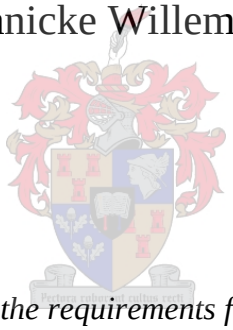
Regulation of iron-sulphur cluster biogenesis in *Mycobacterium tuberculosis*

by

Danicke Willemse



Dissertation presented in fulfilment of the requirements for the degree of Doctor of Philosophy (Molecular Biology) in the Faculty of Medicine and Health Sciences at Stellenbosch University

Supervisor: Dr Monique Joy Williams

Co-supervisor: Prof Robin Mark Warren

December 2016

Author's declaration

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

Danicke Willemse

Date: December 2016

Abstract

Iron-sulphur (Fe-S) clusters are protein cofactors that are important for the functioning of many proteins involved in diverse processes in *Mycobacterium tuberculosis*. Complex Fe-S cluster biogenesis systems are required for their synthesis, to protect the clusters from the deleterious effects of reactive oxygen species *in vivo*. The *suf* system is the primary Fe-S cluster biogenesis system in *M. tuberculosis* and the components are encoded in an operon consisting of seven genes (*Rv1460-Rv1461-Rv1462-Rv1463-csd-Rv1465-Rv1466*). The first gene in the operon, *Rv1460*, is predicted to encode a transcriptional regulator based on homology with the cyanobacterial *suf* operon repressor, SufR. This study aimed to determine whether Rv1460 is involved in the regulation of *suf* operon expression and *M. tuberculosis* physiology.

In order to address this knowledge gap, attempts were made to generate three distinct *Rv1460* deletion mutants in *M. tuberculosis* H37Rv. Surprisingly, generation of only one of the mutants, *Rv1460stop*, in which *Rv1460* is truncated by insertion of a premature stop codon, was successful. This suggests that *Rv1460* is not essential for the *in vitro* growth of *M. tuberculosis*. Our inability to generate the *ΔRv1460* and *Rv1460ΔDNAbd* mutants may be due to polar effects on the expression of downstream genes, which make these mutants non-viable. Analysis of the *Rv1460stop* mutant's growth revealed that *Rv1460* is required for normal growth on solid and in liquid media, under standard culture conditions. The *Rv1460stop* mutant was more sensitive to ROS, indicating the importance of Rv1460 in oxidative stress response, and potentially implicating it in intracellular survival and pathogenesis. The *Rv1460stop* mutant's growth was, surprisingly, not impaired under iron limitation.

Gene expression studies done on the wild-type, *Rv1460stop* mutant and complementation strain revealed that Rv1460 acts as a transcriptional repressor of itself and the rest of the *suf* operon, since transcript levels of both *Rv1460* and *Rv1461* increased in the *Rv1460stop* mutant. *Rv1460* was shown to be co-transcribed with the rest of the operon. Transcript levels, however, also suggested that *Rv1460* may be independently transcribed from the rest of the gene cluster. Electrophoretic mobility shift assays done with recombinant Rv1460 demonstrated binding of Rv1460 to the *Rv1460* promoter and within *Rv1461*. This indicates that Rv1460 mediates transcriptional control through direct interaction with the *suf* operon DNA.

Recombinant Rv1460 was shown to form dimers in solution and to coordinate an Fe-S cluster *in vitro*, which has important implications for its function as a regulator, because the affinity of Fe-S cluster containing regulators for DNA is often influenced by the presence and redox state of their cluster. The role of three conserved cysteine residues (C203, C216, C244), predicted to be involved in Fe-S cluster coordination in Rv1460, could not be confirmed, as replacing these residues with serine residues did not alter their ability to coordinate an Fe-S cluster. The regulation of the *suf* operon in *M. tuberculosis* is multi-faceted, probably because Fe-S cluster biogenesis needs to be fine-tuned to allow survival within the host. This study indicates that Rv1460 plays a key role in this regulation and in *M. tuberculosis* physiology.

Opsomming

Yster-swawel (Fe-S) groepe is protein ko-faktore wat belangrik is vir die funksionering van baie proteïne betrokke in ‘n wye verskeidenheid prosesse in *Mycobacterium tuberculosis*. Omdat dit sensitief is vir reaktiewe suurstof groepe in die sel, word komplekse Fe-S sintese sisteme vir die sintese van Fe-S groepe benodig. Die *suf* sisteem is die hoof Fe-S sintese sisteem in *M. tuberculosis* en word deur ‘n operon wat uit sewe gene bestaan (*Rv1460-Rv1461-Rv1462-Rv1463-csd-Rv1465-Rv1466*) enkodeer. Die eerste geen in die operon, *Rv1460*, enkodeer ‘n transkripsionele reguleerder gebaseer op homologie met SufR, ‘n *suf* sisteem repressor in cyanobacteria. Die doel van hierdie studie is om te bepaal of Rv1460 by die regulering van die *suf* operon se uitdrukking en die fisiologie van *M. tuberculosis* betrokke is.

Om hierdie kennis gaping aan te spreek, is daar gepoog om drie mutante, waarin dele van *Rv1460* weggelaat is, in *M. tuberculosis* H37Rv te genereer. Verbasend kon net een van die mutante, *Rv1460stop*, waarin die proteïen as gevolg van ‘n vroegtydige stop kodon nie gemaak word nie, gegenereer word. Die groei van die *Rv1460stop* mutant is negatief beïnvloed. Dit is moontlik dat die effek van die weglatings in die ander twee mutante, *ΔRv1460* en *Rv1460ΔDNAbd*, die uitdrukking van die daaropvolgende gene negatief beïnvloed en veroorsaak het dat die mutante nie gegenereer kon word nie. Die *Rv1460stop* mutant was meer sensitief vir reaktiewe suurstof groepe, wat impliseer dat Rv1460 belangrik is vir die intrasellulêre groei en patogenese van *M. tuberculosis*.

Geen uitdrukking studies het getoon dat Rv1460 die hele *suf* operon se uitdrukking onderdruk omdat transkripsie van *Rv1460* en *Rv1461* toegeneem het in die *Rv1460stop* mutant in vergelyking met die H37Rv stamvader. Variasie in die transkripsie vlakke van *Rv1460* en *Rv1461*, kan daarop dui dat Rv1460 appart van die res van die gene getranskribeer kan word. Dit dui daarop dat daar ‘n promoter voor die *Rv1461* geen kan voorkom. Elektroforetiese mobiliteits verskuiwing assays is met gesuiwerde Rv1460 proteïen gedoen en het getoon dat Rv1460 proteïen instaat is om aan die *Rv1460* promoter sowel as binne-in die *Rv1461* geen te bind. Dit wys dat Rv1460 die gene direk reguleer deur aan die DNA te bind.

Gesuiwerde Rv1460 proteïen form dimere in oplossing en koördineer ‘n Fe-S groep. Dit het belangrike implikasies vir die funksionering van Rv1460 as ‘n reguleerde, omdat die affiniteit van ander reguleerders vir DNA beïnvloed word deur die Fe-S groep wat dit koördineer. Die Fe-S groep van Rv1460 fasiliteer waarskynlik die suf operon se reaksie op interne Fe-S groep aanvraag. Die rol

van die drie sisteïen residue (C203, C216, C244) in Rv1460 se koördinering van ‘n Fe-S groep is steeds onbekend omdat die Rv1460 proteïen variante wat gemuteerde sisteïen residue bevat steeds in staat was om die Fe-S groep te koördineer. Dit is nie verbasend dat die regulering van die *suf* sisteem in *M. tuberculosis* kompleks is aangesien die uitdrukking van die *suf* sisteem fyn gereguleer moet word om oorlewing in die gasheer te verseker. Hierdie studie dui daarop dat Rv1460 vir die regulasie van die *suf* sisteem en die fisiologie van *M. tuberculosis* belangrik is.

Acknowledgements

I would like to acknowledge and thank the following people and institutions for their contribution to this study:

- My supervisor Dr Monique Williams for entrusting me with this research and her help, support, encouragement and belief in my ability to successfully complete this research. Also, for teaching me many laboratory skills, thoughtful discussion of my work and critical reading of my thesis. Having the opportunity to work with and learn from you has made me grow into a true scientist.
- My supervisor Dr Monique Williams for setting up and doing the Fe-S cluster reconstitution experiments (results indicated in Figure 22 of this thesis) at the Maurice Wohl Clinical Neuroscience Institute in London.
- My co-supervisor Prof RM Warren for his help and thoughtful discussion of my work as well as critical reading of my thesis.
- Prof Trevor Sewel at the Electron Microscope Unit, University of Cape Town for allowing me to use the protein purification equipment.
- Dr Brandon Weber at the Electron Microscope Unit, University of Cape Town for allowing me to use the protein purification equipment, teaching me how to do large-scale protein purification and thoughtful discussions on optimising protein purification and yield. Thank you for your willingness to take so much time to spend entire days helping me. Also, for allowing us to use your Oceans Optics spectrophotometer to assess Fe-S cluster reconstitution. Without you this part of my project would not have been possible.
- Prof. Annalisa Pastore and Dr Salvatore Adinolfi at the Maurice Wohl Clinical Neuroscience Institute in London for allowing Monique to use your laboratory and equipment and expertise for Fe-S cluster reconstitution and providing the vectors and protocols for production and purification of *Escherichia coli* IscS and IscU. Also for guidence and help setting up experiments and critical discussion of results. Without you this part of my project would not have been possible.
- Prof Shez Reid and Dr Brian Kullin at the Department Molecular and Cell Biology University of Cape Town for allowing the use of your anaerobic box and Brian for taking time out to help with our setting up. Without you this part of my project would not have been possible.
- Dr Bavesh Kana at University of Witwatersrand for help with and thoughtful discussion about RT-qPCR. Without you I would still be struggling to get repeatable curves.

- Dr Thomas Oelgeschlager for providing valuable input about electrophoretic mobility shift assays and thoughtful discussion of the electrophoretic mobility shift assay results.
- Prof Gerhard Tromp for valuable advice and help with statistical analyses used in this study.
- Prof Samantha Sampson for critcal reading of my literature review and valuable advice about my project.
- Prof Paul van Helden for providing such a well rounded, productive and helpful environment within the DST/NRF Centre of Excellence in Biomedical Tuberculosis Research at Stellenbosch University and for providing the research funding for this project.
- Harry Crossley Grants for providing research funding for this project.
- NRF Innovation Doctoral Scholarship, NRF Extension Scholarship Support for Doctoral Studies and Ernst and Ethel Eriksen trust for bursaries provided, making it possible for me to continue with my studies.
- The Newton Fund Researchers Link Travel Grant awarded to Dr Monique Williams to travel to London to do the Fe-S cluster reconstitution experiments at the Maurice Wohl Clinical Neuroscience Institute in London.
- The Fe-S cluster cluster, Monique, Lucinda, Jessie and Nandi, for your encouragement, thoughtful discussion of my work and lengthy journal clubs. I am proud to have been part of this group.
- Carine Sao Emani for you encouragement, help and support, thoughtful discussion of results and prayers. Even though you were always so busy, you always made time for me. Your best stories will come from your struggles. The seeds of your successes are in your failures. Your praises will be birthed from your pains. Keep standing. I have never seen a storm last forever. Seasons change. Be encouraged – unknown.
- My mother (Marlene), father (Wikus) and sister (Lodene) for their unwavering support, encouragement, love and help.
- Last, but not least, I want to thank the Lord for providing me with the intellectual capacity to be able to study and for guiding me and providing me with strength, perseverance and help to complete this research. Faith is not knowing what the future holds, but knowing who holds the future – unknown.

Many of life's failures are people who did not realise how close they were to success when they gave up – Thomas Edison

Table of contents

Page

Author's declaration..... ii

Abstract……iii

Opsomming..... v

Acknowledgements..... vii

Table of contents……...ix

List of figures……...xii

List of tables……xiii

List of abbreviations……...xiv

1. Literature review: Regulation of iron-sulphur cluster biogenesis....1

1.1 Fe-S cluster biogenesis systems in prokaryotes……... 2

1.2 The Fe-S cluster biogenesis system in mycobacteria……...5

1.3 Regulation of Fe-S cluster biogenesis in prokaryotes……..8

1.3.1 IscR……9

1.3.2 Fur……10

1.3.3 OxyR and IHF……12

1.3.4 RyhB and FnrS……12

1.3.5 SufR……13

1.3.6 SigM and DtxR……14

1.4 Regulation of Fe-S cluster biogenesis in *M. tuberculosis*……...15

1.5 Rv1460, a possible *suf* operon regulator in *M. tuberculosis*……...16

1.6 Rv1460 could function as a global regulator……..16

1.7 Other Fe-S cluster biogenesis regulators in *M. tuberculosis*……..17

1.8 Future prospects……...18

2. Hypotheses, aims and objectives……22

3. Methods……24

Page

3.1 Bacterial strains and culture conditions……24

3.2 Cloning……...24

3.3 Generation of *M. tuberculosis* mutants harbouring unmarked deletions in *Rv1460*……27

3.3.1 Generation of suicide delivery vectors……..27

3.3.2 Generation of unmarked Rv1460 deletions by two-step allelic exchange……28

3.4 Complementation……32

3.5 RNA extraction……32

3.6 RT-qPCR……...33

3.7 Cumene hydroperoxide exposure……... 34

3.8 Bipyridyl exposure……... 35

3.9 Generation of vectors for the production of Rv1460 and Rv1460 C→S protein variants in *E. coli*……..35

3.10 Optimisation of Rv1460 protein production in *E. coli*……... 37

3.11 Large-scale purification of Rv1460 protein……... 37

3.12 Generation of vectors for the production of $IscS_{Mtb}$ protein in *E. coli*…… 39

3.13 Fe-S cluster reconstitution……...39

3.14 Electrophoretic mobility shift assay……... 40

3.15 Circular dichroism……... 40

3.16 Statistical analyses……41

3.17 Bioinformatic analysis……..41

4. Results…….. 42

4.1 The Rv1460 protein is highly conserved among mycobacteria…….... 42

4.2 Generation of *M. tuberculosis* mutants harbouring unmarked deletions in *Rv1460*……... 44

4.2.1 Generation of suicide delivery and complementation vectors……44

4.2.2 Generation of unmarked Rv1460 deletion mutants by two-step allelic exchange…… 44

4.3 Phenotypic characterisation of the *Rv1460stop* mutant……49

Page

4.3.1 Growth under standard culture conditions……49

4.3.2 Growth in the presence of oxidative stress……51

4.3.3 Growth under iron limiting conditions…… 53

4.4 Effect of *Rv1460stop* mutation on *suf* operon expression…… 55

4.5 Generation of vectors for the production of Rv1460 protein in *E. coli*…... 56

4.6 Optimisation of Rv1460 protein production in *E. coli*……... 56

4.7 Generation of vectors for the production of Rv1460 C→S protein variants in *E. coli*……59

4.8 Large-scale purification of Rv1460 and Rv1460 C203,216,244S proteins.. 59

4.9 Generation of vector for the production of $IscS_{Mtb}$ protein in *E. coli*…….. 62

4.10 Optimisation of $IscS_{Mtb}$ protein production in *E. coli*…….... 62

4.11 Large-scale purification of $IscS_{Mtb}$ protein……64

4.12 Fe-S cluster reconstitution……...65

4.13 Electrophoretic mobility shift assay……... 69

5. Discussion……72

5.1 Rv1460 is required for *M. tuberculosis in vitro* growth……72

5.2 Loss of *Rv1460* alters *M. tuberculosis* growth *in vitro* and survival in the presence of oxidative stress…… 77

5.3 Rv1460 is not required for growth under conditions of iron limitation…..78

5.4. Rv1460 is a transcriptional repressor of the *suf* operon……...78

5.5. Production and purification of recombinant Rv1460 protein in *E. coli*…..79

5.6 Rv1460 binds to its own promoter and within *Rv1461*…….. 81

5.7. Rv1460 coordinates an Fe-S cluster……82

6. Conclusions……..86

7. Future studies……..88

8. References……90

9. Supplementary information……107

List of figures

Page

Figure 1. Schematic representation of Fe-S clusters indicating possible oxidation states and valences. 1

Figure 2. Fe-S cluster biogenesis by the *E. coli* (A) isc and (B) suf systems. 4

Figure 3. The mycobacterial *suf* operon compared to the *suf* operon in other bacteria. 7

Figure 4. Regulation of Fe-S cluster biogenesis in *E. coli*. 11

Figure 5. Representation of the three *Rv1460* deletion mutants to be generated compared to the wild-type. 31

Figure 6. Generation of *M. tuberculosis Rv1460* deletion mutants by two-step allelic exchange. 32

Figure 7. Rv1460 is conserved in various mycobacterial species. 42

Figure 8. Rv1460 has conserved cysteine residues present in SufR homologues. 43

Figure 9. PCR screening to determine the configuration of *ΔRv1460* and *Rv1460ΔDNAbd* SCOs. 46

Figure 10. Genotypic characterisation of the *M. tuberculosis ΔRv1460*:: pMV*Rv1460compl* and *Rv1460ΔDNAbd*:: pMV*Rv1460compl* mutants. 47

Figure 11. Genotypic characterisation of the *M. tuberculosis Rv1460stop* mutants 48

Figure 12. Growth curves of *Rv1460stop* mutants under standard culture conditions. 49

Figure 13. Growth of the *Rv1460stop* 5.20 mutant on solid media. 51

Figure 14. Growth and survival of *Rv1460stop* mutants in the presence of oxidative stress. 52

Figure 15. Growth curves of the *Rv1460stop* 5.20 mutant under iron limitation 53

Figure 16. Relative expression of *Rv1460* and *Rv1461* 55

Figure 17. SDS-PAGE of Rv1460 production in Rosetta-gami 2 and Arctic express (DE3). 57

Figure 18. Large-scale purification of His-tagged Rv1460 and Rv1460C203,216,244S. 60

Figure 19. Far-UV CD spectrum of Rv1460. 62

Figure 20. SDS-PAGE of $IscS_{Mtb}$ production in Rosetta 2 and Arctic express (DE3). 63

Figure 21. Large-scale purification of His-tagged $IscS_{Mtb}$. 64

Figure 22. *In vitro* reconstitution on purified Rv1460. 67

Figure 23. EMSA of "as purified" Rv1460 with the *Rv1460* promoter and *Rv1461* binding regions. 70

Figure 24. Re-annotated *Rv1460* start site and position of transposon insertion sites and predicted ChIP-seq peak centers for Rv1460 binding. 73

List of tables

Page

Table 1. Possible *suf* operon global regulators 19

Table 2. List of bacterial strains used and generated during this study 25

Table 3. List of plasmids used and generated during this study 26

Table 4. List of primers and strategy used for the generation of suicide delivery and complementation vectors 29

Table 5. List of primers and strategy for SCO configuration screening and genotyping of DCOs 30

Table 6. List of primers used for RT and amplification of *Rv1460*, *Rv1461* and *SigA* for RT-qPCR 34

Table 7. Primers used for the generation of Rv1460 and IscS protein production vectors and EMSA 36

List of abbreviations

A	Absorbance
AD	Albumin dextrose
ADC	Albumin dextrose catalase
amp	Ampicillin
A. vinelandii	*Azotobacter vinelandii*
bp	Base pairs
C	Cysteine
CD	Circular dichroism
cDNA	Complementary deoxyribonucleic acid
°C	Degrees celcius
CFUs	Colony forming units
C. glutamicum	*Corynebacterium glutamicum*
CHP	Cumene hydroperoxide
Cm	Chloramphenicol
CV	Column volumes
Da	Dalton
DCOs	Double cross overs
DNA	Deoxyribonucleic acid
DNAbd	DNA-binding domain
DTT	Dithiothreitol
E. chrysanthemi	*Erwinia chrysanthemi*
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
E. faecalis	*Enterococcus faecalis*
Fe-S	Iron-sulphur
Fe^{+3}	Ferric iron
Fe^{+2}	Ferrous iron
Fdx	Ferredoxin
FNR	Fumarate nitrate reductase
Fur	Ferric uptake regulator
×g	Centrifugal force
genta	Gentamicin
g	Gram
h	Hours
hyg	Hygromycin
His	Histidine
HIV	Human immunodeficiency virus
H. pylori	*Helicobacter pylori*
IHF	Integration host factor
IPTG	Isopropyl β-D-thiogalactoside
isc system	Fe-S cluster system
kana	Kanamycin
KatG	Catalase-peroxidase
kDa	Kilo Dalton
L	Litres
lacZ	β-galactosidase encoding gene
LB	Luria-Bertani
M	Molar

µ	Micro
min	Minutes
ml	Millilitres
mM	Millimolar
MmpL	Mycobacterial membrane protein large
mRNA	Messenger RNA
mV	Millivolt
M. avium	*Mycobacterium avium*
M. tuberculosis/ Mtb	*Mycobacterium tuberculosis*
M. marinum	*Mycobacterium marinum*
M. smegmatis	*Mycobacterium smegmatis*
M. bovis	*Mycobacterium bovis*
M. leprae	*Mycobacterium leprae*
nm	Nanometers
NO	Nitric oxide
Ni	Nickel
nif	Nitrogen fixation
OADC	Oleic albumin dextrose catalase
OD	Optical density
P. aeruginosa	*Pseudomonas aeruginosa*
PCR	Polymerase chain reaction
pmol	Picomol
RBS	Ribosomal binding site
RE	Restriction enzyme
ROS	Reactive oxygen species
RNA	Ribonucleic acid
RNS	Reactive nitrogen species
RT	Reverse transcription
RT-qPCR	Real-time quantitative PCR
s	Seconds
S	Sulphur
S^{-2}	Sulphide
SCO	Single cross overs
SigA	Sigma factor A
SNPs	Single nucleotide polymorphisms
TB	Tuberculosis
Tris	Tris(hydroxymethyl)aminomethane
Tween 80	Polyoxyethylene sorbitan monooleate
UV	Ultraviolet
V	Volt
WHO	World Health Organisation
X-gal	5-Bromo-4 Chloro-3 Indolyl β-D-galactosidase

1. Literature review: Regulation of iron-sulphur cluster biogenesis

Iron-sulphur (Fe-S) clusters are versatile, chemically reactive protein cofactors that occur in all living organisms. Since the function of many proteins is dependent on the presence of Fe-S cluster cofactors, these clusters are involved in multiple important cellular processes including DNA replication and repair, RNA modification, gene expression, metabolism, respiration and redox signalling (Py and Barras, 2010; Roche *et al.*, 2013). Fe-S clusters occur mainly in the rhombic [2Fe-2S] or cubic [4Fe-4S] cluster forms and contain either ferric (Fe^{+3}) or ferrous (Fe^{+2}) iron and sulphide (S^{-2}) (Figure 1). Within proteins, the clusters are mostly coordinated by sulphur atoms provided by cysteine residues, although they can be coordinated by nitrogen provided by histidine or arginine residues, or by oxygen provided by aspartate, glutamine or tyrosine residues or even exogenous ligands such as water (Beinert, 2000).

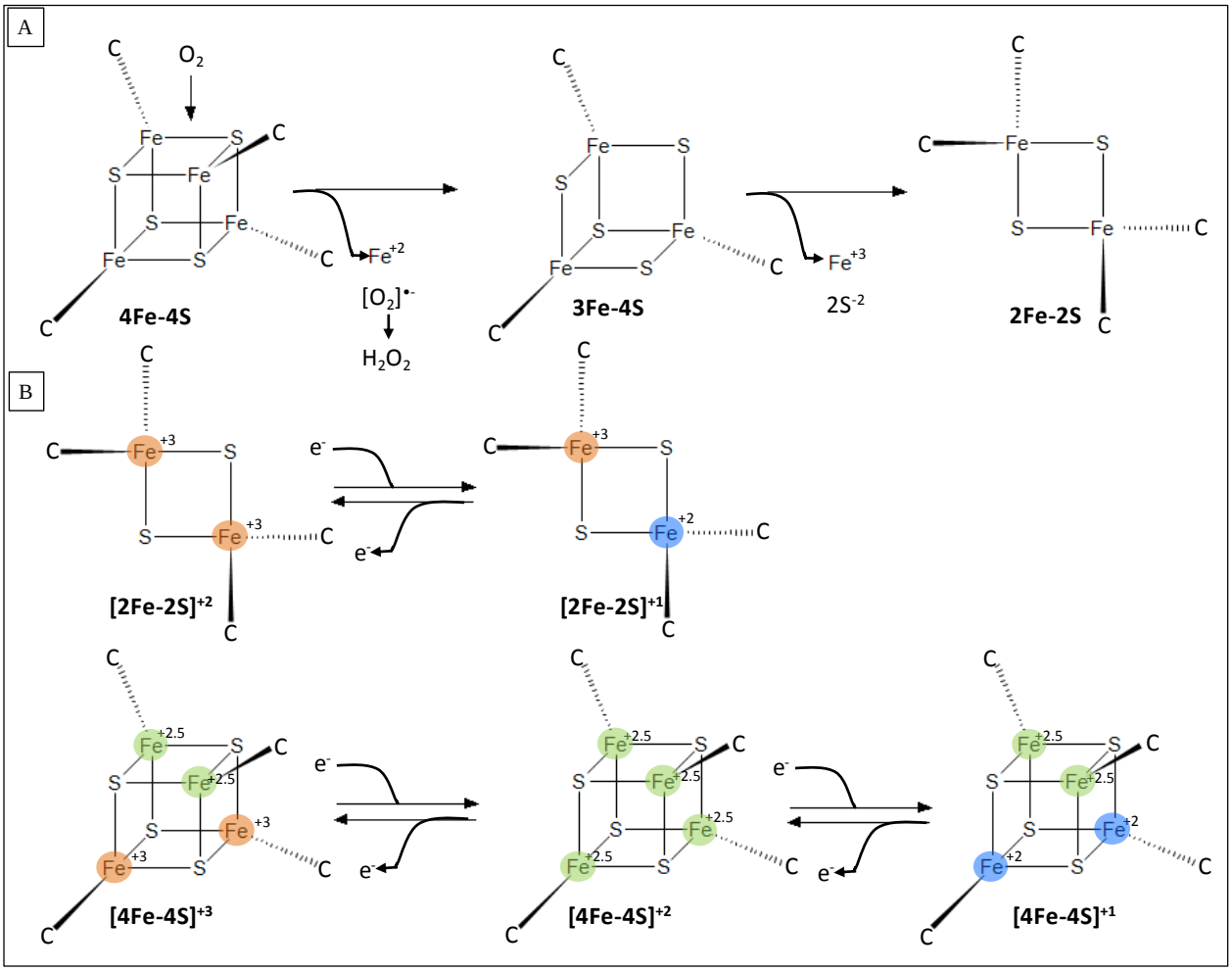


Figure 1. Schematic representation of Fe-S clusters indicating possible oxidation states and valences. (A) Oxidation of the 4Fe-4S cluster to form a 2Fe-2S cluster. The structure of the cubic 4Fe-4S cluster, intermediate butterfly shaped 3Fe-4S cluster and rhombic 2Fe-2S cluster and cluster coordination by cysteine residues (C) from the protein are indicated. This Figure is adapted from Outten, (2007). (B) The valences of 4Fe-4S and 2Fe-2S

clusters in various oxidation states (localised Fe^{+3} (orange), Fe^{+2} (blue) and delocalized $Fe^{+2.5}$ $Fe^{+2.5}$ pairs (green)). This Figure is adapted from Beinert *et al*., (1997).

1.1 Fe-S cluster biogenesis systems in prokaryotes

Fe-S cluster formation and transfer of the cluster to the apo-proteins cannot occur spontaneously *in vivo*, as complex Fe-S biogenesis systems are required to protect the clusters from the deleterious effects of reactive oxygen species (ROS) and free Fe^{+2}, Fe^{+3} and S^{-2} ions present in the cell. Fe^{+2} within the $[4Fe\text{-}4S]^{+2}$ clusters can easily be oxidized to Fe^{+3} in the presence of oxidants; a process during which Fe^{+2} is released from the cluster and inactivation of the protein occurs. This released Fe^{+2} drives the production of additional ROS which in turn interact with and damage biological molecules (Wardman and Candeias, 1996).

Three Fe-S biogenesis systems, the nitrogen fixation (*nif*) (Jacobson *et al.*, 1989), Fe-S cluster (*isc*) (Zheng *et al.*, 1998) and sulphur mobilization (*suf*) systems (Takahashi and Tokumoto, 2002) have been identified in prokaryotes. The number, type, composition and organisation of the operons encoding the Fe-S biogenesis systems vary between prokaryotes. For instance, *Azotobacter vinelandii* (the bacteria in which the *nif* system was first identified) contains the *nif* and *isc* systems (Jacobson *et al.*, 1989; Zheng *et al.*, 1998), *Escherichia coli* contains the *isc* and *suf* systems (Takahashi and Tokumoto, 2002; Tokumoto and Takahashi, 2001), *Erwinia chrysanthemi* contains all three (Rincon-Enriquez *et al.*, 2008) and *Enterococcus faecalis* contains only the *suf* system (Riboldi *et al.*, 2011). Gram-positive bacteria mostly contain only the *suf* system, although there are exceptions (Santos *et al.*, 2014). In some bacteria the systems are atypical, as they contain a combination of both *suf* and *iscU*-like genes (Huet *et al.*, 2005) or significantly reduced as genes encoding some of the components are missing as in the case of the *nifSU* operon in *Helicobacter pylori* (Olson *et al.*, 2000). Fe-S cluster biogenesis systems can either be present on separate locations within the genome or co-localize, forming part of the same operon as in the case of the *isc-suf* operon in *Rhodobacter sphaeroides* (Remes *et al.*, 2015).

In *E. coli*, the *isc* system is the housekeeping system and the *suf* system is used under stress conditions (Dai and Outten, 2012; Jang and Imlay, 2010; Outten *et al.*, 2004). This is one of the few bacteria in which a distinct role for the two systems has been established. The *isc* and *suf* systems can, however, compensate for each other under standard culture conditions in *E. coli*, while the loss of both is lethal (Takahashi and Tokumoto, 2002). Interestingly, the function of the systems seems to be conserved among bacterial species as the *H. pylori nif* system is able to take over some of the functions of the

isc and *suf* system when cloned into *E. coli* (Tokumoto *et al.*, 2004) and the *suf* system of *E. faecalis* (*sufCDSUB*) is able to complement the *E. coli suf* system, but not the *isc* system in *E. coli* or *A. vinelandii* (Riboldi *et al.*, 2011).

Fe-S cluster biogenesis systems mobilize iron and sulphur, assemble Fe-S clusters on scaffold proteins and transfer the clusters to apo-proteins. The *nif* system is mainly involved in Fe-S cluster biogenesis in nitrogenases in nitrogen-fixing bacteria (Jacobson *et al.*, 1989) and is beyond the scope of this review. The *isc* and *suf* systems in *E. coli* have been extensively researched and Fe-S cluster biogenesis by these systems is described in Figure 2. For the isc system (Figure 2 A), IscS is a cysteine desulphurase which provides sulphur from cysteine where it is stored in the form of a persulphide (Schwartz *et al.*, 2000). IscS transfers the sulphur to the scaffold protein, IscU, (Agar *et al.*, 2000) which serves as the primary substrate for IscS (Kim *et al.*, 2012). The iron source for cluster assembly remains controversial, although CyaY (Yoon and Cowan, 2003), a bacterial frataxin homologue, and IscX (Kim *et al.*, 2015) have been proposed to play this role. CyaY negatively regulates Fe-S cluster formation *in vitro* (Adinolfi *et al.*, 2009). IscU accepts iron from the iron donor and assembles a 2Fe-2S cluster (Agar *et al.*, 2000) while ferredoxin (Fdx) provides the reducing equivalent to couple two $[2Fe\text{-}2S]^{+2}$ clusters to a $[4Fe\text{-}4S]^{+2}$ on IscU (Chandramouli *et al.*, 2007). IscU interacts with the chaperone and co-chaperone, HscA and HscB, which facilitates the transfer of the Fe-S cluster from IscU to the apo-protein. This process is energy dependent, provided by the ATPase activity of HscA (Hoff *et al.*, 2000). IscA's role in Fe-S cluster biogenesis is controversial as it could act as an alternative scaffold (Ollagnier-de Choudens *et al.*, 2001) or intermediate A-type Fe-S cluster carrier (Vinella *et al.*, 2009).

In the suf system (Figure 2 B), sulphur is transferred from the cysteine desulphurase, SufS, in the form of a protein-bound persulphide to the sulphur transfer protein SufE. SufE transfers the sulphur to SufB (Layer *et al.*, 2007). The SufBCD complex (Hirabayashi *et al.*, 2015; Outten *et al.*, 2003) occurs mainly in a 1:2:1 (B:C:D) stoichiometry (Wollers *et al.*, 2010) and is able to function as a scaffold for *de novo* Fe-S cluster assembly (Tian *et al.*, 2014; Wollers *et al.*, 2010). SufD may be important in this iron acquisition although its exact role remains unclear (Saini *et al.*, 2010). SufC, a member of the ABC family of ATPase proteins (Nachin *et al.*, 2003), provides the energy for conformational changes in the SufB/D interface (Hirabayashi *et al.*, 2015). The SufA A-type carrier protein is responsible for the transfer of Fe-S clusters to apo-proteins (Chahal *et al.*, 2009; Gupta *et al.*, 2009). Interestingly, in some Gram-positive bacteria including *Bacillus*, another scaffold, SufU, accepts the sulphur from SufS and iron from an unknown source to assemble the Fe-S cluster (Albrecht *et al.*, 2010).

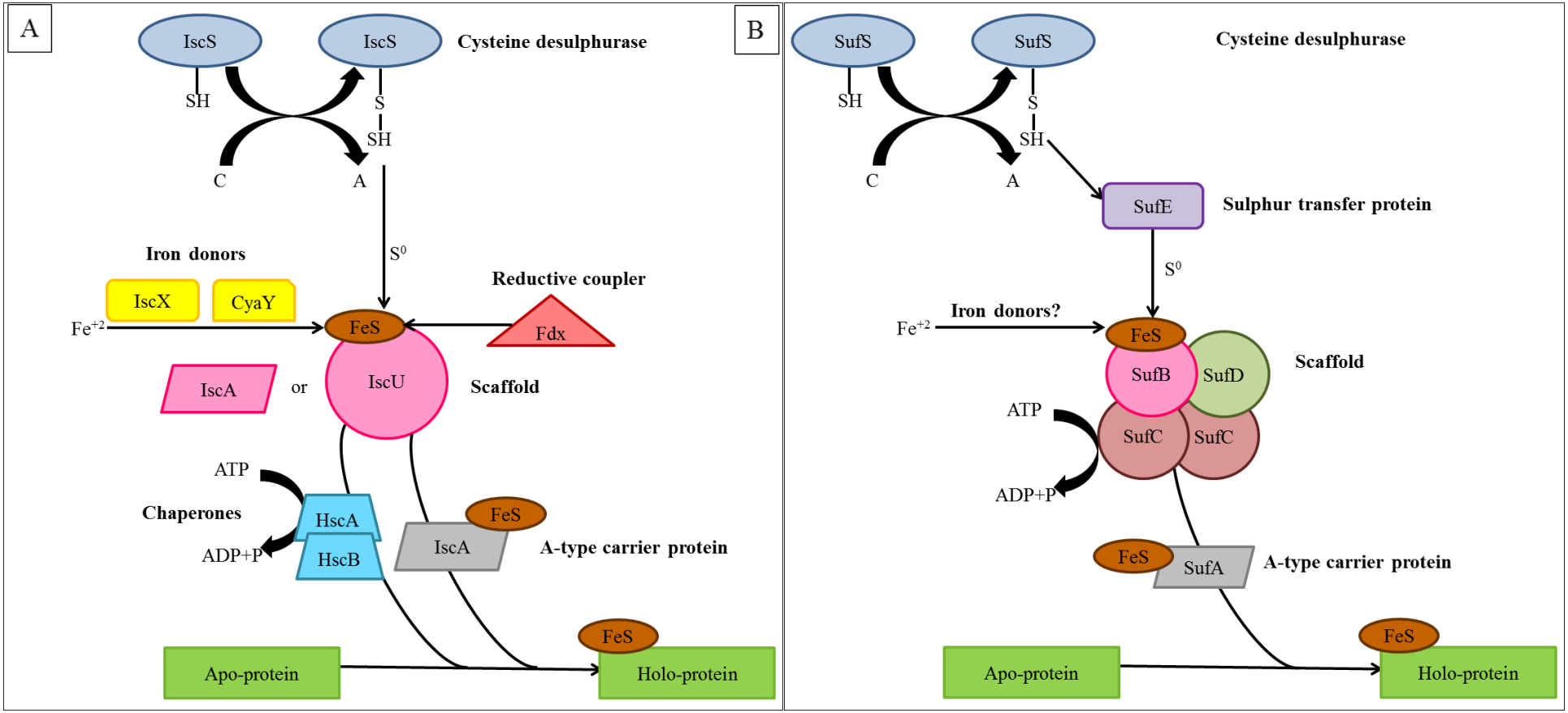


Figure 2. Fe-S cluster biogenesis by the *E. coli* (A) isc and (B) suf systems. Cysteine desulphurase (IscS and SufS) provides the sulphur from cysteine where it is stored as persulphide and transfers the sulphur to the scaffold protein (IscU or IscA, a possible alternative scaffold protein, and the SufBCD complex respectively) via transfer proteins (SufE). Iron is incorporated from an unknown source (potentially involving the IscX and CyaY iron donors in the isc system). Ferredoxin (Fdx) reductively couples two $[2Fe\text{-}2S]^{+2}$ clusters to form a $[4Fe\text{-}4S]^{+2}$ cluster in the isc system. The Fe-S cluster is transferred via A-type carrier proteins (IscA and SufA) to the apo-proteins. This transfer is facilitated by the chaperone and co-chaperone HscA and HscB in the isc system. Proteins are colour coded according to their functions. This figure does not take into account the oligomerization state of proteins. This figure is adapted from Roche *et al.*, (2013).

The presence of multiple systems may allow pathogenic bacteria to survive under a range of intracellular conditions, broadening their potential hosts (Py and Barras, 2010). The plant pathogen *E. chrysanthemi*, with all three systems, can for instance infect a wide range of plant hosts (Rincon-Enriquez *et al.*, 2008). Alternatively the frequency of exposure to oxidative stress may drive selection for the system(s). Bacteria encountering oxidative stress on a regular basis generally dispense of the *isc* system (Jang and Imlay, 2010), since the isc system is susceptible to oxidative stress (Dai and Outten, 2012; Lee *et al.*, 2004; Outten *et al.*, 2004; Runyen-Janecky *et al.*, 2008).

1.2 The Fe-S cluster biogenesis system in mycobacteria

Mycobacteria belong to the Actinobacteria phylum and the *Mycobacteriaceae* family – many of which causes disease. The *suf* system is conserved among members of the mycobacterial genus as the same genes and gene organisation occurs in *Mycobacterium tuberculosis* (Cole *et al.*, 1998; Huet *et al.*, 2005), *Mycobacterium leprae* (Kapopoulou *et al.*, 2011), *Mycobacterium bovis* (Garnier *et al.*, 2003), *Mycobacterium marinum* (Kapopoulou *et al.*, 2011), *Mycobacterium avium* (Lew *et al.*, 2011) and *Mycobacterium smegmatis* (Kapopoulou *et al.*, 2011) (Figure 3). *M. tuberculosis* is the causative agent of the devastating disease tuberculosis (TB), a leading cause of death worldwide alongside Human immunodeficiency virus (HIV) (WHO, 2015). TB treatment is complicated by the emergence of *M. tuberculosis* drug resistance and the remarkable ability of this pathogen to survive within the host and evade the host immune response. Once inhaled, *M. tuberculosis* is phagocytosed by alveolar macrophages where it encounters numerous stresses such as nutrient limitation, low pH, ROS and reactive nitrogen species (RNS) (Schnappinger *et al.*, 2003). One of the limiting nutrients during infection is iron, critical to *M. tuberculosis* survival and pathogenesis (Banerjee *et al.*, 2011). Oxygen is presumably also limiting within the host, especially within hypoxic granulomas (Aly *et al.*, 2006; Via *et al.*, 2008). These limiting factors may explain why the *suf* system is the primary Fe-S cluster biogenesis system in *M. tuberculosis*. In addition to the *suf* system, genes encoding IscS and IscA-like proteins (*Rv3025c* and *Rv2204c*) are present elsewhere in the *M. tuberculosis* genome, however no additional *isc* or *nif* system components were identified (Huet *et al.*, 2005).

The *M. tuberculosis suf* system is encoded in a single operon, containing overlapping start and stop codons and small intergenic regions, consisting of seven co-transcribed genes (*Rv1460-Rv1461-Rv1462-Rv1463-csd-Rv1465-Rv1466*) (Huet *et al.*, 2005; Roback *et al.*, 2007) (Figure 3). As intact MSMEG_3122-MSMEG_3127 was essential while MSMEG_3121 could be separated from the rest of the operon, the *M. smegmatis Rv1460* homologue may be transcribed independently of the

downstream *suf* operon genes (Huet *et al.*, 2005). No promoters have, however, been identified upstream or within the *suf* operon (Huet *et al.*, 2005).

The first gene in the *suf* operon, *Rv1460*, encodes a probable transcriptional regulator that shares 42% amino acid identity with the cyanobacterial SufR. In cyanobacteria, SufR functions as a transcriptional repressor of the *suf* system (Huet *et al.*, 2005; Wang *et al.*, 2004). *Rv1461*, *Rv1462* and *Rv1463* encode SufB, SufC and SufD homologues respectively, and their ability to physically interact has been demonstrated by Yeast-2-hybrid assays (Huet *et al.*, 2005). Rv1461 contains an intein (Topilina *et al.*, 2015). Inteins are mobile genetic elements which are responsible for self-splicing at protein level, and splicing of the Rv1461 intein is essential for its function, as the unspliced protein is unable to interact with its suf partners Rv1462 and Rv1463 (Huet *et al.*, 2006). Intein splicing is sensitive to oxidative and nitrosative stress and is thought to be involved in post-translational regulation (Topilina *et al.*, 2015). The intein insertion site differ for *Rv1461* homologues in *M. tuberculosis*, *M. leprae* and *M.* bovis, but is absent in other mycobacteria (Saves *et al.*, 2002) (Figure 3). Its absence in non-pathogenic mycobacteria suggests that it may be linked to pathogenesis (Huet *et al.*, 2005).

Csd (*Rv1464)* encodes a SufS orthologue and interacts with Rv1461, suggesting that sulphur transfer occurs directly from the cysteine desulphurase to this SufB homologue (Huet *et al.*, 2005). Interestingly, splicing of the Rv1461 intein is not required for Rv1461 to interact with Rv1464 (Huet *et al.*, 2005). IscS is also a cysteine desulphurase, the deletion of which causes impaired Fe-S cluster dependent enzyme activity in *M. tuberculosis*, implying participation in Fe-S cluster biogenesis. It is possible that IscS is able to transfer Fe-S clusters directly to apo-proteins without the need for a scaffold as interaction between IscS and several Fe-S cluster containing proteins was observed (Rybniker *et al.*, 2014). IscS was also able to reconstitute the 4Fe-4S cluster of WhiB3 *in vitro* without the need for providing a scaffold protein (Singh *et al.*, 2009, 2007). This direct interaction could also be indicative of a role for IscS in Fe-S cluster repair (Rybniker *et al.*, 2014). An *iscS* deletion mutant had a slow growth rate on plates and altered colony morphology. This altered growth rate could be partially restored by growth under microaerophilic conditions. IscS deletion mutants were also more sensitive to ROS, indicative of an important role for IscS in *M. tuberculosis* physiology and growth (Rybniker *et al.*, 2014). The need for multiple cysteine desulphurases for optimal growth of *M. tuberculosis* warrants further investigation.

Rv1465 encodes a NifU/IscU-like protein orthologue. In the *isc* system IscU functions as a scaffold protein and receives the protein-bound persulphide directly from IscS (Agar *et al.*, 2000). It is

currently unclear whether Rv1465 similarly functions as a scaffold protein in mycobacteria as it was not observed to interact with IscS from *M. tuberculosis* in Yeast-2-hybrid assays (Rybniker *et al.*, 2014). Rv1466 is thought to be a metal-sulphur cluster biosynthesis enzyme due to the localization of its gene in the *suf* operon. However it does not share homology with any of the *E. coli* or *E. chrysanthemi suf* genes. Rv1465 and Rv1466 similarly do not interact with any proteins encoded by the operon (Huet *et al.*, 2005).

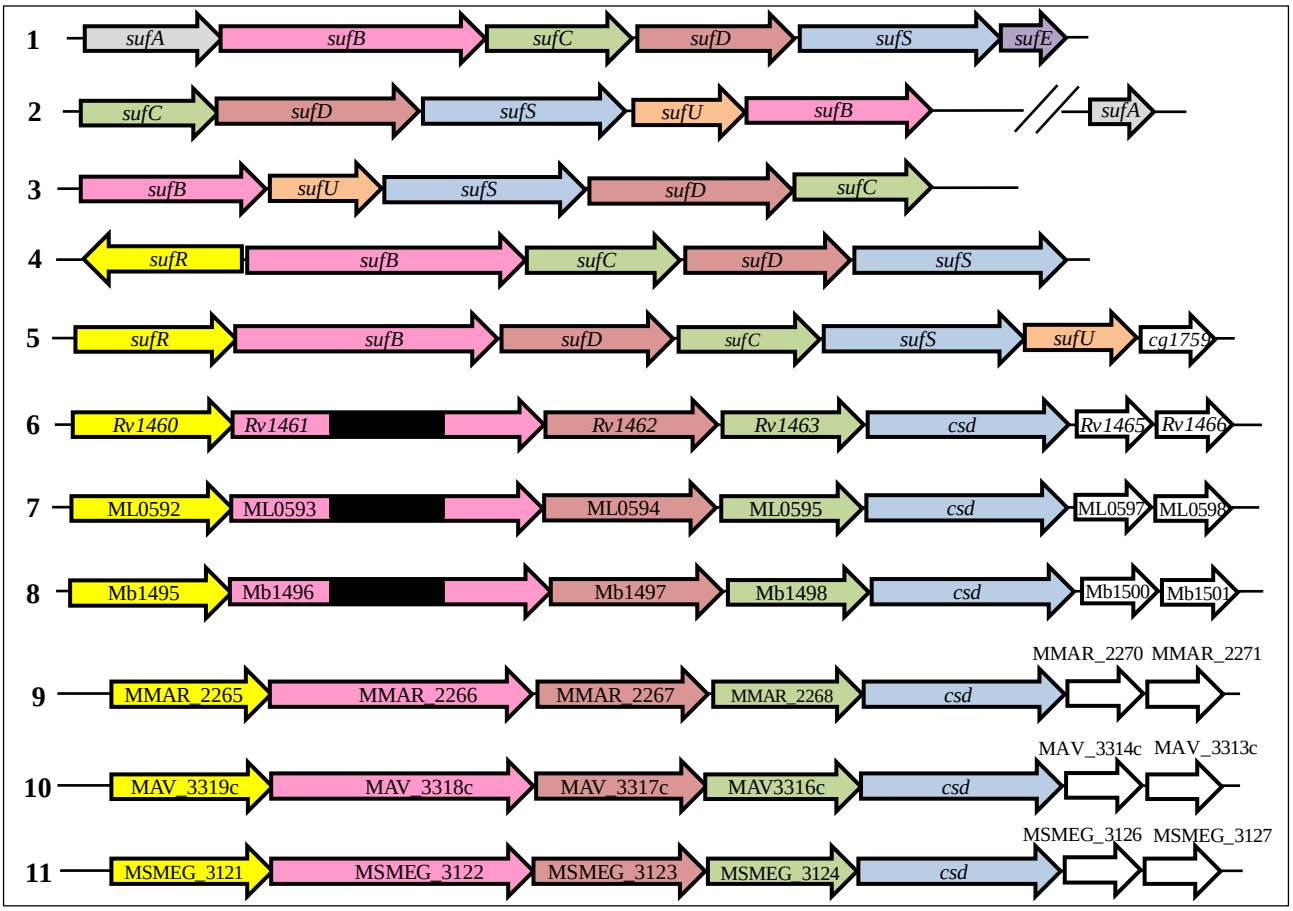


Figure 3. The mycobacterial *suf* operon compared to the *suf* operon in other bacteria. The *suf* operon in (1) *E. coli* (Gram-negative bacteria from the Proteobacteria phylum), (2) *Bacillus subtilis* (Gram-positive bacteria from the Firmicutes phylum) (Albrecht *et al.*, 2010), (3) *E. faecalis* (Riboldi *et al.*, 2009) (Gram-positive bacteria from the Firmicutes phylum), (4) cyanobacteria (Wang *et al.*, 2004), (5) *Corynebacterium glutamicum* (Nakunst *et al.*, 2007) (Gram-positive bacteria from the Actinobacteria phylum) (6) *M. tuberculosis*, (7) *M. leprae*, (8) *M. bovis*, (9) *M. marinum*, (10) *M. avium* and (11) *M. smegmatis.* Genes are colour coded according to their homology with the *E. coli suf* operon components illustrated in Figure 2. Double lines indicate separation in the genome. Notice the lack of a *sufR* gene in the Firmicutes (2,3) and the divergently transcribed *sufR* gene in cyanobacteria (4). The intein in the *M. tuberculosis*, *M. leprae* and *M. bovis Rv1461* homologues are indicated by black blocks. The operons and genes are not drawn to scale.

No orthologues of the *E. coli sufA* gene, encoding an A-type carrier protein, are present, although *Rv2204c* encodes an IscA-like conserved hypothetical protein (Galagan *et al.*, 2010; Reddy *et al.*,

2009). A *sufE* homologue, *Rv3284*, which could act as a sulphur transfer protein, is present elsewhere in the mycobacterial genome (Finn *et al.*, 2016). Even though it is tempting to extrapolate the functions of the mycobacterial suf proteins based on homology with suf components in other bacteria, their functions in Fe-S cluster biogenesis in mycobacteria have not been confirmed experimentally.

Forward genetic screens predict that all the genes in the *suf* operon, except *Rv1460*, are essential for *in vitro* growth of *M. tuberculosis* and *M. marinum* (Griffin *et al.*, 2011; Sassetti *et al.*, 2003; Weerdenburg *et al.*, 2015). Because the *Ms1461-Ms1466* part of the operon needs to be intact to allow survival of *M. smegmatis*, the whole operon (*Ms1461-Ms1466*) was predicted to be essential for *in vitro* growth in *M. smegmatis* (Huet *et al.*, 2005). *Rv1460* was however predicted to be essential for *in vivo* growth in mice (Sassetti and Rubin, 2003), underscoring that all Fe-S cluster containing proteins depend on Fe-S cluster biogenesis. Fifty Fe-S proteins involved in transcriptional regulation, metabolism, energy production and conversion, cell envelope biosynthesis and defence mechanisms have been identified in *M. tuberculosis* (Saini *et al.*, 2012). This list may be incomplete as some Fe-S cluster binding domains may not have been identified yet.

The WhiB family of transcriptional regulators are the main group of Fe-S cluster containing regulators with DNA-binding and activity depending on Fe-S cluster presence and state (Chawla *et al.*, 2012; Singh *et al.*, 2009; Smith *et al.*, 2012, 2010). They are physiologically important in *M. tuberculosis* as they are involved in a wide array of processes including cell division (Gomez and Bishai, 2000; Raghunand and Bishai, 2006), metabolism and metabolic switching upon infection (Singh *et al.*, 2009, 2007), virulence and reactivation (Banaiee *et al.*, 2006; Casonato *et al.*, 2012; Chawla *et al.*, 2012; Singh *et al.*, 2009; Steyn *et al.*, 2002), quorum sensing (Banaiee *et al.*, 2006), drug resistance (Burian *et al.*, 2012; Morris *et al.*, 2005), protein refolding (Konar *et al.*, 2012) and stress response (Geiman *et al.*, 2006). Fe-S cluster homeostasis must therefore be maintained to ensure mycobacterial metabolism.

1.3 Regulation of Fe-S cluster biogenesis in prokaryotes

Overexpression of the *suf* system causes toxic levels of Fe-S clusters and inhibition of bacterial growth (Takahashi and Tokumoto, 2002). Insufficient Fe-S cluster biogenesis is similarly deleterious as Fe-S cluster proteins depend on these cofactors for activity. Optimal regulation of Fe-S cluster biogenesis is thus critical as environmental conditions change to allow adaptation and survival during infection. The fine-tuning of Fe-S cluster biogenesis involves a number of regulators. The most studied bacterial regulators are IscR, which regulates both the *isc* and *suf* systems (Giel *et al.*, 2006;

Lee *et al.*, 2008), ferric uptake regulator (Fur), OxyR and integration host factor (IHF) (Outten *et al.*, 2004).

1.3.1 IscR

IscR, is a member of the Rrf family of transcription factors, has a typical winged helix-turn-helix DNA-binding domain (DNAbd) and forms a dimer in solution (Nesbit *et al.*, 2009; Rajagopalan *et al.*, 2013; Santos *et al.*, 2014). It is a global regulator (Giel *et al.*, 2006; Remes *et al.*, 2015) regulating genes involved in virulence (Kim *et al.*, 2009; Miller *et al.*, 2014; Rincon-Enriquez *et al.*, 2008), pathogenesis (Lim and Choi, 2014; Runyen-Janecky *et al.*, 2008) motility, chemotaxis and adhesion (Lim and Choi, 2014; Wu and Outten, 2009), biofilm formation (Wu *et al.*, 2014; Wu and Outten, 2009) and co-infection (Wong *et al.*, 2013). It regulates Fe-S cluster biogenesis in Gram-negative as opposed to Gram-positive bacteria (Santos *et al.*, 2014). Conserved cysteine residues coordinate a 2Fe-2S cluster, which occurs predominantly in the reduced cluster form ($[2Fe\text{-}2S]^{+}$) (Fleischhacker *et al.*, 2012; Schwartz *et al.*, 2001). IscR primarily receives its Fe-S cluster from the isc system though an overexpressed suf system can act as Fe-S cluster donor (Giel *et al.*, 2006; Mettert and Kiley, 2014; Schwartz *et al.*, 2001).

Two distinct IscR binding site motifs have been identified: Types 1 and 2, upstream of genes repressed or activated by IscR. Promoter regions of genes regulated by IscR contain either a single Type 1 or 2 site. However, the promoter region of IscR contains two Type 1 motifs (Giel *et al.*, 2006; Miller *et al.*, 2014; Romsang *et al.*, 2014) and another IscR binding site unique to the *iscRSUA* promoter (Giel *et al.*, 2006). In *E. coli*, apo-IscR only binds to Type 2 motifs. Fe-S cluster ligation to IscR remodels the protein DNA interface and repositions a key glutamic acid 43 side chain, to broaden the DNA-binding specificity to Type 1 in addition to Type 2 motifs (Rajagopalan *et al.*, 2013; Santos *et al.*, 2014). Replacing glutamic acid 43 with alanine enables apo-IscR to recognize the Type 1 motif, indicating its importance in DNA recognition (Rajagopalan *et al.*, 2013; Santos *et al.*, 2014). Interestingly, glutamic acid 43 is conserved in *Pseudomonas aeruginosa*, but it is not required for holo-IscR-mediated repression of the *isc* promoter (Romsang *et al.*, 2014). Holo-IscR is thus required for regulation of genes in which the Type 1 IscR motif is present, while apo-IscR regulates genes in which the Type 2 motif is present (Giel *et al.*, 2006; Giel *et al.*, 2013; Nesbit *et al.*, 2009; Remes *et al.*, 2015; Romsang *et al.*, 2014) (Figure 4 (1)).

Holo-IscR binds independently to all three IscR binding sites within the *isc* operon, but binding to both Type 1 motifs is necessary to completely repress the *isc* operon (Giel *et al.*, 2013) (Figure 4 (1)).

The *isc* system is more tightly repressed under anaerobic growth conditions presumably because mainly holo-IscR is present and a lower demand for Fe-S clusters exists under anaerobic conditions. Under aerobic conditions oxygen and ROS damage Fe-S clusters increasing the Fe-S cluster demand (Figure 4 (2)). IscR has a low affinity for Fe-S clusters, allowing it to respond to a lack of Fe-S cluster substrates (Giel *et al.*, 2013). This is due to the unusual Fe-S cluster ligands (namely three cysteine and one histidine residue) making IscR a poor competitor for Fe-S clusters in comparison to other isc system components (Fleischhacker *et al.*, 2012; Romsang *et al.*, 2014). These ligands are not conserved among all IscR homologues (Remes *et al.*, 2015). Loss of the Fe-S cluster by IscR leads to increased expression of the *isc* system (Giel *et al.*, 2013; Romsang *et al.*, 2014) (Figure 4 (3)). *IscR* expression is also de-repressed in the absence of holo-IscR binding, which increases apo-IscR protein levels (Mettert and Kiley, 2014). The lack of IscR Fe-S cluster affinity ensures that Fe-S clusters are incorporated into all other isc system substrates before allowing cluster incorporation into IscR (Giel *et al.*, 2013) (Figure 4 (8,15)). As more Fe-S clusters are made and restored by the isc system, less competition between IscR and other isc substrates for the Fe-S clusters occurs and the Fe-S cluster is incorporated into IscR, allowing repression of the *isc* system by holo-IscR (Giel *et al.*, 2013; Romsang *et al.*, 2014) (Figure 4 (8)). The *isc* system is also induced in mutants defective in Fe-S cluster assembly (Schwartz *et al.*, 2001) and under iron- (Outten *et al.*, 2004) and sulphur-limiting conditions (Gyaneshwar *et al.*, 2005).

Three IscR binding sites were identified within the *suf* promoter, one of which is a Type 2 IscR binding site (Giel *et al.*, 2006; Lee *et al.*, 2008; Miller *et al.*, 2014) (Figure 4). Interestingly, both apo-IscR and holo-IscR bind within the *suf* promoter and IscR binding to the *suf* promoter seems to be dependent on increased IscR concentration as opposed to absence or presence of the cluster (Mettert and Kiley, 2014). Apo-IscR binds to the *suf* operon promoter and activates *suf* operon transcription under iron-limiting and oxidative stress conditions (Figure 4 (6)) (Mettert and Kiley, 2014; Rincon-Enriquez *et al.*, 2008; Runyen-Janecky *et al.*, 2008; Yeo *et al.*, 2006). IscR thus functions as an Fe-S cluster biogenesis regulator that allows switching between the *isc* and *suf* systems in *E. coli*.

1.3.2 Fur

Fur is a global regulator acting as both an activator and repressor and directly regulates genes involved in iron assimilation in response to iron availability. Additionally, Fur regulates genes involved in metabolism, oxidative stress response and virulence (Gao *et al.*, 2008; McHugh *et al.*, 2003). Fur senses iron availability through its own Fe-S cluster and only holo-Fur binds to DNA. The Fur-binding site and IscR Type 2 binding site within the *suf* promoter overlap (Figure 4) (Lee *et al.*, 2008;

Outten *et al.*, 2004). Under standard growth conditions holo-Fur, but not apo-Fur, binds to the Fur binding site within the *suf* promoter and inhibits transcription of the *suf* operon (Figure 4 (4)) (Lee *et al.*, 2008; Runyen-Janecky *et al.*, 2008). The transcription of the *suf* operon is not completely repressed as *suf* transcripts are present at basal level (Mettert and Kiley, 2014). Holo-Fur acts as an anti-activator of the *suf* operon as it must dissociate before apo-IscR can bind (Lee *et al.*, 2008).

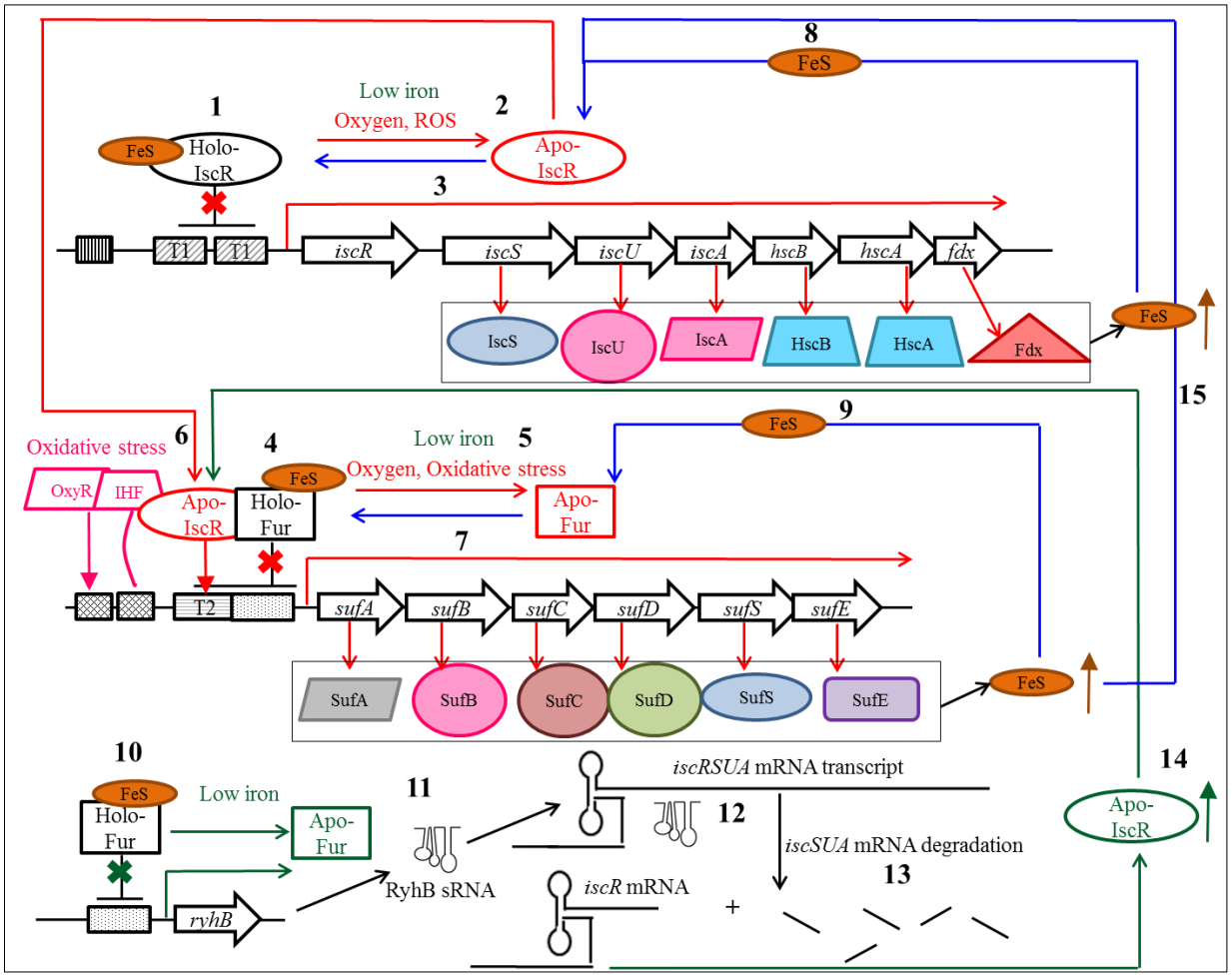


Figure 4. Regulation of Fe-S cluster biogenesis in *E. coli*. (1) Holo-IscR represses *isc* operon expression. (2) Under aerobic, oxidative stress and iron limiting conditions, holo-IscR loses its cluster and (3) de-repression of the *isc* operon occurs. (4) Holo-Fur which acts as repressor of the *suf* operon, (5) loses its cluster and dissociates from the *suf* operon promoter (6) allowing binding of apo-IscR (7) which activates *suf* operon expression. (8) Restoration of Fe-S clusters leads to incorporation of Fe-S clusters into IscR which subsequently represses *isc* operon expression. (9) An Fe-S cluster is also incorporated into Fur and holo-Fur binds to the *suf* promoter to repress *suf* operon expression. (10) Holo-Fur normally represses small RNA *ryhB* expression, (11) but under iron limiting conditions loss of its cluster occurs causing *ryhB* expression. (12) RyhB base-pairs with the *iscRSUA* mRNA transcript and (13) causes degradation of the *iscSUA* mRNA transcript, but not the *iscR* mRNA transcript and (14) apo-IscR expression is increased which activates *suf* operon expression until Fe-S homeostasis is restored. (15) An Fe-S cluster is incorporated into IscR and Fur and holo-IscR and holo-Fur repress the *isc* and *suf* systems respectively. This Figure is adapted from Py and Barras, (2010).

Dissociation of Fur from the *suf* promoter de-represses the *suf* system (Figure 4 (7)), but apo-IscR is required to activate the *suf* operon (Mettert and Kiley, 2014) (Figure 4 (6,7)). Formation of Fe-S clusters by the suf system incorporates an Fe-S cluster into Fur and eventually IscR. The resultant holo-Fur and holo-IscR, in turn repress the *suf* and *isc* operons (Figure 4 (9,15)).

1.3.3 OxyR and IHF

OxyR belongs to the LysR family of transcriptional regulators. In addition to regulating its own expression, OxyR globally regulates genes involved in the oxidative stress response. While OxyR is an activator of the oxidative stress response genes in *E. coli*, in corynebacteria by contrast it represses oxidative stress response genes under standard culture conditions. This reversal of activator versus repressor functionality means that *oxyR* disruption causes hydrogen peroxide sensitivity in *E. coli*, but increased resistance to hydrogen peroxide in corynebacteria. Even though the OxyR regulon varies greatly between bacteria, the regulation of catalases and hydroperoxidases is universally conserved (Milse *et al.*, 2014).

The *suf* system is part of the stress response genes in the OxyR regulon. In corynebacteria *SufR* is upregulated as part of the initial response to oxidative stress and the rest of the operon within the prolonged oxidative stress response. OxyR binds to the promoter region of *sufR* while oxidative stress causes its dissociation from the promoter enabling *suf* operon transcription (Figure 4). Interestingly, OxyR binds to the *suf* promoter only weakly compared to other stress response genes (Milse *et al.*, 2014). IHF is a histone like DNA-binding and -bending protein which plays a role in DNA replication, site-specific recombination and gene expression. It is a global regulator as it alters the expression of numerous genes (Freundlich *et al.*, 1992). IHF is required for OxyR to induce *sufS*, as IHF presumably mediates by binding and bending the DNA to allow OxyR interaction with the promoter complex (Lee *et al.*, 2004).

1.3.4 RyhB and FnrS

In addition to transcriptional regulation, post-transcriptional regulation of *iscRSUA* enables differential expression of IscR and IscSUA in *E. coli* (Figure 4). Under iron-replete conditions, holo-Fur represses *ryhB* by binding within its promoter region (Figure 4 (10)). RyhB is a small non-coding regulatory RNA and a global post-transcriptional regulator which causes degradation of specific mRNA transcripts (Massé *et al.*, 2005). When iron becomes limiting, Fur dissociates from the

promoter and RyhB synthesis is de-repressed (Massé and Gottesman, 2002) (Figure 4 (11)). RyhB in turn base-pairs with the ribosomal binding site (RBS) of *iscS* mRNA between *iscR* and *iscS* (Figure 4 (12)), recruiting RNA degrading enzymes that cleave the polycistronic *iscRSUA* mRNA upstream of the RyhB binding site. The *iscSUA* mRNA is degraded while the upstream *iscR* region is not because it is protected by secondary structure in the intergenic region (Figure 4 (13)). RyhB thus promotes the expression of *iscR* transcript while reducing the *iscSUA* transcript levels. Translation of the *iscR* transcript still occurs and IscR is produced in its apo-form (Figure 4 (14)) (Desnoyers *et al.*, 2009). Thus, even though *iscRSUA* transcript levels are increased under iron-limiting conditions, post-transcriptional degradation of the *iscSUA* and not *iscR* enables continued translation of only *iscR* producing apo-IscR (Figure 4 (14)). Apo-IscR in turn activates *suf* operon expression which produces Fe-S clusters and eventually restores the holo-IscR levels (Figure 4 (15)) enabling repression of Fe-S cluster biogenesis as Fe-S homeostasis is reached (Figure 4 (1)).

FnrS is another small regulatory RNA which acts as global post-transcriptional regulator causing down-regulation of mRNA expression. Expression of FnrS is upregulated under anaerobic conditions through the fumarate nitrate reductase (FNR) transcriptional regulator (Durand and Storz, 2010). FnrS interacts with the *iscRSUAB* mRNA transcript base-pairing with the region upstream of *iscR* presumably down-regulating the *isc* system expression (Wright *et al.*, 2013).

1.3.5 SufR

SufR belongs to the DeoR family of transcriptional regulators containing a helix-loop-helix DNAbd near the N-terminus (Wang *et al.*, 2004). SufR homologues have been identified in many pathogenic bacteria, but are not present in all *suf* system containing bacteria. The *suf* system in some Gram-positive bacteria, specifically of the Firmicutes phylum, does not include a *sufR* gene even though the *suf* system is the only Fe-S cluster biogenesis system in these bacteria (Albrecht *et al.*, 2010; Riboldi *et al.*, 2009) (Figure 3). It is not known which regulators have taken over its function. *SufR* can either form part of the *suf* operon, be divergently transcribed from the rest of the *suf* operon or be present at a separate position within the genome completely separated from the rest of the *suf* operon (Wang *et al.*, 2004).

SufR homologues have four conserved cysteine residues in a C-X_{12}-C-X_{13}-C-X_{14}-C motif near the C-terminus, three of which are involved in Fe-S cluster coordination (C164, C171 and C206) in cyanobacteria, suggesting that another residue or exogenous molecule acts as the fourth ligand for cluster coordination (Shen *et al.*, 2007; Wang *et al.*, 2004). Both apo- and holo-SufR proteins form

homodimers, and each homodimer coordinates two oxygen sensitive $[4Fe\text{-}4S]^{+2,+1}$clusters (Shen *et al.*, 2007; Wang *et al.*, 2004).

In cyanobacteria, *suf* system expression (*sufB* and *sufS*) is increased under oxidative stress conditions, though *sufR* expression itself is not regulated by oxidative stress or iron starvation (Shen *et al.*, 2007). The DNA-binding affinity of SufR depends on the presence and redox state of the SufR Fe-S cluster, with apo-SufR and reduced holo-SufR ($[4Fe\text{-}4S]^{1+}$) binding weakly and holo-SufR ($[4Fe\text{-}4S]^{2+}$) binding with a high affinity to the *suf* promoter region (Shen *et al.*, 2007). The intergenic region between *sufR* and *sufBCDS* in cyanobacteria contains two *SufR* binding sites. Both sites possess identical motifs containing inverted repeats, with one of the motifs being flanked by an extended palindromic sequence. Holo-SufR binds with high affinity to the site with the extended palindromic flanking sequence controlling *sufBCDS* expression, and with low affinity to the shorter inverted repeat motif controlling *sufR* gene expression. The difference in DNA-binding affinity allows holo-SufR to repress *sufBCDS* in response to changing environmental conditions and autoregulate its own expression (Shen *et al.*, 2007; Wang *et al.*, 2004). Under stress conditions the capacity for Fe-S cluster production decreases, changing the redox state of SufR or causing cluster loss, thereby lowering the affinity for binding within the *suf* promoter. De-repression of *sufBCDS* occurs until Fe-S cluster homeostasis is restored. A cluster is then inserted into SufR and holo-SufR binds to the *suf* promoter and blocks *sufBCDS* transcription (Shen *et al.*, 2007).

1.3.6 SigM and DtxR

SigM is an alternative sigma factor of the extracytoplasmic function sigma factor subfamily and is involved in stress response. SigM expression is upregulated upon heat- and cold shock and disulphide stress and in turn activates the *suf* operon in response to disulphide stress (Nakunst *et al.*, 2007). *SigM* is expressed from a SigH dependent promoter and causes an indirect downregulation of the *suf* operon in a *sigH* deletion mutant (Ehira *et al.*, 2009). The sequences recognized by SigM and SigH are very similar and it was suggested that they can recognize each other's binding sites depending on certain conditions, making these promoters dual responsive (Pátek and Nešvera, 2011). Interestingly, diamide stress led to a switch in *suf* operon promoter usage as two different *suf* mRNA transcripts were detected under standard culture conditions versus diamide exposed conditions (Nakunst *et al.*, 2007). DtxR in *C. glutamicum* is an IdeR orthologue and is a global iron responsive regulator of genes involved in iron metabolism. The *suf* system expression is decreased in a *dtxR* deletion mutant suggesting that DtxR could act as an activator of the *suf* operon. This regulation is likely direct since DtxR binds upstream of *sufR* (Wennerhold and Bott, 2006).

1.4 Regulation of Fe-S cluster biogenesis in *M. tuberculosis*

Several studies have demonstrated differential expression of the *suf* operon in *M. tuberculosis* in response to various stresses. Genes from the *suf* operon were upregulated by ROS and RNS (Schnappinger *et al.*, 2003; Voskuil *et al.*, 2011). The *suf* operon genes were induced upon infection of naive and activated macrophages (Rohde *et al.*, 2007; Schnappinger *et al.*, 2003), suggesting a role for the *suf* operon genes in sustained intracellular survival of *M. tuberculosis*. The *suf* operon was also regulated in response to starvation (Betts *et al.*, 2002) and induced by iron limitation (Rodriguez *et al.*, 2002) and upon entry into anaerobic stationary phase (Murugasu-Oei *et al.*, 1999). Rv1460 is induced in bacteria present in sputum samples from pulmonary TB patients, suggesting a role during infection (Kumar *et al.*, 2011). The role of the *suf* system in mycobacterial physiology is however not well understood.

Although regulation of Fe-S cluster biogenesis has been studied extensively in other bacteria, not much is known about this process in *M. tuberculosis*. Many of the known regulators of Fe-S cluster biogenesis in other bacteria are not involved in *suf* system regulation in *M. tuberculosis*. IscR is responsible for switching between the *isc* and *suf* systems upon exposure to stress conditions in other bacteria, but since no switching between systems is needed in *M. tuberculosis*, it is not unexpected that no IscR homologue is present in mycobacteria. FurA has a more specialised function in *M. tuberculosis* since it regulates expression of *katG*, which encodes catalase-peroxidase (KatG) (Pym *et al.*, 2001), as well as its own expression, although it has been suggested that it could regulate other genes (Sala *et al.*, 2003). FurA has however not been directly linked to the regulation of the *suf* operon in *M. tuberculosis* and no FurA binding site within the promoter of the *suf* operon was identified during ChIP-seq analyses (Minch *et al.*, 2015). Like *Rv1460*, *furA* expression is induced by ROS and RNS (Schnappinger *et al.*, 2003; Voskuil *et al.*, 2011) and within macrophages (Schnappinger *et al.*, 2003). *FurA* expression is also upregulated in *M. tuberculosis* under iron-replete conditions and virtually absent under iron-limiting conditions (Wong *et al.*, 1999), indicating a link with iron responsiveness.

O*xyR* is a pseudogene and inactivated in *M. tuberculosis*, with other regulators presumably taking over its regulon (Deretic *et al.*, 1997). SigM has also not been linked to the regulation of the *suf* operon in *M. tuberculosis* although another sigma factor, SigK has been predicted to bind within the *Rv1460* promoter region (Guo *et al.*, 2009). The regulation of the *suf* operon genes in response to iron are also not dependent on IdeR (a DtxR homologue) in *M. tuberculosis* (Rodriguez *et al.*, 2002) and no IdeR binding site has been identified upstream of the *suf* operon (Prakash *et al.*, 2005), indicating

that IdeR is not linked to the direct regulation of the *suf* operon in *M. tuberculosis*. A SufR homologue was however identified in *M. tuberculosis* and predicted to be involved in the regulation of Fe-S cluster biogenesis (Huet *et al.*, 2005; Shen *et al.*, 2007).

1.5 Rv1460, a possible *suf* operon regulator in *M. tuberculosis*

Rv1460 is a cyanobacterial SufR homologue and has a predicted winged helix-turn-helix DNAbd near its N-terminus, suggesting specificity and affinity for DNA-binding. Rv1460 contains the three conserved cysteine residues, C203, C216 and C244, in *M. tuberculosis* and may bind an Fe-S cluster (Cole *et al.*, 1998; Shen *et al.*, 2007; Wang *et al.*, 2004). The presence and role of an Fe-S cluster in Rv1460 have not been investigated.

Rv1460 may repress the *suf* system since *Rv1460* overexpression led to decreased *Rv1461-Rv1466* expression (although this decrease was not significant for all of the genes) (Rustad *et al.*, 2014). Correspondingly, ChIP-seq experiments indicated binding sites for Rv1460 within the *Rv1460* promoter region as well as within *Rv1461* upstream of the intein (Minch *et al.*, 2015). The implications of two independent binding sites within the *suf* operon have not been investigated, but could indicate that Rv1460 binding to the respective sites drive differential regulation of *Rv1460* expression versus the expression of the rest of the downstream operon (*Rv1461-Rv1466*).

1.6 Rv1460 could function as a global regulator

Several Fe-S cluster biogenesis regulators are also global regulators; due iron metabolism being intertwined with many other biological processes. Rv1460 may similarly function as a global regulator as Rv1460 binding sites within the promoter regions of multiple genes were identified by ChIP-seq. Genes with Rv1460 binding sites in their promoter region are *Rv0456A*, *Rv1116*, *Rv2107* and *Rv3597c* (Minch *et al.*, 2015). *Rv0456A* encodes a possible mazF1 toxin and forms part of toxin-antitoxin operon with *Rv0456B*. Toxin-antitoxin systems are stress-response elements that help bacteria survive under unfavourable growth conditions (Zhu *et al.*, 2006). *Rv1116* encodes a hypothetical protein of unknown function (Lew *et al.*, 2011). Rv2107 (PE22) is a member of mycobacterial PE family (Lew *et al.*, 2011). Although the specific functions of most PE proteins are unknown, they have been implicated in the modulation of host immune response (Sampson, 2011). Interestingly, a probable binding site for SigK was identified in the *Rv1116* and *Rv2107* promoter regions (Rodrigue *et al.*, 2007) while SigK was also predicted to bind within the *Rv1460* promoter

(Guo *et al.*, 2009). This suggests that the *suf* operon and *Rv1116* and *Rv2107* genes can be co-regulated by SigK.

Rv3597c encodes Lsr2, a histone-like nucleoid structuring protein involved in regulation of gene expression by controlling chromosomal topology. It regulates a wide range of DNA-interacting enzymes and genes involved in several pathways by binding to an AT-rich sequence (Colangeli *et al.*, 2007). Lsr2 protects *M. tuberculosis* against ROS by binding to and physically shielding the DNA (Colangeli *et al.*, 2009). Lsr2 protein expression is upregulated under high iron conditions (Wong *et al.*, 1999) and in response to starvation (Betts *et al.*, 2002).

Rv1460 overexpression did not significantly affect the expression of any of the aforementioned genes under high aeration culture conditions (Rustad *et al.*, 2014), although this may be due to the fact that Rv1460 expression was only threefold higher. However, *Rv2824c*, encoding a hypothetical protein of unknown function (Galagan *et al.*, 2010; Reddy *et al.*, 2009), was upregulated by *Rv1460* overexpression, although its regulation is probably not direct as no Rv1460 binding site was identified in its promoter (Rustad *et al.*, 2014).

1.7 Other Fe-S cluster biogenesis regulators in *M. tuberculosis*

In addition to the possible local regulation of the *suf* operon by Rv1460, global regulators are predicted to have a binding site within the *suf* operon, which suggests possible regulation of the *suf* operon by these regulators in *M. tuberculosis*. This is not surprising as global regulation of the *suf* operon could enable faster response times and the ability to respond to a broader range of environmental conditions. Seven possible global regulators that bind within the 297 bp region upstream of *Rv1460* were identified by bacterial one-hybrid screens (Guo *et al.*, 2009). Three other possible regulators that bind within the region -150 to +70 of *Rv1460* have been identified by ChIP-seq analysis (Minch *et al.*, 2015). They are listed in Table 1 and include global regulators involved in metabolism, cell wall biosynthesis, metal cofactor biosynthesis, intracellular survival and virulence. Interestingly, there was no overlap in the regulators identified by the two studies indicating the importance of the methodology used to predict regulator binding sites and confirmation of these binding sites by binding assays. Overexpression of the predicted possible regulators did not affect the expression of the *suf* operon significantly and their possible role in regulation of the *suf* operon thus remains unclear (Rustad *et al.*, 2014).

1.8 Future prospects

Fe-S cluster biogenesis is a key part of metabolism and potentially allows *M. tuberculosis* to adapt to intracellular conditions. It is therefore important that this process be tightly regulated. Although Fe-S cluster biogenesis and the regulation thereof are well studied in other bacteria, it is an understudied area of *M. tuberculosis* metabolism. It is tempting to extrapolate the function of Rv1460 from its SufR homologue, but its precise function in Fe-S cluster biogenesis needs to be established. It is therefore important to determine if Rv1460 regulates the *suf* operon and if so, whether the regulation is direct. The presence of a Rv1460 Fe-S cluster and the possible effect of cluster coordination on DNA-binding need to be investigated. Could the complex regulation of the *suf* operon by SufR in cyanobacteria, driven by the cluster presence and state, be similarly true of Rv1460's regulation of the *suf* operon in *M. tuberculosis*? Understanding Fe-S cluster biogenesis regulation could provide insight into the physiology and pathogenesis of this deadly pathogen.

Table 1. Possible *suf* operon global regulators

Regulator	Name	Function and expression conditions	Position of predicted binding site	Method of prediction
Rv0081	-	Transcriptional regulator of the ArsR family (Galagan *et al.*, 2010; Reddy *et al.*, 2009) Induced upon nitric oxide (NO) exposure as part of the initial hypoxic response and the DosR regulon (Voskuil *et al.*, 2003) Represses its own expression and that of downstream genes encoding components of the formate hydrogenylase enzyme complex (*Rv0081-Rv0088*) following activation by DosR in response to NO (He *et al.*, 2011) Induced during infection of macrophage-like THP-1 cells (Fontán *et al.*, 2008) Induced response to lung surfactant (Schwab *et al.*, 2009)	Within promoter of *Rv1463*	ChIP-seq (Minch *et al.*, 2015)
Rv0260c	-	Transcriptional regulator with a uroporphyrinogen-III synthase domain (Galagan *et al.*, 2010; Reddy *et al.*, 2009) Suggested to be involved in the tetrapyrrole biosynthesis I pathway.Tetrapyrroles are metal-binding cofactors (such as heme and siroheme) formed by methylating or decarboxylating uroporphyrinogen-III (Galagan *et al.*, 2010; Reddy *et al.*, 2009) Induced in naive and activated macrophages (Schnappinger *et al.*, 2003)	Within promoter of *Rv1460*	Bacterial one-hybrid screen (Guo *et al.*, 2009)
Rv0445c	SigK	One of thirteen sigma factors, a dissociable subunit of RNA polymerase (Sachdeva *et al.*, 2010) Global regulator involved in response to specific environmental signals (Sachdeva *et al.*, 2010) Induced as part of the initial hypoxic response (Rustad *et al.*, 2008) Induced in macrophages (Saïd-Salim *et al.*, 2006)	Within promoter of *Rv1460*	Bacterial one-hybrid screen (Guo *et al.*, 2009)
Rv0678	-	MarR family transcriptional regulator (Radhakrishnan *et al.*, 2014) Regulates expression of mmpS5-mmpL5 (mycobacterial membrane protein small and large) efflux system (Milano *et al.*, 2009; Radhakrishnan *et al.*, 2014), involved in cell wall lipid biosynthesis and virulence (Delmar *et al.*, 2015) Mutations in *Rv0678* associated with azole resistance (Milano *et al.*, 2009)	Within promoter of *Rv1463* and *Rv1466*	ChIP-seq (Minch *et al.*, 2015)
Rv0691c	-	TetR transcriptional regulator Probable mycobactocin system regulator, as it occurs in genome context near mycofactocin system (Galagan *et al.*, 2010; Reddy *et al.*, 2009)	Within promoter of *Rv1460*	ChIP-seq (Minch *et al.*, 2015)
Rv0818	GlnR	Global transcriptional regulator (Jenkins *et al.*, 2013) Involved in nitrogen metabolism in *M. smegmatis* as it activates nitrite reductase expression responsible for second step in nitrate assimilation (Jenkins *et al.*, 2013; Malm *et al.*, 2009) Nitrate reductase in other bacteria are siroheme-dependent (Malm *et al.*, 2009)	Within promoter of *Rv1460*	Bacterial one-hybrid screen (Guo *et al.*, 2009)

Continued overleaf

Table 1. Possible *suf* operon global regulators (Continued)

Regulator	Name	Function and expression conditions	Position of predicted binding site	Method of prediction
Rv0967	CsoR	Copper specific copper-responsive transcriptional regulator (Ma *et al.*, 2009) Binds DNA in its copper bound form to de-repression expression of copper chaperones and efflux systems (Liu *et al.*, 2007)	Within promoter of *Rv1461*	ChIP-seq (Minch *et al.*, 2015)
Rv1033c	TrcR	Response regulator in two-component system (TrcRS) TrcS encodes for a histidine kinase Autoactivates its own expression (Haydel *et al.*, 2002) Global regulator (Haydel and Clark-Curtiss, 2006) Induced upon initial and prolonged macrophage infection (Haydel *et al.*, 2002)	Within promoter of *Rv1460*	ChIP-seq (Minch *et al.*, 2015)
Rv1353c	-	TetR transcriptional regulator (Galagan *et al.*, 2010; Reddy *et al.*, 2009) Part of large gene deletions associated with Haarlem strains (Cubillos-Ruiz *et al.*, 2010)	Within promoter of *Rv1463*	ChIP-seq (Minch *et al.*, 2015)
Rv1359	-	Transcriptional regulator involved in metabolism (Galagan *et al.*, 2010; Reddy *et al.*, 2009) Contains an adenylate cyclase domain. Adenylate cyclases catalyze the synthesis of cAMP from ATP (Flores-Valdez *et al.*, 2012). Induced after 24 h starvation and remains upregulated at 96 h (Betts *et al.*, 2002) Induced during macrophage infection (Homolka *et al.*, 2010) Expression affected by iron (Galagan *et al.*, 2010; Reddy *et al.*, 2009)	Within promoter of *Rv1460*	Bacterial one-hybrid screen (Guo *et al.*, 2009)
Rv1816	-	Binds saturated fatty acids (Delmar *et al.*, 2015) Fatty acids binding changes DNA-binding causing de-repression of *mmpL* expression encoding mycobacterial membrane protein large (MmpL) involved in cell wall lipid biosynthesis and virulence (Delmar *et al.*, 2015).	Within promoter of *Rv1461*	ChIP-seq (Minch *et al.*, 2015)
Rv1931	-	AraC family transcriptional activator (Frota *et al.*, 2004) Increase expression in virulent versus avirulent strain (Gao *et al.*, 2004) Involved in survival in macrophages and in mouse infection (Frota *et al.*, 2004)	Within promoter of *Rv1460*	Bacterial one-hybrid screen (Guo *et al.*, 2009)
Rv2610c	PimA	α-mannosyltransferase Involved in phosphatidylinositol mannosides production (such as lipomannan and lipoarabinomannan) which are cell wall components (Korduláková *et al.*, 2002)	Within promoter of *Rv1460*	ChIP-seq (Minch *et al.*, 2015)
Rv3416	WhiB3	Global transcriptional regulator (Singh *et al.*, 2009) Contains an oxygen and NO sensitive Fe-S cluster (Singh *et al.*, 2007) Fe-S cluster presence and state changes binding to DNA (Singh *et al.*, 2009) Acts as a protein disulphide reductase, only in its clusterless form (Alam and Agrawal, 2008) Involved in metabolic switching upon infection (Singh *et al.*, 2009, 2007) Involved in virulence and reactivation (Banaiee *et al.*, 2006; Singh *et al.*, 2009) Involved in quorum sensing (Banaiee *et al.*, 2006)	Within promoter of *Rv1460*	Bacterial one-hybrid screen (Guo *et al.*, 2009)

Continued overleaf

Table 1. Possible *suf* operon global regulators (Continued)

Regulator	Name	Function and expression conditions	Position of predicted binding site	Method of prediction
Rv3557c	KstR2	TetR transcriptional regulator Repressor of the genes involved in cholesterol metabolism (Kendall *et al.*, 2010) Responds to the presence of cholesterol with de-repression of its regulon in the presence of cholesterol (Kendall *et al.*, 2010) Induced during macrophage infection (Schnappinger *et al.*, 2003)	Within promoter of *Rv1460*	Bacterial one-hybrid screen (Guo *et al.*, 2009)

2. Hypotheses, aims and objectives

Hypothesis 1: Rv1460 is involved in regulating the *suf* operon and important in *M. tuberculosis* physiology.

Aim 1: To determine whether Rv1460 is involved in regulating the *suf* operon expression and in *M. tuberculosis* physiology.

Objective 1: To generate three *M. tuberculosis* H37Rv mutants harbouring distinct deletions in the *Rv1460* gene and determine the impact of the *Rv1460* deletions on bacterial viability and sensitivity to oxidative stress and iron limitation and on the expression of the *suf* operon.
Approach: Two-step allelic exchange will be used to generate *Rv1460* deletions by homologous recombination. Growth of the mutants will be monitored at optical density (OD_{600nm}) under standard culture conditions or by determining survival by colony forming units (CFUs) during iron limitation and after exposure to an oxidative stress generator. Mutants will be cultured under standard culture conditions, total RNA extracted at mid-log growth and RT-qPCR done to determine the level of *suf* operon expression.

Hypothesis 2: Rv1460 binds directly to the *suf* operon promoter region.

Aim 2: To determine whether Rv1460 binds directly within the *suf* operon promoter.

Objective 2: To produce and purifying recombinant Rv1460 and investigate the binding of Rv1460 to DNA using electrophoretic mobility shift assay (EMSA).
Approach: Vectors for the production of Rv1460 as a His-tag fusion protein will be generated and the Rv1460 protein will be produced in *E. coli*. Recombinant Rv1460 will be purified by nickel affinity chromatography followed by gel filtration and the His-tag removed by thrombin cleavage. Binding of the Rv1460 protein to the *Rv1460* promoter region and within *Rv1461* will be investigated as binding sites for Rv1460 were predicted to be present within the *Rv1460* promoter and within *Rv1461* by ChIP-seq (Minch *et al.*, 2015).

Hypothesis 3: Rv1460 is able to coordinate an Fe-S cluster via its conserved cysteine residues.

Aim 3: To determine whether Rv1460 coordinates an Fe-S cluster and which residues are involved in this coordination.

Objective 3: To reconstitute an Fe-S cluster on recombinant Rv1460 protein *in vitro* and investigate the role of three conserved cysteine residues in coordination of the Rv1460 Fe-S cluster.

Approach: Fe-S cluster reconstitution will be done by adding a cysteine desulphurase, scaffold protein, cysteine and iron to purified Rv1460 under reducing, anaerobic conditions and monitoring the change in spectrum over a range of wavelengths by UV-visible spectroscopy. The cysteine residues, predicted to coordinate the Fe-S cluster in Rv1460 (Shen *et al.*, 2007), will be replaced with serine residues and the ability of the protein variants to coordinate an Fe-S cluster determined by Fe-S cluster reconstitution and UV-visible spectroscopy.

3. Methods

3.1 Bacterial strains and culture conditions

M. tuberculosis H37Rv was cultured in 7H9 broth (Difco) with 0.05% Tween 80 supplemented with 0.2% glycerol and albumin dextrose catalase (ADC) (Bovine Albumin fraction V (50 g/L), Dextrose (20 g/L) and catalase (0.0375%)) or Middlebrook Oleic albumin dextrose catalase (OADC) or on Middlebrook 7H10 (Difco) supplemented with ADC or OADC. Kanamycin (kana) (50 µg/ml), hygromycin (hyg) (50 µg/ml), 5-bromo-4 chloro-3 indolyl β-D-galactosidase (X-gal) (40 µg/ml) and sucrose (5%) were used for selection purposes where applicable. Growth was monitored by OD_{600nm} readings or by determining CFUs after plating serial dilutions on supplemented Middlebrook 7H10 media. The *E. coli* XL1 Blue strain was used for all cloning purposes. Culturing was done in Luria-Bertani (LB) broth and selection done using ampicillin (amp) (100 µg/ml), kana (50 µg/ml), hyg (150 µg/ml) and X-gal (40 µg/ml) when applicable. *E. coli* strain Arctic express (DE3) was cultured in LB broth and gentamicin (genta) (20 µg/ml) to maintain the endogenous plasmid. *E. coli* strain Rosetta 2 pLysS (DE3) was cultured in LB or terrific broth (tryptone (12 g/L), yeast extract (24 g/L), glycerol (0.004%), 0.017 M KH_2PO_4 and 0.072M K_2HPO_4) and chloramphenicol (Cm) (34 µg/ml) to maintain the endogenous plasmid. *E. coli* strain Rosetta-gami 2 (DE3) was cultured in LB broth with Cm (34 µg/ml), tetracycline (tet) (12.5 µg/ml) and streptomycin (strep) (50 µg/ml) to maintain the endogenous plasmids. The bacterial strains used and generated during this study are listed in Table 2.

3.2 Cloning

Electrocompetent or rubidium chloride competent *E. coli* was prepared as described (Sambrook *et al.*, 1998; Sanchez-Puelles *et al.*, 1986). For ligations either Promega or Roche T4 DNA ligase overnight at 4 °C were used according to manufacturer's instructions. Vector DNA was purified using the NucleoBond PC100 Plasmid DNA Purification kit (Macherey-Nagel) and Wizard Plus SV Minipreps DNA Purification System (Promega) according to the manufacturer's instructions. DNA fragments from PCR reactions or gels were purified using the Wizard SV Gel and PCR Clean-Up System (Promega). All PCR products were cloned into sequencing vectors, pJET1.2 or pGEMTeasy, and sequenced using commercially available sequencing primers for these vectors (pJET1.2For/Rev or pGEMTeasyT7/SP6) before restriction enzyme (RE) digestion (enzymes listed in brackets in the

Tables) and sub-cloning. The plasmids used and generated in this study are indicated in Table 3. Marker IV (Roche) or 1 kbp plus DNA ladder (Fermentas) were used as DNA ladders.

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

Bacterial strains	Description	Source/ reference
M. tuberculosis		
H37Rv	H37RvMA (ATCC: 27294)	(Ioerger *et al.*, 2010)
Δ*Rv1460*:: pMVΔ*Rv1460compl*	Derivative of H37Rv carrying a 528 bp unmarked in frame deletion in *Rv1460* and the pMV*Rv1460compl* vector integrated into the mycobacterial *attB* site (Hyg^R)	This study
*Rv1460*Δ*DNAbd*:: pMV*Rv1460compl*	Derivative of H37Rv carrying a 237 bp unmarked in frame deletion of the predicted *Rv1460* DNAbd and the pMV*Rv1460compl* vector integrated into the mycobacterial *attB* site (Hyg^R)	This study
Rv1460stop	Derivative of H37Rv carrying a 236 bp unmarked deletion of the predicted *Rv1460* DNAbd and a frame-shift causing a premature stop and truncation of the Rv1460 protein	This study
Rv1460stop:: pMVΔ*int1460*	Derivative of *Rv1460stop* with the pMVΔ*int1460* vector stably integrated into the mycobacterial *attB* site (Hyg^R)	This study
E. coli		
XL1Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F´ *proAB lacI*^{q}*ZΔM15* Tn*10*] (Tet^R)	Stratagene
Arctic express (DE3)	B F^- *ompT hsdS*(r_B^- m_B^-) *dcm*$^+$ Tet^r *gal* λ(DE3) *endA* Hte [*cpn10 cpn60*] ($Genta^R$)	Agilent Technologies
Rv1460NpET28	Derivative of Arctic express (DE3) containing the pET28*Rv1460*N plasmid for production of Rv1460 as a His-tagged protein ($Genta^R$, $Kana^R$)	This study
Rv1460 C203S	Derivative of Arctic express (DE3) containing the pET28*Rv1460*C203S plasmid for production of Rv1460 C203S as a His-tagged protein ($Genta^R$, $Kana^R$)	This study
Rv1460 C216S	Derivative of Arctic express (DE3) containing the pET28*Rv1460*C216S plasmid for production of Rv1460 C216S as a His-tagged protein ($Genta^R$, $Kana^R$)	This study
Rv1460 C242S	Derivative of Arctic express (DE3) containing the pET28*Rv1460*C242S plasmid for production of Rv1460 C242S as a His-tagged protein ($Genta^R$, $Kana^R$)	This study
Rv1460 C244S	Derivative of Arctic express (DE3) containing the pET28*Rv1460*C244S plasmid for production of Rv1460 C244S as a His-tagged protein ($Genta^R$, $Kana^R$)	This study
Rv1460 C203,216,244S	Derivative of Arctic express (DE3) containing the pET28*Rv1460*C203,216,244S plasmid for production of Rv1460 C203,216,244S as a His-tagged protein variant ($Genta^R$, $Kana^R$)	This study
Rosetta 2 (DE3) pLysS	F^- *ompT hsdS*$_B$(r_B^- m_B^-) *gal dcm* (DE3) pLysSRARE2 (Cm^R)	Novagen
Rosetta-gami 2 (DE3)	Δ(*ara-leu*)*7697* Δ*lacX74* Δ*phoA Pvu*II *phoR araD139 ahpC galE galK rpsL* (DE3) F'[*lac*$^+$ *lacI*^{q} *pro*] *gor522*::Tn*10 trxB* pRARE2 (Cm^R, $Strep^R$, Tet^R)	Novagen

Table 3. List of plasmids used and generated during this study

Plasmids	Description	Source/ reference
Sequencing:		
pJET1.2	Linearized blunt end cloning vector (Amp^R)	CloneJet
pGEM-T-easy	Linearized vector with 3´-T overhangs with a multiple cloning region within the α-peptide coding region of the enzyme β-galactosidase allowing blue/white screening for transformants on indicator plates (Amp^R)	Promega
Mutant generation:		
p2NIL	Cloning vector ($Kana^R$)	(Parish and Stoker, 2000)
pGOAL17	Plasmid with the *lacZ* and *sacB* genes flanked by two PacI restriction sites (Amp^R)	(Parish and Stoker, 2000)
p2NIL17ΔRv1460	Suicide delivery vector containing the *upstream* and *downdel* regions (Table 4) and *lacZ* and *sacB* genes from pGOAL17, $Kana^R$	This study
p2NIL17Rv1460ΔDNAbd	Suicide delivery vector containing the *upstream* and *downDNAbddel* regions (Table S2) and *lacZ* and *sacB* genes from pGOAL17 ($Kana^R$)	This study
p2NIL17Rv1460stop	Suicide delivery vector containing the *upstream* and *downDNAbddelstop* regions (Table 4) and *lacZ* and *sacB* genes from pGOAL17 ($Kana^R$)	This study
Complementation:		
pMV306H	Mycobacterial integrating shuttle vector pMV306 (Stover *et al.*, 1991) derivative containing a hyg resistance gene (Hyg^R)	H. Boshoff
pMVsuf01	Derivative of pMV306H containing a 14.7 kbp fragment that includes the *M. tuberculosis suf* operon. The fragment was obtained by digesting BAC-Rv30 (Brosch *et al.*, 1998) with EcoRVand HindIII (Hyg^R)	M. Williams
pMV*Rv1460compl*	Derivative of pMV306H containing *Rv1460compl* (Table 4) (Hyg^R)	This study
pMV*Δint*1460	Derivative of pMV*Rv1460compl* which has a 926 bp region of the integrase gene between the PstI and AvrII restriction sites removed by restriction enzyme digestion and re-ligation of the blunted vector, creating stable integration into the mycobacterial *attB* site (Springer *et al.*, 2001) (Hyg^R)	This study
pBluescriptint (pBSint)	Used in combination with pMV*Δint1460* to provide integrase for integration at the *attB* site. This vector does not have a mycobacterial origin of replication and is therefore lost upon replication thereby removing the integrase and ensuring stable integration of the pMV*Δint1460* vector when it is co-transformed	(Springer *et al.*, 2001) S. Sampson
Rv1460 production:		
pJET*Rv1460*N	Derivative of pJET1.2 containing the *Rv1460*N region (Table 7) cloned into the blunt end cloning position	This study
pET28a	*E. coli* protein production plasmid with a T7 *lac* promoter and a multiple cloning site for creating N- or C- terminally His-tag fusion protein with a thrombin cleavage site for the removal of the N-terminal His-tag ($Kana^R$)	Novagen
pET28*Rv1460*N	Derivative of pET28a with the *Rv1460*N fragment (Table 7) cloned into the multiple cloning site between NdeI and HindIII cut sites	This study

Continued overleaf

Table 3. List of plasmids used and generated during this study (Continued)

Plasmids	Description	Source/ reference
pET28*Rv1460*C203S	Derivative of pET28a with Rv1460 containing a C203S mutation cloned into the multiple cloning site between NdeI and HindIII cut sites	This study
pET28*Rv1460*C216S	Derivative of pET28a with Rv1460 containing a C216S mutation cloned into the multiple cloning site between NdeI and HindIII cut sites	This study
pET28*Rv1460*C242S	Derivative of pET28a with Rv1460 containing a C242S mutation cloned into the multiple cloning site between NdeI and HindIII cut sites	This study
pET28*Rv1460*C244S	Derivative of pET28a with Rv1460 containing a C244S mutation cloned into the multiple cloning site between NdeI and HindIII cut sites	This study
pET28*Rv1460*C203,216,244S	Derivative of pET28a with Rv1460 containing C203S, C216S and C244S mutations cloned into the multiple cloning site between NdeI and HindIII cut sites	This study
pET21	*E. coli* protein production plasmid with a T7 *lac* promoter and a multiple cloning site for creating C- terminally His-tag fusion protein (Amp^R)	This study
pET21*Rv1460*C	Derivative of pET21 with the *Rv1460*C fragment (Table 7) cloned into the multiple cloning site between NdeI and HindIII cut sites	This study
Fe-S cluster reconstitution:		
pET28iscS	Derivative of pET28a with the *iscSN* fragment (Table 7) cloned into the multiple cloning site between NdeI and HindIII cut sites	This study
IscSpET	Plasmid for the production of *E. coli* IscS as a His-tagged protein, containing a TEV site enabling cleavage of the His-tag	(Prischi *et al.*, 2010) Prof A Pastore, King's College UK
IscUpET	Plasmid for the production of *E. coli* IscU as a His and GST-tagged protein containing a TEV site enabling cleavage of both tags	(Prischi *et al.*, 2010) Prof A Pastore, King's College UK

3.3 Generation of *M. tuberculosis* mutants harbouring unmarked deletions in *Rv1460*

3.3.1 Generation of suicide delivery vectors

Suicide delivery vectors for the generation of *Rv1460* deletion mutants were generated by amplifying the *Rv1460 upstream*, *downdel*, *downDNAbddel* and *downDNAbddelstop* regions from H37Rv genomic DNA using the primers and strategy indicated in Table 4 (Phusion HotStart DNA polymerase (Finnzymes): initial denaturation at 98 °C for 1 minute (min), amplification for 30 cycles at 98 °C for 15 s, 60 °C for 30 s, 72 °C for 45 s, final extension at 72 °C for 10 min) and sub-cloned

into the p2NIL vector as indicated in Table 4. Selectable markers from the pGOAL17 vector were ligated into the PacI site. The final vectors were Sanger sequenced using primers indicated in Table 4.

3.3.2 Generation of unmarked Rv1460 deletions by two-step allelic exchange

The suicide delivery vectors were used for the generation of three different *Rv1460* deletions (Figure 5); Δ*Rv1460*, which has an in frame deletion of the bulk of the *Rv1460* gene (codons 36 to 211 deleted), *Rv1460ΔDNAbd*, which has an in frame deletion of the predicted *Rv1460* DNAbd (codons 36 to 114 deleted) and *Rv1460stop* which has an out of frame deletion of the DNAbd (codon 36 to 114 deleted) and a premature stop codon causing truncation of the Rv1460 protein, were generated by two-step allelic exchange (Parish and Stoker, 2000) as described in Figure 6.

M. tuberculosis was transformed with the suicide delivery vectors (p2NIL17*ΔRv1460*, p2NIL17Rv1460ΔDNAbd and p2NIL17Rv1460stop (Table 3) and the two-step selection procedure done as previously described (Parish and Stoker, 2000). Selection was done on solid media containing kana and X-gal for a single homologous recombination event giving rise to single cross overs (SCOs) (Figure 6). The sucrose sensitivity of SCOs was tested by spotting re-suspended colonies on solid media containing 5% sucrose. The configuration of SCOs was determined by PCR screening of colonies (FastStart Taq DNA polymerase kit (Roche): initial denaturation at 95 °C for 4 min, amplification for 30 cycles at 95 °C for 45 s, 58 °C for 45 s, 72 °C for 1 min, final extension at 72 °C for 10 min) amplifying SCO region 1 and 2 using the primers indicated in Table 5 and separating PCR products by agarose gel electrophoresis. Both upstream and downstream SCOs, which were kana resistant, sucrose sensitive and blue on X-gal, were cultured and sub-cultured into media without kana and plated on 5% sucrose and X-gal plates without kana to allow for a second recombination event giving rise to double cross overs (DCOs) (Figure 6).

White colonies were screened by PCR for the deletion using the primers indicated in Table 5 (FastStart Taq DNA polymerase kit (Roche): initial denaturation at 95 °C for 4 min, amplification for 30 cycles at 95 °C for 30 s, 60 °C for 30 s, 72 °C for 30 s, final extension at 72 °C for 7 min). White colonies were re-suspended in 30 µl of water and boiled for 30 min to release DNA to use as PCR template. Southern blotting was done to confirm the presence of the *Rv1460* deletions in DCOs as previously described (Warren *et al.*, 2009). The presence of the frame-shift mutation in the *Rv1460stop* mutants was confirmed by Sanger sequencing using the primers indicated in Table 4.

Table 4. List of primers and strategy used for the generation of suicide delivery and complementation vectors

Region	Primer name	Sequence (5' → 3')[a]	Description[b]	Mutation[c]
Rv1460 upstream	Rv1460upF	GGATCCGGGTGAGTGACAACACG (BamHI)	1482 bp product, including a 1365 bp region upstream of *Rv1460* and 105 bp of the 5' end of *Rv1460* and 12 bp for RE cut sites	
	Rv1460upR	GGTACCGGGGACAGTGGTAGAGACC (Asp718)		
Rv1460 downdel	Rv1460delF	GGTACCTTCCCCGAATTGTGCG (Asp718)	1106 bp product, including 174 bp of the 3' end of *Rv1460* and 919 bp downstream of *Rv1460* and 13 bp for RE cut sites	Used in combination with the *Rv1460 upstream* region to delete 175 of the 268 amino acids of the Rv1460 protein (codons 36 to 211 deleted) (*ΔRv1460*)
	Rv1460delR	GAAGCTTGACATCGGTGTGAAGGTGAA (HindIII)		
Rv1460 downDNAbddel	Rv1460DNAbddelF	GGTACCAAGCTCGACCACTCCTAT (Asp718)	1397 bp product including 465 bp of the 3' end of *Rv1460* and 919 bp downstream of *Rv1460* and 13 bp for RE cut sites	Used in combination with the *Rv1460 upstream* region to delete part of the predicted DNAbd of Rv1460 (codons 36 to 113 deleted) (*Rv1460ΔDNAbd*)
	Rv1460delR	GAAGCTTGACATCGGTGTGAAGGTGAA (HindIII)		
Rv1460 downDNAbddelstop	Rv1460stopF[d]	GGTACC**C**AAGCTCGACCACTCCTAT (Asp718)	1398 bp product including 466 bp of the 3' end of *Rv1460* and 919 bp downstream of *Rv1460* and 13 bp for RE cut sits	Used in combination with Rv1460 *upstream* region to delete part of the predicted DNAbd of Rv1460 (codons 36 to 113 deleted) and introduce a premature stop codon at codon 121 (*Rv1460stop*)
	Rv1460delR	GAAGCTTGACATCGGTGTGAAGGTGAA (HindIII)		
Suicide delivery vectors	SeqF0	TACCGGCATAACCAAGCCTA	Sequencing final vectors	NA
Suicide delivery vectors	SeqF1	TGCCTGACTGCGTTAGCAAT	Sequencing final vectors	NA
Suicide delivery vectors	SeqF2	TGTCCTCTGGCAAATCATCC	Sequencing final vectors	NA
Suicide delivery vectors	SeqF3	TGATCGGTCTACCAGGGATG	Sequencing final vectors	NA
Suicide delivery vectors	SeqF3b	ATTTGCACACCGTGAATCG	Sequencing final vectors	NA

Continued overleaf

Table 4. List of primers and strategy used for generation of suicide delivery and complementation vectors (Continued)

Region	Primer name	Sequence (5' → 3')[a]	Description[b]	Mutation[c]
Suicide delivery vectors	SeqF4	CATATGGTCCGCTGTTCTTGT	Sequencing final vectors	NA
Rv1460compl	ComplF ComplR	GGATATCAGCTCGGTAGTGGTCAGCG (EcoRv) GAAGCTTCCTCTGGGGTGAGTGTCATC (HindIII)	975 bp product including 139 bp upstream of *Rv1460*, presumably including the promoter region, up to 15 bp downstream of *Rv1460* and 14 bp for RE cut sites	NA

[a] RE cut sites are underlined and RE indicated in parenthesis

[b] Annotated according to the annotated start site of *Rv1460* specified on TBDB and Tuberculist

[c] For a description of the *Rv1460* mutants generated refer to Figure 5

[d] Extra nucleotide causing frame-shift mutation indicated in bold

Table 5. List of primers and strategy for SCO configuration screening and genotyping of DCOs

Region	Primer name	Sequence (5' → 3')	Description	Expected product sizes[a]				SCO configuration[a]	
				WT	***ΔRv1460***	***Rv1460 ΔDNAbd***	***Rv1460stop***	**Upstream SCO**	**Downstream SCO**
SCO region 1	SCOupup SCOindel	CAATAACAGCCAGCACACCA GCAGTCTCCGTTGACGATG	Binds 49 bp upstream of *upstream* region and within *downdel* region	2142 bp	1624 bp	1915 bp	1916 bp	1624 bp (*ΔRv1460*) 1915 bp (*Rv1460 ΔDNAbd*)	2142 bp
SCO region 2	SCOinup SCOdowndel	CATCCCTGGTAGACCGATCA ACTTCAGCCTTCCACCCATT	Binds 76 bp from the 3' end of the *upstream* region up to 105 bp downstream of *downdel* region	1803 bp	1287 bp	1578 bp	1579 bp	1803 bp	1287 bp (*ΔRv1460*) 1578 bp (*Rv1460 ΔDNAbd*)

Continued overleaf

Table 5. List of primers and strategy for SCO configuration screening and genotyping of DCOs

Region	Primer name	Sequence (5' → 3')	Description	Expected product sizes[a]				SCO configuration[a]	
				WT	***ΔRv1460***	***Rv1460 ΔDNAbd***	***Rv1460stop***	**Upstream SCO**	**Downstream SCO**
DCO genotyping[a]	Screen 1	GATTCGTGACGGCAGATTGAGC	Binds within the *upstream* region	408 bp	843 bp	1134 bp	1135 bp	NA	NA
	Screen 2	GAGTGTGACCGTCCGAGACTG	Binds within the *DNAbd* region deleted in all mutants					NA	NA
	Screen 3	CGATGCCATCGAGGTTGGAGC	Binds within the undeleted 174 bp region at 3' end of *Rv1460*					NA	NA

[a] Refer to Figure 9 and 4 for visual representation of SCO configuration determination and DCO genotyping using these primers

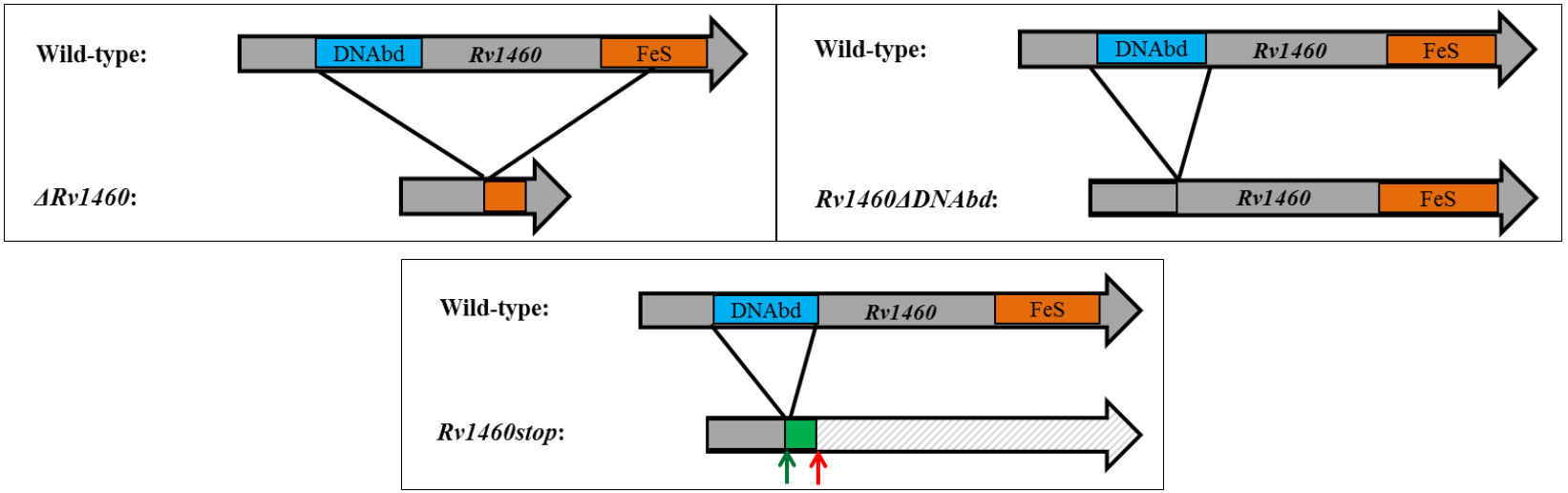


Figure 5. Representation of the three *Rv1460* deletion mutants to be generated compared to the wild-type. The green and red arrows indicate the position of the frameshift mutation and stop codon in the *Rv1460stop* mutant. The green region indicates different amino acids introduced as a result of the frameshift mutation. The regions are not drawn to scale.

3.4 Complementation

A complementation vector was generated by amplifying the *Rv1460compl* region from the pMVsuf01 vector using the primers indicated in Table 4 (FastStart Taq DNA polymerase kit (Roche): initial denaturation at 95 °C for 4 min, amplification for 25 cycles at 95 °C for 30 s, 60 °C for 30s, 72 °C for 30 s, final extension at 72 °C for 7 min) and cloning it into the pMV306H vector. This vector was converted to a stable integrating vector by removing the integrase gene by RE digestion with AvrII and PstI and re-ligation of the blunted product (Springer *et al.*, 2001) as indicated in Table 3. A construct containing the whole *suf* operon, pMVsuf01 (Table 3), was also used for complementation where indicated.

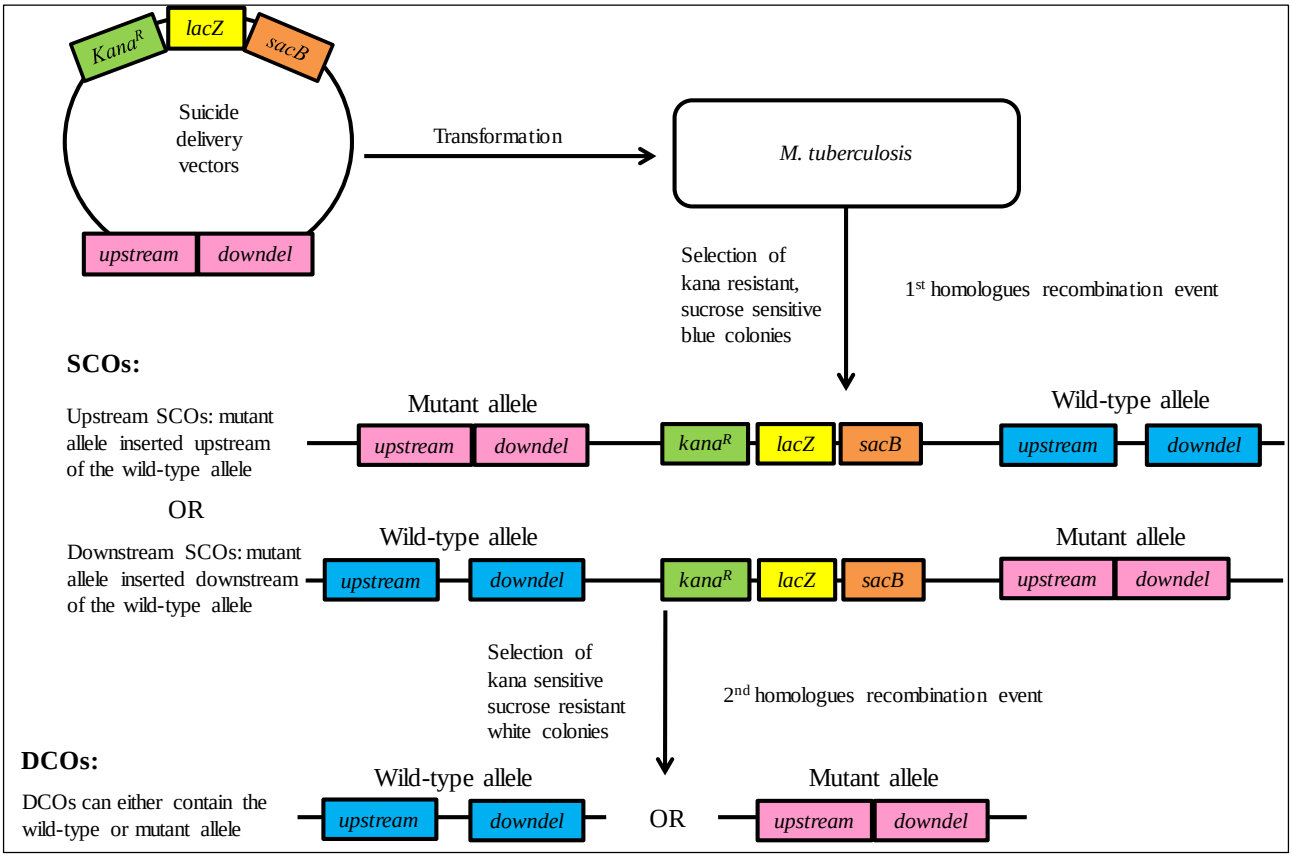


Figure 6. Generation of *M. tuberculosis Rv1460* deletion mutants by two-step allelic exchange.

3.5 RNA extraction

M. tuberculosis was cultured to mid-log growth phase, 10 ml culture pelleted and re-suspended in 1 ml RNAProBlue solution (Iepsa) and transferred to a 2 ml tube with glass beads. After ribolysation for four 25 s cycles at 4.5 Watts using the FastPrep-24 ribolyser (MP Biomedicals), lysates were frozen at -80 °C. Lysates were thawed, spun down at 2 190×g for 15 min to pellet the cellular debris.

RNA was extracted in 300 µl chloroform and 500 µl ethanol added to precipitate the RNA prior to loading on a Nucleospin RNA kit column (Macherey-Nagel). The column was desalted, sample treated on column with rDNase provided with the kit (15 min) and column washed as indicated according to manufacturer's instructions. RNA was eluted in 30 µl nuclease free water and column silica removed by centrifugation at 2 190×g for 15 min at 4 °C. The quality of the RNA was assessed by Bioanalyser (CPGR) and nanodrop. The RNA was treated with TURBO DNase provided in the TURBO DNA-free Kit (Ambion) using the rigorous treatment option and DNase inactivated with the DNase inactivation reagent (Ambion). The DNase inactivation reagent was removed by centrifugation at 2 190×g for 10 min at 4 °C. RNA was stored at -80 °C. The RNA concentration was taken as half of the concentration determined by Bioanalyser before treatment with TURBO DNase and inactivation reagent, as the inactivation reagent could interfere with the Bioanalyser results.

3.6 RT-qPCR

RNA (500 ng) was reverse transcribed with 1 mM reverse transcription (RT) primers indicated in Table 6 using the Transcriptor first strand cDNA synthesis kit (Roche) according to the manufacturer's instructions. cDNA was diluted 10 fold and stored at -80 °C. The FastStart Essential DNA Green Master mix (Roche) was used with 0.1 mM PCR primer to amplify *Rv1460*, *Rv1461* and *sigA* (housekeeping gene) according to the manufacturer's instructions. RT-qPCR was done using the LightCyler 96 machine (Roche) with the following protocol: pre-incubation: 4.4 °C/s ramp rate, for 600 s at 95 °C, 3 step amplification: denaturing at 4.4 °C/s ramp rate for 20 s at 95 °C, amplification at 2.2 °C/s ramp rate for 20 s touchdown from 67 °C to 62 °C with 1 °C change per 2 cycles for the first 12 cycles followed by an additional 33 cycles (total 45 cycles) at 62 °C and elongation at 4.4 °C ramp rate for 20s at 72 °C. Melt curve analysis was done at 4.4 °C ramp rate for 10 s at 95 °C. SYBR green 1 detection was done with a threshold value of 0.2.

Log dilutions of *M. tuberculosis* genomic DNA (10^6 to 10^1 genome copies per 2 µl) were prepared. The log of genomic DNA copies was plotted against the C_t values to generate a standard curve for each gene (*Rv1460*, *Rv1461* and *sigA*). Each standard curve fit a straight line and the formula of the line was used to determine the log copy number for each C_t value (representing an average of technical duplicate C_t readings). The log copy numbers were converted to copy numbers and the expression of *Rv1460* and *Rv1461* were normalised by dividing by *sigA* copy numbers for each sample to determine the relative *Rv1460* and *Rv1461* expression.

Table 6. List of primers used for RT and amplification of *Rv1460*, *Rv1461* and *SigA* for RT-qPCR

Gene	Primer name	Sequence (5' → 3')	Binding position[a,b]	Product size
Rv1460	Rv1460F	GAATTGTGCGAAACCGAGCAGCAG	640 bp from start of Rv1460	106 bp product
	Rv1460R	CAGGGGTACGTGGGTGGTG	732 bp from start of Rv1460	
	Rv1460RT	TCATCGGGACGCTCCTTCGG	788 bp from start of Rv1460	NA
Rv1461	Rv1461F	CAAGAGCGTTGCCCAGCCC	24 bp from start of Rv1461	93 bp product
	Rv1461R	GAGTCCGCCCAGCCGTAG	99 bp from start of Rv1461	
	Rv1461RT	GCTCGTTCTTCTTCGCGGA	175 bp from start of Rv1461	NA
SigA	SigAF	TGCAGTCGGTGCTGGACAC	1379 bp from start of SigA	195 bp product
	SigAR	CGCGCAGGACCTGTGAGCGG	1552 bp from start of SigA	
	SigART	CTGACATGGGGGCCCGCTACGTTG	19 bp downstream of SigA	NA

[a] Annotated according to the annotated start site of *Rv1460* specified on TBDB and Tuberculist

[b] The gene specific *Rv1460* primers bind outside the region deleted in the *Rv1460stop* mutant and the transcript levels of *Rv1460* can thus still be assessed for the *Rv1460stop* mutant

3.7 Cumene hydroperoxide exposure

The *M. tuberculosis* wild-type H37Rv progenitor, *Rv1460stop* mutant and the complementation strain were cultured in 7H9 OADC for six days, pelleted and washed twice in 1x PBS containing 0.05% Tween 80. Washed cells were re-suspended, filtered through a Falcon cell strainer (Corning) and sonicated for 12 min in a water bath sonicator to remove clumps. Cells were diluted to an OD_{600nm} of 0.2 in 1× PBS containing 0.05% Tween 80 before aliquoting 5 ml into small culture flasks. Cumene hydroperoxide (CHP) solution (5.4 M) (Sigma-Aldrich) was diluted (100×) in 100% ethanol. Fresh dilutions were prepared for every experiment. Cultures were exposed to 250 µM CHP or 100% ethanol (diluent of CHP) for 3 hours (h). Serial dilutions of cultures were prepared in 1× PBS containing 0.5% Tween 80 and plated in triplicate on 7H10 OADC. No hygromycin was added to the 7H10 plates for the complementation strains as it caused a growth delay. CFUs were determined after three to four weeks. The percentage survival was calculated by dividing the CFUs/ml of the CHP exposed culture by the CFUs/ml of the culture exposed to the diluent (unexposed).

3.8 Bipyridyl exposure

A stock solution of bipyridyl (Sigma-Aldrich) was prepared in ethanol (40 mM) and stored at -20 °C. The concentration of bipyridyl which did not significantly inhibit the growth of the wild-type stain was determined by exposing the H37Rv wild-type progenitor to 0, 40, 60, 80 and 100 µM bipyridyl. Starter cultures of *M. tuberculosis* wild-type H37Rv progenitor, *Rv1460stop* mutant and the complementation strain were sub-cultured to an OD_{600nm} of 0.05 in 70 ml media and the culture split into two 30 ml cultures. After two days, cultures were exposed to bipyridyl (at a concentration which did not affect growth of the wild-type H37Rv progenitor) or diluent (100% ethanol). Growth was monitored for a further 5 days by plating dilutions of cultures on 7H10 OADC and determining CFUs/ml after three to four weeks. No hygromycin was added to the 7H10 plates for the complementation strains.

3.9 Generation of vectors for the production of Rv1460 and Rv1460 C→S protein variants in *E. coli*

A vector for the production of Rv1460 as a N-terminally His-tagged protein was generated by amplifying the *Rv1460N* region from the pMVsuf01 vector using the primers indicated in Table 7 (Phusion HostStart polymerase: initial denaturation at 98 °C for 1 min, amplification for 25 cycles at 98 °C for 15 s, 64 °C for 30 s, 72 °C for 30 s, final extension 72 °C for 10 min) and cloning into the pET28 vector. A vector for the production of Rv1460 as a C-terminally His-tagged protein was generated by amplifying the *Rv1460C* region from the pMVsuf01 vector using the primers indicated in Table 7 (FastStart polymerase (Roche): initial denaturation at 95 °C for 4 min, amplification for 30 cycles at 95 °C for 30 s, 61 °C for 30 s, 72 °C for 30 s, final extension 72 °C for 7 min respectively) and cloning it into the pET21 vector. The three cysteine residues (C203, C216 and C244) predicted to coordinate an Fe-S cluster in Rv1460 (Cole *et al.*, 1998; Shen *et al.*, 2007) and an additional cysteine (C242) located in the vicinity of the third cysteine were replaced by serine residues to create Rv1460 C203S, C216S, C242S and C244S protein variants. This was done by amplifying the pJETRv1460N plasmid with primers indicated in Table 7 and the site-directed mutagenesis procedure previously described (Edelheit *et al.*, 2009) and sub-cloning the resulting mutated *Rv1460N* fragment into pET28. Multiple rounds of site-directed mutagenesis were done to create a vector containing all three Rv1460 C203S, C216S and C244S variants (Rv1460 C203,216,244S).

Table 7. Primers used for the generation of Rv1460 and IscS protein production vectors and EMSA

Region	Primer name	Sequence (5' → 3')[a]	Description[b]	Mutation
*Rv1460*N	Rv1460pETNF	CGCATATGGTGACCAGCACAACCCTGCCG (NdeI)	821 bp product including the 807 bp *Rv1460* gene with its stop codon and 14 bp for RE cut sites	NA
	Rv1460pETNR[c]	CAAGCT**TCA**TCGGGACGCTCCTTCGGTGCTG (HindIII)		
*Rv1460*C	Rv1460pETF	CGCATATGGTGACCAGCACAACCCTGCCG (NdeI)	819 bp product including 804 bp of *Rv1460* gene excluding the stop codon and 15 bp for RE cut sites	NA
	Rv1460pETCR	GAAGCTTTCGGGACGCTCCTTCGGTG (HindIII)		
C203S[d]	SNP1F	CAGCACCAT**TCC**CCGGTATCC	Site-directed mutagenesis	C→S at codon 203 in Rv1460
	SNP1R	GGATACCGG**GGA**ATGGTGCTG		
C216S[d]	SNP2F	CCCGAATTG**TCC**GAAACCGAG	Site-directed mutagenesis	C→S at codon 216 in Rv1460
	SNP2R	CTCGGTTTC**GGA**CAATTCGGG		
C242S[d]	SNP3F	AACGGAGAC**TCC**GCCTGCACC	Site-directed mutagenesis	C→S at codon 242 in Rv1460
	SNP3R	GGTGCAGGC**GGA**GTCTCCGTT		
C244S[d]	SNP4F	GACTGCGCC**TCC**ACCACCCAC	Site-directed mutagenesis	C→S at codon 244 in Rv1460
	SNP4R	GTGGGTGGT**GGA**GGCGCAGTC		
iscSN[c]	IscSF	CATATGGCCTACCTGGATCACG (NdeI)	1194 bp region including 1182 bp of *iscS* gene including the stop codon and 12 bp for RE cut sites	NA
	IscSR	AAGCTT**TCA**TCGGGATGCTCC (HindIII)		
Rv1460 promoter region	EMSARv1460F[e]	GGGCCCAGCTCGGTAGTGGTCA (ApaI)	325 bp product containing 190 bp upstream and 122 bp downstream of the ChIP-seq peak center (at 20 bp upstream of the newly annotated *Rv1460* start codon) and 12 bp for RE cut sites	NA
	BioEMSARv1460F[e]	Biotin-AGCTCGGTAGTGGTCA		
	EMSARv1460R	GGGCCCGGTGATCGATCCGGATTC (ApaI)		
Rv1461 binding region	EMSARv1461F[e]	GGGCCCATCCGCAACACCTACGAC (ApaI)	348 bp region including 166 bp upstream and 169 bp downstream of the ChIP-seq peak center (at 524 bp downstream of the start codon) within *Rv1461* and 12 bp for RE cut sites	NA
	BioEMSARv1461F[e]	Biotin-ATCCGCAACACCTACGAC		
	EMSARv1461R	GGGCCCCATGTTCTCGGTGTTGATTCGG (ApaI)		

[a] RE cutting sites are underlined and RE indicated in parenthesis

[b] Annotated according to the annotated start site of *Rv1460* specified on TBDB and Tuberculist

[c] Bold region indicates a stop codon and [d] Double underlined letter indicates the base pair change G→C introduced and bold letters indicate the codon which is mutated from C→S

[e] Either the biotin-labelled primer or unlabelled primer was used to prepare labelled DNA or specific competitor for EMSA experiment

3.10 Optimisation of Rv1460 protein production in *E. coli*

Production of Rv1460 as an N- and C-terminally His-tagged protein was tested in Rosetta-gami 2 and Arctic express (DE3) protein production strains. Pre-cultures were cultured at 37 °C, sub-cultured and grown at 37 °C and 30 °C respectively to an OD_{600nm} of ~0.5-0.6. A volume of 20 ml unexposed culture was pelleted and stored at -80 °C. The cultures were then exposed to different isopropyl β-D-thiogalactoside (IPTG) concentrations at different temperatures for varying times. At each time point 20 ml culture was collected and pelleted and stored at -80 °C. Each 20 ml pellet was re-suspended in 2 ml sonication buffer (20 mM Tris, 500 mM NaCl, 20 mM imidazole, 1% glycerol, pH 8), cells lysed by sonication on ice (three 30 s cycles at 50 Watts) and cellular debris (insoluble fraction) pelleted by centrifugation for 30 min at 2 190×g. Cellular debris was re-suspended in sonication buffer and both insoluble and soluble fractions examined by denaturing SDS-PAGE (12% acrylamide). Western blot, using His-probe (G-18) rabbit (primary) and goat anti-rabbit IgG-HRP (secondary) antibodies (Santa Cruz), was used to confirm the protein identity. Proteins were purified on small-scale using the MagneHis purification kit (Promega) according to the manufacturer's instructions and the protein separated on an SDS-PAGE gel and analysed by mass spectrometry at the Central Analytical Facility at Stellenbosch University to confirm its identity. The His-tag was subsequently cleaved using the Thrombin cleavage capture kit (Novagen) according to manufacturer's instructions.

3.11 Large-scale purification of Rv1460 protein

Starter cultures containing appropriate antibiotics were cultured at 37 °C and sub-cultured into five 2 L Erlenmeyer flasks containing 500 ml LB with antibiotics. Zinc sulphate (8.3 µM) was added to the LB when producing Rv1460 because zinc ions (Zn^{+2}) can potentially occupy an Fe^{+2} binding site within a protein thereby stabilizing the protein during folding and improving its solubility. Recombinant protein was produced and cells pelleted by centrifugation at 5000×g for 10 min and stored at -80 °C. The cell pellet (from 2.5 L culture) was resuspended in 50 ml sonication buffer (20 mM Tris, 500 mM NaCl, 20 mM imidazole, 1% glycerol, pH 8) containing 1 mM dithiothreitol (DTT) (freshly added) and half of a protease inhibitor tablet (Sigma-Aldrich). Cells were lysed by sonication on ice using a QSonica probe sonicator (3/4 tip diameter) for 4 min at 50 Watts with 15 s on and off intervals. Cellular debris was pelleted at 20 200×g for 30 min at 4 °C. The supernatant was filtered through a 0.45 micron surfactant free cellulose acetate filter (Merck), before loading onto a

HiTrap 5 ml nickel-NTA (Ni-NTA) column (GE Healthcare Life Sciences) coupled to an AKTA protein purification system (GE Healthcare Life Sciences).

The column was equilibrated with 10 column volumes (CV) of wash buffer (20 mM Tris, 500 mM NaCl, 20 mM imidazole, pH 8) at a flow rate of 5 ml per min before loading samples. The column was washed with 10 CV of wash buffer and 5 ml flow through fractions collected. Protein was eluted with an imidazole gradient from 100% wash buffer to 70% elution buffer (20 mM Tris, 500 mM NaCl, 500 mM imidazole, pH 8) in 100 ml and 2.5 ml fractions collected. The remaining protein was eluted with 100% elution buffer. Protein elution was monitored at 280 nm.

Alternatively, a 5 ml nickel column was packed using His-Pur Ni-NTA Superflow agarose (Thermo Scientific) in a Poly-prep chromatography column (20 ml) (Bio-Rad). This column was pre-washed with water and equilibrated with 10 CV of wash buffer containing 1 mM DTT. Culture pellet (1 L culture) was re-suspended in 25 ml sonication buffer containing 100 mg lysozyme, sonicated using a QSonica probe sonicator with microtip for 4 min at 50 Watts with 15 s on and off intervals and cellular debris pelleted by centrifugation at 10 000×g for 1 h. The supernatant (25 ml) was loaded onto the column and the flow through collected by gravity flow. The column was washed with five CV of wash buffer containing 1 mM DTT, where after the protein was eluted in ten CV of 10% elution buffer containing 1 mM DTT followed by five CV of 30% elution buffer containing 1 mM DTT and fractions collected by gravity flow. The remaining protein was eluted in five CV of 100% elution buffer.

SDS-PAGE and protein staining using AquaBlue protein stain (Vacutec) and western blotting was used to determine which fractions contained the protein of interest and the appropriate fractions pooled and concentrated to 5 ml using an Amicon flow cell (Bioseparations, Millipore) with Ultrafiltration disk membranes (10 kDa cut off) (Millipore) before filtering through a 0.22 micron filter (Merck). The sample was loaded onto a Sephadex S200 column (16/600) pre-equilibrated with gel filtration buffer (20 mM Tris, 150 mM NaCl, 10% glycerol, pH 8 with 1 mM DTT) at a flow rate of 1 ml per min and 1.5 ml fractions collected. The fractions were analysed by SDS-PAGE and selected fractions pooled and concentrated using the Amicon Ultra Centrifugal Filter units (3 kDa cut off) (Millipore) and centrifugation at 3 220×g. Protein samples were aliquoted and flash frozen in liquid nitrogen before storing at -80 °C.

Thrombin cleavage was done using the Thrombin cleavage capture kit (Novagen) according to the manufacturer's instructions. The MagneHis nickel beads were used to bind the uncleaved His-tagged

protein while the cleaved protein was collected in the flow through using the wash buffer previously described. The biotinylated thrombin was removed by incubation with streptavidin agarose which was pelleted by centrifugation at 2 190×g for 15 min at 4 °C. Alternatively, the protein sample purified by affinity purification (containing high imidazole concentrations) was dialysed overnight in dialysis tubing (10 kDa cut off) (Sigma-Aldrich) in dialysis buffer (20 mM Tris, 150 mM NaCl, pH 8 with 1 mM DTT) and the His-tag concurrently cleaved using thrombin (Sigma-Aldrich) (5 units/20mg protein) and removed by gel filtration. The protein concentration was determined either by Bradford assays or spectrophotometrically by determining the absorbance at 280 nm (A_{280nm}) using a Spectro UV-16 spectrophotometer (Mrc) and using the extinction coefficient for reduced Rv1460, $\varepsilon_{reduced} = 9970\ M^{-1}cm^{-1}$, and 1 cm light path to calculate the protein concentrations $C = \frac{A_{280nm}}{\varepsilon \times 1cm}$ before it was used for subsequent experiments.

3.12 Generation of vectors for the production of IscS*Mtb* protein in *E. coli*

The *M. tuberculosis iscS* gene (*IscSN*) was amplified from H37Rv genomic DNA using the primers indicated in Table 7 (Phusion HostStart DNA polymerase: initial denaturation at 98 °C for 1 min, amplification for 30 cycles at 98 °C for 15 s, 60 °C for 30 s, 72 °C for 45 s, final extension at 72 °C for 10 min) and cloned into pET28 to create the pET28iscS vector. Production of recombinant IscS*Mtb* was optimised by using different *E. coli* protein production strains (Rosetta 2 and Arctic express (DE3)), culture media and inducing at different temperatures for varying times. Large-scale IscS*Mtb* production was done at the optimal production conditions determined. Cell pellets (0.5 L) were re-suspended in 50 ml sonication buffer and the same protocol as for large-scale purification of Rv1460 used for IscS*Mtb* purification.

3.13 Fe-S cluster reconstitution

Vectors for the production of *E. coli* IscS, IscU and TEV were provided by Prof Annalisa Pastore at King's College, London. *E. coli* IscS, IscU and modified TEV protease (which is used for cleavage of the His-tag) were produced and large-scale purification done as previously described (Prischi *et al.*, 2010). IscS and IscU were concentrated and aliquots flash frozen in liquid nitrogen until use in reconstitution experiments.

Fe-S reconstitution reactions were set up in an anaerobic glove box (Belle Technology at King's College, London) under nitrogen atmosphere. Reconstitution reactions contained 25 µM Rv1460,

12.5 µM Fe^{2+}, 1.5 µM DTT, 0.5 µM IscS and 2.5 µM IscU, 20 mM Tris-HCl (pH 8), 150 mM NaCl, in a total volume of 800 µl and were set up in a 1 ml sealable quartz cuvette. The reaction was initiated by addition of 125 µM cysteine and monitored over time by recording UV-visible spectra over a range of wavelengths (250 – 600 nm) using a Varian Spectrophotometer (at King's College, London).

3.14 Electrophoretic mobility shift assay

The regions surrounding the peak center of the Rv1460 binding sites within *Rv1460* and *Rv1461* predicted by ChIP-seq (Minch *et al.*, 2015) were amplified using a biotin-labelled forward primer (Table 6) and unlabelled reverse primer or unlabelled forward and reverse primers (Phusion High-Fidelity DNA polymerase (New England Biolabs): activation at 98 °C for 1 min, 20 cycles of amplification at 98 °C for 15 s, 64 °C for 30 s, 72 °C for 30 s, final extension at 72 °C for 10 min). The concentration of biotin-labelled PCR product was determined using a nanodrop. Binding reactions were set up using the components of the LightShift Chemiluminescent EMSA kit (Thermo Scientific) according to manufacturer's instructions. Binding reactions contained biotin-labelled DNA (20 fmol), mixed with different ratios of Rv1460 protein (His-tag cleaved) in a buffer containing 10 mM Tris, 50 mM KCl, 1 mM DTT, 5 mM $MgCl_2$, pH 7.5. Specific competitor (unlabelled DNA) was added in order to compete with biotin-labelled DNA for binding to the Rv1460 protein. After 30 min incubation at room temperature, loading dye was added to the samples. TBE gels (8%) (Life Technologies or self-made) were used to separate the DNA by electrophoresis at 100 Volt for approximately 2.5 h in 1× TBE buffer (89.2 mM Tris, 82.3 mM Boric acid and 20 mM EDTA). The DNA was transferred to a nitrocellulose membrane (GE Healthcare) by blotting for 1 h and cross-linked to the membrane at 254 nm at 0.12 mJ/cm^2. The LightShift Chemiluminescent Nucleic acid detection module kit (Thermo Scientific) was used to detect biotin-labelled DNA according to the manufacturer's instructions.

3.15 Circular dichroism

Far-UV circular dichroism (CD) was used to check whether Rv1460 was properly folded after large-scale purification using a Chirascan CD Spectrometer (Applied Photophysics) (at the Central Analytical Facility of Stellenbosch University) and a 1 ml quartz cuvette at 25°C. Near-UV CD was used to determine whether Fe-S cluster reconstitution on Rv1460 occurred using a Jasco J-715 spectropolarimeter (at King's College, London) and a 1 ml sealable quartz cuvette at 25°C.

3.16 Statistical analyses

GraphPad Prism software was used for all statistical analyses. Multiple t-tests were done without assuming equal standard deviation among samples and without correction for multiple comparisons. A p-value of smaller than 0.05 was considered statistically significant. Each data set for the growth curves was fitted to a sigmoidal curve and straight line and best fit determined using a least squares (ordinary) fit and comparing the fit using Akaike's Information Criteria to select the model that most likely represents the data. The y-intercept was specified for each data set according to the measured value on day 0. The HillSlope of the curves fitting a sigmoidal model and slope of the curves fitting a straight line for the three replicates were used for t-test analyses as previously mentioned. Obvious outliers were excluded from the analyses. Growth under iron deprivation was compared to that of the unexposed cultures by doing a t-test for each day post-exposure to bipyridyl.

3.17 Bioinformatic analysis

The protein sequence of Rv1460 was obtained from the TBDB database (Kapopoulou *et al.*, 2011) and homologs were identified by BLASTP analysis (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Protein alignments were done using the Clustal Omega software available online (http://www.ebi.ac.uk/Tools/msa/clustalo/).

4. Results

4.1 The Rv1460 protein is highly conserved among mycobacteria

To assess the level of Rv1460 conservation among mycbocterial species, the protein encoded by *Rv1460* was aligned with homologues from selected mycobacterial species. This revealed that the protein is highly conserved among mycobacteria (Figure 7 A). The three cysteines at positions 203, 216 and 244, predicted to coordinate an Fe-S cluster, are particularly conserved (Figure 7 A).

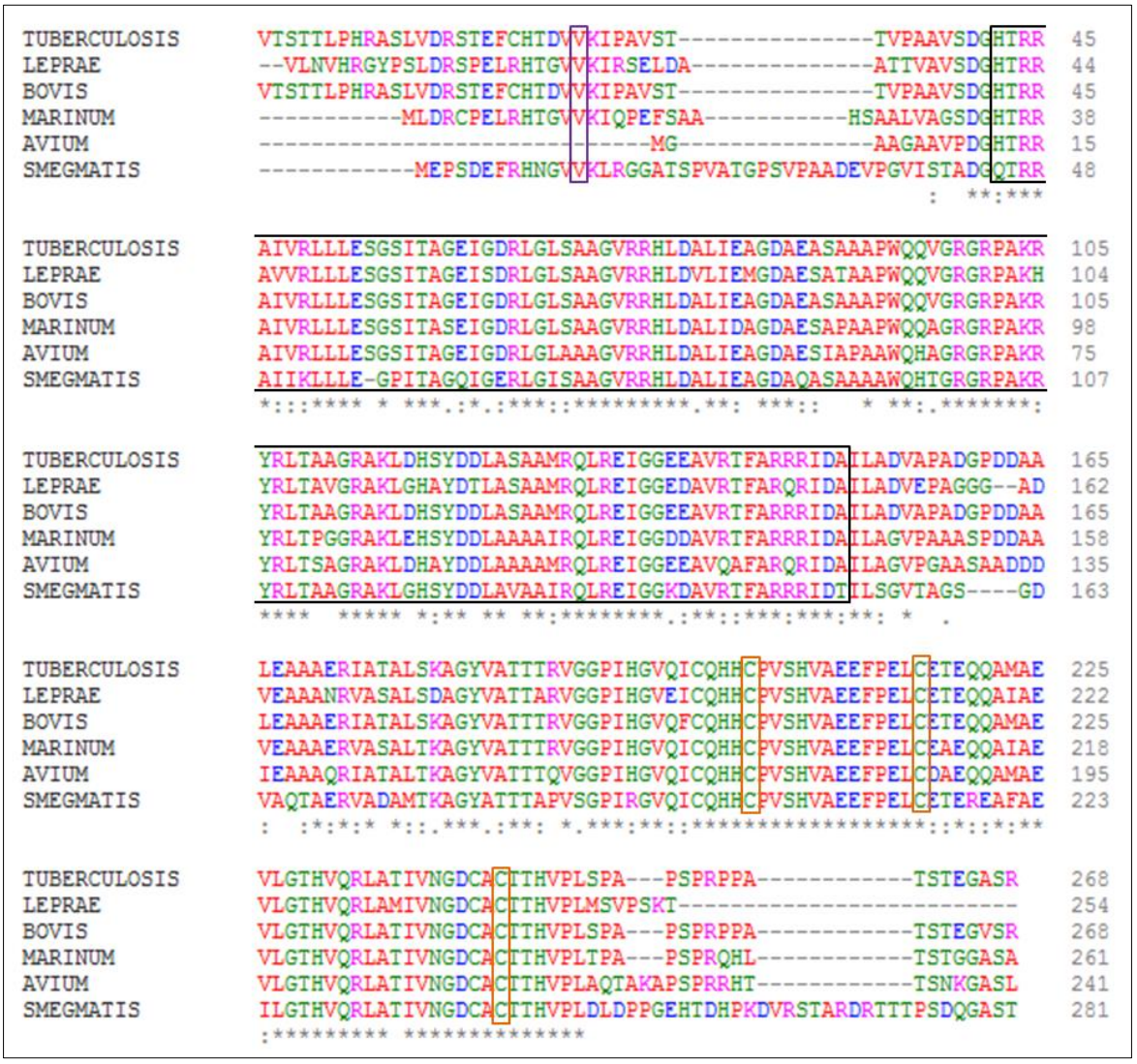


Figure 7. Rv1460 is conserved in various mycobacterial species. Alignment of Rv1460 protein sequences to homologues in various mycobacterial species using Clustal Omega software (http://www.ebi.ac.uk/Tools/msa/clustalo/). The purple box indicates the newly annotated start codon for *M. tuberculosis Rv1460* (Cortes *et al.*, 2013). The black box indicates the winged helix-turn-helix DNA-binding motif (IPR011991 – InterProScan) spanning amino acids 42 to 150 according to the *M. tuberculosis* Tuberculist annotated start

codon. Orange boxes indicate the positions of conserved cysteine residues predicted to be involved in Fe-S cluster coordination (Shen *et al.*, 2007).

Comparison of the Rv1460 protein sequence with SufR homologues from *Synechocyistis*, *Corynebacterium*, *Rhodococcus* and *Nocardia* revealed conservation of these cysteine residues among bacterial species (Figure 8). Interestingly, a N-terminal extension is present in most mycobacterial homologues, but absent in other species. This extension varies in both length and sequence in the mycobacterial homologues (Figure 7). Genome-wide transcription start-site mapping

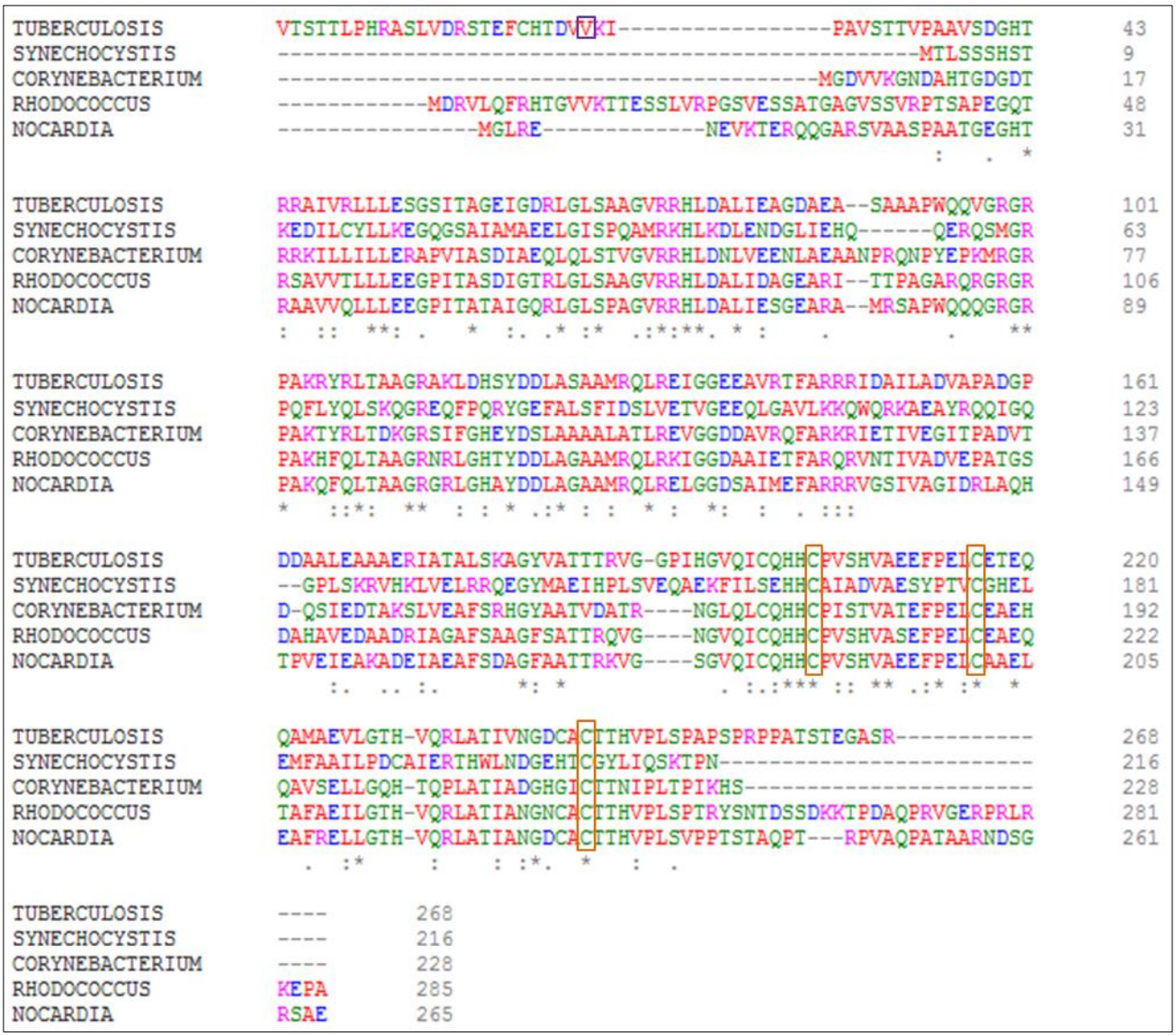


Figure 8. Rv1460 has conserved cysteine residues present in SufR homologues. Alignment of Rv1460 protein sequences to SufR homologues in *Synechocystis*, *Corynebacterium glutamicum* (ATCC 13032), *Rhodococcus fascians* and *Nocardia veterana* using Clustal Omega software (http://www.ebi.ac.uk/Tools/msa/clustalo/). The purple box indicates the newly annotated start codon for *M. tuberculosis Rv1460* (Cortes *et al.*, 2013). Orange boxes indicate the positions of conserved cysteine residues involved in Fe-S cluster coordination in *Synechocystis* (Shen *et al.*, 2007).

re-annotated the start site of Rv1460 at +73 and identified a 120 bp 5' untranslated region, suggesting that only part of the extension may be translated (Cortes *et al.*, 2013) (Figure 7). ChIP-seq analysis identified a binding site for Rv1460 20 bp upstream of the +73 start codon, implicating it in regulation of transcription through direct binding to this region (Minch *et al.*, 2015).

4.2 Generation of *M. tuberculosis* mutants harbouring unmarked deletions in *Rv1460*

4.2.1 Generation of suicide delivery and complementation vectors

To assess the role of *Rv1460* in *M. tuberculosis* physiology, attempts were made to generate three *Rv1460* deletion mutants: an in-frame, unmarked deletion which removes the bulk of the *Rv1460* gene (*ΔRv1460*), in-frame unmarked deletion of the predicted *Rv1460* DNAbd (*Rv1460ΔDNAbd*) and a truncation mutant where a frameshift mutation results in a premature stop codon and truncation of the Rv1460 protein (*Rv1460stop*) (Figure 5). Mutants were generated according to the Tuberculist and TBDB annotated start codon. The *upstream*, *downdel*, *downDNAbddel* and *downDNAbddelstop* regions (Table 4) were amplified from H37Rv genomic DNA (Supplementary Figure 1) and cloned into the p2NIL vector. The selectable markers from pGOAL17 were cloned into the PacI site to create the three suicide delivery vectors p2NIL17ΔRv1460, p2NIL17Rv1460ΔDNAbd and p2NIL17Rv1460stop (Table 3) which were confirmed to be correct by RE mapping (Supplementary Figure 2 A-C) and Sanger sequencing using the sequencing primers indicated in Table 4.

Complementation vectors, containing *Rv1460* and presumably its promoter region, were generated by amplifying *Rv1460compl* (Table 4) (Supplementary Figure 3 A) and cloning the resulting product into the pMV306 vector to generate the pMV*Rv1460compl* vector which was confirmed by RE mapping (Supplementary Figure 3 B). A stable integrating vector, pMV*Δint1460*, was obtained by removing the integrase gene (Springer *et al.*, 2001) by RE digestion of the pMV*Rv1460compl* vector with PstI and AvrII and was also confirmed by RE mapping (Supplementary Figure 3 C).

4.2.2 Generation of unmarked Rv1460 deletion mutants by two-step allelic exchange

M. tuberculosis H37Rv was transformed with suicide delivery vectors and two-step allelic exchange (Figure 6) done to generate SCOs. Both upstream and downstream SCOs were obtained as confirmed by PCR using primers indicated in Table 5 (Figure 9), indicating efficient homologous recombination

in both configurations without bias for either configuration. Six attempts to generate the Δ*Rv1460* and *Rv1460ΔDNAbd* mutants were unsuccessful as only colonies with the wild-type allele were obtained following selection and screening of 296 (*ΔRv1460*) and 269 (*Rv1460ΔDNAbd*) DCOs respectively. It was suspected that culturing under anaerobic conditions could affect mutant viability favourably as fewer ROS are generated under anaerobic conditions (Vatansever *et al.*, 2013). Attempts to generate the Δ*Rv1460* and *Rv1460ΔDNAbd* mutants under anaerobic conditions, using an anaerobic chamber for the incubation of 7H10 plates to recover DCOs, were however also unsuccessful.

To determine whether *Rv1460* is essential for *in vitro* growth under the conditions investigated, the pMV*Rv1460compl* vector was introduced at the *attB* site in the chromosome of SCOs and selection for DCOs done. Both *ΔRv1460*:: pMV*Rv1460compl* and *Rv1460ΔDNAbd*:: pMV*Rv1460compl* mutants were successfully generated and their genotype confirmed by PCR and Southern blotting (Figure 10). The Δ*Rv1460* and *Rv1460ΔDNAbd* mutations were therefore lethal in the absence of a functional copy of *Rv1460* suggesting that *Rv1460* is essential under the conditions investigated. Interestingly, when the entire *suf* operon instead of an *Rv1460* copy was introduced at the *attB* site of SCOs prior to DCO selection, neither mutant could be generated.

In contrast, *Rv1460stop* mutants were successfully generated. The genotype of the *Rv1460stop* mutants was confirmed by PCR and Southern blotting (Figure 11) and the insertion of an extra nucleotide causing a frame-shift mutation confirmed by Sanger sequencing. This is suprising given that Rv1460 lacks both functional domains in these strains (Figure 5). *Rv1460stop* mutants were complemented using the pMV*Δint1460* vector in combination with the pBSint vector which provides the integrase in trans for stable integration of the pMV*Δint1460* vector. When pBSint was not co-transformed with the pMV*Δint1460* vector, no integration of the pMV*Δint1460* vector was observed, indicating stable integration of pMV*Δint1460* at the *attB* site, enabling culturing in the absence of the selectable antibiotic.

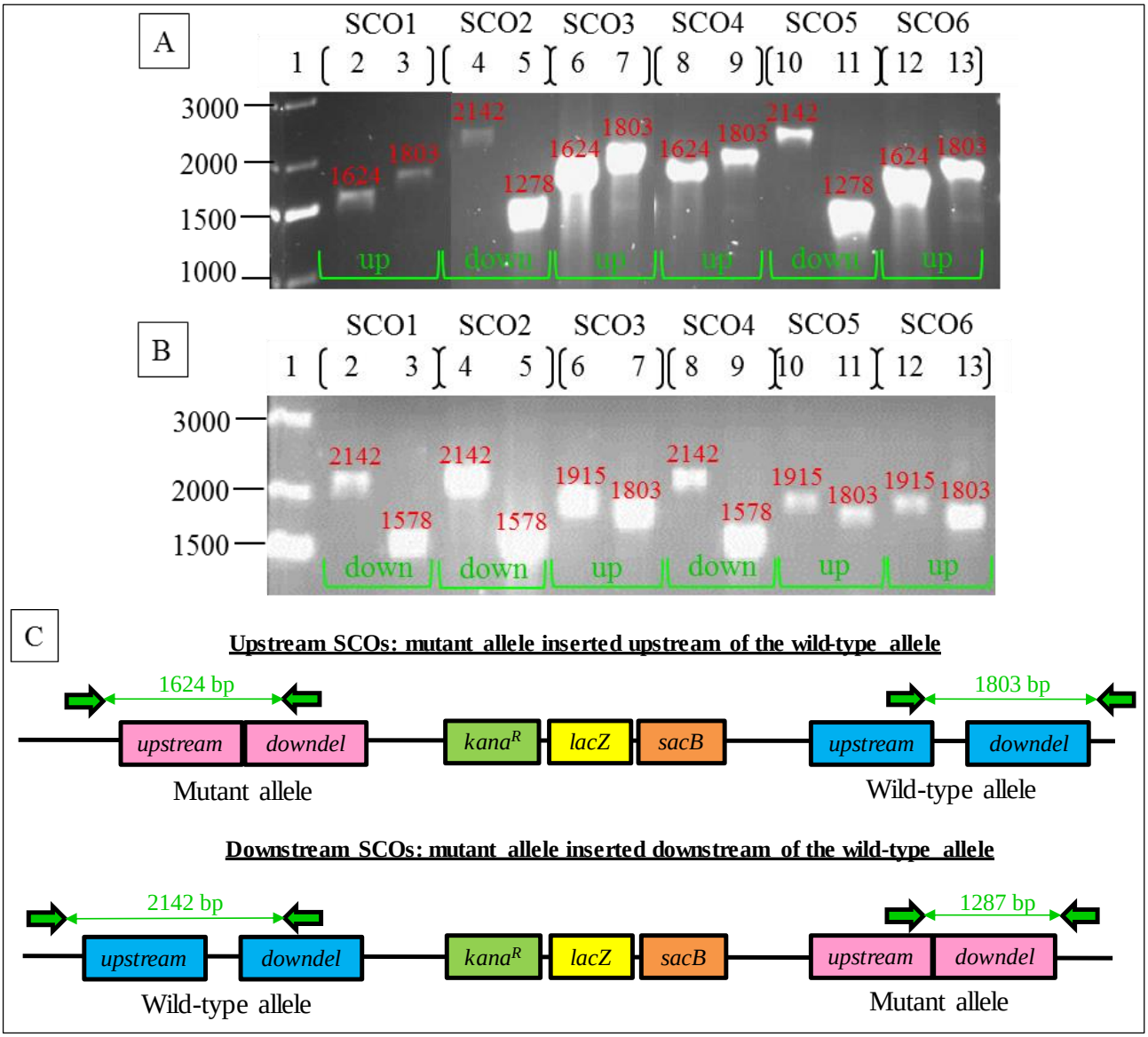


Figure 9. PCR screening to determine the configuration of *ΔRv1460* and *Rv1460ΔDNAbd* SCOs.

PCR products for (A) *ΔRv1460* and (B) *Rv1460ΔDNAbd* SCOs indicating whether the mutant *Rv1460* allele was inserted upstream or downstream of the wild-type *Rv1460* allele (Table 5): lane (1) DNA ladder (2) SCO1 region 1 and (3) 2, (4) SCO2 region 1 and (5) 2, (6) SCO3 region 1 and (7) 2, (8) SCO4 region 1 and (9) 2, (10) SCO5 region 1 and (11) 2 and (12) SCO6 region 1 and (13) 2. The green "up" and "down" on gels indicate whether the SCO is a upstream or downstream SCO. (C) Graphical representation of expected PCR products for SCO region 1 and 2 (Figure 6) in upstream and downstream SCOs of *ΔRv1460*.

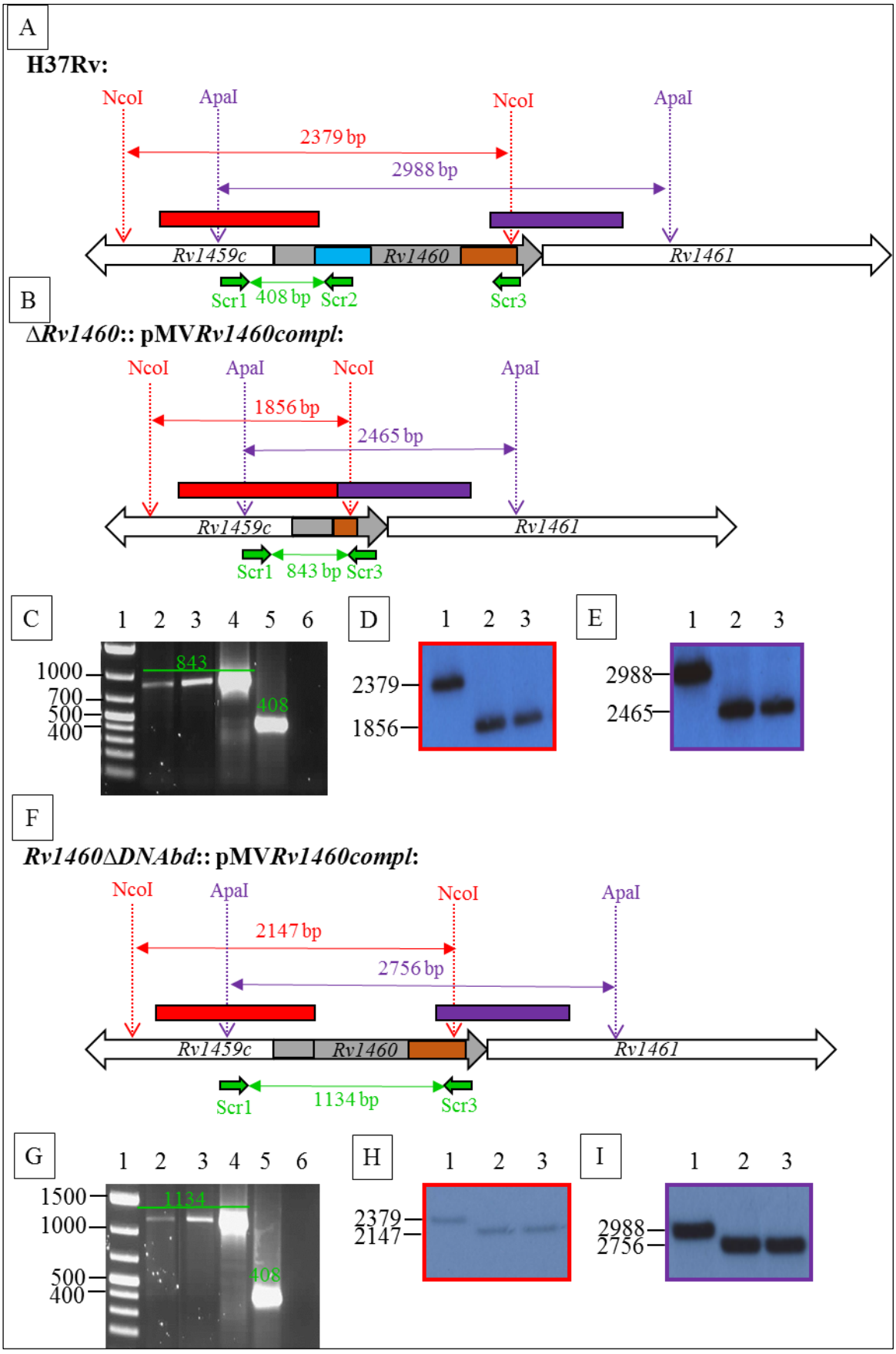


Figure 10. Genotypic characterisation of the *M. tuberculosis ΔRv1460*:: pMV*Rv1460compl* and *Rv1460ΔDNAbd*:: pMV*Rv1460compl* mutants. Graphical representation of the strategy for PCR and Southern blot genotyping of (A) H37Rv wild-type progenitor, (B) Δ*Rv1460*:: pMV*Rv1460compl* and (F) *Rv1460ΔDNAbd::*

pMV*Rv1460compl*. Genomic DNA was digested with NcoI and ApaI of which RE cut sites are indicated by red and purple arrows. The *upstream* and *downdel* regions (Table 4) were used as probes as indicated by the red and purple boxes above the genes. NcoI and ApaI digested fragments were detected using the *upstream* and *downdel* regions respectively as probes. The position of the Screen 1-3 primers and expected PCR product sizes (Table 5) are indicated. (C) PCR screening: lane (1) DNA ladder, (2) H37Rv, (3-4) two Δ*Rv1460*:: pMV*Rv1460compl* mutants, (5) p2NIL17*ΔRv1460* vector control and (6) no DNA control. (D) Southern blot analysis: lane (1) H37Rv, (2-3) two Δ*Rv1460*:: pMV*Rv1460compl* mutants using NcoI RE and *upstream* region (red box) as probe and (E) using ApaI RE and *downdel* region (purple box) as probe. (G) PCR screening: lane (1) DNA ladder (2) H37Rv, (3-4) two *Rv1460ΔDNAbd::* pMV*Rv1460compl* mutants, (5) p2NIL17*Rv1460ΔDNAbd* vector control and (6) no DNA control. (H) Southern blot analysis: lane (1) H37Rv, (2-3) two *Rv1460ΔDNAbd::* pMV*Rv1460* mutants using NcoI RE and *upstream* region (red box) as probe and (I) using ApaI RE and *downdel* region (purple box) as probe. Genes are not drawn to scale.

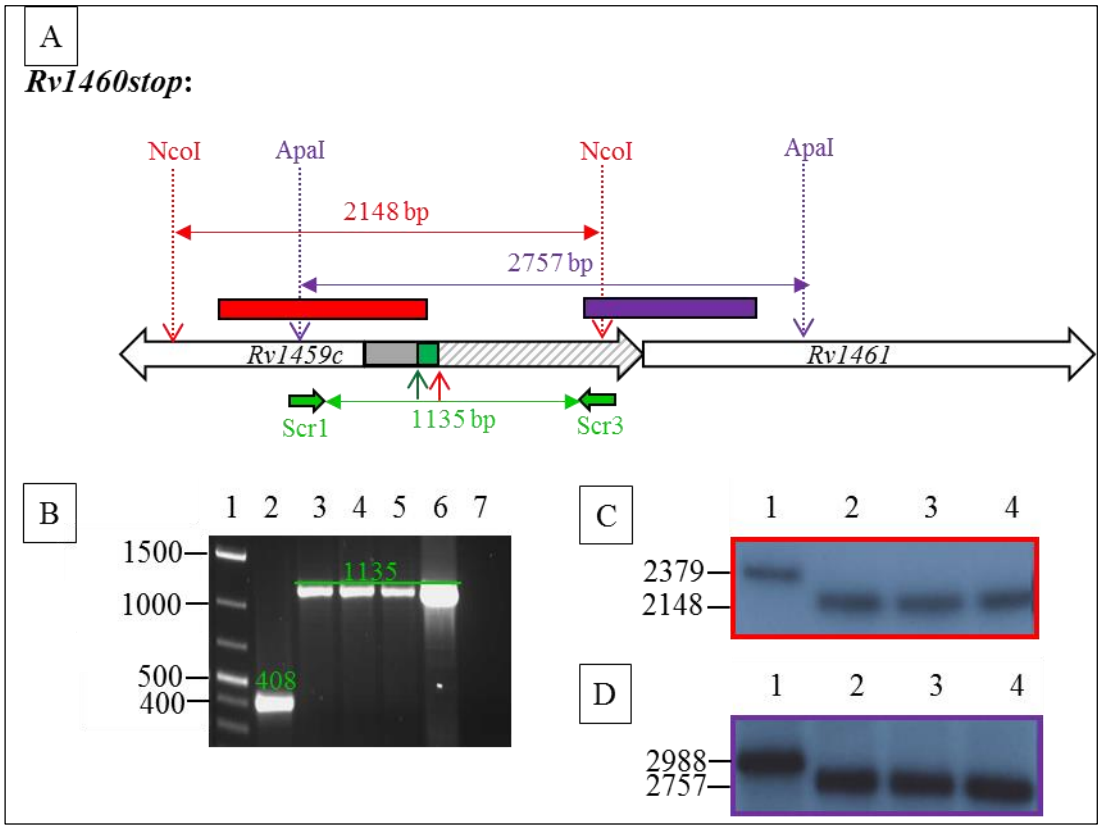


Figure 11. Genotypic characterisation of the *M. tuberculosis Rv1460stop* mutants. Graphical representation of the strategy for PCR and Southern blot genotyping of the (A) *Rv1460stop* mutant. Genomic DNA was digested with NcoI and ApaI of which RE cut sites are indicated by red and purple arrows. The *upstream* and *downdel* regions (Table 4) were used as probes as indicated by the red and purple boxes above the genes. NcoI and ApaI digested fragments were detected using the *upstream* and *downdel* regions respectively as probes. The position of the Screen 1-3 primers and expected PCR product sizes (Table 5) are indicated. (B) PCR screening: lane (1) DNA ladder, (2) H37Rv, (3-5) three *Rv1460stop* mutants (1.19, 5.19, 5.20), (6) p2NIL17*Rv1460stop* vector control and (7) no DNA control. (C) Southern blot analysis: lane (1) H37Rv, (2-4) three *Rv1460stop* mutants (1.19, 5.19, 5.20) using NcoI RE and *upstream* region (red box) as probe and (D) using ApaI RE and *downdel* region (purple box) as probe. Genes are not drawn to scale.

4.3 Phenotypic characterisation of the *Rv1460stop* mutant

4.3.1 Growth under standard culture conditions

To assess the effect of the truncation mutation on growth, growth of H37Rv wild-type progenitor, three independently isolated *Rv1460stop* mutants (1.19, 5.19, 5.20) and their complementation strains were monitored in liquid culture under standard culture conditions. All *Rv1460stop* mutants had a growth defect in liquid culture under standard culture conditions compared to the H37Rv wild-type progenitor, although the severity of the phenotype differed, with the *Rv1460stop* 5.20 mutant showing a much less severe defect (Figure 12 A-C). In all the *Rv1460stop* mutants, the growth defect could be reversed by genetic complementation with the pMV*Δint1460* vector at the *attB* site (Figure 12 A-C). The growth of the *Rv1460stop* 5.20 mutant on solid media (7H10 OADC) was also delayed (Figure 13). Since the *Rv1460* copy was inserted at the *attB* site and not within its natural native gene context, the loss of Rv1460 protein was responsible for the growth phenotype observed in all three mutants. Subsequent experiments were performed using the *Rv1460stop* 5.20 mutant which showed the least severe growth defect, since the more severe phenotype of the 1.19 and 5.19 *Rv1460stop* mutants could potentially mask a survival defect that results from exposure to oxidative stress. In addition, it was anticipated that the growth defect of the 1.19 and 5.19 mutants would be even more pronounced (even though this was not assessed) and the delayed appearance of colonies on solid media could make it difficult to accurately determine the CFUs for these strains. The growth of the Δ*Rv1460*:: pMV*Rv1460compl* and *Rv1460ΔDNAbd*:: pMV*Rv1460compl* strains were also compared to that of H37Rv wild-type progenitor and did not show any growth defect (Figure 12 D).

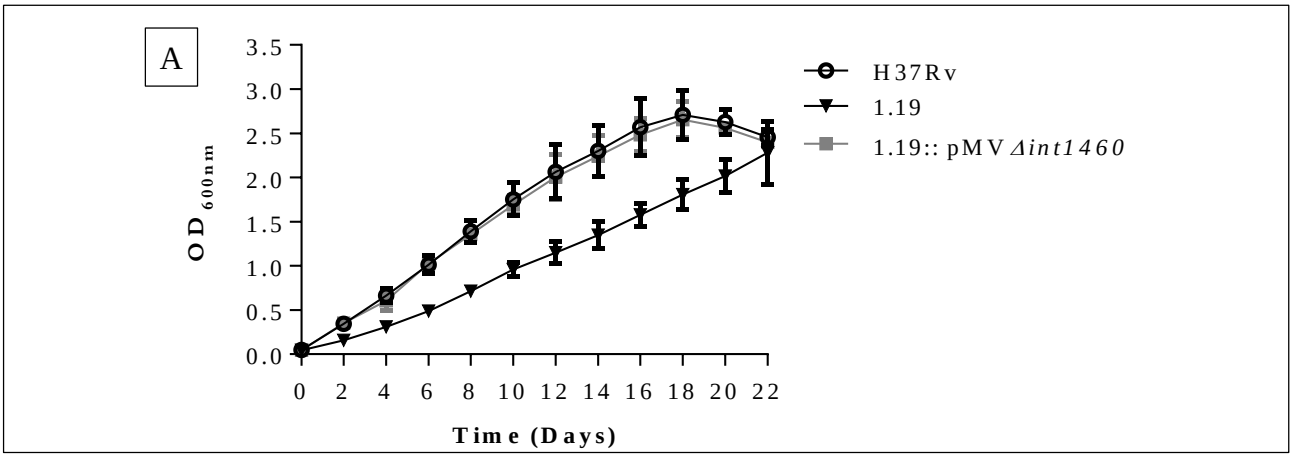


Figure 12. Growth curves of *Rv1460stop* mutants under standard culture conditions.

Continued overleaf

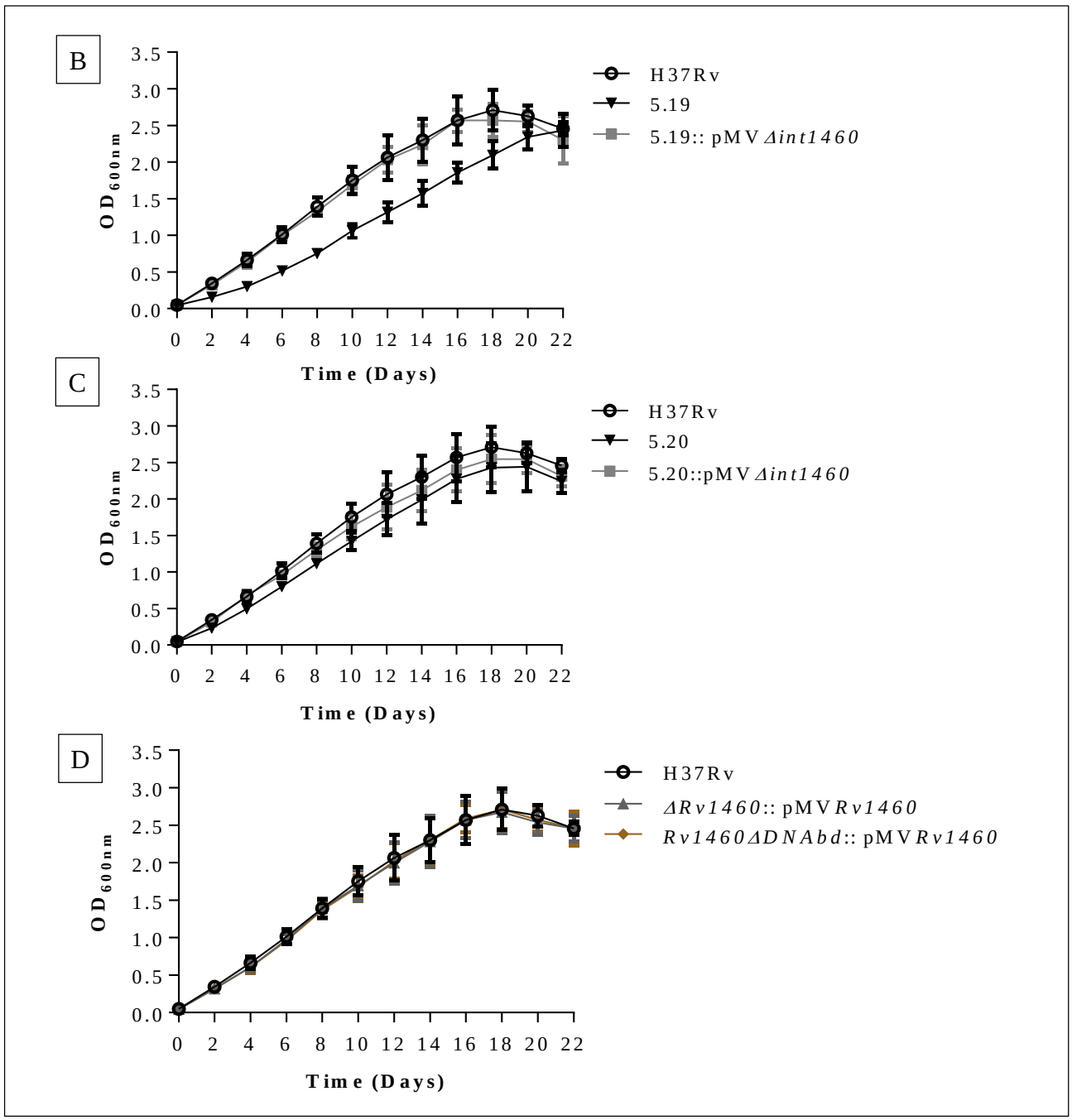


Figure 12. Growth curves of *Rv1460stop* mutants under standard culture conditions. Growth of three *Rv1460stop* mutants (A) 1.19, (B) 5.19 and (C) 5.20 and their complementation strains and (D) Δ*Rv1460*:: pMV*Rv1460compl* and *Rv1460ΔDNAbd*:: pMV*Rv1460compl* strains were monitored by determining OD_{600nm} readings over 22 days. Values represent the mean and the error bars represent standard deviation for three independent experiments (n=3). The *Rv1460stop* 1.19 mutant fit a straight line model (R^2 >0.98) and *Rv1460stop* 5.19 and 5.20 mutants fit a sigmoidal model (R^2>0.99 and >0.98 respectively). The HillSlopes and slopes were compared using multiple t-tests as detailed in section 3.16. *Rv1460stop* 1.19 (p=0.004) and 5.19 (p=0.009) grew significantly different compared to the H37Rv wild-type progenitor while 5.20 did not (p=0.196). Growth of the Δ*Rv1460*:: pMV*Rv1460compl* (p=0.839) and *Rv1460ΔDNAbd*:: pMV*Rv1460compl* (p=0.454) strains were not significantly different to the H37Rv wild-type progenitor.

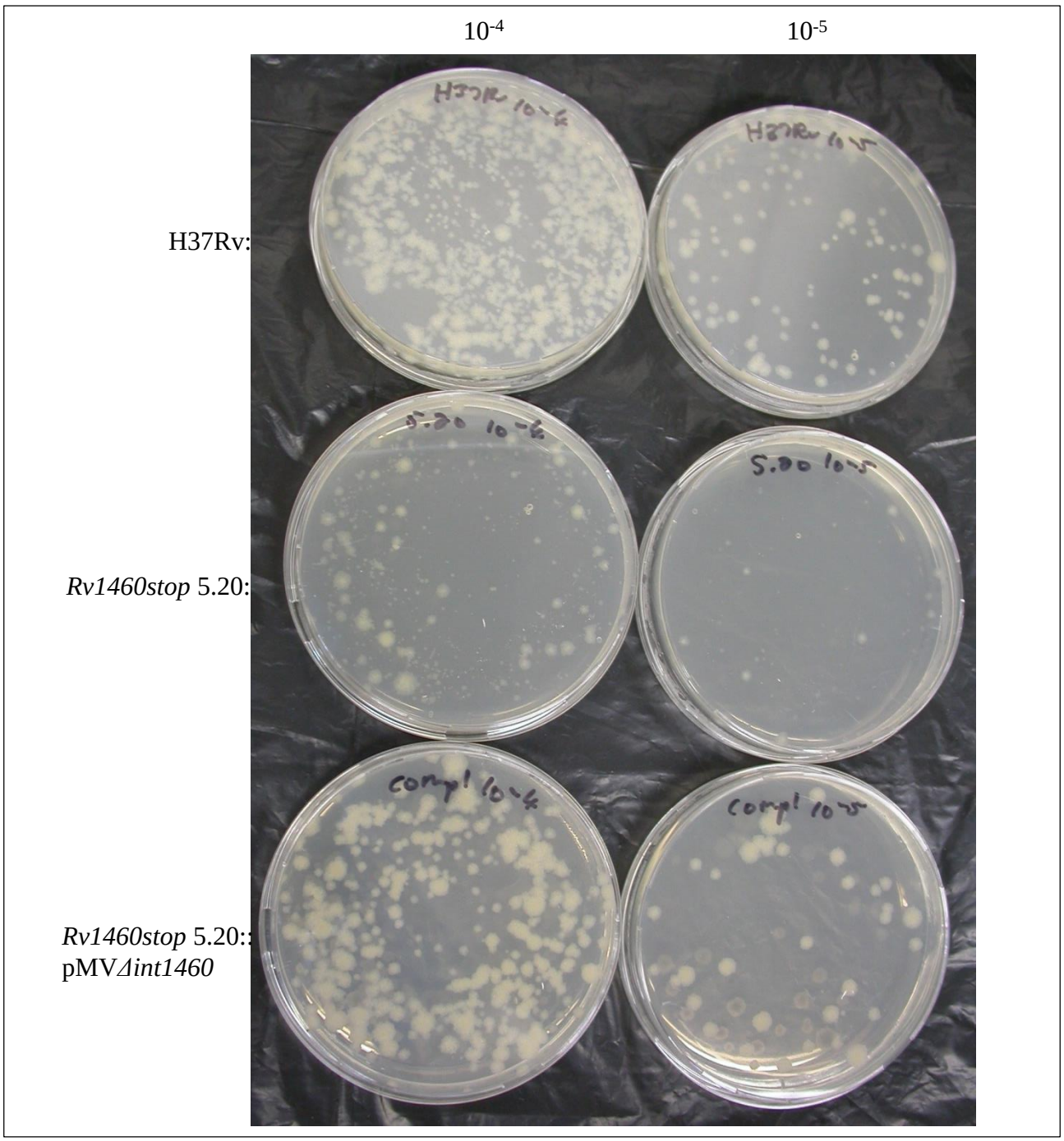


Figure 13. Growth of the *Rv1460stop* 5.20 mutant on solid media. The H37Rv wild-type, *Rv1460stop* 5.20 mutant and complementation strains were cultured on 7H10 OADC. Growth after 18 days on 7H10 OADC is indicated. Cultures were filtered through a cell strainer to remove clumps, diluted to an OD of 0.2 in 1× PBS 0.05% Tween 80, sonicated for 12 min in a waterbath sonicator and a dilution series made in 1× PBS 0.5% Tween 80. Dilutions were plate in duplicate on 7H10 OADC and growth monitored.

4.3.2 Growth in the presence of oxidative stress

During *in vitro* growth toxic peroxides may be generated during respiration and *M. tuberculosis* is therefore routinely cultured in the presence of catalase. In order to assess the effect of low level

endogenous ROS generated during respiration on the growth of the *Rv1460stop* mutant, the growth of the *Rv1460stop* 5.20 mutant was monitored in 7H9 AD without catalase. Culturing in the absence of catalase exacerbated the growth defect observed under standard culture conditions, indicating that the *Rv1460stop* mutant is more sensitive to low level ROS (Figure 14 A). To further investigate the *Rv1460stop* mutants' apparent sensitivity to ROS, it was exposed to the ROS generator, CHP, and its survival compared to that of H37Rv wild-type progenitor by determining CFUs of exposed and unexposed cultures (Figure 14 B). The *Rv1460stop* 5.20 mutant showed a reduction in survival compared to the H37Rv wild-type progenitor although not statistically significant (p=0.06) and survival was restored upon genetic complementation with Rv1460. This indicates that the *Rv1460stop* mutant is more sensitive to higher levels of ROS.

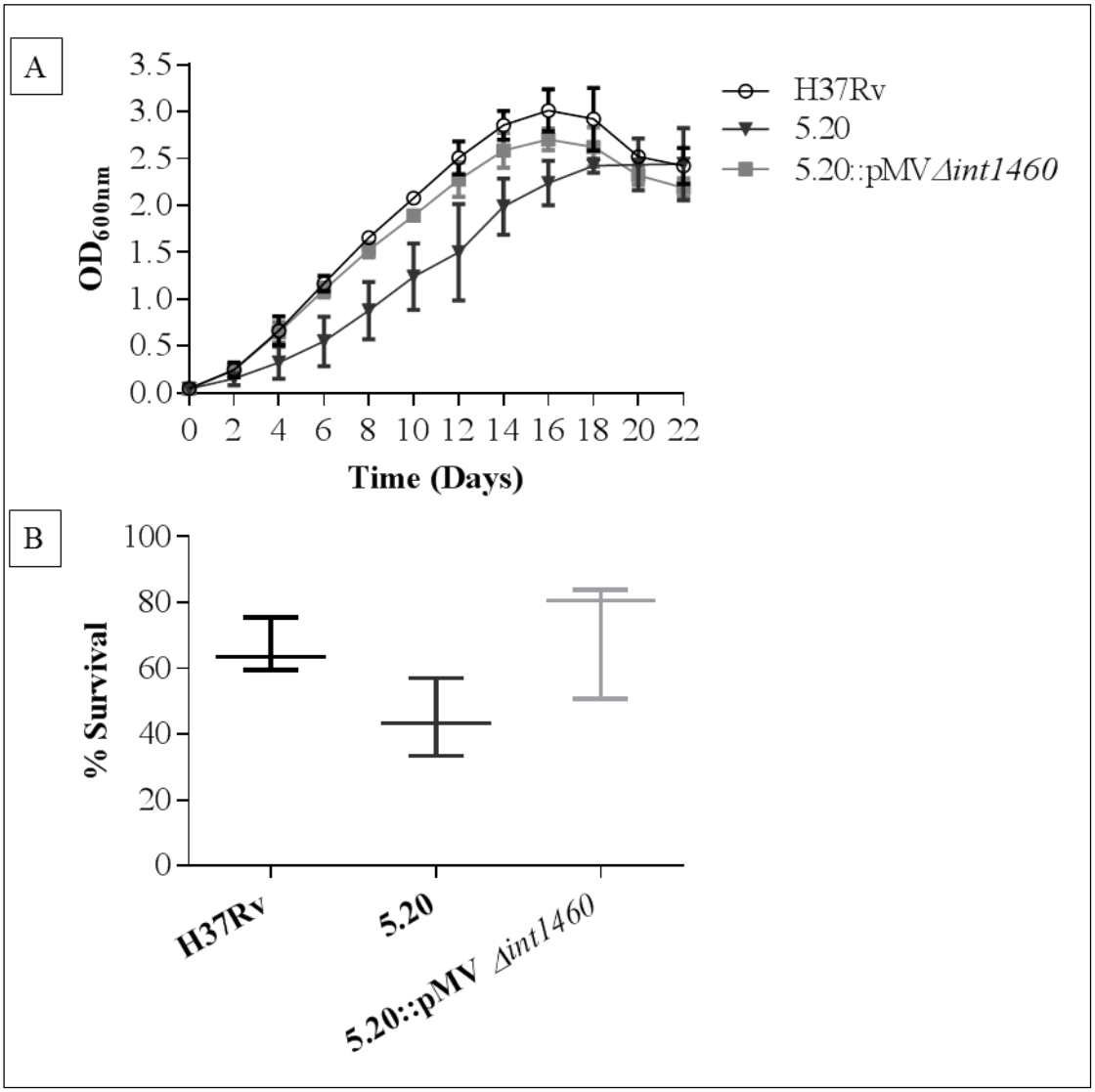


Figure 14. Growth and survival of *Rv1460stop* mutants in the presence of oxidative stress. (A) Growth of *Rv1460stop* 5.20 mutant and its complementation strain were monitored by determining OD_{600nm} readings over 22 days in 7H9 AD containing no catalase. Values represent the mean and the error bars represent standard deviation of three independent experiments (n=3). All growth curves fit a sigmoidal growth model (R^2>0.97 for H37Rv and 5.20::

pMV*Δint1460* and R^2>0.98 for *Rv1460stop* 5.20). To determine the differences in HillSlopes of the growth curves, t-tests were done and indicated that the *Rv1460stop* 5.20 mutant grew significantly different compared to the H37Rv wild-type progenitor (p=0.022). (B) Percentage survival after exposure to 250 µM CHP for 3 h. Percentage survival was calculated by dividing the CFUs average of three plates for CHP exposed cultures by the average of three plates for unexposed cultures (technical triplicate). The box and whiskers plot represents the minimum, median and maximum values for three independent experiments (n=3). No significant difference in percentage survival was observed between the H37Rv and *Rv1460stop* 5.20 mutant (p=0.06), H37Rv and complementation (p=0.65) or *Rv1460stop* 5.20 mutant and complementation (p=0.97) strains.

4.3.3 Growth under iron limiting conditions

M. tuberculosis is subjected to iron limitation within the host (Banerjee *et al.*, 2011) and the expression of the *suf* system is induced upon iron limitation in *M. tuberculosis* (Rodriguez *et al.*, 2002). We thus investigated the effects of iron limitation on *Rv1460stop* mutant growth by exposing cultures to the intracellular iron chelator, bipyridyl. The concentration of bipyridyl to use for this experiment was first established by identifying the concentration at which growth of the H37Rv wild-type progenitor was not severely affected (Figure 15 A). Growth of H37Rv was not severely affected when exposed to 60 µM bipyridyl at three or eight days after exposure and this concentration was thus used for subsequent experiments (Figure 15 A). Growth of the *Rv1460stop* 5.20 mutant was not significantly altered by bipyridyl exposure (60 µM) compared to the unexposed culture (Figure 15 B-D), indicating that the *Rv1460stop* mutant is not sensitive to iron limitation under the conditions investigated.

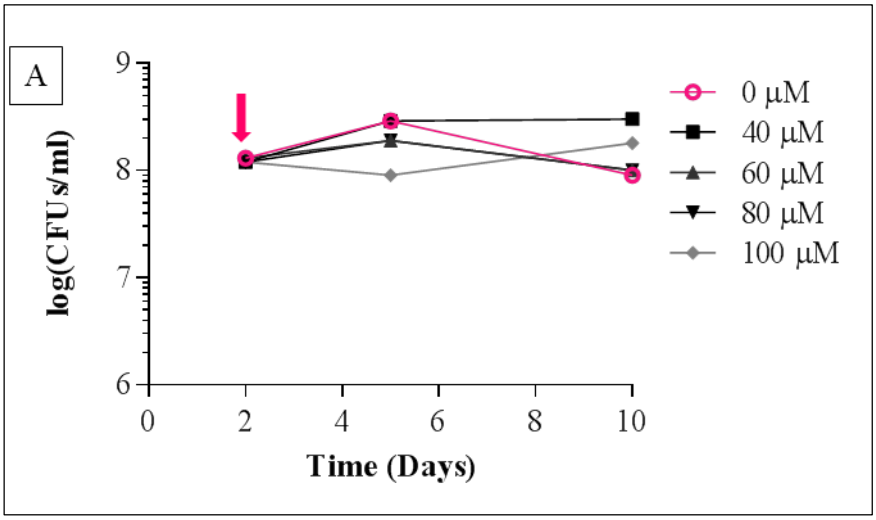


Figure 15. Growth curves of the *Rv1460stop* 5.20 mutant under iron limitation. (A) Growth of the H37Rv wild-type progenitor when exposed to varying bipyridyl concentrations. Values represent only one experiment (n=1).

Continued overleaf

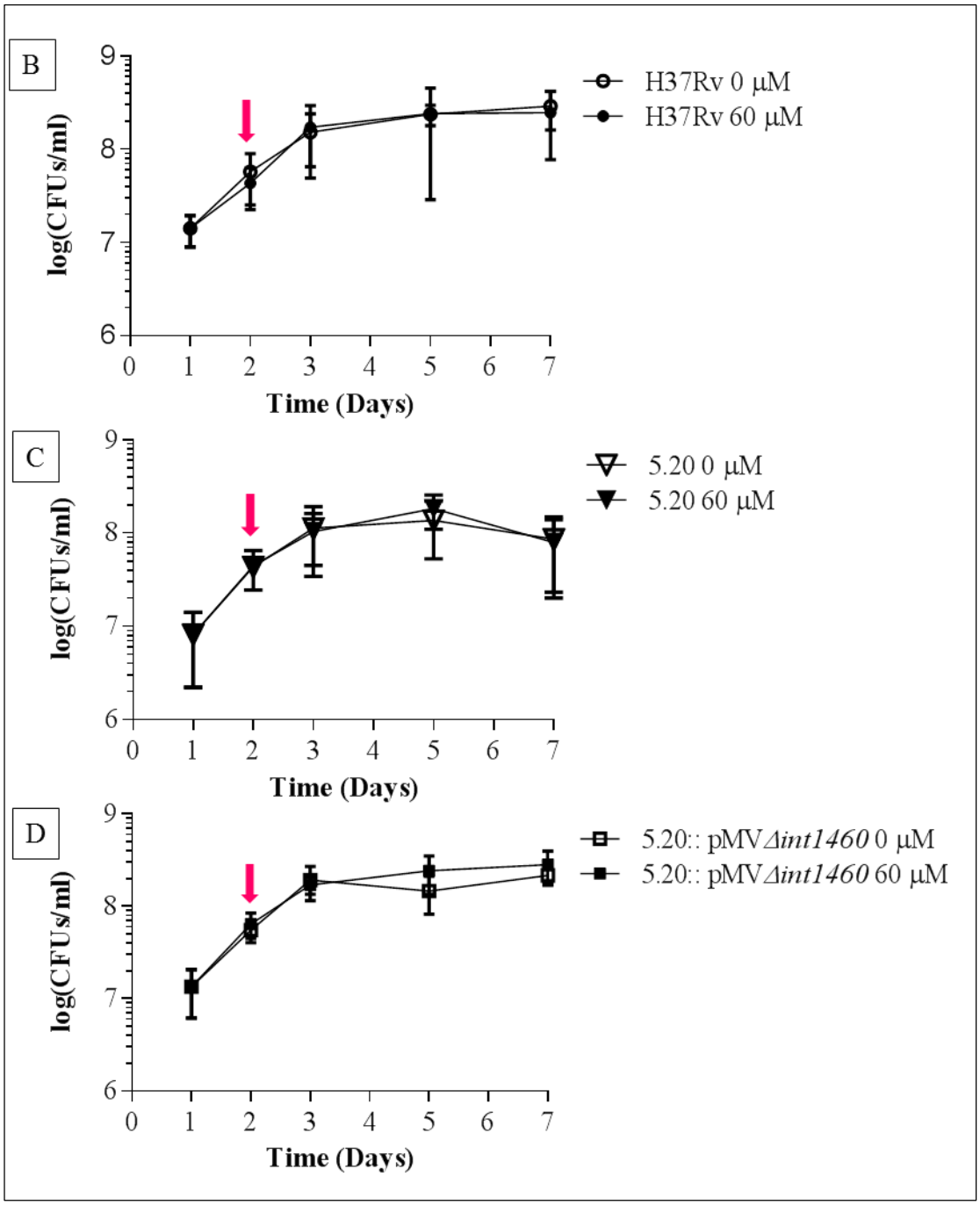


Figure 15. Growth curves of the *Rv1460stop* 5.20 mutant under iron limitation. (A) Growth of the H37Rv wild-type progenitor when exposed to varying bipyridyl concentrations. Values represent only one experiment (n=1). Growth of the (B) H37Rv wild-type progenitor, (C) *Rv1460stop* 5.20 mutant and (D) complementation strain when exposed to 60 µM bipyridyl compared to unexposed (0 µM) (exposed to the diluent of bipyridyl). Cultures were exposed to bipyridyl on day 2 as indicated by the arrow. Values represent the average of the CFUs/ml for three independent experiments (n=3) plotted on a log scale. Multiple t-tests for each day post-exposure to bipyridyl indicated no significant differences between the growth of exposed and unexposed cultures for all the strains ($p>0.05$).

4.4 Effect of *Rv1460stop* mutation on *suf* operon expression

To assess the role of Rv1460 in *suf* operon expression, the expression of *Rv1460* and *Rv1461* (as genes representative of the *suf* operon) were examined by RT-qPCR for the *Rv1460stop* 5.20 mutant. Two RT primers, one which binds within *Rv1460* and one within *Rv1461* were used (Table 6). This allowed differentiation of transcription of the native *Rv1460* gene (evaluated using the Rv1461RT primer and *Rv1460* specific PCR primers) and the total *Rv1460* trancript copies including *Rv1460* integrated at the *attB* site (evaluated using the Rv1460RT primer and *Rv1460* specific PCR primers). It should be noted that the gene specific *Rv1460* primers bind outside the region deleted in the *Rv1460stop* mutant and the *Rv1460* transcript levels could thus still be assessed for the *Rv1460stop* mutant. *Rv1460* expression was significantly higher (~5 fold) in the *Rv1460stop* 5.20 mutant compared to the wild-type H37Rv progenitor when using the Rv1460RT ($p=0.00095$) and Rv1461RT ($p<0.0001$) primers suggesting that Rv1460 acts as a repressor of its own expression. *Rv1461* expression was also significantly higher (~2.6 fold) in the mutant compared to the wild-type ($p=0.0027$) (Figure 16) suggesting that Rv1460 represses *suf* operon expression.

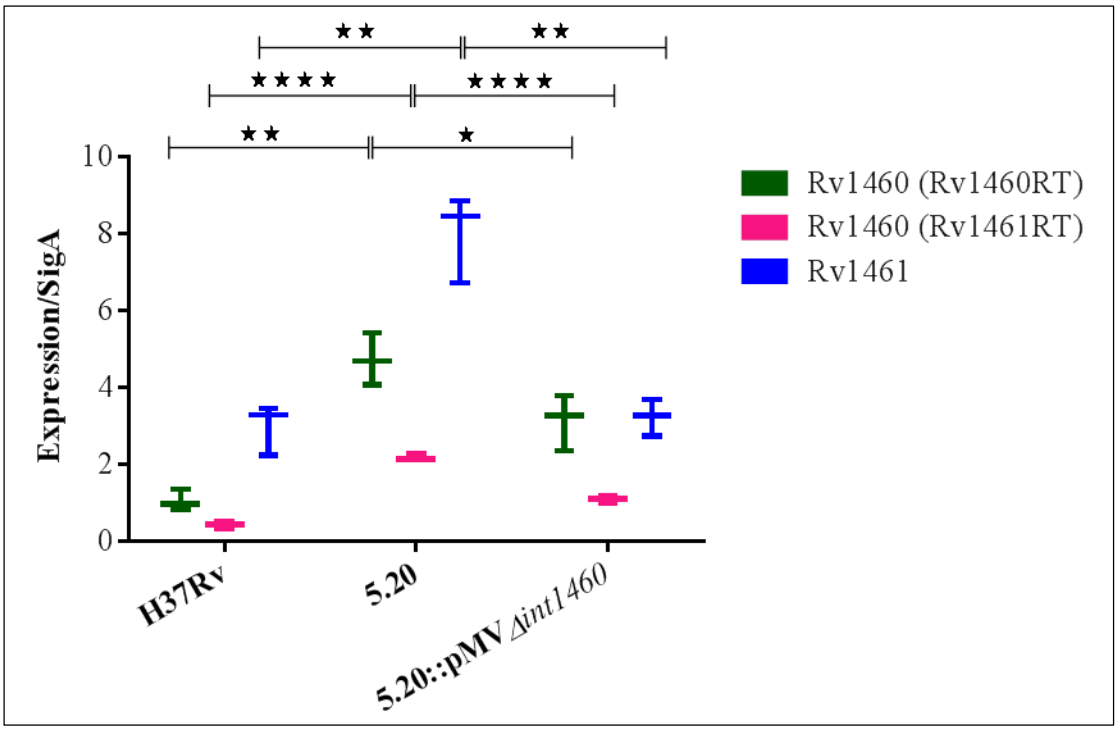


Figure 16. Relative expression of *Rv1460* and *Rv1461*. Total number of transcript copies determined by RT-qPCR for *Rv1460* and *Rv1461* divided by total number of transcript copies of the housekeeping genes, *sigA*. *Rv1460* (Rv1461RT) represents the native *Rv1460* transcripts while *Rv1460* (Rv1460RT) represents the total transcripts including the expression from the complementation vector at the *attB* site. The box and whiskers plot shows the minimum, median and maximum values of three independent experiments (n=3). Stars indicate the significant differences and level of significance between only some of the expression levels as determined by multiple t-tests.

Complementation partially restored the repression of the native copy of *Rv1460* and restored the repression of *Rv1461* to wild-type levels, thus confirming that the loss of *Rv1460* was responsible for the variation in gene expression (Figure 16). The amplification of *Rv1460* from the Rv1461RT primer product also indicates that *Rv1460* is trancribed as part of the *suf* operon. The level of *Rv1460* transcripts detected using the Rv1460RT primer was significantly higher (~2 fold) than with the Rv1461RT primer in all the strains (H37Rv: p=0.022, 5.20: p=0.003, complementation: p=0.008) (Figure 16). The expression of *Rv1461* was also significantly higher than the native copy of *Rv1460* (Rv1460 (Rv1461RT)) in all the strains (H37Rv: p=0.003 ~7 fold higher, 5.20: p=0.0009 ~3 fold higher, complementation: p=0.001 ~3 fold higher). This may indicate that a promoter element is present upstream of *Rv1461*.

4.5 Generation of vectors for the production of Rv1460 protein in *E. coli*

To characterise Rv1460 and investigate its function as a transcriptional regulator, we sought to produce and purify the recombinant protein in *E. coli*. To produce *Rv1460* in *E. coli*, the *Rv1460N* and *Rv1460C* regions (Table 7), which include the whole gene (excluding the stop codon in *Rv1460C*), were amplified (Supplementary Figure 4 A-B) and cloned into pET28 and pET21 vectors respectively and confirmed by RE mapping (Supplementary Figure 4 C-D) and Sanger sequencing. Protein production strains were transformed with the resulting vectors.

4.6 Optimisation of Rv1460 protein production in *E. coli*

Production of His-tagged Rv1460 using the pET28*Rv1460N* and pET21*Rv1460C* vectors was initially tested in the Rosetta-gami 2 strain. No visible protein production was observed in this strain from either vector (Figure 17 A-D). The temperature and IPTG concentration was decreased from 37 °C to 20 °C and 1 mM to 0.4 mM respectively and the time of induction increased from 4 to 24 h. A band was observed at ~34 kDalton (kDa) instead of the expected size of 30. 4 kDa. This band was present in the soluble and insoluble fractions, but predominates in the insoluble fraction (Figure 17 E). Culturing in a richer broth (terrific broth) did not increase the proportion of Rv1460 in the soluble fraction (Figure 17 F). As a lower temperature seemed to improve Rv1460 production, production was tested in Arctic express (DE3) from the pET28*Rv1460N* and pET21*Rv1460C* vectors. His-tagged Rv1460 was observed in both fractions, but predominantly in the insoluble fraction (Figure 17 G-J). There was no difference in protein production between the 24 and 48 h induction at 13 °C (Figure 17 G-J). A further decrease in temperature from 13 °C to 10 °C and culturing in terrific broth did not

increase the proportion of soluble protein (Figure 17 K-L). Induction with 1 mM versus 0.4 mM IPTG also did not alter the protein production (results not shown). Protein was subsequently produced using the pET28*Rv1460N* vector in Arctic express (DE3) induced with 0.4 mM IPTG at 13 °C for 24 h.

The His-tag was detected during western blot using an anti-His antibody (data not shown). The protein was purified on small-scale using the MagneHis purification kit, the product separated on a SDS-PAGE gel and analysed using mass spectrometry to confirm the protein identity.

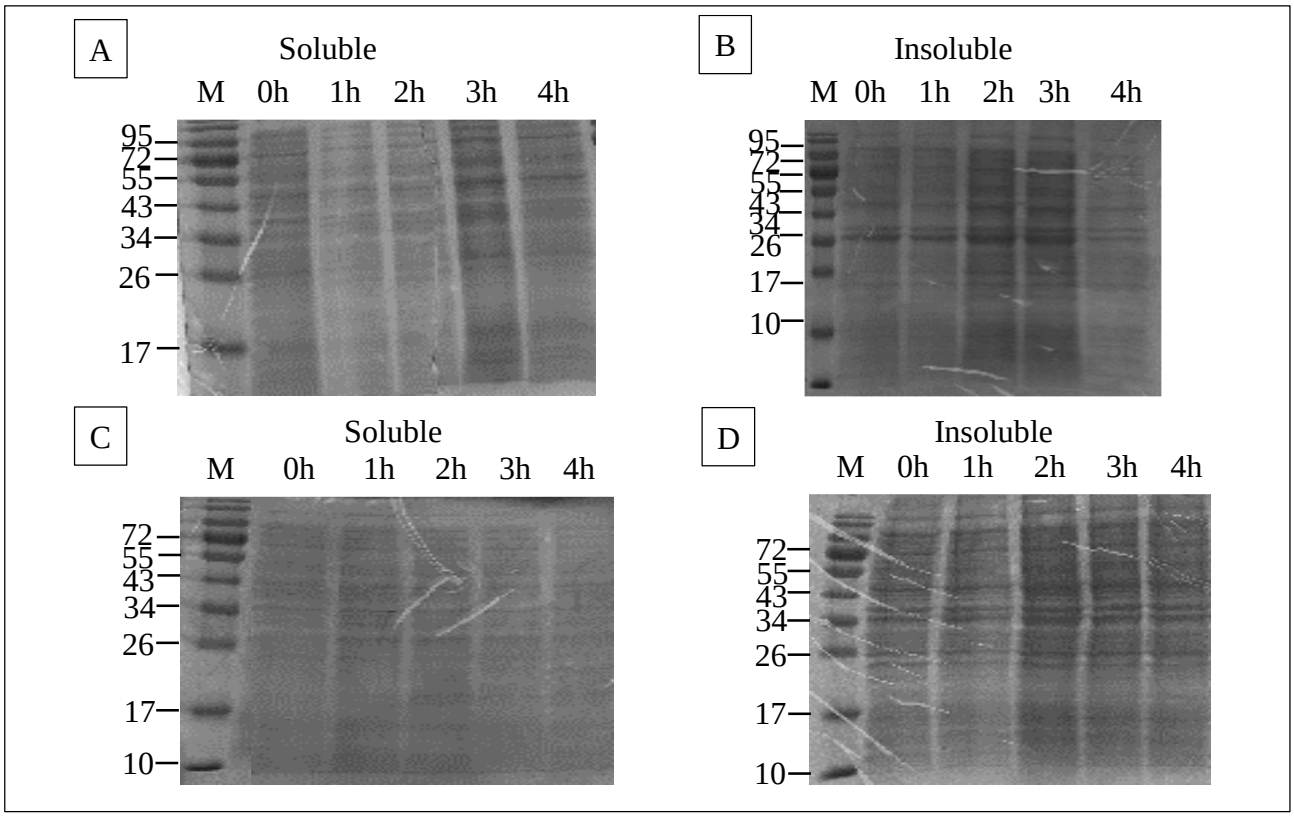


Figure 17. SDS-PAGE of Rv1460 production in Rosetta-gami 2 and Arctic express (DE3). Rv1460 production using the pET28*Rv1460N* ((A) soluble and (B) insoluble) and pET21*Rv1460C* ((C) soluble and (D) insoluble) vectors in Rosetta-gami 2 cultured in LB broth and induced for 4 h with 1 mM IPTG at 37 °C. Rv1460 production using the pET28*Rv1460N* vector in Rosetta-gami 2 cultured in (E) LB and (F) terrific broth and induced for 24 h with 0.4 mM IPTG at 20 °C. Rv1460 production using the pET28*Rv1460N* ((G) soluble and (H) insoluble) and pET21*Rv1460C* ((I) soluble and (J) insoluble) vectors in Arctic express (DE3) cultured in LB broth and induced for 48 h with 1 mM IPTG at 13 °C. Rv1460 production of Rv1460 from the pET28*Rv1460N* vector ((K) soluble and (L) insoluble) in Arctic express (DE3) cultured in LB and terrific broth and induced for 48 h with 1 mM IPTG at 10 °C.

Continued overleaf

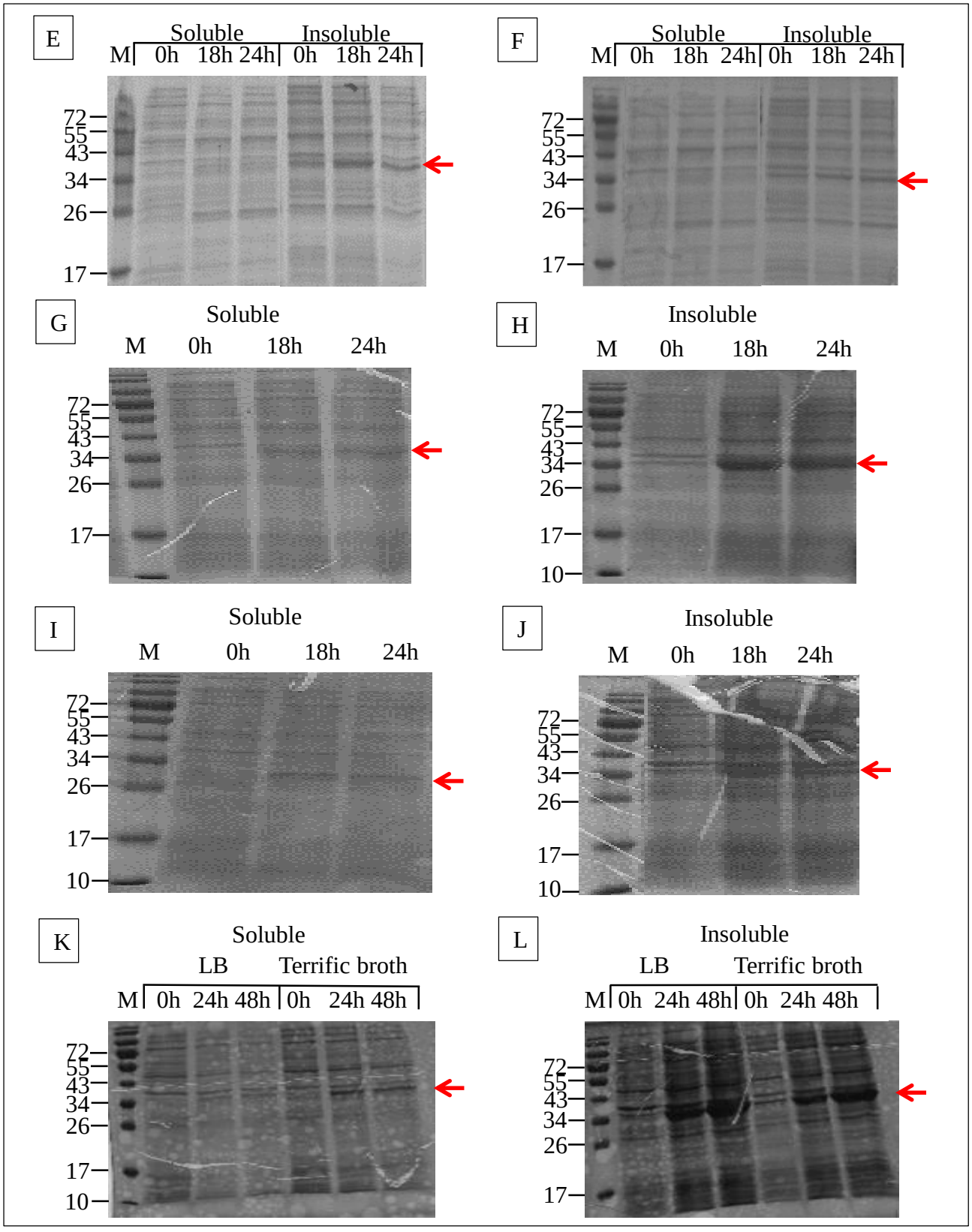


Figure 17. SDS-PAGE of Rv1460 production in Rosetta-gami 2 and Arctic express (DE3).

Rv1460 production using the pET28*Rv1460N* ((A) soluble and (B) insoluble) and pET21*Rv1460C* ((C) soluble and (D) insoluble) vectors in Rosetta-gami 2 cultured in LB broth and induced for 4 h with 1 mM IPTG at 37 °C. Rv1460 production using the pET28*Rv1460N* vector in Rosetta-gami 2 cultured in (E) LB and (F) terrific broth and induced for 24 h with 0.4 mM IPTG at 20 °C. Rv1460 production using the pET28*Rv1460N* ((G) soluble and (H) insoluble) and pET21*Rv1460C* ((I) soluble and (J) insoluble) vectors in Arctic express (DE3) cultured in LB broth and induced for 48 h

with 1 mM IPTG at 13 °C. Rv1460 production of Rv1460 from the pET28*Rv1460N* vector ((K) soluble and (L) insoluble) in Arctic express (DE3) cultured in LB and terrific broth and induced for 48 h with 1 mM IPTG at 10 °C.

4.7 Generation of vectors for the production of Rv1460 C→S protein variants in *E. coli*

Three conserved cysteine residues (C203, C216 and C244) predicted to coordinate an Fe-S cluster in Rv1460 (Cole *et al.*, 1998; Shen *et al.*, 2007) and an additional cysteine located in the vicinity (C242) were replaced by serine residues by site-directed mutagenesis to create Rv1460 C203S, C216S, C242S, C244S and C203,216,244S protein variants. The presence of the single nucleotide changes in each pET28 construct was confirmed by Sanger sequencing. The Rv1460 cysteine to serine protein variants were produced as N-terminally His-tagged proteins in Arctic express (DE3) under the same conditions as determined optimal for Rv1460 production.

4.8 Large-scale purification of Rv1460 and Rv1460 C203,216,244S proteins

As larger quantities of protein were essential for subsequent experiments, the purification of the protein on large-scale was optimised. Only the results for the large-scale purification of Rv1460 and the Rv1460 C203,216,244S variant protein are shown. His-tagged Rv1460 (from a 2.5 L culture) and Rv1460 C203,216,244S variant were eluted from a Ni-NTA column using a 20 mM to 350 mM imidazole gradient. In both cases, two peaks were observed on the Ni-NTA elution profile (Figure 18 A-B). Both monomeric and higher molecular weight species that correspond to the size of a dimer were identified by denaturing SDS-PAGE (Figure 18 C-D). A western blot using an anti-His antibody detected a major band at ~34 kDa, as well as a minor band at ~70 kDa, corresponding to the size of a dimer which was only present in the second peak (Figure 18 E-F). The fractions corresponding to the second peak were pooled, concentrated and further purified by gel filtration.

The gel filtration elution profile of Rv1460 showed a complex peak with at least two shoulders that could correspond to a tetramer, dimer and monomer according to the gel filtration standards (Figure 18 G-H). A more defined “tetramer” shoulder was observed for Rv1460 C203,216,244S, while the “monomer” shoulder was less pronounced (Figure 18 G-H). The smaller bands observed on the gels and blots represent degraded protein (Figure 18 I-L). Protein variants Rv1460, Rv1460 C203S, C216S, C242S, C244S and C203,216,244S were also purified using a 5 ml Ni-NTA column packed using His-Pur Ni-NTA Superflow agarose using step-wise elution (results not shown). The His-tag

was subsequently cleaved with thrombin and His-tag cleaved protein used for subsequent experiments.

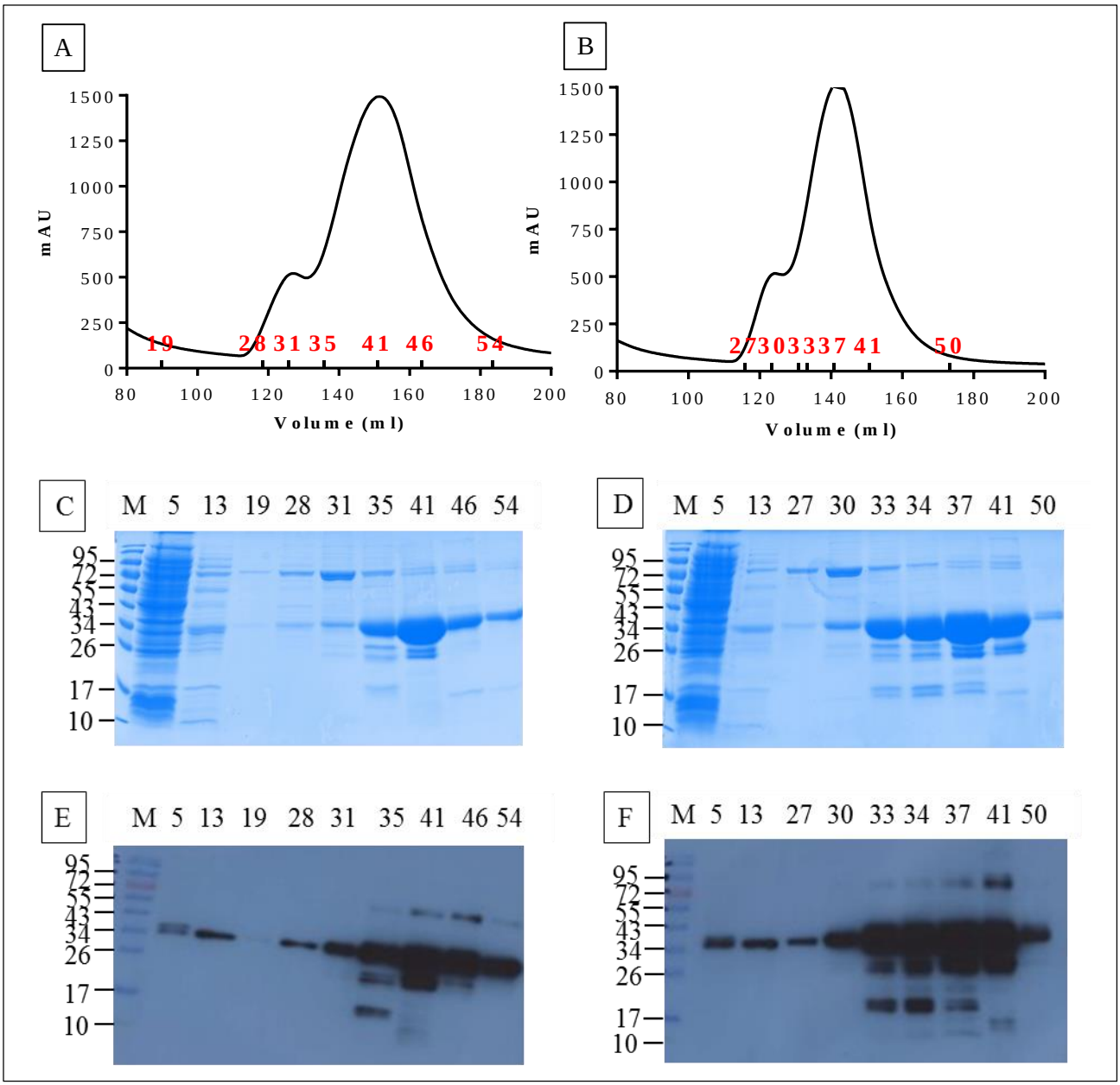


Figure 18. Large-scale purification of His-tagged Rv1460 and Rv1460C203,216,244S. Ni-NTA elution profile for (A) Rv1460 and (B) Rv1460C203,216,244S variant protein. SDS-PAGE of (C) Rv1460 and (D) Rv1460C203,216,244S Ni-NTA fractions (fractions are indicated in red on the inside of the axis). Western blots using an anti-His antibody of (E) Rv1460 and (F) Rv1460C203,216,244S Ni-NTA fractions. Sephadex S200 (16/600) elution profiles of (G) Rv1460 and (H) Rv1460 C203,216,244S. The position of elution of the gel filtration standards (in kDa) are indicated in red on the inside of the axis. SDS-PAGE of (I) Rv1460 and (J) Rv1460 C203,216,244S and western blot using an anti-His antibody of (K) Rv1460 and (L) Rv1460 C203,216,244S for gel filtration fractions (in time (min)). Continued overleaf

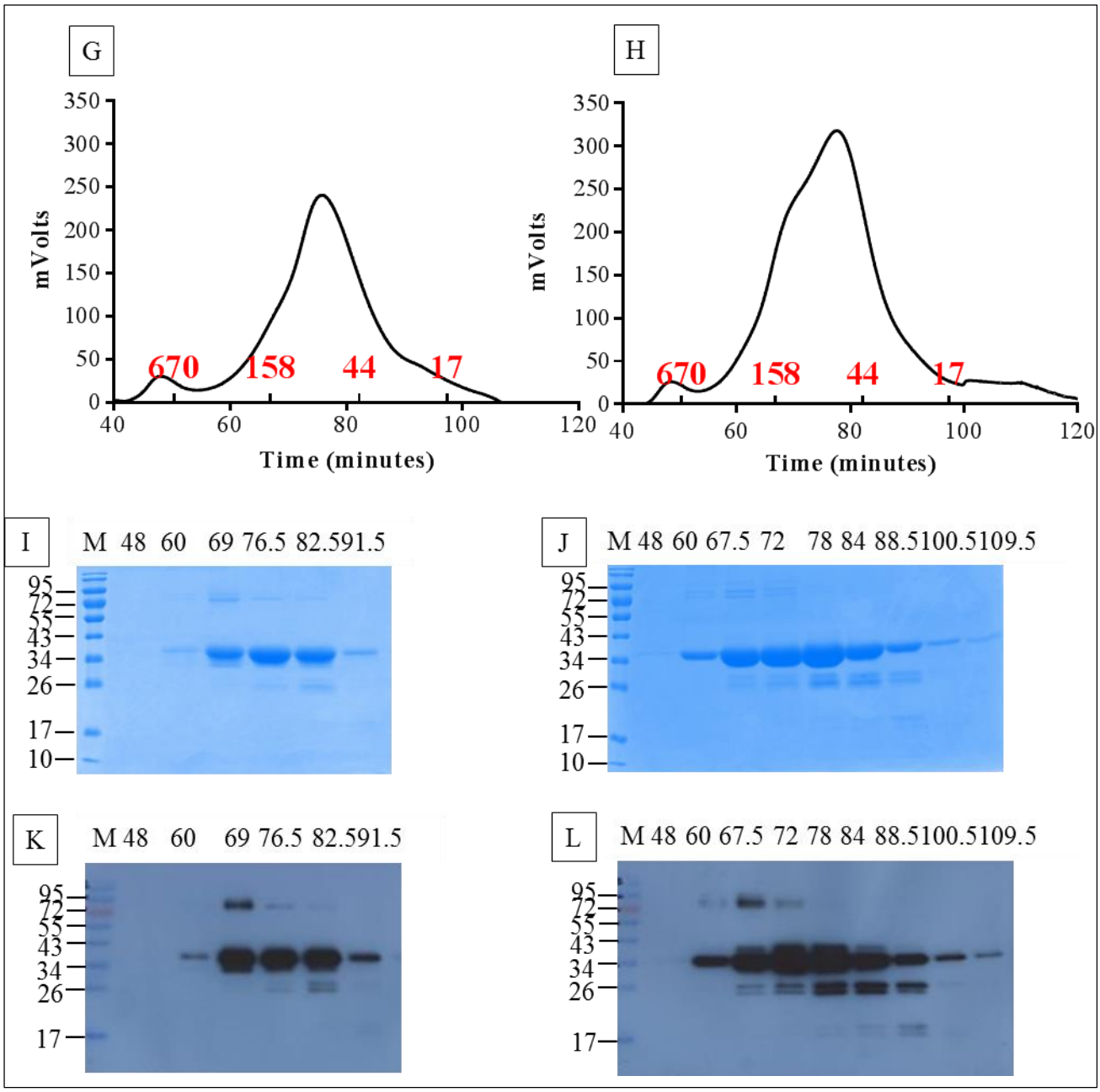


Figure 18. Large-scale purification of His-tagged Rv1460 and Rv1460C203,216,244S. Ni-NTA elution profile for (A) Rv1460 and (B) Rv1460C203,216,244S variant protein. SDS-PAGE of (C) Rv1460 and (D) Rv1460C203,216,244S Ni-NTA fractions (fractions are indicated in red on the inside of the axis). Western blots using an anti-His antibody of (E) Rv1460 and (F) Rv1460C203,216,244S Ni-NTA fractions. Sephadex S200 (16/600) elution profiles of (G) Rv1460 and (H) Rv1460 C203,216,244S. The position of elution of the gel filtration standards (in kDa) are indicated in red on the inside of the axis. SDS-PAGE of (I) Rv1460 and (J) Rv1460 C203,216,244S and western blot using an anti-His antibody of (K) Rv1460 and (L) Rv1460 C203,216,244S for gel filtration fractions (in time (min)).

Since protein solubility does not necessarily correlate with proper folding, the presence of secondary structure of His-tag cleaved Rv1460 was investigated by CD to investigate proper folding of Rv1460. CD is a low resolution tool to identify the overall structural features of a protein. Individual secondary structure elements result in characteristic CD spectra in far-UV (Kelly *et al.*, 2005). The far-UV CD spectrum of the purified Rv1460 protein was predominated by two negative bands at 208 and 222 nm

(Figure 19). This indicates the presence of an α-helix and supports the idea that the protein is properly folded (Kelly *et al.*, 2005).

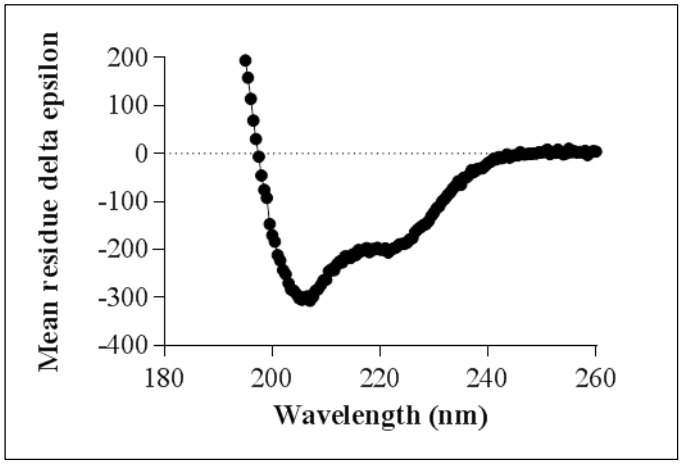


Figure 19. Far-UV CD spectrum of Rv1460. His-tag cleaved Rv1460 (0.17 mg/ml) was subjected to far-UV CD using a Chirascan CD Spectrometer and step wise (0.5 nm step size) measurements over a range of wavelengths (195 nm – 260 nm) using a 1 cm quartz cuvette.

4.9 Generation of vector for the production of IscS$_{Mtb}$ protein in *E. coli*

M. tuberculosis IscS (IscS$_{Mtb}$) is a cysteine desulphurase able to provide sulphur for Fe-S cluster reconstitution (Rybniker *et al.*, 2014). The ability of IscS$_{Mtb}$ to reconstitute an Fe-S cluster on Rv1460 *in vitro* was investigated. *IscS$_{Mtb}$* was amplified from H37Rv genomic DNA (Supplementary Figure 5 A) and cloned into the pET28 vector to create pET28iscS and confirmed by RE mapping (Supplementary Figure 5 B). IscS$_{Mtb}$ was subsequently produced as a N-terminally His-tagged protein and purified for use in reconstitution experiments.

4.10 Optimisation of IscS$_{Mtb}$ protein production in *E. coli*

Production of N- terminally His-tagged IscS$_{Mtb}$ was optimised by testing production in Rosetta 2 and Arctic express (DE3) protein production strains in different growth media at varying temperatures. Production of IscS$_{Mtb}$ was successful in Rosetta 2 in LB broth at 37 °C for 4 h, although the bulk of the protein was insoluble (Figure 20 A). Production of IscS$_{Mtb}$ in Arctic express (DE3) in LB broth at 13 °C was poor, but improved after culturing in terrific broth (Figure 20 B). Protein production in Rosetta 2 at a lower temperature for extended periods in LB and terrific broth increased the production of IscS$_{Mtb}$ (Figure 20 C-D). Optimal production of recombinant IscS$_{Mtb}$ was achieved using Rosetta 2 strains cultured in terrific broth and induced with 0.4 mM IPTG for 24 h at 18 °C. Production in

Arctic express (DE3) in terrific broth induced with 0.4 mM IPTG for 24 h at 13 °C was also successful (Figure 20 E).

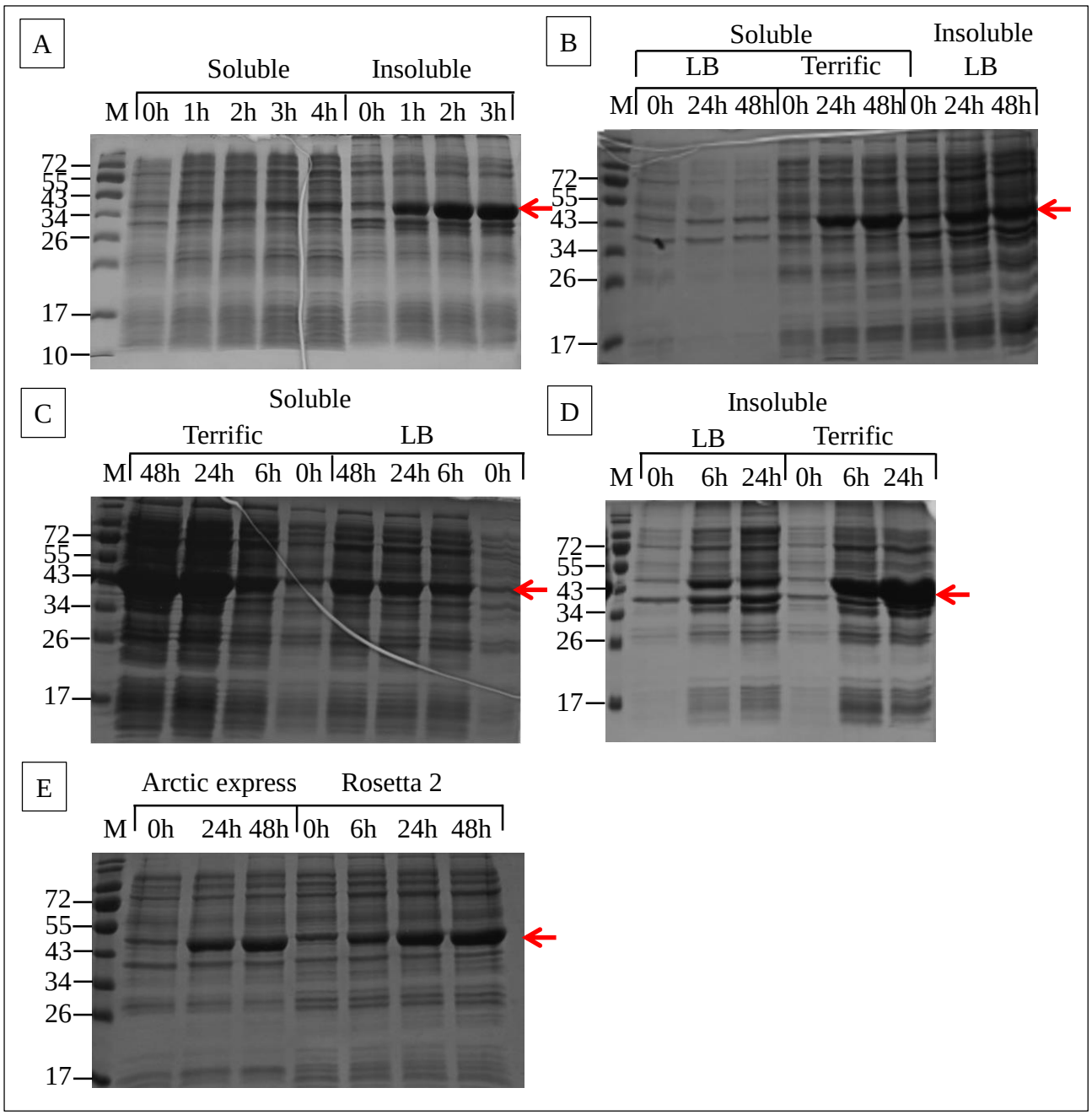


Figure 20. SDS-PAGE of IscS$_{Mtb}$ production in Rosetta 2 and Arctic express (DE3). IscS$_{Mtb}$ production using the pET28iscS vector in (A) Rosetta 2 cultured in LB broth and induced with 0.4 mM IPTG for 4 h at 37 °C and (B) Arctic express (DE3) cultured in LB and terrific broth and induced with 0.4 mM IPTG for 48 h at 13 °C. IscS$_{Mtb}$ from pET28iscS in Rosetta 2 cultured in LB and terrific broth and induced with 0.4 mM IPTG for 48 h at 18 °C in the (C) soluble and (D) insoluble fractions. (D) Comparison of the production levels of IscS$_{Mtb}$ in Arctic express (DE3) and Rosetta 2 cultured in terrific broth at 13 °C and 18 °C for 48 h. Equal amounts of protein were loaded following Bradford protein concentration determination.

4.11 Large-scale purification of IscS*Mtb* protein

Because high concentrations of IscS*Mtb* were required for Fe-S cluster reconstitution experiments, larger amounts of His-tagged IscS*Mtb* were purified (from 0.5 L culture) by nickel affinity purification resulting in two peaks on the Ni-NTA elution profile (Figure 21 A). Monomeric IscS*Mtb* (~43 kDa) was detected in both fractions with denaturing SDS-PAGE (Figure 21 B) and western blot using an anti-His antibody (Figure 21 C). Both peaks were pooled and concentrated for gel filtration. His-tagged IscS*Mtb* eluted as a single band from the gel filtration column and the size corresponded to dimer according to the gel filtration column standards (Figure 21 D) Analysis of the gel filtration peak fractions by denaturing SDS-PAGE (Figure 21 E) and western blot (Figure 21 F) revealed only the monomer. The His-tag of IscS*Mtb* was subsequently cleaved and the uncleaved IscS*Mtb* removed by binding to a Ni-NTA column. His-tag cleaved protein collected in the flow through was used for subsequent experiments. The activity of IscS*Mtb* was confirmed by measuring conversion of cysteine to alanine.

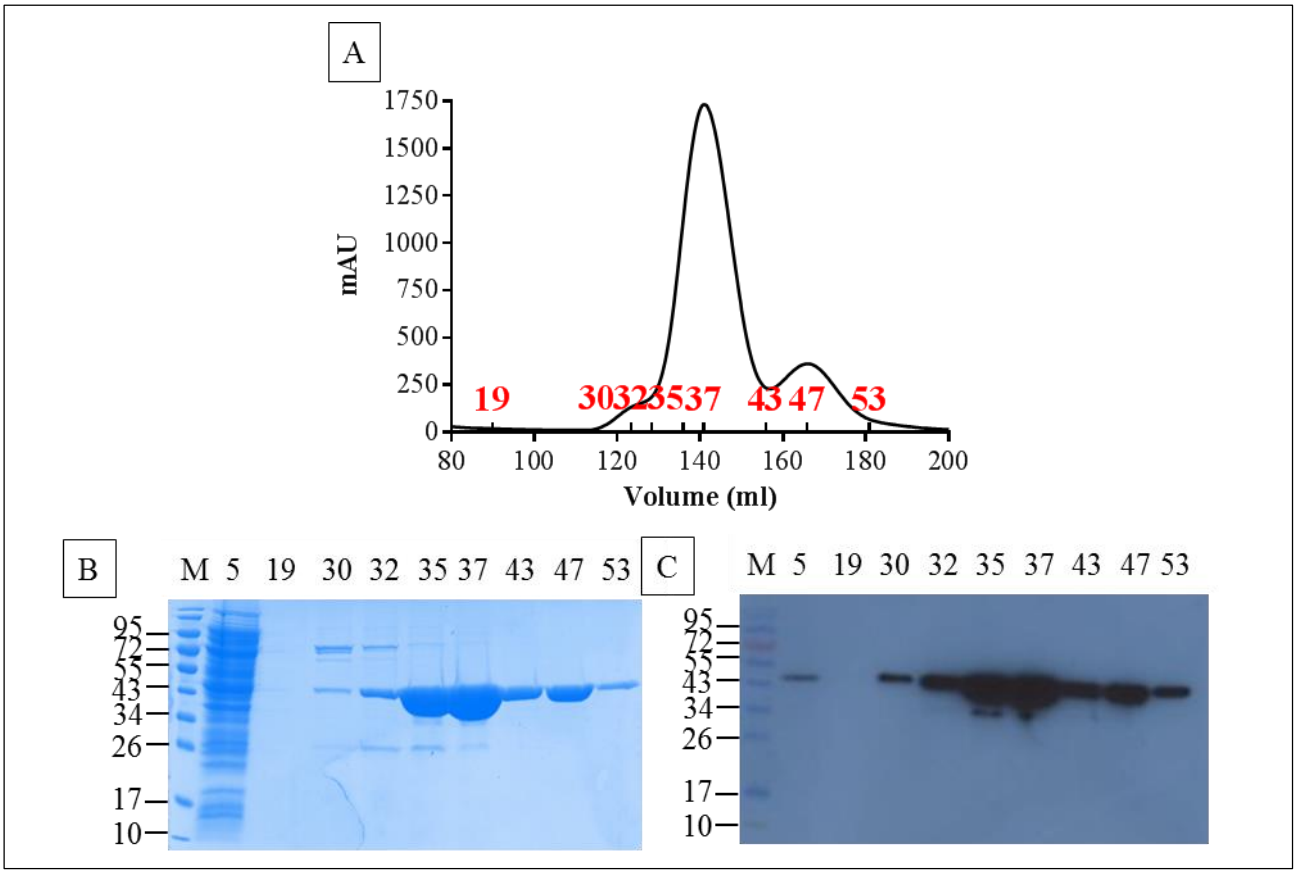


Figure 21. Large-scale purification of His-tagged IscS*Mtb*. Ni-NTA elution profile for the purification of (A) IscS*Mtb* (fractions are indicated in red on the inside of the axis). (B) SDS-PAGE and (C) western blot using an anti-His antibody of Ni-NTA fractions. (D) Sephadex S200 (16/600) elution profile of IscS*Mtb*. (E) SDS-PAGE and (F) western blot using an anti-His antibody of IscS*Mtb* gel filtration fractions. The position of elution of the gel filtration standards (in kDa) are indicated in red on the inside of the axis. Continued overleaf.

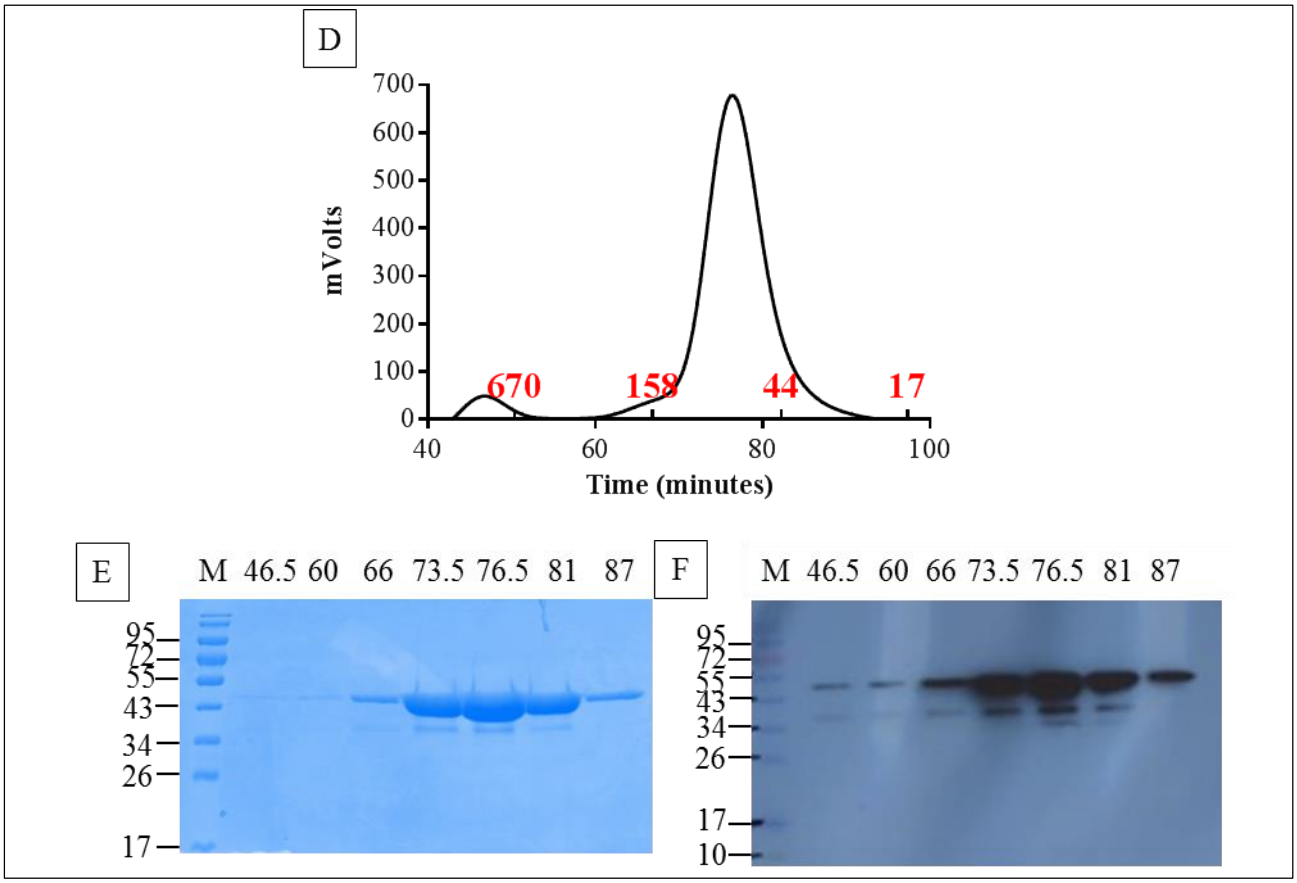


Figure 21. Large-scale purification of His-tagged IscS$_{Mtb}$. Ni-NTA elution profile for the purification of (A) IscS$_{Mtb}$ (fractions are indicated in red on the inside of the axis). (B) SDS-PAGE and (C) western blot using an anti-His antibody of Ni-NTA fractions. (D) Sephadex S200 (16/600) elution profile of IscS$_{Mtb}$. (E) SDS-PAGE and (F) western blot using an anti-His antibody of IscS$_{Mtb}$ gel filtration fractions. The position of elution of the gel filtration standards (in kDa) are indicated in red on the inside of the axis.

4.12 Fe-S cluster reconstitution

The ability of purified IscS$_{Mtb}$ to reconstitute an Fe-S cluster on purified Rv1460 *in vitro* under anaerobic conditions was investigated. The reconstitution reaction contained cysteine, iron and DTT and was monitored by UV-visible spectroscopy over time. No increase in absorbance between 410 and 450 nm was observed suggesting that no Fe-S cluster was formed on Rv1460 (Figure 22 A). This suggests that IscS$_{Mtb}$ is not able to directly donate sulphur for Fe-S cluster assembly on Rv1460.

Subsequently we attempted to use an established *in vitro* reconstitution system that utilises purified IscS$_{E.coli}$ and IscU$_{E.coli}$ for reconstitution (Adinolfi *et al.*, 2009). An increase in absorbance was initially observed at 410 nm and 450 nm, suggesting Fe-S cluster reconstitution on Rv1460. Over time the absorbance transitioned to a single maximum at 420 nm (Figure 22 B). Absorbance peaks within this range is suggestive of the initial formation of a 2Fe-2S, which is then converted to a 4Fe-

4S cluster (Adinolfi *et al.*, 2009; Agar *et al.*, 2000; Ryle *et al.*, 1996). No increase in absorbance was observed when both $IscS_{E.coli}$ and $IscU_{E.coli}$, or $IscU_{E.coli}$ alone were omitted from the reaction. To further characterise the cluster, CD was used. The near-UV CD spectrum recorded for Rv1460 after the reconstitution reaction contained maximum values at 340, 420 and 510 nm and minimum values at 370 nm and 560 nm (Figure 22 C) indicating the presence of a 2Fe-2S cluster (Adinolfi *et al.*, 2009; Bonomi *et al.*, 2005). Addition of the $IscU_{E.coli}$ to the $IscS_{Mtb}$ reconstitution reaction did not assist in cluster reconstitution (data not shown).

The role of the conserved cysteine residues at positions 203, 216 and 244 in co-ordinating the Fe-S cluster in Rv1460 was investigated by the reconstitution experiments with purified Rv1460 variants C203S, C216S, C244S, C242S and C203,216,244S. UV-visible spectroscopy of reconstitution reactions for C203S, C216S and C203,216,244S variants are shown (Figure 22 D-F). In all cases, cluster reconstitution was observed, suggesting that the serine residues coordinate the cluster *in vitro* under anaerobic conditions (Figure 22 D-F).

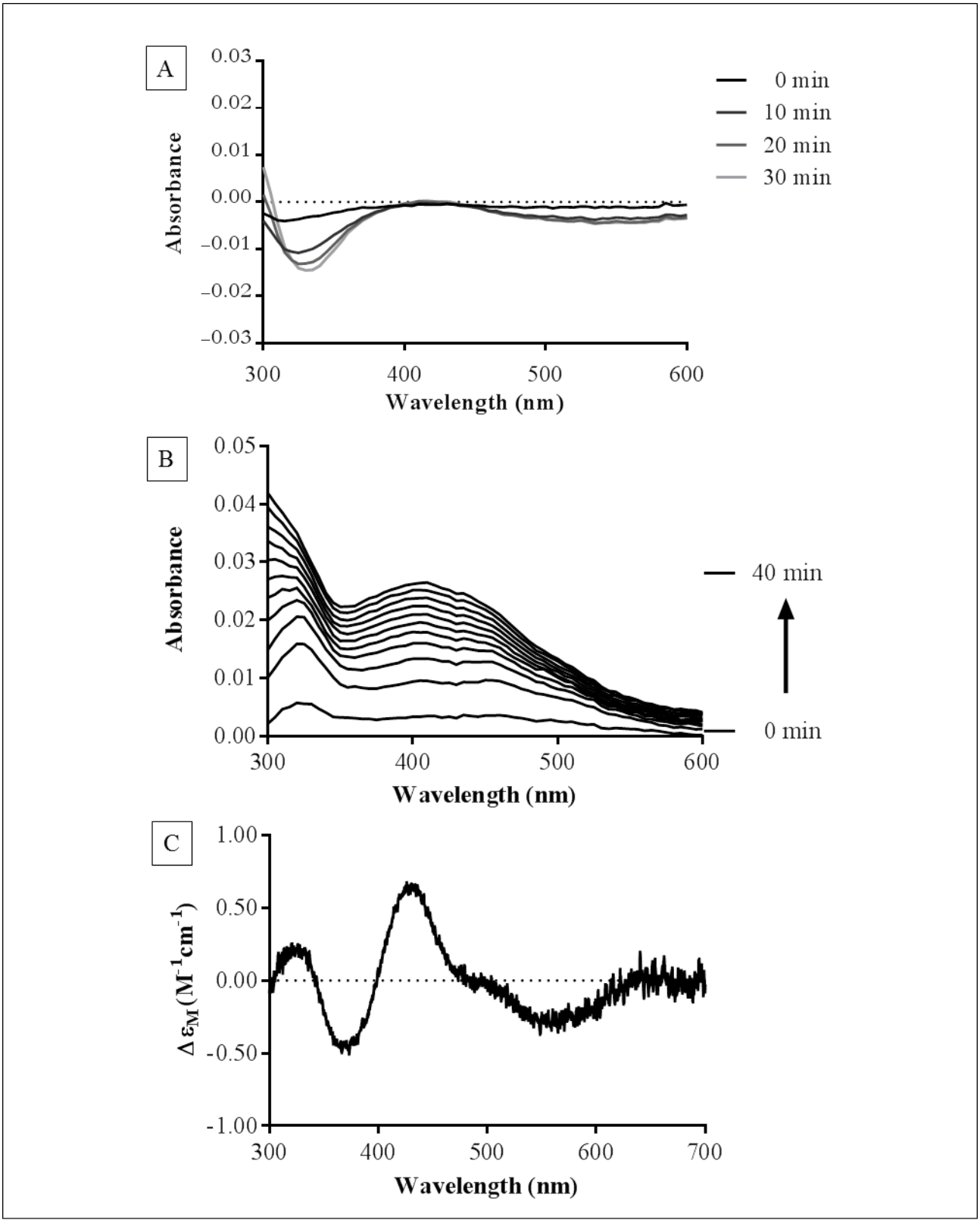


Figure 22. *In vitro* reconstitution on purified Rv1460. UV-visible spectrum monitored over time for reconstitution of Rv1460 (A) using $IscS_{Mtb}$ and (B) $IscS_{E.coli}$ and $IscU_{E.coli}$. (C) Near-UV CD spectrum of reconstituted Rv1460 indicating the presence of a 2Fe-2S cluster. UV-visible spectrum monitored over time for reconstitution of (D) Rv1460 C203, (E) Rv1460 C242 and (F) Rv1460 C203,216,244S variants Fe-S clusters using $IscS_{E.coli}$ and $IscU_{E.coli}$. Continued overleaf

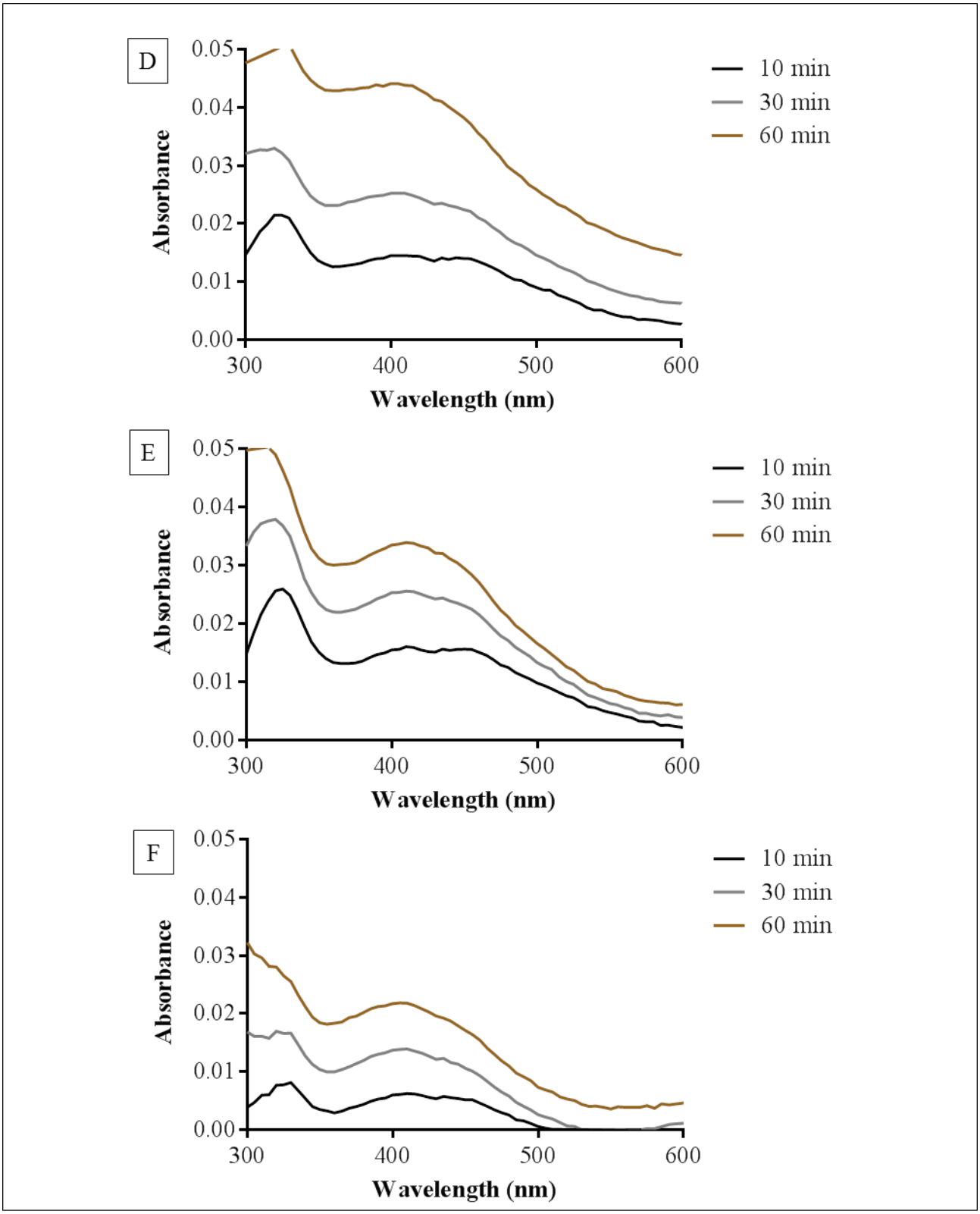


Figure 22. *In vitro* reconstitution on purified Rv1460. UV-visible spectrum monitored over time for reconstitution of Rv1460 (A) using $IscS_{Mtb}$ and (B) $IscS_{E.coli}$ and $IscU_{E.coli}$. (C) Near-UV CD spectrum of reconstituted Rv1460 indicating the presence of a 2Fe-2S cluster. UV-visible spectrum monitored over time for reconstitution of (D) Rv1460 C203, (E) Rv1460 C242 and (F) Rv1460 C203,216,244S variants Fe-S clusters using $IscS_{E.coli}$ and $IscU_{E.coli}$.

4.13 Electrophoretic mobility shift assay

ChIP-seq data indicated binding sites for Rv1460 within the *Rv1460* promoter region and within *Rv1461*. The ChIP-seq peak centers were present 20 bp upstream of the newly annotated start codon of *Rv1460* and 524 bp downstream of the *Rv1461* start codon (Minch *et al.*, 2015). Because the exact binding motif for Rv1460 is unknown, extended regions around the two ChIP-seq centers were investigated for Rv1460 binding. The region upstream of *Rv1460* (including 190 bp upstream and 122 bp downstream of the ChIP-seq peak center) and a region within *Rv1461* (including 166 bp upstream and 169 bp downstream of the ChIP-seq peak center) were amplified using biotin-labelled forward primers indicated in Table 7. These regions are subsequently referred to as the *Rv1460* promoter and *Rv1461* binding regions.

The binding of "as purified" Rv1460 (without a His-tag) to these regions were investigated by EMSA. Since Rv1460 was purified under aerobic conditions, it may be assumed that the protein does not contain an Fe-S cluster and is thus in its apo-Rv1460 form. A gradual increase in the shift of the *Rv1460* promoter region DNA was observed with increasing protein concentrations indicating that Rv1460 binds to the *Rv1460* promoter region in a concentration dependent manner (Figure 23 A). This binding is specific as increasing amounts of unlabelled *Rv1460* promoter region DNA increased the amount of unbound biotin-labelled *Rv1460* promoter DNA (Figure 23 B). Even at a 200-fold molar excess of unlabelled DNA versus biotin-labelled DNA, the interaction between Rv1460 (2.5 μM) and biotin-labelled *Rv1460* promoter region DNA could not be out-competed, indicating a strong interaction between Rv1460 and the *Rv1460* promoter region (Figure 23 B).

A less gradual increase in the shift of biotin-labelled *Rv1461* binding region DNA was observed with increased Rv1460 concentrations, indicating that Rv1460 binds within the *Rv1461* binding region in a concentration dependent manner (Figure 23 C). A higher concentration of Rv1460 was needed to fully shift the biotin-labelled *Rv1461* binding region DNA which could indicate that the affinity of Rv1460 for the *Rv1461* binding region is lower than for the *Rv1460* promoter region. Kinetic studies could provide additional data on the binding affinity of Rv1460 for these DNA regions since a binding constant can be determined. Interestingly, unlabelled *Rv1461* binding region DNA did not out-compete the biotin-labelled *Rv1461* binding region DNA even at a 200-fold molar excess (Figure 23 D) indicating a strong interaction between Rv1460 (5 μM) and the *Rv1461* binding region.

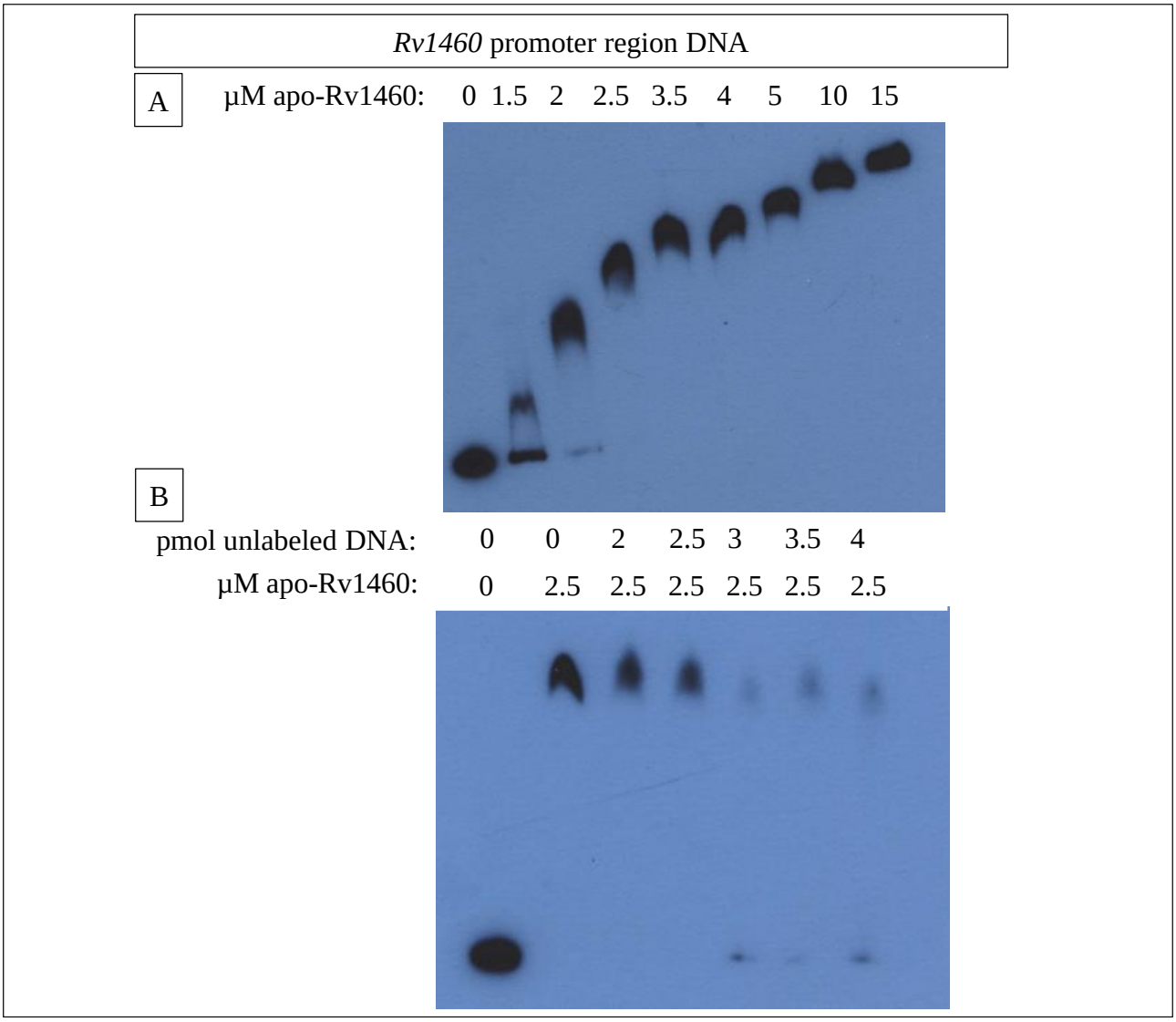


Figure 23. EMSA of "as purified" Rv1460 with the *Rv1460* promoter and *Rv1461* binding regions. All binding reactions contained 20 fmol biotin-labelled *Rv1460* promoter or *Rv1461* binding region DNA prepared by PCR using a 5' biotin labelled forward primer. Increasing concentrations of "as purified" Rv1460 were added to each binding reaction and incubated at room temperature for 30 min. Subsequently increasing concentrations of unlabelled *Rv1460* promoter or *Rv1461* binding region DNA were added as specific competitor. (A) EMSA blot indicating binding of Rv1460 to the *Rv1460* promoter region DNA in a concentration dependent manner and (B) unlabelled *Rv1460* promoter region DNA added as specific competitor competing for binding to Rv1460 (2.5 µM protein). (C) EMSA blot indicating binding of Rv1460 to the *Rv1461* binding region DNA in a concentration dependent manner and (D) unlabelled *Rv1461* DNA-binding region added as specific competitor competing for binding to Rv1460 (5 µM protein).

Continued overleaf

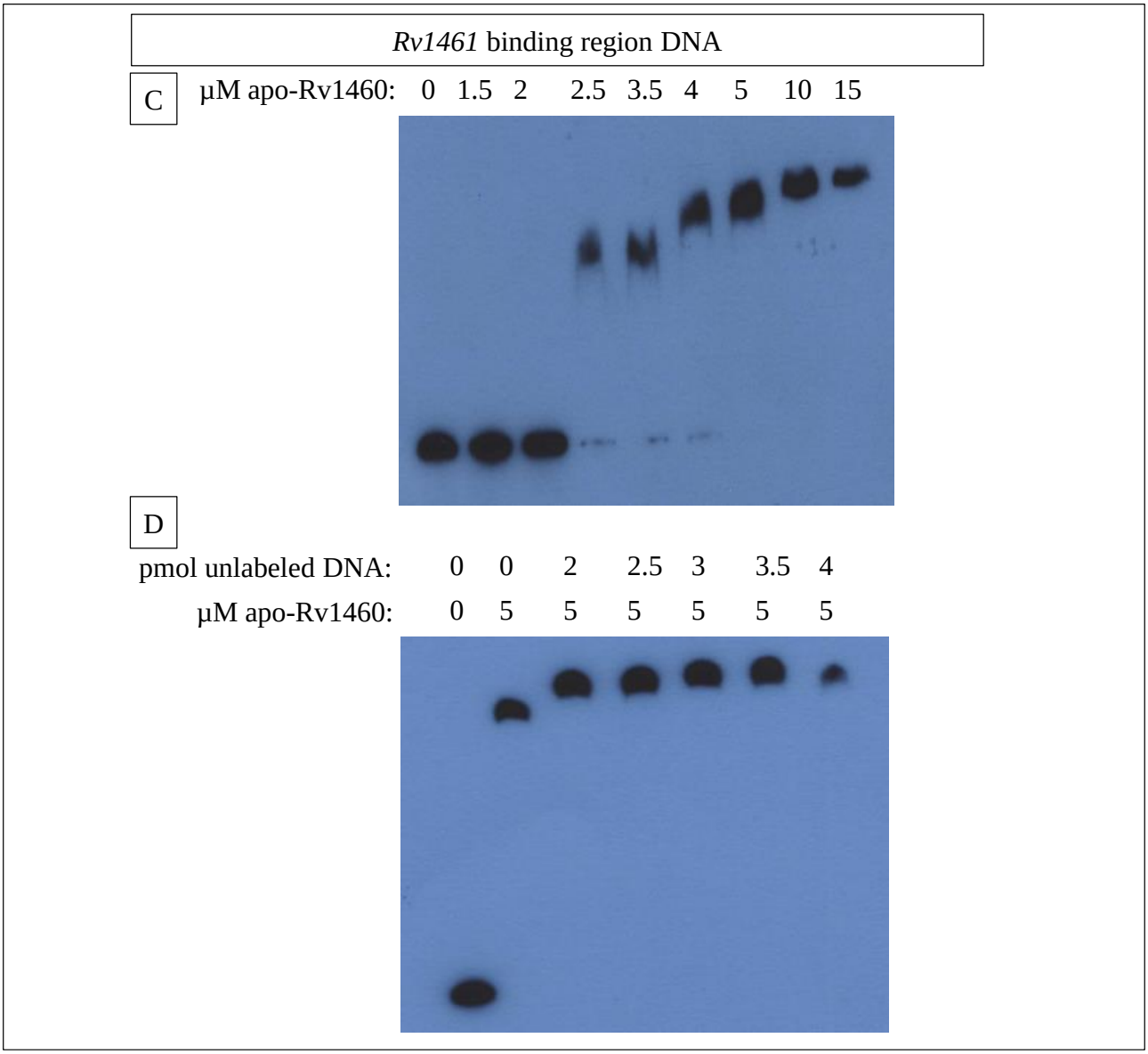


Figure 23. EMSA of "as purified" Rv1460 with the *Rv1460* promoter and *Rv1461* binding region. All binding reactions contained 20 fmol biotin-labelled *Rv1460* promoter or *Rv1461* binding region DNA prepared by PCR using a 5' biotin labelled forward primer. Increasing concentrations of "as purified" Rv1460 were added to each binding reaction and incubated at room temperature for 30 min. Subsequently increasing concentrations of unlabelled *Rv1460* promoter or *Rv1461* binding region DNA were added as specific competitor. (A) EMSA blot indicating binding of Rv1460 to the *Rv1460* promoter region DNA in a concentration dependent manner and (B) unlabelled *Rv1460* promoter region DNA added as specific competitor competing for binding to Rv1460 (2.5 µM protein). (C) EMSA blot indicating binding of Rv1460 to the *Rv1461* binding region DNA in a concentration dependent manner and (D) unlabelled *Rv1461* DNA-binding region added as specific competitor competing for binding to Rv1460 (5 µM protein).

5. Discussion

Fe-S clusters are protein cofactors that are synthesised by complex Fe-S cluster biogenesis systems (Py and Barras, 2010; Roche *et al.*, 2013), which protect the clusters from the deleterious effects of ROS present in the cell during their assembly (Wardman and Candeias, 1996). Fifty Fe-S cluster containing proteins, including proteins involved in transcriptional regulation, metabolism, energy production and conversion, cell division, virulence, drug resistance and stress response, have been identified in *M. tuberculosis*, indicating the importance of Fe-S clusters in *M. tuberculosis* physiology (Saini *et al.*, 2012). The *suf* system is the primary Fe-S cluster biogenesis system in *M. tuberculosis* (Cole *et al.*, 1998; Huet *et al.*, 2005) and is conserved among members of the mycobacterial genus (Garnier *et al.*, 2003; Kapopoulou *et al.*, 2011; Lew *et al.*, 2011). The *suf* system genes are differentially expressed during starvation (Betts *et al.*, 2002) and induced during iron limitation (Rodriguez *et al.*, 2002). The *suf* system is also upregulated upon entry into anaerobic stationary phase (Murugasu-Oei *et al.*, 1999) and by ROS and RNS (Schnappinger *et al.*, 2003; Voskuil *et al.*, 2011), indicating a role in stress response. Upregulation of the *suf* system in macrophages (Rohde *et al.*, 2007; Schnappinger *et al.*, 2003) and pulmonary TB patients (Kumar *et al.*, 2011) indicates the importance of the *suf* system in the intracellular survival of *M. tuberculosis* and during infection. Although the first gene in the *suf* operon, *Rv1460*, encodes a probable transcriptional regulator with homology to a *suf* operon regulator in cyanobacteria, it is not known how this important system is regulated in mycobacteria. The aim of this study was to investigate the role of *Rv1460* in *suf* operon expression and *M. tuberculosis* physiology.

5.1 Rv1460 is required for *M. tuberculosis in vitro* growth

Generation of three distinct *Rv1460* deletion mutants in *M. tuberculosis* was attempted. All three mutants were expected to be generated, as *Rv1460* was predicted to be non-essential for *in vitro* growth of *M. tuberculosis* and *M. marinum* by forward genetic screens (Griffin *et al.*, 2011; Sassetti *et al.*, 2003; Weerdenburg *et al.*, 2015). Six attempts to generate the Δ*Rv1460* and *Rv1460ΔDNAbd* mutants were however unsuccessful, which is not consistent with the ability to generate transposon mutants harbouring insertions in this gene.

During transposon mutagenesis, a *himarI*-based transposon delivered by transducing bacteriophages is inserted relatively randomly at any TA dinucleotide present within the gene. All genes with TA dinucleotides should contain insertions, except those essential for growth (Sassetti *et al.*, 2003).

Although nine TA insertion sites are present in the *Rv1460* gene (Figure 24), transposon insertion only occured at three of these sites. Transposon insertion occured at 23 bp downstream of the Tuberculist annotated start codon (Griffin *et al.*, 2011). *Rv1460* has, however, been re-annotated to start 72 bp downstream of the Tuberculist annotation (Cortes *et al.*, 2013); invalidating this transposition as it is upstream of the gene. Interestingly, mapping of the transcriptional start site confirmed that *Rv1460* is transcribed from the newly annotated start codon producing a leaderless transcript (unpublished results, M. Williams) (Figure 24).

A *Rv1460* transposon mutant harbouring a transposon at 472 bp downstream of the newly annotated start codon was viable when mutants were selected on glycerol, but not cholesterol as a carbon source. Since cholesterol is thought to be used as a carbon source by *M. tuberculosis* within the host (Griffin *et al.*, 2011), this may partially account for the predicted essentiality of *Rv1460* for *in vivo* growth in mice (Sassetti and Rubin, 2003). Mutants harbouring a transposon 64 bp from the 3' end of the gene were also viable (Griffin *et al.*, 2011). It is possible that this transposition is tolerated because it is at the end of the gene and is therefore less likely to affect gene function (Weerdenburg *et al.*, 2015).

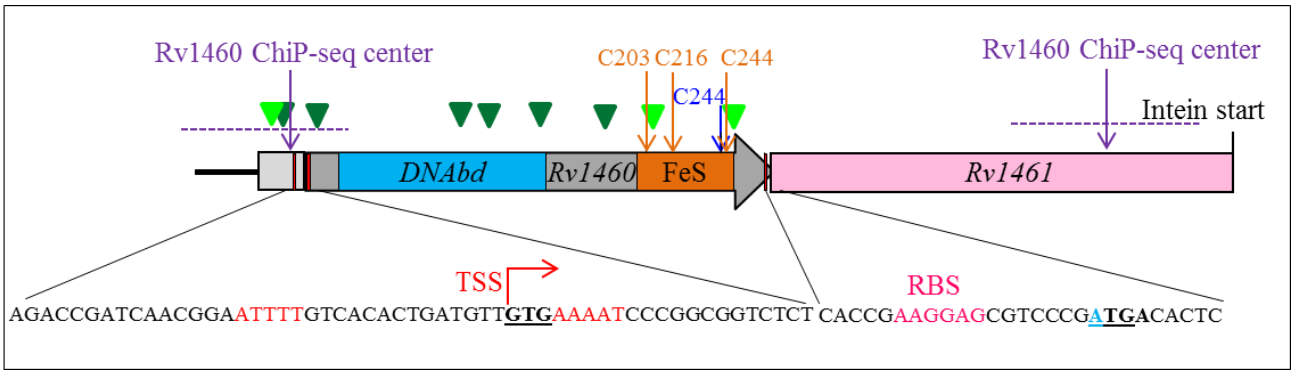


Figure 24. Re-annotated *Rv1460* start site and position of transposon insertion sites and predicted ChIP-seq peak centers for Rv1460 binding. The light grey region represents the region excluded from the Tuberculist and TBDB annotated sequence when Rv1460 was re-annotated to start at +72 bp (Cortes *et al.*, 2013). The *Rv1460* transcriptional start site is indicated by the red arrow and coincides with the newly annotated translational start site (start codon indicated in bold and underlined), creating a leaderless transcript (unpublished results, M. Williams). A palindromic sequence flanking the transcriptional start site is indicated in red. The position of the DNAbd is indicated by the blue box. The predicted RBS is indicated in pink and overlapping start and stop codons of *Rv1460* and *Rv1461*, which could be responsible for translational coupling, are indicated in bold and underlined. The green arrows indicate the nine possible transposon insertion sites, with the light green indicating the three positions where transposition occured (Griffin *et al.*, 2011). Orange arrows indicate the codons of the conserved cysteine residues (C203, C216 and C244) predicted to coordinate the Fe-S cluster and the orange box indicates the Fe-S cluster binding region. The dark blue arrow indicates the codon of the additional cysteine (C242). Purple arrows indicate the position of the peak center for the Rv1460 binding site predicted by ChIP-seq within the *Rv1460* promoter region and within *Rv1461* and purple dashed lines indicate the positions of the regions used for EMSA. The map is drawn to scale.

The high stringency criteria used to predict essential genes from transposon data led to the annotation of some genes previously proven to be essential for *in vitro* growth, as non-essential (Sassetti *et al.*, 2003). *Rv1460* mutants may have been annotated as non-essential despite being attenuated in their growth. The data suggest that the position of the transposon insertion in *Rv1460* has an impact on the resulting phenotype, and may explain the inability to isolate the *ΔRv1460* and *Rv1460ΔDNAbd* mutants.

The *ΔRv1460* mutant would contain a large deletion (528 bp) including the DNAbd and part of the predicted Fe-S cluster binding domain (Figure 5). It is possible that unidentified promoter and or regulatory elements (involved in transcription, translation, post-transcriptional and post-translational regulation) of the downstream or upstream genes are present within this large region. Deletion of the regulatory elements would be detrimental to the survival of *M. tuberculosis*, since *Rv1459c* and the *suf* operon genes are predicted to be essential for *in vitro* growth (Griffin *et al.*, 2011; Huet *et al.*, 2005; Sassetti *et al.*, 2003). The positions of the *Rv1459c* and *Rv1461* promoters have not as yet been identified and no regulator binding sites were predicted to fall within the deleted region (Minch *et al.*, 2015). It is therefore unclear whether this is the reason for the inability to generate these mutants.

The introduction of the pMV*Rv1460compl* vector at the *attB* site of SCOs, prior to DCOs selection, enabled the generation of both *ΔRv1460*:: pMV*Rv1460compl* and *Rv1460ΔDNAbd*:: pMV*Rv1460compl* mutants (Figure 10). Although loss of promoter and regulatory elements may contribute to the inability to generate the *ΔRv1460* and *Rv1460ΔDNAbd* mutants, this argument is contradicted by the ability to generate both mutants when providing the *Rv1460* gene at a separate genomic location, which essentially provides functional Rv1460 protein. Growth of the *ΔRv1460*:: pMV*Rv1460compl* and *Rv1460ΔDNAbd*:: pMV*Rv1460compl* strains was comparable to the wild-type H37Rv progenitor under standard culture conditions (Figure 12 D). It is possible that the lethality of the mutations in the *ΔRv1460* and *Rv1460ΔDNAbd* mutants is due to removal of both the functional protein and promoter/regulatory elements, while removal of either is not lethal. Mutant generation was unsuccessful when introducing the entire *suf* operon at the *attB* site, prior to DCO generation. This suggests that *M. tuberculosis* does not tolerate two copies of the *Rv1461-Rv1462-Rv1463-csd-Rv1465-Rv1466* genes in its genome, presumably due to increased expression levels.

Given that the generation of the Δ*Rv1460* and *Rv1460ΔDNAbd* mutants were unsuccessful, it was surprising that the *Rv1460stop* mutant could be generated (Figure 11), especially as it differs from the *Rv1460ΔDNAbd* mutant by a single base pair. This suggests that *Rv1460* may not truly be essential for the survival of *M. tuberculosis* under the conditions investigated. One possibility is that the growth

of the Δ*Rv1460* and *Rv1460ΔDNAbd* mutants may be more impaired than the *Rv1460stop* mutant. The *Rv1460stop* mutants had a significant growth defect in liquid media under standard culture conditions (Figure 12 A-C) and the *Rv1460stop* 5.20 mutant also had a growth defect on solid media (7H10 OADC) (Figure 13), which was restored to wild-type levels upon genetic complementation with *Rv1460*. *Rv1460* is therefore important for the optimal growth of *M. tuberculosis*. This growth defect is most likely due to the dysregulation of Fe-S cluster biogenesis which affects the function of numerous Fe-S cluster containing proteins involved in diverse cellular processes, making it a complex phenotype. DCO colonies of all sizes, including those with a growth delay, were screened during the mutant generation process to prevent exclusion of mutants that were viable but impaired for growth. It would be interesting to determine whether the Δ*Rv1460* and *Rv1460ΔDNAbd* mutants could be generated as marked deletions, which involves inserting a hyg resistance casette in *Rv1460*. This would prevent DCOs containing the wild-type allele from growing on the selection plates (containing hyg), potentially enabling selection of mutants with a more severe growth defect.

The inability to isolate the Δ*Rv1460* and *Rv1460ΔDNAbd* mutants despite being able to isolate mutants with a truncated non-function Rv1460 (*Rv1460stop* mutant), suggests that the loss of Rv1460 is not the only factor influencing the growth of these mutants. It is possible that in the *ΔRv1460* and *Rv1460ΔDNAbd* mutants, polar effects alter the expression of the downstream genes. This altered expression, together with the loss of Rv1460, may be lethal or causes an even more severe growth phenotype making the selection of these mutants difficult. A single nucleotide difference between the *Rv1460ΔDNAbd* mutant and the *Rv1460stop* mutant may be significant if the translation of *Rv1460* and *Rv1461* is coupled. Translational coupling is a mechanism that makes translation of the distal gene dependent on translation of the gene immediately upstream and adds an additional level of control of gene expression to maintain stoichiometry among gene products (Govantes *et al.*, 1998). The independent translation of the distal gene may in some systems be partially inhibited by the lack of a translation initiation start site or mRNA secondary structure making the translational start sites of the distal gene inaccessible to ribosomes. Translation of the upstream gene melts the mRNA secondary structure providing access to the RBS in the distal gene to the ribosomes. This makes the translation of the distal gene dependent on the translation of the upstream gene (Govantes *et al.*, 1998).

The mycobacterial *suf* genes are in an operon consisting of genes with overlapping start and stop codons and small intergenic regions (Cole *et al.*, 1998; Huet *et al.*, 2005). *Rv1460's* stop codon

overlaps with the *Rv1461* start codon and *Rv1461* is translated in a different reading frame to *Rv1460* (Figure 24). Genes with overlapping start and stop codons are typically translationally coupled (Govantes *et al.*, 1998).

Translational coupling is linked to translational frameshifting, which is an additional regulatory mechanism programmed by *cis* regulatory elements within the mRNA nucleotide sequence and allows the distal gene that is present in a different reading frame to be co-translated with the upstream gene (Engelberg-Kulka and Schoulaker-Schwarz, 1994). This can occur through programmed ribosomal frameshifting, where the ribosome is displaced by one nucleotide to change the initial reading frame, or programmed transcriptional realignment in which a tRNA reads an mRNA in a reading frame shifted by +1 or -1 from the reading frame of the initiation codon thereby allowing consecutive translation of both genes (Antonov *et al.*, 2013; Engelberg-Kulka and Schoulaker-Schwarz, 1994). Translational coupling is especially important in genes that have apposing functions where the product levels of both genes need to be fine-tuned for survival, such as toxin-antitoxin systems. Translational coupling was for instance shown to be important for the expression of the toxin-antitoxin *vapBC* operon in *M. smegmatis*. VapB acts as a transcription factor and antitoxin by repressing VapC activity. Deletion of *vapB* increases transcript levels of *vapC*, but despite the increase in transcription in the absence of the transcriptional repressor, no VapC protein was detected in the *vapB* mutant because translational coupling between the genes caused inhibition of VapC translation (Robson *et al.*, 2009).

In the Δ*Rv1460* and *Rv1460ΔDNAbd* mutants, the ribosome would undergo translational frameshifting between *Rv1460* and *Rv1461* in order to continue translation of the distal *Rv1461* gene. However, in the *Rv1460stop* mutant the addition of an extra nucleotide would cause a premature stop codon in Rv1460. Since the *Rv1460* stop codon and *Rv1461* start codon no longer overlap, their translation would be uncoupled. In a previous study, insertion of a premature stop codon in the upstream gene caused translational uncoupling, and greatly reduced the translation of the downstream gene (Adachi *et al.*, 2012). Since a predicted RBS is present upstream of *Rv1461* (Lew *et al.*, 2011) (Figure 24), its expression may normally be inhibited by mRNA secondary structure. If premature termination of translation occurs, enough of the mRNA may be exposed allowing re-initiation of translation from the Rv1461 RBS (Buckley and Hayashi, 1987). In light of this possible regulation of the *suf* operon's expression on translational level, it would be interesting to investigate the protein levels of the suf operon components in the *Rv1460stop* mutant compared to the H37Rv wild-type progenitor.

A compensatory mutation somewhere else within the genome, could possibly have arrisen during the culturing or mutant generation process, that enabled compensation for the growth deficit in the *Rv1460stop* mutant. The Δ*Rv1460* and *Rv1460ΔDNAbd* mutant generation was attempted six times (each time from a different culture) and was unsuccessful. The *Rv1460stop* mutant generation was however successful during multiple rounds of mutant generation, with the *Rv1460stop* 1.19 mutant arising from a seperate original culture than the 5.19 and 5.20 mutants. It is therefore extremely unlikely that compensatory mutations arising during the mutant generation process enabled the generation of the *Rv1460stop* mutant and not the Δ*Rv1460* and *Rv1460ΔDNAbd* mutants. The reasons for the ability to generate only the *Rv1460stop* mutants and not the *ΔRv1460* and *Rv1460ΔDNAbd* mutants warrants further investigation. The *Rv1460stop* mutant does however provide a valuable tool for investigating the phenotype associated with the loss of *Rv1460* and may provide insight into the role of *Rv1460* in *M. tuberculosis* physiology.

5.2 Loss of *Rv1460* alters *M. tuberculosis* growth *in vitro* and survival in the presence of oxidative stress

Truncation of Rv1460 resulted in a growth defect on both liquid and solid media under standard growth conditions that was restored by genetic complementation with *Rv1460* (Figure 12 and 13). The varying levels of growth defects observed for the three *Rv1460stop* mutants in liquid culture is possibly due to compensatory mutations that arrose in the *Rv1460stop* 5.20 mutant during the mutant generation process. This is currently being investigated by whole genome sequencing. Truncation of Rv1460 was also associated with increased sensitivity to low (Figure 14 A) and high levels of oxidative stress (Figure 14 B). The survival in the presence of ROS was completely restored to wild-type levels by genetic complementation with *Rv1460*, suggesting that the increased sensitivity was due to truncation of Rv1460. If unique single nucleotide polymorphisms (SNPs) are identified within the three strains it may be worth evaluating the survival of the other *Rv1460stop* mutants (1.19 and 5.19) in the presence of oxidative stress, to evaluate their contribution to the phenotypes observed.

Fe-S clusters are sensitive to oxidation and Fe-S biogenesis systems are therefore required to protect the Fe-S clusters from oxidation during *de novo* Fe-S cluster synthesis and insertion, and also to repair damaged clusters. Upon inhalation, *M. tuberculosis* is phagocytosed by alveolar macrophages where it encounters numerous stresses including ROS and RNS (Schnappinger *et al.*, 2003). *Suf* operon expression is induced in macrophages after sustained infection indicating a role for the *suf* operon in sustained intracellular survival (Rohde *et al.*, 2007; Schnappinger *et al.*, 2003). The *suf* operon is also

upregulated by ROS (Schnappinger *et al.*, 2003; Voskuil *et al.*, 2011) and RNS (Voskuil *et al.*, 2011). A binding site for Rv1460 within the *lsr2* promoter is predicted, suggesting regulation of *lsr2* by Rv1460 (Minch *et al.*, 2015). Lsr2 binds to and physically shields *M. tuberculosis* DNA to protect it from ROS (Colangeli *et al.*, 2009). ROS and RNS furthermore inhibit intein splicing required to generate active Rv1461 and providing a post-translational regulatory mechanism to respond to these stress conditions (Topilina *et al.*, 2015). The *suf* operon genes are also predicted to be regulated by other regulators that are linked to intracellular survival (Table 1). Loss of Rv1460 was therefore expected to alter the *Rv1460stop* mutant's survival upon ROS exposure, due to dysregulation of Fe-S cluster biogenesis in response to oxidative stress.

5.3 Rv1460 is not required for growth under conditions of iron limitation

Iron is critical for the survival and pathogenesis of *M. tuberculosis* and is limited intracellularly (Banerjee *et al.*, 2011). The *suf* system genes are accordingly induced under iron limitation in *M. tuberculosis* (Rodriguez *et al.*, 2002) and other bacteria (Desnoyers *et al.*, 2009; Massé and Gottesman, 2002) and regulated by Fur and IscR in response to the presence of their Fe-S clusters in other bacteria (Lee *et al.*, 2008; Mettert and Kiley, 2014; Runyen-Janecky *et al.*, 2008), indicating a link between the *suf* operon and iron responsiveness. The effect of iron limitation on the growth of the *Rv1460stop* mutant was therefore investigated. Interestingly, iron limitation did not affect the growth of the *Rv1460stop* mutant or the complementation strain compared to the H37Rv wild-type progenitor (Figure 15) suggesting that the *Rv1460stop* mutant may be able to scavenge iron just as effectively as the wild-type strain and Fe-S cluster biogenesis may not be affected by the limited iron under the conditions investigated. It is also possible that the iron levels were not sufficiently depleted under the conditions investigated. The experimental conditions would therefore need to be optimised by evaluating the iron status of the cells in order to determine at which point iron becomes limiting.

5.4. Rv1460 is a transcriptional repressor of the *suf* operon

All *suf* operon genes could be amplified from cDNA synthesised by RT primers binding within downstream genes, suggesting that the genes in the *suf* operon are co-transcribed (Huet *et al.*, 2005; Roback *et al.*, 2007). The *M. smegmatis Rv1460* homologue can by contrast be transcribed independently of the rest of the operon since only SCOs with a wild-type copy of *Rv1460* and intact

(uninterupted) *Rv1461-Rv1466* genes could be isolated during attempts to generate marked deletions of the *M. smegmatis Rv1461* homologue (Huet *et al.*, 2005).

The role of *Rv1460* in *suf* operon expression was investigated using the *Rv1460stop* 5.20 mutant. Transription of *Rv1460* and *Rv1461* was increased in the *Rv1460stop* 5.20 mutant compared to the wild-type H37Rv progenitor (Figure 16) indicating that Rv1460 functions as a transcriptional repressor of its own expression as well as the rest of the *suf* operon. This agrees with decreased *Rv1461-Rv1466* expression when *Rv1460* was overexpressed (Rustad *et al.*, 2014). The cyanobacterial Rv1460 homologue, SufR, similarly represses its own expression and that of the *sufBCDS* operon (Wang *et al.*, 2004). *Rv1460* and *Rv1461* transcript levels returned to wild-type levels following genetic complementation with *Rv1460*, indicating that polar effects on the expression of downstream genes were not responsible for the varying transcript levels (Figure 16).

Amplification of *Rv1460* from the Rv1461RT primer product confirmed co-transcription of *Rv1460* with the rest of the operon (Figure 16). *Rv1460* expression was however higher using the Rv1460RT primer compared to the Rv1461RT primer in all the strains (Figure 16) indicating that all three transcripts, containing *Rv1460-Rv1466*, *Rv1461-Rv1466* and only *Rv1460* were made in all the strains. Alternatively RT efficiency may be higher when using the Rv1460RT than the Rv1461RT primer leading to increased cDNA synthesis. Stronger PCR amplification of all the *suf* operon genes was previously observed for cDNA generated from a Rv1466RT primer (Huet *et al.*, 2005). Expression from multiple promoters has been demonstrated for other genes in *M. tuberculosis.* The *gyrBA* genes in *M. tuberculosis* are for instance primarily co-transcribed from a promoter upstream of the *gyrB* gene. Other promoters upstream of *gyrA* were identified and predicted to regulate *gyr* gene expression, presumably by enabling the binding of different sigma factors (Unniraman *et al.*, 2002). Different promoters also allow responses to diverse environmental conditions. For example, diamide stress switches *suf* operon promoter usage in *C. glutamicum*, resulting in two different *suf* mRNA transcripts from standard culture and diamide exposed conditions (Nakunst *et al.*, 2007).

5.5. Production and purification of recombinant Rv1460 protein in *E. coli*

To investigate the properties of Rv1460 that contribute to its role as *suf* operon regulator and in *M. tuberculosis* physiology, Rv1460 was produced and purified as a recombinant His-tagged protein in *E. coli*. Production of soluble recombinant Rv1460 in *E. coli* increased with decreasing temperature and extent of induction (Figure 17). Although protein production was better in the Arctic express than Rosetta-gami 2 strains, recombinant Rv1460 was still predominantly insoluble (Figure 17 G-J). Poor

solubility was not unexpected as solubility prediction software (University of Oklahoma School of Chemical, Biological, and Materials Engineering, Recombinant Protein Solubility Prediction - http://www.biotech.ou.edu/) predicted a 0% chance of solubility for the recombinant Rv1460 protein based on numerous factors including amino acid composition, hydrophilicity index and protein charge (Diaz *et al.*, 2010). Interestingly, SufR, with 42% amino acid identity to Rv1460 (Huet *et al.*, 2005; Wang *et al.*, 2004), was also purified from inclusion bodies (Wang *et al.*, 2004).

Solubility correlates with proper protein folding, and refolding of proteins isolated from the insoluble fraction is often difficult (Francis and Page, 2010). However even soluble produced protein may occasionally be incorrectly folded (González-Montalbán *et al.*, 2007). Increased aeration using baffled flasks during culturing, increased the proportion of soluble protein (results not shown). Protein precipitation reduced the overall yield during large-scale Rv1460 purification possibly due to aggregation caused by disulphide crosslinking between Rv1460 as reducing conditions stabilised the protein increasing its solubility (Francis and Page, 2010). Correspondingly, higher molecular weight species were observed during protein purification (Figure 18).

The inclusion of a 24 aa N-terminal extension may have contributed to poor solublity as it does not form part of the Rv1460 protein according to the new start site annotation (Cortes *et al.*, 2013). This region contains five threonine residues and two arginine residues that may lead to poor protein solubility (Price *et al.*, 2011; Trevino *et al.*, 2007) although exclusion of this 24 amino acid region from the protein sequence did not improve the solubility prediction (Diaz *et al.*, 2010). Preliminary results suggest that the solubility of an N-terminally truncated version of Rv1460 (starting at codon 36) is not improved compared to the full-length protein used in this study (unpublished results, M. Williams).

Large-scale soluble recombinant Rv1460 was successfully purified by nickel affinity and gel filtration chromatography under aerobic conditions. The major peak on the gel filtration elution profile corresponded to the dimeric form of Rv1460 (Figure 18), consistent with the formation of SufR dimers in cyanobacteria (Shen *et al.*, 2007). Interestingly, apo-SufR purified by gel filtration chromatography was entirely dimeric, while holo-SufR formed a small proportion of tetramers (Huet *et al.*, 2005). Rv1460 and Rv1460 C203,216,244S protein variant were expected to be in the apo-Rv1460 form since they were purified under aerobic conditions. A more pronounced “tetramer” shoulder was present in the C203,216,244S variant protein compared to Rv1460 protein, and may indicate a structural change in the variant protein (Figure 18). SDS-PAGE gels were denaturing and the gels and western blots thus indicated only dimer and monomer even though higher order structures

may be formed by Rv1460. Far-UV CD indicated that Rv1460 contained secondary structure (Figure 19). Since a crystal structure for Rv1460 is not available, the degree to which the protein is properly folded cannot be determined. Determining the structure of Rv1460 by nuclear magnetic resonance and X-ray crystallography will provide valuable information about the manner in which it functions (Kelly *et al.*, 2005).

5.6 Rv1460 binds to its own promoter and within *Rv1461*

Rv1460 was previously predicted to bind within its own promoter and within *Rv1461* by ChIP-seq, and therefore EMSA was performed to confirm that regulation of the *suf* operon occurs through direct binding to the DNA (Minch *et al.*, 2015). Binding of "as purified" Rv1460 within its own promoter was shown by EMSA (Figure 23). In addition, binding of Rv1460 to the *Rv1461* binding region within *Rv1461* was also confirmed (Figure 23). The role of a Rv1460 binding site within *Rv1461* is unknown since it is located downstream of the *Rv1461* start site and RBS. The structure of the promoter(s) of *Rv1461* needs to be investigated as it could help determine whether this binding site has a role in *Rv1461* expression.

Unexpectedly, the affinity of "as purified" Rv1460 for the *Rv1460* promoter region appeared higher than for the *Rv1461* binding region as higher concentrations of "as purified" Rv1460 were needed to fully shift the *Rv1461* binding region DNA. Since the electrophoretic resolution for EMSA is better when using smaller DNA fragments, the larger fragment used for *Rv1461* binding region (325 bp versus 348 bp) may have impacted this result (Hellman and Fried, 2007). EMSA is only semi-quantitative and can therefore not be used to accurately assess the affinity of Rv1460 for the DNA. The binding constant of "as purified" Rv1460 for the *Rv1460* promoter and *Rv1461* binding regions needs to be determined. This can be done using MicroScale Thermophoresis, which works by measuring the motion of molecules along microscopic temperature gradients and detecting changes in their size, hydration shell and charge to provide information on their interaction. This information is complementary to that of EMSA.

The intergenic region between the divergently transcribed *sufR* and *sufBCDS* genes in cyanobacteria contains two identical *SufR* binding sites with inverted repeats. One motif, that controls *sufBCDS* expression, is flanked by an extended palindromic flanking sequence (Shen *et al.*, 2007; Wang *et al.*, 2004) which changes the affinity of holo-SufR for the binding site. (Shen *et al.*, 2007; Wang *et al.*, 2004). The increased affinity for the site upstream of *sufBCDS* allows SufR to differentially regulate this operon relative to its own expression (Shen *et al.*, 2007; Wang *et al.*, 2004). Apo-SufR and

reduced holo-SufR bind to the shorter inverted repeat site with low affinity while its binding to the extended palindromic flanking sequence was not investigated (Shen *et al.*, 2007). Nevertheless, the affinity of apo-SufR for these binding sites possibly also differs, emphasising the importance of the binding sequences and the SufR Fe-S cluster and cluster redox state in the regulation of the *suf* operon in cyanobacteria.

Since large DNA fragments (325 bp and 348 bp for the *Rv1460* promoter region and *Rv1461* binding region respectively) were used for EMSA binding assays, it is important to narrow down the exact Rv1460 binding sites by EMSA competition assays using smaller DNA fragments spanning the ChIP-seq peak centers (Minch *et al.*, 2015) and DNase I footprinting. Comparing the DNA sequence of the two binding sites would provide insight into the molecular interactions between the protein and DNA and the DNA-specific factors that influence the strength of this interaction.

Once the exact DNA-binding motif of Rv1460 has been determined, it would be interesting to model the structure of Rv1460 and its binding to this motif. Although there is currently no published crystal structure for Rv1460 or the SufR homologue, the winged helix-turn-helix DNAbd domain of the closest structural homologue available can be used for the modelling. This will provide additional information about the affinity of Rv1460 for the DNA-binding motif and may explain the apparent difference in binding affinity for the *Rv1460* promoter region and *Rv1461* region. Additionally, amino acids important for the interaction of Rv1460 with its DNA-binding motif and their role in this interaction can be predicted by the modelled structure by investigating alterations in binding affinity when these amino acids are replaced. The genes previously predicted to have Rv1460 binding sites by ChiP-seq (Minch *et al.*, 2015) can also be searched for the identified DNA-binding motif and the binding of Rv1460 to these regions investigated by EMSA.

5.7. Rv1460 coordinates an Fe-S cluster

Although Rv1460 contains conserved cysteine residues indicative of an Fe-S cluster binding domain (Cole *et al.*, 1998; Shen *et al.*, 2007), its ability to coordinate an Fe-S cluster and the role of these cysteines in cluster coordination have not been investigated previously. To investigate this, the *M. tuberculosis* cysteine desulphurase, $IscS_{Mtb}$, was produced and purified for use in Fe-S cluster reconstitution experiments. $IscS_{Mtb}$ was produced in Rosetta 2 strains both solubly and insolubly (Figure 20 C-D). The Ni-NTA elution profile for $IscS_{Mtb}$ revealed two peaks, both containing recombinant $IscS_{Mtb}$ (Figure 21 A). Two peaks were previously observed on the gel filtration elution profile when purifying the *E. coli* $IscS_{E.coli}$ with the first peak corresponded to the PLP-bound, active

enzyme and the second peak represents the PLP-free, unfolded, inactive form (Prischi *et al.*, 2010). The two peaks observed in the Ni-NTA elution profile in this study could similarly represent the PLP-bound and PLP-free forms of $IscS_{Mtb}$ suggesting that the cofactor influences the accessibility of the His-tag and affects binding to the affinity resin. When these two peaks were combined and analysed on a S200 gel filtration column, only a single peak resulted, suggesting that they share a similar size. The second peak was much smaller than the first and may have been masked during gel filtration. The size of the gel filtration peak was interpreted to correspond to dimeric $IscS_{Mtb}$ (Figure 21 D), in agreement with previous studies (Rybniker *et al.*, 2014).

Fe-S cluster reconstitution on Rv1460 using $IscS_{Mtb}$ was unsuccessful (Figure 22 A), despite proven cysteine desulphurase activity of $IscS_{Mtb}$. $IscS_{Mtb}$ was previously used to reconstitute the 4Fe-4S cluster on WhiB3 (Singh *et al.*, 2009, 2007). $IscS_{Mtb}$ may only be able to reconstitute the Fe-S cluster on some proteins and/or Rv1460 receives its Fe-S cluster primarily from the *suf* system. $IscS_{Mtb}$ was predicted to transfer Fe-S clusters directly to apo-proteins without the need for a scaffold protein (Rybniker *et al.*, 2014). Possibly though $IscS_{Mtb}$ does need a scaffold protein for Fe-S cluster reconstitution on some proteins, such as Rv1460. $IscU_{E.coli}$ was unable to function as a scaffold for $IscS_{Mtb}$ in the reconstitution reaction (data not shown), consistent with the observation that these proteins are unable to form a stable complex. *Rv1465* encodes a NifU/IscU-like scaffold protein, however, Yeast-two-hybrid studies did not reveal any interaction between $IscS_{Mtb}$ and Rv1465 (Rybniker *et al.*, 2014), suggesting that it does not function as a scaffold for reconstitution by $IscS_{Mtb}$.

Using an established Fe-S cluster reconstitution system that utilises purified cysteine desulphurase ($IscS_{E.coli}$) and scaffold protein ($IscU_{E.coli}$) from *E. coli*, the Fe-S cluster on Rv1460 was successfully reconstituted *in vitro* (Figure 22 B-C), demonstrating that $IscU_{E.coli}$ is able to transfer Fe-S clusters to apo-proteins in a non-specific manner. Although $IscS_{Mtb}$ and $IscS_{E.coli}$ share an overall fold and active site, their surface amino acids are distinct and they interact differently with mycobacterial Fe-S cluster proteins (Rybniker *et al.*, 2014). Complementation of *M. tuberculosis* $iscS_{Mtb}$ with *E. coli* $iscS_{E.coli}$ did not restore the growth defect observed for the *M. tuberculosis iscS* deletion mutant indicating that $IscS_{E.coli}$ requires its scaffold protein, $IscU_{E.coli}$, for reconstitution (Rybniker *et al.*, 2014).

The UV-visible absorbance spectrum observed during *in vitro* reconstitution indicates the formation of a 2Fe-2S cluster that transitions to a 4Fe-4S cluster. The near-UV CD spectrum of reconstituted Rv1460 similarly indicatives a 2Fe-2S cluster (Adinolfi *et al.*, 2009; Bonomi *et al.*, 2005). *In vitro* reconstitution of the cyanobacterial SufR revealed that each dimer coordinates two oxygen sensitive 4Fe-4S clusters (Shen *et al.*, 2007; Wang *et al.*, 2004). While these *in vitro* systems confirm that both

regulators coordinate Fe-S clusters, they cannot identify the type of cluster bound *in vivo*, due to excess iron and sulphur in the reconstitution reactions. The type of cluster bound *in vivo* can only be determined by purifying *in vivo* reconstituted proteins under anaerobic conditions.

Aligning the amino acid sequence of Rv1460 with SufR homologues from other organisms revealed three conserved cysteine residues, C203, C216, C244, predicted to coordinate the Fe-S cluster. Rv1460 protein variants in which the conserved cysteine residues at positions 203, 216 and 244S were replaced with serine residues were still able to coordinate the Fe-S cluster (Figure 22 C-E). Serine residues were chosen as they differ from cysteine by only one atom (-OH versus -SH group), and presumably cause less structural variation than changing to alanine (-H side chain). Serine residues in some cases are able to coordinate Fe-S clusters, while altering the chemical properties of the cluster (Evans and Leigh, 1991). Replacing the coordinating cysteine residues with serine residues did not significantly affect the coordination and geometry of the Fe-S cluster in ferredoxin NADP reductase (FNR) from *Anabaena* PCC 7120 and High potential Fe-S protein from *Chromatium vinosum*, but some cysteine to serine residue mutations did change the cluster redox state (Babini *et al.*, 1996; Hurley *et al.*, 1997).

The importance of the Fe-S cluster coordinating residues was underscored by none of the Rv1460 C203S, C216S or C244S protein variants complementing the *M. tuberculosis ΔRv1460* mutant (unpublished results, M. Williams). These variants presumably alter the function of the Rv1460 protein in *M. tuberculosis*, thereby affecting its survival negatively. Interestingly, Rv1460 C242S did complement the mutant, indicating that it is not critical to Rv1460 function. Since C242 is predicted not to be involved in Fe-S cluster coordination (unpublished results, M. Williams), the altered function of C203S, C216S and C244S mutants is presumably due to altered Fe-S cluster coordination or redox state of the Fe-S cluster (Babini *et al.*, 1996; Hurley *et al.*, 1997). The role of these cysteine residues (C203, C216, C244) in cluster coordination should be confirmed by replacing them with alanine and investigating the effect on Fe-S cluster coordination.

Since Rv1460 coordinates an Fe-S cluster, it is probable that holo-Rv1460 binds with a different affinity to the Rv1460 binding sites in the *Rv1460* promoter and within *Rv1461*, and that this differential binding may affect *suf* operon expression (as established for SufR and its binding sites in cyanobacteria). The influence of the Fe-S cluster on the affinity of Rv1460 for DNA will be investigated during future studies to understand the mechanism by which Rv1460 regulates Fe-S cluster biogenesis in *M. tuberculosis.*

This study provides evidence that Rv1460 is important in regulating the *suf* operon and *M. tuberculosis* physiology. Other global regulators have also been predicted to bind and regulate the *suf* operon (Guo *et al.*, 2009; Minch *et al.*, 2015). In some bacteria, regulator binding sites overlap in the *suf* operon promoter, with one regulator acting as anti-activator by blocking the binding of the activator. Dissociation of the anti-activator under stress conditions allows binding of the activator, ensuring efficient expression of the *suf* operon (Lee *et al.*, 2008; Mettert and Kiley, 2014; Outten *et al.*, 2004). Yet another mechanism of regulation of *suf* operon expression thus exists allowing responses to numerous environmental conditions. Potential *suf* operon regulators in *M. tuberculosis* include global regulators involved in metabolism, cell wall biosynthesis, metal cofactor biosynthesis, intracellular survival and virulence (Table 1). The importance of the *suf* system to such diverse processes mirrors the importance of Fe-S clusters in the function of proteins involved in diverse processes (Saini *et al.*, 2012) and indicates the interconnectedness of the *suf* system with the metabolism and survival of *M. tuberculosis*.

6. Conclusions

Fe-S cluster biogenesis is an understudied area of *M. tuberculosis* metabolism. The *suf* system is the primary Fe-S cluster biogenesis system in mycobacteria and its components are encoded in an operon consisting of seven genes (*Rv1460-Rv1461-Rv1462-Rv1463-csd-Rv1465-Rv1466*) (Cole *et al.*, 1998; Huet *et al.*, 2005). The first gene in the operon, *Rv1460*, is predicted to encode a transcriptional regulator due to homology to the cyanobacterial *suf* operon repressor, SufR (Shen *et al.*, 2007; Wang *et al.*, 2004). This study aimed to investigate the role of *Rv1460* in the regulation of *suf* operon expression and in *M. tuberculosis* physiology.

Attempts were made to generate three distinct *Rv1460* deletion mutants in *M. tuberculosis* H37Rv. The *ΔRv1460* and *Rv1460ΔDNAbd* mutants could not be generated while the *Rv1460stop* mutant, in which Rv1460 is truncated by insertion of a premature stop codon, was successful. This suggests that Rv1460 is not essential for *in vitro* growth of *M. tuberculosis*, but that a combination of the loss of Rv1460 and polar effects on the expression of downstream genes could cause the *ΔRv1460* and *Rv1460ΔDNAbd* mutants to be non-viable. Presumably these polar effects are not present in the *Rv1460stop* mutant as they are viable even in the absence of Rv1460. The *Rv1460stop* mutant provided a valuable tool to investigate whether Rv1460 is involved in *suf* operon regulation and *M. tuberculosis* physiology. The *Rv1460stop* mutants had a growth defect under standard culture conditions, and was more sensitive to ROS, indicating the importance of *Rv1460* in oxidative stress response, and implicating it in intracellular survival and pathogenesis. Interestingly, growth of the *Rv1460stop* mutant was not affected by iron limiting conditions.

Transcript levels of both *Rv1460* and *Rv1461* increased in the *Rv1460stop* mutant, indicating that Rv1460 transcriptionally represses its own gene and the rest of the *suf* operon. Rv1460 was co-transcribed with the rest of the operon. However, transcript levels also suggested that *Rv1460* may be transcribed independently from the rest of the gene cluster. Multiple promoters in this region may provide an added level of *suf* operon regulation in *M. tuberculosis*.

The observed binding of Rv1460 to the *Rv1460* promoter region and within *Rv1461* indicates that it transcriptionally controls through direct interaction with the *suf* operon DNA. The affinity of Rv1460 appeared to be higher for the *Rv1460* promoter than the *Rv1461* binding site. Although the exact binding motif/s involved in Rv1460 binding remains undetermined, the presence of two binding sites

and apparent different affinities for the binding sites could have important implications for the *suf* operon's regulation, and warrants further investigation.

The ability of Rv1460 to coordinate an Fe-S cluster *in vitro* was confirmed. The role of three conserved cysteine residues (C203, C216, C244), predicted to be involved in Fe-S cluster coordination in Rv1460, could not be confirmed, as mutating these residues to serine residues did not alter their ability to coordinate an Fe-S cluster. The coordination of an Fe-S cluster by Rv1460 has potentially important implications for its role as a regulator, as Fe-S cluster presence and redox state has been shown to alter binding specificity and affinity of regulators for DNA. This would in turn facilitate regulation of the Fe-S cluster biogenesis system by Rv1460 in response to the Fe-S clusters demand within the cell.

In addition to regulation of the *suf* operon by Rv1460, other global regulators have also been predicted to bind and regulate the *suf* operon (Guo *et al.*, 2009; Minch *et al.*, 2015). Overlapping regulator binding sites and competition between regulators for these binding sites may provide yet another mechanism for fine-tuning the *suf* operon expression. Regulation of the *suf* operon in *M. tuberculosis* is multi-faceted, presumably because Fe-S cluster biogenesis must be fine-tuned to allow survival within the host. This study indicates that Rv1460 has a key role in the regulation of the *suf* operon and in *M. tuberculosis* physiology.

7. Future studies

To our knowledge, this is the first study investigating the regulation of Fe-S cluster biogenesis in *M. tuberculosis* and provides a role for Rv1460 in this process. Additional questions about the regulation of Fe-S cluster biogenesis in *M. tuberculosis* are however raised in this study. The ability to generate the *Rv1460stop* mutant while numerous attempts to generate the *ΔRv1460* and *Rv1460ΔDNAbd* mutants were unsuccessful, warrants further investigation. It would be interesting to generate marked deletions of *Rv1460* instead of unmarked deletions and determine whether selection of these mutants, that could have a more severe phenotype, is possible. The possibility that polar effects on the expression of downstream genes and translational coupling between *Rv1460* and *Rv1461* could be involved in *suf* operon regulation, should be explored.

The variation in the growth defect observed for the three *Rv1460stop* mutants under standard culture conditions is currently being investigated by whole genome sequencing to determine whether compensatory mutations occurred and in which genes. This could provide information about genes encoding proteins potentially involved in Fe-S cluster biogenesis in *M. tuberculosis*. It would be interesting to further explore the role of Rv1460 in oxidative stress response by investigating the effect of additional ROS and RNS generators on the survival of the *Rv1460stop* mutant, since the *Rv1460stop* mutant was more sensitive to CHP. The survival of the *Rv1460stop* mutants within macrophages should also be investigated and will be interesting because *Rv1460* was predicted to be essential for *in vivo* growth in mice. Since varying transcript levels could indicate a promoter element upstream of *Rv1461*, the presence of a promoter needs to be determined.

EMSA assays indicated binding within the promoter regions upstream of *Rv1460* and within *Rv1461*. It would be interesting to narrow down the exact binding positions using EMSA competition experiments and DNaseI footprinting, since the DNA regions used in the EMSA experiments were large. Determination of the exact binding site will provide information about the DNA sequence(s) recognized by Rv1460 and possibly about its mechanism of regulation. The binding constant of Rv1460 for these regions also needs to be determined as it will provide information about the affinity of Rv1460 for these sites. A high resolution structure of Rv1460 should also be determined by nuclear magnetic resonance and X-ray crystallography and will provide valuable information about the residues involved in DNA-binding and Rv1460's function as a regulator. It would be interesting to use the structure of the winged helix-turn-helix DNAbd domain of the closest Rv1460 homologue to

model its binding to the identified Rv1460 DNA-binding motif and identify amino acids important for this interaction.

The role of the cysteine residues, predicted to be involved in Fe-S cluster biogenesis, should also be investigated further by replacing the cysteine residues with alanine instead of serine residues (since alanine residues do not have an intrinsic ability to coordinate Fe-S clusters) and doing Fe-S cluster reconstitution experiments. Binding of the clusterless Rv1460 and holo-Rv1460 to the *Rv1460* promoter and *Rv1461* binding regions needs to be investigated by EMSA, to determine whether the cluster presence affects the DNA-binding ability of Rv1460.

Although numerous regulators have been predicted to bind within the *suf* operon, no role of these regulators in *suf* operon regulation has been established. Binding of these regulators to the *suf* operon should be confirmed by EMSA assays and DNaseI footprinting used to determine their exact binding sites. The role of each regulator in *suf* operon regulation should also be determined. Further investigation of the mechanism of regulation of the *suf* operon by Rv1460 and the role of Rv1460 in *M. tuberculosis* physiology will allow a better understanding of Fe-S cluster biogenesis in *M. tuberculosis*. Since this is an understudied area of *M. tuberculosis* metabolism, it will provide a better understanding of this deadly pathogen.

8. References

1. Adachi, Y., Kuroda, H., Yukawa, Y., Sugiura, M., 2012. Translation of partially overlapping *psbD-psbC* mRNAs in chloroplasts: The role of 5′-processing and translational coupling. *Nucleic Acids Res.* 40, 3152–3158.
2. Adinolfi, S., Iannuzzi, C., Prischi, F., Pastore, C., Iametti, S., Martin, S.R., Bonomi, F., Pastore, A., 2009. Bacterial frataxin CyaY is the gatekeeper of iron-sulfur cluster formation catalyzed by IscS. *Nat. Struct. Mol. Biol.* 16, 390–396.
3. Agar, J.N., Krebs, C., Frazzon, J., Huynh, B.H., Dean, D.R., Johnson, M.K., 2000. IscU as a scaffold for iron−sulfur cluster biosynthesis: sequential assembly of [2Fe-2S] and [4Fe-4S] clusters in IscU. *Biochemistry* 39, 7856–7862.
4. Alam, S., Agrawal, P., 2008. Matrix-assisted refolding and redox properties of WhiB3/Rv3416 of *Mycobacterium tuberculosis* H37Rv. *Protein Expr. Purif.* 61, 83–91.
5. Albrecht, A.G., Netz, D.J.A., Miethke, M., Pierik, A.J., Burghaus, O., Peuckert, F., Lill, R., Marahiel, M.A., 2010. SufU is an essential iron-sulfur cluster scaffold protein in *Bacillus subtilis. J. Bacteriol.* 192, 1643–1651.
6. Aly, S., Wagner, K., Keller, C., Malm, S., Malzan, A., Brandau, S., Bange, F.C., Ehlers, S., 2006. Oxygen status of lung granulomas in *Mycobacterium tuberculosis*-infected mice. *J. Pathol.* 210, 298–305.
7. Antonov, I., Coakley, A., Atkins, J.F., Baranov, P.V., Borodovsky, M., 2013. Identification of the nature of reading frame transitions observed in prokaryotic genomes. *Nucleic Acids Res.* 41, 6514–6530.
8. Babini, E., Bertini, I., Borsari, M., Capozzi, F., Dikiy, A., Eltis, L.D., Luchinat, C., 1996. A serine → cysteine ligand mutation in the high potential iron−sulfur protein from *Chromatium vinosum* provides insight into the electronic structure of the [4Fe−4S] cluster. *J. Am. Chem. Soc.* 118, 75–80.
9. Banaiee, N., Jacobs, W.R., Ernst, J.D., 2006. Regulation of *Mycobacterium tuberculosis whiB3* in the mouse lung and macrophages. *Infect. Immun.* 74, 6449–6457.
10. Banerjee, S., Farhana, A., Ehtesham, N.Z., Hasnain, S.E., 2011. Iron acquisition, assimilation and regulation in mycobacteria. *Infect. Genet. Evol.* 11, 825–838.
11. Beinert, H., 2000. Iron-sulfur proteins: Ancient structures, still full of surprises. *J. Biol. Inorg. Chem.* 5, 2–15.
12. Beinert, H., Holm, R.H., Münck, E., 1997. Iron-sulfur clusters: Nature's modular, multipurpose structures. *Science*. 277, 653-659.

13. Betts, J.C., Lukey, P.T., Robb, L.C., McAdam, R.A., Duncan, K., 2002. Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling. *Mol. Microbiol.* 43, 717–731.
14. Bonomi, F., Iametti, S., Ta, D., Vickery, L.E., 2005. Multiple turnover transfer of [2Fe2S] clusters by the iron-sulfur cluster assembly scaffold proteins IscU and IscA. *J. Biol. Chem.* 280, 29513–29518.
15. Brosch, R., Gordon, S.V., Billault, A., Garnier, T., Eiglmeier, K., Soravito, C., Barrell, B.G., Cole, S.T., 1998. Use of a *Mycobacterium tuberculosis* H37Rv bacterial artificial chromosome library for genome mapping, sequencing, and comparative genomics. *Infect. Immun.* 66, 2221–2229.
16. Buckley, K.J., Hayashi, M., 1987. Role of premature translational termination in the regulation of expression of the *φX174* lysis gene. *J. Mol. Biol.* 198, 599–607.
17. Burian, J., Ramón-García, S., Sweet, G., Gómez-Velasco, A., Av-Gay, Y., Thompson, C.J., 2012. The mycobacterial transcriptional regulator *whiB7* gene links redox homeostasis and intrinsic antibiotic resistance. *J. Biol. Chem.* 287, 299–310.
18. Casonato, S., Cervantes Sánchez, A., Haruki, H., Rengifo González, M., Provvedi, R., Dainese, E., Jaouen, T., Gola, S., Bini, E., Vicente, M., Johnsson, K., Ghisotti, D., Palù, G., Hernández-Pando, R., Manganelli, R., 2012. WhiB5, a transcriptional regulator that contributes to *Mycobacterium tuberculosis* virulence and reactivation. *Infect. Immun.* 80, 3132–3144.
19. Chahal, H.K., Dai, Y., Saini, A., Ayala-Castro, C., Outten, F.W., 2009. The SufBCD Fe-S scaffold complex interacts with SufA for Fe-S cluster transfer. *Biochemistry* 48, 10644–10653.
20. Chandramouli, K., Unciuleac, M.C., Naik, S., Dean, D.R., Huynh, B.H., Johnson, M.K., 2007. Formation and properties of [4Fe-4S] clusters on the IscU scaffold protein. *Biochemistry* 46, 6804–6811.
21. Chawla, M., Parikh, P., Saxena, A., Munshi, M., Mehta, M., Mai, D., Srivastava, A.K., Narasimhulu, K.V., Redding, K.E., Vashi, N., Kumar, D., Steyn, A.J.C., Singh, A., 2012. *Mycobacterium tuberculosis* WhiB4 regulates oxidative stress response to modulate survival and dissemination *in vivo*. *Mol. Microbiol.* 85, 1148–1165.
22. Colangeli, R., Haq, A., Arcus, V.L., Summers, E., Magliozzo, R.S., McBride, A., Mitra, A.K., Radjainia, M., Khajo, A., Jacobs, W.R., Salgame, P., Alland, D., 2009. The multifunctional histone-like protein Lsr2 protects mycobacteria against reactive oxygen intermediates. *Proc. Natl. Acad. Sci.* 106, 4414–4418.

23. Colangeli, R., Helb, D., Vilchèze, C., Hazbón, M.H., Lee, C.G., Safi, H., Sayers, B., Sardone, I., Jones, M.B., Fleischmann, R.D., Peterson, S.N., Jacobs, W.R., Alland, D., 2007. Transcriptional regulation of multi-drug tolerance and antibiotic-induced responses by the histone-like protein Lsr2 in *M. tuberculosis. PLoS Pathog.* 3. doi:10.1371/journal.ppat.0030087

24. Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry, C.E., Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M.A., Rajandream, M.A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J.E., Taylor, K., Whitehead, S., Barrell, B.G., 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 393, 537–544.

25. Cortes, T., Schubert, O.T., Rose, G., Arnvig, K.B., Comas, I., Aebersold, R., Young, D.B., 2013. Genome-wide mapping of transcriptional start sites defines an extensive leaderless transcriptome in *Mycobacterium tuberculosis. Cell Rep.* 5, 1121–1131.

26. Cubillos-Ruiz, A., Sandoval, A., Ritacco, V., López, B., Robledo, J., Correa, N., Hernandez-Neuta, I., Zambrano, M.M., Portillo, P.D., 2010. Genomic signatures of the Haarlem lineage of *Mycobacterium tuberculosis*: Implications of strain genetic variation in drug and vaccine development. *J. Clin. Microbiol.* 48, 3614–3623.

27. Dai, Y., Outten, F.W., 2012. The *E. coli* SufS-SufE sulfur transfer system is more resistant to oxidative stress than IscS-IscU. *FEBS Lett.* 586, 4016–4022.

28. Delmar, J.A., Chou, T.H., Wright, C.C., Licon, M.H., Doh, J.K., Radhakrishnan, A., Kumar, N., Lei, H.T., Bolla, J.R., Rajashankar, K.R., Su, C.C., Purdy, G.E., Yu, E.W., 2015. Structural basis for the regulation of the MmpL transporters of *Mycobacterium tuberculosis. J. Biol. Chem.* 290, 28559–28574.

29. Deretic, V., Song, J., Pagán-Ramos, E., 1997. Loss of *oxyR* in *Mycobacterium tuberculosis. Trends Microbiol.* 5, 367–372.

30. Desnoyers, G., Morissette, A., Prévost, K., Massé, E., 2009. Small RNA-induced differential degradation of the polycistronic mRNA *iscRSUA. EMBO J.* 28, 1551–1561.

31. Diaz, A.A., Tomba, E., Lennarson, R., Richard, R., Bagajewicz, M.J., Harrison, R.G., 2010. Prediction of protein solubility in *Escherichia coli* using logistic regression. *Biotechnol. Bioeng.* 105, 374–383.

32. Durand, S., Storz, G., 2010. Reprogramming of anaerobic metabolism by the FnrS small RNA. *Mol. Microbiol.* 75, 1215–1231.

33. Edelheit, O., Hanukoglu, A., Hanukoglu, I., 2009. Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies. *BMC Biotechnol.* 9, 61–68.

34. Ehira, S., Teramoto, H., Inui, M., Yukawa, H., 2009. Regulation of *Corynebacterium glutamicum* heat shock response by the extracytoplasmic-function sigma factor SigH and transcriptional regulators HspR and HrcA. *J. Bacteriol.* 191, 2964–2972.

35. Engelberg-Kulka, H., Schoulaker-Schwarz, R., 1994. Regulatory implications of translational frameshifting in cellular gene expression. *Mol. Microbiol.* 11, 3–8.

36. Evans, D.J., Leigh, G.J., 1991. The coordination of esters of amino acids to $[Fe4S4]^{2+}$ clusters. *J. Inorg. Biochem.* 42, 25–35.

37. Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G.A., Tate, J., Bateman, A., 2016. The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Res.* 44, 279–285.

38. Fleischhacker, A.S., Stubna, A., Hsueh, K.L., Guo, Y., Teter, S.J., Rose, J.C., Brunold, T.C., Markley, J.L., Münck, E., Kiley, P.J., 2012. Characterization of the [2Fe-2S] cluster of the *Escherichia coli* transcription factor IscR. *Biochemistry* 51, 4453–4462.

39. Flores-Valdez, M.A., Barba, J., Alvarez, A.H., 2012. *Mycobacterium tuberculosis* signaling via c-AMP. In: Cardona, P.J., (Ed.). *Understanding tuberculosis: Deciphering the secret life of the Bacilli.* pp. 103–117. Rijeka: InTech.

40. Fontán, P., Aris, V., Ghanny, S., Soteropoulos, P., Smith, I., 2008. Global transcriptional profile of *Mycobacterium tuberculosis* during THP-1 human macrophage infection. *Infect. Immun.* 76, 717–725.

41. Francis, D.M., Page, R., 2010. Strategies to optimize protein expression in *E. coli*. *Curr. Protoc. Protein Sci* 24, 1–29.

42. Freundlich, M., Ramani, N., Mathew, E., Sirko, A., Tsui, P., 1992. The role of integration host factor in gene expression in *Escherichia coli. Mol. Microbiol.* 6, 2557–2563.

43. Frota, C.C., Papavinasasundaram, K.G., Davis, E.O., Colston, M.J., 2004. The AraC family transcriptional regulator Rv1931c plays a role in the virulence of *Mycobacterium tuberculosis. Infect. Immun.* 72, 5483–5486.

44. Galagan, J.E., Sisk, P., Stolte, C., Weiner, B., Koehrsen, M., Wymore, F., Reddy, T.B.K., Zucker, J.D., Engels, R., Gellesch, M., Hubble, J., Jin, H., Larson, L., Mao, M., Nitzberg, M., White, J., Zachariah, Z.K., Sherlock, G., Ball, C.A., Schoolnik, G.K., 2010. TB database 2010: Overview and update. *Tuberculosis* 90, 225–235.

45. Gao, H., Zhou, D., Li, Y., Guo, Z., Han, Y., Song, Y., Zhai, J., Du, Z., Wang, X., Lu, J., Yang, R., 2008. The iron-responsive Fur regulon in *Yersinia pestis*. *J. Bacteriol.* 190, 3063–3075.

46. Gao, Q., Kripke, K., Arinc, Z., Voskuil, M., Small, P., 2004. Comparative expression studies of a complex phenotype: Cord formation in *Mycobacterium tuberculosis*. *Tuberculosis* 84, 188–196.

47. Garnier, T., Eiglmeier, K., Camus, J.C., Medina, N., Mansoor, H., Pryor, M., Duthoy, S., Grondin, S., Lacroix, C., Monsempe, C., Simon, S., Harris, B., Atkin, R., Doggett, J., Mayes, R., Keating, L., Wheeler, P.R., Parkhill, J., Barrell, B.G., Cole, S.T., Gordon, S.V., Hewinson, R.G., 2003. The complete genome sequence of Mycobacterium bovis. *Proc. Natl. Acad. Sci.* 100, 7877–7882.

48. Geiman, D.E., Raghunand, T.R., Agarwal, N., Bishai, W.R., 2006. Differential gene expression in response to exposure to antimycobacterial agents and other stress conditions among seven *Mycobacterium tuberculosis whiB*-like genes. *Antimicrob. Agents Chemother.* 50, 2836–2841.

49. Giel, J.L., Nesbit, A.D., Mettert, E.L., Fleischhacker, A.S., Wanta, B.T., Kiley, P.J., 2013. Regulation of iron-sulfur cluster homeostasis through transcriptional control of the isc pathway by [2Fe-2S]-IscR in *Escherichia coli*. *Mol. Microbiol.* 87, 478–492.

50. Giel, J.L., Rodionov, D., Liu, M., Blattner, F.R., Kiley, P.J., 2006. IscR-dependent gene expression links iron-sulphur cluster assembly to the control of O_2-regulated genes in *Escherichia coli*. *Mol. Microbiol.* 60, 1058–1075.

51. Gomez, J.E., Bishai, W.R., 2000. *whmD* is an essential mycobacterial gene required for proper septation and cell division. *Proc. Natl. Acad. Sci.* 97, 8554–8559.

52. González-Montalbán, N., García-Fruitós, E., Villaverde, A., 2007. Recombinant protein solubility - does more mean better? *Nat. Biotechnol.* 25, 718–720.

53. Govantes, F., Andújar, E., Santero, E., 1998. Mechanism of translational coupling in the *nifLA* operon of *Klebsiella pneumoniae*. *EMBO J.* 17, 2368–2377.

54. Griffin, J.E., Gawronski, J.D., DeJesus, M.A., Ioerger, T.R., Akerley, B.J., Sassetti, C.M., 2011. High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. *PLoS Pathog.* 7. doi:10.1371/journal.ppat.1002251

55. Guo, M., Feng, H., Zhang, J., Wang, W., Wang, Y., Li, Y., Gao, C., Chen, H., Feng, Y., He, Z.G., 2009. Dissecting transcription regulatory pathways through a new bacterial one-hybrid reporter system. *Genome Res.* 19, 1301–1308.

56. Gupta, V., Sendra, M., Naik, S.G., Chahal, H.K., Huynh, B.H., Outten, F.W., Fontecave, M., Ollagnier-de Choudens, S., 2009. Native *Escherichia coli* SufA, coexpressed with

SufBCDSE, purifies as a [2Fe−2S] protein and acts as an Fe−S transporter to Fe−S target enzymes. *J. Am. Chem. Soc.* 131, 6149–6153.

57. Gyaneshwar, P., Paliy, O., McAuliffe, J., Popham, D.L., Jordan, M.I., Kustu, S., 2005. Sulfur and nitrogen limitation in *Escherichia coli* K-12: Specific homeostatic responses. *J. Bacteriol.* 187, 1074–1090.
58. Haydel, S.E., Benjamin, W.H., Dunlap, N.E., Clark-Curtiss, J.E., 2002. Expression, autoregulation, and DNA-binding properties of the *Mycobacterium tuberculosis* TrcR response regulator. *J. Bacteriol.* 184, 2192–2203.
59. Haydel, S.E., Clark-Curtiss, J.E., 2006. The *Mycobacterium tuberculosis* TrcR response regulator represses transcription of the intracellularly expressed *Rv1057* gene, encoding a seven-bladed β-propeller. *J. Bacteriol.* 188, 150–159.
60. He, H., Bretl, D.J., Penoske, R.M., Anderson, D.M., Zahrt, T.C., 2011. Components of the *Rv0081-Rv0088* locus, which encodes a predicted formate hydrogenlyase complex, are coregulated by Rv0081, MprA, and DosR in *Mycobacterium tuberculosis*. *J. Bacteriol.* 193, 5105–5118.
61. Hellman, L.M., Fried, M.G., 2007. Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nat. Protoc.* 2, 1849–1861.
62. Hirabayashi, K., Yuda, E., Tanaka, N., Katayama, S., Iwasaki, K., Matsumoto, T., Kurisu, G., Outten, F.W., Fukuyama, K., Takahashi, Y., Wada, K., 2015. Functional dynamics revealed by the structure of the SufBCD complex, a novel ATP-binding cassette (ABC) protein that serves as a scaffold for iron-sulfur cluster biogenesis. *J. Biol. Chem.* 290, 29717–29731.
63. Hoff, K.G., Silberg, J.J., Vickery, L.E., 2000. Interaction of the iron-sulfur cluster assembly protein IscU with the Hsc66/Hsc20 molecular chaperone system of *Escherichia coli*. *Proc. Natl. Acad. Sci.* 97, 7790–7795.
64. 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. doi:10.1371/journal.ppat.1000988
65. Huet, G., Castaing, J.P., Fournier, D., Daffé, M., Saves, I., 2006. Protein splicing of SufB is crucial for the functionality of the *Mycobacterium tuberculosis* SUF Machinery. J. Bacteriol. 188, 3412–3414.
66. Huet, G., Daffé, M., Saves, I., 2005. Identification of the *Mycobacterium tuberculosis* SUF machinery as the exclusive mycobacterial system of [Fe-S] cluster assembly: Evidence for its implication in the pathogen's survival. *J. Bacteriol.* 187, 6137–6146.

67. Hurley, J.K., Weber-Main, A.M., Hodges, A.E., Stankovich, M.T., Benning, M.M., Holden, H.M., Cheng, H., Xia, B., Markley, J.L., Genzor, C., Gomez-Moreno, C., Hafezi, R., Tollin, G., 1997. Iron-sulfur cluster cysteine-to-serine mutants of *Anabaena* 2Fe-2S-ferredoxin exhibit unexpected redox properties and are competent in electron transfer to ferredoxin $NADP^+$ reductase. *Biochemistry* 36, 15109–15117.

68. Ioerger, T.R., Feng, Y., Ganesula, K., Chen, X., Dobos, K.M., Fortune, S., Jacobs, W.R., Mizrahi, V., Parish, T., Rubin, E., Sassetti, C., Sacchettini, J.C., 2010. Variation among genome sequences of H37Rv strains of *Mycobacterium tuberculosis* from multiple laboratories. *J. Bacteriol.* 192, 3645–3653.

69. Jacobson, M.R., Cash, V.L., Weiss, M.C., Laird, N.F., Newton, W.E., Dean, D.R., 1989. Biochemical and genetic analysis of the *nifUSVWZM* cluster from *Azotobacter vinelandii*. *Mol. Gen. Genet.* 219, 49–57.

70. Jang, S., Imlay, J.A., 2010. Hydrogen peroxide inactivates the *Escherichia coli* isc iron-sulfur assembly system, and OxyR induces the suf system to compensate. *Mol. Microbiol.* 78, 1448–1467.

71. Jenkins, V.A., Barton, G.R., Robertson, B.D., Williams, K.J., 2013. Genome wide analysis of the complete GlnR nitrogen-response regulon in *Mycobacterium smegmatis*. *BMC Genomics* 14, 1–15.

72. Kapopoulou, A., Lew, J.M., Cole, S.T., 2011. The MycoBrowser portal: A comprehensive and manually annotated resource for mycobacterial genomes. *Tuberculosis* 91, 8–13.

73. Kelly, S.M., Jess, T.J., Price, N.C., 2005. How to study proteins by circular dichroism. *Biochim. Biophys. Acta* 1751, 119–139.

74. Kendall, S.L., Burgess, P., Balhana, R., Withers, M., Ten Bokum, A., Lott, J.S., Gao, C., Uhia-Castro, I., Stoker, N.G., 2010. Cholesterol utilization in mycobacteria is controlled by two TetR-type transcriptional regulators: KstR and KstR2. *Microbiology* 156, 1362–1371.

75. Kim, J.H., Bothe, J.R., Alderson, T.R., Markley, J.L., 2015. Tangled web of interactions among proteins involved in iron–sulfur cluster assembly as unraveled by NMR, SAXS, chemical crosslinking, and functional studies. *Biochim. Biophys. Acta* 1853, 1416–1428.

76. Kim, J.H., Tonelli, M., Markley, J.L., 2012. Disordered form of the scaffold protein IscU is the substrate for iron-sulfur cluster assembly on cysteine desulfurase. *Proc. Natl. Acad. Sci.* 109, 454–459.

77. Kim, S.H., Lee, B.Y., Lau, G.W., Cho, Y.H., 2009. IscR modulates catalase A (KatA) activity, peroxide resistance and full virulence of *Pseudomonas aeruginosa PA14*. *J. Microbiol. Biotechnol.* 19, 1520–1526.

78. Konar, M., Alam, M.S., Arora, C., Agrawal, P., 2012. WhiB2/Rv3260c, a cell division-associated protein of *Mycobacterium tuberculosis* H37Rv, has properties of a chaperone. *FEBS J.* 279, 2781–2792.

79. Korduláková, J., Gilleron, M., Mikušová, K., Puzo, G., Brennan, P.J., Gicquel, B., Jackson, M., 2002. Definition of the first mannosylation step in phosphatidylinositol mannoside synthesis: PimA is essential for growth of mycobacteria. *J. Biol. Chem.* 277, 31335–31344.

80. Kumar, M., Khan, F.G., Sharma, S., Kumar, R., Faujdar, J., Sharma, R., Chauhan, D.S., Singh, R., Magotra, S.K., Khan, I.A., 2011. Identification of *Mycobacterium tuberculosis* genes preferentially expressed during human infection. *Microb. Pathog.* 50, 31–38.

81. Layer, G., Gaddam, S.A., Ayala-Castro, C.N., Ollagnier-de Choudens, S., Lascoux, D., Fontecave, M., Outten, F.W., 2007. SufE transfers sulfur from SufS to SufB for iron-sulfur cluster assembly. *J. Biol. Chem.* 282, 13342–13350.

82. Lee, J.H., Yeo, W.S., Roe, J.H., 2004. Induction of the *sufA* operon encoding Fe-S assembly proteins by superoxide generators and hydrogen peroxide: Involvement of OxyR, IHF and an unidentified oxidant-responsive factor. *Mol. Microbiol.* 51, 1745–1755.

83. Lee, K.C., Yeo, W.S., Roe, J.H., 2008. Oxidant-responsive induction of the suf operon, encoding a Fe-S assembly system, through Fur and IscR in *Escherichia coli. J. Bacteriol.* 190, 8244–8247.

84. Lew, J.M., Kapopoulou, A., Jones, L.M., Cole, S.T., 2011. TubercuList - 10 years after. *Tuberculosis* 91, 1–7.

85. Lim, J.G., Choi, S.H., 2014. IscR is a global regulator essential for pathogenesis of *Vibrio vulnificus* and induced by host cells. *Infect. Immun.* 82, 569–578.

86. Liu, T., Ramesh, A., Ma, Z., Ward, S.K., Zhang, L., George, G.N., Talaat, A.M., Sacchettini, J.C., Giedroc, D.P., 2007. CsoR is a novel *Mycobacterium tuberculosis* copper-sensing transcriptional regulator. *Nat. Chem. Biol.* 3, 60–68.

87. Ma, Z., Cowart, D.M., Scott, R.A., Giedroc, D.P., 2009. Molecular insights into the metal selectivity of the Cu(I)-sensing repressor CsoR from *Bacillus subtilis. Biochemistry* 48, 3325–3334.

88. 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. Microbiol.* 155, 1332–1339.

89. Massé, E., Gottesman, S., 2002. A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli. Proc. Natl. Acad. Sci.* 99, 4620–4625.

90. Massé, E., Vanderpool, C.K., Gottesman, S., 2005. Effect of RyhB small RNA on global iron use in *Escherichia coli. J. Bacteriol.* 187, 6962–6971.

91. McHugh, J.P., Rodríguez-Quiñones, F., Abdul-Tehrani, H., Svistunenko, D.A., Poole, R.K., Cooper, C.E., Andrews, S.C., 2003. Global iron-dependent gene regulation in *Escherichia coli*: A new mechanism for iron homeostasis. *J. Biol. Chem.* 278, 29478–29486.

92. Mettert, E.L., Kiley, P.J., 2014. Coordinate regulation of the *suf* and *isc* Fe-S cluster biogenesis pathways by IscR is essential for viability of *Escherichia coli. J. Bacteriol.* 196, 4315–4323.

93. Milano, A., Pasca, M.R., Provvedi, R., Lucarelli, A.P., Manina, G., Ribeiro, A.L. de J.L., Manganelli, R., Riccardi, G., 2009. Azole resistance in *Mycobacterium tuberculosis* is mediated by the MmpS5-MmpL5 efflux system. *Tuberculosis* 89, 84–90.

94. Miller, H.K., Kwuan, L., Schwiesow, L., Bernick, D.L., Mettert, E., Ramirez, H.A., Ragle, J.M., Chan, P.P., Kiley, P.J., Lowe, T.M., Auerbuch, V., 2014. IscR is essential for *Yersinia pseudotuberculosis* Type III secretion and virulence. *PLoS Pathog.* 10. doi:10.1371/journal.ppat.1004194

95. Milse, J., Petri, K., Rückert, C., Kalinowski, J., 2014. Transcriptional response of *Corynebacterium glutamicum* ATCC 13032 to hydrogen peroxide stress and characterization of the OxyR regulon. *J. Biotechnol.* 190, 40–54.

96. Minch, K.J., Rustad, T.R., Peterson, E.J.R., Winkler, J., Reiss, D.J., Ma, S., Hickey, M., Brabant, W., Morrison, B., Turkarslan, S., Mawhinney, C., Galagan, J.E., Price, N.D., Baliga, N.S., Sherman, D.R., 2015. The DNA-binding network of *Mycobacterium tuberculosis. Nat. Commun.* 6. doi:10.1038/ncomms6829

97. Morris, R.P., Nguyen, L., Gatfield, J., Visconti, K., Nguyen, K., Schnappinger, D., Ehrt, S., Liu, Y., Heifets, L., Pieters, J., Schoolnik, G., Thompson, C.J., 2005. Ancestral antibiotic resistance in *Mycobacterium tuberculosis. Proc. Natl. Acad. Sci.* 102, 12200–12205.

98. Murugasu-Oei, B., Tay, A., Dick, T., 1999. Upregulation of stress response genes and ABC transporters in anaerobic stationary-phase *Mycobacterium smegmatis. Mol. Gen. Genet.* 262, 677–682.

99. Nachin, L., Loiseau, L., Expert, D., Barras, F., 2003. SufC: An unorthodox cytoplasmic ABC/ATPase required for [Fe-S] biogenesis under oxidative stress. *EMBO J.* 22, 427–437.

100. Nakunst, D., Larisch, C., Hüser, A.T., Tauch, A., Pühler, A., Kalinowski, J., 2007. The extracytoplasmic function-type sigma factor SigM of *Corynebacterium glutamicum* ATCC

13032 is involved in transcription of disulfide stress-related genes. *J. Bacteriol.* 189, 4696–4707.

101. Nesbit, A.D., Giel, J.L., Rose, J.C., Kiley, P.J., 2009. Sequence-specific binding to a subset of IscR regulated promoters does not require IscR Fe-S cluster ligation. *J. Mol. Biol.* 387, 28–41.
102. Ollagnier-de Choudens, S., Mattioli, T., Takahashi, Y., Fontecave, M., 2001. Iron-sulfur cluster assembly: Characterization of IscA and evidence for a specific and functional complex with ferredoxin. *J. Biol. Chem.* 276, 22604–22607.
103. Olson, J.W., Agar, J.N., Johnson, M.K., Maier, R.J., 2000. Characterization of the NifU and NifS Fe−S cluster formation proteins essential for viability in *Helicobacter pylori*. *Biochemistry* 39, 16213–16219.
104. Outten, F.W., 2007. Iron-sulfur clusters as oxygen-responsive molecular switches. *Nat. Chem. Biol*. 3, 206-207.
105. Outten, F.W., Djaman, O., Storz, G., 2004. A *suf* operon requirement for Fe–S cluster assembly during iron starvation in *Escherichia coli. Mol. Microbiol.* 52, 861–872.
106. Outten, F.W., Wood, M.J., Munoz, F.M., Storz, G., 2003. The SufE protein and the SufBCD complex enhance SufS cysteine desulfurase activity as part of a sulfur transfer pathway for Fe-S cluster assembly in *Escherichia coli. J. Biol. Chem.* 278, 45713–45719.
107. Parish, T., Stoker, N.G., 2000. Use of a flexible cassette method to generate a double unmarked *Mycobacterium tuberculosis tlyA plcABC* mutant by gene replacement. *Microbiology* 146, 1969–1975.
108. Pátek, M., Nešvera, J., 2011. Sigma factors and promoters in *Corynebacterium glutamicum. J. Biotechnol.* 154, 101–113.
109. Prakash, P., Yellaboina, S., Ranjan, A., Hasnain, S.E., 2005. Computational prediction and experimental verification of novel IdeR binding sites in the upstream sequences of *Mycobacterium tuberculosis* open reading frames. *Bioinforma.* 21, 2161–2166.
110. Price, W.N., Handelman, S.K., Everett, J.K., Tong, S.N., Bracic, A., Luff, J.D., Naumov, V., Acton, T., Manor, P., Xiao, R., Rost, B., Montelione, G.T., Hunt, J.F., 2011. Large-scale experimental studies show unexpected amino acid effects on protein expression and solubility *in vivo* in *E. coli. Microb. Inform. Exp.* 1. doi:10.1186/2042-5783-1-6
111. Prischi, F., Pastore, C., Carroni, M., Iannuzzi, C., Adinolfi, S., Temussi, P., Pastore, A., 2010. Of the vulnerability of orphan complex proteins: The case study of the *E. coli* IscU and IscS proteins. *Protein Expr. Purif.* 73, 161–166.
112. Py, B., Barras, F., 2010. Building Fe-S proteins: Bacterial strategies. *Nat. Rev. Microbiol.* 8, 436–446.

113. Pym, A.S., Domenech, P., Honoré, N., Song, J., Deretic, V., Cole, S.T., 2001. Regulation of catalase-peroxidase (katG) expression, isoniazid sensitivity and virulence by FurA of *Mycobacterium tuberculosis*. *Mol. Microbiol.* 40, 879–889.

114. Radhakrishnan, A., Kumar, N., Wright, C.C., Chou, T.H., Tringides, M.L., Bolla, J.R., Lei, H.T., Rajashankar, K.R., Su, C.C., Purdy, G.E., Yu, E.W., 2014. Crystal structure of the transcriptional regulator Rv0678 of *Mycobacterium tuberculosis*. *J. Biol. Chem.* 289, 16526–16540.

115. Raghunand, T.R., Bishai, W.R., 2006. Mapping essential domains of *Mycobacterium smegmatis* WhmD: Insights into WhiB structure and function. *J. Bacteriol.* 188, 6966–6976.

116. Rajagopalan, S., Teter, S.J., Zwart, P.H., Brennan, R.G., Phillips, K.J., Kiley, P.J., 2013. Studies of IscR reveal a unique mechanism for metal-dependent regulation of DNA-binding specificity. *Nat. Struct. Mol. Biol.* 20, 740–747.

117. Reddy, T.B.K., Riley, R., Wymore, F., Montgomery, P., DeCaprio, D., Engels, R., Gellesch, M., Hubble, J., Jen, D., Jin, H., Koehrsen, M., Larson, L., Mao, M., Nitzberg, M., Sisk, P., Stolte, C., Weiner, B., White, J., Zachariah, Z.K., Sherlock, G., Galagan, J.E., Ball, C.A., Schoolnik, G.K., 2009. TB database: An integrated platform for tuberculosis research. *Nucleic Acids Res.* 37, 499–508.

118. Remes, B., Eisenhardt, B.D., Srinivasan, V., Klug, G., 2015. IscR of *Rhodobacter sphaeroides* functions as repressor of genes for iron-sulfur metabolism and represents a new type of iron-sulfur-binding protein. *MicrobiologyOpen* 4, 790–802.

119. Riboldi, G.P., Larson, T.J., Frazzon, J., 2011. *Enterococcus faecalis sufCDSUB* complements *Escherichia coli sufABCDSE*. *FEMS Microbiol. Lett.* 320, 15–24.

120. Riboldi, G.P., Verli, H., Frazzon, J., 2009. Structural studies of the *Enterococcus faecalis* SufU [Fe-S] cluster protein. *BMC Biochem.* 10. doi:10.1186/1471-2091-10-3

121. Rincon-Enriquez, G., Crété, P., Barras, F., Py, B., 2008. Biogenesis of Fe/S proteins and pathogenicity: IscR plays a key role in allowing *Erwinia chrysanthemi* to adapt to hostile conditions. *Mol. Microbiol.* 67, 1257–1273.

122. Roback, P., Beard, J., Baumann, D., Gille, C., Henry, K., Krohn, S., Wiste, H., Voskuil, M.I., Rainville, C., Rutherford, R., 2007. A predicted operon map for *Mycobacterium tuberculosis*. *Nucleic Acids Res.* 35, 5085–5095.

123. Robson, J., McKenzie, J.L., Cursons, R., Cook, G.M., Arcus, V.L., 2009. The *vapBC* operon from *Mycobacterium smegmatis* is an autoregulated toxin-antitoxin module that controls growth via inhibition of translation. *J. Mol. Biol.* 390, 353–367.

124. Roche, B., Aussel, L., Ezraty, B., Mandin, P., Py, B., Barras, F., 2013. Iron/sulfur proteins biogenesis in prokaryotes: Formation, regulation and diversity. *Biochim. Biophys. Acta* 1827, 455–469.

125. Rodrigue, S., Brodeur, J., Jacques, P.É., Gervais, A.L., Brzezinski, R., Gaudreau, L., 2007. Identification of mycobacterial σ factor binding sites by chromatin immunoprecipitation assays. *J. Bacteriol.* 189, 1505–1513.

126. Rodriguez, G.M., Voskuil, M.I., Gold, B., Schoolnik, G.K., Smith, I., 2002. *ideR*, an essential gene in *Mycobacterium tuberculosis*: Role of IdeR in iron-dependent gene expression, iron metabolism, and oxidative stress response. *Infect. Immun.* 70, 3371–3381.

127. Rohde, K.H., Abramovitch, R.B., Russell, D.G., 2007. *Mycobacterium tuberculosis* invasion of macrophages: Linking bacterial gene expression to environmental cues. *Cell Host Microbe* 2, 352–364.

128. Romsang, A., Duang-Nkern, J., Leesukon, P., Saninjuk, K., Vattanaviboon, P., Mongkolsuk, S., 2014. The iron-sulphur cluster biosynthesis regulator IscR contributes to iron homeostasis and resistance to oxidants in *Pseudomonas aeruginosa*. *PLoS ONE* 9. doi:10.1371/journal.pone.0086763

129. Runyen-Janecky, L., Daugherty, A., Lloyd, B., Wellington, C., Eskandarian, H., Sagransky, M., 2008. Role and regulation of iron-sulfur cluster biosynthesis genes in *Shigella flexneri* virulence. *Infect. Immun.* 76, 1083–1092.

130. Rustad, T.R., Harrell, M.I., Liao, R., Sherman, D.R., 2008. The enduring hypoxic response of *Mycobacterium tuberculosis. PLoS ONE* 3. doi:10.1371/journal.pone.0001502

131. Rustad, T.R., Minch, K.J., Ma, S., Winkler, J.K., Hobbs, S., Hickey, M., Brabant, W., Turkarslan, S., Price, N.D., Baliga, N.S., Sherman, D.R., 2014. Mapping and manipulating the *Mycobacterium tuberculosis* transcriptome using a transcription factor overexpression-derived regulatory network. *Genome Biol.* 15. doi:10.1186/s13059-014-0502-3

132. Rybniker, J., Pojer, F., Marienhagen, J., Kolly, G.S., Chen, J.M., van Gumpel, E., Hartmann, P., Cole, S.T., 2014. The cysteine desulfurase IscS of *Mycobacterium tuberculosis* is involved in iron-sulfur cluster biogenesis and oxidative stress defence. *Biochem. J.* 459, 467–478.

133. Ryle, M.J., Lanzilotta, W.N., Seefeldt, L.C., Scarrow, R.C., Jensen, G.M., 1996. Circular dichroism and X-ray spectroscopies of *Azotobacter vinelandii* nitrogenase iron protein: MgATP and MgADP induced protein conformational changes affecting the [4Fe-4S] cluster and characterization of a [2Fe-2S] form. *J. Biol. Chem.* 271, 1551–1557.

134. Sachdeva, P., Misra, R., Tyagi, A.K., Singh, Y., 2010. The sigma factors of *Mycobacterium tuberculosis*: Regulation of the regulators. *FEBS J.* 277, 605–626.

135. Saïd-Salim, B., Mostowy, S., Kristof, A.S., Behr, M.A., 2006. Mutations in *Mycobacterium tuberculosis Rv0444c*, the gene encoding anti-SigK, explain high level expression of MPB70 and MPB83 in *Mycobacterium bovis*. *Mol. Microbiol.* 62, 1251–1263.

136. Saini, A., Mapolelo, D.T., Chahal, H.K., Johnson, M.K., Outten, F.W., 2010. SufD and SufC ATPase activity are required for iron acquisition during *in vivo* Fe-S cluster formation on SufB. *Biochemistry* 49, 9402–9412.

137. Saini, V., Farhana, A., Glasgow, J.N., Steyn, A.J., 2012. Iron sulfur cluster proteins and microbial regulation: Implications for understanding tuberculosis. *Curr. Opin. Chem. Biol.* 16, 45–53.

138. Sala, C., Forti, F., Di Florio, E., Canneva, F., Milano, A., Riccardi, G., Ghisotti, D., 2003. *Mycobacterium tuberculosis* FurA autoregulates its own expression. *J. Bacteriol.* 185, 5357–5362.

139. Sambrook, J., Fritsch, E.F., Maniatis, T., 1998. Molecular cloning: A laboratory manual. 2nd Edition. New York: Cold Spring Harbor Laboratory.

140. Sampson, S.L., 2011. Mycobacterial PE/PPE proteins at the host-pathogen interface. *Clin. Dev. Immunol.* doi:10.1155/2011/497203

141. Sanchez-Puelles, J.M., Ronda, C., Garcia, J.L., Garcia, P., Lopez, R., Garcia, E., 1986. Searching for autolysin functions. Characterization of a pneumococcal mutant deleted in the *lytA* gene. *Eur. J. Biochem.* 158, 289–293.

142. Santos, J.A., Alonso-García, N., Macedo-Ribeiro, S., Pereira, P.J.B., 2014. The unique regulation of iron-sulfur cluster biogenesis in a Gram-positive bacterium. *Proc. Natl. Acad. Sci.* 111, 2251-2260.

143. 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.

144. Sassetti, C.M., Rubin, E.J., 2003. Genetic requirements for mycobacterial survival during infection. *Proc. Natl. Acad. Sci.* 100, 12989–12994.

145. Saves, I., Lewis, L.A., Westrelin, F., Warren, R., Daffé, M., Masson, J.M., 2002. Specificities and functions of the *recA* and *pps1* intein genes of *Mycobacterium tuberculosis* and application for diagnosis of tuberculosis. *J. Clin. Microbiol.* 40, 943–950.

146. Schnappinger, D., Ehrt, S., Voskuil, M.I., Liu, Y., Mangan, J.A., Monahan, I.M., Dolganov, G., Efron, B., Butcher, P.D., Nathan, C., Schoolnik, G.K., 2003. Transcriptional adaptation of *Mycobacterium tuberculosis* within macrophages. *J. Exp. Med.* 198, 693–704.

147. Schwab, U., Rohde, K.H., Wang, Z., Chess, P.R., Notter, R.H., Russell, D.G., 2009. Transcriptional responses of *Mycobacterium tuberculosis* to lung surfactant. *Microb. Pathog.* 46, 185–193.

148. Schwartz, C.J., Djaman, O., Imlay, J.A., Kiley, P.J., 2000. The cysteine desulfurase, IscS, has a major role in *in vivo* Fe-S cluster formation in *Escherichia coli*. *Proc. Natl. Acad. Sci.* 97, 9009–9014.

149. Schwartz, C.J., Giel, J.L., Patschkowski, T., Luther, C., Ruzicka, F.J., Beinert, H., Kiley, P.J., 2001. IscR, an Fe-S cluster-containing transcription factor, represses expression of *Escherichia coli* genes encoding Fe-S cluster assembly proteins. *Proc. Natl. Acad. Sci.* 98, 14895–14900.

150. Shen, G., Balasubramanian, R., Wang, T., Wu, Y., Hoffart, L.M., Krebs, C., Bryant, D.A., Golbeck, J.H., 2007. SufR coordinates two $[4Fe\text{-}4S]^{2+,\ 1+}$ clusters and functions as a transcriptional repressor of the *sufBCDS* operon and an autoregulator of *sufR* in cyanobacteria. *J. Biol. Chem.* 282, 31909–31919.

151. Singh, A., Crossman, D.K., Mai, D., Guidry, L., Voskuil, M.I., Renfrow, M.B., Steyn, A.J.C., 2009. *Mycobacterium tuberculosis* WhiB3 maintains redox homeostasis by regulating virulence lipid anabolism to modulate macrophage response. *PLoS Pathog.* 5. doi:10.1371/journal.ppat.1000545

152. Singh, A., Guidry, L., Narasimhulu, K.V., Mai, D., Trombley, J., Redding, K.E., Giles, G.I., Lancaster, J.R., Steyn, A.J.C., 2007. *Mycobacterium tuberculosis* WhiB3 responds to O_2 and nitric oxide via its [4Fe-4S] cluster and is essential for nutrient starvation survival. *Proc. Natl. Acad. Sci.* 104, 11562–11567.

153. Smith, L.J., Stapleton, M.R., Buxton, R.S., Green, J., 2012. Structure-function relationships of the *Mycobacterium tuberculosis* transcription factor WhiB1. *PloS ONE* 7. doi:10.1371/journal.pone.0040407

154. Smith, L.J., Stapleton, M.R., Fullstone, G.J.M., Crack, J.C., Thomson, A.J., Le Brun, N.E., Hunt, D.M., Harvey, E., Adinolfi, S., Buxton, R.S., Green, J., 2010. *Mycobacterium tuberculosis* WhiB1 is an essential DNA-binding protein with a nitric oxide-sensitive iron–sulfur cluster. *Biochem. J.* 432, 417–427.

155. Springer, B., Sander, P., Sedlacek, L., Ellrott, K., Böttger, E.C., 2001. Instability and site-specific excision of integration-proficient mycobacteriophage L5 plasmids: Development of stably maintained integrative vectors. *Int. J. Med. Microbiol.* 290, 669–675.

156. Steyn, A.J.C., Collins, D.M., Hondalus, M.K., Jacobs, W.R., Kawakami, R.P., Bloom, B.R., 2002. *Mycobacterium tuberculosis* WhiB3 interacts with RpoV to affect host survival but is dispensable for *in vivo* growth. *Proc. Natl. Acad. Sci.* 99, 3147–3152.

157. Stover, C.K., Cruz, V.F. de la, Fuerst, T.R., Burlein, J.E., Benson, L.A., Bennett, L.T., Bansal, G.P., Young, J.F., Lee, M.H., Hatfull, G.F., Snapper, S.B., Barletta, R.G., Jacobs, W.R., Bloom, B.R., 1991. New use of BCG for recombinant vaccines. *Nature* 351, 456–460.

158. Takahashi, Y., Tokumoto, U., 2002. A third bacterial system for the assembly of iron-sulfur clusters with homologs in archaea and plastids. *J. Biol. Chem.* 277, 28380–28383.

159. Tian, T., He, H., Liu, X.Q., 2014. The SufBCD protein complex is the scaffold for iron-sulfur cluster assembly in *Thermus thermophiles HB8. Biochem. Biophys. Res. Commun.* 443, 376–381.

160. Tokumoto, U., Kitamura, S., Fukuyama, K., Takahashi, Y., 2004. Interchangeability and distinct properties of bacterial Fe-S cluster assembly systems: Functional replacement of the *isc* and *suf* operons in *Escherichia coli* with the *nifSU*-like operon from *Helicobacter pylori. J. Biochem.* 136, 199–209.

161. Tokumoto, U., Takahashi, Y., 2001. Genetic analysis of the isc operon in *Escherichia coli* involved in the biogenesis of cellular iron-sulfur proteins. *J. Biochem.* 130, 63–71.

162. Topilina, N.I., Green, C.M., Jayachandran, P., Kelley, D.S., Stanger, M.J., Piazza, C.L., Nayak, S., Belfort, M., 2015. SufB intein of *Mycobacterium tuberculosis* as a sensor for oxidative and nitrosative stresses. *Proc. Natl. Acad. Sci.* 112, 10348–10353.

163. Trevino, S.R., Scholtz, J.M., Pace, C.N., 2007. Amino acid contribution to protein solubility: Asp, Glu, and Ser contribute more favorably than the other hydrophilic amino acids in RNase. *Sa. J. Mol. Biol.* 366, 449–460.

164. Unniraman, S., Chatterji, M., Nagaraja, V., 2002. DNA *gyrase* genes in *Mycobacterium tuberculosis*: A single operon driven by multiple promoters. *J. Bacteriol.* 184, 5449–5456.

165. Vatansever, F., de Melo, W.C.M.A., Avci, P., Vecchio, D., Sadasivam, M., Gupta, A., Chandran, R., Karimi, M., Parizotto, N.A., Yin, R., Tegos, G.P., Hamblin, M.R., 2013. Antimicrobial strategies centered around reactive oxygen species: Bactericidal antibiotics, photodynamic therapy, and beyond. *FEMS Microbiol. Rev.* 37, 955–989.

166. Via, L.E., Lin, P.L., Ray, S.M., Carrillo, J., Allen, S.S., Eum, S.Y., Taylor, K., Klein, E., Manjunatha, U., Gonzales, J., Lee, E.G., Park, S.K., Raleigh, J.A., Cho, S.N., McMurray, D.N., Flynn, J.L., Barry, C.E., 2008. Tuberculous granulomas are hypoxic in guinea pigs, rabbits, and nonhuman primates. *Infect. Immun.* 76, 2333–2340.

167. Vinella, D., Brochier-Armanet, C., Loiseau, L., Talla, E., Barras, F., 2009. Iron-sulfur (Fe/S) protein biogenesis: Phylogenomic and genetic studies of A-type carriers. *PLoS Genet.* 5. doi:10.1371/journal.pgen.1000497

168. Voskuil, M.I., Bartek, I.L., Visconti, K., Schoolnik, G.K., 2011. The response of *Mycobacterium tuberculosis* to reactive oxygen and nitrogen species. *Front. Microbiol.* 2. doi:10.3389/fmicb.2011.00105

169. 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, 705–713.

170. Wang, T., Shen, G., Balasubramanian, R., McIntosh, L., Bryant, D.A., Golbeck, J.H., 2004. The *sufR* gene (*sll0088* in *Synechocystis sp.* strain *PCC 6803*) functions as a repressor of the *sufBCDS* operon in iron-sulfur cluster biogenesis in cyanobacteria. *J. Bacteriol.* 186, 956–967.

171. Wardman, P., Candeias, L.P., 1996. Fenton chemistry: An introduction. *Radiat. Res.* 145, 523–531.

172. Warren, R., Van Helden, P., Van Pittius, N.G., 2009. Insertion element IS6110-based restriction fragment length polymorphism genotyping of *Mycobacterium tuberculosis*. In: Parish, T., Brown, A.C. (Eds.). *Methods in molecular biology: Mycobacteria protocols*. pp. 353–370. New Jersey: Humana Press.

173. Weerdenburg, E.M., Abdallah, A.M., Rangkuti, F., Ghany, M.A.E., Otto, T.D., Adroub, S.A., Molenaar, D., Ummels, R., Veen, K. Ter, Stempvoort, G. Van, Sar, A.M. Van der, Ali, S., Langridge, G.C., Thomson, N.R., Pain, A., Bitter, W., 2015. Genome-wide transposon mutagenesis indicates that *Mycobacterium marinum* customizes its virulence mechanisms for survival and replication in different hosts. *Infect. Immun.* 83, 1778–1788.

174. Wennerhold, J., Bott, M., 2006. The DtxR regulon of *Corynebacterium glutamicum. J. Bacteriol.* 188, 2907–2918.

175. WHO., 2015. *Global tuberculosis report*. 20th Edition. Geneva: WHO. Available from: http://www.who.int/tb/publications/global_report/en/ (Accessed on 2 February 2016).

176. Wollers, S., Layer, G., Garcia-Serres, R., Signor, L., Clemancey, M., Latour, J.M., Fontecave, M., Ollagnier-de Choudens, S., 2010. Iron-sulfur (Fe-S) cluster assembly: The SufBCD complex is a new type of Fe-S scaffold with a flavin redox cofactor. *J. Biol. Chem.* 285, 23331–23341.

177. Wong, D.K., Lee, B.Y., Horwitz, M.A., Gibson, B.W., 1999. Identification of Fur, aconitase, and other proteins expressed by *Mycobacterium tuberculosis* under conditions of low and high concentrations of iron by combined two-dimensional gel electrophoresis and mass spectrometry. *Infect. Immun.* 67, 327–336.

178. Wong, S.M., Bernui, M., Shen, H., Akerley, B.J., 2013. Genome-wide fitness profiling reveals adaptations required by *Haemophilus* in coinfection with Influenza A virus in the murine lung. *Proc. Natl. Acad. Sci.* 110, 15413–15418.

179. Wright, P.R., Richter, A.S., Papenfort, K., Mann, M., Vogel, J., Hess, W.R., Backofen, R., Georg, J., 2013. Comparative genomics boosts target prediction for bacterial small RNAs. *Proc. Natl. Acad. Sci.* 110, 3487–3496.

180. Wu, C.C., Wang, C.K., Chen, Y.C., Lin, T.H., Jinn, T.R., Lin, C.T., 2014. IscR regulation of capsular polysaccharide biosynthesis and iron-acquisition systems in *Klebsiella pneumoniae* CG43. *PLoS ONE* 9. doi:10.1371/journal.pone.0107812

181. Wu, Y., Outten, F.W., 2009. IscR controls iron-dependent biofilm formation in *Escherichia coli* by regulating Type I fimbria expression. *J. Bacteriol.* 191, 1248–1257.

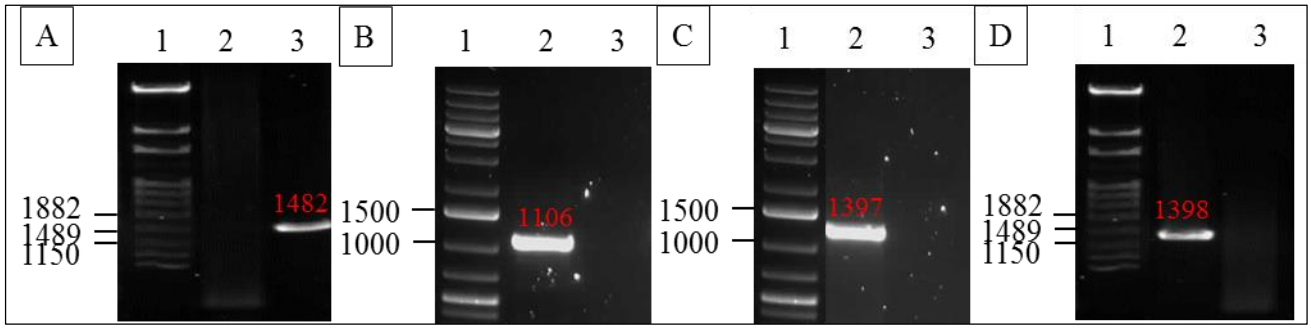
182. Yeo, W.S., Lee, J.H., Lee, K.C., Roe, J.H., 2006. IscR acts as an activator in response to oxidative stress for the *suf* operon encoding Fe-S assembly proteins. *Mol. Microbiol.* 61, 206–218.

183. Yoon, T., Cowan, J.A., 2003. Iron-sulfur cluster biosynthesis: Characterization of frataxin as an iron donor for assembly of [2Fe-2S] clusters in ISU-type proteins. *J. Am. Chem. Soc.* 125, 6078–6084.

184. Zheng, L., Cash, V.L., Flint, D.H., Dean, D.R., 1998. Assembly of iron-sulfur clusters: Identification of an *iscSUA-hscBA-fdx* gene cluster from *Azotobacter vinelandii. J. Biol. Chem.* 273, 13264–13272.

185. Zhu, L., Zhang, Y., Teh, J.S., Zhang, J., Connell, N., Rubin, H., Inouye, M., 2006. Characterization of mRNA interferases from *Mycobacterium tuberculosis. J. Biol. Chem.* 281, 18638–18643.

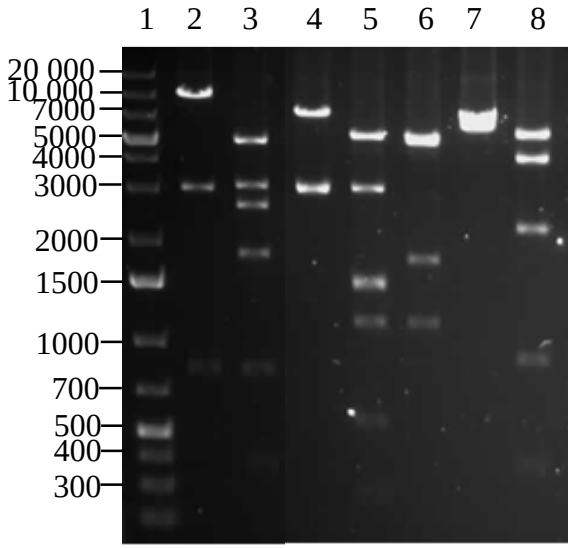
9. Supplementary information

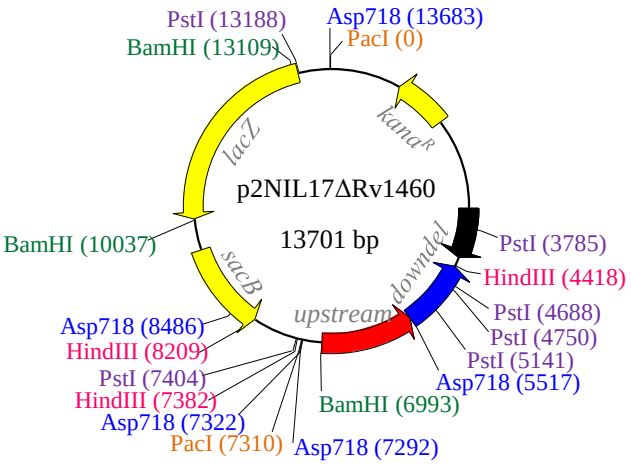


Supplementary Figure 1. Amplification of regions for the generation of suicide delivery vectors.

PCR products for (A) *Rv1460 upstream*, (B) *downdel*, (C) *downDNAbddel* and (D) *downDNAbddelstop*. Refer to Table 4 for a detailed description of the regions.

A



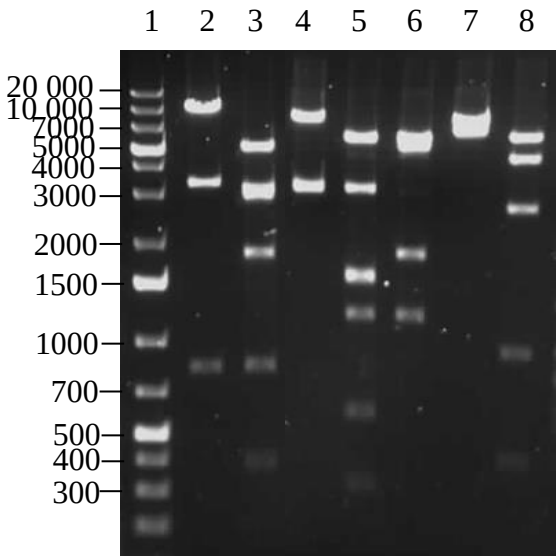


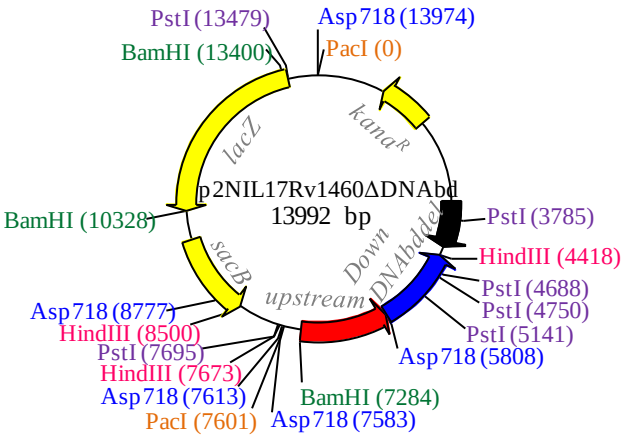
2	3	4	5	6	7	8
HindIII	HindIII + BamHI	BamHI	Asp718 + BamHI	Asp718	PacI	PstI
9910	5010	7285	5535	5535	7310	5784
2964	3072	3072	3072	5197	6391	4298
827	2575	3044	1551	1775		2263
	1828		1476	1164		903
	827		1164			391
	389		574			
			299			

Supplementary Figure 2. Generation of suicide delivery vectors used for two-step allelic exchange mutagenesis of *Rv1460*. RE mapping of suicide delivery vectors (A) p2NIL17ΔRv1460, (B) p2NIL17Rv1460ΔDNAbd and (C) p2NIL17Rv1460stop: lane (1) DNA ladder, (2) HindIII, (3) HindIII + BamHI, (4) BamHI, (5) Asp718 + BamHI, (6) Asp718, (7) PacI and (8) PstI. RE maps and tables of the expected product sizes for each RE digest are included.

Continued overleaf

B



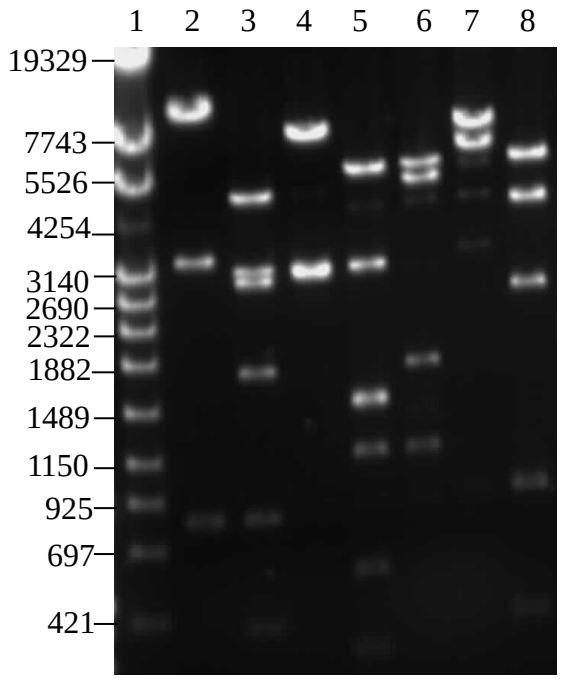


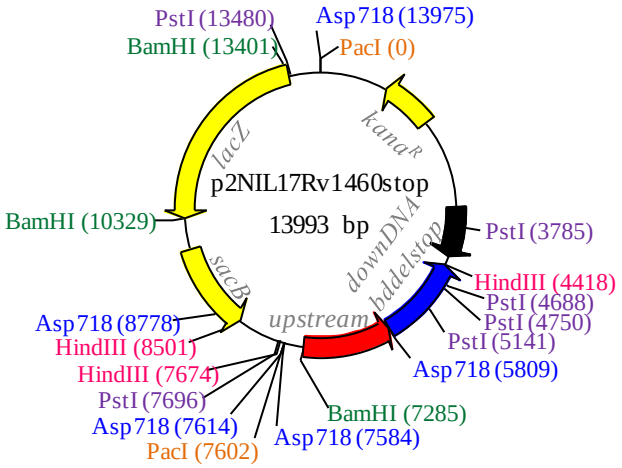
2	3	4	5	6	7	8
HindIII	HindIII + BamHI	BamHI	Asp718 + BamHI	Asp718	PacI	PstI
9910	5010	7876	5826	5826	7601	5784
3255	3072	3072	3072	5197	6391	4298
827	2866	3044	1551	1775		2554
	1828		1476	1164		903
	827		1164			391
	389		574			
			299			

Supplementary Figure 2. Generation of suicide delivery vectors used for two-step allelic exchange mutagenesis of *Rv1460*. RE mapping of suicide delivery vectors (A) p2NIL17ΔRv1460, (B) p2NIL17Rv1460ΔDNAbd and (C) p2NIL17Rv1460stop: lane (1) DNA ladder, (2) HindIII, (3) HindIII + BamHI, (4) BamHI, (5) Asp718 + BamHI, (6) Asp718, (7) PacI and (8) PstI. RE maps and tables of the expected product sizes for each RE digest are included.

Continued overleaf

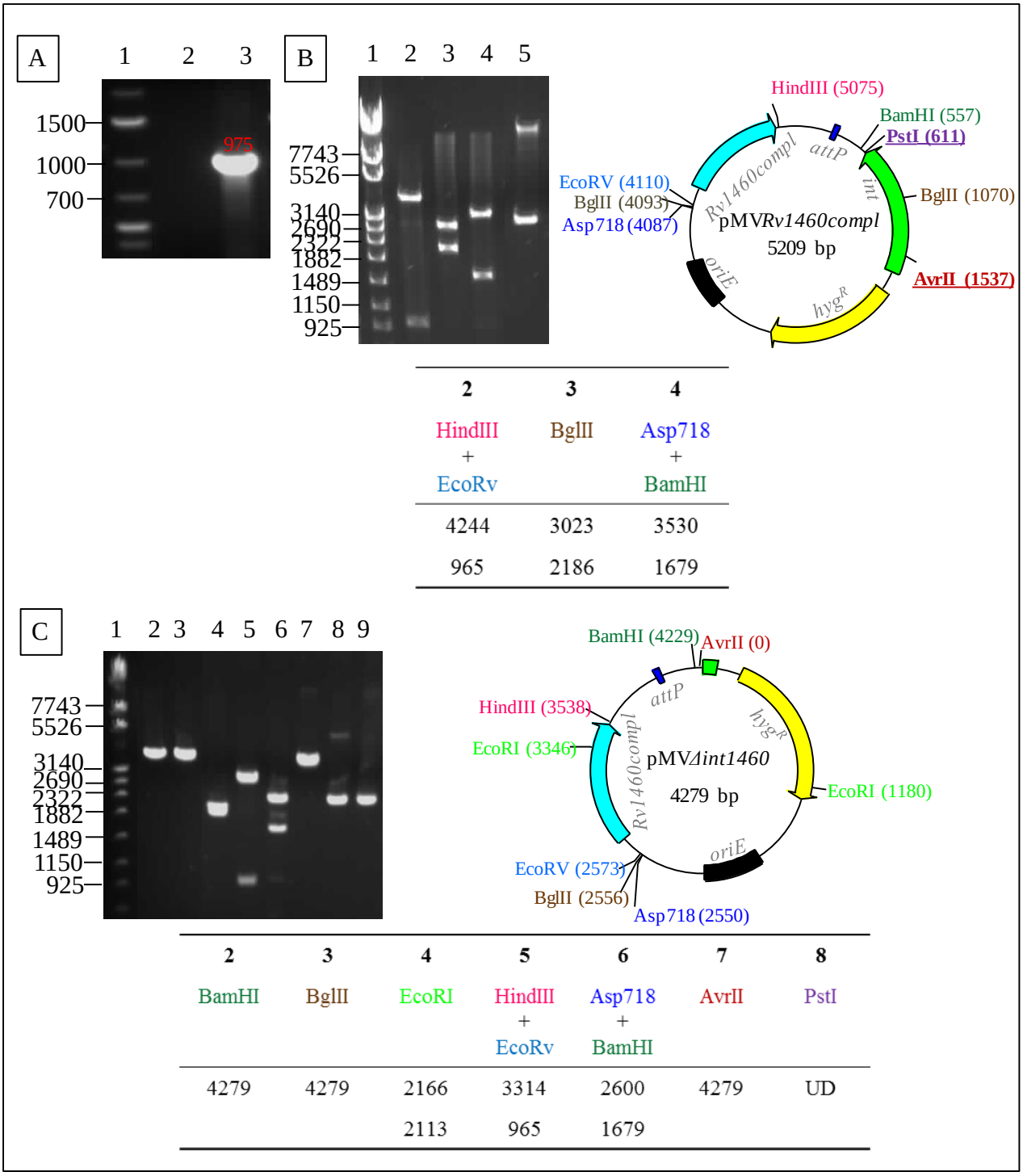
C





2	3	4	5	6	7	8
HindIII	HindIII + BamHI	BamHI	Asp718 + BamHI	Asp718	PacI	PstI
9910	5010	7877	5827	5827	7602	5784
3256	3072	3072	3072	5197	6391	4298
827	2867	3044	1551	1775		2555
	1828		1476	1164		903
	827		1164			391
	389		574			
			299			

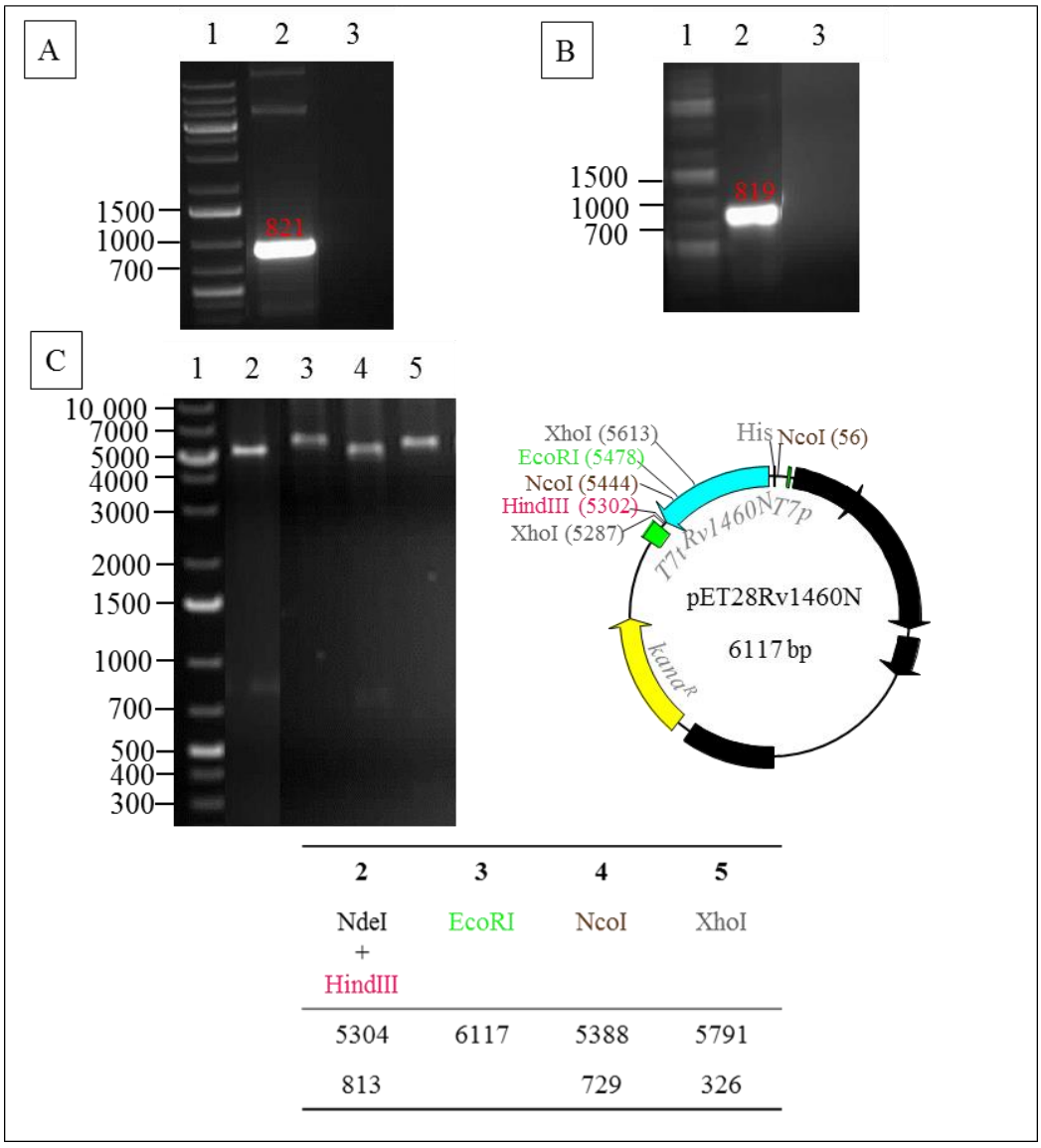
Supplementary Figure 2. Generation of suicide delivery vectors used for two-step allelic exchange mutagenesis of *Rv1460*. RE mapping of suicide delivery vectors (A) p2NIL17ΔRv1460, (B) p2NIL17Rv1460ΔDNAbd and (C) p2NIL17Rv1460stop: lane (1) DNA ladder, (2) HindIII, (3) HindIII + BamHI, (4) BamHI, (5) Asp718 + BamHI, (6) Asp718, (7) PacI and (8) PstI. RE maps and tables of the expected product sizes for each RE digest are included.



2	3	4
HindIII + EcoRv	BglII	Asp718 + BamHI
4244	3023	3530
965	2186	1679

2	3	4	5	6	7	8
BamHI	BglII	EcoRI	HindIII + EcoRv	Asp718 + BamHI	AvrII	PstI
4279	4279	2166	3314	2600	4279	UD
		2113	965	1679		

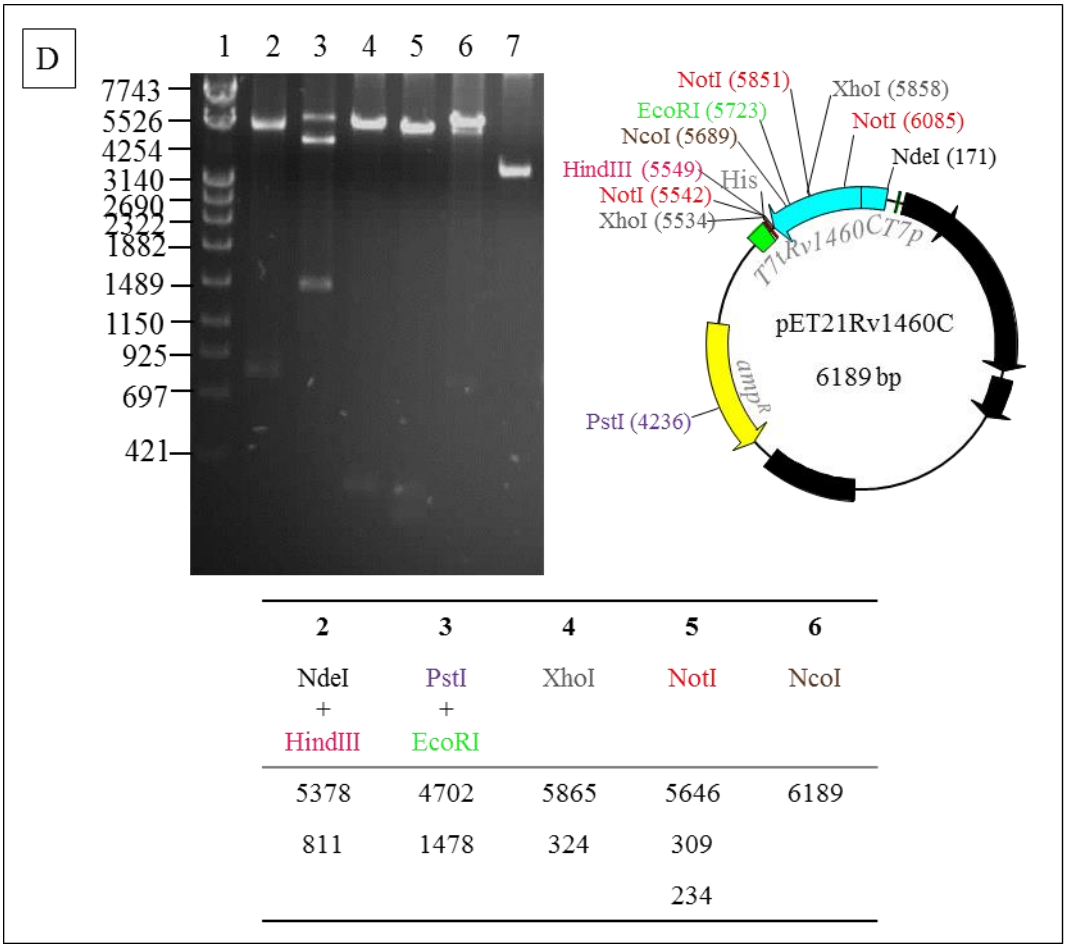
Supplementary Figure 3. Generation of *Rv1460* complementation vectors. (A) PCR product for *Rv1460compl* (Table 4). RE mapping of (B) pMV*Rv1460compl* vector: lane (1) DNA ladder (2) HindIII + EcoRV (3) BglII, (4) Asp718 + BamHI and (5) undigested and (C) pMV*Δint1460* vector: lane (1) DNA ladder, (2) BamHI, (3) BglII, (4) EcoRI, (5) HindIII + EcoRV, (6) Asp718 + BamHI, (7) AvrII, (8) PstI and (9) undigested. RE maps and tables of the expected product sizes for each RE digest are included.



2	3	4	5
NdeI + HindIII	EcoRI	NcoI	XhoI
5304	6117	5388	5791
813		729	326

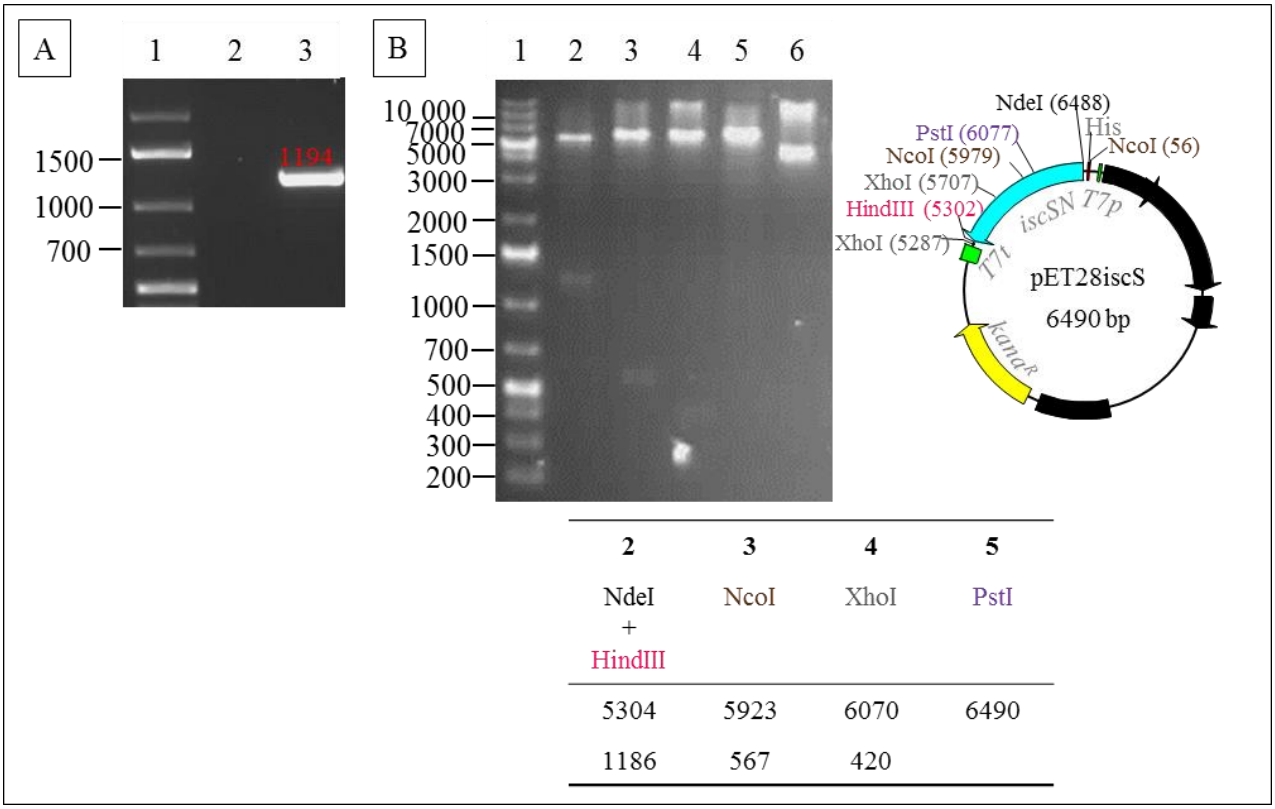
Supplementary Figure 4. Generation of pET28*Rv1460N* and pET21*Rv1460C* protein production vectors. (A) PCR product for *Rv1460N* and (B) *Rv1460C* (Table 7). RE mapping of (C) pET28*Rv1460N* vector: lane (1) DNA ladder, (2) NdeI + HindIII, (3) EcoRI, (4) NcoI and (5) XhoI, and (D) pET21*Rv1460C* vector: lane (1) DNA ladder, (2) NdeI + HindIII, (3) PstI + EcoRI, (4) XhoI, (5) NotI, (6) NcoI and (7) undigested. RE maps and tables of the expected product sizes for each RE digest are included.

Continued overleaf



2	3	4	5	6
NdeI + HindIII	PstI + EcoRI	XhoI	NotI	NcoI
5378	4702	5865	5646	6189
811	1478	324	309	
			234	

Supplementary Figure 4. Generation of pET28*Rv1460N* and pET21*Rv1460C* protein production vectors. (A) PCR product for *Rv1460N* and (B) *Rv1460C* (Table 7). RE mapping of (C) pET28*Rv1460N* vector: lane (1) DNA ladder, (2) NdeI + HindIII, (3) EcoRI, (4) NcoI and (5) XhoI, and (D) pET21*Rv1460C* vector: lane (1) DNA ladder, (2) NdeI + HindIII, (3) PstI + EcoRI, (4) XhoI, (5) NotI, (6) NcoI and (7) undigested. RE maps and tables of the expected product sizes for each RE digest are included.



2	3	4	5
NdeI + HindIII	NcoI	XhoI	PstI
5304	5923	6070	6490
1186	567	420	

Supplementary Figure 5. Generation of pET28iscS protein production vector. (A) PCR product for *iscSN* (Table 7). (B) RE mapping of the pET28iscS vector: lane (1) DNA ladder, (2) NdeI + HindIII, (3) NcoI, (4) XhoI, (5) PstI and (6) undigested. RE maps and tables of the expected product sizes for each RE digest are included.