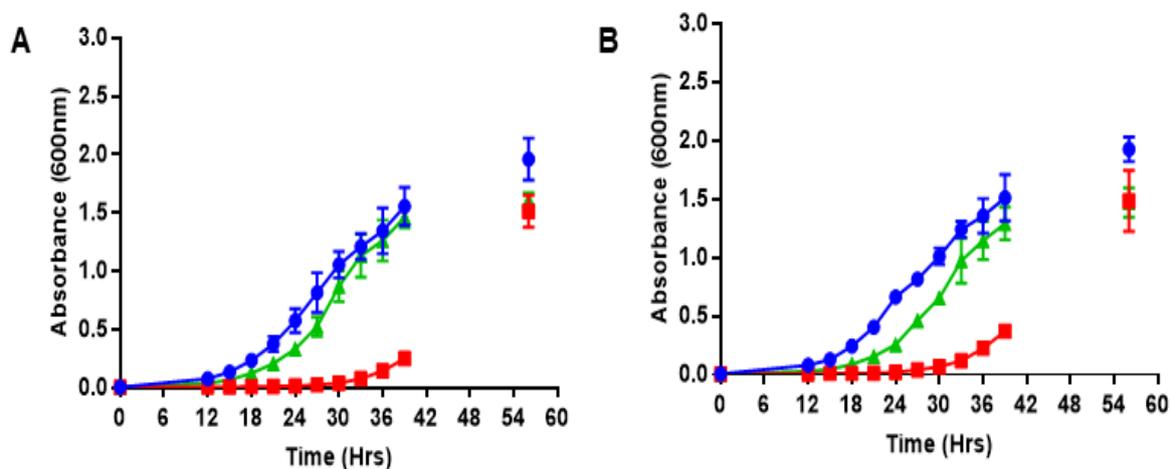


Figure 3. 17. Growth curves of *M. smegmatis* mc²¹⁵⁵ (blue circles), $\Delta sufT$ (red squares) and $\Delta sufT$ (pSE3127) (green triangles) after, (A) one growth cycle and (B) two growth cycles in MM with 0 μ M supplemental iron. The data represents the average and standard deviation (shown by error bars) of three biological replicates (A) and of two biological replicates (B).

These results confirm that while the deletion of SufT has no impact on growth under standard culture conditions, SufT plays a role in adaptation to growth under iron limiting conditions in *M. smegmatis*.

3.3.6 Standard culture (Sauton's media) pellicle biofilm formation

Biofilms have become synonymous with drug resistance in bacteria as they have been shown to harbour drug tolerant bacteria (Costerton et al., 1999; Ohja et al., 2008). Mycobacteria have a propensity to form biofilms, although the significance of these biofilms during infection is not understood (Ohja et al., 2008). Understanding biofilm formation in mycobacteria could be important to fully understanding drug resistance and pathogenicity in *M. tuberculosis*. In mycobacteria, iron metabolism has been linked to the formation of biofilms that grow on top of the media at the air liquid interface called pellicle biofilms (Ohja & Hatful, 2007). The formation of pellicle biofilms is a sequential process that starts with cell aggregation, then formation of a smooth extracellular matrix, then matrix surface attachment and finally maturation of the extracellular matrix to form a reticulated pellicle (Ohja et al., 2005). The impact of iron on the ability of the $\Delta sufT$ mutant to form pellicle biofilms will therefore be investigated.

3.3.6.1 Method optimisation

The growth of mycobacterial pellicle biofilms in tissue culture plates in Sauton's media has previously been described for *M. tuberculosis* (Kulka et al., 2012). The same protocol was optimised for the growth of *M. smegmatis* pellicle biofilms in 24 well plates using the wild-type mc²155-1 strain. The previous protocol specified a 1/10 head space ratio when growing biofilms in 12 well plates, therefore the biofilms were cultured in 2.4 ml in the 24 well plates. There was also need, to verify that biofilm growth was not dependent on well position, i.e. growth differences in outer wells versus inner wells of the plate due to increased drying of outer wells. Pellicle biofilms that grew evenly on the plate regardless of well position were cultured successfully (Fig. 3.18A). It was observed that mc²155-1 formed pellicle biofilms that attached to the wall of the well, then spread up the wall to the top of the well and had an extracellular matrix that formed a pattern of folds that seemed to converge at the centre of the biofilm (Fig. 3.18B).

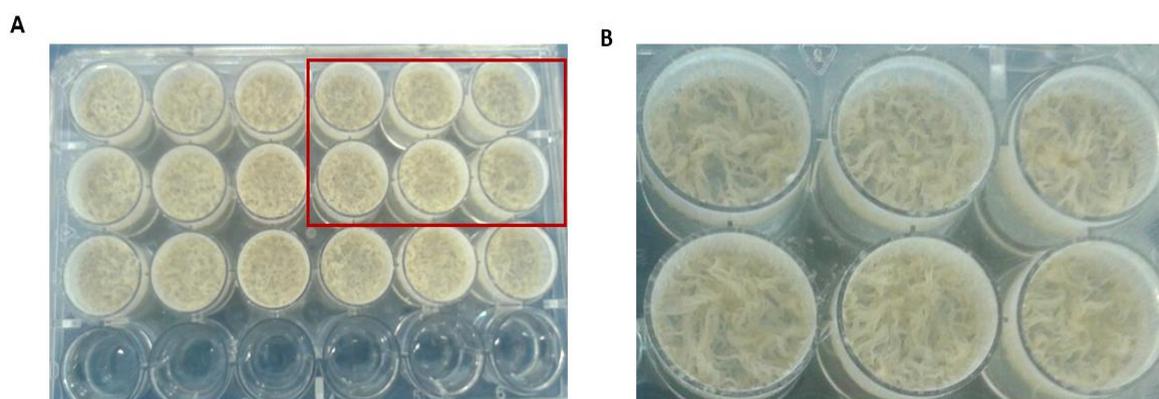


Figure 3. 18. Optimised growth of *M. smegmatis* mc²155-1 pellicle biofilms in 24 well plates. (A). Pellicle biofilms growing indiscriminately on the plate and the bottom wells without growth are the negative controls (Sauton's media) (B). Enlarged image of biofilms in the red box in (A) showing morphology of *M. smegmatis* mature pellicle biofilms typically characterised by the folded patterns formed by the extracellular matrix, the attachment and spreading of the biofilm up the walls of the plate. The displaced media can also be seen between the wells at the bottom of the plate.

The biofilm biomass was quantified indirectly by measuring the total protein for each well. Total protein is measured using the Bradford assay which is a colorimetric assay based on the binding of proteins to the dye Coomassie Blue G-250 causing a shift in dye absorbance from 465 to 595 nm. This shift in absorbance is directly proportional to the concentration of protein in the sample (Bradford, 1976). Triton X-100 is used to resuspend the biofilm biomass before quantification, however, high concentrations of detergents such as Triton X-100 produce excessive colour changes which interfere with the accuracy of the test (Bradford,

1976). To eliminate the inference of Triton X-100 with the assay, the final concentration of Triton X-100 in the samples was limited to 0.01%. It was observed that as the biofilms were forming at the air liquid interface, they were displacing some of the media in the plates making it difficult to calculate the volume of Triton X-100 needed to reach a final concentration of 0.01% in each sample. Since the spilling happened at the end stage of the biofilm formation process and did not affect the biomass, the volume below the biofilms was aspirated and the biomass resuspended in 1 ml of 0.01% Triton X-100.

3.3.6.2 Biofilm formation under standard culture conditions (Sauton's media)

The impact of loss of *sufT* on pellicle biofilm formation under standard culture conditions was subsequently investigated. It was observed that Δ *sufT* formed pellicle biofilms that were phenotypically indistinguishable from those of *mc*²155 and that they progressed through the stages of biofilm formation at the same rate and both formed an extracellular matrix with the characteristic folded pattern morphology after 5 days (Fig. 3.18A). Quantification of the biofilm biomass showed that there was no statistically significant difference in their biofilm biomasses either (Fig 3.18B). Taken together, these results confirmed that SufT is dispensable for biofilm formation under standard culture condition in Sauton's media.

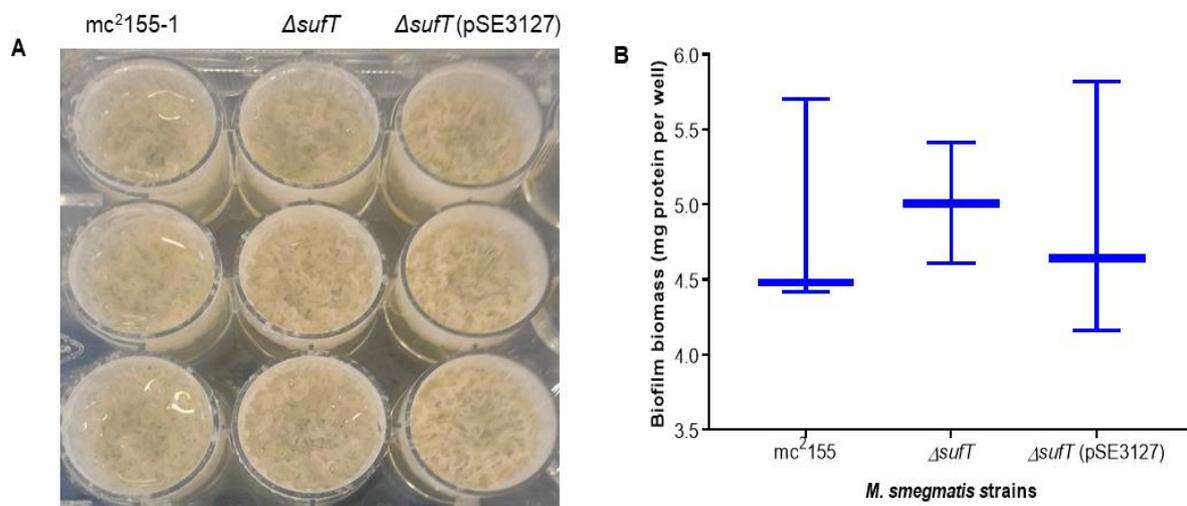


Figure 3. 19. Growth of pellicle biofilms in *M. smegmatis* *mc*²155-1, Δ *sufT* and Δ *sufT* (pSE3127) under standard culture conditions in Sauton's media. (A). A representative plate showing the morphology of the pellicle biofilms formed by the three strains. (B). Quantification of the pellicle biofilm biomass in *M. smegmatis* *mc*²155-1, Δ *sufT* and Δ *sufT* (pSE3127) under standard culture conditions. The data represents the average and standard deviation (shown by error bars) of three biological replicates. P-values were determined by unpaired t-tests using GraphPad Prism software with statistical significance indicated by p<0.05 (*), p<0.01 (**), and p<0.001 (***).

3.3.7 Iron limitation biofilm formation

Maturation of a pellicle biofilm in *M. smegmatis* was described previously as a transition from a smooth to a reticulated appearance with ridges and folds (Ohja et al., 2005; Ohja and Hatful, 2007). The study showed that a supplemental iron concentration of at least 2 μM is needed for the formation of mature pellicle biofilms in *M. smegmatis* (Ohja & Hatful, 2007). The impact of iron limitation on the ability of the ΔsufT mutant to form pellicle biofilms was investigated.

3.3.7.1 Impact of media components on pellicle biofilm formation

Growth under iron limitation is done in Chelex 100 resin treated mineral defined media (MM) which has a different composition to the standard culture Sauton's media. The MM is minimal in terms of minerals added, but it is a nutrient rich media with two carbon sources (glycerol and dextrose). In contrast, Sauton's has one carbon source (glycerol) and is rich in metals including ferric iron (Fe^{3+}). The culture of *M. smegmatis* pellicle biofilms in MM was therefore first optimised in mc²155-1, and the impact of media components on pellicle biofilm formation in *M. smegmatis* was also investigated. Both the Sauton's media and MM media supplemented with a high concentration of iron (50 μM) iron supported the growth of biofilms with a fully developed extracellular matrix after five days (Fig. 3.20a), with Sauton's media resulting in the greatest biomass (Fig. 3.20a, lane 1). In addition, sub-culturing the bacteria in an iron rich medium (7H9) prior to biofilm culture in MM media without supplemental iron did not impact the development of the extracellular matrix (Fig. 3.20a lane 2). This suggests that the iron present within the bacteria, and residual iron in the MM media, is sufficient to support biofilm formation. To investigate this further, the bacteria were first starved of iron by sub-culturing in MM without supplemental iron for 27 hrs prior to biofilm culture in MM. The bacteria starved of iron only formed a very thin biofilm which did not develop further to form an extracellular matrix (Fig. 3.20b, lane 1). These results confirm that biofilm maturation is halted below a certain intracellular iron threshold.

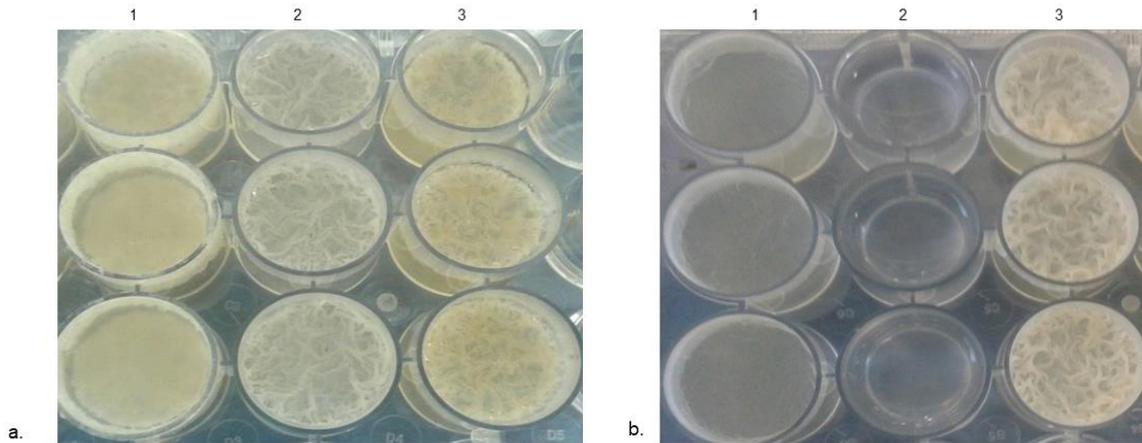


Figure 3. 20. Impact of media components on pellicle biofilm formation in *M. smegmatis* mc²155-1. (A) *M. smegmatis* mc²155-1 pellicle biofilms sub-cultured in 7H9 prior to growth in (1) Sauton's media, (2) chelex treated metal defined media (MM) without supplemental iron and (3) MM with 50 μ M supplemented iron. (B) *M. smegmatis* mc²155-1 pellicle biofilms sub-cultured in MM without supplemental iron prior to growth in (1) MM without supplemental iron (2) Negative control (media), and (3) MM media supplemented with 50 μ M iron

3.3.7.2 Biofilm formation under iron limiting conditions

To determine the iron concentration needed to initiate the maturation of pellicle biofilms in *M. smegmatis*, mc²155-1 was cultured in varying iron concentrations up to 2 μ M. After 5 days of growth, a macroscopic observation of the biofilms showed mature biofilms were not formed in less than 2 μ M iron and that progress through the steps of pellicle biofilm formation i.e. cell aggregation, matrix formation and matrix maturation were halted earlier when less iron was present (Fig 3.21). At 0 μ M the cells had aggregated at the air-liquid interface to form a very thin biofilm, at 0.1 μ M iron the biofilm started attaching to the walls of the wells and at 0.15 μ M iron it had fully attached and spread to the top of the well. At 0.2 μ M iron biofilm maturation was initiated (Fig 3.21 B3-4) and the biofilm thickened with increasing iron concentration and at 2 μ M iron the pellicle biofilm formed was comparable to standard culture biofilms. Since *M. smegmatis* displays no growth defect during planktonic culture at these iron concentrations, these results confirmed that iron impacted pellicle biofilm maturation in wild-type *M. smegmatis* independent of growth.

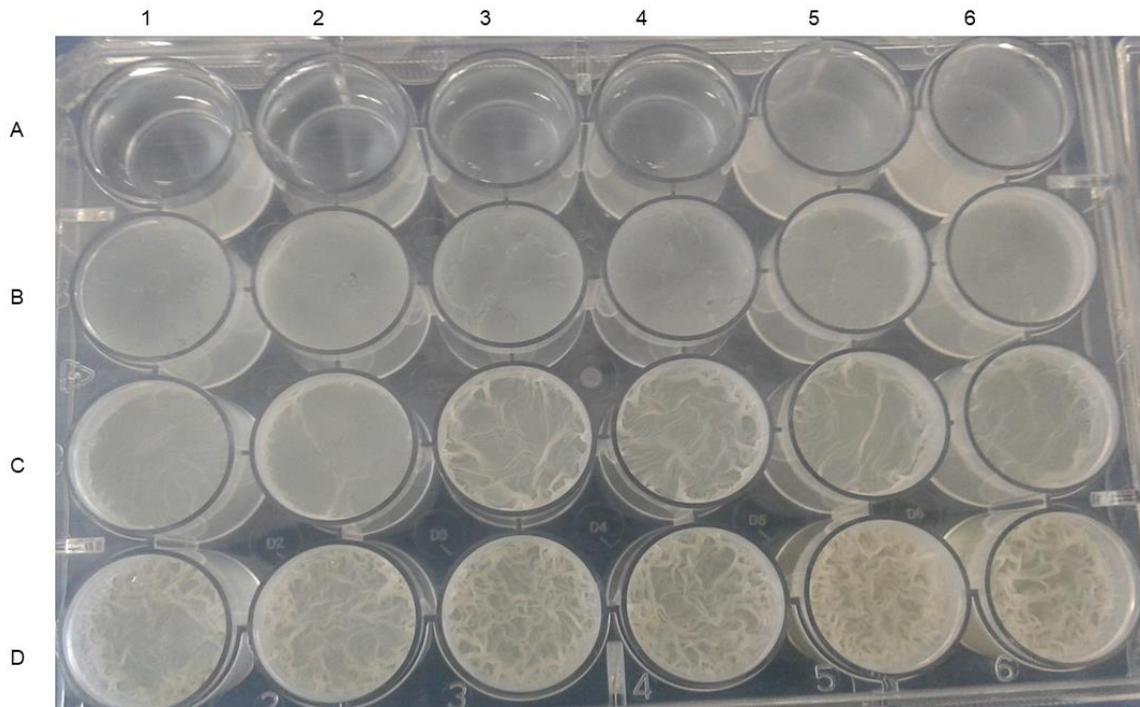


Figure 3. 21. Pellicle biofilm formation of *M. smegmatis* mc²155-1 in increasing iron (Fe³⁺) concentrations from left to right: (A1-2) 0 μ M; (A3-4) 0.05 μ M; (A5-6) 0.1 μ M; (B1-2) 0.15 μ M; (B3-4) 0.2 μ M; (B5-6) 0.25 μ M; (C1-2) 0.3 μ M; (C3-4) 0.4 μ M; (C5-6) 0.5 μ M; (D1-2) 0.75 μ M; (D3-4) 1 μ M; (D5-6) 2 μ M.

Since the Δ *sufT* mutant displayed a growth defect under iron limiting conditions, the impact of iron on the ability of the Δ *sufT* mutant to form pellicle biofilms was also investigated. After 5 days of growth, a macroscopic comparison of the biofilms showed that mc²155-1 and Δ *sufT* (Fig. 3.22) proceeded through the same stages of pellicle biofilm maturation. However, while iron concentration impacted biofilm formation in both strains, the effects were more severe in the Δ *sufT* strain which required more iron to progress through the stages of biofilm formation (Fig. 3.22). At iron concentrations less than 0.1 μ M, the Δ *sufT* strain formed a thin biofilm and most of the cells has settled at the bottom of the well. From 0.1 μ M to 0.25 μ M the smooth biofilm layer increased in thickness, but it did not attach to the walls of the wells and appeared brittle. At an iron concentration of 0.3 μ M, the pellicle started attaching to the walls of the wells and at 0.4 μ M it had fully attached to the walls and spread up to the top of the well. Unlike mc²155-1 which started maturation immediately after full wall attachment, Δ *sufT* had a lag phase where the biofilm continued to thicken but still maintained the smooth texture. Maturation was initiated at 0.75 μ M iron where the first ridges and folds appeared (Fig. 3.22 D1-2). The biofilm continued to thicken with increasing iron concentration and the 2 μ M supplemental iron biofilm had the thickest pellicle.

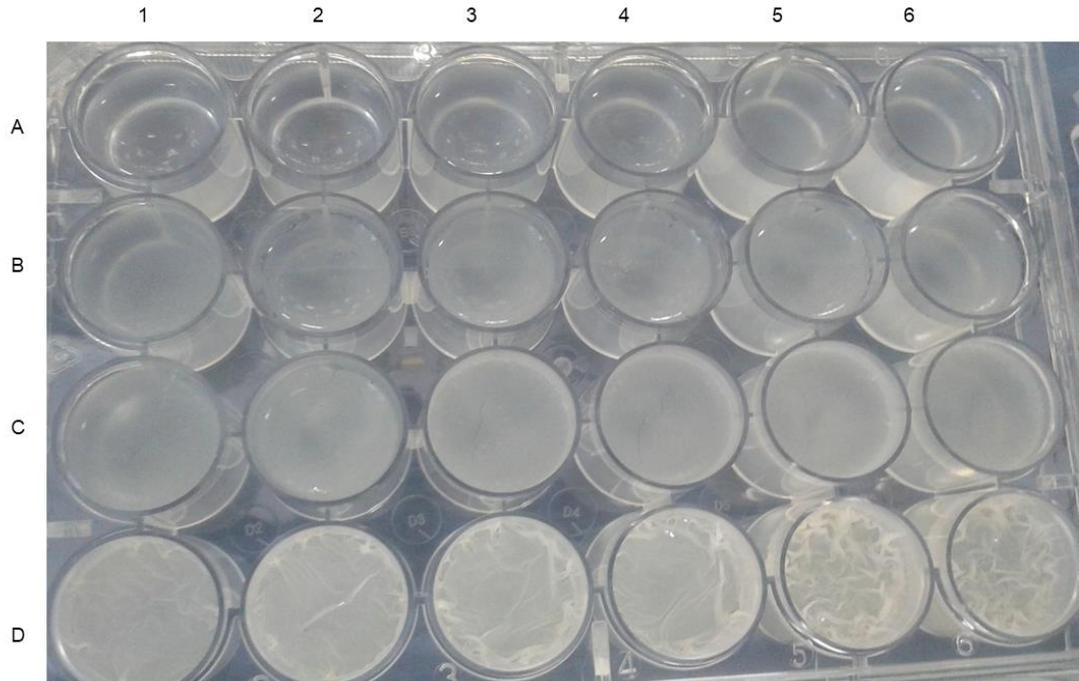


Figure 3. 22. Pellicle biofilm formation of *M. smegmatis* Δ *sufT* in increasing iron (Fe^{3+}) concentrations from left to right: (A1-2) 0 μM ; (A3-4) 0,05 μM ; (A5-6) 0,1 μM ; (B1-2) 0,15 μM ; (B3-4) 0,2 μM ; (B5-6) 0,25 μM ; (C1-2) 0,3 μM ; (C3-4) 0,4 μM ; (C5-6) 0,5 μM ; (D1-2) 0,75 μM ; (D3-4) 1 μM ; (D5-6) 2 μM .

The Δ *sufT* (pSE3127) strain also had the same maturation stages as *mc*²¹⁵⁵⁻¹ (Fig. 3.23). After 5 days of growth, at 0.1 μM iron the biofilm had already attached to walls and at 0.15 μM iron it had fully attached and spread to the top of the well. Like *mc*²¹⁵⁵⁻¹, Δ *sufT* (pSE3127) also started maturation at 0.2 μM iron (Fig. 3.23 B3-4). The 2 μM iron culture had the thickest biofilm after 5 days but the thickness was not comparable to standard culture biofilms.

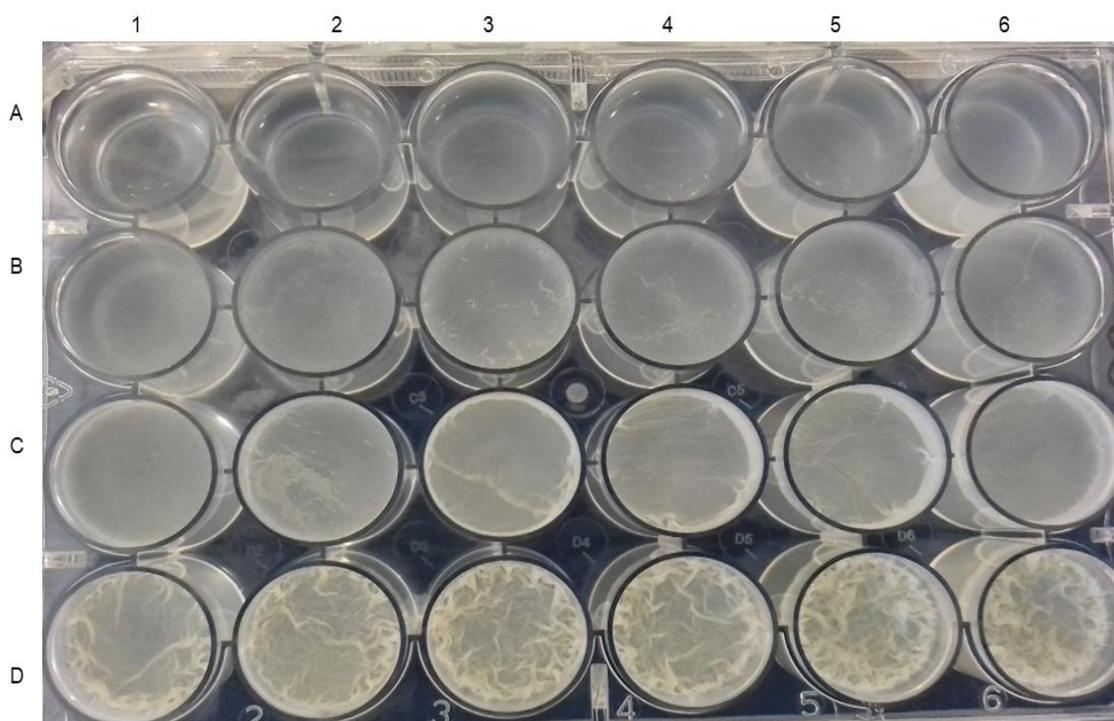


Figure 3. 23. Pellicle biofilm formation of *M. smegmatis* Δ *sufT* (pSE3127) in increasing iron (Fe^{3+}) concentrations from left to right: (A1-2) 0 μM ; (A3-4) 0,05 μM ; (A5-6) 0,1 μM ; (B1-2) 0,15 μM ; (B3-4) 0,2 μM ; (B5-6) 0,25 μM ; (C1-2) 0,3 μM ; (C3-4) 0,4 μM ; (C5-6) 0,5 μM ; (D1-2) 0,75 μM ; (D3-4) 1 μM ; (D5-6) 2 μM .

It was also observed macroscopically that the Δ *sufT* biofilms appeared to be less thick than the mc²155-1 biofilms at the same concentration of iron. To confirm this, the biomass in the three strains was quantified. Four representative iron concentrations (0, 0.1, 0.5 and 2 μM) were selected because they represented distinct stages in the maturation process in the mc²155-1 strain. These concentrations were also selected so that comparisons with planktonic growth could be made.

Since early stage biofilms do not have a well-formed matrix, they are fragile, and therefore the liquid below the biofilm could not be aspirated prior to biofilm quantitation. To overcome this, Triton X-100 was added directly to each well at a final concentration of 0,01% to solubilise the biofilm. This was feasible since no spillage was observed for biofilms grown in MM over 5 days. Using this method the biomass in 0 μM Fe and 0.1 μM Fe was comparable for all the strains, even though visually differences could be observed (Fig. 3.21-3.23). This discrepancy is presumably because planktonic cells were present in the media in the wells with early stage biofilms, and these increased the total biomass. Crystal violet staining was considered as an

alternative method, but it was not suitable for early stage biofilms either because the method also involves aspirating the liquid below the biofilm.

Macroscopic observation suggested that the biofilm biomass produced by the $\Delta sufT$ strain were lower than those of the mc^2155-1 strain at the same concentration of iron however, no statistically significant difference between the biomass formed by the two strains was observed. There was however a statistically significant positive correlation between iron concentration and quantity of biofilm biomass when comparing the biomass formed without supplemental iron to those formed in supplemental iron (Fig. 3.24).

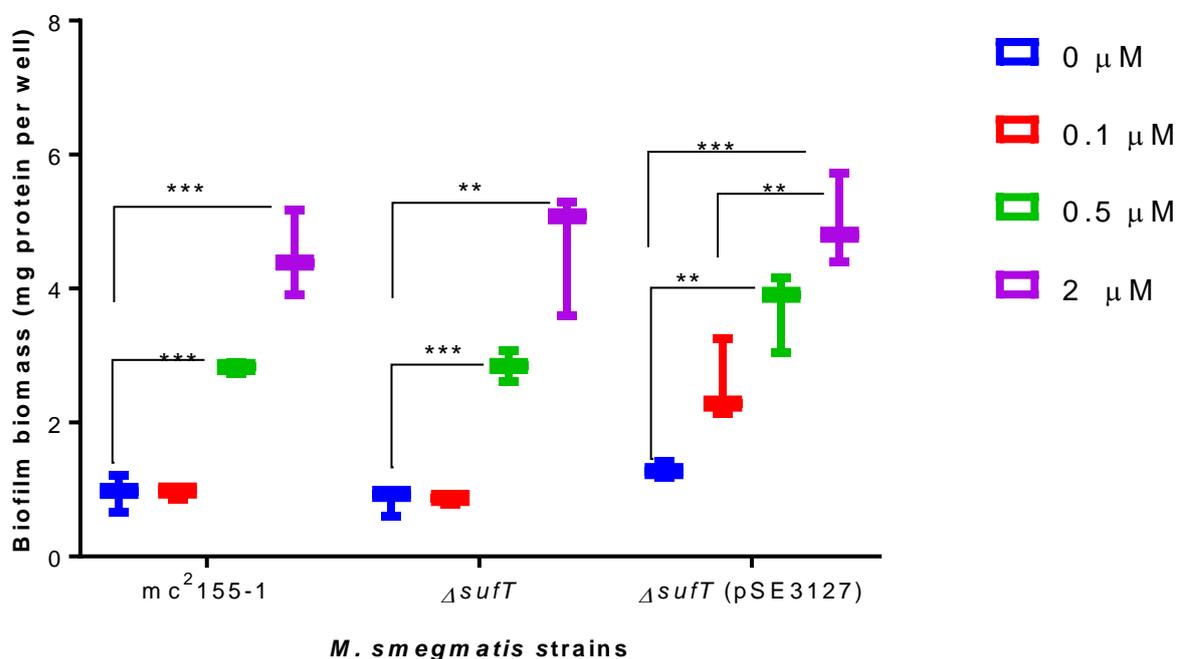


Figure 3. 24. Quantification of the biofilm biomass in *M. smegmatis* mc^2155-1 , $\Delta sufT$ and $\Delta sufT$ (pSE3127) under iron limitation with varying concentrations of supplemental iron. Total protein (mg) was used as an indicator of biomass. The data represents the average and standard deviation (shown by error bars) of three biological replicates. P-values were determined by unpaired t-tests using GraphPad Prism 7.03 with statistically significant results indicated by $p < 0.05$ (*), $p < 0.01$ (**) and $p < 0.001$ (***).

3.3.4.3 Growth of pellicle biofilms under iron limitation in a mycosin 3 mutant

To better understand how iron was influencing biofilm maturation, we investigated the impact of loss of the iron acquisition protein mycosin 3 (encoded by the gene *mycP₃*) on biofilm formation. Mycosin 3 was previously shown to be dispensable for planktonic growth in iron limiting conditions (Fang et al., 2017) and growth in MM without supplemental iron confirmed

that *MycP*₃ was dispensable for planktonic growth under iron limitation in *M. smegmatis* (Fig. 3.25). Since mycosin 3 is involved in iron metabolism but loss of *mycP*₃ does not cause a growth defect in planktonic growth under iron limitation (Fig. 3.25), the $\Delta MycP_{3ms}$ mutant provided a good model for elucidating a mechanism through which iron metabolism could be influencing biofilm maturation in *M. smegmatis*.

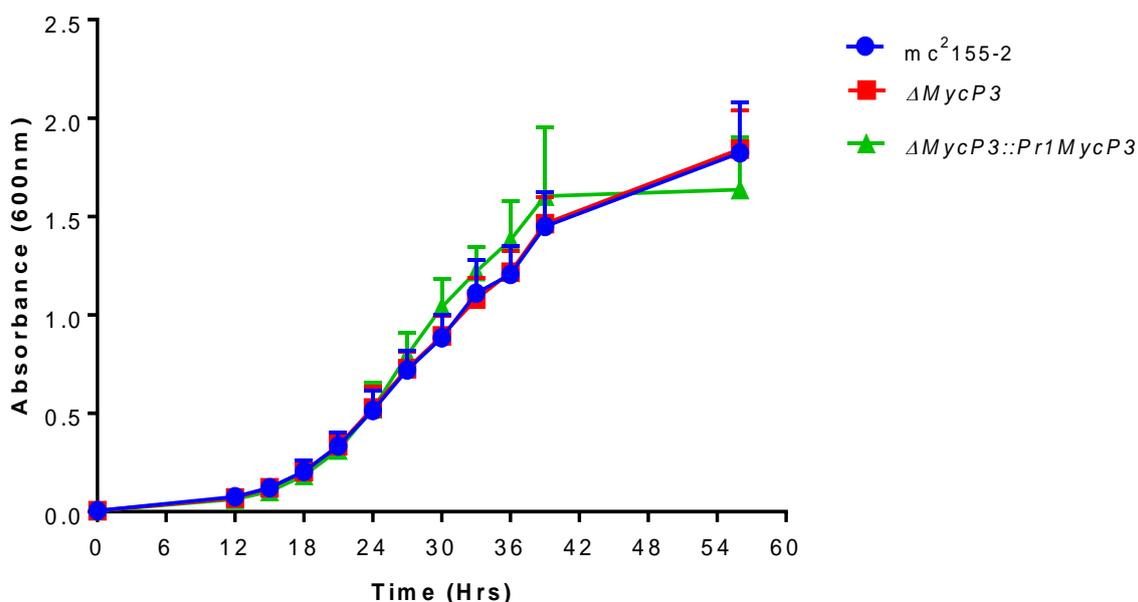


Figure 3. 25. Growth curve of *M. smegmatis* mc²155-2 (blue circles), $\Delta MycP_{3ms}$ (red squares) and $\Delta MycP_{3ms}::Pr1MycP_{3ms}$ (green triangles) under iron limitation without the addition of supplemental iron. The data represents the average and standard deviation (shown by error bars) of three biological replicates.

The impact of loss of *mycosin 3* on the ability of *M. smegmatis* to form pellicle biofilms was therefore investigated. After 5 days of growth, macroscopic observation showed that mc²155-2 (the mc²155 strain used to generate the $\Delta MycP_{3ms}$) and $\Delta MycP_{3ms}$ formed pellicle biofilms that were macroscopically indistinguishable at the different iron concentrations (Fig. 3.26). The pellicle biofilm formation of $\Delta MycP_{3ms}$ under iron limitation mirrored that of the mc²155-2 strain, confirming that $\Delta MycP_{3ms}$ was dispensable for both planktonic growth and biofilm formation in *M. smegmatis*.

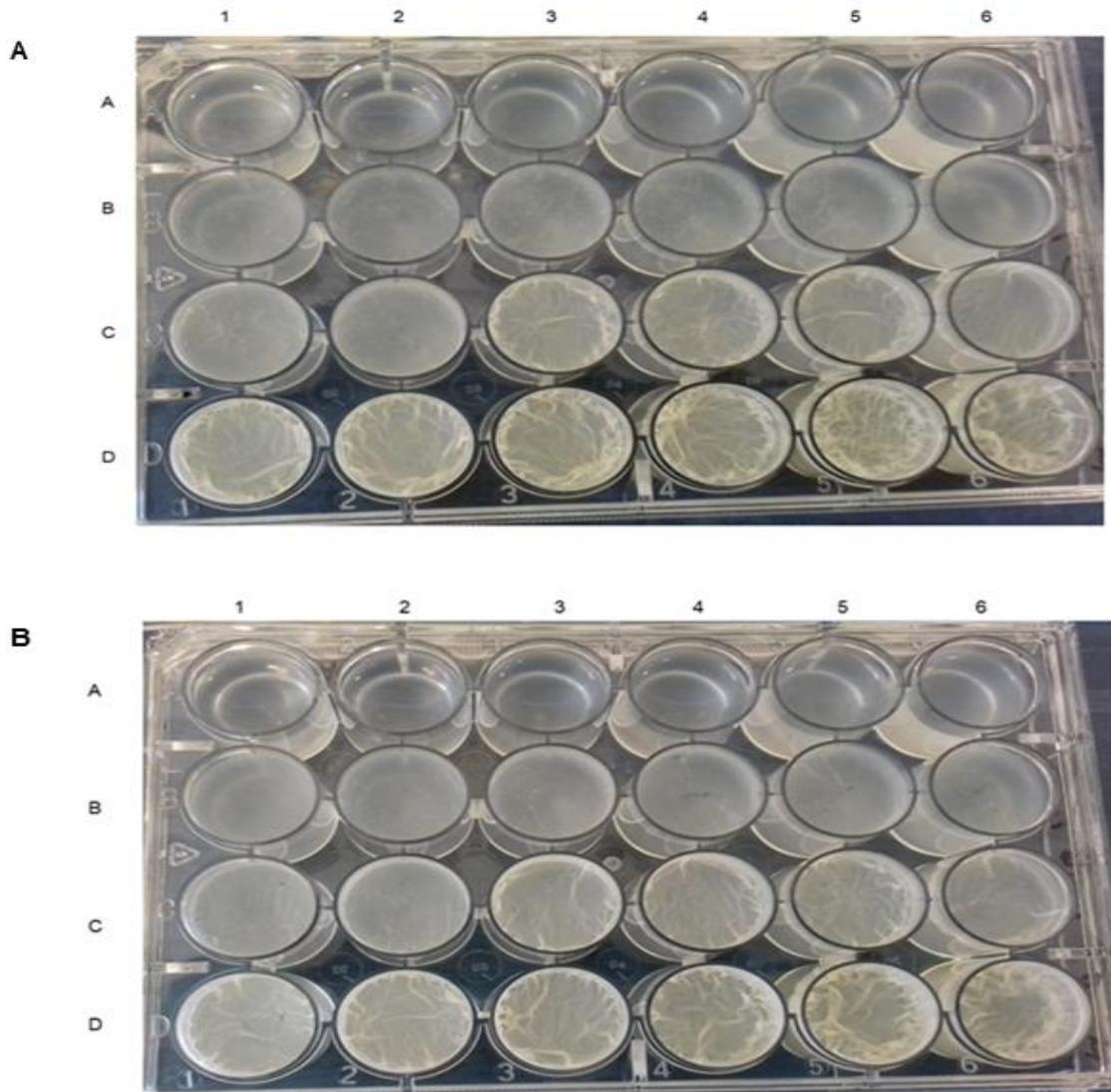


Figure 3. 26. Pellicle biofilm formation of *M. smegmatis* (A) $\Delta MycP3ms$ (B) mc^2155-2 in increasing iron (Fe^{3+}) concentrations from left to right: (A1-2) 0 μM ; (A3-4) 0,05 μM ; (A5-6) 0,1 μM ; (B1-2) 0,15 μM ; (B3-4) 0,2 μM ; (B5-6) 0,25 μM ; (C1-2) 0,3 μM ; (C3-4) 0,4 μM ; (C5-6) 0,5 μM ; (D1-2) 0,75 μM ; (D3-4) 1 μM ; (D5-6) 2 μM .

Chapter 4: Discussion and Conclusion

4.1 Discussion

Proteins drive cellular processes and each protein has (a) specific role(s) to play in an organisms' metabolism (Bateman et al., 2010, Goodacre et al., 2013). Some proteins require co-factors to become functionally active (Lill & Mühlenhoff, 2006), and Fe-S clusters are co-factors to proteins involved in an array of fundamental cellular processes such as respiration and DNA repair (Castro et al., 2008; Py et al., 2011). Fe-S clusters do not form spontaneously in the cell, instead they are synthesised *de novo* in a controlled way using complex assembly systems (Lill & Mühlenhoff, 2006; Fontcave & Ollagnier-de-Choudens, 2007; Blanc et al., 2014). The *sufR-sufB-sufD-sufC-csd-sufU-sufT* operon (*suf* operon) encodes the primary Fe-S biogenesis pathway in mycobacteria and the *suf* operon is essential to the survival of *M. tuberculosis* (Huet et al., 2005). Understanding the role of each protein in the operon is crucial to fully understanding how it contributes to pathogenesis (Outten, 2015). This study used targeted mutant generation and phenotypic characterisation to investigate the role of SufT in the physiology of mycobacteria using *M. smegmatis* as a model organism.

An *M. smegmatis* Δ *sufT* knockout mutant strain harbouring an unmarked deletion in *sufT* was successfully generated using allelic exchange mutagenesis (Fig. 3.6). The fact that a Δ *sufT* knockout mutant could be recovered suggested that *sufT* was not essential for growth under standard aerobic culture conditions. Forward genetic screens have made contradictory essentiality predictions for *sufT* in mycobacteria. Transposon mutagenesis using microarray hybridisation (TraSH) initially predicted *sufT* to be an essential gene in *M. tuberculosis* (Sassetti et al., 2003). However, a subsequent transposon mutagenesis study using next generation sequencing (Tn-seq) predicted *sufT* to be a non-essential gene (Griffin et al., 2011), and the non-essentiality prediction was supported by data generated using CRISPRi in *M. smegmatis* (de Wet et al., 2018). TraSH had discordant results possibly it lacks sensitivity to detect low mutant abundance associated with growth defects resulting in miss-classification as essential (de Wet et al., 2018).

Transposon mutagenesis and CRISPRi screening is specific to the growth condition that was tested (Sassetti et al., 2003). Targeted gene deletion is required to confirm essentiality predictions made by the forward genetic screens and has the advantage of having capacity to test essentiality under multiple growth conditions. The results from this study therefore confirmed previous findings (Griffin et al., 2011; de Wet et al., 2018) that *sufT* is a non-essential gene in mycobacteria. In addition, loss of SufT did not affect planktonic growth under

standard culture, a further confirmation that SufT is dispensable for growth under these conditions. This is consistent with previous studies showing that SufT is dispensable for growth under standard culture conditions in *S. aureus* (Mashruwala et al., 2016a) and *S. meliloti* (Sasaki et al., 2016). Essentiality in *M. smegmatis* cannot always be directly inferred to *M. tuberculosis* without factoring in physiologic relevance. This is because even though *M. smegmatis* is a surrogate for *M. tuberculosis*, their genomes differ in size because they occupy different environmental niches. *M. tuberculosis* has undergone genome down-sizing to become a specialised pathogen whereas *M. smegmatis* has retained a larger genome which allows it to survive in the soil. Therefore, any essentiality predictions made in *M. smegmatis* need to be confirmed in *M. tuberculosis*.

Based on sequence homology and the ability of the *M. tuberculosis* *sufT* to complement loss of *sufT* in *S. aureus*, it was inferred that *M. smegmatis* *sufT* would have a similar function as *S. aureus*. In *S. aureus* SufT is postulated to be an accessory protein for the maturation of Fe-S clusters under conditions of high Fe-S demand (Mashruwala et al., 2016a; Mashruwala et al., 2016b). To determine whether mycobacterial SufT was also involved in Fe-S biogenesis, the effect of loss of SufT on the formation of Fe-S clusters was investigated by measuring the activities of the Fe-S containing enzymes succinate dehydrogenase (SDH) and aconitase (ACN). Loss of SufT significantly affected SDH activity (Fig. 3.10) but did not impact ACN activity (Fig 3.11). The negative impact on SDH activity was unexpected because there was no observed growth defect in planktonic cultures under standard aerobic conditions, suggesting that SufT was not needed for Fe-S biogenesis under those conditions.

In *S. aureus* cells lacking SufT were shown to have decreased activity of Fe-S containing enzymes when entering respiratory growth, suggesting increased demand for Fe-S clusters during respiratory growth (Mashruwala et al., 2016a). In this study, enzyme activity was measured in cells at mid-log growth which is characterised by optimal cellular metabolism. Optimal cellular metabolism increases the demand for Fe-S clusters because fundamental cell processes like aerobic respiration are catalysed by enzymes that require Fe-S clusters (Cheng et al., 2006). SDH is involved in both the tricarboxylic acid (TCA) cycle and electron transport chain and requires three Fe-S clusters (2Fe-2S, 3Fe-4S and 4Fe-4S) (Cheng et al., 2006), while ACN is involved in the TCA cycle and needs a single 4Fe-4S cluster (Djaman et al., 2004). We could therefore speculate that as a putative Fe-S maturation protein, SufT is needed for the maturation of Fe-S clusters targeted for SDH and not for ACN. The specificity for target proteins by maturation proteins has previously been shown in SufT homologues in eukaryotes where Fam96a and Fam96b both functioned as CIA maturation proteins, but with distinct target apoproteins (Stehling et al., 2013). To determine whether SufT functions as a

maturation protein, its' ability to bind Fe-S clusters will need to be determined. This would involve the expression and purification of recombinant SufT protein in a heterologous host and performing Fe-S cluster reconstitutions on SufT or the ability of SufT to transfer clusters to other proteins. The specificity of SufT for SDH maturation can be tested using the preferential use model (Mashruwala et al., 2016b) and measuring the activity of SDH under condition of low demand for Fe-S clusters (fermentative growth) and high demand for Fe-S clusters (aerobic respiration) to see if loss of SufT has an impact on enzyme activity. In addition, co-immunoprecipitation assays can be done to determine if there are protein-protein interactions between SufT and SDH.

Fe-S clusters are susceptible to damage by reactive oxygen species (ROS), (Varghese et al., 2003; Djaman et al., 2004; Imlay, 2006) and since it was proposed that SufT was involved in *de novo* synthesis of Fe-S clusters as a maturation protein, the function of SufT in Fe-S biogenesis under oxidative stress was investigated. Loss of SufT did not impact survival post exposure to DMNQ (Fig. 3.12) and these results suggested that either SufT was not involved in the repair or protection of Fe-S clusters from ROS or that the impact of ROS on Fe-S cluster containing enzymes did not cause a growth defect in cells experiencing oxidative stress. In addition, it is proteins like aconitase which have 4Fe-4S that are most affected by oxidative stress because their clusters are solvent exposed whereas in SDH the Fe-S clusters are protected from damage by oxidative stress (Djaman et al., 2004). SufT's role may be limited to maturation of proteins not affected by oxidative stress. One of the limitations of this experiments is that the colony forming units (CFUs) based-method used to enumerate percentage survival was not very accurate due to clumping of the cells upon exposure to DMNQ. For future studies, a method based on determining total protein might be a more accurate measure of the number of bacteria present under oxidative stress.

Loss of SufT did not impact the minimum inhibitory concentration (MIC) of the anti-tuberculosis drugs clofazimine (CFZ), isoniazid (INH) and rifampicin (RIF). One of the limitations of the broth microdilution method is that the sensitivity is determined by the range of the 2-fold dilution where growth is inhibited e.g. 32-64 µg/ml for INH and 1-4 µg/ml for CFZ. To refine these ranges, it may be useful to test intermediate concentrations, which may reveal differences. Since CFZ is postulated to kill by generating bactericidal levels of ROS (Yano et al., 2011) and loss of SufT did not impact growth under oxidative stress, this result was not unexpected. The activation of INH by the catalase peroxidase KatG is iron dependent and since SufT is proposed to play a role in Fe-S biogenesis and changes in Fe-S biogenesis can affect iron metabolism, the impact of loss of SufT on INH activity was investigated. However, since loss of SufT did not cause a growth defect in iron replete conditions, it may not impact iron homeostasis under these conditions. In future studies, the impact of loss of SufT on the

activity of INH under iron limitation needs to be investigated. RIF kills mycobacteria by inhibiting the transcription of bacterial DNA (Wehrli, 1983). Loss of SufT had no impact on the RIF MIC. DUF59 containing proteins in eukaryotes have been linked with Fe-S biogenesis targeted for DNA repair proteins (Luo et al., 2012), it would be interesting to see if loss of SufT influences DNA repair in mycobacteria. This could be done by inducing DNA damage using ultra violet (UV) then plating on RIF to determine the rate of spontaneous RIF resistance. The mutation rate will be used as an indirect measure of efficient DNA repair.

SufT was confirmed as non-essential under standard aerobic culture however, essentiality is condition-specific because bacteria require different proteins under different growth conditions (Caglar et al., 2017; Rodriguez et al., 2002; Sassetti et al., 2002). The expression of *sufT* in *M. tuberculosis* is upregulated under iron limitation suggesting a role for SufT in iron metabolism (Rodriguez et al., 2002), therefore the impact of the loss of SufT on planktonic growth under iron limitation was evaluated. Loss of SufT caused a growth defect in *M. smegmatis* growing under iron limiting conditions, but the growth defect was only observed during the lag phase. Interestingly, adding increasing concentrations of iron (up to 2 μ M) did not correct the growth defect (Fig. 3.13), suggesting that SufT was needed for adaptation to growth under iron limitation.

However, the bacteria eventually entered log growth and grew to comparable cell densities (Fig. 3.13). A possible explanation for this observation was that the bacteria had acquired spontaneous mutations however, the same growth phenotype was observed in subsequent growth cycles ruling out this possibility (Fig 3.17). Another possible explanation is that induction of a protein that compensates for the loss of *sufT* occurs, allowing the bacteria to adapt to the iron limiting conditions and grow like the wild type. This explanation is plausible if there is a protein in *M. smegmatis* that is a functional homologue of SufT, since functional homologues have redundant functions and can compensate for loss of each other (Vinella et al., 2009). Proteins from the *nif* system have been shown to be share functional redundancy with *suf* proteins (Yokoyama et al., 2018) and in *S. aureus* Nfu was able to compensate for loss of SufT (Mashruwala et al., 2016b). Additionally, Nfu in *S. aureus* was confirmed to be an Fe-S carrier that delivers Fe-S clusters to target apoproteins (Mashruwala et al., 2015). The *suf* operon of mycobacteria has a putative NifU-like protein immediately upstream of *sufT*, namely *sufU* (Fig 1.5). The protein encoded by *sufU* could be compensating for loss of SufT in planktonic growth under iron limiting conditions. However, the function of the NifU-like protein in mycobacteria has not been characterised and its ability to compensate for the loss of SufT would need to be confirmed. This could be done by complementing the Δ *sufT* mutant with *sufU* to see if the phenotype can be corrected.

This study confirmed previous findings that the iron needs of the cell differ significantly during planktonic growth and biofilm formation (Ohja & Hatful, 2007). The intracellular demand for iron during biofilm formation far exceeds that for planktonic growth, particularly during the maturation of the biofilms to form an extracellular matrix (Ohja et al., 2005, Ohja & Hatful, 2007). The extracellular matrix of mycobacteria is postulated to be made up of long chain (C₅₆ – C₆₈) mycolic acids (Ohja et al., 2005; Ohja et al., 2008). The synthesis of these long chain mycolic acids is regulated by the heat shock protein GroEL1 (Ohja et al., 2005) using the fatty acid synthase II (FASII) pathway (Schaeffer et al., 2001) and catalysed by β -ketoacyl-ACP synthase (KasA) (Ohja et al., 2005). The synthesis of these mycolic acids places a high energy demand on the cell (Cook et al., 2014; Jamet et al., 2015). Energy is produced by respiration and oxidative phosphorylation and the enzymes that catalyse these processes such as citrate synthase, isocitrate dehydrogenase and succinate dehydrogenase are Fe-S proteins (Djaman et al., 2004). If the synthesis of long chain mycolic acids needed for maturation of the extracellular matrix places a high demand for Fe-S clusters on the cell, this could explain why iron limitation inhibits the formation of an extracellular matrix in *mc*²¹⁵⁵⁻¹ (Fig. 3.21). Accordingly, the biofilm maturation growth defect seen in the Δ *sufT* mutant (Fig. 3.22) is plausible if loss of SufT is impacting the biogenesis of Fe-S clusters for proteins involved in vital metabolic pathways, resulting in a growth defect. A similar role was proposed for SufT in *S. meliloti* during symbiosis under iron limiting conditions because symbiosis also increases the intracellular demand for Fe-S proteins (Sasaki et al., 2016). The researchers hypothesized that *sufT* was needed for the biogenesis of Fe-S clusters under iron limitation making it indispensable for symbiosis (Sasaki et al., 2016). The impact of iron limitation on Fe-S biogenesis will need to be investigated further, to determine if there is a link between Fe-S biogenesis and formation of long chain mycolic acids in mycobacteria under these conditions.

Mycosin 3 is an iron uptake protein, and iron limitation did not cause a biofilm growth defect in the mycosin 3 mutant. During biofilm formation there is increased expression of iron uptake genes, which is postulated to be an adaptive response driven by an increased metabolic demand for iron (Ohja & Hatful, 2007; Yang et al., 2017). It would follow then that if iron uptake was impaired, then increasing iron would not impact biofilm maturation because the cell would not be able to access the iron, as a result Fe-S homeostasis would be impaired and mycolic acids synthesis would be inhibited. However, the results in this study support previous findings that loss of mycosin 3 does not impact the iron uptake ability of *M. smegmatis*, presumably because there are other iron uptake proteins that are compensating for loss of mycosin 3 (Fang et al., 2017). The impact of loss of iron acquisition on biofilm maturation will need to be investigated further.

Quantifying the biofilm biomass was challenging. The method used was not suitable for pre-maturation biofilms causing discrepancies between the macroscopic observations and the biomass quantification. In addition, quantifying the biomass did not add value especially when trying to distinguish small changes in the biofilm such as moving from aggregation to surface attachment. Since the interest in studying biofilms lies in investigating the physical changes that are happening to the bacteria and the extracellular matrix during the biofilm formation steps, using another qualitative method would potentially provide a better understanding of the pellicle biofilm process in mycobacteria. In this study, the biofilms were examined macroscopically but perhaps microscopic observation of the biofilms using a combination of scanning electron microscopy (SEM) and transmission electron microscopy (TEM) can elucidate better the physical changes happening to the biofilm at the different stages of growth. Future studies will therefore involve use of SEM and TEM.

4.2 Conclusion

Fe-S biogenesis is a vital process in cellular physiology yet the functioning of the primary Fe-S biogenesis machinery in mycobacteria is not fully understood. The last gene in the mycobacterial *suf* operon encodes a SufT homologue. SufT was predicted to be involved in the maturation of Fe-S clusters even though its role in mycobacteria has not been determined. In this study the role of SufT in mycobacterial physiology and Fe-S cluster biogenesis was investigated by creating a Δ *sufT* mutant. Deletion of *sufT* did not affect growth under standard culture conditions but affected adaptation of *M. smegmatis* to iron limitation. Loss of SufT affected the activity of SDH but not ACN, indicating a level of specificity in its putative role in Fe-S cluster maturation. Survival upon exposure to ROS and ROS generating drugs was not affected by the loss of SufT, suggesting that the impact of ROS on Fe-S containing proteins does not cause a growth defect under the conditions tested. Deletion of SufT did not affect the formation of biofilms under standard culture conditions, but a growth defect was seen under iron limitation possibly due to the impact of impaired Fe-S biogenesis on bacterial growth. Mycobacterial SufT therefore seems to function analogously to the cyanobacterial SufT and even though the findings in this study cannot definitively assign a function to SufT, they will guide in the accurate design of future experiments.

References

- Akaike, H., 2011. Akaike's Information Criterion, in: Lovric, M. (Ed.), *International Encyclopedia of Statistical Science*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 25–25. https://doi.org/10.1007/978-3-642-04898-2_110
- Akinola, R.O., Maz, G.K., u, Mulder, N.J., 2013. A Systems Level Comparison of *Mycobacterium tuberculosis*, *Mycobacterium leprae* and *Mycobacterium smegmatis* Based on Functional Interaction Network Analysis. *Journal of Bacteriology & Parasitology* 4, 1–11. <https://doi.org/10.4172/2155-9597.1000173>
- Alland, D., Steyn, A.J., Weisbrod, T., Aldrich, K., Jacobs, W.R., 2000. Characterization of the *Mycobacterium tuberculosis* iniBAC promoter, a promoter that responds to cell wall biosynthesis inhibition. *J. Bacteriol.* 182, 1802–1811.
- Almeida, M.S., Herrmann, T., Peti, W., Wilson, I.A., Wüthrich, K., 2005. NMR structure of the conserved hypothetical protein TM0487 from *Thermotoga maritima*: Implications for 216 homologous DUF59 proteins. *Protein Sci* 14, 2880–2886. <https://doi.org/10.1110/ps.051755805>
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215, 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Apweiler, R., Attwood, T.K., Bairoch, A., Bateman, A., Birney, E., Biswas, M., Bucher, P., Cerutti, L., Corpet, F., Croning, M.D.R., Durbin, R., Falquet, L., Fleischmann, W., Gouzy, J., Hermjakob, H., Hulo, N., Jonassen, I., Kahn, D., Kanapin, A., Karavidopoulou, Y., Lopez, R., Marx, B., Mulder, N.J., Oinn, T.M., Pagni, M., Servant, F., Sigrist, C.J.A., Zdobnov, E.M., 2001. The InterPro database, an integrated documentation resource for protein families, domains and functional sites. *Nucleic Acids Res* 29, 37–40. <https://doi.org/10.1093/nar/29.1.37>
- Atkinson, H.J., Morris, J.H., Ferrin, T.E., Babbitt, P.C., 2009. Using Sequence Similarity Networks for Visualization of Relationships Across Diverse Protein Superfamilies. *PLOS ONE* 4, e4345. <https://doi.org/10.1371/journal.pone.0004345>
- Ayala-Castro, C., Saini, A., Outten, F.W., 2008. Fe-S Cluster Assembly Pathways in Bacteria. *Microbiol Mol Biol Rev* 72, 110–125. <https://doi.org/10.1128/MMBR.00034-07>
- Balk, J., Pilon, M., 2011. Ancient and essential: the assembly of iron–sulfur clusters in plants. *Trends in Plant Science* 16, 218–226. <https://doi.org/10.1016/j.tplants.2010.12.006>
- Bateman, A., Coggill, P., Finn, R.D., 2010. DUFs: families in search of function. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 66, 1148–1152. <https://doi.org/10.1107/S1744309110001685>
- Benjamini, Y., Krieger, A.M., Yekutieli, D., 2006. Adaptive linear step-up procedures that control the false discovery rate. *Biometrika* 93, 491–507. <https://doi.org/10.1093/biomet/93.3.491>
- Blanc, B., Gerez, C., Ollagnier de Choudens, S., 2015. Assembly of Fe/S proteins in bacterial systems: Biochemistry of the bacterial ISC system. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, SI: Fe/S proteins 1853, 1436–1447. <https://doi.org/10.1016/j.bbamcr.2014.12.009>

- Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* 72, 248–254. [https://doi.org/10.1016/0003-2697\(76\)90527-3](https://doi.org/10.1016/0003-2697(76)90527-3)
- Caglar, M.U., Houser, J.R., Barnhart, C.S., Boutz, D.R., Carroll, S.M., Dasgupta, A., Lenoir, W.F., Smith, B.L., Sridhara, V., Sydykova, D.K., Vander Wood, D., Marx, C.J., Marcotte, E.M., Barrick, J.E., Wilke, C.O., 2017. The *E. coli* molecular phenotype under different growth conditions. *Sci Rep* 7. <https://doi.org/10.1038/srep45303>
- Chassy, B.M., Mercenier, A., Flickinger, J., 1988. Transformation of bacteria by electroporation 6, 7.
- Chen, K.-E., Richards, A.A., Ariffin, J.K., Ross, I.L., Sweet, M.J., Kellie, S., Kobe, B., Martin, J.L., 2012. The mammalian DUF59 protein Fam96a forms two distinct types of domain-swapped dimer. *Acta Cryst. D* 68, 637–648. <https://doi.org/10.1107/S0907444912006592>
- Cheng, V.W.T., Ma, E., Zhao, Z., Rothery, R.A., Weiner, J.H., 2006. The Iron-Sulfur Clusters in *Escherichia coli* Succinate Dehydrogenase Direct Electron Flow. *J. Biol. Chem.* 281, 27662–27668. <https://doi.org/10.1074/jbc.M604900200>
- Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Iii, C.E.B., Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M.A., Rajandream, M.-A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J.E., Taylor, K., Whitehead, S., Barrell, B.G., 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 393, 537–544. <https://doi.org/10.1038/31159>
- Cook, G.M., Hards, K., Vilchèze, C., Hartman, T., Berney, M., 2014. Energetics of Respiration and Oxidative Phosphorylation in Mycobacteria. *Microbiol Spectr* 2. <https://doi.org/10.1128/microbiolspec.MGM2-0015-2013>
- Costerton, J.W., Stewart, P.S., Greenberg, E.P., 1999. Bacterial biofilms: a common cause of persistent infections. *Science* 284, 1318–1322.
- Craig, N.L., 1997. Target site selection in transposition. *Annu. Rev. Biochem.* 66, 437–474. <https://doi.org/10.1146/annurev.biochem.66.1.437>
- Cromie, G.A., Connelly, J.C., Leach, D.R.F., 2001. Recombination at Double-Strand Breaks and DNA Ends: Conserved Mechanisms from Phage to Humans. *Molecular Cell* 8, 1163–1174. [https://doi.org/10.1016/S1097-2765\(01\)00419-1](https://doi.org/10.1016/S1097-2765(01)00419-1)
- Daniel, T.M., 2006. The history of tuberculosis. *Respiratory Medicine* 100, 1862–1870. <https://doi.org/10.1016/j.rmed.2006.08.006>
- Djaman, O., Outten, F.W., Imlay, J.A., 2004. Repair of oxidized iron-sulfur clusters in *Escherichia coli*. *J. Biol. Chem.* 279, 44590–44599. <https://doi.org/10.1074/jbc.M406487200>
- Donoghue, H.D., 2009. Human tuberculosis – an ancient disease, as elucidated by ancient microbial biomolecules. *Microbes and Infection, Forum on Chikungunya* 11, 1156–1162. <https://doi.org/10.1016/j.micinf.2009.08.008>

- Dunbrack, R.L., 2006. Sequence comparison and protein structure prediction. *Current Opinion in Structural Biology, Nucleic acids/Sequences and topology* 16, 374–384. <https://doi.org/10.1016/j.sbi.2006.05.006>
- Eddy, S.R., 1996. Hidden Markov models. *Current Opinion in Structural Biology* 6, 361–365. [https://doi.org/10.1016/S0959-440X\(96\)80056-X](https://doi.org/10.1016/S0959-440X(96)80056-X)
- Fang, Z., Newton-Foot, M., Sampson, S.L., Gey van Pittius, N.C., 2017. Two promoters in the *esx-3* gene cluster of *Mycobacterium smegmatis* respond inversely to different iron concentrations in vitro. *BMC Research Notes* 10, 426. <https://doi.org/10.1186/s13104-017-2752-0>
- Faulds-Pain, A., Wren, B.W., 2013. Improved Bacterial Mutagenesis by High-Frequency Allele Exchange, Demonstrated in *Clostridium difficile* and *Streptococcus suis*. *Appl. Environ. Microbiol.* 79, 4768–4771. <https://doi.org/10.1128/AEM.01195-13>
- Fidler, D.R., Murphy, S.E., Courtis, K., Antonoudiou, P., El-Tohamy, R., lent, J., Levine, T.P., 2016. Using HHsearch to tackle proteins of unknown function: A pilot study with PH domains. *Traffic* 17, 1214–1226. <https://doi.org/10.1111/tra.12432>
- Finn, R.D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E.L.L., Tate, J., Punta, M., 2014. Pfam: the protein families database. *Nucleic Acids Res* 42, D222–D230. <https://doi.org/10.1093/nar/gkt1223>
- Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G.A., Tate, J., Bateman, A., 2016. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res* 44, D279–D285. <https://doi.org/10.1093/nar/gkv1344>
- Fontecave, M., Ollagnier-de-Choudens, S., 2008. Iron–sulfur cluster biosynthesis in bacteria: Mechanisms of cluster assembly and transfer. *Archives of Biochemistry and Biophysics, Special Issue: Enzymology in Europe* 474, 226–237. <https://doi.org/10.1016/j.abb.2007.12.014>
- Galperin, M.Y., Koonin, E.V., 2004. “Conserved hypothetical” proteins: prioritization of targets for experimental study. *Nucleic Acids Res.* 32, 5452–5463. <https://doi.org/10.1093/nar/gkh885>
- Gerlt, J.A., 2016. Tools and strategies for discovering novel enzymes and metabolic pathways. *Perspectives in Science, Proceedings of the Beilstein ESCEC Symposium 2015* 9, 24–32. <https://doi.org/10.1016/j.pisc.2016.07.001>
- Goodacre, N.F., Gerloff, D.L., Uetz, P., 2014. Protein Domains of Unknown Function Are Essential in Bacteria. *mBio* 5, e00744-13. <https://doi.org/10.1128/mBio.00744-13>
- Goude, R., Parish, T., 2008. Electroporation of Mycobacteria. *JoVE (Journal of Visualized Experiments)* e761–e761. <https://doi.org/10.3791/761>
- Griffin, J.E., Gawronski, J.D., DeJesus, M.A., Ioerger, T.R., Akerley, B.J., Sasseti, C.M., 2011. High-Resolution Phenotypic Profiling Defines Genes Essential for Mycobacterial Growth and Cholesterol Catabolism. *PLoS Pathog* 7. <https://doi.org/10.1371/journal.ppat.1002251>
- Grishin, A.M., Cygler, M., 2015. Structural Organization of Enzymes of the Phenylacetate Catabolic Hybrid Pathway. *Biology* 4, 424–442. <https://doi.org/10.3390/biology4020424>

- Guerrero, G., Peralta, H., Aguilar, A., Díaz, R., Villalobos, M.A., Medrano-Soto, A., Mora, J., 2005. Evolutionary, structural and functional relationships revealed by comparative analysis of syntenic genes in Rhizobiales. *BMC Evol Biol* 5, 55. <https://doi.org/10.1186/1471-2148-5-55>
- Hausmann, A., Netz, D.J.A., Balk, J., Pierik, A.J., Mühlenhoff, U., Lill, R., 2005. The eukaryotic P loop NTPase Nbp35: An essential component of the cytosolic and nuclear iron–sulfur protein assembly machinery. *PNAS* 102, 3266–3271. <https://doi.org/10.1073/pnas.0406447102>
- Hensel, M., Holden, D.W., 1996. Molecular genetic approaches for the study of virulence in both pathogenic bacteria and fungi. *Microbiology* 142, 1049–1058. <https://doi.org/10.1099/13500872-142-5-1049>
- Horan, K., Jang, C., Bailey-Serres, J., Mittler, R., Shelton, C., Harper, J.F., Zhu, J.-K., Cushman, J.C., Gollery, M., Girke, T., 2008. Annotating Genes of Known and Unknown Function by Large-Scale Coexpression Analysis. *Plant Physiol* 147, 41–57. <https://doi.org/10.1104/pp.108.117366>
- Huet, G., Daffé, M., Saves, I., 2005. Identification of the *Mycobacterium tuberculosis* SUF Machinery as the Exclusive Mycobacterial System of [Fe-S] Cluster Assembly: Evidence for Its Implication in the Pathogen's Survival. *J Bacteriol* 187, 6137–6146. <https://doi.org/10.1128/JB.187.17.6137-6146.2005>
- Imlay, J.A., 2006. Iron-sulphur clusters and the problem with oxygen. *Molecular Microbiology* 59, 1073–1082. <https://doi.org/10.1111/j.1365-2958.2006.05028.x>
- Jamet, S., Quentin, Y., Coudray, C., Texier, P., Laval, F., Daffé, M., Fichant, G., Cam, K., 2015. Evolution of Mycolic Acid Biosynthesis Genes and Their Regulation during Starvation in *Mycobacterium tuberculosis*. *J Bacteriol* 197, 3797–3811. <https://doi.org/10.1128/JB.00433-15>
- Jang, S., Imlay, J.A., 2010. Hydrogen peroxide inactivates the *Escherichia coli* Isc iron-sulfur assembly system, and OxyR induces the Suf system to compensate. *Mol Microbiol* 78, 1448–1467. <https://doi.org/10.1111/j.1365-2958.2010.07418.x>
- Jaroszewski, L., Li, Z., Krishna, S.S., Bakolitsa, C., Wooley, J., Deacon, A.M., Wilson, I.A., Godzik, A., 2009. Exploration of Uncharted Regions of the Protein Universe. *PLoS Biol* 7. <https://doi.org/10.1371/journal.pbio.1000205>
- Johnson, D.C., Dos Santos, P.C., Dean, D.R., 2005. NifU and NifS are required for the maturation of nitrogenase and cannot replace the function of isc-gene products in *Azotobacter vinelandii*. *Biochem. Soc. Trans.* 33, 90–93. <https://doi.org/10.1042/BST0330090>
- Kulka, K., Hatfull, G., Ojha, A.K., 2012. Growth of *Mycobacterium tuberculosis* biofilms. *J Vis Exp*. <https://doi.org/10.3791/3820>
- Lamichhane, G., 2011. *Mycobacterium Tuberculosis* Response to Stress from Reactive Oxygen and Nitrogen Species. *Front Microbiol* 2. <https://doi.org/10.3389/fmicb.2011.00176>
- Lill, R., Mühlenhoff, U., 2006. Iron-Sulfur Protein Biogenesis in Eukaryotes: Components and Mechanisms. *Annual Review of Cell and Developmental Biology* 22, 457–486. <https://doi.org/10.1146/annurev.cellbio.22.010305.104538>
- Liu, Y., Eisenberg, D., 2002. 3D domain swapping: As domains continue to swap. *Protein Sci* 11, 1285–1299.

- Lundblad, V., Struhl, K., 2008. Yeast. *Current Protocols in Molecular Biology* 82, 13.0.1-13.0.4. <https://doi.org/10.1002/0471142727.mb1300s82>
- Loewenstein, Y., Raimondo, D., Redfern, O.C., Watson, J., Frishman, D., Linial, M., Orengo, C., Thornton, J., Tramontano, A., 2009. Protein function annotation by homology-based inference. *Genome Biol* 10, 207. <https://doi.org/10.1186/gb-2009-10-2-207>
- Luo, D., Bernard, D.G., Balk, J., Hai, H., Cui, X., 2012. The DUF59 Family Gene *AE7* Acts in the Cytosolic Iron-Sulfur Cluster Assembly Pathway to Maintain Nuclear Genome Integrity in *Arabidopsis*[C][W][OA]. *Plant Cell* 24, 4135–4148. <https://doi.org/10.1105/tpc.112.102608>
- Mao, C., Shukla, M., Larrouy-Maumus, G., Dix, F.L., Kelley, L.A., Sternberg, M.J., Sobral, B.W., de Carvalho, L.P.S., 2013. Functional assignment of *Mycobacterium tuberculosis* proteome revealed by genome-scale fold-recognition. *Tuberculosis, Featuring Reports From The Tuberculosis Community Annotation Project Jamboree held at Virginia Tech, USA, March 7-8, 2012* 93, 40–46. <https://doi.org/10.1016/j.tube.2012.11.008>
- Mashruwala, A.A., Bhatt, S., Poudel, S., Boyd, E.S., Boyd, J.M., 2016a. The DUF59 Containing Protein SufT Is Involved in the Maturation of Iron-Sulfur (FeS) Proteins during Conditions of High FeS Cofactor Demand in *Staphylococcus aureus*. *PLoS Genet* 12. <https://doi.org/10.1371/journal.pgen.1006233>
- Mashruwala, A.A., Pang, Y.Y., Rosario-Cruz, Z., Chahal, H.K., Benson, M.A., Mike, L.A., Skaar, E.P., Torres, V.J., Nauseef, W.M., Boyd, J.M., 2015. Nfu facilitates the maturation of iron-sulfur proteins and participates in virulence in *Staphylococcus aureus*. *Molecular Microbiology* 95, 383–409. <https://doi.org/10.1111/mmi.12860>
- Mashruwala, A.A., Roberts, C.A., Bhatt, S., May, K.L., Carroll, R.K., Shaw, L.N., Boyd, J.M., 2016b. *Staphylococcus aureus* SufT: an essential iron-sulphur cluster assembly factor in cells experiencing a high-demand for lipoic acid. *Mol. Microbiol.* 102, 1099–1119. <https://doi.org/10.1111/mmi.13539>
- Mudgal, R., Sandhya, S., Chandra, N., Srinivasan, N., 2015. De-DUFing the DUFs: Deciphering distant evolutionary relationships of Domains of Unknown Function using sensitive homology detection methods. *Biology Direct* 10, 38. <https://doi.org/10.1186/s13062-015-0069-2>
- Mulder, N.J., Kersey, P., Pruess, M., Apweiler, R., 2008. In Silico Characterization of Proteins: UniProt, InterPro and Integr8. *Mol Biotechnol* 38, 165–177. <https://doi.org/10.1007/s12033-007-9003-x>
- Munujos, P., Coll-Cantí, J., González-Sastre, F., Gella, F.J., 1993. Assay of succinate dehydrogenase activity by a colorimetric-continuous method using iodinitrotetrazolium chloride as electron acceptor. *Anal. Biochem.* 212, 506–509.
- Neumann, E., Schaefer-Ridder, M., Wang, Y., Hofschneider, P.H., 1982. Gene transfer into mouse lymphoma cells by electroporation in high electric fields. *EMBO J* 1, 841–845.
- Niehaus, T.D., Thamm, A.M.K., de Crécy-Lagard, V., Hanson, A.D., 2015. Proteins of Unknown Biochemical Function: A Persistent Problem and a Roadmap to Help Overcome It1. *Plant Physiol* 169, 1436–1442. <https://doi.org/10.1104/pp.15.00959>
- Ojha, A., Anand, M., Bhatt, A., Kremer, L., Jacobs Jr., W.R., Hatfull, G.F., 2005. GroEL1: A Dedicated Chaperone Involved in Mycolic Acid Biosynthesis during Biofilm Formation in *Mycobacteria*. *Cell* 123, 861–873. <https://doi.org/10.1016/j.cell.2005.09.012>

- Ojha, A., Hatfull, G.F., 2007. The role of iron in *Mycobacterium smegmatis* biofilm formation: the exochelin siderophore is essential in limiting iron conditions for biofilm formation but not for planktonic growth. *Mol Microbiol* 66, 468–483. <https://doi.org/10.1111/j.1365-2958.2007.05935.x>
- Ojha, A.K., Baughn, A.D., Sambandan, D., Hsu, T., Trivelli, X., Guerardel, Y., Alahari, A., Kremer, L., Jacobs, W.R., Hatfull, G.F., 2008. Growth of *Mycobacterium tuberculosis* biofilms containing free mycolic acids and harbouring drug-tolerant bacteria. *Mol Microbiol* 69, 164–174. <https://doi.org/10.1111/j.1365-2958.2008.06274.x>
- Ouyang, B., Wang, Lei, Wan, S., Luo, Y., Wang, Lu, Lin, J., Xia, B., 2013. Solution structure of monomeric human FAM96A. *J. Biomol. NMR* 56, 387–392. <https://doi.org/10.1007/s10858-013-9746-6>
- 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. <https://doi.org/10.1099/00221287-146-8-1969>
- Park, J., Karplus, K., Barrett, C., Hughey, R., Haussler, D., Hubbard, T., Chothia, C., 1998. Sequence comparisons using multiple sequences detect three times as many remote homologues as pairwise methods. *Journal of Molecular Biology* 284, 1201–1210. <https://doi.org/10.1006/jmbi.1998.2221>
- Pearson, W.R., Sierk, M.L., 2005. The limits of protein sequence comparison? *Curr Opin Struct Biol* 15, 254–260. <https://doi.org/10.1016/j.sbi.2005.05.005>
- Potter, H., Heller, R., 2003. Transfection by Electroporation. *Curr Protoc Mol Biol* CHAPTER, Unit-9.3. <https://doi.org/10.1002/0471142727.mb0903s62>
- Prakash, A., Yogeeshwari, S., Sircar, S., Agrawal, S., 2011. Protein Domain of Unknown Function 3233 is a Translocation Domain of Autotransporter Secretory Mechanism in Gamma proteobacteria. *PLOS ONE* 6, e25570. <https://doi.org/10.1371/journal.pone.0025570>
- Punta, M., Coghill, P.C., Eberhardt, R.Y., Mistry, J., Tate, J., Boursnell, C., Pang, N., Forslund, K., Ceric, G., Clements, J., Heger, A., Holm, L., Sonnhammer, E.L.L., Eddy, S.R., Bateman, A., Finn, R.D., 2012. The Pfam protein families database. *Nucleic Acids Res* 40, D290–D301. <https://doi.org/10.1093/nar/gkr1065>
- Py, B., Barras, F., 2010. Building Fe-S proteins: bacterial strategies. *Nat. Rev. Microbiol.* 8, 436–446. <https://doi.org/10.1038/nrmicro2356>
- Py, B., Moreau, P.L., Barras, F., 2011. Fe-S clusters, fragile sentinels of the cell. *Curr. Opin. Microbiol.* 14, 218–223. <https://doi.org/10.1016/j.mib.2011.01.004>
- Ramakrishnan, G., Ochoa-Montaño, B., Raghavender, U.S., Mudgal, R., Joshi, A.G., Chandra, N.R., Sowdhamini, R., Blundell, T.L., Srinivasan, N., 2015. Enriching the annotation of *Mycobacterium tuberculosis* H37Rv proteome using remote homology detection approaches: Insights into structure and function. *Tuberculosis* 95, 14–25. <https://doi.org/10.1016/j.tube.2014.10.009>
- Robbins, A.H., Stout, C.D., 1989. Structure of activated aconitase: formation of the [4Fe-4S] cluster in the crystal. *Proc Natl Acad Sci U S A* 86, 3639–3643.

- Rodriguez, G.M., Voskuil, M.I., Gold, B., Schoolnik, G.K., Smith, I., 2002. IdeR, an Essential Gene in *Mycobacterium tuberculosis*: Role of IdeR in Iron-Dependent Gene Expression, Iron Metabolism, and Oxidative Stress Response. *Infect Immun* 70, 3371–3381. <https://doi.org/10.1128/IAI.70.7.3371-3381.2002>
- Rubin, E.J., Akerley, B.J., Novik, V.N., Lampe, D.J., Husson, R.N., Mekalanos, J.J., 1999. In vivo transposition of mariner-based elements in enteric bacteria and mycobacteria. *Proc. Natl. Acad. Sci. U.S.A.* 96, 1645–1650.
- Rybniker, J., Pojer, F., Marienhagen, J., Kolly, G.S., Chen, J.M., Gumpel, E. van, Hartmann, P., Cole, S.T., 2014. The cysteine desulfurase IscS of *Mycobacterium tuberculosis* is involved in iron–sulfur cluster biogenesis and oxidative stress defence. *Biochemical Journal* 459, 467–478. <https://doi.org/10.1042/BJ20130732>
- Sambrook, J., and Russell, R.W.(2001). *Molecular cloning: A laboratory manual*, 3rd ed. Cold spring harbor laboratory press, cold spring harbor, N.Y.
- Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. *Molecular cloning: a laboratory manual*. Molecular cloning: a laboratory manual.
- Santos, J.A., Alonso-García, N., Macedo-Ribeiro, S., Pereira, P.J.B., 2014. The unique regulation of iron-sulfur cluster biogenesis in a Gram-positive bacterium. *Proc. Natl. Acad. Sci. U.S.A.* 111, E2251-2260. <https://doi.org/10.1073/pnas.1322728111>
- Sasaki, S., Minamisawa, K., Mitsui, H., 2016. A *Sinorhizobium meliloti* RpoH-Regulated Gene Is Involved in Iron-Sulfur Protein Metabolism and Effective Plant Symbiosis under Intrinsic Iron Limitation. *J. Bacteriol.* 198, 2297–2306. <https://doi.org/10.1128/JB.00287-16>
- Sassetti, C.M., Boyd, D.H., Rubin, E.J., 2001. Comprehensive identification of conditionally essential genes in mycobacteria. *Proc. Natl. Acad. Sci. U.S.A.* 98, 12712–12717. <https://doi.org/10.1073/pnas.231275498>
- Sassetti, C.M., Boyd, D.H., Rubin, E.J., 2003. Genes required for mycobacterial growth defined by high density mutagenesis. *Molecular Microbiology* 48, 77–84. <https://doi.org/10.1046/j.1365-2958.2003.03425.x>
- Schaeffer, M.L., Agnihotri, G., Volker, C., Kallender, H., Brennan, P.J., Lonsdale, J.T., 2001. Purification and biochemical characterization of the *Mycobacterium tuberculosis* beta-ketoacyl-acyl carrier protein synthases KasA and KasB. *J. Biol. Chem.* 276, 47029–47037. <https://doi.org/10.1074/jbc.M108903200>
- Schirawski, J., Unden, G., 1998. Menaquinone-dependent succinate dehydrogenase of bacteria catalyzes reversed electron transport driven by the proton potential. *European Journal of Biochemistry* 257, 210–215. <https://doi.org/10.1046/j.1432-1327.1998.2570210.x>
- Schwenkert, S., Netz, D.J.A., Frazzon, J., Pierik, A.J., Bill, E., Gross, J., Lill, R., Meurer, J., 2009. Chloroplast HCF101 is a scaffold protein for [4Fe-4S] cluster assembly. *Biochem. J.* 425, 207–214. <https://doi.org/10.1042/BJ20091290>
- Selbach, B.P., Chung, A.H., Scott, A.D., George, S.J., Cramer, S.P., Dos Santos, P.C., 2014. Fe-S Cluster Biogenesis in Gram-Positive Bacteria: SufU Is a Zinc-Dependent Sulfur Transfer Protein. *Biochemistry* 53, 152–160. <https://doi.org/10.1021/bi4011978>

- Sigrist, C.J.A., Cerutti, L., de Castro, E., Langendijk-Genevoux, P.S., Bulliard, V., Bairoch, A., Hulo, N., 2010. PROSITE, a protein domain database for functional characterization and annotation. *Nucleic Acids Res* 38, D161–D166. <https://doi.org/10.1093/nar/gkp885>
- Snapper, S.B., Melton, R.E., Mustafa, S., Kieser, T., Jacobs, W.R., 1990. Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Mol. Microbiol.* 4, 1911–1919.
- Sritharan, M., Yeruva, V.C., Sivasailappan, S.C., Duggirala, S., 2006. Iron enhances the susceptibility of pathogenic mycobacteria to isoniazid, an antitubercular drug. *World J Microbiol Biotechnol* 22, 1357. <https://doi.org/10.1007/s11274-006-9183-8>
- Stehling, O., Mascarenhas, J., Vashisht, A.A., Sheftel, A.D., Niggemeyer, B., Rösser, R., Pierik, A.J., Wohlschlegel, J.A., Lill, R., 2013. Human CIA2A-FAM96A and CIA2B-FAM96B Integrate Iron Homeostasis and Maturation of Different Subsets of Cytosolic-Nuclear Iron-Sulfur Proteins. *Cell Metabolism* 18, 187–198. <https://doi.org/10.1016/j.cmet.2013.06.015>
- Tamames, J., 2001. Evolution of gene order conservation in prokaryotes. *Genome Biol* 2, research0020.1-research0020.11.
- Tatusov, R.L., Galperin, M.Y., Natale, D.A., Koonin, E.V., 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res* 28, 33–36.
- Tsaousis, A.D., Gentekaki, E., Eme, L., Gaston, D., Roger, A.J., 2014. Evolution of the Cytosolic Iron-Sulfur Cluster Assembly Machinery in Blastocystis Species and Other Microbial Eukaryotes. *Eukaryot Cell* 13, 143–153. <https://doi.org/10.1128/EC.00158-13>
- Vinella, D., Brochier-Armanet, C., Loiseau, L., Talla, E., Barras, F., 2009. Iron-Sulfur (Fe/S) Protein Biogenesis: Phylogenomic and Genetic Studies of A-Type Carriers. *PLOS Genetics* 5, e1000497. <https://doi.org/10.1371/journal.pgen.1000497>
- Voskuil, M.I., Bartek, I.L., Visconti, K., Schoolnik, G.K., 2011. The Response of *Mycobacterium Tuberculosis* to Reactive Oxygen and Nitrogen Species. *Front Microbiol* 2. <https://doi.org/10.3389/fmicb.2011.00105>
- Wardman, P., Candeias, L.P., 1996. Fenton Chemistry: An Introduction. *Radiation Research* 145, 523–531. <https://doi.org/10.2307/3579270>
- Warren, R.M., van Helden, P.D., Gey van Pittius, N.C., 2009. Insertion element IS6110-based restriction fragment length polymorphism genotyping of *Mycobacterium tuberculosis*. *Methods Mol. Biol.* 465, 353–370. https://doi.org/10.1007/978-1-59745-207-6_24
- Watanabe, N., Forman, H.J., 2003. Autoxidation of extracellular hydroquinones is a causative event for the cytotoxicity of menadione and DMNQ in A549-S cells. *Arch Biochem Biophys* 411, 145–157.
- Wayne Outten, F., 2015. Recent advances in the Suf Fe–S cluster biogenesis pathway: Beyond the Proteobacteria. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, SI: Fe/S proteins 1853, 1464–1469. <https://doi.org/10.1016/j.bbamcr.2014.11.001>
- Wehrli, W., 1983. Rifampin: mechanisms of action and resistance. *Rev. Infect. Dis.* 5 Suppl 3, S407-411.

- Welch, M.D., 2015. Why should cell biologists study microbial pathogens? *Mol Biol Cell* 26, 4295–4301. <https://doi.org/10.1091/mbc.E15-03-0144>
- Wet, T.J. de, Gobe, I., Mhlanga, M.M., Warner, D.F., 2018. CRISPRi-Seq for the Identification and Characterisation of Essential Mycobacterial Genes and Transcriptional Units. *bioRxiv* 358275. <https://doi.org/10.1101/358275>
- Wilson, K., 2001. Preparation of genomic DNA from bacteria. *Curr Protoc Mol Biol* Chapter 2, Unit 2.4. <https://doi.org/10.1002/0471142727.mb0204s56>
- Winterbourn, C.C., 1995. Toxicity of iron and hydrogen peroxide: the Fenton reaction. *Toxicol. Lett.* 82–83, 969–974.
- World Health Organisation. (2017) Global Tuberculosis Report 2017. (accessed 5 June 2018). Available at: http://www.who.int/tb/publications/global_report/MainText_13Nov2017
- Wu, N., Matand, K., Kebede, B., Acquaaah, G., Williams, S., 2010. Enhancing DNA electrotransformation efficiency in *Escherichia coli* DH10B electrocompetent cells. *Electronic Journal of Biotechnology* 13, 0–0. <https://doi.org/10.2225/vol13-issue5-fulltext-11>
- Xu, D., 2004. Protein Databases on the Internet. *Curr Protoc Mol Biol* CHAPTER, Unit-19.4. <https://doi.org/10.1002/0471142727.mb1904s68>
- Yang, J., Bitoun, J.P., Ding, H., 2006. Interplay of IscA and IscU in Biogenesis of Iron-Sulfur Clusters. *J. Biol. Chem.* 281, 27956–27963. <https://doi.org/10.1074/jbc.M601356200>
- Yang, Y., Thomas, J., Li, Y., Vilchèze, C., Derbyshire, K.M., Jacobs, W.R., Ojha, A.K., 2017. Defining a temporal order of genetic requirements for development of mycobacterial biofilms. *Mol. Microbiol.* 105, 794–809. <https://doi.org/10.1111/mmi.13734>
- Yano, T., Kassovska-Bratinova, S., Teh, J.S., Winkler, J., Sullivan, K., Isaacs, A., Schechter, N.M., Rubin, H., 2011. Reduction of Clofazimine by Mycobacterial Type 2 NADH:Quinone Oxidoreductase. A pathway for the generation of bactericidal levels of reactive oxygen species. *J. Biol. Chem.* 286, 10276–10287. <https://doi.org/10.1074/jbc.M110.200501>
- Yeruva, V.C., Sundaram, C.S., Sritharan, M., 2005. Effect of iron concentration on the expression and activity of catalase-peroxidases in mycobacteria. *IJBB* Vol.42(1) [February 2005].
- Zhang, H., Zhu, F., Yang, T., Ding, L., Zhou, M., Li, J., Haslam, S.M., Dell, A., Erlandsen, H., Wu, H., 2014. The highly conserved domain of unknown function 1792 has a distinct glycosyltransferase fold. *Nature Communications* 5, 4339. <https://doi.org/10.1038/ncomms5339>
- Zhang, X., Carter, M.S., Vetting, M.W., San Francisco, B., Zhao, S., Al-Obaidi, N.F., Solbiati, J.O., Thiaville, J.J., de Crécy-Lagard, V., Jacobson, M.P., Almo, S.C., Gerlt, J.A., 2016. Assignment of function to a domain of unknown function: DUF1537 is a new kinase family in catabolic pathways for acid sugars. *Proc. Natl. Acad. Sci. U.S.A.* 113, E4161-4169. <https://doi.org/10.1073/pnas.1605546113>
- Zhao, S., Sakai, A., Zhang, X., Vetting, M.W., Kumar, R., Hillerich, B., Francisco, B.S., Solbiati, J., Steves, A., Brown, S., Akiva, E., Barber, A., Seidel, R.D., Babbitt, P.C., Almo, S.C., Gerlt, J.A., Jacobson, M.P., 2014. Prediction and characterization of enzymatic activities guided by sequence similarity and genome neighborhood networks [WWW Document]. *eLife*. <https://doi.org/10.7554/eLife.03275>

