

https://dspace.org/

Functional characterisation of molbdopterin synthase-encoding genes in mycobacteria

Narrandes, Nicole Collette

2014-02-18
http://hdl.handle.net10539/13845

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

FUNCTIONAL CHARACTERISATION OF MOLYBDOPTERIN SYNTHASE-ENCODING GENES IN MYCOBACTERIA

Nicole Collette Narrandes

A dissertation submitted to the Faculty of Health Science, University of the Witwatersrand, Johannesburg, in fulfillment of the requirements for the degree of Master of Science in Medicine.

Johannesburg, 2013

Declaration

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



(Nicole C Narrandes)

28th day of May 2013

Presentations

Parts of this work have been presented at the following conferences:

1. University of the Witwatersrand Cross Faculty Symposium 2010. **Poster presentation**
2. University of the Witwatersrand Faculty of Health Science Research Day 2010. **Poster presentation**
3. Medical Research Council Research Day 2010. **Oral presentation**
4. Molecular Biosciences Research Thrust Research Day 2010. **Poster presentation**
5. Medical Research Council Research Day 2011. **Oral presentation**
6. SASBMB/FASBMB Conference 2012. **Oral presentation**
7. EMBO Tuberculosis 2012: Biology, pathogenesis and Intervention strategies. **Poster presentation**
8. University of the Witwatersrand Faculty of Health Science Research Day 2012. **Oral presentation**
9. 4th Cross Faculty Graduate Symposium: Showcasing Postgraduate Research at Wits 2012. **Poster presentation**
10. Molecular Medicine and Haematology Seminar Series 2012. **Oral presentation**
11. Centre of Excellence for Biomedical TB Research Retreat 2013. **Oral presentation**

Abstract

Mycobacterium tuberculosis (*Mtb*) possesses a complete repertoire of genes for the biosynthesis of molybdopterin cofactor (MoCo). The multi-step biosynthetic pathway in *Mtb* is distinguished by the fact that it displays a multiplicity of homologues of several genes, most notably those involved in the second step, which include *moaD1*, *moaD2*, *moaE1*, *moaE2* and *moaX*. The *moaD* and *moaE* genes encode the small and large subunits of the molybdopterin (MPT) synthase enzyme respectively, whereas *moaX* encodes a novel fused MPT synthase which contains both MoaD and MoaE functional domains. This study aimed to assess the function of these multiple homologues and their relative contributions to MoCo biosynthesis in *Mtb* and to investigate the role of post-translational processing in MoaX function. In addition, the contribution of two *Mycobacterium smegmatis* MoCo-dependent nitrate reductase (NR) enzymes, the putative assimilatory NarB and the respiratory NarGHI, to nitrate assimilation was investigated. Derivatives of the MoCo-deficient *M. smegmatis* Δ*moaD2* Δ*moaE2* double mutant were generated carrying all possible combinations of the *Mtb moaD* and *moaE* genes to assess the ability of these genes to complement the growth phenotype when expressed in this heterologous host. MoCo biosynthesis was monitored by the ability to grow in minimal media containing nitrate as a sole nitrogen source (MPLN), facilitated by a MoCo dependent assimilatory NR. Complementation studies showed that only the *moaD2 moaE2* combination of *Mtb* genes are able to restore growth of the *M. smegmatis* double mutant in MPLN when introduced on multi-copy plasmid, pointing to a functional hierarchy in MPT synthase encoding genes in *Mtb*. Furthermore, the fused MPT synthase, MoaX, was shown to be cleaved at a glycine residue (Gly81), corresponding to the penultimate glycine in MoaD homologues; this

process is essential for MPT synthase activity. Site-directed mutagenesis was used to show that another glycine residue in MoaX (Gly82), corresponding to the terminal glycine residue of MoaD homologues, is crucial for MoaX function. Together, these data suggest that MoaX functions as a canonical MPT synthase. Phenotypic characterization of the NR-deficient mutants, Δ*narB*, Δ*narGHJI* and Δ*narB* Δ*narGHJI*, revealed that the loss of both NarB and NarGHI did not alter the organisms ability to grow in MPLN, suggesting either that *M. smegmatis* possesses additional MoCo-dependent enzymes which are able to catalyze the reduction of nitrate to nitrite or an alternate nitrate assimilation pathway exists. In summary, this study has provided new insights into the biosynthesis of a key mycobacterial cofactor, which may contribute to the development of improved strategies to combat tuberculosis.

Acknowledgments

Firstly I would like to thank all the institutions that provided me with funding throughout this MSc, without which it would not have been possible: the National Research Foundation (NRF) through the DST/NRF Centre of Excellence for Biomedical TB Research, the South African Medical Research Council (MRC), the University of the Witwatersrand (Postgraduate Merit Award and Postgraduate Merit Scholarship) and the Belgian Technical Corporation.

I would like to acknowledge my co-supervisor, Prof Valerie Mizrahi for her much valued advice in guiding the research.

My supervisor, Dr Bavesh Kana- I can't thank you enough for your unwavering support and guidance for my project and life as a whole. Your scientific skills and knowledge are unmatched, much like your compassion and kindness.

I would like to thank Dr Monique Williams for providing me with strains and vectors.

My thanks go out to all the past and present members of the CBTBR who I had the pleasure of working with, particularly my lunch-time and Nando's buddies: Germar, Sibu, Chris, Rukaya and Farzanah. The advice, laughs and food kept me sane and full.

I would also like to thank my family, Narrandes, Cardoso and Budhu for all your support in all aspects. Especially Warr, Lu and Aunty Annie- words cannot express my gratitude.

And finally to my best friend, my love and my "Roc": Darrin. I don't have enough words or time to express how much I love you and how grateful I am for everything you have done, you continue to do and everything you are. I will spend the rest of my life trying though.

Table of contents

Declaration ii
Presentations iii
Abstract iv
Acknowledgments vi
Table of contents vii
List of figures xi
List of tables xiii
Nomenclature xiv
1 Introduction 1
1.1. Tuberculosis: Prevention and treatment 1
1.2. *Mtb* infection and the host environment 4
1.2.1. *Mtb* adaptations for survival 6
1.3. Molybdenum 8
1.3.1. Molybdoenzymes 8
1.4. MoCo-dependent enzymes in mycobacteria 10
1.4.1. *Mtb* molybdoenzymes and pathogenesis 10
1.4.2. *M. smegmatis* molybdoenzymes 13
1.5. MoCo biosynthesis 13
1.5.1. Molybdenum uptake 14
1.5.2. MoCo biosynthetic pathway 15
1.6. MoCo and *Mtb* pathogenesis 17
1.7. Expansion of MoCo biosynthetic genes in *Mtb* 18
1.8. MPT-synthase 19
1.8.1. *Mtb* MPT synthase 21
1.9. Aims 22
2 Methods 24
2.1 Bioinformatic tools and software 24
2.2 Chemicals and reagents 24
2.3 Bacterial strains and culture conditions 24

2.4 Bacterial transformations 25
2.4.1 *E. coli* transformations 25
2.4.2 *M. smegmatis* electroporation 26
2.5 DNA extraction methods 26
2.5.1 Mini-prep plasmid DNA extraction 26
2.5.2 Maxi-prep plasmid DNA extraction 27
2.5.3 Small scale genomic DNA extraction 28
2.5.4 Large scale genomic DNA extraction 28
2.6 DNA quantification 29
2.7 DNA manipulation methods 29
2.7.1 DNA amplification-Polymerase chain reaction (PCR) 29
2.7.2 Restriction digestion 30
2.7.3 Modification of DNA overhangs 31
2.7.4 Dephosphorylation of DNA 31
2.7.5 DNA ligation 32
2.8 Visualisation of DNA 33
2.9 DNA fragment purification 33
2.10 DNA sequencing 33
2.11 Construction of integrating vectors carrying *Mtb moaD*, *moaE* and *moaX* homologues 34
2.12 Generation of *M. smegmatis* strains carrying integrating complementation vectors 35
2.13 Construction of episomal vectors carrying *Mtb moaD* and *moaE* homologues 36
2.14 Generation of *M. smegmatis* strains carrying episomal complementation vectors 37
2.15 MoCo biosynthesis measurement: Heterologous complementation assay 37
2.15.1 Growth curve in nitrate minimal media 38
2.16 Construction of FLAG-tagged derivatives of *moaX* 38
2.17 Generation of *M. smegmatis* strains carrying FLAG-tagged MoaX 42
2.18 MoaX mutagenesis 43
2.19 MoaX protein analyses 44
2.19.1 Protein induction 44
2.19.2 *M. smegmatis* protein extractions 45
2.19.3 Protein quantification 47
2.19.4 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) 47

2.19.5 Western-blotting 47
2.20 Generation of *M. smegmatis* knock-out mutants 49
2.20.1 Construction of *narB* and *narGHJI* suicide vectors 50
2.20.2 Generation of Δ*narB* knock-out mutant 52
2.20.3 Generation of Δ*narGHJI* and Δ*narB* Δ*narGHJI* knock-out mutants 53
2.21 Southern blot analysis 53
2.21.1 Electro-blotting 53
2.21.2 Probe labeling 54
2.21.3 Hybridization 54
2.21.4 Immunological detection 55
2.22 Phenotypic characterization of knock-out mutants 56
3 Results 57
3.1 Assessment of *moaD* and *moaE* gene function with single copy integrating vectors 57
3.1.1 Strain generation and genotypic confirmation 57
3.1.2 MoCo biosynthesis in Δ*moaD2* Δ*moaE2* strains complemented with integrating vectors 62
3.2 A single copy of *moaX* can restore MoCo biosynthesis in *M. smegmatis* Δ*moaD2* Δ*moaE2* 67
3.3 Operonic expression of *Mtb moaD* and *moaE* genes from episomal vectors 69
3.3.1 *Mtb moaE1* is toxic when expressed in a synthetic operon 69
3.4 MoaX is a fused MPT synthase 72
3.5 FLAG™-tagged *moaX* 73
3.6 FLAG-tagging does not abrogate the function of *moaX* 75
3.7 MoaX processing 76
3.8 Essential MoaX residues 79
3.9 Gly81 and Gly82 are both essential for MoaX function 80
3.10 Gly81 is important for MoaX cleavage 81
3.11 MoaX is not functional in *E. coli* due to incorrect cleavage 82
3.12 Generation of *M. smegmatis* Δ*narB* knock-out mutant 83
3.13 *narB* is dispensable for growth in nitrate minimal media 86
3.14 Generation of Δ*narGHJI* and Δ*narB* Δ*narGHJI* knock-out mutants 87
3.15 Both *narB* and *narGHJI* are dispensable for growth in nitrate minimal media 90
4 Discussion 92
4.1 Concluding remarks 100

5 Appendices 101
Appendix A- Bioinformatic tools 101
A 1. BLAST 101
A 2. Genolist 101
A 3. KEGG Pathway Database 101
A 4. Sequence alignments 101
Appendix B- Media and solution preparation 103
Appendix C- Molecular weight markers 107
Appendix D- Plasmids and primers 129
Appendix E- Generation and restriction confirmation of vectors 133
E 1. Restriction analyses of integrating vectors 133
E 2. Restriction mapping of pTmoaX 135
E 3. Restriction analyses of episomal vectors 136
E 4. Construction of pFLAGEM vectors carrying *moaX* 139
E 5. Construction of pFLAGEM vectors carrying mutated *moaX* 140
E 6. Construction of Δ*narB* suicide vector 145
E 7. Generation of Δ*narGHJI* suicide vector 147
6 References 149

List of figures

Figure 1.1: Categories of molybdenum-containing enzymes and the structure of the cofactors present in each 10
Figure 1.2: Schematic representation of the molybdate ion transport system.....14
Figure 1.3: The highly conserved multi-step MoCo biosynthetic pathway in bacteria.....15
Figure 1.4: Chromosomal distribution of *Mtb* genes involved in MoCo biosynthesis 18
Figure 1.5: Schematic representation of the structure of MPT synthase.. 19
Figure 2.1: Schematic representation of the induction of *moaX* in the Tet system......42
Figure 2.2: Diagram depicting the Megaprimer method of generating site-directed mutations in *moaX*.. 43
Figure 2.3: Schematic depiction of two-step allelic exchange mutagenesis using *narB* as the example gene.....50
Figure 2.4: Schematic representation of the generation the suicide vector p∆*narB* 51
Figure 2.5: Schematic representation of the generation the suicide vector p∆*narGHJI*. 52
Figure 3.1: PCR confirmation of *M. smegmatis* double mutant strains complemented with different combinations of *Mtb moaD1, moaD2, moaE1* and *moaE2* genes carried on integrating vectors.....58
Figure 3.2: Schematic representation of integration of pHINT into the chromosome of *M. smegmatis* 59
Figure 3.3: PCR confirmation of site-specific integration of pHINT carrying *Mtb moaD1* and *moaD2* into the *M. smegmatis* chromosome at the attB site, $tRNA^{Glycine}$ 60
Figure 3.4: Schematic representation of the integration of pTT1b into the chromosome of *M. smegmatis* 61
Figure 3.5: Growth curve of *M. smegmatis* ∆*moaD2*∆*moaE2* complemented with different combinations of *Mtb moaD1, moaD2, moaE1* and *moaE2* carried on integrating vectors 63
Figure 3.6: PCR confirmation of *M. smegmatis* single mutant strains complemented with *Mtb* genes on integrating or episomal vectors.....64
Figure 3.7: Confirmation of site specific integration of pHINT carrying *Mtb moaD1* or *moaD2* and pTT1b carrying *moaE1* or *moaE2* into the chromosome of the *M. smegmatis* single mutants.....65
Figure 3.8: Growth curve of *M. smegmatis* single mutants,∆*moaD2*and ∆*moaE2* complemented with either *Mtb moaD1, moaD2, moaE1* or *moaE2* carried on integrating and episomal vectors. 67
Figure 3.9: PCR confirmation of ∆*moaD2* ∆*moaE2*:: pTX.....68
Figure 3.10: Growth curve comparing complementation with a single copy of the gene vs multiple copies. 68
Figure 3.11: PCR confirmation of double mutant strains complemented with different combinations of *Mtb moaD1, moaD2, moaE1* and *moaE2* carried on episomal vectors. 71
Figure 3.12: Growth curve of strains complemented with episomal vectors carrying different combinations of *Mtb moaD1, moaD2, moaE1*and *moaE2* genes. 72
Figure 3.13: Sequence alignment of *E. coli* MoaD and *Mtb* MoaD1, MoaD2 and MoaX proteins. 73
Figure 3.14: Schematic representation of the cleavage of MoaX showing the predicted site of cleavage and the expected sizes of each subunit once MoaX is processed at this site. 73
Figure 3.15: PCR confirmation of the site-specific integration of pMC1s 74
Figure 3.16: PCR confirmation of the presence of *moaX* in strains complemented with pFLAGmoaXN and pFLAGmoaXC. 75
Figure 3.17: Growth curve analysis of strains carrying FLAG-tagged *moaX* 76
Figure 3.18: Western blot showing the post-translational cleavage of MoaX 77
Figure 3.19: MoaX is cleavage is not altered by media composition 79
Figure 3.20: Growth curve analysis of strains carrying FLAG-tagged derivatives of *moaX* with either a 242G>C or 245G>C mutation 80
Figure 3.21: Western blot analysis of protein extracts from strains carrying mutated copies of *moaX*. 81
Figure 3.22: Western blot analysis of FLAG-tagged MoaX protein samples extracted from *E. coli* and *M. smegmatis*. 82
Figure 3.23: Screening and genotypic confirmation of ∆*narB*. 86

Figure 3.24: Growth curve analysis of Δ*narB* in nitrate minimal media shows that it is dispensable for growth 86
Figure 3.25: Screening and genotypic confirmation of Δ*narGHJI* and Δ*narB ΔnarGHJI*. 88
Figure 3.26: Growth curve analysis of Δ*narB*, Δ*narGHJI* and Δ*narB* Δ*narGHJI* in nitrate minimal media shows that both genes are dispensable for growth in nitrate minimal media. 91
Figure 4.1: Crystal structure of *E. coli* MPT synthase enzyme. 95
Figure E 1: Restriction analysis of integrating vector, pHD1 carrying *Mtb moaD1* driven off the constitutive *hsp60* promoter. 133
Figure E 2: Restriction analysis of integrating vector, pHD2 carrying *Mtb moaD2* driven off the constitutive hsp60 promoter 134
Figure E 3: Restriction analysis of integrating vector, pTE1 carrying *Mtb moaE1* driven off the constitutive *hsp60* promoter. 134
Figure E 4: Restriction analysis of integrating vector, pTE2 carrying *Mtb moaE2* driven off the constitutive *hsp60* promoter. 135
Figure E 5: Restriction analysis of integrating vector, pTmoaX carrying a single copy of *moaX* driven off the constitutive *hsp60* promoter 136
Figure E 6: Restriction analysis of episomal vector carrying *Mtb moaD1* and *moaE1* genes driven off the constitutive *hsp60* promoter as an operon. 137
Figure E 7: Restriction analysis of episomal vector carrying *Mtb moaD1* and *moaE2* genes driven off the constitutive *hsp60* promoter as an operon. 137
Figure E 9: Restriction analysis of episomal vector carrying *Mtb moaD2* and *moaE2* genes driven off the constitutive *hsp60* promoter as an operon. 138
Figure E 8: Restriction analysis of episomal vector carrying *Mtb moaD2* and *moaE1* genes driven off the constitutive *hsp60* promoter as an operon 138
Figure E 10: Restriction analysis of pFLAG vector carrying C-terminally FLAG-tagged *Mtb moaX* under the control of the *tet* operator. 139
Figure E 11: Restriction analysis of pFLAG vector carrying N-terminally FLAG-tagged *Mtb moaX* under the control of the tet operator. 140
Figure E 12: Generation of megaprimers carrying point mutations to be incorporated into *moaX*.. 140
Figure E 13: Generation of full length *moaX* with point mutations 141
Figure E 14: Re-amplification of *moaX* carrying point mutations 142
Figure E 15: *Sac*II screening of full length *moaX* with either 242G→C or 245G→C point mutations incorporated 143
Figure E 16: Confirmation of the incorporation of point mutation 245G→C into *moaX*. 143
Figure E 17: Restriction mapping of pFLAGga1C carrying a C-terminally FLAG-tagged derivative of *moaX* with point mutation 242G>C 144
Figure E 18: Restriction mapping of pFLAGga2C carrying a C-terminally FLAG-tagged derivative of *moaX* with point mutation 245G>C 144
Figure E 19: Image of chromatogram showing the incorporation of the point mutations 242G→C and 245G→C into *moaX*. 145
Figure E 20: Confirmation of p2nilnarB clone by restriction digestiont. 146
Figure E 21: Restriction digest confirmation of pΔnarB. 147
Figure E 22: Confirmation of suicide vector pΔ*narGHJI* by restriction digestion 148

List of tables

Table 2.1: Criteria used for the selection of oligonucleotide sequences on Primer3 29
Table 2.2: Primers used for the amplification of *Mtb moaD1* and *moaD2* with vector DNA as a template 36
Table 2.3: Primers used to generate FLAG-tagged derivatives of *moaX* 39
Table 2.4: Strains assessed for MoCo biosynthesis using the heterologous complementation assay 40
Table 2.5: Primers used to introduce point mutations in *moaX* 44
Table 2.6: List of strains carrying FLAG-tagged derivatives of *Mtb moaX* 46
Table 2.7: List of *M. smegmatis* knock-out mutant strains generated in this study 56
Table 3.1: Simplified names assigned to strains carrying integrating vectors 58
Table 3.2: Episomal vectors pMD1E1 and pMD2E1 are toxic to *M. smegmatis* cells. 70
Table 3.3: Simplified names assigned to strains carrying episomal vectors 70
Table 4.1: List of possible nitrate reduction catalyzing enzymes 100
Table B 1: Recipes of media used for bacterial growth 103
Table B 2: Recipes for media supplementation stocks 103
Table B 3: Solutions used for preparation of chemically competent *E. coli* cells 103
Table B 4: Solutions used for extraction of genomic DNA from *M. smegmatis* 103
Table B 5: Solutions used for plasmid extractions from *E. coli* 104
Table B 6: Solutions used for DNA precipitation 104
Table B 7: Solutions used for protein extractions 104
Table B 8: DNA electrophoresis solutions 104
Table B 9: Recipe for agarose gels 104
Table B 10: Protein electrophoresis solutions 105
Table B 11: Recipe for two SDS-PAGE gels (10 ml) 105
Table B 12: Southern blot solutions 105
Table B 13: Western blot solutions 106
Table D 1: List of plasmids used and generated throughout this study 129
Table D 2: Primers used to assess site specific intergration of L5-based vectors, pHINT and pMC1s 130
Table D 3: List of primers used for screening and confirmation of *M. smegmatis* complemented strains carrying different *Mtb* genes 131
Table D 4: Primers used to amplify upstream and downstream regions of *narB* and *narGHJI* for the generation of knock out mutants 131
Table D 5: Primers used for PCR screening of *ΔnarGHJI* mutants 132
Table D 6: Primers used for PCR screening of *ΔnarB* mutants 132

Nomenclature

ABC	ATP-binding cassette
Amp	Ampicilin
APCs	Antigen presenting cells
Arg	Arginine
ART	Antiretroviral therapy
Asp	Asparagine
ATc	Anhydrotetracycline
ATP	Adenosine triphosphate
BCG	Bacille Calmette-Guérin
***bis*-MGD**	*bis*-molybdopterin guanine dinucleotide
bp	Base pairs
BSA	Bovine serum albumin
CDH	Carbon monoxide dehydrogenase
CO	Carbon monoxide
cPMP	Cyclic pyranopterin monophosphate
CTAB	Cetyltrimethylammonium bromide
CTP	Cytosine triphosphate
DIG-dUTP	Digoxygenin labeled –dUTP
DMSO	Dimethyl sulfoxide
DMSOR	Dimethylsulfoxide reductase
DNA	Deoxyribo-nucleic acid
dNTPs	Deoxynucleotide triphosphates
DOTS	*d*irectly *o*bserved *t*herapy *s*hortcourse
dTTP	Deoxy-Tyrosine triphosphate
dUTP	Deoxy-Uracil triphosphate
EDTA	Ethylenediaminetetraacetic acid
EMB	Ethambutol
FdhF	Formate dehydrogenase
FeMo-co	Iron-molybdenum cofactor
Fe-S	Iron-sulfur
g	Gravitational acceleration
Gly	Glycine
GMP	Guanosine monophosphate
GTP	Guanosine triphosphate
HCl	Hydrochloric acid
His	Histidine
HIV	Human Immunodeficiency Virus
hr	Hours
Hyg	Hygromycin
IFNγ	Interferon gamma
INH	Isoniazid
Kan	Kanamycin

kb	Kilo base pairs
kDa	Kilo Daltons
LA	Luria Bertani agar
LB	Luria Bertani broth
LTBI	Latent TB infection
Lys	Lysine
M	Molar
MCD	Molybdopterin-cytosine dinucleotide cofactor
MDR	Multi-drug resistant
$MgCl_2$	Magnesium chloride
min	Minutes
Mo	Molybdenum
MoCo	Molybdenum cofactor
MPLN	*M. phlei* media
MPT	Molybdopterin or metal-containing pterin
Mtb	*Mycobacterium tuberculosis*
MTBC	*Mtb* complex
NaCl	Sodium Chloride
NaOH	Sodium hydroxide
NEB	New England Biolabs
NO	Nitric oxide
NR	Nitrate reductase
OADC	Oleic acid-albumin-dextrose-catalase
OD	Optical density
oriM	Origin of replication
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PZA	Pyrazinamide
RIF	Rifampicin
RNI	Reactive nitrogen intermediate
ROI	Reactive oxygen intermediate
SAM	S-adenosylmethionine
sdH_2O	Sterile distilled water
SDS	Sodium dodecyl sulfate
sec	Seconds
SO	Sulfite oxidase
Ta	Annealing temperature
TB	Tuberculosis
TetR	Tet repressor
TLRs	Toll-like receptors
TNF-α	Tumor necrosis factor alpha
Tyr	Tyrosine
WHO	World Health Organization
XDR	Extensively drug resistant
XO	Xanthine oxidase

1 Introduction

Mycobacterium tuberculosis (*Mtb*), the causative agent of tuberculosis (TB), is one of the most devastating human pathogens and is currently responsible for the largest number of deaths attributable to a single bacterium (Corbett *et al.*, 2003). In 2011 alone, 8.7 million new cases of infection were reported and 1.4 million people worldwide died due to the disease (WHO, 2012). According to the World Health Organization (WHO), South Africa has the third highest incidence of TB in the world. However, when one considers the incidence per capita, South Africa moves up to the highest burden country, a situation worsened by the fact that the country also has the highest rate of HIV co-infection with 65% of patients tested for TB being HIV positive (WHO, 2012). Further exacerbating the TB epidemic is the spread of drug resistant strains which makes a prolonged treatment program more difficult to manage. Estimates indicate that at the end of 2011, between 2.1-5.2% of new cases and 13-26% of previous cases globally were multi-drug resistant (MDR) TB (WHO, 2012), defined as resistant to isoniazid (INH) and rifampicin (RIF) (WHO, 2006). In addition, extensively drug resistant (XDR) TB, classified as MDR and also resistant to a fluoroquinolone as well as one or more of the second-line injectable drugs, kanamycin, capreomycin and amikacin (WHO, 2006), has been identified in 84 countries (WHO, 2012). The ability of *Mtb* to circumvent killing by the host immune response, rapidly develop drug resistance and persist during a prolonged state of clinical latency attest to its capacity to adapt to the variable conditions encountered during host infection (Warner and Mizrahi, 2008).

1.1. Tuberculosis: Prevention and treatment

When one considers the TB epidemic, it is astounding to note that TB continues to be a global health crisis despite the fact that there is a multiplicity of drugs for chemotherapeutic intervention and a pre-exposure vaccine. Currently, the Bacille Calmette-Guérin (BCG) vaccine, a live attenuated form of *Mycobacterium bovis*, is used for the prevention of TB with the WHO recommending administration of the vaccine to all neonates in high burden countries (WHO, 2004 and WHO, 2012). The vaccine has been shown to provide protection against TB meningitis and miliary TB in children, with protective estimates between 67-79% and 75-87%, respectively (Trunz *et al.*, 2006). However, the vaccine leads to complications in HIV-infected or otherwise immune-compromised children with the development of BCG disease and in many cases death, which further contributes to the high mycobacterial-associated infant mortality rates (Hesseling *et al.*, 2006). In addition, the protection against primary infection in adults is variable, at best, and the vaccine does not provide protection against reactivation in latently infected individuals who serve as reservoirs of disease in the community (WHO, 2004). This demonstrates that the use of the BCG vaccine alone is not a sufficient preventative measure – particularly in the South African context, where the infection pressure is very high –and highlights the need for a new and/or improved TB vaccine (Tameris *et al*, 2013).

The major control strategy for drug susceptible TB is *d*irectly *o*bserved *t*herapy *s*hortcourse (DOTS) which involves the controlled administration of drugs to patients over a period of 6 months. The first two months of intensive treatment involves a four-drug combination of first-line drugs, RIF, INH, ethambutol (EMB) and pyrazinamide (PZA); followed by a four-month continuation phase with RIF and INH (WHO, 2010). The treatment regimen for cases of MDR-TB is slightly different and highly dependent on the individual resistance profile of *Mtb*

obtained for each patient. However, the WHO recommends a minimum treatment period of 18 months, involving a combination of first-line drugs, injectables, fluoroquinolones and second-line bacteriostatic oral drugs (WHO, 2010). Although comprehensive TB control programmes are in place, several obstacles are still faced for TB treatment. The length of treatment as well as drug toxicity contributes to patient non-adherence, which fuels the development of drug resistance in circulating strains (Gandhi *et al*., 2006). In addition, interactions between TB drugs and antiretroviral therapy (ART) have made treating patients with dual infection very complicated (WHO, 2012).

As mentioned previously, a major outcome of infection with *Mtb* is clinical latency (latent TB infection, LTBI), which is a state in which *Mtb* is able to persist in a host without causing symptoms of disease, but with an associated risk of recrudescent infection to give rise to reactivation or post-primary TB. This phenomenon has been extensively debated in the field and a great deal of controversy surrounds the mechanisms underlying latency and reactivation. It was previously hypothesized that latent infection consisted of a viable population of *Mtb* cells which were not actively growing or "dormant"; however, in recent years a pool of evidence has gathered that disputes this notion. It has been shown that INH, which targets actively growing bacilli is effective in treating LTBI suggesting that active bacterial growth is ongoing during this type of infection (Barry *et al*., 2009). In addition, a pioneering study demonstrated that chromosomal mutations still arose in a population of *Mtb* during latent infection, suggesting that growth is taking place since DNA replication is required for mutations to be fixed (Ford *et al*., 2011). It has also been hypothesized that instead of a simple binary separation between latent and active disease, *Mtb* infection outcomes are better described as a spectrum ranging from complete clearance to subclinical active disease to full-

blown granulomatous disease (Robertson *et al.*, 2012). It therefore stands to reason that the growth of *Mtb* is dynamic over this spectrum and at any given point, a heterogenous bacterial population exists comprising actively growing, slow growing and dormant bacteria which would have implications for drug tolerance and the emergence of resistance.

As discussed above, a vaccine is available for the prevention of TB and comprehensive treatment plans are recommended for individuals who develop TB. However, these strategies have inherent problems which are a consequence of the limitations of the drugs and/or vaccine used. Consequently, a better vaccine and new drugs with novel targets need to be developed in order to address the issue of LTBI and attempt to curb the ongoing evolution and spread of drug resistance. However; that *Mtb* will be able to adapt and respond to any new intervention is without doubt. The best way to ensure effective eradication of TB, is to thoroughly understand how the tubercle bacillus survives under various conditions by studying its basic physiology and metabolism as these are inextricably linked to pathogenesis (Warner and Mizrahi, 2008)

1.2. *Mtb* infection and the host environment

Mtb infection begins by the inhalation of aerosol particles containing the infectious agent into the lung alveoli where the cells are met by the first line of defense and are phagocytosed by alveolar macrophages and dendritic cells. When exposed to *Mtb*, some individuals will completely clear the pathogen via a strong innate immune response. However, infection becomes established in other individuals with two outcomes being possible: either the development of active disease as observed in < 10% of cases, or the development of LTBI as is the case for the majority (90-95%) of infected individuals (Bhatt and Salgame, 2007). The

group who become latently infected has a 10% lifetime risk of developing active disease which is heightened to a risk of 10% per annum in the case of HIV infected or otherwise immune-compromised individuals. *Mtb* primarily replicates in resting macrophages where it is able to halt the phagolysosomal maturation thus preventing the conventional microbicidal method used by these professional killer cells. The pathogen is able to interact with toll-like receptors (TLRs) on the surface of macrophage and dendritic cells for phagocytosis. TLR-dependent uptake of *Mtb* induces cytokine and chemokine secretions thus recruiting other immune cells to the site of infection and initiating granuloma formation. In addition, this mechanism of uptake induces maturation of dendritic cells allowing them to become efficient antigen presenting cells (APCs), which then migrate to the draining lymph nodes where antigen is presented to naïve T cells thus initiating the adaptive immune response (Bhatt and Salgame, 2007).

Granuloma formation is directly dependent on the adaptive immune response with the ratio of pro- and anti-inflammatory reactions determining the outcome. Three common types of granulomas have been identified and are characterized based on their structure and immune cell composition. Caseous granulomas are the classic type found in both active and latent disease and are thought to arise due to an increased pro-inflammatory response. This type of granuloma is made up of epithelial macrophages, neutrophils, lymphocytes ($CD4^+$ and $CD8^+$ T cells and B cells) and fibroblasts. The center of this granuloma comprises a hypoxic, caseous environment made up of dead macrophages and other cells. *Mtb* is usually contained within this environment unless the individual becomes immune compromised, in which case active disease develops rapidly and non-necrotizing granulomas are observed. These granulomas are primarily made up of macrophages and some lymphocytes but are not structured like the

caseous granulomas and allow for the dissemination of *Mtb* in the lung. The third type of granuloma is the fibrotic granuloma mostly observed in LTBI. It is hypothesized that this type of granuloma arises when there is a balance between the pro- and anti-inflammatory immune responses and is made up primarily of fibroblasts and some macrophages. Calcification is also observed in this type of granuloma wherein *Mtb* is able to survive for extended periods of time (Barry *et al.,* 2009; Scanga and Flynn, 2010).

1.2.1. *Mtb* adaptations for survival

Upon infection, the human immune system mounts an attack on *Mtb* involving both the innate and adaptive immune responses; however, the pathogen is able to counter these attacks and survive in the host highlighting how well equipped *Mtb* is for the harsh conditions encountered during infection. As mentioned above, *Mtb* is able to replicate in resting macrophages due to its ability to arrest phagosome maturation which is characterized by acidification and the production of reactive oxygen and nitrogen intermediates (ROI and RNI respectively). Upon activation of the macrophage with interferon gamma (IFNγ), this process is allowed to continue (MacMicking *et al.,* 2003) suggesting that the *Mtb* cells contained within would be killed. It is widely accepted that *Mtb* is challenged with a variety of factors in the host including oxygen depletion, an acidic pH, reduction in the amount and availability of nutrients as well as oxidative and nitrosative stress (Baek *et al.*, 2011).

Oxidative and nitrosative stress production in the phagosome is a particularly efficient method for macrophages to deal with pathogens because the ROIs and RNIs target a range of macromolecules including DNA, proteins, lipids and carbohydrates (Ehrt and Schnappinger, 2009). *Mtb* however has evolved several mechanisms to deal with this by detoxification of

ROI and RNI, as well as DNA and protein repair when detoxification is insufficient (Ehrt and Schnappinger, 2009). *Mtb* possesses several mechanisms for the detoxification of ROI and RNI including the decomposition of H_2O_2 by the KatG catalase-peroxidase into water and oxygen (Ng *et al*., 2004). This mechanism was validated by the fact that the *katG* mutant is attenuated in wild type mice but virulent in mice unable to generate ROIs (Ng *et al*., 2004). In addition to the direct detoxification of ROI and RNI, *Mtb* also has mechanisms which allow for the repair of DNA damage caused by these molecules. In a study carried out by Darwin and Nathan (2005) the *Mtb* nucleotide excision repair pathway was identified as being important for resistance to RNI and ROI with a *uvrB* mutant showing reduced growth in wild type mice but full virulence in mice defective for RNI and ROI production.

One of the other conditions encountered by *Mtb* in the granuloma is hypoxia which is thought to be the driving force behind the reduced metabolic activity that is postulated to occur during LTBI (Boshoff and Barry, 2005). Using *in vitro* models of hypoxia, most notably the Wayne Model, it has been shown that *Mtb* is able to persist under anaerobic conditions when oxygen is gradually depleted from the system and it is thought that this same phenomenon takes place *in vivo* (Wayne, 1994). The survival of *Mtb* under conditions of the Wayne Model has been attributed to the pathogen's ability to switch from aerobic respiration, where oxygen serves as the terminal electron acceptor, to anaerobic respiration where nitrate/fumarate serves as terminal electron acceptor (Wayne and Hayes, 1998). In addition to allowing for survival during hypoxia, nitrate respiration has also been shown to provide protection against acidity and RNIs (Tan *et al.*, 2010).

The inherent characteristics of *Mtb* discussed above highlight the versatility of the pathogen's basic physiology and metabolism which allows for it to adapt to the various conditions

encountered in the host during infection. It is therefore reasonable to assume that numerous metabolic pathways, and the interplay between them, are important during pathogenesis and a greater understanding of these would aid the quest of developing better control strategies for TB. This study focuses on the biosynthesis of molybdenum cofactor (MoCo), which has been predicted to be important for survival of the organism during pathogenesis, particularly when nitrate is available for growth and/or anaerobic respiration.

1.3. Molybdenum

Molybdenum (Mo) is a trace element required for the activation of several enzymes in organisms across all three orders of life. The chemical versatility of Mo makes it ideal for use as a cofactor, forming the catalytic centre of several enzymes which catalyze redox reactions involved in global carbon, nitrogen and sulfur metabolism (Williams *et al*., 2011). Mo serves as an electron sink and is able to change oxidation states under physiologically relevant conditions ranging from oxidation states VI to IV allowing for one- and two-electron reduction-oxidation reactions to be catalyzed (Hille, 2002).

1.3.1. Molybdoenzymes

Mo acquires biological activity once it is incorporated into a cofactor and there are two main categories of Mo-containing cofactors: iron-molybdenum cofactor (FeMo-co) which is unique to bacterial nitrogenases; and molybdenum cofactor (MoCo) which is a pterin based cofactor found in the remaining molybdoenzymes (Schwarz *et al.*, 2009). MoCo-dependent enzymes are further categorized into three subfamilies based on the coordination of MPT to the metal and the presence of additional side groups. The sulfite oxidase (SO), xanthine oxidase (XO)

and dimethylsulfoxide reductase (DMSOR) families as well as the structures of cofactors present in each family are shown in Figure 1.1.

FeMo-co nitrogenase is responsible for the biological fixation of nitrogen via the reduction of atmospheric dinitrogen (N_2) to ammonia (NH_3) (Hernandez *et al.*, 2009). The active site of FeMo-co is a complex structure with a central light atom (C, N or O) coordinated by iron-sulfur (Fe-S) clusters and capped by a Mo atom which is further coordinated by a homocitrate ligand (Hernandez *et al.*, 2009). The basic form of MoCo is a tricyclic pyranopterin with a Mo atom coordinated to the dithiolene side chain group of molybdopterin or metal-containing pterin (MPT) (Iobbi-Nivol and Leimkühler, 2012). The SO family cofactor has the same basic structure of MoCo with an additional cysteine ligand on the metal and members of this family catalyze the transfer of an oxygen atom to or from a substrate (Hille, 1996; Iobbi-Nivol and Leimkühler, 2012). The cofactor from the XO family is also characterized as having the basic MoCo structure with a sulfide group in place of oxygen and an additional hydroxyl group on Mo and enzymes from this family catalyze hydroxylation and oxo-transfer reactions with water as the source of oxygen (Iobbi-Nivol and Leimkühler, 2012). In addition to all forms of MoCo in the XO family being sulfurated, another modification is observed with the incorporation of cytosine to form molybdopterin-cytosine dinucleotide cofactor (MCD) (Neumann *et al*., 2009a). The DMSOR family of MoCo-dependent enzymes is the largest and most diverse group in terms of structure and function. A modified form of the cofactor is present in this family with two MPT molecules carrying a guanine nucleotide on each carboxy terminus being coordinated around a single Mo atom thus referred to as *bis*-molybdopterin guanine dinucleotide (*bis*-MGD) (Hille, 1996).

1.4. MoCo-dependent enzymes in mycobacteria

Bioinformatic analysis reveals that *Mtb* possesses eight molybdoenzymes, the majority of which are predicted to utilize the *bis*-MGD form, whereas the proteome of the model organism, *Mycobacterium smegmatis* contains twenty identifiable molybdoenzymes (Williams *et al*., 2013).

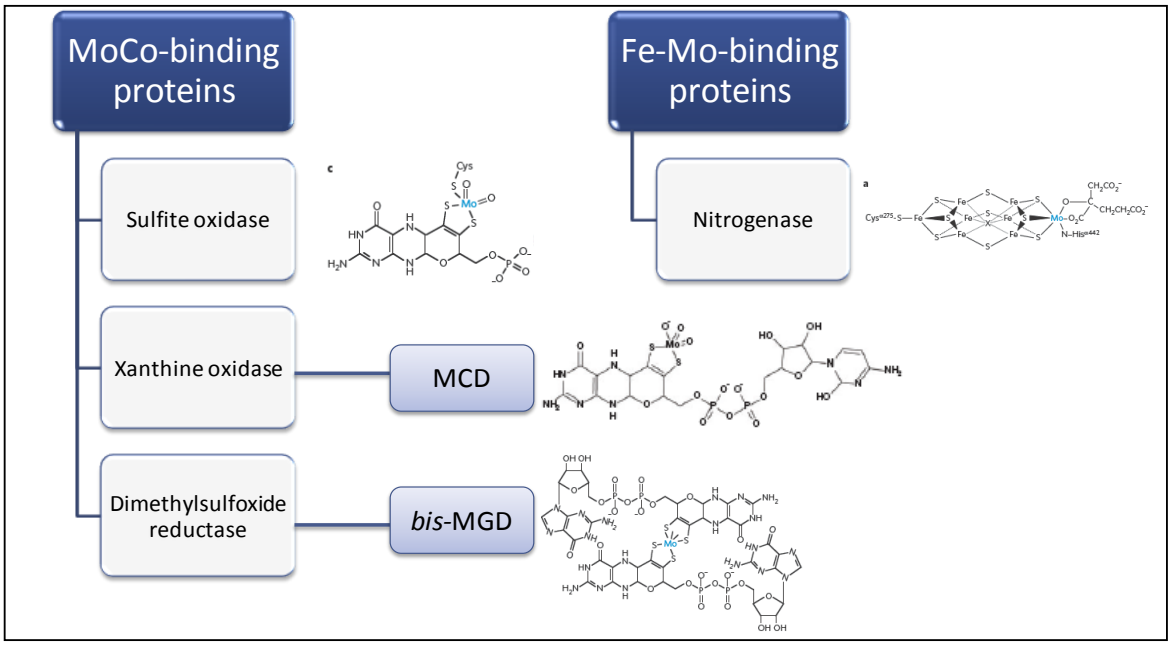


Figure 1.1: Categories of molybdenum-containing enzymes and the structure of the cofactors present in each. Molybdopterin-cytosine dinucleotide cofactor (MCD), *bis*-molybdopterin guanine dinucleotide cofactor (*bis*-MGD).

1.4.1. *Mtb* molybdoenzymes and pathogenesis

Two recent reviews have summarized the molybdoenzymes in mycobacteria and their roles in the physiology and pathogenesis of *Mtb* (Shi and Xie, 2011; Williams *et al.*, 2013). Molybdoenzymes in mycobacteria, as in most other organisms, catalyze diverse reactions highlighting a role for the cofactor in several areas of metabolism and physiology. It is noteworthy that of the eight molybdoenzymes in *Mtb*, four are implicated in pathogenesis (Williams *et al*., 2013). The most well characterized of these is the respiratory and assimilatory nitrate reductase (NR) enzyme encoded by *narGHI* (Malm *et al*., 2009). As

discussed above, nitrate respiration by MoCo-dependent NR enables *Mtb* to survive under hypoxia and provides protection against acidity and RNIs (Tan *et al.*, 2010). In addition, several lines of evidence exist implicating this molybdoenzyme in virulence: (I) The reduced fitness observed for clinical isolates of *Mtb* and *Mycobacterium africanum* during macrophage infections was associated with under-expression of *narG* and a lack of induction of the *narGHJI* operon (Homoloka *et al.*, 2010) (II) There is an increase in NR activity when the pathogen is exposed to hypoxia, nitric oxide (NO) or carbon monoxide (CO) which is attributed to the induction of *narK2*, the nitrate transporter that forms part of the Dos regulon which is important for adaptation to harsh conditions and long term survival in the host (Sohaskey and Wayne, 2003; Voskuil *et al*., 2003; Sohaskey, 2005; Shiloh *et* al., 2008; Sohaskey and Modesti, 2005). The MoCo-dependent NuoG is one of fourteen subunits of the Type I NADH dehydrogenase complex and has been shown to be involved in *Mtb* virulence due to its ability to inhibit macrophage apoptosis thus prolonging survival in the host (Velmurugan *et al*., 2007). Through a 'gain-of-function' assay it was shown that *nuoG* was able to increase the virulence of *Mycobacterium kansasii*, a facultative-pathogenic strain of mycobacteria with this increased virulence being attributed to increased inhibition of apoptosis (Velmurugan *et al*., 2007). In addition, the study also showed that the *Mtb* Δ*nuoG* mutant was less virulent than wild type with Δ*nuoG*-infected SCID mice surviving twice as long as their wild type infected counterparts. In a separate investigation, it was shown that macrophages infected with the *Mtb* Δ*nuoG* mutant were able to accumulate NOX-2 mediated toxic ROI which leads to TNF-α secretion and ultimately host cell death. It was therefore hypothesized that the mechanism of NuoG-mediated apoptosis inhibition in macrophages was due to an interference with NOX-2 sensing of intracellular *Mtb* thus inhibiting an inflammatory response (Miller *et al*., 2010). More recently it was shown that *nuoG* plays a role in inhibition

of neutrophil apoptosis (Blomgran *et al*., 2012). In that study, infection of neutrophils with the *Mtb* Δ*nuoG* mutant resulted in accelerated neutrophil death, increased trafficking of *Mtb*-containing dendritic cells to the lymph nodes, and thus, faster *Mtb* specific CD4 T cell priming, leading to a delay in the onset of the adaptive immune response and prolonging survival of *Mtb* (Blomgran *et al*., 2012).

Another molybdoenzyme implicated in *Mtb* pathogenesis is carbon monoxide dehydrogenase (CDH) which catalyzes the oxidation of CO to CO_2 which can be used for assimilation with the electrons being fed into the respiratory electron transport chain (Oh *et al*., 2010). A study carried out by Shiloh *et al.* (2008) showed that *Mtb* is exposed to CO during macrophage infections and this in turn induces the dormancy regulon in the pathogen. CO can be toxic to bacterial cells, however, *Mtb* is able to grow aerobically in the presence of CO, and other mycobacterial organisms (*M. smegmatis*, *M. bovis* BCG and *Mtb* H37Ra) are able to oxidize the gas at physiological concentrations, thus implicating a role for CDH in mycobacterial physiology and possibly pathogenesis (Park *et al.*, 2003; King, 2003). Further highlighting a role for CDH in pathogenesis is the ability of the *Mycobacterium* sp. strain JC1 enzyme to function as a NO dehydrogenase thus protecting cells against the bactericidal activity of NO (Park *et al*., 2007).

Biotin sulfoxide reductase, another MoCo-dependent enzyme encoded by *bisC* in *Mtb*, is responsible for the reduction of spontaneous oxidation products of biotin (del Campillo-Campbell and Campbell, 1982) which is an important cofactor required for the synthesis of fatty acids in the cell envelope of *Mtb* (Woong Park *et al*., 2011). A recent study investigated the biotin biosynthetic pathway, and with the use of a Δ*bioA* mutant found that *de novo* biosynthesis of the cofactor is required for establishing and maintaining infection in a mouse

model (Woong Park *et al*., 2011) thus identifying a role for biotin in pathogenesis. The remaining molybdoenzymes in *Mtb* include the fused NR encoded by *narX*, the possible oxidoreductase Rv0197, the probable transmembrane protein Rv0218 and formate dehydrogenase (FdhF), the functions of which have yet to be determined (Williams *et al*., 2013).

1.4.2. *M. smegmatis* molybdoenzymes

As mentioned previously, *M. smegmatis* contains far more putative molybdoenzymes than *Mtb*. In addition to harboring three homologues of CDH (MSMEG_0746, MSMEG_2949 and MSMEG_2462), *M. smegmatis* also possesses homologues for each of the other molybdoenzymes implicated in *Mtb* pathogenesis. The additional enzymes include six putative oxidoreductase enzymes, a competence damage-inducible protein (MSMEG_3521), anaerobic dehydrogenase (MSMEG_2237), nicotine dehydrogenase (MSMEG_5880), MSMEG_0684 annotated as aldehyde oxidase and xanthine dehydrogenase and a putative assimilatory NR, NarB (MSMEG_2837) (Williams *et al*., 2013). Using specific inhibitors, previous studies demonstrated that *M. smegmatis* possesses both respiratory (Khan and Sarkar, 2006) and assimilatory NR activity, with NarB being identified as the putative assimilatory enzyme which facilitates growth of *M. smegmatis* on media with nitrate as sole nitrogen source (Khan *et al.*, 2008).

1.5. MoCo biosynthesis

MoCo biosynthesis has been studied extensively in *Escherichia coli* and unless otherwise stated, discussions on the biosynthetic pathway are based on experiments carried out in, and results obtained from, this organism. Both *Mtb* and *M. smegmatis* possess the full set of genes

required for MoCo biosynthesis and this pathway is highly conserved between these mycobacterial species and *E. coli*, suggesting that some of the conclusions drawn from *E. coli* studies can be extrapolated to these mycobacteria.

1.5.1. Molybdenum uptake

Mo is bioavailable as molybdate (MoO_4^{2-}) and a specialized transport system exists in bacteria for its uptake from the environment. The transport system is encoded by the *modABC* operon which is also present in *Mtb*. The ATP-binding cassette (ABC) transport system, shown in Figure 1.2 is made up of ModA which is the periplasmic molybdate binding protein; ModB which with its numerous hydrophobic regions forms a homodimer across the membrane making the channel through which molybdate is transported into the cytoplasm; and ModC that has an ATP binding motif and functions as an ATPase, providing energy for molybdate transport (Grunden and Shanmugam, 1997). Transport of the ion is regulated by ModE which acts as a repressor of the *modABC* operon when bound to molybdate (Grunden *et al*., 1996)

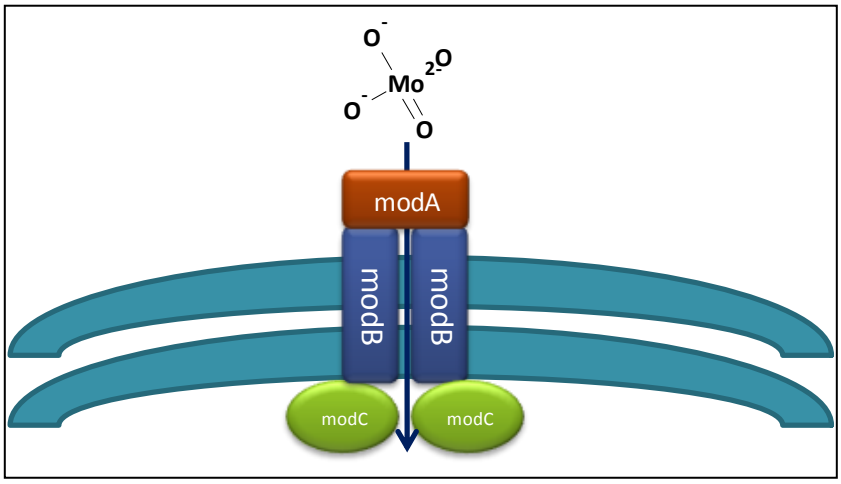


Figure 1.2: Schematic representation of the molybdate ion transport system. The periplasmic molybdate binding protein, ModA is depicted in red, the ModB homodimer membrane channel is shown in blue and the ModC ATPase subunits are shown in green.

Once in the cell, molybdate is incorporated into the cofactor as described below.

1.5.2. MoCo biosynthetic pathway

The MoCo biosynthesis pathway is a highly conserved multi-step pathway which requires the input of several gene products at each step in order to be functional. In bacteria, the process

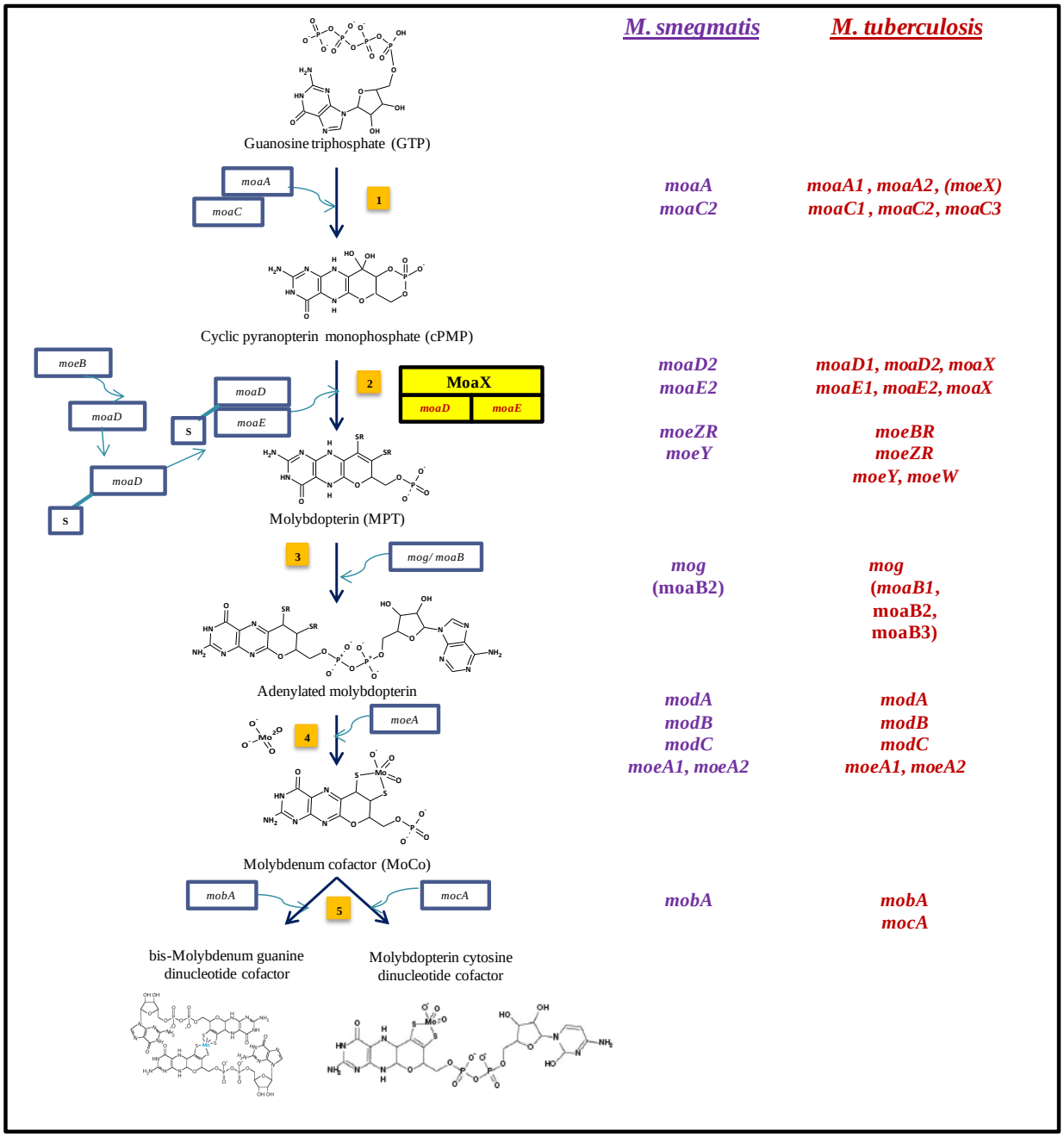


Figure 1.3: The highly conserved multi-step MoCo biosynthetic pathway in bacteria. Steps 1-5 are labeled in yellow boxes and the structures and names of each substrate and product are shown. The pathway is labeled according to the *E. coli* nomenclature for each of the genes involved (boxed in blue) while the *M. smegmatis* and *Mtb* homologues are listed in purple and red respectively. The black-bordered yellow box highlights *Mtb* MoaX which has been identified as a fused MPT synthase with homology to MoaD and MoaE (Williams *et al*., 2011).

takes place over five steps shown in Figure 1.3, namely: (1) conversion of guanosine triphosphate (GTP) to cyclic pyranopterin monophosphate (cPMP); (2) generation of molybdopterin (MPT); (3) MPT adenylation; (4) insertion of Mo to generate MoCo and (5) MoCo maturation. The first step of the pathway is catalyzed by the products of the *moaA* and *moaC* genes. MoaA is a member of the S-adenosylmethionine (SAM)-dependent superfamily of enzymes which catalyze the production of radical products by the reductive cleavage of SAM (Sofia *et al.*, 2001), while the function of MoaC has yet to be completely characterized. Through the action of these two gene products, GTP is cleaved and carbon atoms are rearranged to form the pyranopterin structure cPMP (Iobbi-Nivol and Leimkühler, 2012). As mentioned previously MoCo has a tricyclic pyranopterin structure and therefore the formation of MPT is critical for MoCo biosynthesis. cPMP is structurally similar to MPT but is lacking the dithiolene groups essential for the attachment of Mo to the cofactor. For the conversion of cPMP to MPT at the second step of the pathway, two sulfur atoms need to be incorporated at positions C1' and C2' and this reaction is catalyzed by the enzyme MPT synthase, encoded by the *moaD* and *moaE* genes. Additional reactions involving *moeB* and the adenylation and re-sulfuration of MoaD are required for the continual functioning of MPT synthase and are discussed more thoroughly below. The third and fourth steps of the pathway involve the activation of MPT and incorporation of molybdate respectively. MPT activation occurs as a result of adenylation by MogA in *E. coli*, however under high molybdate concentrations this step is not required and the metal can be directly inserted into MPT with the aid of MoeA, while under low concentrations MPT adenylation doubles the rate of metal insertion, thus providing evidence that MoeA mediates molybdate insertion and MogA enhances it (Neumann and Leimkühler, 2008; Nichols and Rajagopalan, 2005). Once formed, MoCo can undergo further modifications in bacteria with the addition of either a cytosine or two guanine

nucleotides to generate MCD or *bis*-MGD forms of the cofactor respectively. MCD formation is catalyzed by the gene product of *mocA* which encodes the cytidylyltransferase enzyme that specifically incorporates CMP onto the C4' phosphate of MoCo (Neumann *et al*., 2009b). The study also showed that the *E. coli mocA* enzyme was specific for CTP and was unable to utilize the nucleotides ATP or GTP (Neumann *et al*., 2009b). Recently it has been shown that *Mtb* possesses a *mocA* homologue suggesting that the pathogen is able to synthesize this form of the cofactor for CDH (Williams *et al*., 2013). *bis*-MGD is the more commonly found form of the cofactor in molybdoenzymes and is synthesized by the incorporation of GMP onto the C4' phosphate of MPT, as catalyzed by MobA. In *E. coli mobA* is expressed as part of an operon with *mobB*, the product of which was thought to function as an adapter protein for efficient *bis*-MGD synthesis (Iobbi-Nivol *et al*., 1995). However, evidence shows that MobA alone is sufficient for *bis*-MGD formation suggesting that the role of MobB in this reaction is not essential (Palmer *et al*., 1996). *Mtb* possesses only a single *mobA* homologue and together with the presence of *bis*-MGD-dependent enzymes, this further validates that MobA alone is required for *bis*-MGD synthesis.

1.6. MoCo and *Mtb* pathogenesis

The interest in MoCo biosynthesis is further fueled by the fact that in addition to selected molybdoenzymes being implicated in pathogenesis, several forward genetic, genome-wide mutagenesis screens have identified genes directly involved in MoCo biosynthesis as important for survival, pathogenesis and virulence of *Mtb*. In a study carried out by Brodin *et al*. (2010), independent insertions in *moaD1* and *moaC1* resulted in the inability of these mutants to arrest phagosome maturation, which, as discussed above, is important for *Mtb* pathogenesis. A similar phenotype was observed for an insertion mutant of *moeB1* which was

also defective for growth in macrophages (Macgurn and Cox, 2007), much like the *moaD1* and *moaC1* mutants. The *moaC1* mutant has also been shown to be attenuated in its ability to parasitize macrophages (Rosas-Magallanes *et al.*, 2007) as well as being attenuated in primate lungs (Dutta *et al.*, 2010). In the study published by Rosas-Magallanes *et al*. (2007), it was also shown that a *moaX* insertion mutant was unable to parasitize macrophages. Most recently, *moaD1* has been identified as playing a role in resistance to ROI (Mestre *et al.*, 2013). The study showed that the *moaD1* mutant was hyper-susceptible to oxidative stress from H_2O_2 and was attenuated for growth in macrophages (Mestre *et al.*, 2013). Another gene involved in MoCo biosynthesis and implicated in *Mtb* pathogenesis is *modA*, which when mutated, is attenuated for growth in mice (Camacho *et al.*, 1999). These findings highlight the importance of MoCo for *Mtb* pathogenesis.

1.7. Expansion of MoCo biosynthetic genes in *Mtb*

The MoCo biosynthetic genes in *Mtb* are distributed throughout the chromosome and in some cases, are on operons (Figure 1.4).

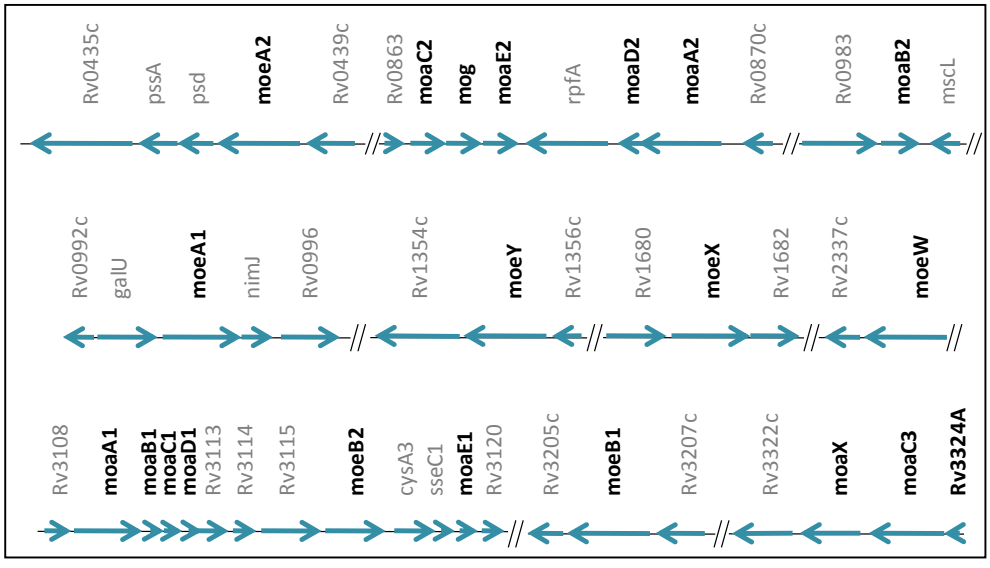


Figure 1.4: Chromosomal distribution of *Mtb* genes involved in MoCo biosynthesis. Genes shown in bold are predicted by homology to be involved inMoCo biosynthesis.

Although the MoCo biosynthetic pathway is highly conserved among most organisms, the pathway in *Mtb* and members of the *Mtb* complex (*MTB*C) is distinguished in that there are multiple homologues for several genes involved in the first and second steps of the pathway (Figure 1.3). Also notable is that the repertoire of MoCo biosynthesis genes in *M. smegmatis* is much simpler than *Mtb* although it possesses a larger complement of putative MoCo-dependent enzymes. This raises the question of whether the Mtb homologues are all in fact functional, and if so, what cellular function is served by the expansion.

1.8. MPT-synthase

The second step of the biosynthetic pathway is catalyzed by the enzyme MPT synthase. The canonical MPT synthase is a heterotetrameric structure made up of two large subunits of MoaE and two small subunits of MoaD which embed into pockets of MoaE to form the active site of the enzyme (Rudolph *et al*., 2001). A schematic image of MPT synthase is shown in Figure 1.5. In its active form, MoaD is thiocarboxylated at the terminal glycine (Gly) residue (Pitterle and Rajagopalan, 1993; Rudolph *et al.*, 2003). During the formation of MPT, the sulfur atom on the C-terminus of MoaD is transferred to the cPMP substrate in the pocket of MoaE, the volume of which is sufficient to accommodate both the C-terminus of MoaD and MPT or cPMP (Rudolph *et al*., 2003).

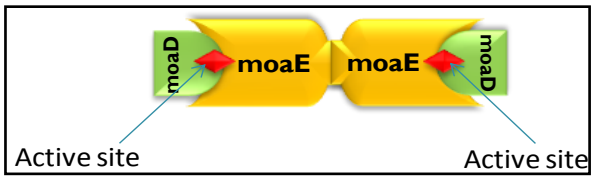


Figure 1.5: Schematic representation of the structure of MPT synthase. MoaE subunits are shown in yellow and MoaD in green. The red diamond represents the active site of the enzyme found embedded in the pocket of MoaE. Each MPT synthase heterotetramer has the capacity for two catalytic reactions.

In a study carried out by Wuebbens and Rajagopalan (2003), MoaE residues important for MPT synthase formation and function were identified. The pocket of MoaE which forms the active site of MPT synthase is lined with highly conserved arginine (Arg), lysine (Lys) and histidine (His) residues, of which Lys119 has been shown to be essential for MPT synthase activity with His103 and Arg104 being important for the electrostatic stability of the C-terminal thiosulfate in MoaD (Rudolph *et al*., 2001; Wuebbens and Rajagopalan, 2003). MoaD, like ubiquitin, contains a C-terminal Gly-Gly motif which has been shown to be important for the function and stability of MPT synthase (Schmitz *et al*., 2007). Deletion, insertion and substitution mutants of the MoaD terminal Gly residues revealed that the terminal Gly81, but not Gly80, was essential for MPT synthase activity. In addition, substitution of either residue did not affect the ability to form a MoaD-MoaE heterodimer complex, although the G81A substitution significantly slowed down the process by 60 %, thus identifying the terminal Gly81 as essential for optimal MPT synthase function (Schmitz *et al*., 2007).

In order for MPT synthase to remain catalytically active, the sulfurs on the C-terminus of MoaD need to be regenerated in a series of reactions catalyzed by the *moeB* gene product (Pitterle and Rajagopalan, 1993; Leimkühler *et al.*, 2001). This procedure involves multiple steps on its own and requires that MoaD forms a complex with MoeB similar to that observed with MoaE (Lake *et al*., 2001). MoeB is responsible for the ATP-dependent activation of the MoaD C-terminus to form MoaD-adenylate (Gutzke *et al*., 2001) which is then sulfurated by the cysteine desulfurase IscS in *E. coli* (Zhang *et al*., 2010). IscS however acts as a sulfur donor in various biosynthetic reactions including biotin, thiamin and lipoic acid (Marquet,

2001). It has recently been shown that IscS does not function alone but in conjunction with a rhodanese-like protein YnjE in order to direct IscS to MoCo biosynthesis (Dahl *et al*., 2011).

1.8.1. *Mtb* MPT synthase

A study carried out in the MMRU attempted to investigate the phenomenon of MoCo biosynthetic gene multiplicity with some intriguing results (Williams *et al*., 2011). Firstly it was shown that the *moaA1-moaB1-moaC1-moaD1* operon was dispensable for growth *in vitro* under standard laboratory conditions contradicting a previous study which suggested that this operon is essential for growth of *Mtb in vitro* (Sassetti *et al*., 2003). This discrepancy could possibly be attributed to the identification of essential genes through saturating transposon mutagenesis and screening which does not account for competitive growth selection (Williams *et al*., 2013). An assay measuring MoCo biosynthesis was developed which relies on the activity of the MoCo-dependent NR enzyme (Williams *et al*., 2011). The premise of this assay is that NR activity relies on the availability of MoCo which, when present, would be incorporated into the enzyme allowing for its respiratory and assimilatory activity to be measured in the presence of nitrate. However, reduced (or abrogated) MoCo production would result in no NR activity being observed. Hence NR activity, as measured by nitrate assimilation (*i.e*., growth on nitrate as sole nitrogen source) serves as a surrogate measure of MoCo biosynthesis. An important point to note is that this assay allows for both respiratory and assimilatory NR activity to be measured when nitrate is provided as a sole nitrogen source. It was observed that the *Mtb* knock-out mutant lacking both *moaD1* and *moaD2* was severely attenuated for growth on nitrate as a sole nitrogen source but still displayed residual respiratory NR activity which was attributed to MoaX, a protein which has both MoaD and MoaE functional domains suggesting that it may act as an MPT synthase (Williams *et al*.,

2011). These data confirm that the NR has greater sensitivity for measuring respiratory compared with assimilatory NR activity. As shown in Figure 1.3, *M. smegmatis* possesses a much simpler MoCo biosynthetic pathway than *Mtb*. *M. smegmatis* knock-out mutants deficient in *moaD2* (Δ*moaD2*), *moaE2* (Δ*moaE2*) or both (Δ*moaD2* Δ*moaE2*) were all defective in MoCo biosynthesis, as measured by assimilatory NR activity. This set of strains thus provided an ideal resource for interrogating the function of the different *Mtb moaD* and *moaE* homologues. Using heterologous complementation, Williams *et al.* (2011) showed that only the *Mtb moaD2*, *moaE1* and *moaE2* were functional MPT synthase-encoding genes, as evidenced by their ability to restore MoCo biosynthesis in the *M. smegmatis* mutant strains. It was hypothesized, at the time that the lack of activity of *moaD1* in *M. smegmatis* was due to the absence of a cognate *moeBR* homologue which may be responsible for its adenylation (Williams *et al*., 2011). However, a subsequent study showed that both *moeBR* and *moeZR* were able to catalyze the sulfuration of *moaD1* and *moaD2* (Voss *et al*., 2011). Williams *et al*. (2011) also identified MoaX as a fused, functional MPT synthase enzyme that was able to restore MoCo biosynthesis in all three MoCo-deficient *M. smegmatis* mutants.

1.9. Aims

The results reported by Williams *et al*. (2011), when considered in context of the canonical MPT synthase structure (Rudolph *et al*., 2001), raised several interesting questions about MoCo biosynthesis in mycobacteria, specifically with respect to the genes encoding MPT synthase. Firstly, do the different *Mtb moaD* and *moaE* homologues combine to form chimeras of the enzyme with differing activities? Furthermore, is the fused MPT synthase encoded by *moaX* cleaved to form a functional enzyme? In addition, the assay used by Williams *et al*. (2011), and throughout this study, relies on the activity of a previously uncharacterized

assimilatory NR. These questions formed the basis of this study, which had the following aims:

- To assess the function and relative contributions of the multiple *Mtb moaD* and *moaE* homologues to MoCo biosynthesis
- To construct a FLAG-tagged derivative of MoaX to determine whether it is post-translationally cleaved to form a functional MPT synthase
- To construct a Δ*narB* mutant in order to confirm that growth in nitrate media is due to the MoCo-dependent NarB

2 Methods

2.1 Bioinformatic tools and software

Several bioinformatic tools and software packages were used throughout this study to identify and analyze the genes and proteins of interest (Appendix A).

2.2 Chemicals and reagents

All enzymes used during this study for molecular DNA modifications were supplied by New England Biolabs, Fermentas and Roche. Unless otherwise stated, all primers were obtained from Inqaba Biotech. Reagents used for protein-based experiments were obtained from Thermo Fischer Scientific and Sigma Aldrich. For a detailed list of chemicals and reagents used, please refer to Appendix B.

2.3 Bacterial strains and culture conditions

E. coli DH5α was used for plasmid propagation and was grown in Luria-Bertani broth (LB) or agar (LA) supplemented with the appropriate antibiotics at concentrations of 100 µg/ml ampicillin (Amp), 200 µg/ml hygromycin (Hyg) and 50 µg/ml kanamycin (Kan). Cultures carrying plasmids < 8 kb were incubated at 37º C and those > 8 kb were propagated at 30º C (New Brunswick Scientific Innova™ 4000), with shaking for liquid cultures. The plasmids used and generated in this study are listed in Table D1.

Wild type *M. smegmatis* mc^2 155 and derivative strains were grown in Middlebrook 7H9 liquid medium (Difco) supplemented with Middlebrook oleic acid-albumin-dextrose-catalase (OADC) enrichment (Difco), 0.2% glycerol and 0.05% Tween80 with shaking or on Middlebrook 7H10 solid medium (Difco) supplemented with 0.085% NaCl, 0.2% glucose and 0.5% glycerol. Media

for growth of *M. smegmatis* was supplemented with antibiotics at concentrations of 50 µg/ml Hyg and/or 25 µg/ml Kan where appropriate. *M. smegmatis* strains used and generated in this study are listed in Table 2.4, 2.6 and 2.7.

2.4 Bacterial transformations

2.4.1 *E. coli* transformations

Chemically competent *E. coli* cells were prepared and transformed as previously described (Sambrook *et al.*, 1989). Briefly, *E. coli* cells were grown to mid-log phase (OD_{600} ~0.4) in 100 ml of 2×TY. Cells were cooled on ice for 15 min and harvested by centrifugation at 4˚ C and 3 901×*g* (Beckman Coulter Allegra™ X-22R Centrifuge) for 5 min. Cell pellets were re-suspended in 0.4 (original culture) volumes of TfbI (30 mM potassium acetate, 100 mM rubidium chloride, 10 mM calcium chloride, 50 mM manganese chloride, 15% v/v glycerol) and cooled on ice for 15 min. Cells were harvested as before and the pellet re-suspended in 0.04 (original culture) volumes of TfbII (10 mM MOPS, 75 mM calcium chloride, 10 mM rubidium chloride, 15 % v/v glycerol) followed by cooling on ice for 15 min. Two hundred and fifty to 500 µl aliquots of the cell suspension were prepared and used immediately or quick-frozen in ethanol prior to storage at - 80˚ C. Plasmid DNA was added to 100 µl of competent cells and incubated on ice for 15 min. This was followed by a heat shock step at 42˚ C for 90 sec, incubation on ice for 3 min, addition of 800 µl 2×TY and incubation at 37˚ C for 1 hour to allow for the phenotypic expression of antibiotic resistance genes. Transformants were subsequently selected on media containing the appropriate antibiotics. The recipes for all the solutions and media used can be found in Appendix B.

2.4.2 *M. smegmatis* electroporation

Electro-competent *M. smegmatis* cells were prepared as previously described by Larsen (2000). Cells were grown in 50-100 ml of media to mid-log phase (OD_{600} 0.5-0.8) and harvested at 2 360 ×*g* and 4˚ C for 10 min. The bacterial cell pellet was re-suspended in 10 ml of cold 10% glycerol and harvested as before; this wash step was repeated twice more. Cells were harvested and pellets re-suspended in 1 ml of glycerol which was then separated into 100 µl aliquots in Eppendorf tubes. A final wash of the cells was carried out at 12 470 × *g* (Beckman Coulter Microfuge 16) for 1 min and pellets re-suspended in 100 µl of 10 % glycerol. These competent cells were used immediately for electroporations. Approximately 1 µg of plasmid DNA was used for electroporations. The DNA was added to 0.2 cm electroporation cuvettes into which 100 µl of electro-competent *M. smegmatis* cells were added and gently mixed. The BIO-RAD Gene Pulser XCell™ system was used to perform electroporations with the following parameters: 2 500 V, 25 µF, 1 000 Ω, 0.2 cm. Immediately after the pulse, 800 µl of 2×TY was added and cells were incubated at 37˚ C for 3- 16 hr to allow for the phenotypic expression of selectable marker genes. Transformed cells were selected on 7H10 plates containing the appropriate antibiotics and/or selective supplements for 3-7 days.

2.5 DNA extraction methods

2.5.1 Mini-prep plasmid DNA extraction

E. coli cultures carrying the plasmid of interest were grown overnight in 2 ml LB with the appropriate antibiotics, 1 ml of which was used for the extraction procedure. Cultures were spun down at 12 470 × *g* and pellets re-suspended in 100 µl of Solution I (50mM glucose, 25mM Tris-HCl (pH 8), 10 mM EDTA) followed by the addition of 200 µl of Solution II (1% SDS, 0.2 M

NaOH) which was mixed by gentle inversion and incubated at room temperature for 5 min. Finally, 150 µl of Solution III (3 M potassium acetate, 11.5% acetic acid) was added and incubated on ice for 5 min. The suspensions were centrifuged for 5 min at 12 470 × *g* after which time the supernatant was decanted into a fresh Eppendorf tube. At this point, 1 µl of RNaseA (10 mg/ml) was added to the supernatant and incubated at 42˚ C for 15 min. Plasmid DNA was then precipitated by the addition of 600 µl of isopropanol, washed with 70% Ethanol and re-suspended in 100 µl of sdH_2O. This was followed by ethanol precipitation of the DNA by addition of 1/10 volume of 3 M sodium acetate (pH 5.3), followed by 3× volume 100% ethanol (-20˚C). DNA was collected by centrifugation at 12 470 × *g* for 20 min, washed with 70% ethanol and dried at 60˚ C for ~10 min in the Eppendorf Concentrator 5301. The dried pellet was then re-suspended in sdH_2O and quantified using the NanoDrop.

2.5.2 Maxi-prep plasmid DNA extraction

Two methods were employed for the bulk extraction of plasmid DNA from *E. coli* cells: either the Machery-Nagel NucleoBond Plasmid extraction kit was used or the standard mini-prep method described above was scaled up. The manufacturer's protocol was followed when the kit was used with the addition of an ethanol precipitation step. The scaled up mini-prep method involved the bulk isolation of plasmid from 50 ml of bacterial cultures. Cells were precipitated by centrifugation at 3 901 × *g* for 15 min at 4° C. Pellets were re-suspended in 600 µl of Solution I and aliquoted equally into three separate Eppendorf tubes. The cells were again precipitated at 12 470 × *g* for 1 min and re-suspended in 200 µl of Solution I with 5 µl of RNaseA. This was followed by the addition of 400 µl of Solution II, followed by 300 µl Solution III, with the same incubation times for each as used during the mini-prep. The mini-prep protocol was then followed until the wash with 70% ethanol with larger volumes of RNaseA (3 µl) and isopropanol

(800 µl) being used. Following the ethanol wash, pellets were dried, re-suspended in 100 µl of sdH_2O and the three tubes were then pooled before the ethanol precipitation.

2.5.3 Small scale genomic DNA extraction

The colony boil method was used for the small scale extraction of DNA from *M. smegmatis*. Briefly, half a colony (~10 mm diameter) was re-suspended in 10 µl of sdH_2O and boiled at 95˚ C for 5 min. Thereafter, 50 µl of chloroform was added to the suspension (a phenol:chloroform mixture is normally used, but it was observed that the phenol was contributing to the inhibition of PCR reactions and was therefore excluded), mixed vigorously and centrifuged at 12 470 × *g* for 5 min. The aqueous suspension above the cell debris interface was decanted into a fresh Eppendorf and was used as a DNA template in PCR reactions or for transformations.

2.5.4 Large scale genomic DNA extraction

The cetyltrimethylammonium bromide (CTAB) method was used for the bulk extraction of chromosomal DNA from *M. smegmatis.* Cells were grown to a lawn on 7H10 plates (with the appropriate antibiotics where necessary) from which four loopfuls of the culture were re-suspended in 500 µl of TE buffer. The cells were killed by heating the suspension at 65˚ C for 20 min after which lysozyme (10 mg/ml) was added and incubated at 37˚ C for an hour. Thereafter, 6 µl of proteinase K and 70 µl of 10% SDS were mixed into the suspension and the mixture was incubated at 65˚ C for 2 hours. This was followed by the addition of 100 µl of NaCl (5M) with mixing and 80 µl of pre-warmed CTAB/NaCl also with mixing. This suspension was then incubated at 65˚ C for 10 min. DNA was purified from this mixture by adding an equal volume of chloroform: isoamyl alcohol (24:1 v/v), mixing vigorously and spinning down at 12 470 × *g* for 5 min. The top aqueous layer was decanted into a fresh Eppendorf tube to which 600 µl of

isopropanol was added in order to precipitate the DNA by centrifugation at 12 470 × g for 20 min. The pellet was washed with 70% ethanol followed by an ethanol precipitation, drying and resuspension in sdH_2O.

2.6 DNA quantification

DNA was quantified on the Nanodrop ND- 100 Spectrophotometer, measured as a function of the absorbance of the sample at a wavelength of 260 nm. The Nanodrop also allowed for the purity of the sample to be measured by assessing the 260/280 ratio which represents RNA contamination as well as the 230/260 ratio which represents contamination with organic salts. Agarose gel electrophoresis was also used to estimate DNA concentrations based on the intensity of the DNA bands which could be compared to the intensity of the molecular weight marker bands (Roche and Fermentas) of known concentrations. The molecular weight markers used throughout this study were λIV, λV and λVI (Appendix C).

2.7 DNA manipulation methods

2.7.1 DNA amplification-Polymerase chain reaction (PCR)

Primers were designed using the online program Primer3 (http://frodo.wi.mit.edu/) which suggests the most appropriate primer sequences from the input region based on the selection criteria stipulated in Table 2.1.

Table 2.1: Criteria used for the selection of oligonucleotide sequences on Primer3

	Size	**Tm**	**% GC**
Minimum	18	55	55
Maximum	25	63	65
Optimum	23	60	62

Routine PCR reactions for screening were performed with non-proof reading DNA polymerase enzymes, either FastStart Taq (Roche) or Maxima HotStartTaq (Fermentas) following the manufacturers' instructions. PCR reactions were set up in 25µl with the following components common to both polymerase enzymes: 1× recommended buffer, dNTPs to a final concentration of 0.2 mM each, forward and reverse primers to a final concentration of 1µM each, DNA template between 10-100 ng/µl and 2U enzyme. FastStart Taq required the addition of 1× GC Rich in the reaction and $MgCl_2$ was added to a final concentration of 2 mM for reactions with Maxima HotStart Taq. Reactions were always made up to volume with sterile distilled nuclease free water. Cycling conditions were carried out as follows: one cycle of an initial denaturation at 94° C for 4 min; 30-35 cycles of 30 sec denaturation at 94° C, 30 sec annealing at 55-65° C and 30-90 sec elongation at 72° C which was followed by a final elongation step at 72° C for 5-7 min.

PCR products to be used for cloning were amplified with the high-fidelity, proof reading enzyme, Phusion polymerase (Finnzymes). The common components used for the non-proofreading enzymes remained the same for Phusion reactions with the addition of 3% dimethyl sulfoxide (DMSO) for GC-rich amplicons. Cycling conditions also remained the same except that the denaturation steps were carried out at 98° C and the annealing temperatures used were 5-10° C higher than the Tm calculated for primer sets, as recommended by the manufacturer.

2.7.2 Restriction digestion

Restriction enzymes used were purchased either from New England Biolabs (NEB) or Fermentas. Restriction digests were carried out as per the manufacturer's instructions with the

recommended buffers and when necessary the addition of bovine serum albumin (BSA) to a final concentration of 10 µg/ml. Double digests were either carried out simultaneously in a compatible buffer or sequentially with the inactivation of the first enzyme at 65˚ C followed by the addition of fresh buffer and the second enzyme when the buffers were incompatible. For plasmid screening approximately 0.5-1 µg of DNA was digested in a reaction volume of 10-20 µl and incubated at 37˚ C for 1 hour, unless otherwise recommended. Bulk digests were carried out for plasmid DNA and PCR products with 1-3 µg of DNA in reaction volumes of 15-30 µl and incubated at the recommended temperature for 1 hour. Approximately 2-5 µg genomic DNA was digested in reaction volumes of 20-50 µl and incubated at the recommended temperature overnight (no more than 16 hours) for Southern blot analysis.

2.7.3 Modification of DNA overhangs

Following restriction digests 3' and/or 5' overhangs were sometimes generated. When required, these fragments were blunted either by removing overhangs or filling in the gaps. T4 DNA Polymerase (NEB) catalyses the synthesis of DNA from primed single stranded DNA and possesses 3'→5' exonuclease activity. As per the manufacturer's instructions, T4 DNA polymerase was used for the blunting of fragments with both 5' and 3' overhangs. When only 5' overhangs were present in the digested DNA fragment, Klenow Fragment (NEB) was used as per the manufacturer's instructions. Reactions were carried out at 37° C in the presence of dNTPs for 10 min and were inactivated by heating at 65° C for 20 min.

2.7.4 Dephosphorylation of DNA

The removal of phosphate groups from the termini of linearized vector DNA fragments is required in order to prevent self-ligation and thus reduce the vector background during cloning.

Antarctic phosphatase (NEB) which catalyzes the removal of 5' phosphate groups from DNA was used for this reaction. The reaction volume varied according to the amount of DNA used but the supplied buffer and enzyme were always used at 1/10 the final reaction volume which was made up with sdH_2O. Reactions were incubated at 37° C for 1 hour and the enzyme was inactivated by heating at 65° C for 20 minutes.

2.7.5 DNA ligation

DNA fragments were ligated together using the T4 DNA ligase enzyme (Fermentas or Epicentre Biotechnologies Fast-Link DNA Ligation Kit), which catalyzes the formation of a phosphodiester bond between free 5'-phosphate and 3'-hydroxyl groups on the termini of DNA fragments. An optimum ratio of vector DNA to insert DNA needs to be used in order for the ligation reaction to be successful. A constant of 50 ng was always used for the vector DNA and the amount of insert DNA required for a 1:1 reaction was calculated using the equation:

$$Amount\ of\ insert\ DNA\ (ng) = \frac{50\ ng \text{ x size of insert (bp)}}{size\ of\ vector\ (bp)}$$

T4 DNA ligase requires ATP as a cofactor which was added to the reaction at a final concentration of 0.5 mM. Ligation reactions contained the appropriate volume of DNA, 1 µl of enzyme, 0.75 µl of ATP, 1 µl of the supplied buffer and were made up to 15 µl with sdH_2O. The reactions were incubated at room temperature for 20 minutes and heat inactivated for 10 minutes at 65° C prior to transformation and viewing on a gel. To assess the extent of ligation, observed as a decrease in the amount of individual fragments and an increase in the amount of circular DNA, 5 µl of the reaction was run on a 1% agarose gel.

2.8 Visualisation of DNA

DNA was viewed and analyzed by agarose gel electrophoresis, which allows for the separation of DNA fragments based on their size. Agarose gels (0.8-2 %) were prepared in TAE buffer with ethidium bromide added to a final concentration of 0.5 µg/ml. DNA was mixed with loading dye prior to being loaded onto the gel and separated in TAE buffer at 80-100 V. Molecular weight markers were always included on the gels to determine the size of the DNA fragments being separated. Gels were viewed and images captured under UV light using the Vacutec G:Box SYNGENE system and software (GeneSnap).

2.9 DNA fragment purification

The NucleoSpin Extract II Kit (Macherey Nagel) was used to purify DNA fragments. This kit allows for the purification of fragments excised from agarose gels as well as the purification of fragments directly from PCR reactions. The protocol provided by the manufacturer was followed. Gel fragments containing DNA were first melted at 45° C before loading onto the column, whereas PCR reactions with the desired product were loaded directly onto the column for binding of the DNA to the matrix. The column was washed and thereafter DNA was eluted in 35-50 µl of sdH_2O.

2.10 DNA sequencing

DNA sequencing was performed for all constructs generated in this study using PCR based cloning techniques to confirm that no mutations had been introduced into the gene/region of interest. Sequencing was outsourced to the DNA Sequencing Facility of Stellenbosch University and was performed using the Big Dye terminator v3.1 Cycle Sequencing kit and Bioline Half

Dye Mix. The EditSeq and SeqMan™ modules of the Lasergene suite of programs were used to analyze the sequencing data.

2.11 Construction of integrating vectors carrying *Mtb moaD*, *moaE* and *moaX* homologues

The development of gene transfer systems allowed for significant progress to be made in gaining a better understanding of *Mtb*. The commonly used systems include the use of plasmids carrying the origin of replication from the naturally occurring mycobacterial episomal plasmid pAL5000 or the chromosomal attachment site for the L5 mycobacteriophage in addition to a selectable antibiotic resistance gene (Garbe *et al*., 1994). Mycobacterial cells are able to maintain vectors integrated into the genome carrying a single copy of the gene of interest more stably when compared to episomal vectors (Pham *et al.*, 2007). pHINT is a mycobacterial integrating vector carrying the L5 attachment site which integrates at the $tRNA^{Gly}$ on the mycobacterial chromosome and the *hyg* resistance gene as shown in Figure 3.2. The integrating vector pTT1B carries the kan^R gene along with the integrase and attachment site from the mycobacteriophage Tweety (Pham *et al*., 2007). This vector integrates at the $tRNA^{Lys}$ gene on the mycobacterial chromosome shown in Figure 3.4 and contains a kan^R selectable marker gene, thereby making it compatible for simultaneous co-transformation with L5-based integrating vectors such as pHINT (Pham *et al.*, 2007). This system allows for the introduction of more than one gene into the genome, each driven off its own promoter and was ideal for the introduction of different combinations of the *Mtb moaD* and *moaE* homologues.

In a previous study carried out in the MMRU, multi-copy episomal vectors (pTBD1, pTBD2, pTBE1 and pTBE2), carrying each of the *Mtb moaD* and *moaE* homologues were constructed (Williams *et al*., 2011). It was shown that of the four vectors constructed, only pTBD2, pTBE1

and pTBE2 were able to complement the growth phenotype of the *M. smegmatis* single mutants, Δ*moaD2* and Δ*moaE2*, while pTBD1 was unable to complement Δ*moaD2* although the gene was being expressed (Williams *et al*., 2011). These validated vectors then formed the foundation of this study and were used during the construction of the integrating plasmids. Using the restriction enzymes *Bgl*II and *Pvu*I, the genes, together with their hsp60 promoters were excised from pTBD1, pTBD2, pTBE1 and pTBE2. 5' and 3' overhangs were filled in and removed respectively with T4 DNA Polymerase prior to the ligation reaction to allow for the blunt cloning of the fragments into their respective vectors. The integrating vectors, pTT1B and pHINT, were linearised with *Sca*I and phosphate groups were removed with Antarctic Phosphatase (NEB) to prevent vector re-ligation. *Mtb moaE1* and *moaE2* fragments were ligated to linearised pTT1B to generate the integrating vectors pTE1 and pTE2 respectively. The *moaD1* and *moaD2* genes were similarly incorporated into pHINT forming the vectors pHD1 and pHD2 respectively. An integrating vector carrying *moaX* was also constructed by removing the gene together with its *hsp60* promoter from pMoaX (Williams *et al*., 2011) with *Sac*II. The *moaX* gene in pMoaX was shown to be able to restore MoCo biosynthesis in a Δ*moaD2* Δ*moaE2* deletion mutant thus confirming that it encodes a novel fused MPT synthase (Williams *et al*., 2011). The 1 195 bp fragment was blunted with T4 DNA Polymerase and ligated to linearised pTT1B to generate the integrating vector pTX. Ligations were all transformed into competent *E. coli* DH5α cells and transformants were selected on LA Kan^{100} (pTT1B) or LA Hyg^{200} (pHINT) plates. Clones were screened and confirmed by restriction digest. Plasmids used and generated are listed in Table D1.

2.12 Generation of *M. smegmatis* strains carrying integrating complementation vectors

Once confirmed by restriction digest and sequencing, the integrating vectors were introduced into different electro-competent *M. smegmatis* strains by electroporation. The two single mutant

strains, Δ*moaD2* and Δ*moaE2*, as well as the double mutant, Δ*moaD2* Δ*moaE2*, were used, all of which are deficient for MoCo biosynthesis. Plasmids pHD1 and pHD2 were introduced into Δ*moaD2* individually and pTE1 and pTE2 were introduced into Δ*moaE2* individually. Combinations of the four integrating vectors (pHD1+pTE1, pHD1+pTE2, pHD2+pTE1 and pHD2+pTE2) were electroporated into Δ*moaD2* Δ*moaE2* simultaneously. The electroporations were spread on 7H10 with the appropriate antibiotics for 3-4 days at 37˚ C, after which time transformants were picked and confirmed by PCR using primers specific for the genes introduced. The strains generated are listed in Table 2.4.

2.13 Construction of episomal vectors carrying *Mtb moaD* and *moaE* homologues

In addition to the integrating vectors, episomal vectors carrying different combinations of the *Mtb moaD* and *moaE* homologues were constructed using a PCR cloning strategy. This strategy allowed for the introduction of *moaD1* and *moaD2* upstream of both *moaE1* and *moaE2* carried on episomal vectors and facilitated the operonic expression of two genes driven off a single *hsp60* promoter. The primers listed in Table 2.2 were used to amplify the *Mtb moaD1* and *moaD2* genes. The purified PCR products as well as the vectors pTBE1 and pTBE2 were digested with *Pst*I and *Hind*III to allow for directional cloning.

Table2.2: Primers used for the amplification of *Mtb* moaD1 and moaD2 with vector DNA as a template

Name	Sequence 5'-3'
mD1F	GGCGCTGCAGAATGATTAAAGTGAATGTTCTTTACTTC (*Pst*I)
mD1R	CGAAGCTTTCAGCCTCCGGCTACCTG (*Hind*III)
mD2F	GGCGCTGCAGAGTGACGCAGGTGTCCGA (*Pst*I)
mD2R	CGAAGCTTTTAGCCGCCGGCGAAAGG (*Hind*III)

[§]Restriction sites are underlined in each primer with the enzyme names shown in parenthesis

The digested fragments were ligated together in different combinations to generate the episomal vectors pMD1E1, pMD1E2, pMD2E1 and pMD2E2 which were transformed into competent *E.*

coli DH5α cells and selected on LA Hyg^{200} plates. Several colonies were picked for each ligation and screened by restriction digest. Clones were confirmed by restriction digest and sequencing of the vectors. Plasmid properties are listed in Table D1.

2.14 Generation of *M. smegmatis* strains carrying episomal complementation vectors

Confirmed vectors were introduced into the electro-competent *M. smegmatis* double mutant MoCo-deficient strain by electroporation. Transformants were selected on 7H10 with Hyg (200 μg/ml) at 37˚ C for 3-4 days and confirmed by PCR using primers specific for the genes introduced (Table D 3). The strains generated are listed in Table 2.4.

2.15 MoCo biosynthesis measurement: Heterologous complementation assay

M. smegmatis possesses several MoCo-dependent enzymes including two putative NR enzymes, the respiratory NarGHI and the assimilatory enzyme, NarB (Khan *et al.*, 2008). As a MoCo-dependent enzyme, NarB requires that the cofactor be available for its activity which would allow growth in media with nitrate as a sole nitrogen source. In the absence of MoCo, NarB is non-functional and is thus expected to render the organism incapable of growth in nitrate minimal media. Since growth can be restored by production of the cofactor, growth in nitrate minimal media thus serves as a surrogate for MoCo biosynthesis. Unlike *Mtb, M. smegmatis* does not possess a multiplicity of MoCo biosynthetic genes (Figure 1.3), is fast-growing and non-pathogenic and thus provides an ideal model in which to evaluate the mycobacterial MoCo biosynthetic pathway. In a previous study carried out in the MMRU, the *M. smegmatis* single mutants Δ*moaD2*, Δ*moaE2* and the double mutant Δ*moaD2* Δ*moaE2* were used to investigate the contribution of certain *Mtb* homologues to MoCo biosynthesis by heterologous complementation (Williams *et al.*, 2011).

2.15.1 Growth curve in nitrate minimal media

Growth curves were carried out in modified MPLN media (Table B1) which is a nitrate minimal medium. In this media, bacteria can only grow if they assimilate nitrate through nitrate reductase activity. Pre-cultures were grown overnight at 37˚ C in 5 ml of 7H9 with the appropriate antibiotics. Cells from pre-cultures were pelleted by centrifugation at 2 360 × *g* for 10 min and re-suspended in 1 ml of modified MPLN. This was followed by two rounds of washing by centrifugation and resuspension in modified MPLN to eliminate nutrient carryover from the 7H9. The final cell suspension was then used to inoculate 10 ml of fresh modified MPLN to a final OD of 0.05 in 50 ml Erlenmeyer flasks. Cultures were then grown at 37˚ C with shaking at 115 rpm for 5 days with OD readings taken every day.

2.16 Construction of FLAG-tagged derivatives of *moaX*

FLAG is a hydrophilic, eight amino acid (Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lys) peptide which can be added to the N- or C- terminus of recombinant proteins as a tag for use in detection and/or purification (Hopp *et al*., 1988). The advantages of this peptide include its small size which decreases the chances of interfering with protein folding and activity, and the availability of commercial antibodies specific for this sequence (Hopp *et al*., 1988). A vector, pFLAGEM (kindly provided by Dr. Edith Machowski, MMRU), containing a modified 3× FLAG sequence (Asp-Tyr-Lys-Asp-His-Asp-Gly-Asp-Tyr-Lys-Asp-His-Asp-Ile-Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lys) was used to construct FLAG-tagged derivatives of MoaX. The pFLAGEM vector allowed *moaX* to be cloned either 5’ or 3’ of the 3×FLAG-encoding sequence to create a fusion protein with the FLAG-tag fused in-frame to the C- or N- terminus of the target protein. The primers listed in Table 2.3 were used to amplify *moaX*.

Table 2.3: Primers used to generate FLAG-tagged derivatives of moaX

Primer name	Sequence 5'-3' [§]	Position
moaX-F	gccgTGTACAGATGATTACTGTCAATGTGCTC (*BsrGI*)	1-21 of *moaX*
moaX-R	gccgCGTACGCCTGGCCGATTGGCCACCCACTC (*BsiWI*)	649-663 of *moaX*

[§]Restriction site sequence is underlined in each primer with the restriction enzyme names shown in parenthesis

The 679 bp PCR product was digested with *Bsr*GI and *Bsi*WI prior to ligation. The vector pFLAGEM was linearised with either *Bsr*GI or *Acc*651 which is an isocaudomer of *Bsi*WI to produce overhangs compatible for cloning. The digested PCR fragment was ligated with the *Bsr*GI linearized vector for the incorporation of the tag on the C-terminus of MoaX to generate pFLAGmoaXC, and with the *Acc*651 linearised vector to produce an N-terminally tagged protein on the vector pFLAGmoaXN. Ligation reactions were transformed into competent *E. coli* DH5α cells and selected on LA Hyg^{200} plates. Several colonies were picked for each ligation and screened by restriction digest. Positive clones were confirmed by restriction digest (Figure E 10 and Figure E 11) as well as sequencing.

Table 2.4: Strains assessed for MoCo biosynthesis using the heterologous complementation assay

Name	Description	Source/ reference
M. smegmatis		
mc^2155	*ept-1* (efficient plasmid transformation) mutant of mc^26	Snapper *et al.*, 1990
Δ*moaD2*	Derivative of mc^2155 carrying an unmarked deletion in *M. smegmatis moaD2*	Williams *et al.*, 2011
Δ*moaD2* (pTBD1)	Derivative of *M. smegmatis* Δ*moaD2* carrying an episomal plasmid expressing *MtbmoaD1*from the *hsp60* promoter; Hyg^r	Williams *et al.*, 2011
Δ*moaD2* (pTBD2)	Derivative of *M. smegmatis* Δ*moaD2* carrying a plasmid expressing *Mtb moaD2* from the *hsp60* promoter; Hyg^r	Williams *et al.*, 2011
Δ*moaD2* (pMoaX)	Derivative of *M. smegmatis* Δ*moaD2* carrying a plasmid expressing *Mtb moaX* from the *hsp60* promoter; Hyg^r	Williams *et al.*, 2011
Δ*moaE2*	Derivative of mc^2155 carrying an unmarked deletion in the *M. smegmatis moaE2* gene	Williams *et al.*, 2011
Δ*moaE2*(pTBE1)	Derivative of *M. smegmatis* Δ*moaE2* carrying an episomal plasmid expressing *MtbmoaE1* from the *hsp60* promoter; Hyg^r	Williams *et al.*, 2011
Δ*moaE2* (pTBE2)	Derivative of *M. smegmatis* Δ*moaE2* carrying an episomal plasmid expressing *MtbmoaE2* from the *hsp60* promoter; Hyg^r	Williams *et al.*, 2011
Δ*moaE2* (pMoaX)	Derivative of *M. smegmatis* Δ*moaE2* carrying an episomal plasmid expressing *MtbmoaX* from the *hsp60* promoter; Hyg^r	Williams *et al.*, 2011
Δ*moaD2* Δ*moaE2*	Derivative of *M. smegmatis* Δ*moaE2* carrying an unmarked deletion in the *M. smegmatis moaD2*gene	Williams *et al.*, 2011
Δ*moaD2* Δ*moaE2* (pMoaX)	Derivative of *M. smegmatis* Δ*moaD2* Δ*moaE2* carrying an episomal plasmid expressing *Mtb moaX* from the *hsp60* promoter; Hyg^r	Williams *et al.*, 2011
Δ*moaD2* Δ*moaE2* (pTmoaX)	Derivative of *M. smegmatis* Δ*moaE2* with an integrating plasmid expressing *Mtb moaX* from the *hsp60* promoter; Kan^r	This work
Δ*moaD2*::pHD1	Derivative of *M. smegmatis* Δ*moaD2* with an integrating plasmid expressing *Mtb moaD1*from the *hsp60* promoter; Hyg^r	This work
Δ*moaD2*::pHD2	Derivative of *M. smegmatis* Δ*moaD2* with an integrating plasmid expressing *Mtb moaD2*from the *hsp60* promoter; Hyg^r	This work
Δ*moaE2*::pTE1	Derivative of *M. smegmatis* Δ*moaE2*with an integrating plasmid expressing	This work

	MtbmoaE1 from the *hsp60* promoter; Kan^r	
Δ*moaE2*::pTE2	Derivative of *M. smegmatis* Δ*moaE2*with an integrating plasmid expressing *Mtb moaE2* from the *hsp60* promoter; Kan^r	This work
Δ*moaD2* Δ*moaE2*::pIntD1E1	Derivative of *M. smegmatis* Δ*moaD2* Δ*moaE2* with integrating plasmids pHD1 and pTE1 expressing *Mtb moaD1* and *moaE1* respectively from the *hsp60* promoter; Hyg^r, Kan^r	This work
Δ*moaD2* Δ*moaE2*::pIntD1E2	Derivative of *M. smegmatis* Δ*moaD2* Δ*moaE2* with integrating plasmids pHD1 and pTE2 expressing *Mtb moaD1* and *moaE2* respectively from the *hsp60* promoter; Hyg^r, Kan^r	This work
Δ*moaD2* Δ*moaE2*::pIntD2E1	Derivative of *M. smegmatis* Δ*moaD2* Δ*moaE2* with integrating plasmids pHD2 and pTE1 expressing *Mtb moaD2* and *moaE1* respectively from the *hsp60* promoter; Hyg^r, Kan^r	This work
Δ*moaD2* Δ*moaE2*::pIntD2E2	Derivative of *M. smegmatis* Δ*moaD2* Δ*moaE2* with integrating plasmids pHD2 and pTE2 expressing *Mtb moaD2* and *moaE2* respectively from the *hsp60* promoter; Hyg^r, Kan^r	This work
Δ*moaD2* Δ*moaE2*::pMD1E1	Derivative of *M. smegmatis* Δ*moaD2* Δ*moaE2* with episomal plasmid pMhsp60D1E1 expressing *Mtb moaD1* and *moaE1* from a single upstream *hsp60* promoter; Hyg^r	This work
Δ*moaD2* Δ*moaE2*::pMD1E2	Derivative of *M. smegmatis* Δ*moaD2* Δ*moaE2* with episomal plasmid pMhsp60D1E2 expressing *Mtb moaD1* and *moaE2* from a single upstream *hsp60* promoter; Hyg^r	This work
Δ*moaD2* Δ*moaE2*::pMD2E1	Derivative of *M. smegmatis* Δ*moaD2* Δ*moaE2* with episomal plasmid pMhsp60D2E1 expressing *Mtb moaD2* and *moaE1* from a single upstream *hsp60* promoter; Hyg^r	This work
Δ*moaD2* Δ*moaE2*::pMD2E2	Derivative of *M. smegmatis* Δ*moaD2* Δ*moaE2* with episomal plasmid pMhsp60D2E2 expressing *Mtb moaD2* and *moaE2* from a single upstream *hsp60* promoter; Hyg^r	This work

2.17 Generation of *M. smegmatis* strains carrying FLAG-tagged MoaX

The above-mentioned pFLAGEM vector, carrying *moaX* was used in these studies. In this vector, in the absence of any repression, *moaX* is under the control of the constitutive promoter-operator fusion, P_{myc1}*tetO* described by Ehrt *et al.*, (2005). The *tetR* gene encoding the transcriptional repressor of P_{myc1}*tetO* is carried on an integrating vector, pMC1s, under the control of the strong promoter, P_{smyc}. The tet repressor (TetR) protein is constitutively expressed from P_{smyc} and binds tightly to the Tet operator thus inhibiting transcription. When available in the cell, tetracycline binds the TetR repressor causing a conformational change and dissociation of the repressor from *tetO* thus inducing expression of the gene/s under its control (Figure 2.1A).

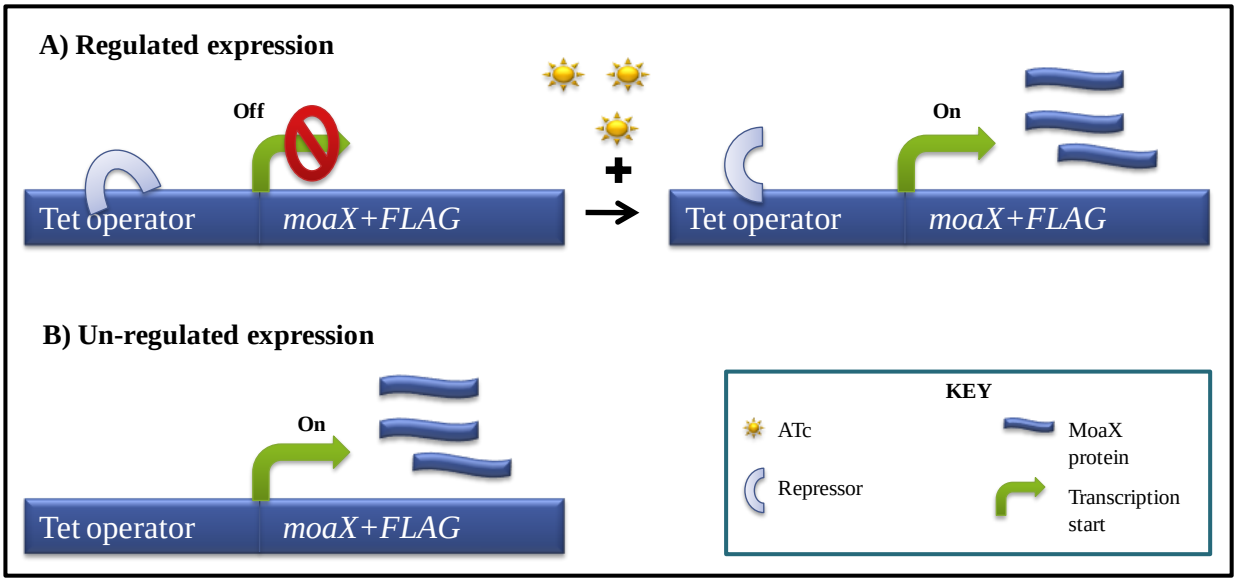


Figure 2.1: Schematic representation of the induction of *moaX* in the Tet system. (A) Regulated expression of *moaX* in the Tet system requires the addition of ATc which causes a conformational change in the repressor allowing transcription to proceed. (B) The unregulated expression of *moaX* takes place when no repressor is being expressed in the cell resulting in the constitutive expression of the gene of interest.

This system allows for the regulated expression of genes under the control of P_{myc1}*tetO*. For regulated expression of FLAG-tagged *moaX*, pFLAGEM carrying either the N- or C- terminally tagged MoaX was co-electroporated into Δ*moaD2* Δ*moaE2* with the pMC1s repressor plasmid. For unregulated expression of *moaX*, pFLAGmoaXN and pFLAGmoaXC were electroporated

into the double mutant individually, without any repressor. In the resulting strains, expression of *moaX* would be constitutive because there is no repressor binding to the operator (Figure 2.1B). These strains were assessed for MoCo biosynthesis using the heterologous complementation assay. In the complemented double mutant strains carrying the strong repressor on pMC1s, induction with anhydrotetracycline (ATc) was required to obtain expression of *moaX*. ATc was added to the media at a final concentration of 50 ng/ml prior to being inoculated with fresh washed culture to a final OD= 0.05.

2.18 MoaX mutagenesis

Two glycine residues in MoaX, corresponding to the terminal glycine residues of MoaD, were mutated by site-directed mutagenesis to evaluate the role they played in MoaX activity and cleavage. The Megaprimer method described by Smith and Klugman (1997) was used to introduce the point mutations. The technique, depicted in Figure 2.2, involves two PCR steps. The first step involves the synthesis of the megaprimer using the forward primer, moaXF and a reverse primer, mutatorR, in which the point mutation has been included. This is followed by a second round of PCR using the megaprimer as the forward primer and moaXR to amplify the full length *moaX* with the mutation incorporated.

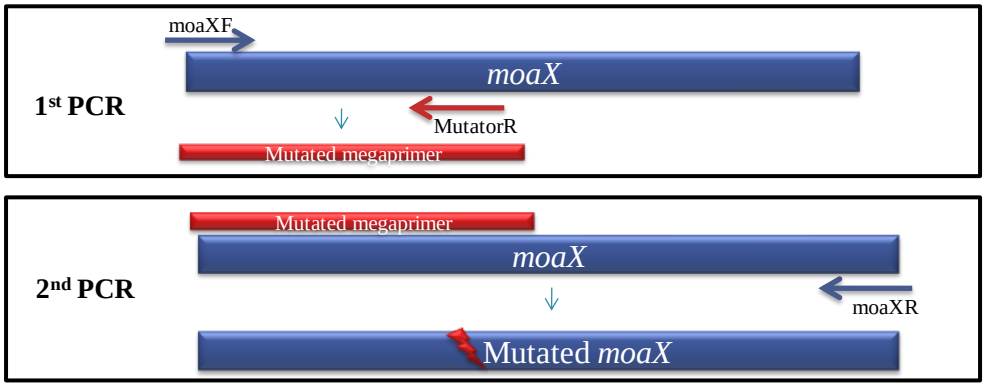


Figure 2.2: Diagram depicting the Megaprimer method of generating site-directed mutations in *moaX*. The first PCR reaction generates a mutated megaprimer using the wild type forward primer and a reverser primer carrying the point mutation. This megaprimer is then used in a second PCR reaction as a forward primer along with the wild type reverse primer generating the full length gene carrying the mutation.

The mutator reverse primers along with the wild type forward primer used to generate the megaprimers are shown in Table 2.5. The reverse primer, moaX-R shown in Table 2.3 was used with each megaprimer to generate full length *moaX* carrying a point mutation.

Table 2.5: Primers used to introduce point mutations in *moaX*

Primer name	Sequence 5'- 3'	Position
moaXga1R	GACATCGGAGCCC**G**CGGCAACCTGC[a]	231-255
moaXga2R	GACATCGGAG**G**CCCCGGCAACCTGC[b]	231-255
moaX-F	gccgTGTACAGATGATTACTGTCAATGTGCTC[c]	1-21

[a]The mutated residue is shown in bold and incorporates the point mutation 242→C into *moaX*. [b]The mutated residue is shown in bold and incorporates the point mutation 245G→C into *moaX*. [c]The restriction site is underlined with the enzyme name shown in parenthesis

Full length mutated *moaX* genes were cloned into pFLAGEM in the same manner as the wild type gene, described in section 2.16 to generate the strains listed in Table 2.6. The mutation 242G→C (G81A) introduced a *Sac*II restriction site into *moaX* and the mutation of 245G→C (G82A) introduced restriction site *Hae*III. These new restriction sites were used to distinguish between wild type and mutant copies of *moaX*. Sequencing was also used to confirm the introduction of the mutations. Strains carrying mutated copies of *moaX* on pFLAGEM were generated as in section 2.17 and assessed for MoCo biosynthesis.

2.19 MoaX protein analyses

2.19.1 Protein induction

Protein expression needed to be induced in strains carrying pMC1s. Cultures were grown to an OD between 0.4-0.8 at which time ATc was added to the culture at a final concentration of 50 ng/ml. The cultures were then grown for 3- 5 hours prior to being harvested for protein extraction.

2.19.2 *M. smegmatis* protein extractions

Cells were grown in 50 or 100 ml cultures and harvested by centrifugation at 2 360 × *g* for 10 min. Pellets were then re-suspended in B-PER (Fischer Scientific) cocktail solution (250 µl/ 50 ml culture) and either stored at - 80° C or lysed immediately. The B-PER solution is a lysis buffer used for the lysing of bacterial cells without the need for mechanical methods. However, due to the complex nature of the mycobacterial cell wall, incubation in the solution alone is not sufficient and further steps are required to obtain adequate yields of protein from the cell. Therefore, cells re-suspended in B-PER cocktail were transferred to Lysing Matrix B (IEPSA) tubes which contain 0.1 µm silica beads for the mechanical shearing of cells. Cells were lysed by ribolysing the tubes in the FastPrep Savant FP-120 Ribolyser for 20 sec at speed 6 with three repeats and 5 min incubations on ice between each run. After a final cooling on ice for 5 min, the tubes were spun down at 12 470 × *g* for 10 sec to pellet the cell debris and silica beads. The supernatant was then transferred to clean 1.5 ml Eppendorf tubes and spun down at 12 470 × *g* for 5 min to separate the soluble and insoluble protein fractions. The protein of interest in this study was to be found in the soluble fraction which was transferred to a clean 1.5 ml Eppendorf tube to be used immediately or stored at -20° C until required. For the extraction of protein from *E. coli* cells, 10- 20 ml cultures were grown in LB with the appropriate antibiotics and harvested by centrifugation at 2 360 × *g* for 10 min. Cell pellets were then re-suspended in 250-500 µl of B-PER cocktail and incubated at room temperature for 10 min. Cell debris was collected by centrifugation at 12 470 × *g* for 5 min and the supernatant was used for downstream processes.

Table 2.6: List of strains carrying FLAG-tagged derivatives of *Mtb* moaX

Strain	Description
Δ*moaD2* Δ*moaE2*::pMC1s	Derivative of Δ*moaD2* Δ*moaE2* with integrating vector pMC1s expressing the *tet* repressor, TetR; Kan^r
Δ*moaD2* Δ*moaE2*::pMC1s (pFLAGmoaXN)	Derivative of Δ*moaD2* Δ*moaE2*::pMC1s with episomal plasmid pFLAGmoaXN expressing *Mtb* N-terminally FLAG-tagged *moaX* under the control of the *tetO*; Kan^r, Hyg^r
Δ*moaD2* Δ*moaE2*::pMC1s (pFLAGmoaXC)	Derivative of Δ*moaD2* Δ*moaE2*::pMC1s with episomal plasmid pFLAGmoaXC expressing *Mtb* C-terminally FLAG-tagged *moaX* under the control of the *tetO*; Kan^r, Hyg^r
Δ*moaD2*Δ*moaE2* (pFLAGmoaXN)	Derivative of Δ*moaD2* Δ*moaE2* with episomal plasmid pFLAGmoaXN expressing *Mtb* N-terminally FLAG-tagged *moaX* under the control of the *tetO*; Hyg^r
Δ*moaD2*Δ*moaE2* (pFLAGmoaXC)	Derivative of Δ*moaD2* Δ*moaE2* with episomal plasmid pFLAGmoaXC expressing *Mtb* C-terminally FLAG-tagged *moaX* under the control of the *tetO*; Hyg^r
Δ*moaD2*Δ*moaE2* (pFLAGga1C)	Derivative of Δ*moaD2* Δ*moaE2* with episomal plasmid pFLAGga1C expressing mutated *Mtb* C-terminally FLAG-tagged *moaX* under the control of the *tetO*; Hyg^r
Δ*moaD2*Δ*moaE2* (pFLAGga2C)	Derivative of Δ*moaD2* Δ*moaE2* with episomal plasmid pFLAGga2C expressing mutated *Mtb* C-terminally FLAG-tagged *moaX* under the control of the *tetO*; Hyg^r

2.19.3 Protein quantification

Prior to being loaded onto SDS gels, protein extracts were quantified to ensure that equal amounts of the different sample extracts were used. For this, the Bradford assay was performed as previously described (Sambrook *et al.,* 1989). BSA was used as the protein standard for the assay. Dilutions of 10, 5, 2.5 and 1.25 µg/ml BSA were made up in 800 µl of sterile distilled protease-free H_2O in duplicate. The dilutions were incubated with 200 µl of Bradford reagent for 5 min and the absorbance of each sample was measured at 595 nm. Values were averaged and a standard curve was then plotted of absorbance (OD_{595}) vs. concentration (µg/ml). Protein extracts were diluted 100× prior to measuring the absorbance values which were used to determine the extract concentrations by extrapolating from the standard curve.

2.19.4 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE is a protein-denaturing electrophoresis technique which allows for the separation of proteins based on size alone. Acrylamide gels were prepared as outlined in Table B 11. Equal amounts of the different protein samples, in loading buffer, were denatured at 95° C for 5 min prior to being loaded onto the gel. Proteins were separated on the gel in running buffer (Table B 10) for ~ 2 hr or until the ladder was sufficiently separated. To view the proteins, gels were stained in Coomasie blue overnight with shaking at room temperature and subsequently de-stained until discrete protein bands were visible.

2.19.5 Western-blotting

2.19.5.1 Electro-blotting

Protein samples were first resolved by SDS-PAGE as described above and then transferred to a membrane in Tris-Glycine transfer buffer (Table B 13) for 1 hr at 4˚ C (Cleaver Omniblot Mini

Transfer System, 100V, 300 mA). Preliminary Western blotting was carried out using polyvinylidene fluoride (PVDF) membranes, which were readily available in the laboratory at the time; however, the transfer efficiency and resolution was variable which resulted in poor, inconsistent images, and the switch was subsequently made to BioTrace™ NT Pure Nitrocellulose Blotting membranes (Pall Life Sciences).

2.19.5.2 Immunological detection

The primary antibody used for all Western blots in this study was ANTI-FLAG M2® Monoclonal Antibody, mouse-purified IgG (Sigma) at a final concentration of 10 µg/ ml and the secondary antibody used was Rabbit Anti-Mouse IgG , Peroxidase Conjugate (Sigma) at a dilution of 1:25 000-40 000. Following transfer, the membrane was incubated for 30 min at room temperature or overnight at 4˚ C in blocking solution (Table B 13) to prevent non-specific binding of the primary antibody to the membrane. Primary antibody was added to the blocking solution and the membrane was incubated for an hour at room temperature or overnight at 4˚ C with gentle agitation. This was followed by three wash steps with TBST of 5 min each with shaking at room temperature. The membrane was then incubated with the secondary antibody in blocking solution for 1 hr at room temperature with gentle agitation. This was followed by 5 wash steps of 5 min each with TBST at room temperature with shaking. The membrane was then transferred to a hybridization bag to which the Chemiluminescent Peroxidase Substrate (CPS) Reagent (Sigma) was added and incubated for 5-10 min at room temperature. Excess substrate was squeezed out and the membrane was exposed to X-ray film (Kodak Biomax Light-Sigma or CL-Xposure™ Film-Thermo Scientific) for the time required to get the desired intensity (between 30 sec and 2 min). X-ray films were then passed through the Axim automated developer for bands to be observed.

2.20 Generation of *M. smegmatis* knock-out mutants

Two-step allelic exchange by homologous recombination, described by Gordhan and Parish (2001), was used to generate *M. smegmatis* knock-out mutants. A schematic representation of this technique is depicted in Figure 2.3, using *narB* as an example. The method involves the construction of a vector unable to replicate in mycobacteria (as it lacks a mycobacterial origin of replication (*oriM*)), and thus termed a suicide vector. The suicide vector carries homologous sequences of the upstream and downstream regions of the gene of interest fused together creating an inactive, truncated copy of the gene. Once introduced into the *M. smegmatis* cell, a single cross-over (SCO) event between the chromosome and one of the regions of homology results in the integration of the suicide vector into the chromosome. The suicide vector carries the selectable marker genes *aph, hyg* and *lacZ* which allow for the selection of SCO homologous recombinants, identified as blue colonies growing in the presence of X-gal, Kan and Hyg. Also carried on the suicide vector is the *sacB* gene which encodes levansucrase. When grown in the presence of sucrose, cells carrying the *sacB* gene produce levansucrase which converts sucrose to fructose polymers that accumulate in the cell envelope and become toxic, resulting in cell death. The *sacB* gene therefore serves as a counter-selectable marker facilitating a second cross-over event which would result in the expulsion of the vector backbone to either generate a knock-out mutant or reconstitute the wild type allele. Cells in which the second cross-over event occurred would be white when grown on X-gal and able to grow in the presence of sucrose. However, since these cells could either be mutants or wild type revertants, a PCR strategy was therefore used to screen for the two different genotypes. A schematic representation of the generation of an unmarked knock-out mutant is shown in Figure 2.3 using *narB* as the gene of interest; however the principle is the same for any gene.

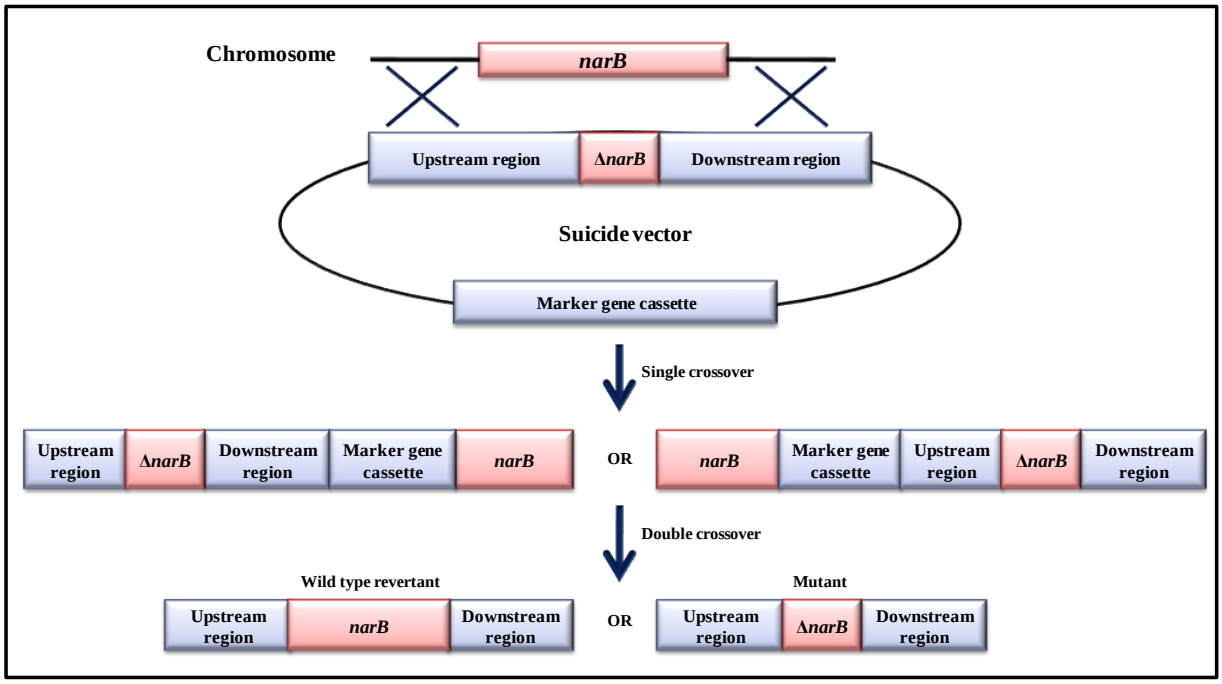


Figure 2.3: Schematic depiction of two-step allelic exchange mutagenesis using *narB* as the example gene.

2.20.1 Construction of *narB* and *narGHJI* suicide vectors

The upstream and downstream regions flanking the *narB* gene and *narGHJI* operon were amplified from genomic DNA using the primers sets listed in Table D4. The amplicons were digested with the appropriate restriction enzymes, *Bgl*II and *Hind*III for the upstream fragments and *Pst*I and *Bgl*II for the downstream fragments, prior to being run on a 1% agarose gel from which they were purified as described in 2.8. A three-way cloning strategy was employed using the *Hind*III and *Pst*I digested 4 753 bp p2NIL fragment to generate intermediate constructs carrying the upstream and downstream regions of the *narB* gene and the *narGHJI* operon, p2NILnarB and p2NILnarGHJI respectively (step 2 of Figure 2.4 and 2.5). These constructs were confirmed by restriction digests and sequencing to ensure that no mutations had been introduced into the homologous regions during PCR. Restriction digest of pGOAL19 with *Pac*I released the 7 939 bp selectable – counter-selectable marker cassette and was subsequently

ligated to the *Pac*I-linearised p2nil construct to generate the suicide vectors pΔnarB (Figure 2.4) and pΔnarGHJI (Figure 2.5). These vectors were validated by restriction digest prior to being used for electroporations.

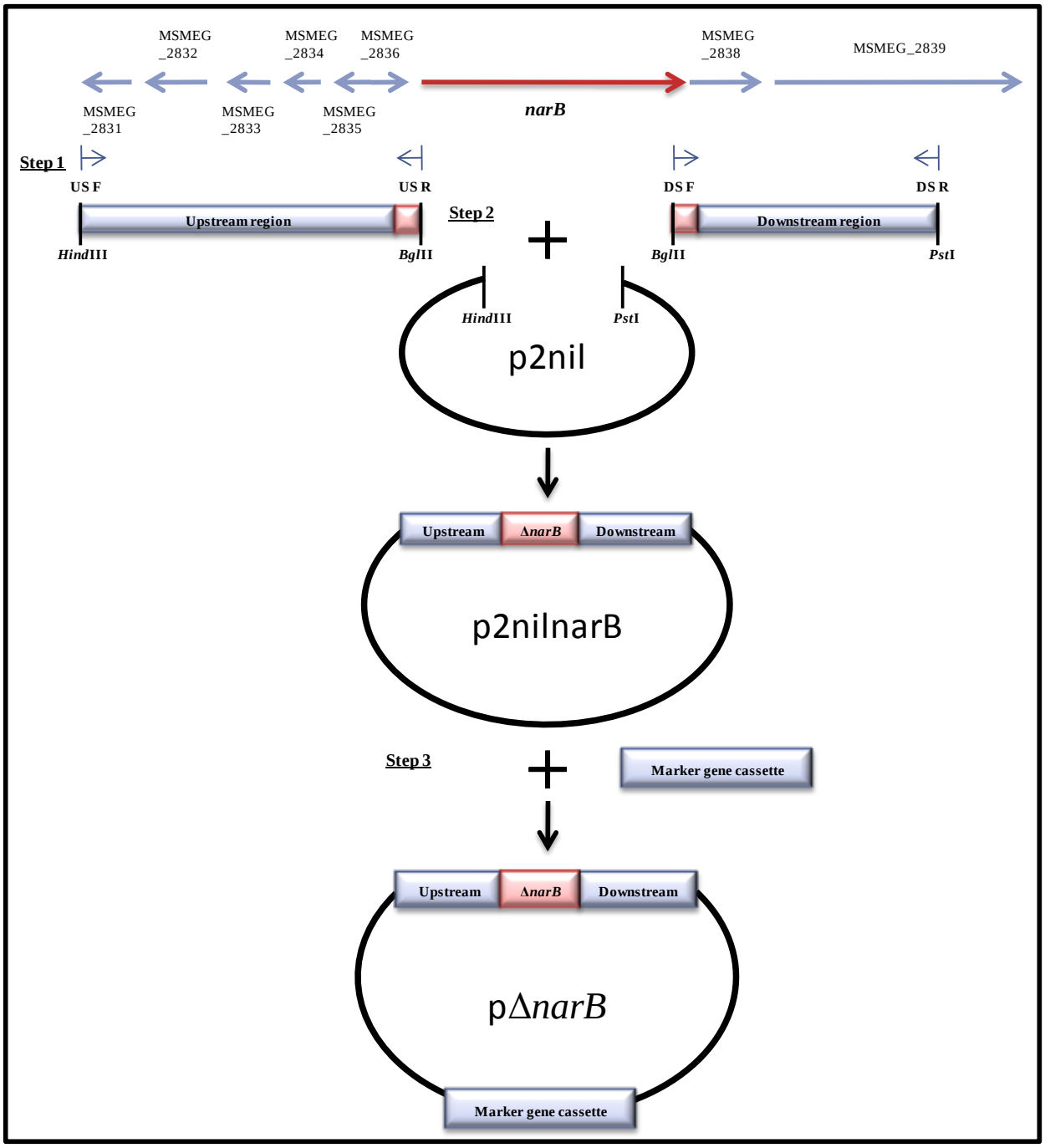


Figure 2.4: Schematic representation of the generation the suicide vector pΔ*narB*. Step 1: Amplification of upstream and downstream regions by PCR with a high fidelity *Taq* polymerase. Step 2: Restriction digestion and ligation of the upstream and downstream regions to p2nil to generate the intermediate vector p2nilnarB. Step 3: Ligation of p2nilnarB with the selectable-counterselectable marker cassette to generate the suicide vector pΔnarB.

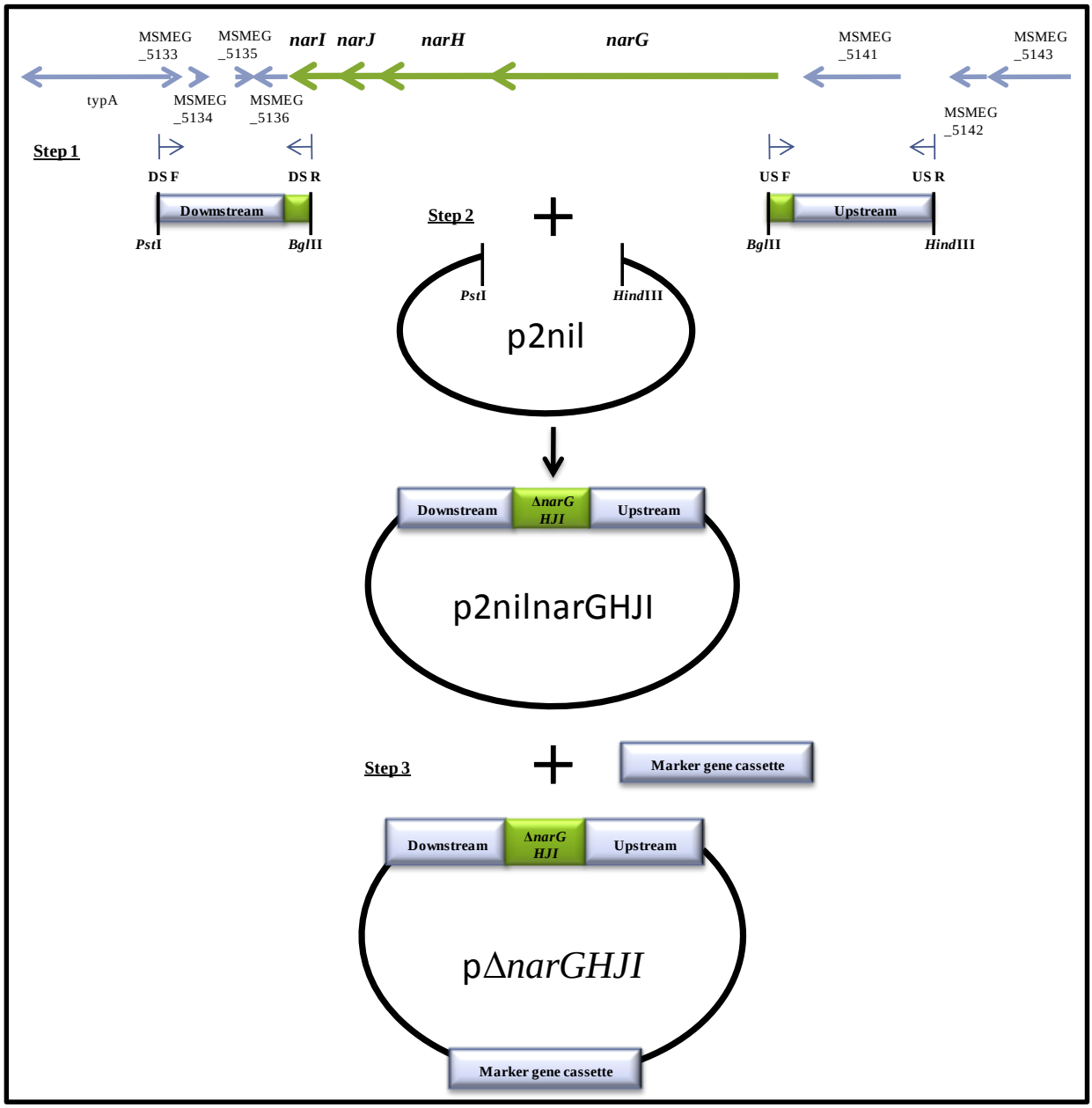


Figure 2.5: Schematic representation of the generation the suicide vector pΔ*narGHJI*. Step 1: Amplification of upstream and downstream regions by PCR with a high fidelity *Taq* polymerase. Step 2: Restriction digestion and ligation of the upstream and downstream regions to p2NIL to generate the intermediate vector p2nilnarGHJI. Step 3: Ligation of p2nilnarGHJI with the selectable-counter-selectable marker cassette to generate the suicide vector pΔnarGHJI.

2.20.2 Generation of Δ*narB* knock-out mutant

Once confirmed by restriction digestion, pΔnarB was electroporated into electro-competent wild type *M. smegmatis* as described in section 2.4.2. Following incubation for 5 days at 30˚ C on 7H10 plates with Kan, Hyg and X-gal, a blue colony was selected and grown overnight at 37˚ C

in 5 ml of 7H9 with Kan. An aliquot of the overnight culture was then used to inoculate fresh 7H9 broth without antibiotic and incubated as before to allow the second cross-over event to occur. Cells from the overnight culture were harvested by centrifugation at 3 901 × *g* and the pellet was re-suspended in 400 µl of 7H9. One hundred µl of the cell suspension was used to prepare a dilution series of 10^{-1}-10^{-7} in 1 ml of 7H9 from which 100 µl aliquots were withdrawn for plating on 7H10 plates with only X-gal, as well as 7H10 plates with X-gal and sucrose. Plates were incubated at 37˚ C for 5 days. White colonies were picked from the plates containing X-gal and sucrose, re-suspended in 20 µl of 7H9 and spotted onto 7H10 X-gal plates with and without Kan to ensure that the vector backbone had been lost. Only white colonies sensitive to Kan were picked to be screened by PCR.

2.20.3 Generation of Δ*narGHJI* and Δ*narB* Δ*narGHJI* knock-out mutants

Once confirmed, the suicide vector pΔnarGHJI was electroporated into competent wild type cells as well as into genotypically confirmed Δ*narB* cells to generate a single Δ*narGHJI* mutant and a double Δ*narB* Δ*narGHJI* mutant, respectively, as described in section 2.20.2.

2.21 Southern blot analysis

Southern blots were performed for the genotypic confirmation of mutants generated in this study.

2.21.1 Electro-blotting

Genomic DNA was digested as described in section 2.7.2. The digested DNA was separated on a 0.8% agarose gel for approximately 2 hr and an image of the gel was captured alongside a ruler. The gel was then incubated in depurination solution (0.25M HCl) for 15 min with mild shaking every 5 min. This was followed by two washes with dH_2O and incubation in denaturation

solution (0.5 M NaOH, 1.5 M NaCl) for 30 min with shaking. Equilibration of the gel was carried out in 1× TBE buffer briefly. The transfer cassette was prepared in 1× TBE with a nylon membrane on the agarose gel and sandwiched between two layers of Whatman filter paper and sponges. The transfer of DNA to the nylon membrane was carried out in a TE 22 Mini Transphor unit (Hoefer) at 0.5 A for 2 hr. Following the transfer the membrane with DNA was cross-linked at 2000 mJ/cm^2 and either used immediately or stored in maleic acid buffer until used.

2.21.2 Probe labeling

The probes used for Southern blots in this study were synthesized using the PCR DIG Probe Synthesis Kit (Roche). This kit allows for the specific labeling of probes by the incorporation of digoxygenin-labeled dUTP (DIG-dUTP), in place of dTTP, into the probe sequence during a PCR reaction. Incorporation of the labeled dNTP into the fragment was confirmed by running the PCR product on a gel alongside an unlabeled amplicon. DIG-dUTP has a higher molecular weight than dTTP and therefore products with the former would run slower than their unlabeled counterparts. Once synthesized, probes were either used immediately or stored at -20° C for no longer than 3 weeks.

2.21.3 Hybridization

Hybridization of probes to DNA-containing nylon membranes was performed using the DIG-High Prime DNA labelling and Detection Starter Kit II (Roche Biochemicals) as per manufacturers' instructions. Hybridization temperatures used were specific for individual probes and were calculated using the following equations where T_m is the melting temperature of the probe, L is the length of the probe sequence and T_{opt} is the optimum hybridization temperature:

$$Tm = 49.82 + 0.41(\%GC) - \frac{600}{L}$$

$$Topt = Tm - 20\ to\ 25^{\circ}\ C$$

Hybridization was carried out in roller bottles (Hybaid HB-OV-BM) in a hybridization oven (Hybaid Micro-4) at the specific temperature for the probes used. A pre-hybridization step was first carried out where the membranes were incubated in ~ 12 ml DIG Easy Hyb solution (Roche) for 20 min. Probes were denatured by boiling at 95° C for 10 min followed by rapid cooling on ice prior to being added to the pre-hybridization solution at 2 µl/ml. Hybridization was carried out overnight at the specific temperature for the probe used. Following hybridization, membranes were subjected to stringency washes to decrease the background on blots. Firstly, two washes were carried out with Solution I (2× SSC, 0.1% SDS) for 5 min each with shaking at room temperature. This was followed by two washes with Solution II (0.5× SSC, 0.1% SDS) for 15 min each at 68° C. Solution recipes are documented in Table B 12.

2.21.4 Immunological detection

Unless otherwise stated, all incubation and wash steps were carried out in a clean plastic container with gentle agitation at room temperature and the recipes for all the solutions used can be found in Table B12. The chemiluminescent detection of positively hybridized bands relies on the activity of the alkaline phosphatase labeled anti-digoxygenin antibody which acts on its substrate, in this case CSPD, to emit a luminescent signal that can be detected on X-ray film. An initial wash of the membrane in wash buffer for 5 min was carried out before being incubated in blocking solution for 30 min. This was followed by incubation in antibody solution for 30 min. Thereafter, two wash steps were carried out for 15 min each with wash buffer. Finally the membrane was equilibrated for 5 min in detection buffer before being placed in a hybridization

bag (with DNA side facing up) with 1 ml CSPD. The substrate was spread evenly on the membrane and care was taken to remove air bubbles. Excess substrate was squeezed out and the membrane incubated at 37° C for 10 min prior to being exposed to X-ray film (Kodak Biomax Light or CL-Xposure™ Film-Thermo Scientific) for the time required to get the desired intensity (between 10 min to overnight). The X-ray films were passed through the Axim automated developer for visualization of bands.

2.22 Phenotypic characterization of knock-out mutants

The knock-out mutants (Table 2.7) were assessed for nitrate assimilation ability by performing growth curves in modified MPLN as described in section 2.15.1.

Table 2.7: List of *M. smegmatis* knock-out mutant strains generated in this study

Strain	Description
Δ*narB*	Derivative of mc^2155 carrying an unmarked deletion in *M. smegmatis narB*
Δ*narGHJI*	Derivative of mc^2155 carrying an unmarked deletion in *M. smegmatis narGHJI* operon
Δ*narB* Δ*narGHJI*	Derivative of Δ*narB* carrying an unmarked deletion in *M. smegmatis narGHJI* operon

3 Results

3.1 Assessment of *moaD* and *moaE* gene function with single copy integrating vectors

The first aim of this study was to investigate whether the different MPT synthase-encoding *Mtb* homologues are able to combine to form chimeras of the enzyme with differing activities. The hypothesis was that isoforms of the MPT synthase, with differing activities, would allow for differential growth in MPLN and thereby provide a mechanism to identify and differentiate functionality of the *Mtb* homologues. To test this hypothesis, integrating vectors carrying each of the *Mtb* homologues (*moaD1*, *moaD2*, *moaE1*, or *moaE2*) were constructed and introduced into the Δ*moaD2* Δ*moaE2* double mutant. Reconstitution of the MPT synthase in these strains would then allow for growth on nitrate minimal media.

3.1.1 Strain generation and genotypic confirmation

Integrating vectors carrying *Mtb moaD* and *moaE* genes were constructed as described in section 2.11. This process yielded vectors pHD1, pHD2, pTE1 and pTE2. The genetic integrity of each of the vectors was confirmed by sequencing and extensive restriction analysis with at least four restriction enzymes and the results of these analyses are reported in Appendix E 1.

After analysis and confirmation, vectors pHD1, pHD2, pTE1 and pTE2 were introduced into Δ*moaD2* Δ*moaE2* mutant in four different combinations as described in section 2.12 to generate the strains Δ*moaD2* Δ*moaE2*:: pIntD1E1, Δ*moaD2* Δ*moaE2*:: pIntD1E2, Δ*moaD2* Δ*moaE2*:: pIntD2E1 and Δ*moaD2* Δ*moaE2*:: pIntD2E2. For easier reading, simpler names were assigned to the strains generated (Table 3.1).

Table 3.1: Simplified names assigned to strains carrying integrating vectors

Strain	Assigned name
Δ*moaD2* Δ*moaE2*:: pIntD1E1	DE::IntD1E1
Δ*moaD2* Δ*moaE2*:: pIntD1E2	DE::IntD1E2
Δ*moaD2* Δ*moaE2*:: pIntD2E1	DE::IntD2E1
Δ*moaD2* Δ*moaE2*:: pIntD2E2	DE::IntD2E2

Following an incubation period of four days on media with antibiotic selection, a single colony was picked for each combination and propagated in 7H9 with Hyg and Kan for subsequent use. To confirm that each of the selected transformants carried the desired combination of vectors, a PCR-based genotyping method was established to specifically amplify each of the *Mtb* genes. This was done by using PCR primers specific for each gene and template DNA that was

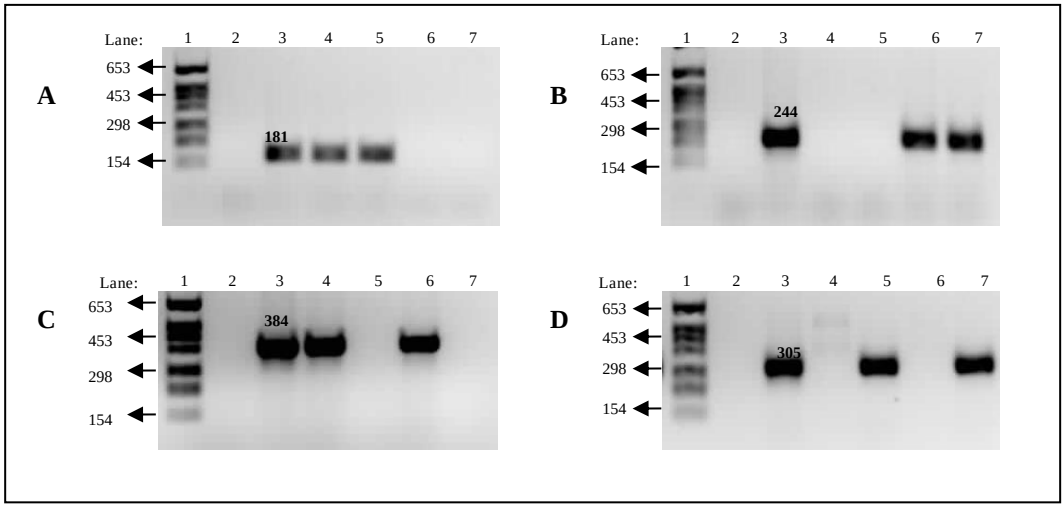


Figure 3.1: PCR confirmation of *M. smegmatis* double mutant strains complemented with different combinations of *Mtb moaD1, moaD2, moaE1* and *moaE2* genes carried on integrating vectors. (A) Amplification with primers TBD1F+TBD1R (B) Amplification with primers moaD2F+moaD2R (C) Amplification with primers TBE1F+TBE1R (D) Amplification with primers TBE2F+TBE2R. Lane 1: Marker λVI, Lane 2: No DNA control, Lane 3: Positive control, Lane 4: DE::IntD1E1, Lane 5: DE::IntD1E2, Lane 6: DE::IntD2E1, Lane 7: DE::IntD2E2

extracted from each of the above-mentioned transformants by colony boil. The primers used are listed in Table D3. The vector DNA used for electroporations was also used as the template for each positive control reaction. The gel images shown in Figure 3.1 confirm that each strain was carrying the correct combination of integrating vectors. In addition to confirming the presence of

the correct gene combination in these strains, the PCR reactions also show the specificity of each primer set for their respective homologue despite the fact that such high homology exists among these genes. After confirming the presence of the correct gene combination in these strains, the site of integration for each of the vectors was also confirmed.

A PCR strategy, currently used in the laboratory and shown in Figure 3.2, was used to confirm the site specific integration of pHINT.

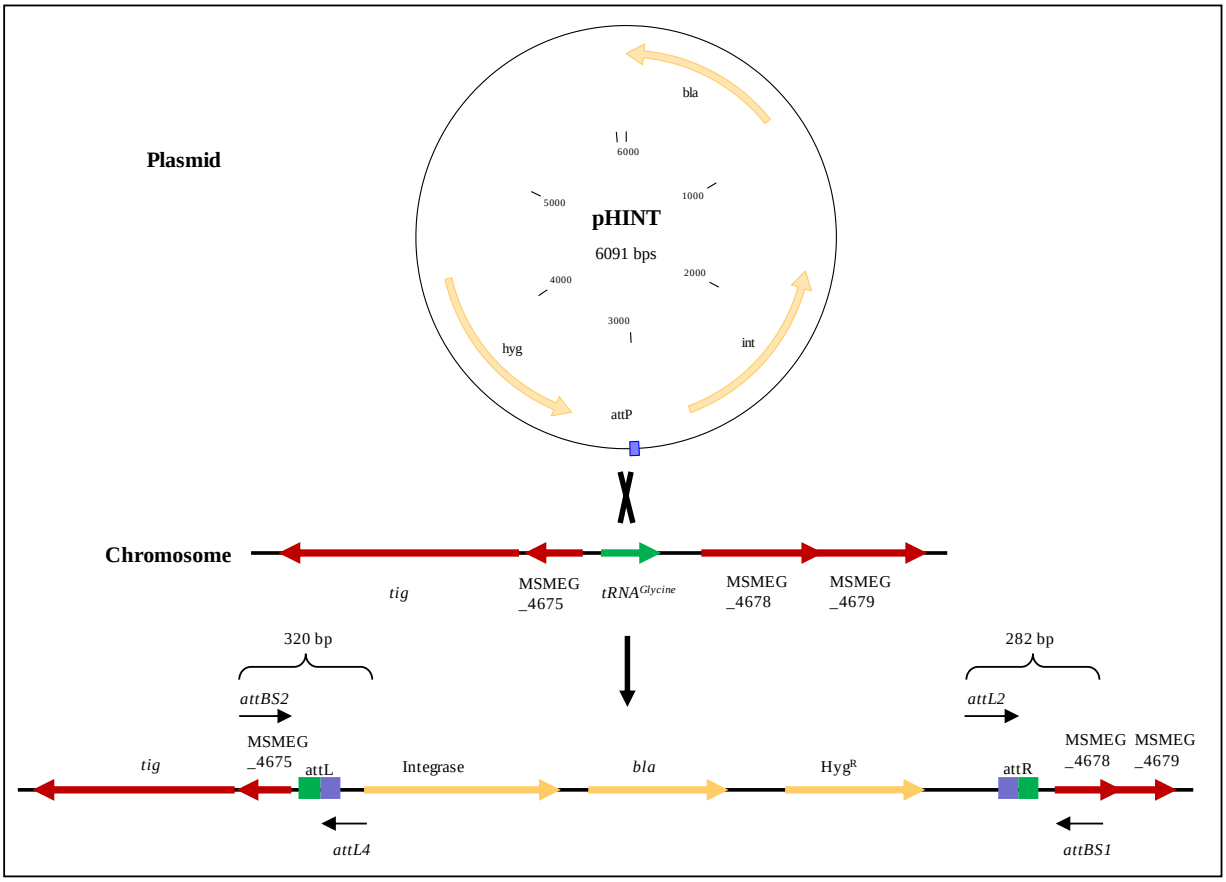


Figure 3.2: Schematic representation of integration of pHINT into the chromosome of *M. smegmatis*. pHINT has an L5 based integration system and the attachment site on the plasmid (blue *attP* block) integrates at the tRNA^{Gly}site on the *M. smegmatis* chromosome (green arrow). Integration of the plasmid results in the reconstitution of the tRNA^{Gly} site on either side of the integrated plasmid, *attL* and *attR*. The primers used to confirm the site specific integration of pHINT and expected amplicons are shown. *attBS2* and *attL4* produce a 320 bp amplicon confirming integration on the left. *attL2* and *attBS1* produce a 282 bp amplicon confirming integration on the right.

PCR analyses assessing the site specific integration of pHINT in the four strains listed in Table 3.1 are shown in Figure 3.3. Figure 3.3A confirms the site-specific integration of the pHINT-based vectors using the region upstream of the integration site as evidenced by the presence of the 320 bp amplicon while Figure 3.3B confirms integration using the downstream region with a 282 bp amplicon being visualized as expected.

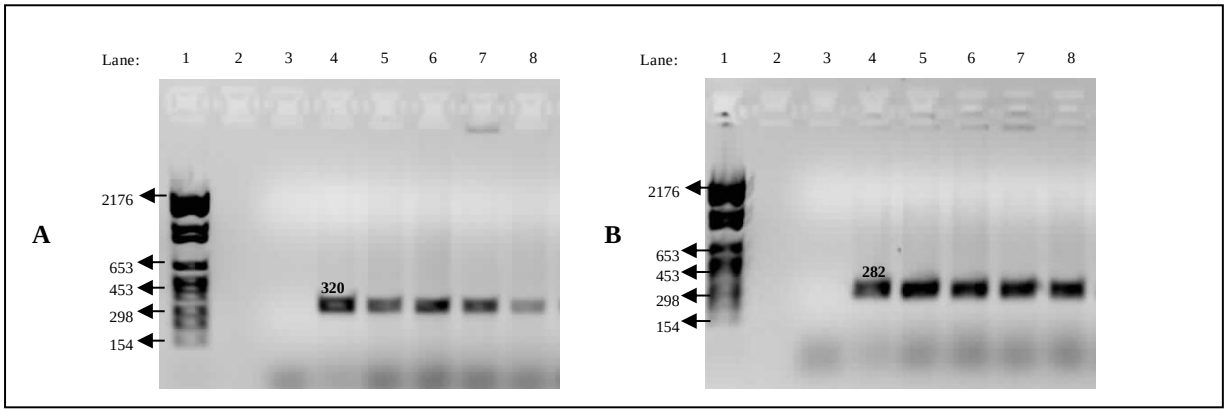


Figure 3.3: PCR confirmation of site-specific integration of pHINT carrying *Mtb moaD1* and *moaD2* into the *M. smegmatis* chromosome at the attB site, $tRNA^{Glycine}$. (A) Amplicons for primer set attBS2+attL4 (B) Amplicons for primer set attL2+attBS1. Lane 1: Marker λVI, Lane 2: Empty, Lane 3: No DNA control, Lane 4: Positive control, Lane 5: DE::IntD1E1, Lane 6: DE::IntD1E2, Lane 7: DE::IntD2E1, Lane 8: DE::IntD2E2

A similar system to assess integration of the pTT1B vector did not exist in the laboratory. Numerous attempts were made to design and test a PCR-based screening system for pTT1B integration but none of these was successful (data not shown). Consequently, a Southern blot analysis strategy, shown in Figure 3.4, was used to confirm the site of integration of pTT1B with probes specific to the vector (pink and orange in Figure 3.4). To confirm site-specific integration, the restriction enzymes used were chosen in a way that allowed one of the sites to be located in the chromosome and the other within the vector. This strategy allowed for the confirmation of each pTT1b-based complementation vector to be confirmed (see in Figure 3.7B for an example of this confirmation).

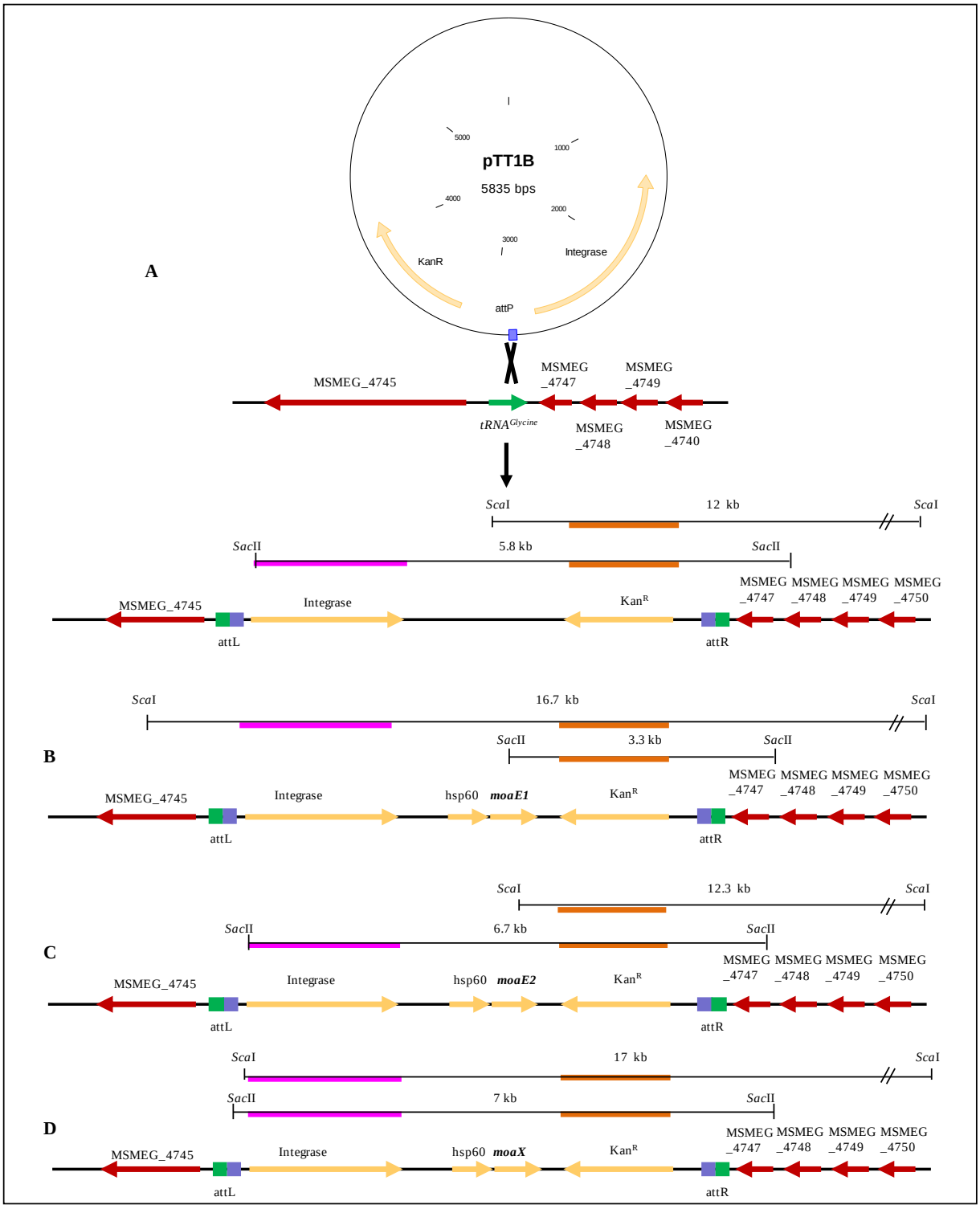


Figure 3.4: Schematic representation of the integration of pTT1b into the chromosome of *M. smegmatis*. The restriction enzymes used for southern blot confirmation are shown as well as the fragment sizes expected with each probe and construct. (A) Genomic map of the integration of pTTT1b into the *M. smegmatis* chromosome and the expected fragment sizes for southern blot analysis. (B) Genomic map of the integration of pTE1 into the *M. smegmatis* chromosome and the expected fragment sizes for southern blot analysis. (C) Genomic map of the integration of pTE2 into the chromosome of *M. smegmatis* and the expected fragment sizes for Southern blot analysis. (D) Genomic map of the integration of pTX into the *M. smegmatis* chromosome and the expected fragment sizes for southern blot analysis. The right probe used for Southern blotting is shown in orange and the left probe in pink.

3.1.2 MoCo biosynthesis in Δ*moaD2* Δ*moaE2* strains complemented with integrating vectors

In prior work, Williams *et al*. (2011) demonstrated that *Mtb moaD2*, *moaE1* and *moaE2* but not *moaD1* were able to restore MoCo biosynthesis in single deletion mutants of *M. smegmatis* lacking these genes. These findings suggest that the *M. smegmatis* MoaE2 is able to interact with the *Mtb* MoaD2 homologue but not the MoaD1, homologue, whereas the *M. smegmatis* MoaD2 is able to interact with both *Mtb* MoaE1 and MoaE2 to form a functional MPT synthase. The ability of the two components of the MPT-synthase, from different organisms, to interact to form a functional enzyme suggests that the multiple *Mtb moaD*- and *moaE*-encoded subunits could associate differentially with differing activities. As mentioned previously, MoCo biosynthesis was measured by monitoring bacterial growth, though nitrate assimilation by the MoCo-dependent NR activity. The strains carrying integrating vectors were therefore assessed for their ability to produce MoCo by assessing growth in MPLN media. It was previously shown that *Mtb moaX* is able to restore growth of the MPT-synthase deficient double mutant in MPLN media (Williams *et al.*, 2011) and therefore, the Δ*moaD2* Δ*moaE2* deletion mutant carrying a vector encoding MoaX, Δ*moaD2* Δ*moaE2* (pMoaX) referred to here as DE(pMX), was included as a positive control (Figure 3.5).

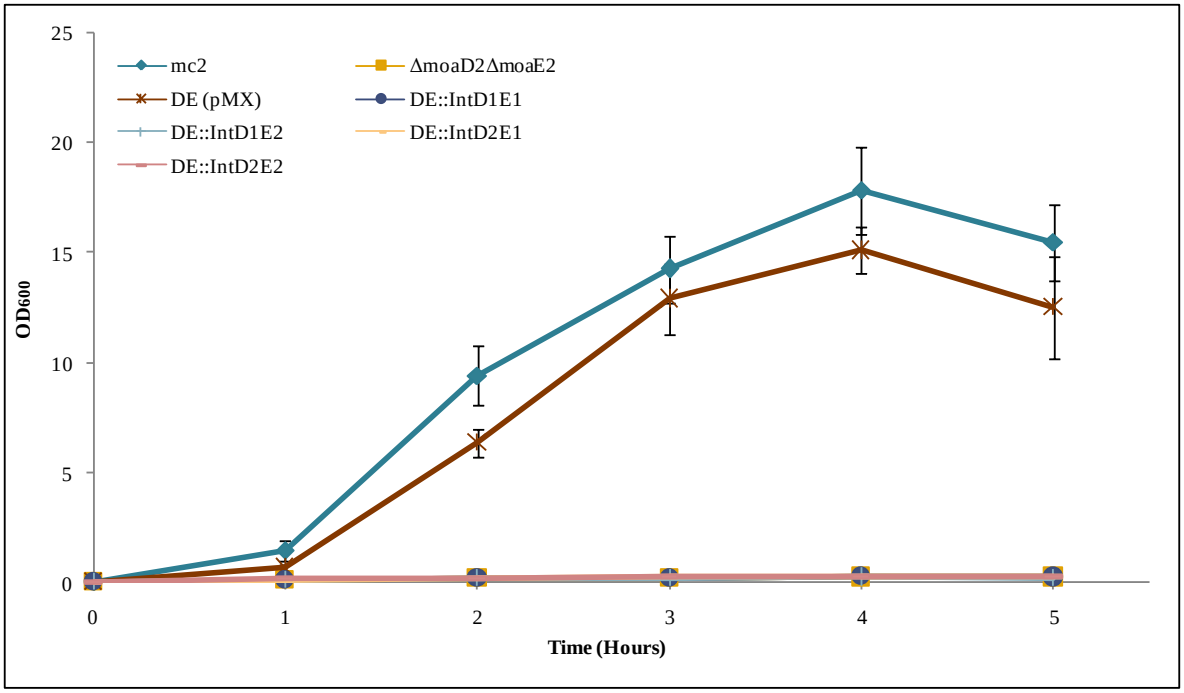


Figure 3.5: Growth curve of *M. smegmatis* Δ*moaD2*Δ*moaE2* complemented with different combinations of *Mtb moaD1, moaD2, moaE1* and *moaE2* carried on integrating vectors. Growth curves were carried out in MPLN with optical density readings taken daily for 5 days. The plotted data points are an average of at least three independent experiments and the standard error for each point is included.

Growth curves in MPLN media, Figure 3.5, show that none of the strains complemented with different combinations of the *Mtb* MPT-synthase encoding genes, on integrating vectors, was able to grow in MPLN media suggesting that a functional MPT synthase was not being generated by any of the combinations. Previous evidence had confirmed the functionality of *Mtb moaD2*, *moaE1* and *moaE2* by complementation of *M. smegmatis* mutants lacking *moaD2* or *moaE2* (Williams *et al.*, 2011). However, the major difference between the complementation strategy employed by Williams *et al*. (2011) and that described in this study is that the former used multi-copy episomal vectors to deliver the complementing gene, whereas integrating vectors that deliver only a single copy of the gene, were used in the present study. The discrepant findings suggested that lack of complementation in the present study could be due to reduced gene dosage. To test this hypothesis, the corresponding integrating vectors were then introduced into

the single *M. smegmatis* mutants, Δ*moaD2* and Δ*moaE2* and assessed for growth in MPLN media. Prior to carrying out growth experiments, each strain was assessed by PCR to confirm the presence of each gene as well as the site specific integration of the vector used. The PCR analysis shown in Figure 3.6 confirmed the genotypes of all strains.

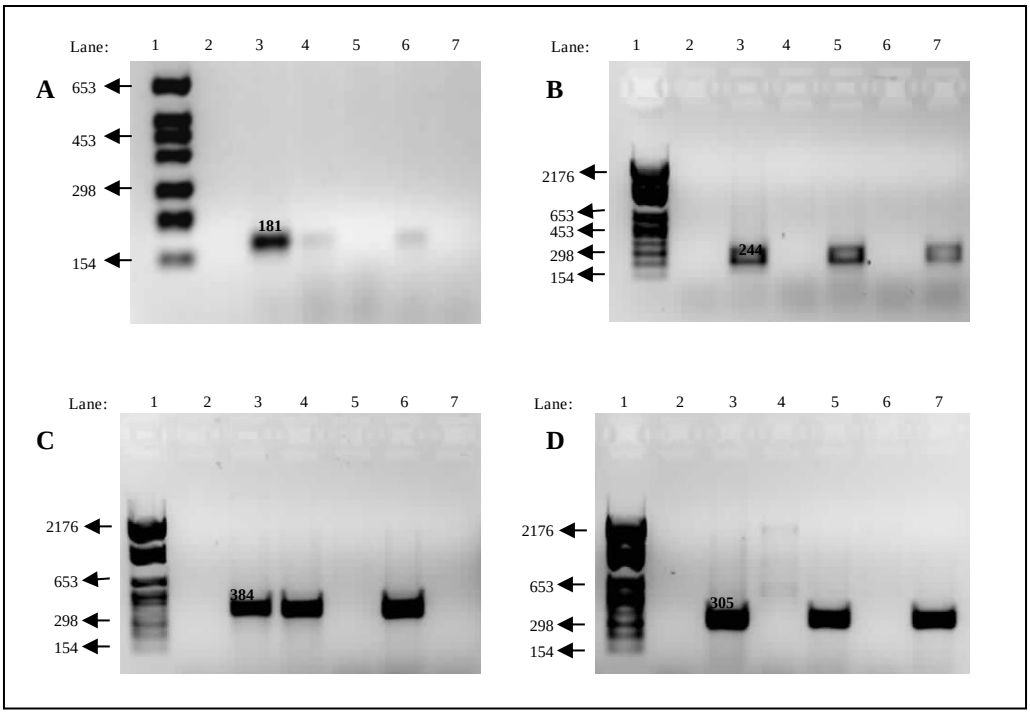


Figure 3.6: PCR confirmation of *M. smegmatis* single mutant strains complemented with *Mtb* genes on integrating or episomal vectors (A) Amplification with primers TBD1F+TBD1R. (B) Amplification with primers moaD2F+moaD2R. The gels in A and B were loaded in the same order- Lane 1: Marker λVI, Lane 2: No DNA control, Lane 3: Positive control, Lane 4: Δ*moaD2* (pMD1), Lane 5: Δ*moaD2* (pMD2), Lane 6: Δ*moaD2*::pHD1, Lane 7: Δ*moaD2*::pHD2 (C) Amplification with primers TBE1F+TBE1R (D) Amplification with primers TBE2F+TBE2R. The gels in C and D were loaded in the same order- Lane 1: Marker λVI, Lane 2: No DNA control, Lane 3: Positive control, Lane 4: Δ*moaE2* (pME1), Lane 5: Δ*moaE2* (pME2), Lane 6: Δ*moaE2*::pTE1, Lane 7: Δ*moaE2*::pTE2

The correct amplicons were observed for each strain thus confirming the presence of each complementing gene. The site specific integration of pHINT and pTT1B carrying their respective genes was also confirmed. Figure 3.7A confirms the correct integration of pHINT by PCR while Figure 3.7B confirms pTT1B integration. The amplicons observed in Lanes 3 and 4 of Figure 3.7A confirm the site-specific integration of pHD1 in the strain Δ*moaD2*::pHD1 and

the amplicons seen in Lanes 6 and 7 confirm the integration of pHD2 in Δ*moaD2*::pHD2. The site specific integration of pTE1 and pTE2 was confirmed by Southern blot analysis using the restriction enzyme *Sac*II and the right probe (orange in Figure 3.4) which allows for the two genes to be differentiated. As per Figure 3.7, the correct band sizes are observed in the blot (Figure 3.7B), confirming the integration of pTE1 and pTE2 into Δ*moaE2*::pTE1 and Δ*moaE2*::pTE2 respectively. After confirming that each strain was carrying the correct gene for complementation studies, the strains were subsequently assessed for MoCo biosynthesis by growth in MPLN media. In this experiment the growth of the strains complemented with single copy integrating vectors carrying the gene of interest was compared to the growth of the same strains complemented with the same gene, driven off the same promoter, only on a multi-copy episomal vector.

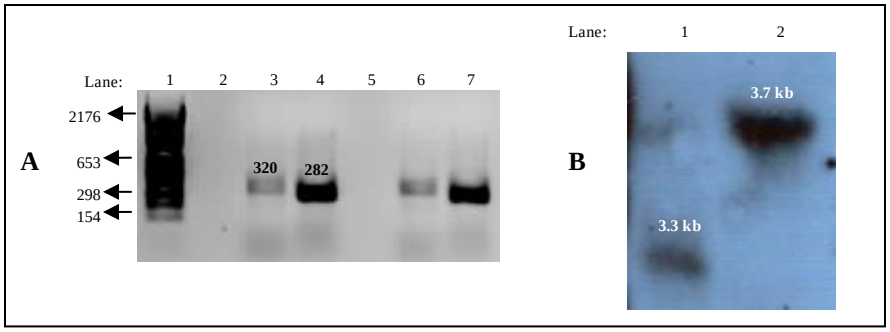


Figure 3.7: Confirmation of site specific integration of pHINT carrying *Mtb moaD1* or *moaD2* and pTT1b carrying *moaE1* or *moaE2* into the chromosome of the *M. smegmatis* single mutants. (A) PCR confirmation of pHINT integration. Lane 1: Marker λVI, Lane 2: No DNA control, Lane 3: Amplification of Δ*moaD2*::pHD1 with primers attBS2 and attL4, Lane 4: Amplification of Δ*moaD2*::pHD1 with primers attL2 and attBS1, Lane 5: Empty, Lane 6: Amplification of Δ*moaD2*::pHD2 with primers attBS2 and attL4, Lane 7: Amplification of Δ*moaD2*::pHD2 with primers attL2 and attBS1. (B) Southern blot confirmation of pTT1b integration using the right probe. Lane 1: *Sac*II digested genomic DNA from Δ*moaE2*::pTE1, Lane 2: *Sac*II digested genomic DNA from Δ*moaE2*::pTE2

Previous work demonstrated that *M. smegmatis* strains Δ*moaD2* (pMD2), Δ*moaE2* (pME1) and Δ*moaE2* (pME2) were able to grow in MPLN media thus confirming the functionality of these *Mtb* genes (Williams *et al*., 2011). A similar result, albeit at a lesser extent, was achieved in this

study (Figure 3.8); however, unusually severe clumping was observed for these strains which prevented a more quantitative measure of growth. Several attempts were made to resolve this problem: (i) the growth temperature was reduced from 37 ˚C to 30˚C but the phenomenon still persisted (data not shown); (ii) substitution of detergent Tween80 with Tyloxapol, an alternate detergent that is not metabolized by mycobacteria and can persist in cultures for longer, did not eliminate it (data not shown). However, the demonstration of some growth in these strains confirms the ability of *Mtb moaD2*, *moaE1* and *moaE2* to restore MoCo biosynthesis when expressed on an episomal plasmid. As observed by Williams *et al.* (2011), Δ*moaD2* (pMD1) was unable to grow in nitrate minimal media in these experiments, and as expected neither was Δ*moaD2*::pHD1. The remaining strains carrying a single copy of *moaD2, moaE1* and *moaE2* were also unable to grow in MPLN media. These data demonstrate that heterologous complementation with some of the MPT-synthase encoding genes can be achieved when these homologues are expressed from a multi-copy as opposed to integration vector. This suggests that there is a threshold of gene expression required to achieve heterologous complementation in this system and that the inability of the different combinations of *Mtb* MPT-synthase encoding genes to restore MoCo biosynthesis (Figure 3.5) may be due to reduced gene expression as opposed to lack of functionality of the reconstituted enzyme *per se*. To further test this, we assessed the ability of a single copy of the fused MPT synthase-encoding gene, *moaX*, to restore MoCo biosynthesis in an *M.smegmatis* MoCo deficient mutant.

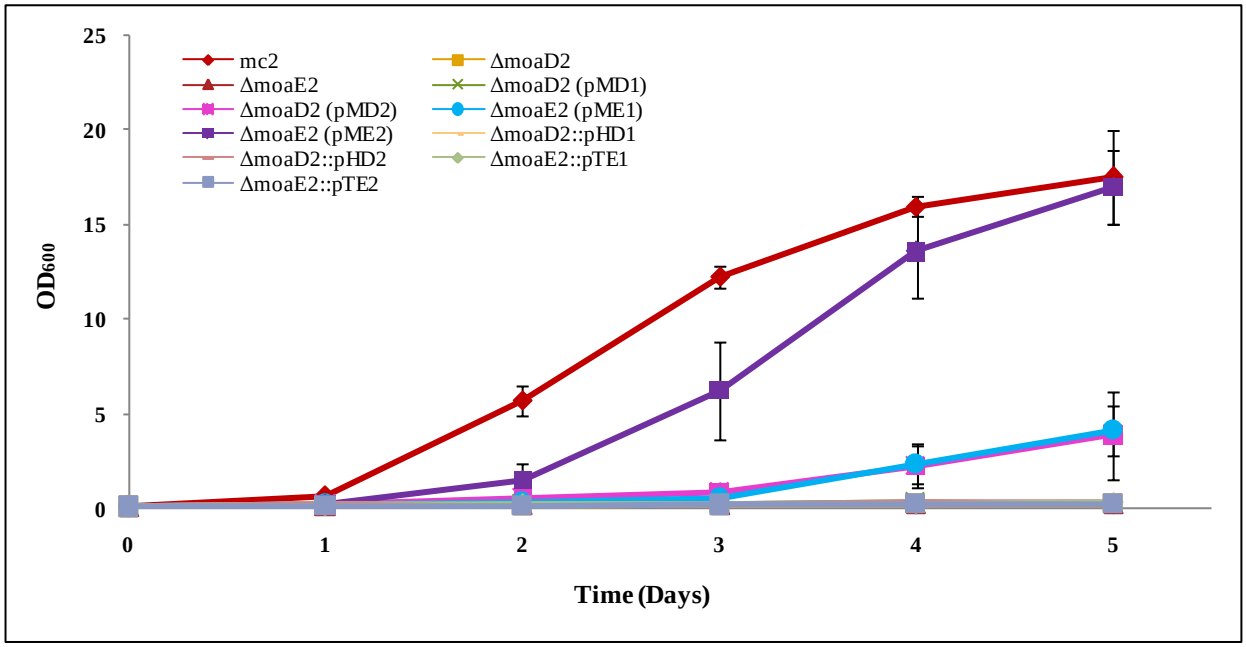


Figure 3.8: Growth curve of *M. smegmatis* single mutants, Δ*moaD2* and Δ*moaE2* complemented with either *Mtb moaD1, moaD2, moaE1* or *moaE2* carried on integrating and episomal vectors. Growth curves were carried out in MPLN media with optical density readings taken daily for 5 days. The plotted data points are an average of at least three independent experiments and the standard error for each point is included.

3.2 A single copy of *moaX* can restore MoCo biosynthesis in *M. smegmatis* Δ*moaD2* Δ*moaE2*

To evaluate the hypothesis that reduced gene dosage from integrating vectors resulted in the inability to restore growth in MPLN media, *moaX* was cloned into an integration vector. The pTT1B vector was chosen for this cloning due to the availability of more restriction enzyme sites. The strategy used to generate the integrating vector pTmoaX and the complemented strain Δ*moaD2* Δ*moaE2*::pTmoaX is described in section 2.11 and 2.12 respectively. The restriction mapping of pTmoaX is shown in Figure E 5, Appendix E 2. Once generated, the strain Δ*moaD2* Δ*moaE2*::pTmoaX, hence referred to as DE::pTX was confirmed by PCR analysis (Figure 3.9) using the primers XscreenF and Xscreen R (Table D 3).

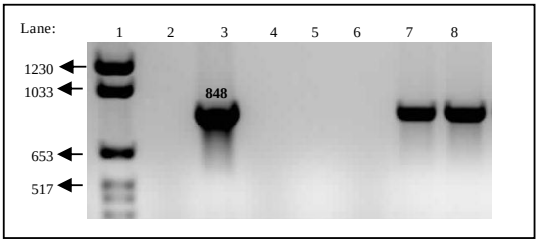


Figure 3.9: PCR confirmation of Δ*moaD2* Δ*moaE2*:: pTX Lane 1: Marker λVI, Lane 2: Empty, Lane 3: Positive control, Lane 4: Empty, Lane 5: No DNA control, Lane 6: Empty, Lane 7: Amplification of DE::pTX, Lane 8: Amplification of DE (pMX)

The presence of an 848 bp amplicon in Lane 7 of Figure 3.9 confirms the presence of *moaX* in DE::pTX and the site-specific integration of the vector was confirmed by Southern blot analysis (data not shown). The growth of DE::pTX in MPLN media was then compared to wild type and DE (pMX) (Figure 3.10).

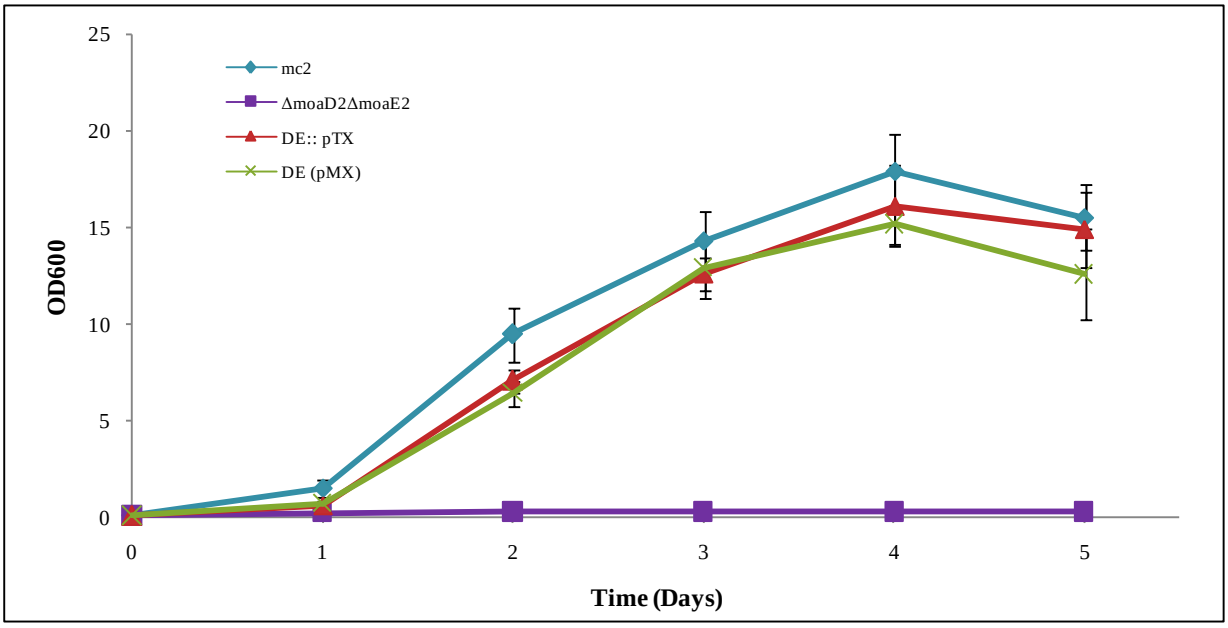


Figure 3.10: Growth curve comparing complementation with a single copy of the gene vs multiple copies. Growth curves were carried out in MPLN media with optical density readings taken daily for 5 days. The plotted data points are an average of at least three independent experiments and the standard error for each point is included.

As shown in Figure 3.10, expression of *moaX* in single copy was able to complement the growth phenotype of the double mutant in MPLN media to a level comparable to the control carrying the same gene on a multi-copy plasmid. These data suggest that the lack of complementation observed for the strains carrying different combinations of the *Mtb moaD* and *moaE* homologues on integrating vectors (Figure 3.5) may be due to reasons other than low expression. Williams *et al*. (2011), reported toxicity effects in *M. smegmatis* strains carrying *Mtb moaD* and *moaE* genes under the control of the *hsp60* promoter, as assessed by the ability of these strains to grow at 37ºC. The underlying mechanism that results in these observations is not clear and similar effects may have prevailed in this study. To further simplify our heterologous expression vectors, we incorporated two genes on a single vector driven off a single promoter as a synthetic operon.

3.3 Operonic expression of *Mtb moaD* and *moaE* genes from episomal vectors

Multi-copy episomal vectors carrying different combinations of the *Mtb moaD* and *moaE* homologues were constructed as described in section 2.13. To achieve this, the Mtb *moaD* homologues were cloned into vectors carrying the *Mtb moaE* genes. This strategy allowed for the introduction of a *moaD* homologue directly between the *hsp60* promoter and a *moaE* homologue on the vector which facilitated the simultaneous expression of both genes from the promoter as an operon. The vectors generated were confirmed by sequencing as well as restriction mapping shown in Figures E 6 to E 9 in Appendix E 3.

3.3.1 *Mtb moaE1* is toxic when expressed in a synthetic operon

The episomal vectors carrying different combinations of the *Mtb* homologues were introduced into the *M. smegmatis* double mutant. However, electroporation results revealed that the two vectors carrying *moaE1* had very low transformation efficiencies, with only a single colony

recovered for each of pMD1E1 and pMD2E1 from the first electroporation. More electroporations were then carried out to investigate this observation further and the results in Table 3.2 show that those vectors were indeed toxic.

Table 3.2: Episomal vectors pMD1E1 and pMD2E1 are toxic to *M. smegmatis* cells.

	ΔDΔE::pMhsp60	ΔDΔE::pMD1E2	ΔDΔE::pMD1E1	ΔDΔE::pMD2E1
Cfu/µg DNA[§]	5.88E+04	7.85E+04	0.3*	0
(± SE)	(1.63E+04)	(2.04E+04)	(0.3)	

*A single colony was observed on one of the electroporation plates from one of three independent experiments. § The data presented are an average of three independent experiments

Included as controls for the electroporation experiments was the empty pMhsp60 vector as well as the pMD1E2 vector, which produced several colonies in the first electroporation. The vector backbone of all the vectors used was exactly the same (*i.e*. pMhsp60) and the transformation efficiency of pMD1E2 was high, suggesting that the toxicity was attributable to the presence of *moaE1*. However, the transformants that were obtained from the initial electroporation experiment were further tested for the presence of the correct gene combination and for heterologous complementation. Once again, for improved readability the resultant strains were assigned simpler names listed in Table 3.3 and confirmed by PCR analysis using the primers listed in Table D 3.

Table 3.3: Simplified names assigned to strains carrying episomal vectors

Strain	Assigned name
Δ*moaD2* Δ*moaE2* (pMD1E1)	DE (pMD1E1)
Δ*moaD2* Δ*moaE2* (pMD1E2)	DE (pMD1E2)
Δ*moaD2* Δ*moaE2* (pMD2E1)	DE (pMD2E1)
Δ*moaD2* Δ*moaE2* (pMD2E2)	DE (pMD2E2)

The amplicons observed for each PCR reaction in Figure 3.11 confirm the presence of the correct genes in each of the strains which were subsequently assessed for MoCo biosynthesis by growth in MPLN.

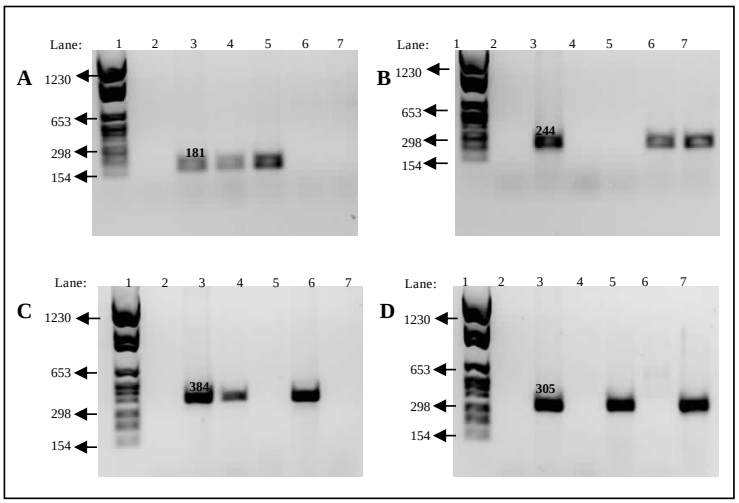


Figure 3.11: PCR confirmation of double mutant strains complemented with different combinations of *Mtb moaD1, moaD2, moaE1* and *moaE2* carried on episomal vectors. (A) Amplification with primers TBD1F+TBD1R (B) Amplification with primers moaD2F+moaD2R (C) Amplification with primers TBE1F+TBE1R (D) Amplification with primers TBE2F+TBE2R. Lane 1: Marker λVI, Lane 2: No DNA control, Lane 3: Positive control, Lane 4: DE (pMD1E1), Lane 5: DE (pMD1E2), Lane 6: DE (pMD2E1), Lane 7: DE (pMD2E2).

Included as controls in the growth experiment, were the wild type and double mutant strain as positive and negative control respectively, and DE (pMX) as a positive complementation control. The results obtained are shown in Figure 3.12. As expected DE (pMX) was able to grow as previously shown. Of the strains complemented with different combinations of the *Mtb* homologues only DE (pMD2E2) was able to grow with nitrate as the sole nitrogen source. No growth was observed for DE (pMD1E1), DE (pMD1E2) or DE (pMD2E1). This result suggests that only *Mtb* MoaD2 and MoaE2 are able to combine to form a functional MPT synthase.

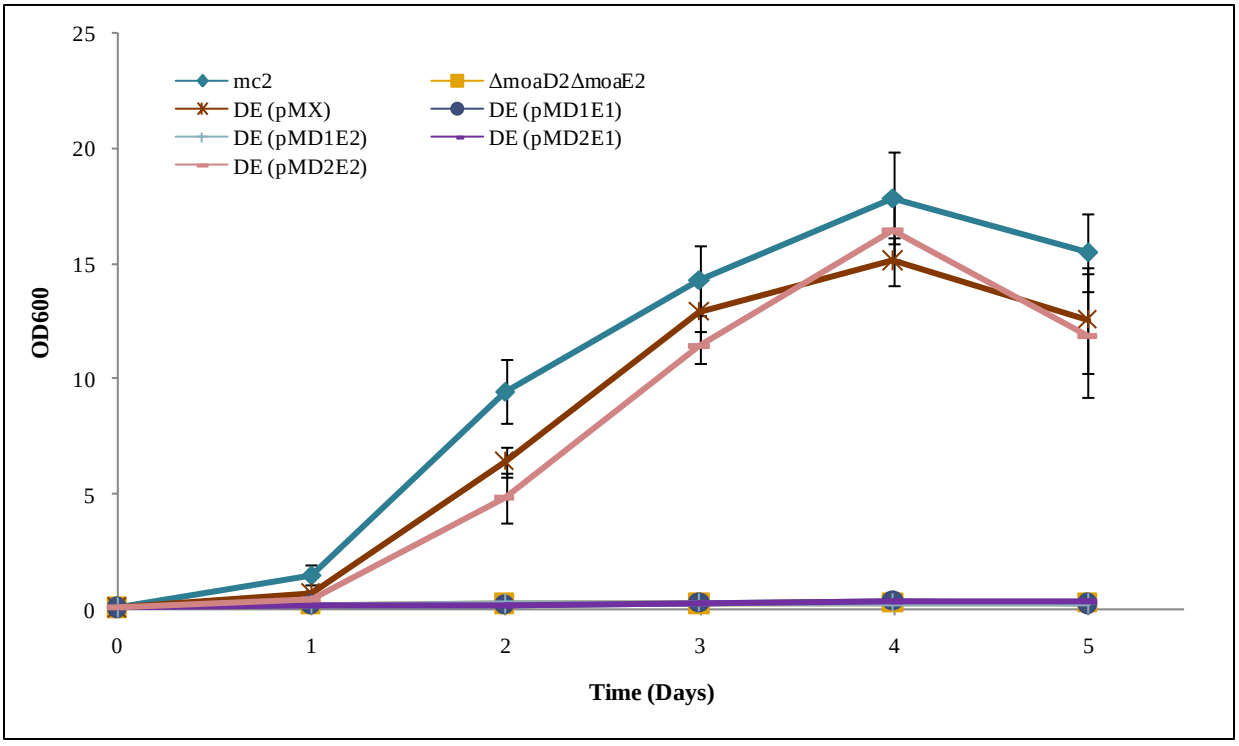


Figure 3.12: Growth curve of strains complemented with episomal vectors carrying different combinations of *Mtb moaD1, moaD2, moaE1*and *moaE2* genes. Cultures were grown in nitrate minimal media for 5 days with OD readings taken daily. The plotted data points are an average of at least three independent experiments and the standard error for each point is included.

3.4 MoaX is a fused MPT synthase

Williams *et al*. (2011) previously demonstrated that the *moaX* gene can fully restore MoCo biosynthesis when expressed in an MPT-synthase deficient mutant. Sequence alignments of the *Mtb* proteins reveal that both MoaD1 and MoaD2 align to the N-terminus of MoaX (Figure 3.13) and MoaE1 and MoaE2 to the C- terminus. The active site of MPT synthase is located within a pocket of MoaE and contains conserved C-terminal Gly residues of MoaD which are directly involved in enzyme activity. It was thus hypothesized that MoaX would have to be post-translationally processed into MoaD and MoaE components to provide access to the residues Gly81 and Gly82 of MoaX for subsequent chemical modification. To monitor the fate of MoaX when expressed in a mycobacterial host, an epitope tagging method was employed (Figure 3.14).

```
MoaD1           ---------MIKVNVLYFGAVREACDETPREEVEVQNGTDVGNLVDQLQQKYPRLRDHCQ 51
MoaX            ---------MITVNVLYFGAVREAC-KVAHEKISLESGTTVDGLVDQLQIDYPPLADFRK 50
MoaD2           VTQVSDESAGIQVTVRYFAAARAAA-GAGSEKVTLRSGATVAELIDGLSVRDVRLATVLS 59
Ecoli           -----------MIKVLFFAQVRELV-GTDATEVAADF-PTVEALRQHMAAQSDRWALALE 47
                            :.* :*. .*     .   ::     . *  * : :          .

MoaD1           RVQMAVN--QFIAPLSTVLGDGDEVAFIPQVAGG-------------------------- 83
MoaX            RVRMAVN--ESIAPASTILDDGDTVAFIPQVAGGSDVYCRLTDEPLSVDEVLNAISGPSQ 108
MoaD2           RCSYLRDG-IVVRDDAVALSAGDTIDVLPPFAGG-------------------------- 92
Ecoli           DGKLLAAVNQTLVSFDHPLTDGDEVAFFPPVTGG-------------------------- 81
                           :      *  ** : .:* .:**

MoaD1           ------------------------------------------------------------
MoaX            GGAVIFVGTVRNNNNGHEVTKLYYEAYPAMVHRTLMDIIEECERQADGVRVAVAHRTGEL 168
MoaD2           ------------------------------------------------------------
Ecoli           ------------------------------------------------------------


MoaD1           ----------------------------------------------------
MoaX            RIGDAAVVIGASAPHRAAAFDAARMCIERLKQDVPIWKKEFALDGVEWVANRP 221
MoaD2           ----------------------------------------------------
Ecoli           ----------------------------------------------------
```

Figure 3.13: Sequence alignment of *E. coli* MoaD and *Mtb* MoaD1, MoaD2 and MoaX proteins. Conserved Gly residues are shown in red. Alignment was generated using sequences obtained from Tuberculist (http://genolist.pasteur.fr/TubercuList/) and the ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/) online alignment tool.

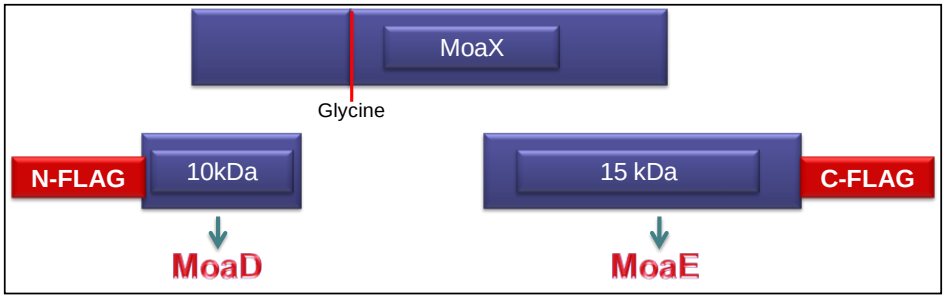


Figure 3.14: Schematic representation of the cleavage of MoaX showing the predicted site of cleavage and the expected sizes of each subunit once MoaX is processed at this site.

3.5 FLAG™-tagged *moaX*

The vectors pFLAGmoaXC and pFLAGmoaXN (Figure E 10 and E 11 respectively, Appendix E 4) were generated as described in section 2.16 and introduced into the *M. smegmatis* double mutant Δ*moaD2* Δ*moaE2* in conjunction with the TetR-expressing vector, pMC1s. A single

colony was picked for each electroporation and the presence of the vector was confirmed by PCR analysis.

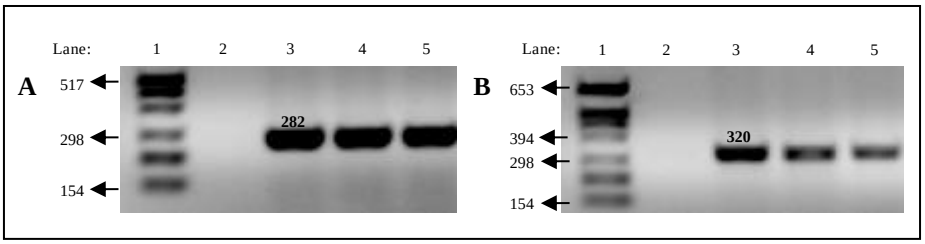


Figure 3.15: PCR confirmation of the site-specific integration of pMC1s (A) PCR amplicons from reaction using the primers *attL2* and *attBS1*. (B) PCR amplicons from reactions using the primer set *attL4* and *attBS2*. Both gels were loaded in the same order Lane 1: Marker λVI, Lane 2: No DNA control, Lane 3: Positive control, Lane 4: Amplicon from transformant carrying N-terminally FLAG-tagged MoaX, Lane 5: Amplicon from transformant carrying C-terminally FLAG-tagged MoaX.

PCR reactions were carried out to validate the site-specific integration of pMC1s (Figure 3.15) as well as the presence of *moaX* carried on the episomal vector pFLAGEM (Figure 3.16) with the primers moaX-F and moaX-R. pMC1s has an L5 based integration system and integrates into the chromosome in the same manner as pHINT shown in Figure 3.2. The correct sizes were observed for each primer set proving that pMC1s was present in each strain and integrated at the correct position. The same template DNA was also used in PCR reactions with primers specific for *moaX* in order to confirm the presence of the gene. Vector DNA was used as a positive control (Figure 3.16, Lane 3) to which the amplicons shown in Lanes 5 and 6 were compared. The correct amplicons sizes were observed thus confirming the strains Δ*moaD2*Δ*moaE2::*pMC1s (pFLAGmoaXN) and Δ*moaD2* Δ*moaE2::*pMC1s (pFLAGmoaXC) which were used in subsequent experiments.

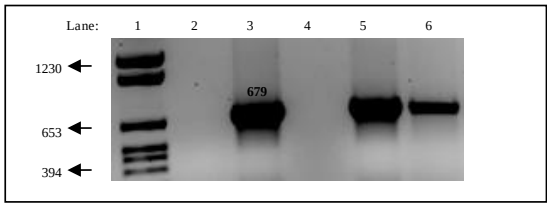


Figure 3.16: PCR confirmation of the presence of *moaX* in strains complemented with pFLAGmoaXN and pFLAGmoaXC. Lane 1: Marker λVI, Lane 2: No DNA control, Lane 3: Positive control, Lane 4: Empty, Lane 5: Amplicon from Δ*moaD2* Δ*moaE2::*pMC1s (pFLAGmoaXN), Lane 6: Amplicon from Δ*moaD2* Δ*moaE2::*pMC1s (pFLAGmoaXC).

3.6 FLAG-tagging does not abrogate the function of *moaX*

In order to assess whether the incorporation of the FLAG tag onto the N- or C-terminus of *moaX* interfered with its function, growth in MPLN media was monitored as previously described, with one modification. In this system, the expression of *moaX* is essential for growth of the strains. As mentioned previously, the expression of *moaX* in Δ*moaD2* Δ*moaE2::*pMC1s (pFLAGmoaXC) and Δ*moaD2* Δ*moaE2::*pMC1s (pFLAGmoaXN) is under the control of the Tet system and therefore requires the addition of ATc for induction. This growth curve was modified by the addition of ATc (50 ng/ ml) to one set of cultures whereas a duplicate set had no ATc. Based on this, it was expected that the cultures lacking ATc would not grow because *moaX* is repressed. As expected, growth was conditional in the presence of ATc inducer (Figure 3.17). This experiment also confirmed that the incorporation of the FLAG tag at either the N- or C-terminus of MoaX did not disrupt MoaX function. Although a lag was observed for the strain carrying the C-terminally tagged MoaX (Figure 3.17B), growth was still observed and eventually reached levels similar to the wild type strain.

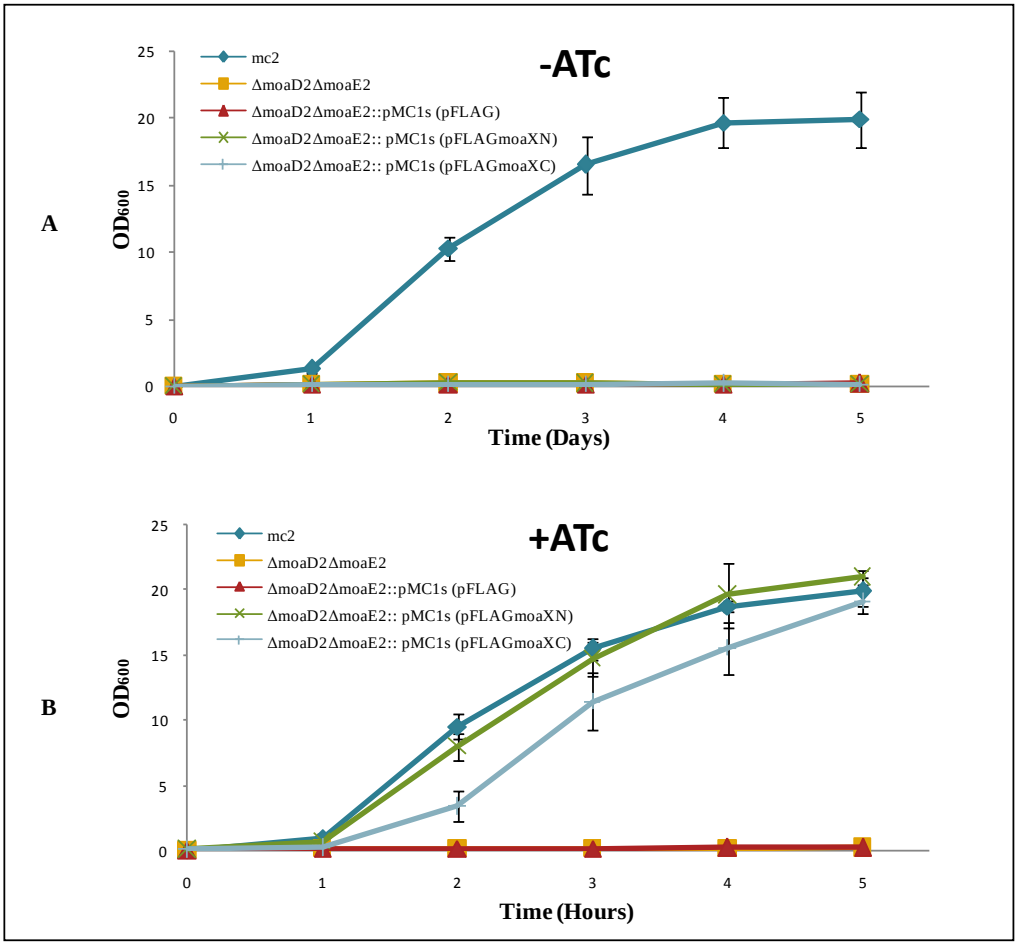


Figure 3.17: Growth curve analysis of strains carrying FLAG-tagged *moaX*. Strains were grown in either the (A) absence or (B) presence of the inducer anhydotetracylcine. Growth curves were performed in nitrate minimal media with OD readings taken daily for 5 days.

3.7 MoaX processing

To assess whether MoaX is cleaved into constituent MoaD and MoaE domains, Western blot analysis of strains carrying tagged MoaX was performed. The first experiment was carried out using protein samples extracted from the cultures of the double mutant carrying N- or C-terminally tagged MoaX, expressed under the control of ATc, using both tight (PMC1s) and intermediate repression (PMC2m). Western blot analysis revealed the presence of two major bands of sizes 25 kDa and ~15 kDa for the C-terminally tagged protein (Figure 3.18). The 25 kDa band corresponds to un-cleaved MoaX while the 15 kDa band corresponds to the expected

MoaE component of MoaX. In contrast, only the 25 kDa, un-cleaved band was detected for N-terminally FLAG-tagged MoaX. This could be due to a number of possibilities including experimental artifacts, the removal of the FLAG tag during processing of MoaX or the inability of the tag to be detected due to protein folding. That these bands correspond to MoaX is supported by the fact that they are only seen in samples where FLAG-tagged *moaX* is present but not in the sample containing the empty pFLAGEM vector. In addition, it can be seen that protein expression is induced in the presence of ATc which significantly abrogates the effects of strong repression. In contrast protein expression is observed under medium repression in the absence of ATc. Every attempt was made to load equal amounts of protein in each lane; however, SDS-PAGE gels shown in Figure 3.18B indicate that some differences in protein concentration were evident. Consequently, no conclusion can be made on levels of protein expression in the repressed/de-repressed/no-repressor strains.

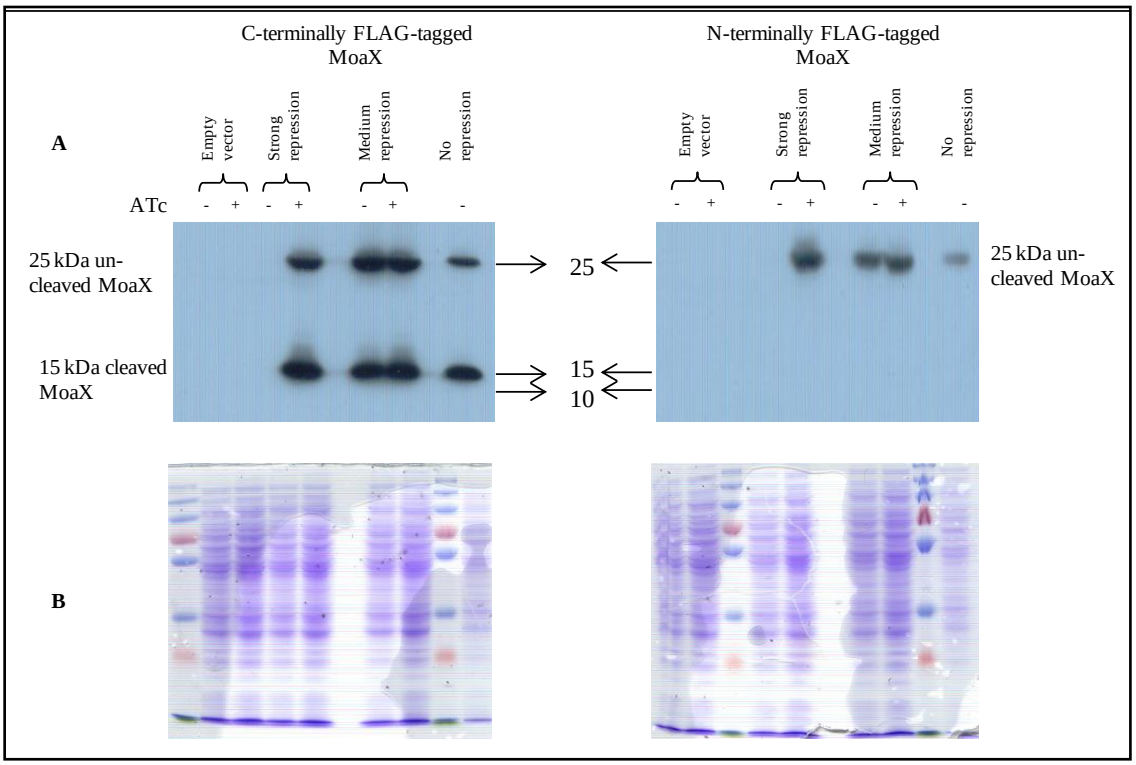


Figure 3.18: Western blot showing the post-translational cleavage of MoaX.(A) Western blots of N- and C-terminally FLAG tagged MoaX from strains grown in 7H9 and the presence or absence ATc (50 ng/ml). (B) Coomasie blue stained gels corresponding to the blots above them. These are images of one of three independent experiments.

The Western blot images shown in Figures 3.18 were performed with protein samples extracted from cultures grown in conventional 7H9 media. The growth curves performed throughout this study were however performed in MPLN media. Therefore, the effect of growth in modified media on MoaX processing was also assessed. As seen in Figure 3.19, processing of MoaX occurs in both 7H9 and MPLN media suggesting that cleavage is not dependent on media conditions. The cleavage of the N- terminally FLAG tagged MoaX could not be detected in either 7H9 or MPLN media, with only the full length product being observed (data not shown). In addition to assessing MoaX cleavage in MPLN from the complemented double mutant strain under strong repression, cleavage was also assessed in the wild type strain conditionally expressing C-terminally tagged MoaX. In this case, three prominent bands are observed in the western blot- 49 kDa, 25 kDa and 15 kDa (Figure 3.19A). The 25 kDa and 15 kDa bands are induced in an ATc-dependent manner and correspond to un-cleaved and cleaved MoaX respectively. The intensity of the 49 kDa band does not change in the presence or absence of repression, suggesting that it is a non-specific band. This is consistent with a previous observation where the same induction system and antibody were used (Ahidjo *et al*., 2011). The Coomasie-blue-stained gels (in Figure 3.19B) were included to show that approximately equivalent amounts of total protein were loaded (except for the un-induced double mutant strain cultured in modified MPLN as this strain is dependent on MoaX for growth in this medium). From the blots it can be seen that MoaX processing occurs in 7H9 media when MoCo biosynthesis is not required to support growth and survival as well as in MPLN media, where it is. This suggests that there is most likely a general protease which recognizes a cleavage signal on the protein when made.

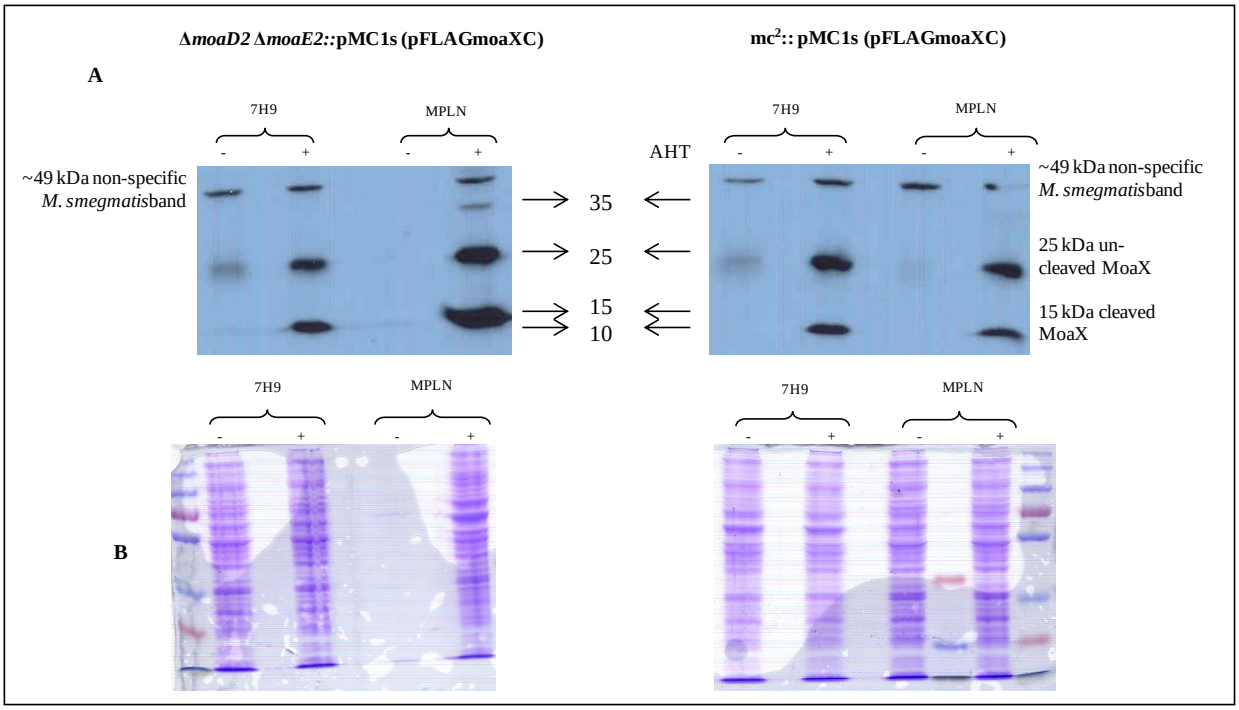


Figure 3.19: MoaX is cleavage is not altered by media composition. (A) Western blot of protein samples extracted from Δ*moaD2* Δ*moaE2::*pMC1s (pFLAGmoaXC) and mc^2:: pMC1s (pFLAGmoaXC) grown in modified *M. phlei*. (B) Gel images of the corresponding blots above.

Further resolution of results from the N-terminally FLAG-tagged MoaX are required to make definitive conclusions. Currently, it is unclear why cleavage of MoaX does not occur when the protein carries an N-terminal FLAG-tag. In this regard, the ability of the N-terminally FLAG-tagged MoaX to complement the Δ*moaD2* Δ*moaE2* mutant suggests that MoaX may also function without proteolytic processing. Further experimentation is required to dissect this and other possibilities.

3.8 Essential MoaX residues

Previous studies aimed at assessing the structure of the MPT synthase and the role of different amino acids in its function have led to the identification of the two terminal Glycine residues in MoaD as being important for MPT synthase activity (Schmitz *et al.*, 2007). The alignment of MoaX with MoaD2 shows that there are also two conserved Gly residues in MoaX, Gly81 and

Gly82 (Figure 3.13). Furthermore, the Western blot analysis using C-terminally tagged MoaX suggested possible processing in this region of the protein. In order to assess whether the two Gly residues of MoaX are important for MoaX function (and/or cleavage), site-directed mutagenesis was performed where each Gly was individually mutated to an Ala. The method involved two rounds of PCR reactions, the results of which are shown in Appendix E5 along with the generation and confirmation of the vectors carrying mutated versions of *moaX* (Figure E 17 and E 18).

3.9 Gly81 and Gly82 are both essential for MoaX function

The point mutations 242G→C and 245G→C encode single amino acid substitutions in MoaX, G81A and G82A respectively. In order to assess the effect of these substitutions on the activity of MoaX, growth curves were performed in MPLN media. Included as positive controls in this experiment were the wild type strain and double mutant strain expressing C-terminally tagged, wild type MoaX, while the double mutant was used as a negative control.

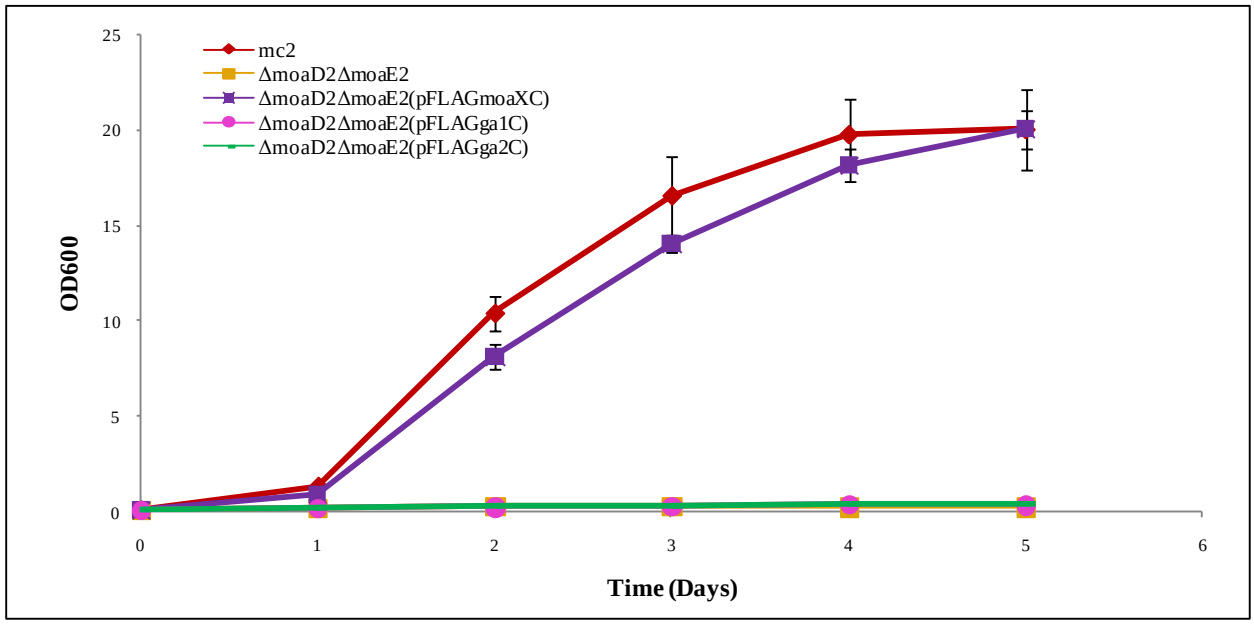


Figure 3.20: Growth curve analysis of strains carrying FLAG-tagged derivatives of *moaX* with either a 242G>C or 245G>C mutation. C-terminally FLAG-tagged MoaX and mc^2 were included as positive controls for growth. Growth curves were performed in MPLN with optical density readings taken daily for five days.

The data shown Figure 3.20 confirm that neither strain carrying a mutated copy of *moaX* was able to grow in nitrate minimal media suggesting that these strains were unable to synthesize a functional MPT synthase enzyme. This result proves that both Gly81 and Gly82 are essential for MoaX activity, consistent with the function of this enzyme as a canonical MPT synthase.

3.10 Gly81 is important for MoaX cleavage

In order to investigate what effect, if any, the mutations had on MoaX cleavage, Western blots were carried out with protein samples extracted from Δ*moaD2* Δ*moaE2* (pFLAGga1C) and Δ*moaD2* Δ*moaE2* (pFLAGga2C) grown in 7H9, where MoaX is not required for growth. Wild type, C-terminally tagged *moaX* was included as a positive control in this experiment. The Coomasie blue stained gel below the blot shows that approximately equivalent amounts of total protein were loaded in each case.

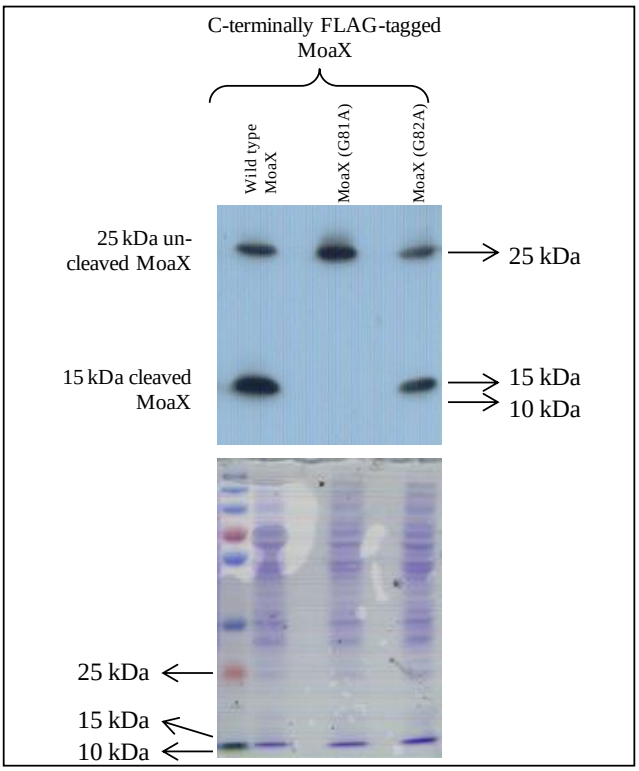


Figure 3.21: Western blot analysis of protein extracts from strains carrying mutated copies of *moaX*. Wild type MoaX serves as a positive control for cleavage. Mutation G81A of C-terminally FLAG-tagged MoaX abolished cleavage, whereas mutation G82A did not interfere with processing.

The blot shows that in addition to being essential for function, Gly81 is also essential for cleavage of MoaX (Figure 3.21) since no measurable cleavage was detected with this protein. However, although essential for function, Gly82 does not appear to be essential for the cleavage of MoaX. This result suggests that proteolytic cleavage of MoaX is required to constitute a functional heterotetramer of MPT synthase however, no definitive conclusion can be made until further resolution of the results obtained with the N-terminally FLAG-tagged protein.

3.11 MoaX is not functional in *E. coli* due to incorrect cleavage

A recent study carried out by Voss *et al.* (2011) reported that *Mtb moaX* was unable to complement *E. coli moaD* and *moaE* single mutants, a result that is in contrast to what is observed in *M. smegmatis.* These authors hypothesized that this could be because *E. coli* cells lack the cleavage machinery required to generate the MoaD and MoaE components of MoaX (Voss *et al*., 2011). In order to evaluate this hypothesis FLAG-tagged MoaX was extracted from wild type *E. coli* cells carrying pFLAGmoaXC and pFLAG and assessed by Western blot (Figure 3.22). In this study, all MoaX clones were propagated in an *E. coli* DH5α strain (cloning host, genotype: *sup*E44 Δ*lac*U169 *hsd*R17 *recA*1 *end*A1*gyrA*96 *thi*-1 *relA*1) which provided an opportunity to test MoaX cleavage in this organism without the presence of an inducer such as ATc for protein expression.

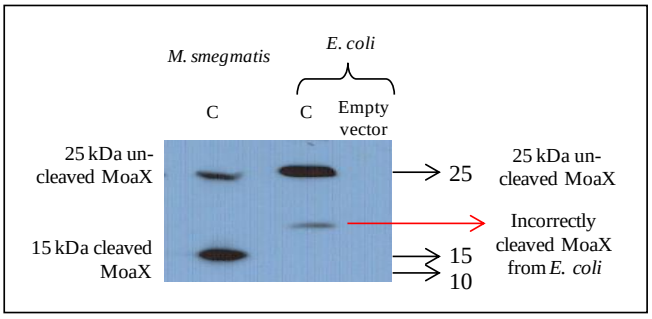


Figure 3.22: Western blot analysis of FLAG-tagged MoaX protein samples extracted from *E. coli* and *M. smegmatis.*C-terminally tagged MoaX from *M. smegmatis* serves as a positive control for cleavage. The blot shown is one of three independent experiments.

From Figure 3.22 it can be seen that the cleavage product of C-terminally FLAG-tagged MoaX samples from *E. coli* is larger than expected. That the bands observed are FLAG-tagged MoaX – related is supported by the absence of these bands in the sample carrying only the empty pFLAGEM vector. This result demonstrates that MoaX is unable to complement the *E. coli* mutants due to incorrect cleavage or processing, further highlighting the importance of accurate processing and cleavage of MoaX in order for it to be functional as a canonical MPT synthase enzyme.

3.12 Generation of *M. smegmatis* Δ*narB* knock-out mutant

To ascertain whether NarB is responsible for growth in MPLN, a knock-out mutant was generated in *M. smegmatis* and assessed for its ability to grow in this media. Using the pΔnarB knock-out construct (Figure E 21, Appendix E 6), a mutant was generated as described in section 2.20.2. Only two blue SCO colonies were obtained from the electroporation reaction with pΔnarB into wild type cells. One of these SCO's was picked and grown in the absence of antibiotic selection, followed by growth in the presence of sucrose to allow for the second recombination event to occur. Eight white colonies were picked from 7H10 plates supplemented with X-gal and sucrose and screened using a PCR strategy.

The PCR strategy used for the screening of knock-out mutants requires that three primers be used in the reaction simultaneously, two flanking the gene and one situated within the region deleted from the gene. With wild type DNA all three primers would be able to anneal to the template; however due to competition for PCR reagents such as dNTPs and the use of a conventional Taq enzyme, the smaller product of Primers 2 and 3 will be preferentially amplified to produce an amplicon of 430 bp. In the knock-out mutant the *narB* gene would be absent

rendering Primer 2 unable to anneal to the template resulting in the production of an amplicon of 270 bp between Primers 1 and 3. The position of the primers used and expected amplicon sizes are shown in Figure 3.23A. Of the eight colonies, six appeared to be mutants, as evidenced by the presence of the 270 bp mutant band observed in Lanes 7-9 and 12-14 of Figure 3.23B. The remaining two colonies were wild type revertants, evidenced by the 420 bp wild type band seen in Lanes 6 and 11. Colony 2 was picked and the genotype was further confirmed by southern blot analysis. Genomic DNA was extracted from wild type *M. smegmatis* and Colony 2, the supposed Δ*narB* mutant, using the CTAB method described in section 2.5.4. The enzymes chosen to perform the Southern blots were *Not*I, *Sac*I and *Nco*I because the differences in fragment sizes between the wild type and mutant strain would be most notable with these enzymes. Restriction digests were set up using those enzymes and 2 µg of the genomic DNA mentioned above for each reaction. The Southern blot protocol described in section 2.21 was then followed using the probes shown in Figure 3.23A which correspond to the upstream and downstream regions used to generate the suicide vector pΔ*narB*. The results of the Southern blots for the upstream and downstream probe are shown in Figure 3.23C and D respectively with the expected sizes shown in Figure 3.23A.

Lane 1 in Figure 3.23C and D show the 3.1 kb band of Marker λIV which serves as a control to show that the procedure has worked in addition to serving as a size control. From the Southern blots it can clearly be seen that there is a size difference between the wild type and mutant strain in the regions probed. An increase in the size of the band observed for the *Not*I digests (5.8 kb vs 3.9 kb and 2.9 kb) in the mutant strain is due to the loss of a restriction site present in the wild type gene confirming that *narB* is no longer present in the mutant. Further confirming the loss of *narB* in the mutant is the reduction in size of the *Sac*I and *Nco*I fragments in the mutant observed

in Lanes 7. The Southern blots definitively prove that *narB* is deleted in this strain, now designated as Δ*narB*.

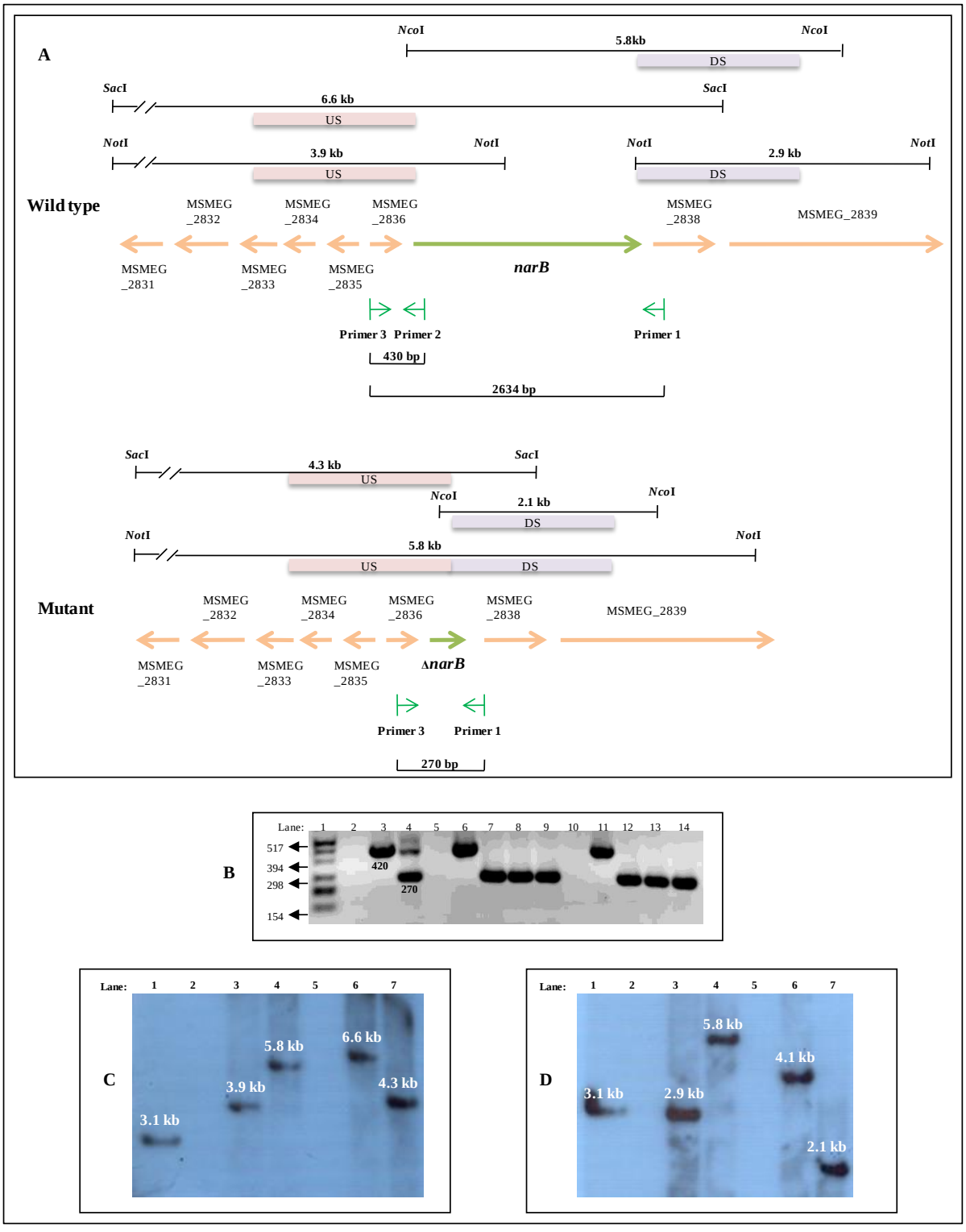


Figure legend on page 86

Figure 3.23: Screening and genotypic confirmation of Δ*narB*. (A) Schematic representation of genomic maps of wild type and mutant *narB* regions. Primer positions (green) and expected amplicons for PCR screening strategy are shown. Restriction enzymes, probes and expected fragment sizes for southern blot confirmation are also depicted. Upstream (UP) and downstream (DS) probes are shown in pink and purple respectively (B) PCR screening of Δ*narB* knock-out mutants. Lane 1: Marker λVI, Lane 2:-, Lane 3: Wild type positive control, Lane 4: SCO positive control, Lane 5: -, Lane 6: Colony 1, Lane 7: Colony 2, Lane 8: Colony 3, Lane 9: Colony 4, Lane 10: -, Lane 11: Colony 5, Lane 12: Colony 6, Lane 13: Colony 7, Lane 14: Colony 8. (C) Southern blot with upstream probe. Lane 1: Marker λIV, Lane 2: Empty, Lane 3: *Not*I digested wild type DNA, Lane 4: *Not*I digested Δ*narB* DNA, Lane 5: Empty, Lane 6: *Sac*I digested wild type DNA, Lane 7: *Sac*I digested Δ*narB* DNA. (D)Southern blot with downstream probe. Lane 1: Marker λIV, Lane 2: Empty, Lane 3: *Not*I digested wild type DNA, Lane 4: *Not*I digested Δ*narB* DNA, Lane 5: Empty, Lane 6: *Nco*I digested wild type DNA, Lane 7: *Nco*I digested Δ*narB* DNA.

3.13 *narB* is dispensable for growth in nitrate minimal media

As discussed previously our hypothesis is that NarB is the putative assimilatory nitrate reductase responsible for the growth of *M. smegmatis* on nitrate minimal media (Khan *et al.*, 2008). NarB is a MoCo- dependent enzyme and the aim of knocking out the gene which encodes the protein was to investigate whether the failure of Δ*moaD2* Δ*moaE2* to grow in nitrate minimal media was due to the inability of the strain to activate NarB with its cofactor, *bis*-MGD. To assess this hypothesis, Δ*narB* was grown in MPLN media with wild type as a positive control and the Δ*moaD2* Δ*moaE2* double mutant as a negative control. The results of this experiment can be seen in the growth curve depicted in Figure 3.24.

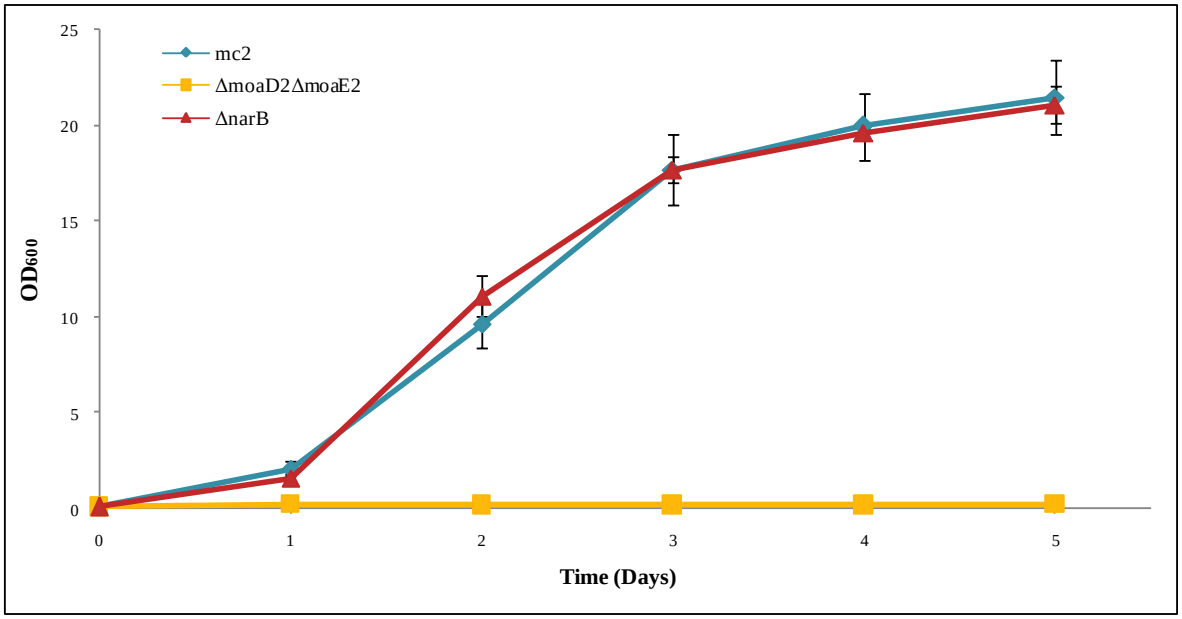


Figure 3.24: Growth curve analysis of Δ*narB* in nitrate minimal media shows that it is dispensable for growth. Averages of at least three independent experiments were plotted for each strain with standard errors included.

Surprisingly, the Δ*narB* strain was able to grow in MPLN media just as well as the wild type strain confirming that NarB is either not involved in, or not solely responsible for nitrate assimilation in *M. smegmatis* under the conditions tested. This result raises the question of which other enzyme/s are responsible for nitrate assimilation in the organism.

In addition to *narB, M. smegmatis* also harbors the *narGHJI* operon, which encodes the respiratory NR comprising NarG, NarH and NarI, which is assembled by the NarJ chaperone. However, Weber *et al.* (2000) had reported that unlike its counterpart in *Mtb*, this *M. smegmatis* enzyme does not display respiratory NR activity. In *Mtb,* NarGHI is responsible for both respiratory and assimilatory NR activity. To explain the lack of effect of the *narB* deletion, it was hypothesized that NarGHI may also be involved in nitrate assimilation in *M. smegmatis*. To test this hypothesis, a Δ*narGHJI* single mutant and Δ*narB* Δ*narGHJI* double mutant were constructed in *M. smegmatis* and assessed for growth in MPLN.

3.14 Generation of Δ*narGHJI* and Δ*narB* Δ*narGHJI* knock-out mutants

The suicide vector pΔ*narGHJI* was generated as described in section 2.20.1 and confirmed by restriction analysis (Figure E 22). Following the strategy outlined in section 2.20.3, pΔ*narGHJI* was introduced into wild type *M. smegmatis* and the Δ*narB* deletion strain to generate Δ*narGHJI* SCO's in each background. Three blue SCO colonies were generated in the wild type background of which one was selected for further counter-selection. Only one single cross-over recombinant was obtained in the Δ*narB* background. These colonies were then processed to identify double cross-over recombinants. Eight white colonies for the wild type background and ten for the Δ*narB* background were picked from 7H10 plates supplemented with X-gal and sucrose to be screened by PCR.

The PCR strategy used for screening was similar to that described in section 3.12. The position of the primers used along with the expected amplicon sizes can be seen in Figure 3.25A. Of the eight colonies screened by PCR, only one appears to be a Δ*narGHJI* mutant as observed by the presence of the ~470 bp amplicon in lane 12 of Figure 3.25B. The band observed was very faint and another PCR reaction was therefore performed on DNA extracted from the strain by the CTAB extraction method to ensure that it was correct, as confirmed in Lane 5 of Figure 3.25C. This mutant colony was therefore selected for subsequent use. Three of the ten colonies screened from the Δ*narB* background strain appear to be mutants, evidenced by the presence of the 470 bp mutant band seen in Lanes 6, 10 and 12 of Figure 3.25D, of which one was selected for further use.

Southern blots were performed to confirm the genotypes of both Δ*narGHJI* and Δ*narB* Δ*narGHJI* (Figure 3.25 E and F). Genomic DNA was extracted from each strain and restriction digests were set up for three restriction enzymes with 2 µg DNA. The restriction enzymes used were *Mlu*I, *Pst*I and *Bam*HI. The probes were the upstream and downstream regions used to generate the suicide vector pΔ*narGHJI* and can be seen in Figure 3.25A along with the expected fragment sizes.

Figure 3.25: Screening and genotypic confirmation of Δ*narGHJI* and Δ*narB ΔnarGHJI*. (A) Schematic representation of genomic maps of wild type and mutant *narGHJI* regions. Primer positions (red arrows) and expected amplicons for PCR screening strategy are shown. Restriction enzymes, probes and expected fragment sizes for southern blot confirmation are also depicted. Upstream (US) and downstream (DS) probes are shown in grey and green respectively. (B) PCR screening of Δ*narGHJI* single mutant. Lane 1: Marker λVI Lane 2: No DNA control, Lane 3: Positive wild type control, Lane 4: Positive mutant control, Lane 5: -, Lane 6- Lane 9 and Lane 11, 13 and 14: Wild type revertant colonies Lane 10: -, Lane 12: Mutant colony. (C) Re-amplification of genomic DNA extracted from the Δ*narGHJI* mutant colony. (D) PCR screening of Δ*narB* Δ*narGHJI* double mutants. Lane 1: Marker λVI, Lane 2: No DNA control, Lane 3: Positive wild type control, Lane 4: Positive mutant control, Lane 5, 7-9, 11 and 13-14: Wild type revertant strains, Lane 6, 10 and 12: Double mutant strains. (E) Southern blot with upstream probe. Lane 1: Marker λIV, Lane 2: *Bam*HI digested wild type DNA, Lane 3: *Bam*HI digested Δ*narGHJI* DNA, Lane 4: *Bam*HI digested Δ*narB* Δ*narGHJI* DNA. (F) Southern blot with downstream probe. Lane 1: Marker λIV, Lane 2: *Pst*I digested wild type DNA, Lane 3: *Pst*I digested Δ*narGHJI* DNA, Lane 4: *Pst*I digested Δ*narB* Δ*narGHJI* DNA, Lane 5: Empty, Lane 6: *Mu*I digested wild type DNA, Lane 7: *Mlu*I digested Δ*narGHJI* DNA, Lane 8: *Mlu*I digested Δ*narB* Δ*narGHJI* DNA.

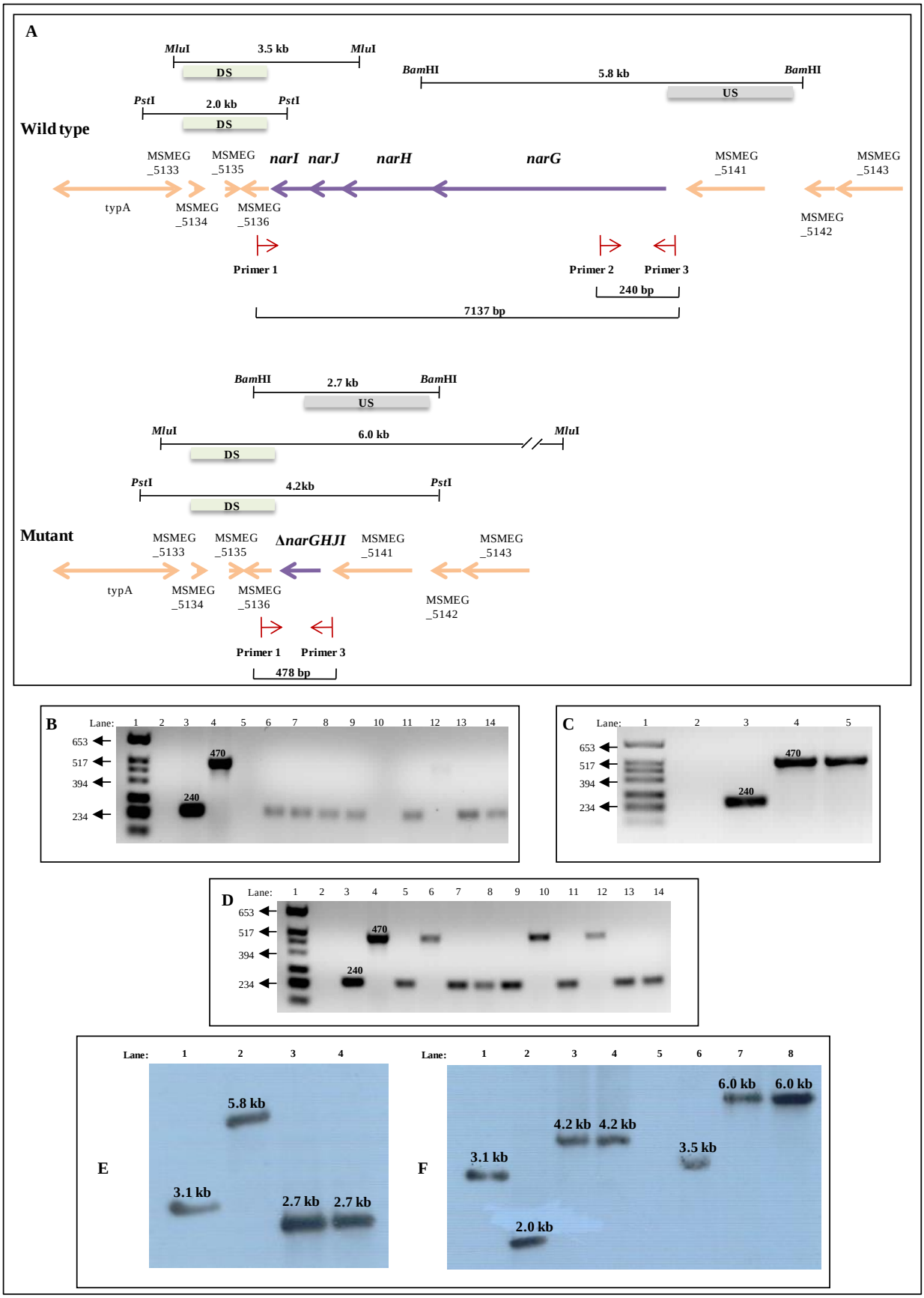


Figure legend on page 88

The *Bam*HI restriction digest shows a reduction in the size of the band probed with the upstream probe from 5.8 kb in wild type to 2.7 kb in the mutant. This reduction in size was due to the loss of a *Bam*HI restriction site found in *narG* when the operon is knocked out. Three *Mlu*I and five *Pst*I restriction sites are present in the wild type *narGHJI* operon with the next closest *Mlu*I sites being 4.3 kb upstream of *narG* and 1.7 kb downstream of *narI* and the next closest *Pst*I sites being 2.4 kb upstream of *narG* and 1.8 kb downstream of *narI*. A band of 6 kb is observed in the mutants (Lanes 7 and 8, Figure 3.25F) for the *Mlu*I digest with the downstream probe corresponding to what would be expected if the operon is lost (4.3 kb + 1.7 kb). For the *Pst*I digest a 4.2 kb band is observed for the mutants (Lanes 3 and 4) which also corresponds to what was expected when the operon is lost (2.4 kb + 1.8 kb). Therefore an increase in the size of the fragments observed for these enzymes when the downstream probe is used confirms the loss of the entire operon.

3.15 Both *narB* and *narGHJI* are dispensable for growth in nitrate minimal media

It was shown in section 3.13 that *narB* is dispensable for growth with nitrate as a sole nitrogen source. It was then hypothesized that *narGHJI* could play a compensatory role in the absence of *narB* to serve as an assimilatory NR as the *Mtb* homologue does. The mutant strains Δ*narGHJI* and Δ*narB* Δ*narGHJI* were thus generated and assessed for growth in MPLN media (Figure 3.26). Wild type was included as a positive control, the double mutant Δ*moaD2* Δ*moaE2* as a negative control and Δ*narB* was included to compare the growth kinetics of the three mutants generated. The results show that both the Δ*narGHJI* and Δ*narB* Δ*narGHJI* mutants are indistinguishable from wild type when grown in the presence of nitrate as the sole nitrogen source. This result, along with that shown in Figure 3.24, suggests that in addition to NarB and NarGHI, *M. smegmatis* possesses another MoCo-dependent enzyme/s for nitrate assimilation.

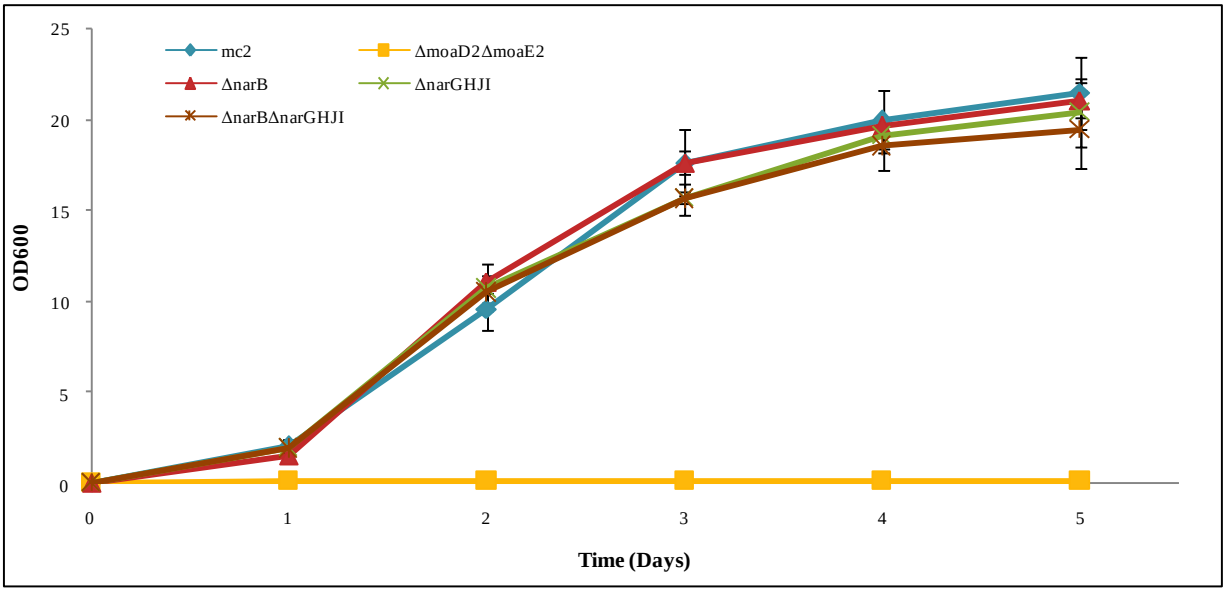


Figure 3.26: Growth curve analysis of Δ*narB*, Δ*narGHJI* and Δ*narB* Δ*narGHJI* in nitrate minimal media shows that both genes are dispensable for growth in nitrate minimal media. Averages of at least three independent experiments were plotted for each strain with standard errors included.

4 Discussion

Mtb has an extraordinary ability to adapt and survive in the host and still maintains its status as a devastating human pathogen despite widespread vaccination and chemotherapeutic intervention. The current treatment protocols and control programs have failed to eradicate TB disease from human society and this is, in part, due to the ability of the tubercle bacillus to rapidly develop drug resistance and persist in the host for a protracted period of time. Considering this, a thorough understanding of the metabolism and physiology of *Mtb* would aid in the development of more effective intervention strategies. This study aimed to gain a better understanding of one such metabolic pathway, MoCo biosynthesis, which has been implicated in the survival of *Mtb* in several forward genetic screens that identify genes/pathways that are essential for survival under various conditions *in vitro* (Sassetti *et al*., 2003), in macrophages (Brodin *et al*., 2010) and in the murine model of TB infection (Camacho *et al*., 1999).

Although highly conserved, the MoCo biosynthetic pathway is notable in *Mtb* in that it displays a multiplicity of homologues for the genes encoding MPT synthase, the enzyme responsible for catalysis of the second step (Figure 1.3). The tubercle bacillus also encodes an expanded genetic repertoire for the genes involved in the first step of the MoCo biosynthetic pathway, but these have not been investigated. With respect to the genes encoding the MPT synthase, *Mtb moaD1*, *moaD2*, *moaE1* and *moaE2* have all been shown to contribute to MoCo biosynthesis (Williams *et al*., 2011); however, it was not known whether those gene products are able to associate in different combinations to form chimeras of the enzyme, possibly with varied activities (Williams *et al.*, 2013), as observed for human-*E. coli* chimeras of MPT synthase, which are able to function *in vitro* (Leimkühler *et al*., 2003). Due to the demonstrated functionality of each

homologue (Williams *et al*., 2011), it was expected that several isoforms of the enzyme might be catalytically proficient. However, the results presented in this study demonstrate that only *moaD2* and *moaE2*, when present on episomal vectors, were able to form an MPT synthase that could complement a double mutant of *M. smegmatis* that lacks these homologues (Δ*moaD2* Δ*moaE2*). All other combinations, on episomal vectors, of *Mtb* homologues were unable to complement the conditional growth phenotype of this *M. smegmatis* mutant on nitrate as sole nitrogen source (Figure 3.12). This contrasted with the previous finding that all three *Mtb moaD* homologues contributed to MoCo biosynthesis in *Mtb* and suggested that the discrepancy may be due to differences in the biology of the heterologous *M. smegmatis* host and nuances in the second step of the MoCo biosynthesis pathway. No heterologous complementation was observed when the *Mtb moaD* and *moaE* homologs were provided on integrating vectors. This could be due to a multitude of factors including reduced gene dosage, the presence of two *hsp60* promoters or other vector toxicity effects. In the absence of any data on gene expression from these vectors, no definitive conclusion can be made regards the lack of functionality in this case.

The MPT synthase-catalyzed second step of MoCo biosynthesis is highly complex, requiring a coordinated series of biochemical reactions driven by the products of several, distinct genes. As discussed previously, MPT synthase is responsible for the transfer of sulfur, carried on the terminal Gly residue of MoaD, to cPMP for the generation of MPT – a reaction that requires constant replenishment of sulfur groups. For sulfuration to take place continuously, the MoaD subunit first needs to be adenylated, followed by thiocarboxylation either by an L-cysteine desulfurase (Zhang *et al*., 2010) or a rhodanese-like protein (Matthies *et al*., 2004). In *E. coli*, MoeB adenylates MoaD which is subsequently sulfurated by IscS with the assistance of the rhodanese-like protein YnjE (Zhang *et* al., 2010; Dahl *et* al., 2011). In mycobacteria the MoeB

proteins contain a rhodanese domain, which was predicted to be directly involved in sulfur transfer and it has recently been shown that in *Mtb* both rhodanese-like proteins, MoeBR and MoeZR are capable of sulfur transfer to both MoaD1 and MoaD2 *in vitro* (Voss *et al*., 2011). The *M. smegmatis* host used in this study only retains a *moeZR* homologue with no detectable *moeBR* gene and this may compromise the ability of the accessory proteins to continuously sulfurate the heterologously expressed *Mtb* homologues. This notion is further supported by the fact that of the four MPT-synthase-encoding genes tested in this study, *moaD1* is not functional in *M. smegmatis*, possibly suggesting that MoaD1 preferentially associates with and is sulfurated by MoeBR, whereas MoeZR may be the preferred interacting partner for MoaD2 (Voss *et al*., 2011; Williams *et al*., 2011; Williams *et al*., 2013). This idea is consistent with the fact that *M. smegmatis* contains only a single *moaD2* homologue, and that *Mtb* acquired the *moaA1-moaB1-moaC1-moaD1* operon together with the downstream *moeBR* by horizontal gene transfer (Williams *et al*., 2011). This provides a plausible explanation for the inability of the *moaD1-moaE1* and *moaD1-moaE2* combinations to complement the *M. smegmatis* Δ*moaD2* Δ*moaE2* mutant. As shown in Figure 3.12, the combination of *moaD2* and *moaE1* was not functional. However, Williams *et al*. (2011) and data from this study, Figure 3.8, confirm that both these genes are individually functional in *M. smegmatis*. The lack of function when both genes are added to a MoCo deficient strain could therefore be due to lack of complex formation or reduced complex formation (Schmitz et al., 2007). These results suggest a functional hierarchy with regards to *Mtb* MPT synthase encoding genes, with *moaD2* and *moaE2* ranking the highest in their ability to function in the heterologous testing system. However, the possibility that the observed hierarchy can be attributed to overall differences in MoCo biosynthetic gene

complements between *Mtb* and *M. smegmatis* cannot be ruled out. The lack of a *moeBR* homologue is particularly relevant in this case.

In addition to *moaD1* and *moaD2*, *Mtb* also encodes *moaX* which is a fused MPT synthase (Williams *et al*., 2011) containing domains of MoaD and MoaE. The crystal structure of MPT synthase (Figure 4.1) is made up of two dimers of MoaD and MoaE which are joined by the two MoaE subunits (Rudolph *et al*., 2001). From the crystal structure, it can be seen that the essential terminal Gly residue of MoaD (yellow in Figure 4.1) is embedded in a pocket of MoaE where the sulfur transfer reaction is hypothesized to occur. Considering this domain organization and the catalysis sequence, it was unclear how the single polypeptide, encoded by *moaX* would be able to function. In this context, an important objective of this study was to assess if post-translational processing of MoaX, in the form of proteolytic cleavage, occurs.

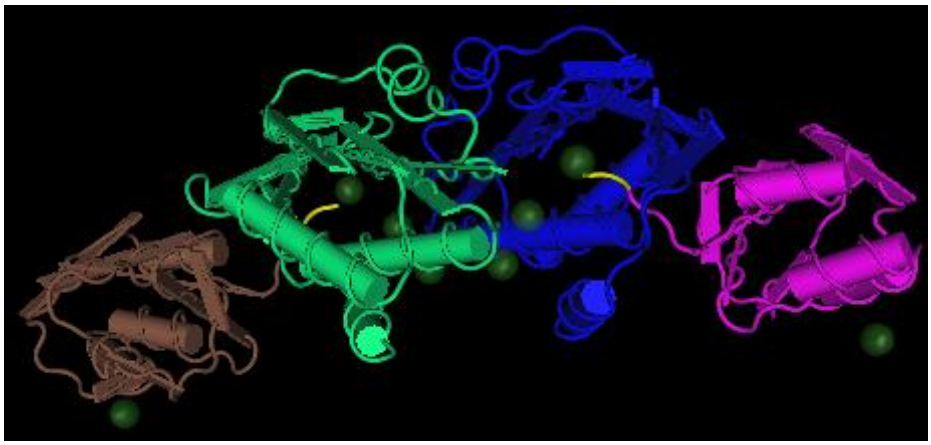


Figure 4.1: Crystal structure of *E. coli* MPT synthase enzyme. MoaD subunits are shown in brown and magenta while MoaE subunits are depicted in cyan and blue. Glycine residues involved in catalysis are shown in yellow.

The data presented for the C-terminally tagged protein demonstrates clearly that cleavage of MoaX does indeed occur and the size of the cleaved product observed suggests that the single peptide is cleaved into its MoaD and MoaE constituent subunits. The Western blot analysis using tagged forms of MoaX suggested that whereas partial processing of MoaX was observed using

C-terminally tagged MoaX, no evidence of processing was seen when the tag was placed at the N-terminus of MoaX. The N-terminus of MoaX aligns to MoaD, and closer inspection of the crystal structure of MPT synthase reveals that the N-terminus of MoaD is folded between an α helix and β sheet of the protein. Furthermore, residue seven of MoaD, a phenylalanine (Phe), forms part of the hydrophobic core of MPT synthase as well as the MoaD-MoeB complex (Rudolph *et al*., 2001; Lake *et al*., 2001). These data indicate that the N-terminus is important for protein stability and addition of the FLAG-tag may have affected protein folding in this region. However, growth curve assays confirmed that incorporation of the FLAG tag on the N-terminus of the protein did not interfere with its function (Figure 3.17). Hence, no clear explanation as to why the FLAG-tag on the N-terminus is not detected after cleavage can be provided and considering this, the possibility that no cleavage occurs in this case cannot be ruled out.

The cleavage of MoaX inferred from C-terminal tagging suggests that the MoaD component would be released, subjected to sulfuration, thereby activating it for catalysis and MPT synthesis. Sequence alignments of MoaX with MoaD1, MoaD2 and *E. coli* MoaD reveal that MoaX (Figure 3.13) also contains conserved C-terminal Gly residues within the MoaD domain (Gly81 and Gly82) which have been implicated in MoaD function. To test whether these residues are essential for MoaX function, these residues were individually mutated to alanines by site-directed mutagenesis. Importantly, both terminal Gly residues were shown to be critical for MoaX function. Studies with the *E. coli* MoaD subunit have shown that the residue corresponding to Gly81 of MoaX is important for the formation of a complex with MoaE while the terminal Gly82 residue is required for adenylation and formation of a complex with MoeB (Schmitz *et al*., 2007). In addition, it was shown that the addition of a further Gly residue to the C-terminus of MoaD resulted in complete abrogation of function suggesting that the two, C-

terminal glycines are important for function and protein stability or the stability of the resulting protein complexes with MoeB and other accessory proteins (Schmitz *et al*., 2007). These observations support the notion that the terminal Gly residues of the MoaD component need to be exposed by proteolytic processing in order for MoaX to be functional. Mutation of Gly81 abrogates the function, and importantly the cleavage of MoaX (Figure 3.21), providing compelling evidence that cleavage is required in order for the protein to be functional. Although full length MoaX is observed even when cleavage does occur, whether the protein is functional in this form is unknown but the mutational analysis suggests that it is not. Purification of recombinant forms of the wild type and mutated versions of MoaX followed by an assessment of their activity and ability to form catalytic complexes with MoeZR and MoeBR *in vitro* would provide some insight in this regard.

Sequence alignments also reveal that MoaX shares more similarity with MoaD1 than with MoaD2. As discussed above, this would suggest that the subunit corresponding to MoaD could be sulfurated by either MoeBR or MoeZR, with a preference for the former. However, MoaX is functional in *M. smegmatis* suggesting that MoeZR would be the protein responsible for sulfuration. In a recent study by Voss *et al.* (2011), it was established that MoaX was not functional in *E. coli* and was unable to interact with either MoeBR or MoeZR from *Mtb*. The authors speculated that this was due to the lack of the MoaX cleavage machinery in *E. coli* as opposed to an inherent inability of these proteins to interact. The results obtained in this study reveal that MoaX is partly cleaved when constitutively expressed of shuttle vectors in *E. coli* DH5α*;* however, the cleavage product was larger than expected, which might explain the inability of MoaX to function in *E. coli*. The protease responsible for the cleavage of MoaX is yet to be identified.

In addition to MoaD1 and MoaD2, MoeZR is also able to transfer sulfur to CysO, a protein involved in cysteine biosynthesis (Voss *et al*., 2011), highlighting a role for MoeZR in both amino acid and MoCo biosynthesis, and linking these metabolic pathways. Cysteine has been implicated in *Mtb* pathogenesis by providing protection against ROI's and RNI's (Senaratne *et al*., 2006); this also highlights a role for MoeZR under these conditions, which is supported by the up-regulation of *cysM, cysO* and *moeZR* in *Mtb* under oxidative stress (Mehra and Kaushal, 2009). The mycobacterial sulfur source for MoCo biosynthesis remains unknown but is most likely L-cysteine (Voss *et al*., 2011), in which case, an L-cysteine desulfurase such as IscS would transfer sulfur to MoeZR. Furthermore, in *E. coli* IscS is implicated in iron-sulfur cluster homeostasis (Giel *et al*., 2012) and may have a similar role in mycobacteria which, through an interaction with MoeZR, would link iron-sulfur cluster homeostasis with the second step of MoCo biosynthesis. It is therefore reasonable to assume that disruption of the second step of the pathway, in the form of mutations in MPT synthase-encoding genes, would not only affect MoCo biosynthesis but also cysteine biosynthesis and sulfur homeostasis in the cell as a whole. This is further evidenced by the large number of mutants, in forward genetic screens, that map to the first two steps of the pathway (Camacho *et al*., 1999; Sassetti *et al*., 2003; Rosas-Magallanes *et al*., 2007; Macgurn and Cox, 2007; Brodin *et al*., 2010; Dutta *et al*., 2010;).

The assay used to measure MoCo biosynthesis in this study relies on the activity of an assimilatory NR. It was shown that the *M. smegmatis* Δ*moaD2*, Δ*moaE2* and Δ*moaD2* Δ*moaE2* mutants were unable to grow in MPLN, suggesting that NR was non-functional in these strains due to MoCo deficiency. To test this, a knock-out mutant of the encoding gene, *narB*, was generated. Growth curve analysis of Δ*narB* revealed that it retained its ability to assimilate nitrate, to levels comparable with wild type (Figure 3.24) suggesting that NarB was not

responsible for or solely involved in the reduction of nitrate to nitrite. To determine whether the *narGHJI*-encoded NR played a role in nitrate assimilation in *M. smegmatis,* as is the case in *Mtb* (Malm *et al*., 2009); additional mutants were generated and assessed for growth in MPLN. Interestingly, the mutant's Δ*narGHJI* and Δ*narB* Δ*narGHJI* both retained the ability to grow in MPLN (Figure 3.26). In these experiments, the possibility that the growth observed was due to nutrient carry-over from growth of the pre-cultures in rich media was ruled out by repeated sub-culture in MPLN media. The continued growth of the NR-deficient mutant strains in MPLN media therefore suggested either that *M. smegmatis* possesses other NR enzymes or that there is an alternate nitrate assimilation pathway in *M. smegmatis* that does not rely on reduction of nitrate to nitrite. Analysis of nitrogen metabolism on KEGG Pathway Database reveals that across all orders of life, the reduction of nitrate to nitrite could be catalyzed by six possible enzymes, five of which are MoCo-dependent (Table 4.1). BLAST searches against the predicted *M. smegmatis* proteome show that this organism possesses possible homologues for all five of the MoCo-dependent enzymes, including NarB and NarGHI. This is in contrast to *Mtb* which possesses a single NR that is able to fulfill both assimilatory and respiratory functions, NarGHI (Malm *et al*., 2009) and one homologue which shares sequence similarity with an NADH nitrate reductase enzyme (Table 4.1). The function and activity of these additional enzymes in *Mtb* and *M. smegmatis* would need to be investigated to determine if they play a role in nitrate assimilation. Nitrate reduction and consequently nitrogen assimilation differ between *Mtb* and *M. smegmatis* and could be due to differences in the natural environments of the two organisms – *i.e*. soil vs. a mammalian host cell – which might differ significantly in terms of nitrogen source availability (Pashley *et al*., 2006; Lin *et al*., 2012).

Table 4.1: List of possible nitrate reduction catalyzing enzymes

Enzyme[§]	**MoCo-dependent**	***M. smegmatis***	***Mtb***
NarB	Yes	MSMEG_2837	-
NarG	Yes	MSMEG_5140	*Rv1161*
NADH nitrate reductase	Yes	MSMEG_4412	-
		MSMEG_2278	-
		MSMEG_0998	*Rv0218*
NADH nitrate oxidoreductase	Yes	MSMEG_4412	-
NADPH nitrate reductase	Yes	MSMEG_4412	-
		MSMEG_2278	-
Ferrocytochrome nitrate oxidoreductase	No	-	-

[§]According to KEGG Pathway Database (http://www.genome.jp/kegg/pathway.html)

4.1 Concluding remarks

Taken together, the results from this study provide insight into the complex MoCo biosynthetic pathway of *Mtb*, and particularly the multiple MPT synthase-encoding genes. Subsequent studies would need to address whether the findings reported here are relevant when MoaX is expressed in its natural *Mtb* host and moreover what role, if any, these multiple MPT-synthase-encoding genes might play in pathogenesis. The results obtained for MoaX provide preliminary evidence for cleavage of MoaX suggesting that it functions as a canonical MPT synthase. However, the observation that a significant amount of un-cleaved protein was also detected suggests that MoaX cleavage may be regulated in the cell. Possible mechanisms for this could include either the binding of an accessory protein that protects the cleavage site or regulation could be achieved through dynamic protein turnover, where the full length protein is produced at a faster rate than the proteolysis that results in cleavage. This study has provided a foundation for these and other future studies which will contribute to a greater understanding of the basic physiology and metabolism of *Mtb*.

5 Appendices

Appendix A- Bioinformatic tools

A 1. BLAST

(http://blast.ncbi.nlm.nih.gov/Blast.cgi)

Basic Local Alignment Search Tool (BLAST) allows for the comparison of nucleotide and protein sequences to those sequenced genomes contained in a database. This tool facilitates the identification of similar DNA regions and proteins among different organisms based on sequence alignments and aids in assigning characteristics to genes and proteins of unknown function.

A 2. Genolist

(http://genolist.pasteur.fr/TubercuList/)

Tuberculist is a database containing genome sequences of various mycobacterial organisms, importantly *Mtb* H37Rv. The database allows for the retrieval of gene and protein sequences and also provides links to functional information associated with the annotations.

A 3. KEGG Pathway Database

(http://www.genome.jp/kegg/pathway.html)

Kyoto Encyclopedia of Genes and Genomes (KEGG) is a database containing a collection of manually drawn pathway maps of several cellular processes from copious organisms. It aids in the global understanding of biological systems from the gene to the organisms environment. This tool allows for the study of specific metabolic pathways.

A 4. Sequence alignments

Sequence alignment tools allow for the alignment of protein or nucleotide sequences to identify homology from which structural and functional similarities or evolutionary relationships can be inferred.

- ClustalW2

(http://www.ebi.ac.uk/Tools/msa/clustalw2/)

ClustalW2, a program developed by the European Bioinformatics Institute is a multiple sequence alignment (MSA) tool used to identify similarities and/or differences among three or more protein or nucleotide sequences of the same length at a time. This tool allows for the identification of conserved residues within sequences and for evolutionary relationships to be studied between the sequences.

- Needle

(http://www.ebi.ac.uk/Tools/psa/emboss_needle/) - protein

(http://www.ebi.ac.uk/Tools/psa/emboss_needle/nucleotide.html) – nucleotide

Needle is a pairwise alignment tool that differs from ClustalW2 in that it only allows for the alignment of two sequences of any length at a time and identifies regions of similarity within the sequences which could point to structural, functional and/or evolutionary relationships shared.

Appendix B- Media and solution preparation

When required, all solutions and media were autoclaved at 121˚C for 10 mins.

Table B 1: Recipes of media used for bacterial growth

Medium	Components
LA	1% tryptone, 0.5% NaCl, 0.5% yeast extract dissolved in dH_2O (autoclaved)
LB	1% tryptone, 0.5% NaCl, 0.5% yeast extract, 1.5% agar dissolved in dH_2O (autoclaved)
7H9	4.9 g powder, Middlebrook oleic acid-albumin-dextrose-catalase (OADC), 0.2% glycerol, 0.05% Tween 80 (filter sterilized)
7H10	19 g powder, 0.085 % NaCl, 0.2% glucose, 0.2% glycerol made up to 1l in dH_2O (autoclaved)
Modified *M. phlei* (MPLN)	5 g KH_2PO_4, 2 g sodium citrate, 0.6 g $MgSO_4$, 0.85 g $NaNO_3$, 20 ml glycerol 5 ml tyloxapol, pH 6.6 (with 10 M NaOH) made up to 1l dH_2O (filter sterilized)
2xTY	2% tryptone, 0.5% NaCl, 1% yeast extract dissolved in dH_2O (autoclaved)

Table B 2: Recipes for media supplementation stocks

Supplement	Components
Glucose salts (100X)	20 g glucose, 8.5 g NaCl dissolved in 100 ml dH_2O (autoclaved)
Tween80 (25 %)	10 ml Tween80 dissolved in 40 mldH_2O (filter sterilized)
Sucrose (75%)	75 g sucrose in 100 ml dH_2O (autoclaved)
X-gal (2%)	1 g X-gal in 50 ml deionised DMF

Table B 3: Solutions used for preparation of chemically competent *E. coli* cells

Solution	Recipe
TfbI	30 mM Potassium acetate, 100 mM Rubidium chloride, 10 mM Calcium chloride, 50 mM Manganese chloride, 15% v/v Glycerol made up in dH_2O and pH 5.8 with dilute acetic acid
TfbII	10 mM MOPS, 75 mM Calcium chloride, 10 mM Rubidium chloride, 15 % v/v Glycerol made up in dH_2O and pH 6.5 with dilute NaOH

Table B 4: Solutions used for extraction of genomic DNA from *M. smegmatis*

Solution	Recipe
CTAB/NaCl	4.1 % NaCl, 10% N-cetyl-N,N, N-trimethyl ammonium bromidedissolved in dH_2O (filter sterilized)
TE buffer	10 mM Tris-HCl (pH 8), 10 mM EDTA dissolved in dH_2O (autoclaved)

Table B 5: Solutions used for plasmid extractions from *E. coli*

Solution	Recipe
Solution I	50mM Glucose, 25mM Tris-HCl (pH 8), 10 mM EDTA dissolved in dH_2O(autoclaved)
Solution II	1% SDS, 0.2 M NaOH dissolved in sdH_2O
Solution III	3 M Potassium acetate, 11.5% Acetic acid dissolved in sdH_2O

Table B 6: Solutions used for DNA precipitation

Solution	Recipe
Chloroform: Isoamyl alcohol	24 ml chloroform, 1 ml isoamyl alcohol
Phenol:chloroform	1 ml phenol, 1 ml chloroform
Sodium acetate	3M sodium acetate dissolved in dH_2O, pH 5.2 (autoclaved)

Table B 7: Solutions used for protein extractions

Solution	Description
B-PER Reagent	Proprietary mild, nonionic protein extraction detergent in 20mM Tris-HCl, pH 7.5 from Thermo Scientific
Protease inhibitor cocktail	1 protease inhibitor cOmplete ULTRA tablet (Roche) in 10ml B-PER solution

Table B 8: DNA electrophoresis solutions

Solution	Recipe
TAE	(50x stock solution): 242 g Tris base, 57.1 ml glacial acetic acid, 100 ml 0.5 M EDTA (pH 8) make up to 1L in dH_2O (1x working solution contains 40 mM Tris-acetate and 1 mM EDTA)
Ethidium bromide	10 mg/ml in sdH_2O

Table B 9: Recipe for agarose gels

Gel percentage	Amount of agarose in 50 ml TAE (g)
0.8	0.4
1	0.5
1.5	0.75

Table B 10: Protein electrophoresis solutions

Solution	Recipe/Description
bis-acrylamide	40 % solution
Loading buffer (5X)	3.8 ml dH_2O, 1 ml 0.5M Tris-HCl (pH 6.8), 0.8 ml glycerol, 1.6 ml 10% SDS, 0.4 ml β-mercaptoethanol, 0.4ml bromophenol blue
SDS (10%)	10 g SDS dissolved in 100 ml in dH_2O (autoclaved)
Tris-HCl	1.5 M (pH 8.8) and 0.5 M (pH 6.8) made up in dH_2O (autoclaved)
APS (10 %)	0.1 g Ammonium persulfate dissolved in 1 ml sdH_2O, stored at 4 °C
TEMED	N,N,N,N -Tetramethyl-Ethylenediamine
Saturated N-butanol	50 ml N-butanol, 5 ml dH_2O
Running buffer	303 g Tris, 144 g glycine, 10 g SDS made up to 1l dH_2O
Coomasie blue	0.1% Coomasie, 40% EtOH and 10% acetic acid made up indH_2O
De-stain	40% EtOH, 10% acetic acid made up in dH_2O

Table B 11: Recipe for two SDS-PAGE gels (10 ml)

Ingredient	Gel percentage	
	10 %	**15 %**
bis-acrylamide (40 %)	2.5 ml	3.75 ml
Tris-HCl pH 8.8	3.75 ml	5.625 ml
dH_2O	3.6 ml	400 ul
SDS (10%)	100 ul	150 ul
APS (10%)	50 ul	75 ul
TEMED	5 ul	7.5 ul

Table B 12: Southern blot solutions

Solution	Recipe/Description
Denaturation solution	0.5 M NaOH, 1.5 M NaCl in dH_2O
Depurination solution	0.25M HCl in dH_2O
TBE (5X)	Tris-Borate-EDTA powder (Sigma) dissolvedin 2l dH_2O
SSC (20X)	3M NaCl, 0.3M sodium citrate in dH_2O
Solution I	2X SSC, 0.1% SDS in dH_2O
Solution II	0.5X SSC, 0.1% SDS in dH_2O
Maleic acid buffer	1M Maleic acid, 1.5M NaCl in dH_2O (adjust to pH 7.5 with NaOH pellets)
Wash buffer	0.1M Maleic acid buffer, 0.3 % Tween20

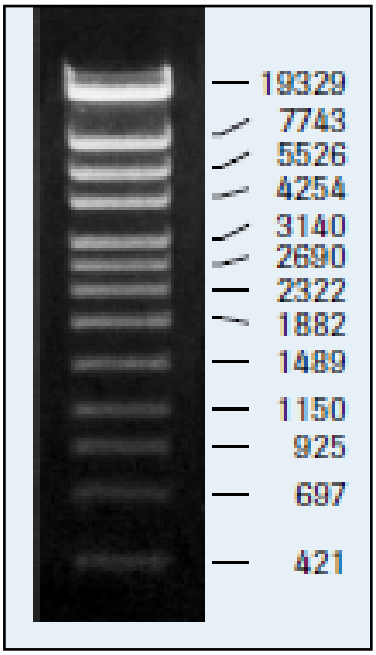
Blocking solution (Roche)	1X blocking solution in maleic acid buffer
Detection buffer	0.1M Tris-HCl, 0.1M NaCl in dH_2O (pH 9.5)
Antibody solution(Roche)	Dilute 1 in 10 000 in blocking solution
CSPD (Roche)	Disodium 2-chloro-5-(4-methoxyspiro (2-dioxetane-3,2 (2-dioxetane-3,2'-(5'-chloro)-tricyclo[3.3.1.1. 3, 7.]decan(-. 4-yl)-1-phenyl phosphate

Table B 13: Western blot solutions

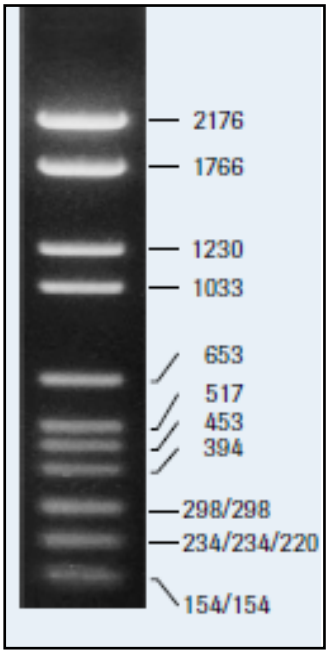
Solution	**Recipe/Description**
Transfer buffer	6 g Tris, 28.8 g glycine, 2 g SDS, 400 ml methanol made up to 2l with dH_2O
TBS (10X)	24.2 g Tris, 80 g NaCl in 1l dH_2O (pH 7.6)
TBST	1X TBS, 0.1 % Tween20 in dH_2O
Blocking solution	5% nonfat dry milk (CellSignal) in TBST
CPS Reagent (Sigma)	Chemiluminescent Peroxidase Substrate

Appendix C- Molecular weight markers

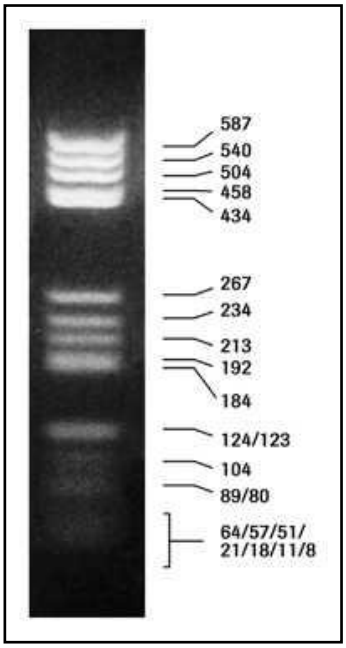
DNA molecular weight Marker IV

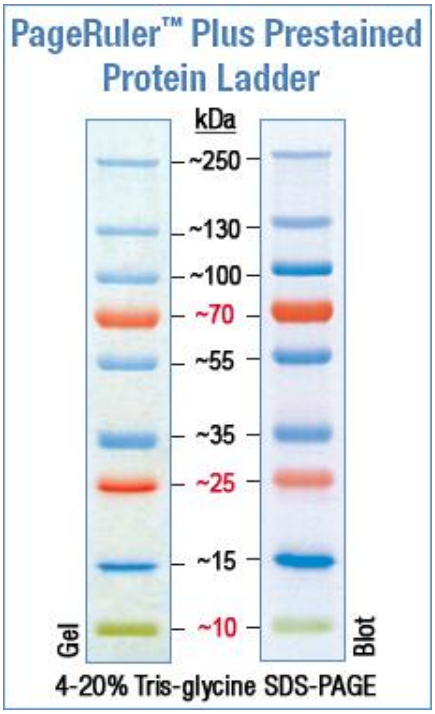


DNA molecular weight Marker VI



DNA molecular weight Marker V





Appendix D- Plasmids and primers

Table D 1: List of plasmids used and generated throughout this study

Name	Description	Source/ reference
Plasmids		
p2NIL	Cloning vector; Km^r	Parish *et al.*, 2000
pGOAL19	Plasmid carrying *hyg, lacZ, and sacB* genes as a *Pac*I cassette; Amp^r, Hyg^r	Parish *et al.*, 2000
pTTP1B	*E. coli-Mycobacterium* integrating shuttle vector Kan^r	Pham *et al.*, 2007
pHINT	*E. coli-Mycobacterium* integrating shuttle vector; Amp^r, Hyg^r	O'Gaora *et al.*, 1997
pTBD1	Derivative of pMhsp60 carrying *Mtb moaD1* expressed under control of the *hsp60* promoter; Hyg^r	Williams *et al.*, 2011
pTBD2	Derivative of pMhsp60 carrying *Mtb moaD2* expressed under control of the *hsp60* promoter; Hyg^r	Williams *et al.*, 2011
pTBE1	Derivative of pMhsp60 carrying *Mtb moaE1* expressed under control of the *hsp60* promoter; Hyg^r	Williams *et al.*, 2011
pTBE2	Derivative of pMhsp60 carrying *Mtb moaE2* expressed under control of the *hsp60* promoter; Hyg^r	Williams *et al.*, 2011
pHD1	Derivative of pHINT carrying *Mtb moaD1* expressed under control of the *hsp60* promoter; Hyg^r	This work
pHD2	Derivative of pHINT carrying *Mtb moaD2* expressed under control of the *hsp60* promoter; Hyg^r	This work
pTE1	Derivative of pTT1B carrying *Mtb moaE1* expressed under control of the *hsp60* promoter; Kan^r	This work
pTE2	Derivative of pTT1B carrying *Mtb moaE2* expressed under control of the *hsp60* promoter; Kan^r	This work
pMhsp60D1E1	Derivative of pMhsp60 carrying *Mtb moaD1*and *moaE1* expressed as an operon under control of the *hsp60* promoter; Hyg^r	This work
pMhsp60D1E2	Derivative of pMhsp60 carrying *Mtb moaD1*and *moaE2* expressed as an operon under control of the *hsp60* promoter; Hyg^r	This work
pMhsp60D2E1	Derivative of pMhsp60 carrying *Mtb moaD2* and *moaE1* expressed as an operon	This work

	under control of the *hsp60* promoter; Hyg^r	
pMhsp60D2E2	Derivative of pMhsp60 carrying *Mtb moaD2* and *moaE2* expressed as an operon under control of the *hsp60* promoter; Hyg^r	This work
pΔnarB	Knock-out vector for creating unmarked deletion in *M. smegmatis narB,* constructed by cloning PCR-amplified upstream and downstream regions of *narB* in p2NIL and insertion of the *hyg-lacZ-sacB* cassette from pGOAL19; Kan^r Hyg^r	This work
pΔnarGHJI	Knock-out vector for creating unmarked deletion in *M. smegmatis narGHJI* operon, constructed by cloning PCR-amplified upstream and downstream regions of *narGHJI* operon in p2NIL and insertion of the *hyg-lacZ-sacB* cassette from pGOAL19; Kan^r Hyg^r	This work
pMC1s	L5-based integration vector carrying P_{smyc}-*tetR*; Kan^r	Ehrt *et al*., 2005
pFLAGEM	*E. coli-Mycobacterium* episomal shuttle vector carrying the 3X FLAG epitope sequence and the Tet-operator; Hyg^r	Dr Edith Machowscki
pFLAGmoaXN	Derivative of pFLAGEM carrying *Mtb moaX* with the 3X FLAG sequence on the N-terminus under the control of the Tet-operator	This work
pFLAGmoaXC	Derivative of pFLAGEM carrying *Mtb moaX* with the 3X FLAG sequence on the C-terminus under the control of the Tet-operator	This work
pFLAGga1C	Derivative of pFLAGEM carrying *Mtb moaX* with a point mutation at position 242 (g→c) and the 3X FLAG sequence on the C-terminus under the control of the Tet-operator	This work
pFLAGga2C	Derivative of pFLAGEM carrying *Mtb moaX* with a point mutation at position 245 (g→c) and the 3X FLAG sequence on the C-terminus under the control of the Tet-operator	This work

Table D 2: Primers used to assess site specific intergration of L5-based vectors, pHINT and pMC1s

Primer name	Sequence 5'-3'	Amplicon
attBS2	ACAGGATTTGAACCTGCGGC	320 bp
attL4	AATTCTTGCAGACCCCTGGA	
attL2	CTTGGATCCTCCCGCTGCGC	282 bp
attBS1	ACGTGGCGGTCCCTACCG	

Table D 3: List of primers used for screening and confirmation of *M. smegmatis* complemented strains carrying different *Mtb* genes

Gene	Primer name	Sequence 5'-3'	Amplicon	Position
MtbmoaD1	moaD1F moaD1R	TACTTCGGTGCCGTTCGT GGCGACCTCATCACCATC	204 bp	22-225 of *moaD1*
MtbmoaD2	moaD2F moaD2R	GCCGGAATTCAGGTGACTG CGAAAGGGGGTAGTACGTCA	244 bp	25-268 of *moaD2*
MtbmoaE1	moaE1F moaE1R	CTGAGTGTGGACGAAGTGCT GTCTATCGCCGACCCATTC	381 bp	58-439 of *moaE1*
MtbmoaE2	moaE2F moaE2R	GATCTTTCTGGCCGAGCAC AACCGAACCCACCCATTC	385 bp	39-423 of *moaE2*
MtbmoaX	XscreenF XscreenR	GGCATAGGCGAGTGCTAAGA CGGCACATCCTGTTTGAG	848 bp	5487-6315 of pTX vector with XscreenR covering positions 592-602 of *moaX*

Table D 4: Primers used to amplify upstream and downstream regions of *narB* and *narGHJI* for the generation of knock out mutants

	Primer name	Sequence 5'-3'*	Amplicon	Position
narB downstream region	narB down Fwd	GGC GCG CTG CAG GCC TGA TCC CAC TGC TTC T (*Pst*I)	1695 bp	From position 2382 of *narB* to +1689 from stop of *narB*
	narB down Rev	GGC GCG AGA TCT CTC TGA GAG GGC CGA TCA T (*Bgl*II)		
narB upstream region	narB up Fwd	GGC GCG CAG ATC TGG TCT GTG CGA GCC ATG AT (*Bgl*II)	1908 bp	-1892 from start of *narB* to 16 of *narB*
	narB up Rev	GGC GCG AAG CTT GGG GTA CAA GCT TGA GGA CA (*Hind*III)		
narGHJI upstream region	narUPF	GCCG AAGCTTGGACTCTACGACGTGCTCAG (*Hind*III)	1672 bp	-1657 from start of *narG* to 14 of *narG*
	narUPR	GCCG AGATCTCAGCAGTTCTTCCACACGTC (*Bgl*II)		
narGHJI downsream region	narDF	GCCG AGATCTCGGCTGGTGACAAGAAGG (*Bgl*II)	1119 bp	From position 722 of *narI* to +1109 from stop of *narI*
	narDR	GCCG CTGCAGGTGATTCTCGCAGGTAGTCGAG (*Pst*I)		

*****Restriction sites underlined with restriction enzymes shown in paranthesis.

Table D 5: Primers used for PCR screening of Δ*narGHJI* mutants

Primer name	Sequence 5'-3'
narScreenF*	GGACGTGTGGAAGAACTGCT
narScreen out R$	GATCCGCACGAAATGGTC
narScreenG R∞	GTAGTCGGTCTCCTGGGTCTC

*Primer 1 in Figure 3.47A, $Primer 2 in Figure 3.47A, ∞Primer 3 in Figure 3.47A

Table D 6: Primers used for PCR screening of Δ*narB* mutants

Primer name	Sequence 5'-3'
narBScreen out*	GGTCATGATCGGCCCTCT
narBScreenF$	GATGCGTCCGTCCTTGAC
narBScreenR∞	TCGTAGCTCAGTGGGAGAGC

*Primer 1 in Figure 3.44A, $Primer 2 in Figure 3.44A, ∞Primer 3 in Figure 3.44A

Appendix E- Generation and restriction confirmation of vectors

E 1. Restriction analyses of integrating vectors

A

Enzyme	Expected fragment sizes (bp)
*Aat*II	2553, 3096
*Eco*RI	565, 1174, 1683, 2227
*Pst*I	374, 1320, 1793, 2216
*Not*I	4091, 1558
*Bam*HI	Linear
*Nae*I	3690, 817, 474, 486, 181

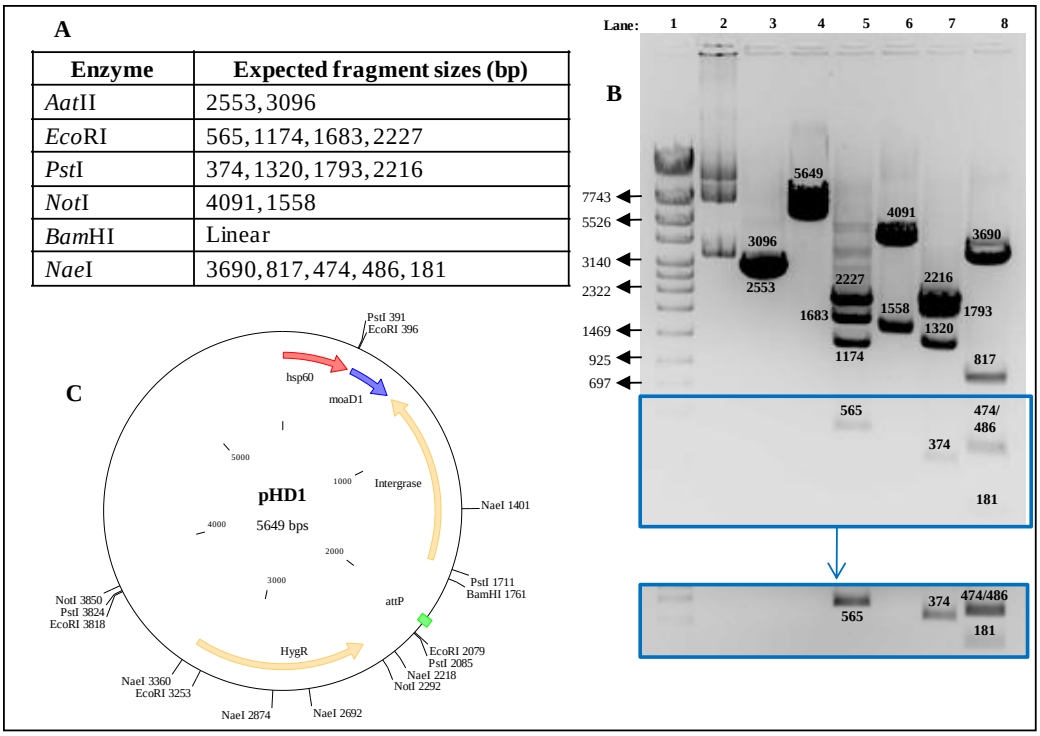


Figure E 1: Restriction analysis of integrating vector, pHD1 carrying *Mtb moaD1* driven off the constitutive hsp60 promoter. (A) Expected fragment sizes for restriction digest. (B) Restriction digests of pHD1 with everal enzymes. Lane 1: Marker λIV, Lane 2: Uncut pHD1, Lane 3: *Aat*II digest, Lane 4: *Bam*HI digest, Lane 5: *Eco*RI digest Lane 6: *Not*I digest, Lane 7: *Pst*I digest, Lane 8: *Nae*I digest (C) pHD1 vector map showing cloned *moaD1* in blue and the hsp60 promoter in red. The *integrase* and *hyg*^{R}genes are shown in yellow and the *attP* site in green.

The restriction patterns observed for each of the vectors in Figures E1B to E4B correspond to the sizes expected (Figures E1A to E4A) thus confirming that the integrating vectors were correct and could be used for electroporations.

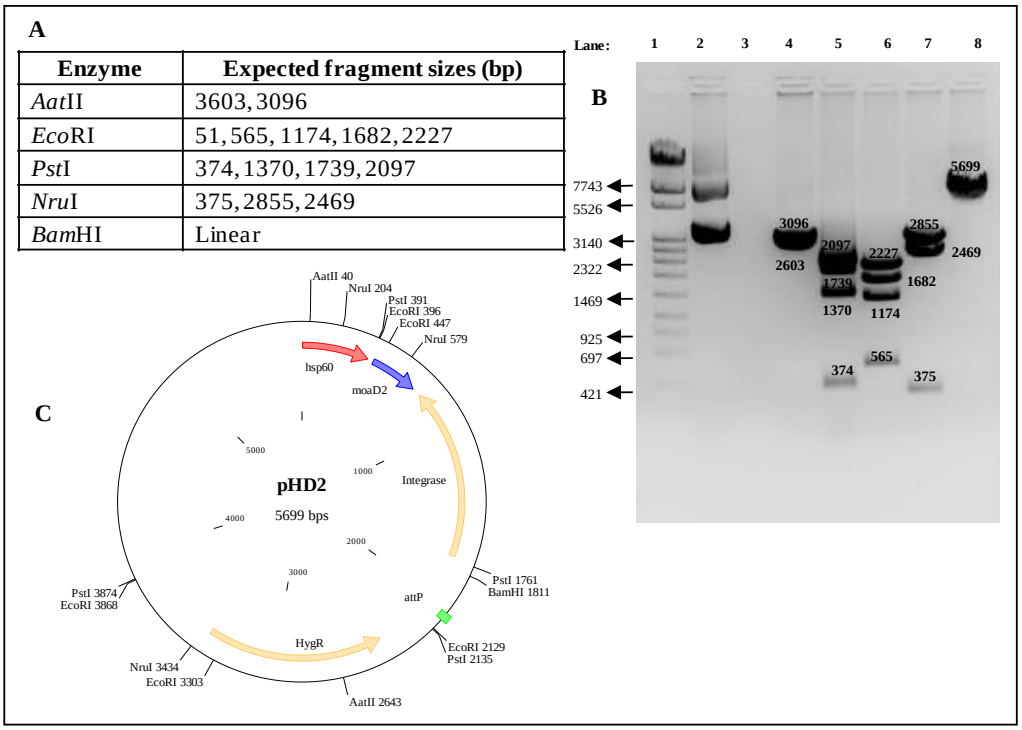


Figure E 2: Restriction analysis of integrating vector, pHD2 carrying *Mtb moaD2* driven off the constitutive hsp60 promoter. (A) Expected fragment sizes for restriction digest. (B) Restriction digests of pHD2 with several enzymes. Lane 1: Marker λIV, Lane 2: Uncut pHD2, Lane 3: Empty, Lane 4: *Aat*II digest, Lane 5: *Pst*I digest Lane 6: *Eco*RI digest, Lane 7: *Nru*I digest, Lane 8: *Bam*HI digest. (C) pHD2 vector map showing cloned *moaD2* in blue and the hsp60 promoter in red. The *integrase* and *hyg*^{R}genes are shown in yellow and the attP site in green.

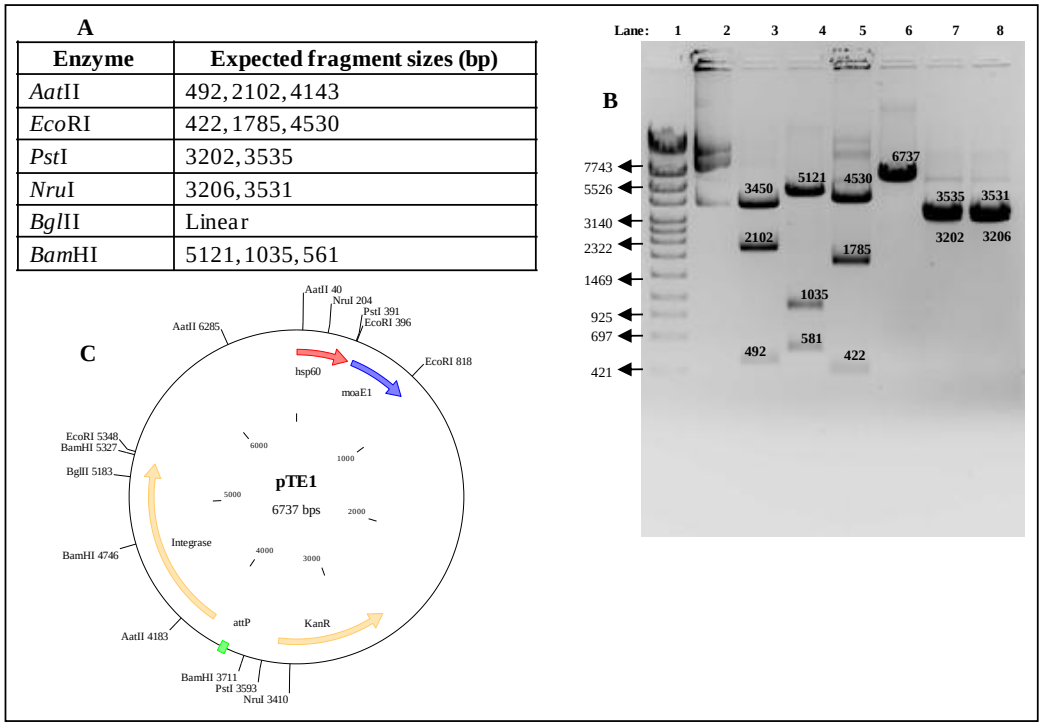


Figure E 3: Restriction analysis of integrating vector, pTE1 carrying *Mtb moaE1* driven off the constitutive hsp60 promoter. (A) Expected fragment sizes for restriction digest. (B) Restriction digests of pTE1 with everal enzymes. Lane 1: Marker λIV, Lane 2: Uncut pTE1, Lane 3: *Aat*II digest, Lane 4: *Bam*HI digest, Lane 5: *Eco*RI digest, Lane 6: *Bgl*II digest, Lane 7: *Pst*I digest, Lane 8: *Nru*I digest (C) pTE1 vector map showing cloned *moaE1* in blue and the hsp60 promoter in red. The *integrase* and *kan*^{R}genes are shown in yellow and the attP site in green.

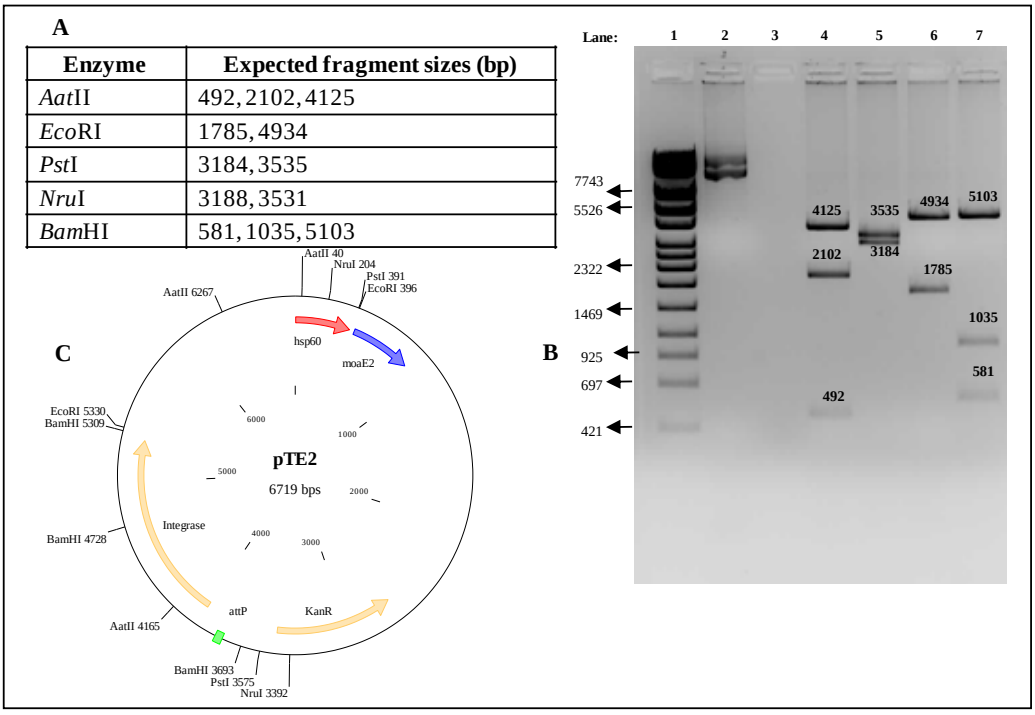


Figure E 4: Restriction analysis of integrating vector, pTE2 carrying *Mtb moaE2* driven off the constitutive hsp60 promoter. (A) Expected fragment sizes for restriction digest. (B) Restriction digests of pTE2 with several enzymes. Lane 1: Marker λIV, Lane 2: Uncut pTE2, Lane 3: Empty, Lane 4: *Aat*II digest, Lane 5: *Pst*I digest Lane 6: *Eco*RI digest, Lane 7: *Nru*I digest, Lane 8: *Bam*HI digest (C) pTE2 vector map showing cloned *moaE1* in blue and the hsp60 promoter in red. The *integrase* and *kan*^{R}genes are shown in yellow and the attP site in green.

E 2. Restriction mapping of pTmoaX

Restriction analysis was carried out in order to confirm the integrity of the vector carrying *moaX* and from the gel image in Figure E 5B it can be seen that the restriction patterns obtained for each digest correspond to the expected sizes (Figure E 5A) and the vector was therefore correct and suitable for use in the heterologous complementation assay.

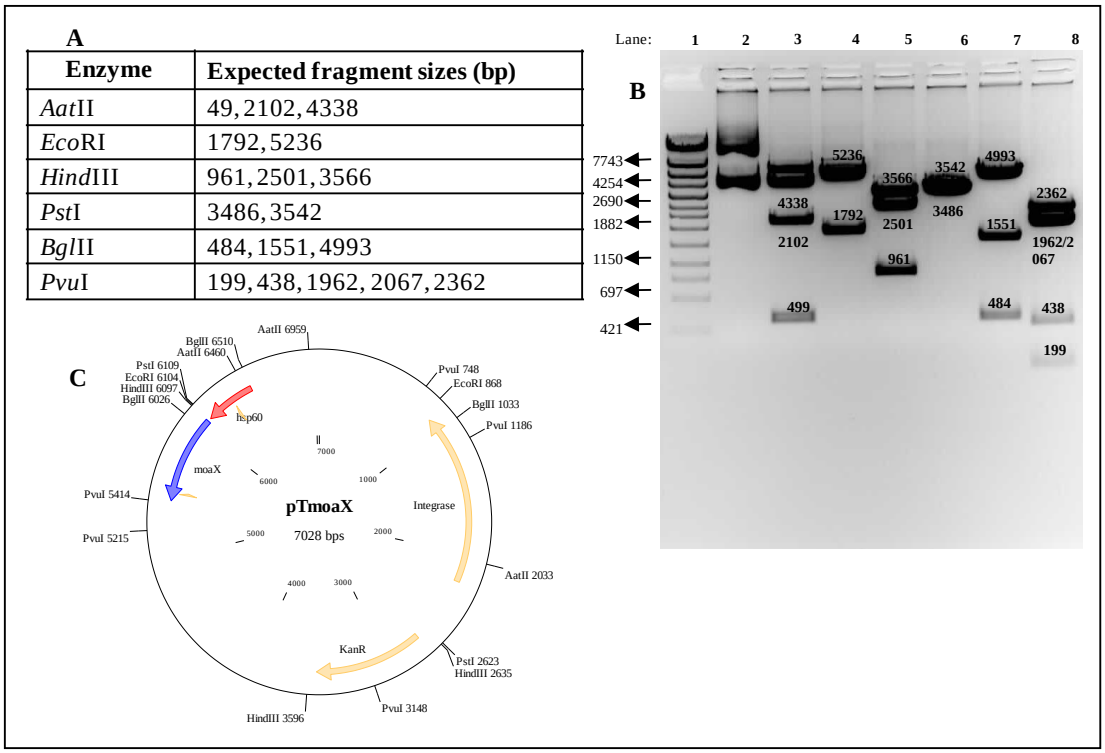


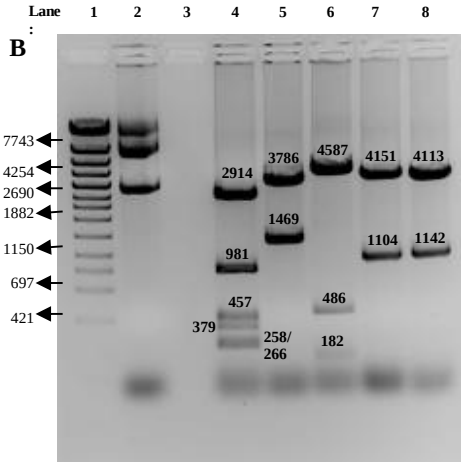
Figure E 5: Restriction analysis of integrating vector, pTmoaX carrying a single copy of *moaX* driven off the constitutive hsp60 promoter (A) Expected fragment sizes for restriction digests (B) Restriction digests of pTmoaX with several enzymes Lane 1: Marker λIV, Lane 2: Uncut pTX, Lane 3: *Aat*II digest, Lane 4: *Eco*RI digest, Lane 5: *Hind*III digest Lane 6: *Pst*I digest, Lane 7: *Bgl*II digest, Lane 8: *Pvu*I digest. (C) pTmoaX vector map showing cloned *moaX* in blue and the hsp60 promoter in red. The *integrase* and *kan*^{R} genes are shown in yellow.

E 3. Restriction analyses of episomal vectors

In addition to sequencing each of the constructed episomal vectors was validated by restriction digestion with five different restriction enzymes shown in Figures E 6 to E 9.

A

Enzyme	Expected fragment sizes (bp)
*Sac*II	258, 266, 379, 457, 981, 2914
*Aat*II	1469, 3786
*Nae*I	182, 486, 4587
*Eco*RI	1104, 4151
*Bgl*II	1142, 4113



C

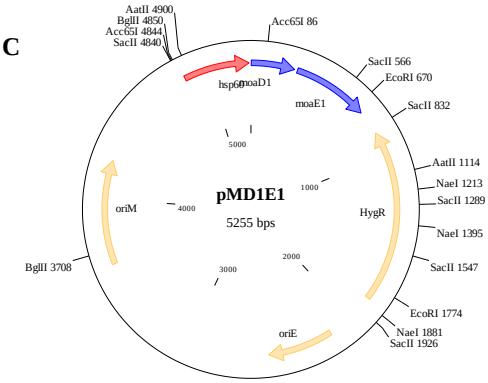
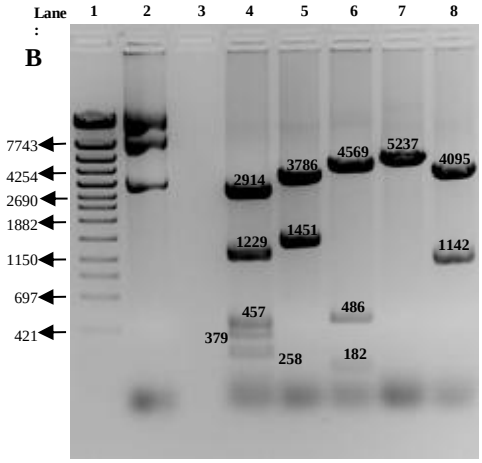


Figure E 6: Restriction analysis of episomal vector carrying *Mtb moaD1* and *moaE1* genes driven off the constitutive hsp60 promoter as an operon. (A) Expected fragment sizes for restriction digests (B) Restriction digests of pMD1E1 with several enzymes Lane 1: Marker λIV, Lane 2: Uncut pMD1E1, Lane 3: Empty, Lane 4: *SacII* digest, Lane 5: *AatII* digest Lane 6: *NaeI* digest, Lane 7: *EcoRI* digest, Lane 8: *BglII* digest. (C) pMD1E1 vector map showing cloned *moaD1* and *moaE1* in blue and the hsp60 promoter in red. The *hyg*^{R} gene, mycobacterial origin of replication (oriM) and *E. coli* origin of replication (oriE) are shown in yellow.

A

Enzyme	Expected fragment sizes (bp)
*Sac*II	258, 379, 457, 1229, 2914
*Aat*II	1451, 3786
*Nae*I	182, 486, 4569
*Eco*RI	5237
*Bgl*II	1142, 4095



C

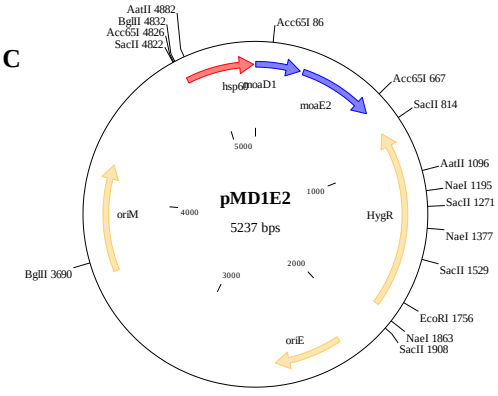
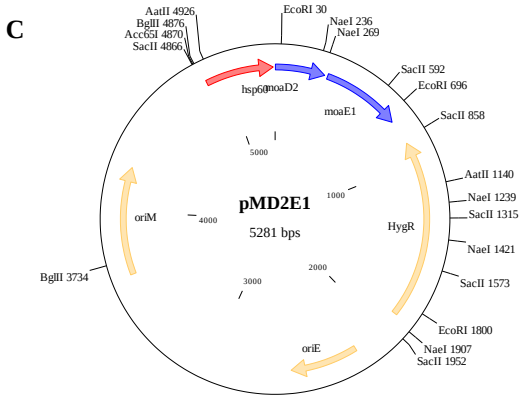


Figure E 7: Restriction analysis of episomal vector carrying *Mtb moaD1* and *moaE2* genes driven off the constitutive hsp60 promoter as an operon. (A) Expected fragment sizes for restriction digests (B) Restriction digests of pMD1E2 with several enzymes Lane 1: Marker λIV, Lane 2: Uncut pMD1E2, Lane 3: Empty, Lane 4: *SacII* digest, Lane 5: *AatII* digest Lane 6: *NaeI* digest, Lane 7: *EcoRI* digest, Lane 8: *BglII* digest. (C) pMD1E2 vector map showing cloned *moaD1* and *moaE2* in blue and the hsp60 promoter in red. The *hyg*^{R} gene, mycobacterial origin of replication (oriM) and *E. coli* origin of replication (oriE) are shown in yellow.

A

Enzyme	Expected fragment sizes (bp)
*Sac*II	258, 266, 379, 457, 1007, 2914
*Eco*RI	666, 1104, 3511
*Acc*651	5281
*Nae*I	486, 970, 3610
*Aat*II	1495, 3786



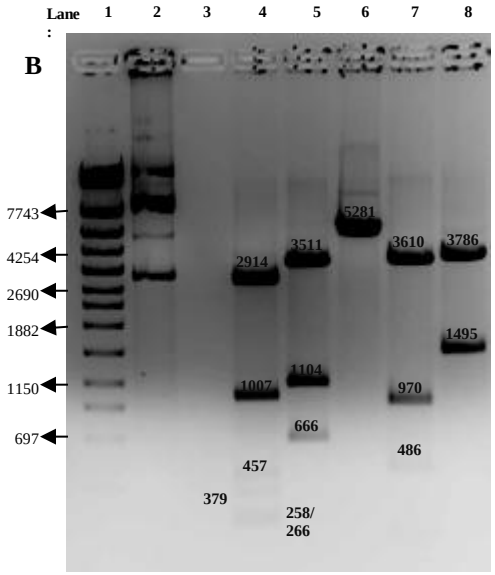
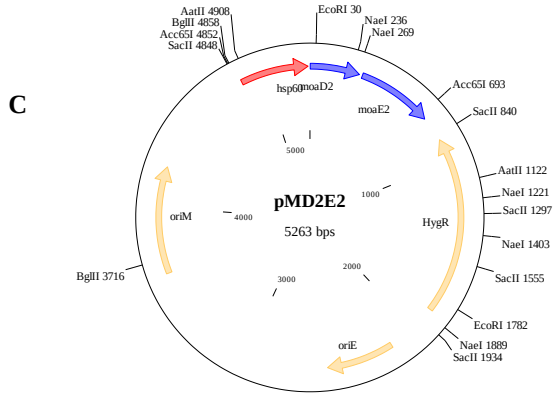


Figure E 9: Restriction analysis of episomal vector carrying *Mtb moaD2* and *moaE1* genes driven off the constitutive hsp60 promoter as an operon. (A) Expected fragment sizes for restriction digests (B) Restriction digests of pMD2E1 with several enzymes Lane 1: Marker λIV, Lane 2: Uncut pMD2E1, Lane 3: Empty, Lane 4: *SacII* digest, Lane 5: *EcoRI* digest Lane 6: *Acc651* digest, Lane 7: *NaeI* digest, Lane 8: *AatII* digest. (C) pMD2E1 vector map showing cloned *moaD2* and *moaE1* in blue and the hsp60 promoter in red. The *hyg*^{R} gene, mycobacterial origin of replication (oriM) and *E. coli* origin of replication (oriE) are shown in yellow.

A

Enzyme	Expected fragment sizes (bp)
*Sac*II	258, 379, 457, 1255, 2914
*Eco*RI	1752, 3511
*Acc*651	1104, 4159
*Nae*I	486, 952, 3610
*Aat*II	1477, 3786



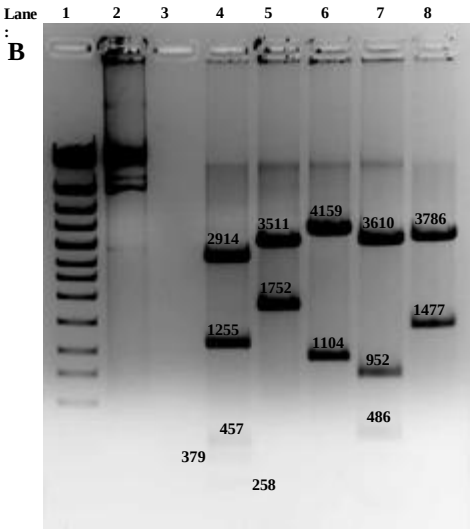


Figure E 8: Restriction analysis of episomal vector carrying *Mtb moaD2* and *moaE2* genes driven off the constitutive hsp60 promoter as an operon. (A) Expected fragment sizes for restriction digests (B) Restriction digests of pMD2E2 with several enzymes Lane 1: Marker λIV, Lane 2: Uncut pMD2E2, Lane 3: Empty, Lane 4: *SacII* digest, Lane 5: *EcoRI* digest Lane 6: *Acc651* digest, Lane 7: *NaeI* digest, Lane 8: *AatII* digest. (C) pMD2E2 vector map showing cloned *moaD2* and *moaE2* in blue and the hsp60 promoter in red. The *hyg*^{R} gene, mycobacterial origin of replication (oriM) and *E. coli* origin of replication (oriE) are shown in yellow.

The restriction patterns observed for all the vectors in Figures E 6B to E 9B corresponded to the expected sizes in each case (Figure E 6A to E 9A), confirming that they were correct. Each of the vectors was also sequenced to ensure that no mutations had been introduced during the PCR amplification of the *moaD1* and *moaD2* homologues. The sequencing data provided further confirmation that the vectors were correct and could be tested in the complementation assay.

E 4. Construction of pFLAGEM vectors carrying *moaX*

In addition to sequencing, restriction analysis was also performed for each of the vectors (Figures E 10 and E 11). For pFLAGmoaXC, fragments from all the restriction digests corresponded with the expected sizes, Figure E 10B, confirming the integrity of the vector. Similarly, restriction analysis of pFLAGmoaXN yielded a fragment pattern that corresponded with the expected sizes, Figure E 11.

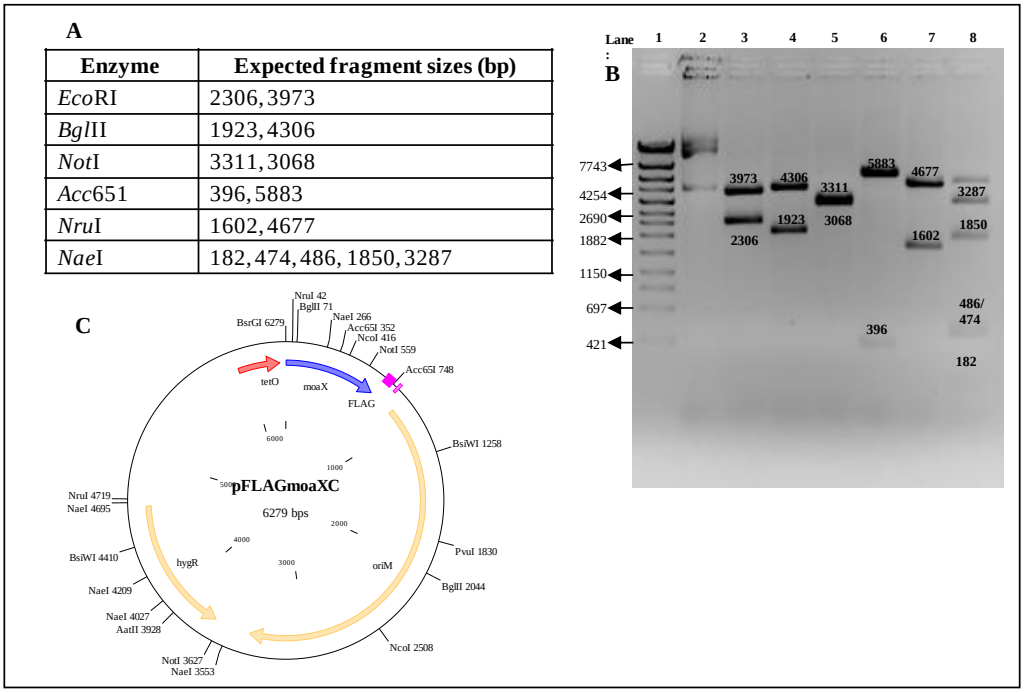
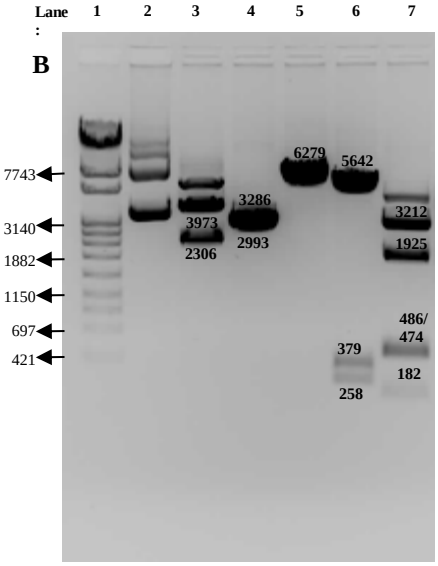


Figure E 10: Restriction analysis of pFLAG vector carrying C-terminally FLAG-tagged *Mtb moaX* under the control of the tet operator. (A) Expected fragment sizes for restriction digests (B) Restriction digests of pFLAGmoaXC with several enzymes Lane 1: Marker λIV, Lane 2: Uncut pMD2E2, Lane 3: *Eco*RI digest, Lane 4: *Bgl*II digest, Lane 5: *Not*I digest Lane 6: *Acc*651 digest, Lane 7: *Nru*I digest, Lane 8: *Nae*Idigest. Incomplete cleavage prodcuts are unlabeled (C) pFLAGmoaXC vector map showing cloned *moaX* in blue and the tet operator in red. The *hyg*^{R} gene and mycobacterial origin of replication (oriM) are shown in yellow

A

Enzyme	Expected fragment sizes (bp)
*Eco*RI	2306, 3973
*Not*I	2993, 3286
*Sac*II	258, 379, 5642
*Acc*651	6279
*Nae*I	182, 474, 486, 1925, 3212



C

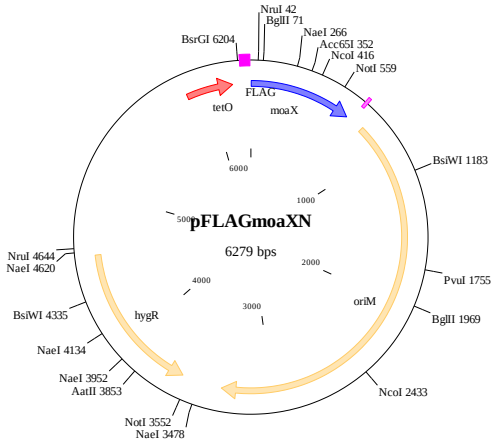


Figure E 11: Restriction analysis of pFLAG vector carrying N-terminally FLAG-tagged *Mtb moaX* under the control of the tet operator. (A) Expected fragment sizes for restriction digests (B) Restriction digests of pFLAGmoaXN with several enzymes Lane 1: Marker λIV, Lane 2: Uncut pMD2E2, Lane 3: *Eco*RI digest, Lane 4: *Not*I digest, Lane 5: *Acc651*digest Lane 6: *Sac*II digest, Lane 7: *Nae*I digest. Incomplete cleavage products are unlabeled (C) pFLAGmoaXN vector map showing cloned *moaX* in blue, the tet operator in red and the FLAG tag in pink. The *hyg*^{R} gene and mycobacterial origin of replication (oriM) are shown in yellow.

E 5. Construction of pFLAGEM vectors carrying mutated *moaX*

The site-directed mutagenesis strategy involved two rounds of PCR (Section 2.18). The first round generated the megaprimers carrying each mutation (Figure E 12).

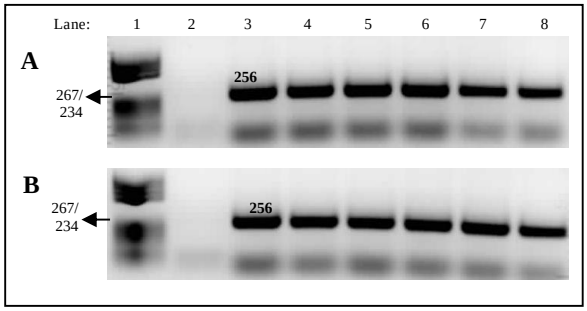


Figure E 12: Generation of megaprimers carrying point mutations to be incorporated into *moaX*. (A) Megaprimer carrying mutation 242G→C. (B) Megaprimer carrying mutation 245G→C. Lane 1: Marker λV, Lane 2: No DNA control, Lane 3: to Lane 8: Megaprimer amplicons

The 256 bp amplicons were then purified from the gel and used in another round of PCR reactions to generate the full length *moaX* with each mutation incorporated. During the second round of PCR reactions, the negative ("no DNA") control reaction always had an amplicon of the correct size (lane 3 of Figure E 13 A and B).

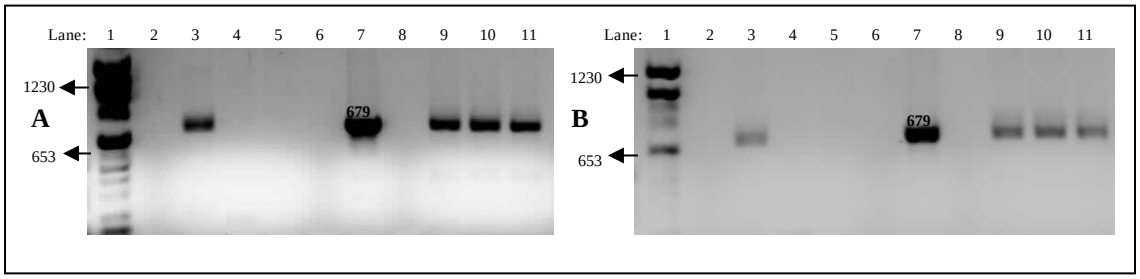


Figure E 13: Generation of full length *moaX* with point mutations (A) 242G→C and (B) 245G→C incorporated. Lane 1: Marker λVI, Lane 2: Empty, Lane 3: No DNA control, Lane 4: Forward primer only, Lane 5: Reverse primer only, Lane 6: Empty, Lane 7: Positive control, Lane 8: Empty, Lane 9 to Lane 11: Full length *moaX* amplicons with point mutations incorporated.

Several attempts were made to get rid of the amplicon being detected in the negative control, however the amplicon would still appear and it was concluded that its presence in the control reaction was attributed to the size of the megaprimer used, 256 bp. Its large size allows for the primer to fold back on itself and act as a template to generate the full length amplicon. In order to test this hypothesis, PCR reactions were performed using the amplicons generated above as template and the primer sets moaX-F+moaX-R, moaXga1F+moaX-R and moaXga2F+moaX-R. As seen in Figure E 14 the amplicon was only observed in the no DNA control when a megaprimer was used as a forward primer (Lanes 5 and 10) confirming the hypothesis. The products from Lanes 4 and 9 were thus used for subsequent experiments.

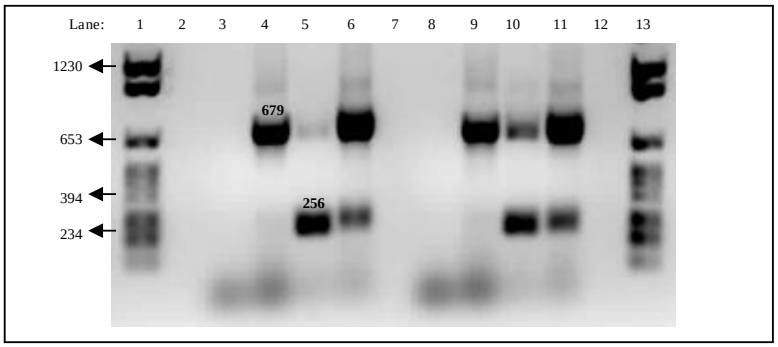


Figure E 14: Re-amplification of *moaX* carrying point mutations. Lane 1: Marker λVI, Lane 2: Empty, Lane 3: No DNA control for primer set moaX-F+moaX-R, Lane 4: *moaX* amplicon with 242G→C mutation incorporated using primer set moaX-F+moaX-R, Lane 5: No DNA control for primer set moaXga1F+moaX-R, Lane 6:*moaX* amplicon with 242G→C mutation incorporated using primer set moaXga1F+moaX-R, Lane 7: Empty, Lane 8: No DNA control for primer set moaX-F+moaX-R, Lane 9: *moaX* amplicon with 245G→C mutation incorporated using primer set moaX-F+moaX-R, Lane 10: No DNA control using primer set moaXga2F+moaX-R , Lane 11: *moaX* amplicon with 245G→C mutation incorporated using primer set moaXga2F+moaX-R, Lane 12: Empty, Lane 13: Marker λVI

Incorporation of each mutation into *moaX* introduced new restriction sites with 242G→C introducing a *Sac*II site and 245G→C introducing a *Hae*III site. These sites allowed for screening and confirmation of the incorporation of the point mutations into the gene sequence. The PCR products from Lanes 4, 6, 9 and 11 of Figure E 14 were digested with *Sac*II to assess whether the mutations had been incorporated (Figure E 15). As expected, the digested fragments which were re-amplified with either moaXga1F or moaXga2F and moaX-R show an extra 256 bp band corresponding to one of the megaprimers, seen in Lanes 4 and 6 of Figure E 15. The expected fragment sizes of 429 and 249 bp are observed in Lane 3 of Figure E 15 for the digested *moaX* with the 242G→C mutation incorporated and a single uncut band of 679 bp is observed for *moaX* carrying the 245G→C mutation, confirming the correct and specific incorporation of the 242G→C mutation.

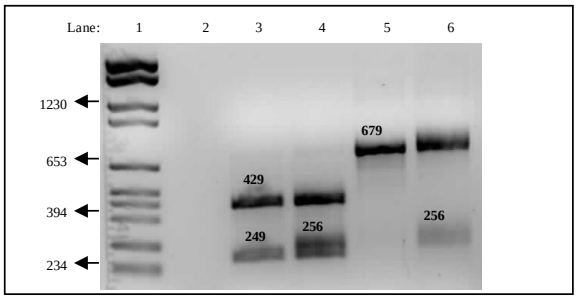


Figure E 15: *Sac*II screening of full length *moaX* with either 242G→C or 245G→C point mutations incorporated. Lane 1: Marker λVI, Lane 2: Empty, Lane 3: *Sac*II digest of *moaX* carrying mutation 242G>C amplified with primer set moaX-F+moaX-R, Lane 4: *Sac*II digest of *moaX* carrying mutation 242G>C amplified with primer set moaXga1F+moaX-R, Lane 5: *Sac*II digest of *moaX* carrying mutation 245G>C amplified with primer set moaX-F+moaX-R, Lane 6: *Sac*II digest of *moaX* carrying mutation 245G>C amplified with primer set moaXga2F+moaX-R

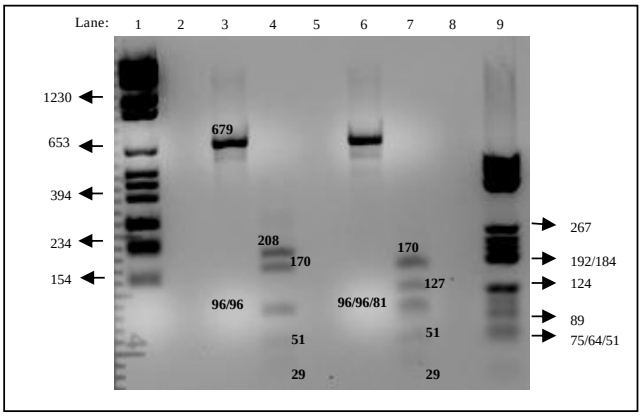


Figure E 16: Confirmation of the incorporation of point mutation 245G→C into *moaX*. Lane 1: Marker λVI, Lane 2: Empty, Lane 3: Uncut *moaX* carrying mutation 242G>C, Lane 4: *Hae*III digest of *moaX* carrying mutation 242G>C, Lane 5: Empty, Lane 6:Uncut *moaX* carrying mutation 245G>C, Lane 7: *Hae*III digest of *moaX* carrying mutation 245G>C, Lane 8: Empty, Lane 9: Marker λV

To confirm the integration of the 245G→C mutation the products from Lanes 4 and 9 from Figure E 14 were digested with *Hae*III. The expected restriction patterns are observed for both *moaX* fragments shown in Lanes 4 and 7 of Figure E 16. A clear difference can be seen between the *Hae*III restriction patterns of *moaX* 242G→C and *moaX* 245G→C confirming that the correct mutation was introduced into each.

Six restriction enzymes were then used for the mapping of each vector and the results from this restriction analysis are shown in Figures E 17 and E 18. The enzymes that gave complete coverage of the vectors and also allow for each point mutation to be identified (shown in bold in Figures E 17 and E 18A and boxed in black in Figures E 17 and E 18B) were chosen. Incomplete digestion was observed for the enzymes *Nar*I and *Nae*I shown in lanes 3 and 7 of Figures E 17B and E 18B respectively. However, the expected fragment sizes as listed in Figures E 17A and E 18A were present for all the digests thus confirming that all the vectors were correct and could be used further.

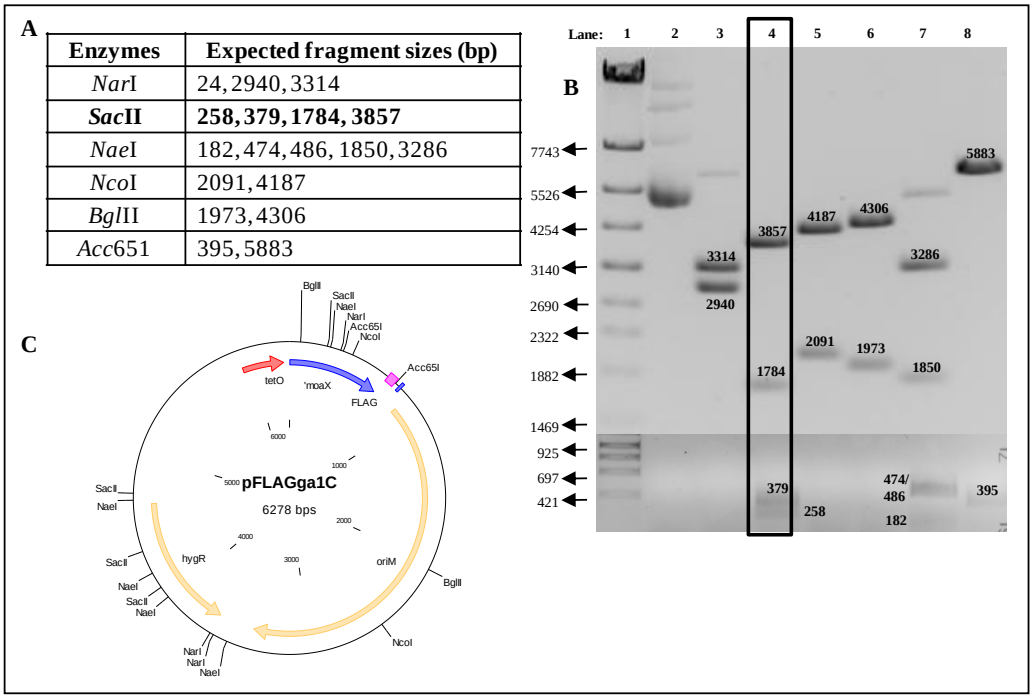


Figure E 17: Restriction mapping of pFLAGga1C carrying a C-terminally FLAG-tagged derivative of *moaX* with point mutation 242G>C. (A) Expected fragment sizes for restriction digests. (B) Restriction digests of pFLAGga1C with several enzymes Lane 1: Marker λIV, Lane 2: uncut vector, Lane 3: *Nar*I digest, Lane 4: *Sac*II digest, Lane 5: *Nco*I digest, Lane 6: *Bgl*II digest, Lane 7: *Nae*I digest, Lane 8: *Acc*651 digest. Incomplete cleavage products are unlabeld (C) pFLAGga1C vector map showing cloned *moaX* in blue, the tet operator in red and the FLAG tag in pink. The *hyg*^{R} gene and mycobacterial origin of replication (oriM) are shown in yellow.

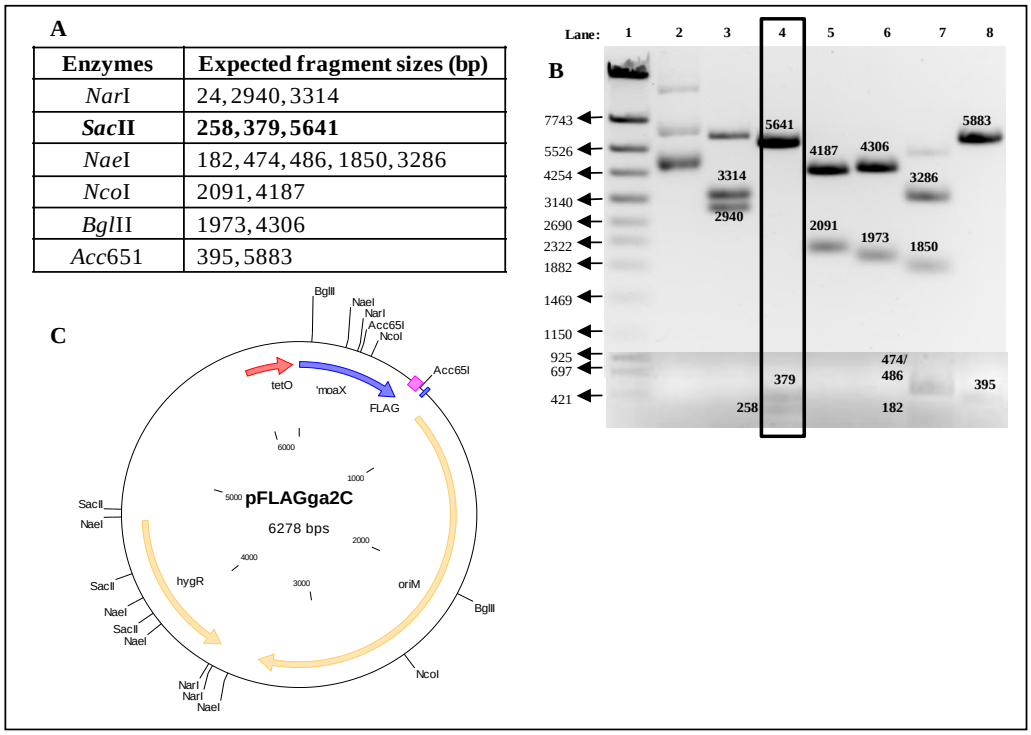


Figure E 18: Restriction mapping of pFLAGga2C carrying a C-terminally FLAG-tagged derivative of *moaX* with point mutation 245G>C. (A) Expected fragment sizes for restriction digests. (B) Restriction digests of pFLAGga2C with several enzymes Lane 1: Marker λIV, Lane 2: uncut vector, Lane 3: *Nar*I digest, Lane 4: *Sac*II digest, Lane 5: *Nco*I digest, Lane 6: *Bgl*II digest, Lane 7: *Nae*I digest, Lane 8: *Acc*651 digest. Incomplete cleavage products are unlabeled(C) pFLAGga2C vector map showing cloned *moaX* in blue, the tet operator in red and the FLAG tag in pink. The *hyg*^{R} gene and mycobacterial origin of replication (oriM) are shown in yellow.

The restriction patterns observed, in conjunction with the sequencing data which showed that no inadvertent, second-site mutations were introduced and that only the correct mutations were incorporated (Figure E 19), confirm the integrity of each vector. These vectors were thus introduced into the double mutant to investigate the effect of the mutations on MoaX activity and cleavage.

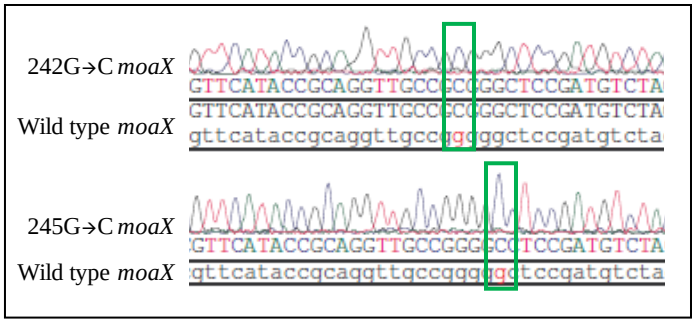


Figure E 19: Image of chromatogram showing the incorporation of the point mutations 242G→C and 245G→C into *moaX*. The wild type base pair is shown in red and the wild type together with the corresponding mutated base pair are boxed in green.

E 6. Construction of Δ*narB* suicide vector

The first step for the generation of a knock-out mutant involved construction of a suicide vector carrying a truncated version of the gene to be deleted. The *narB* suicide vector, pΔ*narB* was generated as described in section 2.20.1 and outlined in Figure 2.4. The upstream (US) and downstream (DS) regions flanking *narB* were amplified by PCR with the high fidelity DNA polymerase Phusion prior to being digested and incorporated into the p2NIL backbone by three-way directional cloning. One positive clone from this cloning was picked and analysed by restriction digest for confirmation that the vector was correct. Restriction digests were performed for vector DNA with two enzymes, *Sal*I and *Pac*I. The empty p2NILvector was included as a control to which the restriction pattern of the clone vector DNA could be compared (Figure E

20). The restriction patterns observed for p2nilnarB confirm that the clone is correct. In addition to restriction digestion, sequencing was performed for this vector in order to ensure that it was correct and that no unwanted mutations had been introduced into the upstream and downstream regions amplified during PCR. This vector, p2nilnarB, was then linearized with *Pac*I and ligated with the selectable marker cassette from pGOAL19 to yield the final knock-out construct (as depicted in Figure 2.4).

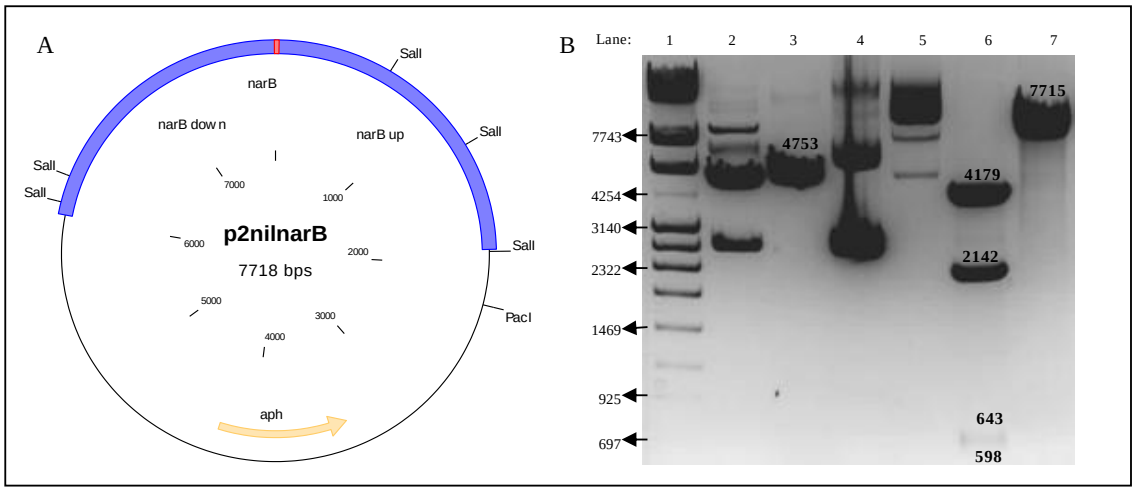


Figure E 20: Confirmation of p2nilnarB clone by restriction digestion. (A) p2nilnarB vector map showing restriction site positions. The upstream and downstream regions are shown in blue, deleted *narB* region in red and kanamycin resistance gene in yellow. (B) Restriction digests of p2nil and p2nilnarB. Lane 1: Marker λIV, Lane 2: p2nil uncut, Lane 3: p2nil *Pac*I digest, Lane 4: p2nil *Sal*I digest, Lane 5: p2nilnarB uncut, Lane 6: p2nilnarB *Sal*I digest, Lane 7: p2nilnarB *Pac*I digest.

Positive pΔnarB clones would be blue, Hyg^R, Kan^R and sucrose sensitive. Selection of transformants was therefore performed on LA plates with Kan (50 ng/ml), Hyg (100 ng/ml) and X-gal. Six blue colonies were picked to be screened by restriction digest with *Eco*RI (data not shown) and one positive clone was picked and confirmed by restriction digestion, Figure E 21, to be used for the generation of the knock-out mutant.

A

Enzymes	Expected fragment sizes (bp)
*Eco*RI	635, 747, 766, 1863, 4539, 7107
*Bam*HI	479, 1231, 1689, 3012, 3072, 6174
*Hind*III	27, 436, 739, 14455
*Pac*I	7718, 7939
*Sma*I	414, 1660, 1906, 5448, 6229
*Bgl*II	5506, 10151

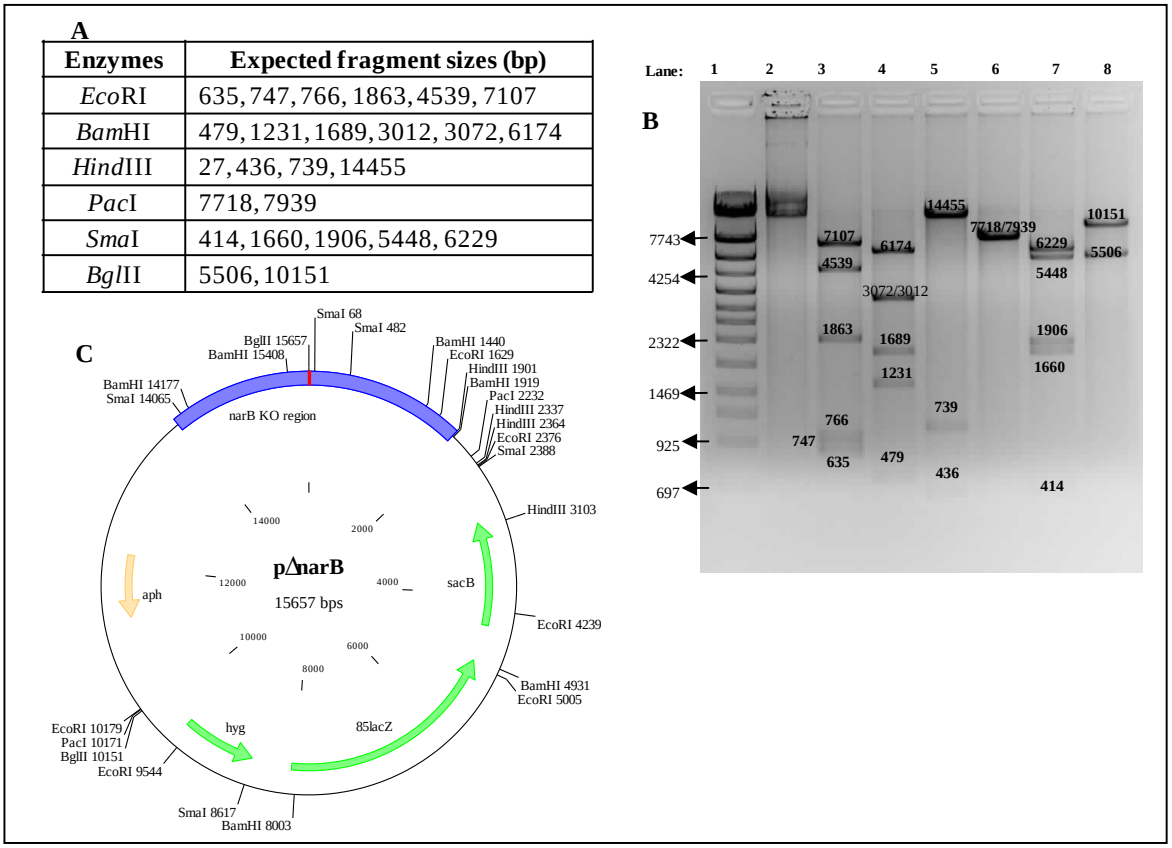


Figure E 21: Restriction digest confirmation of pΔnarB. (A) Expected fragment sizes for restriction digest. (B) Restriction digests of pΔnarB with several enzymes. Lane 1: Marker λIV, Lane 2: Uncut pΔnarB, Lane 3: *Eco*RI digest, Lane 4: *Bam*HI digest, Lane 5: *Hind*III digest, Lane 6: *Pac*I digest, Lane 7: *Sma*I digest, Lane 8: *Bgl*II digest (C) pΔnarB vector map showing upstream and downstream regionsin blue and the deleted *narB* allele in red. The selectable marker genes are shown in green and *kan*^{R} gene is shown in yellow.

The restriction patterns observed for each of the enzymes used was as expected and can clearly be seen in the gel image in Figure E 21B.

E 7. Generation of Δ*narGHJI* suicide vector

The suicide vector pΔ*narGHJI* was generated in the same manner as for pΔ*narB* described in section 2.20.1 and is summarized in Figure 2.5. Once constructed the suicide vector integrity was confirmed by restriction digest analysis, shown in Figure E 22. The fragments observed for each digest on the gel in Figure E 22B correspond to those expected shown in Figure E 22A. This, in

addition to the sequencing performed for p2nilnarGHJI confirms that the vector can be used to generate the knock-out mutant.

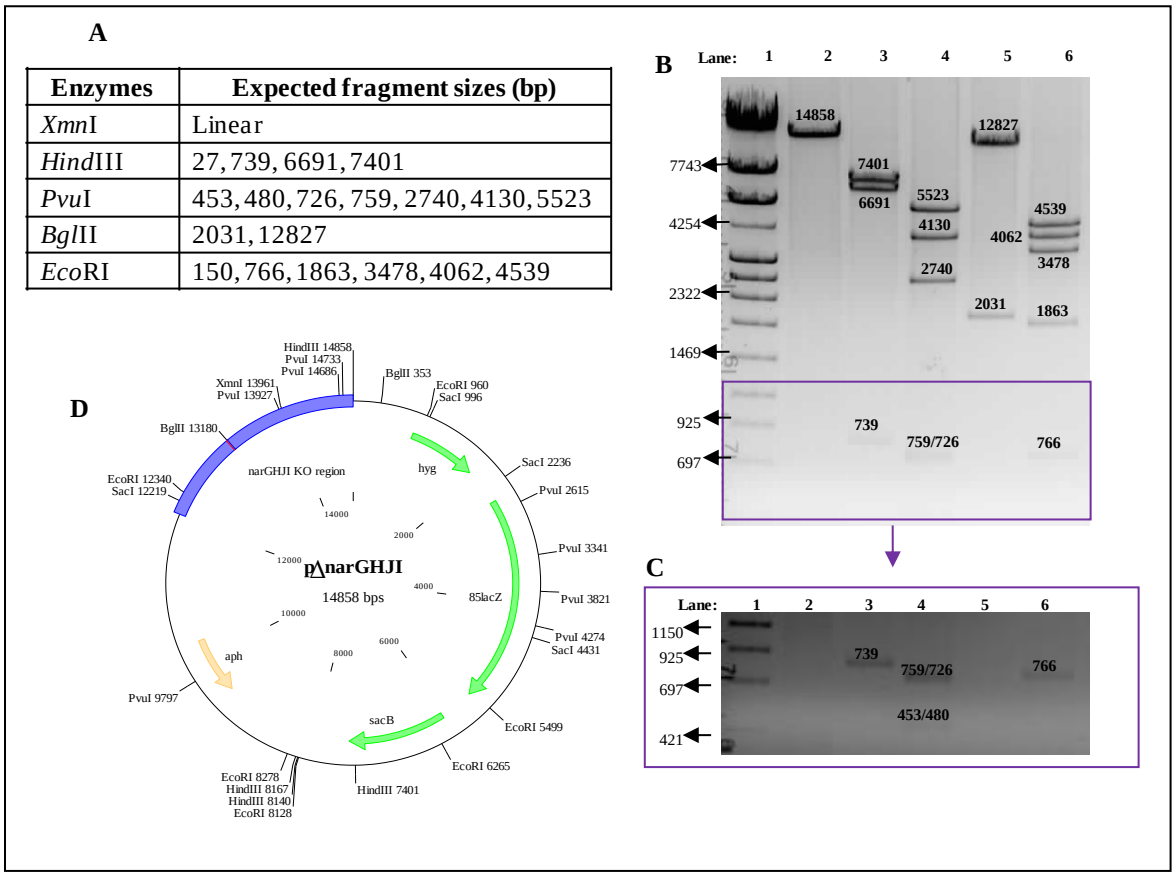


Figure E 22: Confirmation of suicide vector pΔ*narGHJI* by restriction digestion. (A) Expected fragment sizes for restriction digest. (B) Restriction digests of pΔ*narGHJI* with several enzymes. Lane 1: Marker λIV, Lane 2: *Xmn*I digest, Lane 3: *Hind*III digest, Lane 4: *Pvu*I digest, Lane 5: *Bgl*II digest, Lane 6: *Eco*RI digest, (C) Zoomed in image of lower gel showing the smaller restriction fragments. (D) pΔ*narGHJI* vector map showing upstream and downstream regionsin blue and the deleted *narGHJI* region in red. The selectable marker genes are shown in green and *kan*^{R}gene is shown in yellow

Enzymes	Expected fragment sizes (bp)
*Xmn*I	Linear
*Hind*III	27, 739, 6691, 7401
*Pvu*I	453, 480, 726, 759, 2740, 4130, 5523
*Bgl*II	2031, 12827
*Eco*RI	150, 766, 1863, 3478, 4062, 4539

6 References

AHIDJO, B. A., KUHNERT, D., MCKENZIE, J. L., MACHOWSKI, E. E., GORDHAN, B. G., ARCUS, V., ABRAHAMS, G. L. & MIZRAHI, V. (2011) VapC toxins from *Mycobacterium tuberculosis* are ribonucleases that differentially inhibit growth and are neutralized by cognate VapB antitoxins. *PLoS ONE,* 6**,** e21738.

BAEK, S. H., LI, A. H. & SASSETTI, C. M. (2011) Metabolic regulation of mycobacterial growth and antibiotic sensitivity. *PLoS Biol,* 9**,** e1001065.

BARRY, C. E., 3RD, BOSHOFF, H. I., DARTOIS, V., DICK, T., EHRT, S., FLYNN, J., SCHNAPPINGER, D., WILKINSON, R. J. & YOUNG, D. (2009) The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. *Nat Rev Microbiol,* 7**,** 845-55.

BHATT, K. & SALGAME, P. (2007) Host innate immune response to *Mycobacterium tuberculosis*. *J Clin Immunol,* 27**,** 347-62.

BLOMGRAN, R., DESVIGNES, L., BRIKEN, V., ERNST, J. D. (2012) *Mycobacterium tuberculosis* inhibits neutrophil apoptosis, leading to delayed activation of naive CD4 T cells. *Cell Host Microbe,* 11.

BOSHOFF, H. I. & BARRY, C. E., 3RD (2005) Tuberculosis - metabolism and respiration in the absence of growth. *Nat Rev Microbiol,* 3**,** 70-80.

BRODIN, P., POQUET, Y., LEVILLAIN, F., PEGUILLET, I., LARROUY-MAUMUS, G., GILLERON, M., EWANN, F., CHRISTOPHE, T., FENISTEIN, D., JANG, J., JANG, M. S., PARK, S. J., RAUZIER, J., CARRALOT, J. P., SHRIMPTON, R., GENOVESIO, A., GONZALO-ASENSIO, J. A., PUZO, G., MARTIN, C., BROSCH, R., STEWART, G. R., GICQUEL, B. & NEYROLLES, O. (2010) High content phenotypic cell-based visual screen identifies *Mycobacterium tuberculosis* acyltrehalose-containing glycolipids involved in phagosome remodeling. *PLoS Pathog,* 6**,** e1001100.

CAMACHO, L. R., ENSERGUEIX, D., PEREZ, E., GICQUEL, B. & GUILHOT, C. (1999) Identification of a virulence gene cluster of *Mycobacterium tuberculosis* by signature-tagged transposon mutagenesis. *Mol Microbiol,* 34**,** 257-67.

CORBETT, E. L., WATT, C. J., WALKER, N., MAHER, D., WILLIAMS, B. G., RAVIGLIONE, M. C. & DYE, C. (2003) The growing burden of tuberculosis: global trends and interactions with the HIV epidemic. *Arch Intern Med,* 163**,** 1009-21.

DAHL, J. U., URBAN, A., BOLTE, A., SRIYABHAYA, P., DONAHUE, J. L., NIMTZ, M., LARSON, T. J. & LEIMKUHLER, S. (2011) The identification of a novel protein involved in molybdenum cofactor biosynthesis in *Escherichia coli*. *J Biol Chem,* 286**,** 35801-12.

DARWIN, K. H. & NATHAN, C. F. (2005) Role for nucleotide excision repair in virulence of Mycobacterium tuberculosis. *Infect Immun,* 73**,** 4581-7.

DEL CAMPILLO-CAMPBELL, A. & CAMPBELL, A. (1982) Molybdenum cofactor requirement for biotin sulfoxide reduction in *Escherichia coli*. *J Bacteriol,* 149**,** 469-78.

DUTTA, N. K., MEHRA, S., DIDIER, P. J., ROY, C. J., DOYLE, L. A., ALVAREZ, X., RATTERREE, M., BE, N. A., LAMICHHANE, G., JAIN, S. K., LACEY, M. R., LACKNER, A. A. & KAUSHAL, D. (2010) Genetic requirements for the survival of tubercle bacilli in primates. *J Infect Dis,* 201**,** 1743-52.

EHRT, S., GUO, X. V., HICKEY, C. M., RYOU, M., MONTELEONE, M., RILEY, L. W. & SCHNAPPINGER, D. (2005) Controlling gene expression in mycobacteria with anhydrotetracycline and Tet repressor. *Nucleic Acids Res,* 33**,** e21.

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

FORD, C. B., LIN, P. L., CHASE, M. R., SHAH, R. R., IARTCHOUK, O., GALAGAN, J., MOHAIDEEN, N., IOERGER, T. R., SACCHETTINI, J. C., LIPSITCH, M., FLYNN, J. L. & FORTUNE, S. M. (2011) Use of whole genome sequencing to estimate the mutation rate of *Mycobacterium tuberculosis* during latent infection. *Nat Genet,* 43**,** 482-6.

GANDHI, N. R., MOLL, A., STURM, A. W., PAWINSKI, R., GOVENDER, T., LALLOO, U., ZELLER, K., ANDREWS, J. & FRIEDLAND, G. (2006) Extensively drug-resistant tuberculosis as a cause of death in patients co-infected with tuberculosis and HIV in a rural area of South Africa. *Lancet,* 368**,** 1575-80.

GARBE, T. R., BARATHI, J., BARNINI, S., ZHANG, Y., ABOU-ZEID, C., TANG, D., MUKHERJEE, R. & YOUNG, D. B. (1994) Transformation of mycobacterial species using hygromycin resistance as selectable marker. *Microbiology,* 140 (Pt 1)**,** 133-8.

GIEL, J. L., NESBIT, A. D., METTERT, E. L., FLEISCHHACKER, A. S., WANTA, B. T. & KILEY, P. J. (2012) Regulation of iron-sulphur cluster homeostasis through transcriptional control of the Isc pathway by [2Fe-2S]-IscR in *Escherichia coli*. *Mol Microbiol,* 87**,** 478-92.

GORDHAN, B. G. A. P., T (2001) Gene replacement using pre-treated DNA. . IN PARISH, T., STOKER, N. G (Ed.) *Methods Mol Med* Totowa, Human Press.

GRUNDEN, A. M., RAY, R. M., ROSENTEL, J. K., HEALY, F. G., SHANMUGAM, K. T. (1996) Repression of the *Escherichia coli modABCD* (Molybdate transport) operon by ModE. *Journal of Bacteriology* 178**,** 735-744.

GRUNDEN, A. M., SHANMUGAM, K. T. (1997) Molybdate transport and regulation in bacteria. *Arch Microbiol,* 168**,** 345-354.

GUTZKE, G., FISCHER, B., MENDEL, R. R. & SCHWARZ, G. (2001) Thiocarboxylation of molybdopterin synthase provides evidence for the mechanism of dithiolene formation in metal-binding pterins. *J Biol Chem,* 276**,** 36268-74.

HERNANDEZ, J. A., GEORGE, S. J. & RUBIO, L. M. (2009) Molybdenum trafficking for nitrogen fixation. *Biochemistry,* 48**,** 9711-21.

HESSELING, A. C., RABIE, H., MARAIS, B. J., MANDERS, M., LIPS, M., SCHAAF, H. S., GIE, R. P., COTTON, M. F., VAN HELDEN, P. D., WARREN, R. M. & BEYERS, N. (2006) Bacille Calmette-Guerin vaccine-induced disease in HIV-infected and HIV-uninfected children. *Clin Infect Dis,* 42**,** 548-58.

HILLE, R. (1996) The Mononuclear Molybdenum Enzymes. *Chem Rev,* 96**,** 2757-2816.

HILLE, R. (2002) Molybdenum and tungsten in biology. *Trends Biochem Sci,* 27**,** 360-7.

HOMOLKA, S., NIEMANN, S., RUSSELL, D. G. & ROHDE, K. H. (2010) Functional genetic diversity among *Mycobacterium tuberculosis* complex clinical isolates: delineation of conserved core and lineage-specific transcriptomes during intracellular survival. *PLoS Pathog,* 6**,** e1000988.

HOPP, T. P., PRICKETT, K. S., PRICE, V. L., LIBBY, R. T., MARCH, C. J., CERRETTI, D. P., URDAL, D. L AND CONLON, P. J (1988) A short polypeptide marker sequence useful for recombinant protein identification and purification. *Biotechnology,* 6**,** 7.

IOBBI-NIVOL, C., PALMER, T., WHITTY, P. W., MCNAIRN, E. & BOXER, D. H. (1995) The mob locus of *Escherichia coli* K12 required for molybdenum cofactor biosynthesis is expressed at very low levels. *Microbiology,* 141 (Pt 7)**,** 1663-71.

IOBBI-NIVOL, C. & LEIMKUHLER, S. (2012) Molybdenum enzymes, their maturation and molybdenum cofactor biosynthesis in *Escherichia coli*. *Biochim Biophys Acta*.

KHAN, A., SARKAR, D. (2006) Identification of a respiratory-type nitrate reductase and its role for survival of *Mycobacterium smegmatis* in Wayne model. *Microbial Pathogenesis,* 41**,** 90-95.

KHAN, A., AKHTAR, S., AHMAD, J. N., SARKAR, D. (2008) Presence of functional nitrate assimilation pathway in *Mycobacterium smegmatis*. *Microbial Pathogenesis,* 44**,** 71-77.

KING, G. M. (2003) Uptake of carbon monoxide and hydrogen at environmentally relevant concentrations by mycobacteria. *Appl Environ Microbiol,* 69**,** 7266-72.

LAKE, M. W., WUEBBENS, M. M., RAJAGOPALAN, K. V. & SCHINDELIN, H. (2001) Mechanism of ubiquitin activation revealed by the structure of a bacterial MoeB-MoaD complex. *Nature,* 414**,** 325-9.

LARSEN, M. H. (2000) Appendix 1. IN HATFUL, G. F., JACOBS, W. R., JR. (Ed.) *Molecular genetics of mycobacteria.* Washington, D.C., ASM Press.

LEIMKUHLER, S., WUEBBENS, M. M. & RAJAGOPALAN, K. V. (2001) Characterization of *Escherichia coli* MoeB and its involvement in the activation of molybdopterin synthase for the biosynthesis of the molybdenum cofactor. *J Biol Chem,* 276**,** 34695-701.

LEIMKUHLER, S., FREUER, A., ARAUJO, J. A., RAJAGOPALAN, K. V. & MENDEL, R. R. (2003) Mechanistic studies of human molybdopterin synthase reaction and characterization of mutants identified in group B patients of molybdenum cofactor deficiency. *J Biol Chem,* 278**,** 26127-34.

LIN, W., MATHYS, V., ANG, E. L., KOH, V. H., MARTINEZ GOMEZ, J. M., ANG, M. L., ZAINUL RAHIM, S. Z., TAN, M. P., PETHE, K. & ALONSO, S. (2012) Urease activity represents an alternative pathway for *Mycobacterium tuberculosis* nitrogen metabolism. *Infect Immun,* 80**,** 2771-9.

MACGURN, J. A. & COX, J. S. (2007) A genetic screen for *Mycobacterium tuberculosis* mutants defective for phagosome maturation arrest identifies components of the ESX-1 secretion system. *Infect Immun,* 75**,** 2668-78.

MACMICKING, J. D., TAYLOR, G. A. & MCKINNEY, J. D. (2003) Immune control of tuberculosis by IFN-gamma-inducible LRG-47. *Science,* 302**,** 654-9.

MALM, S., TIFFERT, Y., MICKLINGHOFF, J., SCHULTZE, S., JOOST, I., WEBER, I., HORST, S., ACKERMANN, B., SCHMIDT, M., WOHLLEBEN, W., EHLERS, S., GEFFERS, R., REUTHER, J. & BANGE, F. C. (2009) The roles of the nitrate reductase NarGHJI, the nitrite reductase NirBD and the response regulator GlnR in nitrate assimilation of *Mycobacterium tuberculosis*. *Microbiology,* 155**,** 1332-9.

MARQUET, A. (2001) Enzymology of carbon-sulfur bond formation. *Curr Opin Chem Biol,* 5**,** 541-9.

MATTHIES, A., RAJAGOPALAN, K. V., MENDEL, R. R. & LEIMKUHLER, S. (2004) Evidence for the physiological role of a rhodanese-like protein for the biosynthesis of the molybdenum cofactor in humans. *Proc Natl Acad Sci U S A,* 101**,** 5946-51.

MEHRA, S. & KAUSHAL, D. (2009) Functional genomics reveals extended roles of the *Mycobacterium tuberculosis* stress response factor sigmaH. *J Bacteriol,* 191**,** 3965-80.

MESTRE, O., HURTADO-ORTIZ, R., DOS VULTOS, T., NAMOUCHI, A., CIMINO, M., PIMENTEL, M., NEYROLLES, O. & GICQUEL, B. (2013) High Throughput Phenotypic Selection of *Mycobacterium tuberculosis* Mutants with Impaired Resistance to Reactive Oxygen Species Identifies Genes Important for Intracellular Growth. *PLoS ONE,* 8**,** e53486.

MILLER, J. L., VELMURUGAN, K., COWAN, M. J. & BRIKEN, V. (2010) The type I NADH dehydrogenase of *Mycobacterium tuberculosis* counters phagosomal NOX2 activity to inhibit TNF-alpha-mediated host cell apoptosis. *PLoS Pathog,* 6**,** e1000864.

NEUMANN, M. & LEIMKUHLER, S. (2008) Heavy metal ions inhibit molybdoenzyme activity by binding to the dithiolene moiety of molybdopterin in *Escherichia coli*. *FEBS J,* 275**,** 5678-89.

NEUMANN, M., MITTELSTADT, G., IOBBI-NIVOL, C., SAGGU, M., LENDZIAN, F., HILDEBRANDT, P. & LEIMKUHLER, S. (2009a) A periplasmic aldehyde oxidoreductase represents the first molybdopterin cytosine dinucleotide cofactor containing molybdo-flavoenzyme from *Escherichia coli*. *FEBS J,* 276**,** 2762-74.

NEUMANN, M., MITTELSTADT, G., SEDUK, F., IOBBI-NIVOL, C. & LEIMKUHLER, S. (2009b) MocA is a specific cytidylyltransferase involved in molybdopterin cytosine dinucleotide biosynthesis in *Escherichia coli*. *J Biol Chem,* 284**,** 21891-8.

NG, V. H., COX, J. S., SOUSA, A. O., MACMICKING, J. D. & MCKINNEY, J. D. (2004) Role of KatG catalase-peroxidase in mycobacterial pathogenesis: countering the phagocyte oxidative burst. *Mol Microbiol,* 52**,** 1291-302.

NICHOLS, J. D. & RAJAGOPALAN, K. V. (2005) *In vitro* molybdenum ligation to molybdopterin using purified components. *J Biol Chem,* 280**,** 7817-22.

O'GAORA, P., BERNINI, S., HAYWARD, C., FILLEY, E., ROOK, G, YOUNG, G. B AND THOLE, J (1997) Mycobacteria as immunogens. Development of expression vectors for use in multiple species. . *Med Princ Pract* 6**,** 91-96.

OH, J. I., PARK, S. J., SHIN, S. J., KO, I. J., HAN, S. J., PARK, S. W., SONG, T. & KIM, Y. M. (2010) Identification of *trans*- and *cis*-control elements involved in regulation of the carbon monoxide dehydrogenase genes in *Mycobacterium* sp. strain JC1 DSM 3803. *J Bacteriol,* 192**,** 3925-33.

PALMER, T., SANTINI, C. L., IOBBI-NIVOL, C., EAVES, D. J., BOXER, D. H. & GIORDANO, G. (1996) Involvement of the *narJ* and *mob* gene products in distinct steps in the biosynthesis of the molybdoenzyme nitrate reductase in *Escherichia coli*. *Mol Microbiol,* 20**,** 875-84.

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 (Pt 8)**,** 1969-75.

PARK, S. W., HWANG, E. H., PARK, H., KIM, J. A., HEO, J., LEE, K. H., SONG, T., KIM, E., RO, Y. T., KIM, S. W. & KIM, Y. M. (2003) Growth of mycobacteria on carbon monoxide and methanol. *J Bacteriol,* 185**,** 142-7.

PARK, S. W., SONG, T., KIM, S. Y., KIM, E., OH, J. I., EOM, C. Y. & KIM, Y. M. (2007) Carbon monoxide dehydrogenase in mycobacteria possesses a nitric oxide dehydrogenase activity. *Biochem Biophys Res Commun,* 362**,** 449-53.

PASHLEY, C. A., BROWN, A. C., ROBERTSON, D. & PARISH, T. (2006) Identification of the *Mycobacterium tuberculosis* GlnE promoter and its response to nitrogen availability. *Microbiology,* 152**,** 2727-34.

PHAM, T. T., JACOBS-SERA, D., PEDULLA, M. L., HENDRIX, R. W. & HATFULL, G. F. (2007) Comparative genomic analysis of mycobacteriophage Tweety: evolutionary insights and construction of compatible site-specific integration vectors for mycobacteria. *Microbiology,* 153**,** 2711-23.

PITTERLE, D. M. & RAJAGOPALAN, K. V. (1993) The biosynthesis of molybdopterin in *Escherichia coli*. Purification and characterization of the converting factor. *J Biol Chem,* 268**,** 13499-505.

ROBERTSON, B. D., ALTMANN, D., BARRY, C., BISHAI, B., COLE, S., DICK, T., DUNCAN, K., DYE, C., EHRT, S., ESMAIL, H., FLYNN, J., HAFNER, R., HANDLEY, G., HANEKOM, W., VAN HELDEN, P., KAPLAN, G., KAUFMANN, S. H., KIM, P., LIENHARDT, C., MIZRAHI, V., RUBIN, E., SCHNAPPINGER, D., SHERMAN, D., THOLE, J., VANDAL, O., WALZL, G., WARNER, D., WILKINSON, R. & YOUNG, D. (2012) Detection and treatment of subclinical tuberculosis. *Tuberculosis (Edinb),* 92**,** 447-52.

ROSAS-MAGALLANES, V., STADTHAGEN-GOMEZ, G., RAUZIER, J., BARREIRO, L. B., TAILLEUX, L., BOUDOU, F., GRIFFIN, R., NIGOU, J., JACKSON, M., GICQUEL, B. & NEYROLLES, O. (2007) Signature-tagged transposon mutagenesis identifies novel *Mycobacterium tuberculosis* genes involved in the parasitism of human macrophages. *Infect Immun,* 75**,** 504-7.

RUDOLPH, M. J., WUEBBENS, M. M., RAJAGOPALAN, K. V. & SCHINDELIN, H. (2001) Crystal structure of molybdopterin synthase and its evolutionary relationship to ubiquitin activation. *Nat Struct Biol,* 8**,** 42-6.

RUDOLPH, M. J., WUEBBENS, M. M., TURQUE, O., RAJAGOPALAN, K. V. & SCHINDELIN, H. (2003) Structural studies of molybdopterin synthase provide insights into its catalytic mechanism. *J Biol Chem,* 278**,** 14514-22.

SAMBROOK, J., FRITSCH, E. F., MANIATIS, T. (1989) Molecular Cloning. A laboratory manual.

. Second ed. Cold Spring Harbour, New York., Cold Spring Harbour Laboratory Press.

SASSETTI, C. M., BOYD, D. H. & RUBIN, E. J. (2003) Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol,* 48**,** 77-84.

SCANGA, C. A. & FLYNN, J. L. (2010) Mycobacterial infections and the inflammatory seesaw. *Cell Host Microbe,* 7**,** 177-9.

SCHMITZ, J., WUEBBENS, M. M., RAJAGOPALAN, K. V. & LEIMKUHLER, S. (2007) Role of the C-terminal Gly-Gly motif of *Escherichia coli* MoaD, a molybdenum cofactor biosynthesis protein with a ubiquitin fold. *Biochemistry,* 46**,** 909-16.

SCHWARZ, G., MENDEL, R. R. & RIBBE, M. W. (2009) Molybdenum cofactors, enzymes and pathways. *Nature,* 460**,** 839-47.

SENARATNE, R. H., DE SILVA, A. D., WILLIAMS, S. J., MOUGOUS, J. D., READER, J. R., ZHANG, T., CHAN, S., SIDDERS, B., LEE, D. H., CHAN, J., BERTOZZI, C. R. & RILEY, L. W. (2006) 5'-Adenosinephosphosulphate reductase (CysH) protects *Mycobacterium tuberculosis* against free radicals during chronic infection phase in mice. *Mol Microbiol,* 59**,** 1744-53.

SHI, T., XIE, J. (2011) Molybdenum enzymes and molybdenum cofactor in mycobacteria. *Journal of Cellular Biochemistry,* 112**,** 8.

SHILOH, M. U., MANZANILLO, P. & COX, J. S. (2008) *Mycobacterium tuberculosis* senses host-derived carbon monoxide during macrophage infection. *Cell Host Microbe,* 3**,** 323-30.

SMITH, A. M. & KLUGMAN, K. P. (1997) "Megaprimer" method of PCR-based mutagenesis: the concentration of megaprimer is a critical factor. *Biotechniques,* 22**,** 438, 442.

SNAPPER, S. B., MELTON, R. E., MUSTAFA, S., KIESER, T. & JACOBS, W. R., JR. (1990) Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Mol Microbiol,* 4**,** 1911-9.

SOFIA, H. J., CHEN, G., HETZLER, B. G., REYES-SPINDOLA, J. F. & MILLER, N. E. (2001) Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. *Nucleic Acids Res,* 29**,** 1097-106.

SOHASKEY, C. D. & WAYNE, L. G. (2003) Role of *narK2X* and *narGHJI* in hypoxic upregulation of nitrate reduction by *Mycobacterium tuberculosis*. *J Bacteriol,* 185**,** 7247-56.

SOHASKEY, C. D. (2005) Regulation of nitrate reductase activity in *Mycobacterium tuberculosis* by oxygen and nitric oxide. *Microbiology,* 151**,** 3803-10.

SOHASKEY, C. D. & MODESTI, L. (2009) Differences in nitrate reduction between *Mycobacterium tuberculosis* and *Mycobacterium bovis* are due to differential expression of both *narGHJI* and *narK2*. *FEMS Microbiol Lett,* 290**,** 129-34.

TAMERIS, M. D., HATHERILL, M., LANDRY, B. S., SCRIBA, T. J., SNOWDEN, M. A., LOCKHART, S., SHEA, J. E., MCCLAIN, J. B., HUSSEY, G. D., HANEKOM, W. A., MAHOMED, H. & MCSHANE, H. (2013) Safety and efficacy of MVA85A, a new tuberculosis vaccine, in infants previously vaccinated with BCG: a randomised, placebo-controlled phase 2b trial. *Lancet*.

TAN, M. P., SEQUEIRA, P., LIN, W. W., PHONG, W. Y., CLIFF, P., NG, S. H., LEE, B. H., CAMACHO, L., SCHNAPPINGER, D., EHRT, S., DICK, T., PETHE, K. & ALONSO, S. (2010) Nitrate respiration protects hypoxic *Mycobacterium tuberculosis* against acid- and reactive nitrogen species stresses. *PLoS ONE,* 5**,** e13356.

TRUNZ, B. B., FINE, P. & DYE, C. (2006) Effect of BCG vaccination on childhood tuberculous meningitis and miliary tuberculosis worldwide: a meta-analysis and assessment of cost-effectiveness. *Lancet,* 367**,** 1173-80.

VELMURUGAN, K., CHEN, B., MILLER, J. L., AZOGUE, S., GURSES, S., HSU, T., GLICKMAN, M., JACOBS, W. R., JR., PORCELLI, S. A. & BRIKEN, V. (2007) *Mycobacterium tuberculosis nuoG* is a virulence gene that inhibits apoptosis of infected host cells. *PLoS Pathog,* 3**,** e110.

VOSKUIL, M. I., SCHNAPPINGER, D., VISCONTI, K. C., HARRELL, M. I., DOLGANOV, G. M., SHERMAN, D. R. & SCHOOLNIK, G. K. (2003) Inhibition of respiration by nitric oxide induces a *Mycobacterium tuberculosis* dormancy program. *J Exp Med,* 198**,** 705-13.

VOSS, M., NIMTZ, M. & LEIMKUHLER, S. (2011) Elucidation of the dual role of Mycobacterial MoeZR in molybdenum cofactor biosynthesis and cysteine biosynthesis. *PLoS ONE,* 6**,** e28170.

WARNER, D. F., MIZRAHI, V. (2008) *M. tuberculosis* physiology. IN KAUFMANN, S. H. E., RUBIN, E. J. (Ed.) *Handbook of tuberculosis.* Weinheim, Wiley-VCH.

WAYNE, L. G. (1994) Dormancy of *Mycobacterium tuberculosis* and latency of disease. *Eur J Clin Microbiol Infect Dis,* 13**,** 908-14.

WAYNE, L. G. & HAYES, L. G. (1998) Nitrate reduction as a marker for hypoxic shiftdown of *Mycobacterium tuberculosis*. *Tuber Lung Dis,* 79**,** 127-32.

WEBER, I., FRITZ, C., RUTTKOWSKI, S., KREFT, A. & BANGE, F. C. (2000) Anaerobic nitrate reductase (*narGHJI*) activity of *Mycobacterium bovis* BCG *in vitro* and its contribution to virulence in immunodeficient mice. *Mol Microbiol,* 35**,** 1017-25.

WHO (2004) BCG vaccine: WHO position paper. *Weekly Epidemiological Record.*

WHO (2006) Laboratory XDR-TB definitions.

WHO (2010) Treatment of tuberculosis: guidelines. Fourth edition. .

WHO (2012) Global Tuberculosis Report 2012.

WILLIAMS, M. J., KANA, B. D. & MIZRAHI, V. (2011) Functional analysis of molybdopterin biosynthesis in mycobacteria identifies a fused molybdopterin synthase in *Mycobacterium tuberculosis*. *J Bacteriol,* 193**,** 98-106.

WILLIAMS, M., MIZRAHI, V. & KANA, B. D. (2013) Molybdenum cofactor: A key component of *Mycobacterium tuberculosis* pathogenesis? *Crit Rev Microbiol*.

WOONG PARK, S., KLOTZSCHE, M., WILSON, D. J., BOSHOFF, H. I., EOH, H., MANJUNATHA, U., BLUMENTHAL, A., RHEE, K., BARRY, C. E., 3RD, ALDRICH, C. C., EHRT, S. & SCHNAPPINGER, D. (2011) Evaluating the sensitivity of *Mycobacterium tuberculosis* to biotin deprivation using regulated gene expression. *PLoS Pathog,* 7**,** e1002264.

WUEBBENS, M. M. & RAJAGOPALAN, K. V. (2003) Mechanistic and mutational studies of *Escherichia coli* molybdopterin synthase clarify the final step of molybdopterin biosynthesis. *J Biol Chem,* 278**,** 14523-32.

ZHANG, W., URBAN, A., MIHARA, H., LEIMKUHLER, S., KURIHARA, T. & ESAKI, N. (2010) IscS functions as a primary sulfur-donating enzyme by interacting specifically with MoeB and MoaD in the biosynthesis of molybdopterin in *Escherichia coli*. *J Biol Chem,* 285**,** 2302-8.