MRC Centre for Molecular Bacteriology and Infection

Imperial College London

Investigating the role of cAMP in mycobacterial antimicrobial drug tolerance by the discovery of a novel cAMP-phosphodiesterase

Michael Thomson

Thesis submitted for the degree of Doctor of Philosophy of Imperial College London

June 2020

MRC CMBI
Life Sciences Department
Imperial College London
South Kensington Campus
SW7 2AZ

Declaration of Originality

I, Michael Thomson confirm that all work presented in this thesis is either my own or attributed to the respective authors.

Copyright Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution Non-Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Page 4 of 6 Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law

Acknowledgements

I would like to thank my supervisor Dr Gerald Larrouy-Maumus, for the great opportunity to pursue my PhD in his lab. You're a brilliant scientist and I've learnt a huge amount from my time in the group. I am very grateful for the support during my time in the lab and during the thesis writing process. I'd also like to thank the Medical Research Council and the CMBI for funding my PhD.

I want to thank all the past and present members of the CBTB lab group – especially Sonia Rebollo Ramirez and Ashleigh Cheyne. Ever since our hilarious first meeting, you've been great friends and really amazing co-workers. Agnieszka Broda, Yi Liu and Vivian Tounta, it's been great getting to know you over the years, becoming great friends and co-workers – thank you also for all your help. I need to thank Kanokkan Nunta and Krista Grimes, for being some of the best students and friends I could ask for, and for working tirelessly on my project. Thank you also Dr. Nitya Krishnan for your expert Cat III training, general amazing advice and great friendship. I hope to see you all when it doesn't have to be on teams! I am also very grateful to Dr Huw Williams my co-supervisor, and Dr Brian Robertson for your advice and support throughout my PhD.

Thank you everyone on CMBI Flowers Level 3 and 4 – but especially Dr. Charlie Millership and Dr Julia Sanchez Garrido for all your help, advice and lunch recommendations! Thank you also to the amazing Ex-president of the CPG (but always in the constitution), Will Pearson for all the hard work leading our CPG committee to really improve the CMBI. Thanks also for all the great times discussing everything from the intricacies of metabolomics – to the logistics of squirrel leper colonies!

Most importantly, I'd like to say how immensely grateful I am to my family and my incredible fiancé Sabreen. You have all been so helpful, understanding and supportive over my whole PhD. I couldn't have done it without you. I'm really blessed and I hope you know how much I appreciate you!

Abstract

M. tuberculosis, the causative agent of Tuberculosis (TB), is an ancient pathogen that has plagued mankind for over 70,000 years. In 2018, TB was responsible globally for 10 million new infections and 1.5 million deaths - more than any other infectious disease. *M. tuberculosis* bacilli have evolved to thrive in the harsh, nutrient limited environment within the host alveolar macrophage and evade the constant pressure of immune cell mediated killing. This evolution has led to the bacilli developing phenotypic adaptions that concurrently, drastically decrease their susceptibility to many antimicrobials. The basis for phenotypic adaptions is signalling to detect an environmental stimulus and to mediate an appropriate response. To this end, *M. tuberculosis* and other mycobacteria have evolved a robust cyclic AMP (cAMP) signalling system with multiple cAMP producing and cAMP binding effector proteins. Several of these proteins have already been shown to regulate virulence, carbon metabolism and essential gene expression. However, the link between cAMP signalling and antimicrobial susceptibility in mycobacteria has not previously been investigated.

In this project, I identified a new cAMP degrading phosphodiesterase enzyme (Rv1339) and used it as a tool to significantly decrease intrabacterial levels of cAMP in mycobacteria. The effect of this in *M. smegmatis* mc^2155 was to increase antimicrobial susceptibility. By using a combination of metabolomics, RNA-sequencing, antimicrobial susceptibility assays and bioenergetics analysis, I was able to characterise the potential mechanism behind this increased susceptibility. I was also able to begin preliminary work required to investigate this link in *M. tuberculosis* H37Rv. This work represents a proof-of-concept that targeting cAMP signalling is a promising new avenue for antimicrobial development, and expands our understanding of cAMP signalling in mycobacteria.

Table of Contents

Abbreviations 9

Chapter 1 - Introduction 13

1.1 Diseases caused by mycobacteria, their major impact on global health and the dangers of Antimicrobial Resistance/Antimicrobial Tolerance 14

1.1.1 Tuberculosis 14

1.2 Antimicrobial resistance and tolerance 16

1.2.1 Mycobacteria and Drug tolerance 20

1.3 Cell signalling and signal transduction 23

1.4 Cyclic AMP, the first second messenger to be discovered 24

1.6 Pathogens, cyclic nucleotide signalling and virulence 29

1.6.1 Cyclic nucleotides involved in pathogen signaling and virulence 29

1.6.2 Cyclic Di-nucleotides 30

1.6.3 Alarmones 32

1.7 Cyclic AMP toxins 34

1.7.1 Cholera toxin of *Vibrio cholerae* 35

1.7.2 Pertussis toxin and Adenylate cyclase toxin of *Bordetella pertussis* 36

1.7.3 Edema toxin of *Bacillus anthracis* 37

1.7.4 ExoY toxin of Pseudomonas aeruginosa 39

1.7. 5 Labile toxin of *E. coli* 40

1.8 cAMP and gene regulation in different pathogenic bacteria 41

1.8.1 The original cAMP bacterial gene regulation paradigm and current understanding in *E. coli* 41

1.8.2 cAMP mediated virulence gene regulation in *Pseudomonas aeruginosa* 45

1.8.3 cAMP mediated gene regulation of virulence and lifestyle switching in *Vibrio cholerae* 46

1.9 Mycobacteria and cAMP 47

1.10 cAMP intoxication of macrophages and *M. tuberculosis* subsequent evasion of killing 49

1.11 cAMP and mycobacterial response to the host environmental conditions 50

1.12 cAMP mediated gene regulation in mycobacteria 54

1.13 cAMP and mycobacterial metabolism 57

1.14 Modulating cAMP and antimicrobial tolerance in bacterial pathogens and mycobacteria 60

1.15 Aims 65

Chapter 2 - Materials and Methods 66

2.0 Materials 67

2.1 Methods 67

2.1.1 Bacterial strains and growth conditions 67

2.2 Culture Medium 71
2.2.1 7H9 liquid medium 71
2.2.2 Un-buffered 7H9 liquid medium for Seahorse XFP assays 71
2.2.3 7H10 agar 71
2.2.4 7H9 and 7H10 Medium supplement 71
2.2.5 Terrific broth for *E. coli* FX cloning construct purifications 72
2.2.7 Growth curves and Optical density (OD) monitoring 72
2.2.8 Bicinchoninic acid assay (BCA) 72
2.3 SDS-PAGE and Western blot analysis 73
2.3.1 His-tagged protein analysis in mycobacteria 73
2.4 Chromatography 75
2.4.1 Thin Layer Chromatography 75
2.4.2 PDE activity assay 75
2.4.3 Cell lysis 75
2.5 Cloning and generation of mutants 77
2.5.1 Proteomic analysis 77
2.5.2 Fx cloning expression trials and preparations 77
2.5.4: *E. coli* constructs generated as part of this work 79
2.5.5 Primers used during this work 80
2.6 Phenotypic characterisation techniques 81
2.6.1 Minimal Inhibitory Concentration by Resazurin Microtiter Assay (REMA) 81
2.6.2 Time-to-kill Curves 81
2.6.3 Hoechst Staining 82
2.6.4 Crystal violet spotting assay 82
2.6.5 Intracellular ATP Measurement 82
2.6.7 RNA extraction 83
2.6.8 RNA-seq data analysis 83
2.6.9 Seahorse XFP analysis of *M. smegmatis* mc^2155 84
2.7 LC-MS and metabolomics 85
Metabolite extraction experiments 85
2.7.1 Liquid-chromatography-mass spectrometry for targeted and untargeted metabolomics 85
2.7.2 LC/Q-TOF data acquisition using HILIC-Z chromatography 86
U-^{13}C-Labelling analysis 88
2.8 Statistical analysis 88
Chapter 3 - Identification of Rv1339, a conserved actinobacteria enzyme with putative cAMP PDE hydrolase activity 89

3.0 Identification of Rv1339, a conserved actinobacteria enzyme with putative cAMP PDE hydrolase activity 90
3.1 Preliminary work on the project 90
3.2 *M. bovis* BCG displays cAMP PDE activity even in the absence of Rv0805, the only currently annotated cAMP PDE enzyme 91
3.3 Screening for other cAMP PDE enzymes via homology to Rv0805 in *M. bovis* BCG 93
3.4 Optimisation of Rv0805 homologue expression, lysis and further screening for activity 96
3.5 Identification of a novel cAMP PDE enzyme via an unbiased activity led, biochemical approach 100
3.5.1 Summary of the unbiased activity-led biochemical approach 102
3.5.2 CaptoQ purification of BCG ΔRv0805 lysate identifies 2 fractions with cAMP PDE activity 102
3.5.3 Phenyl Sepharose purification of active CaptoQ fractions B11 and B10 104
3.5.4 Size exclusion of the CaptoQ active fraction B11 106
3.5.5 Proteomic analysis 108
3.6 Discussion 110
Chapter 4 - Expression, purification and enzymatic characterisation of Rv1339 112
4.0 Rv1339 is a putative atypical class-II PDE enzymes, from the metallobetalactamase superfamily 113
4.1 Bioinformatic and evolutionary analysis of Rv1339, and atypical Class-II PDE family enzyme 115
4.1.1 The predicted active site of Rv1339 118
4.2 Expression of Rv1339 in *M. smegmatis* mc^{2}155 120
4.3 Full length Rv1339 expression and purifications 125
4.3 Rv1339 purification approaches in *E. coli* 127
4.3.1 Rv1339 expression in BL21 (DE3) STAR *E. coli* leads to toxicity and bacterial death 127
4.3.3 Expression of Rv1339 in the periplasm of *E. coli* 139
4.3.4 Expression of an activity mutant – Rv1339 D180A in *E. coli* 143
4.4 Purifications of Rv1339 in mycobacteria 145
4.4.1 Mycobacterial Fx cloning expression vectors 145
4.5 Purification of Rv1339L in mycobacteria with the constitutive expression vector pVV16 148
4.5.1 Approaches to dissociate the potential chaperone-Rv1339 complex in *M. smegmatis* purifications 155
4.6 Cell free expression 161
4.7 Investigation of Rv1339 substrates specificity from expression of an Rv1339 and an Rv1339 D180A pVV16 construct in *M. smegmatis mc^{2}155* clear soluble lysate 164
4.8 Discussion 168
Chapter 5 - Phenotypic characterisation of Rv1339 expression and decreased cAMP levels in *M. smegmatis* mc^{2}155 and *M. tuberculosis* H37Rv 171

5.0 Phenotypic characterisation of Rv1339 expression and decreased cAMP levels in *M. smegmatis* mc^{2}155 and *M. tuberculosis* H37Rv 172

5.1 Expression of Rv1339 and Rv1339 D180 174

5.1.1 Rv1339 expression leads to significantly decreased cAMP levels, increased turnover and a growth defect 175

5.2 Rv1339 expression leads to increased antimicrobial susceptibility in *M. smegmatis mc^{2}155*, and the D180A catalysis mutant can ablate this increase 180

5.3 Expression of Rv1339 leads to an altered transcriptome, with decreased expression of genes known to be regulated by cAMP 185

5.4 Rv1339 expression decreases succinate dehydrogenase activity, decreases membrane potential and leads to compromised bioenergetics 191

5.5 Untargeted metabolomics analysis reveals that Rv1339 expression alters the metabolome, leading to increased plasticity in peptidoglycan synthesis metabolites 196

5.6 Rv1339 expression leads to increased permeability 200

5.7 Discussion 201

5.8 Preliminary work to validate the findings from *M. smegmatis* mc^{2}155 in *M. tuberculosis* H37Rv 206

5.8.1 Transformation and potential expression in *M. tuberculosis* H37Rv parental, ΔCrp_{MT} and ΔRv0805 with pVV16::*rv1339* full-length 206

5.8.2 Rv1339 expression decreases cAMP levels in the *M. tuberculosis* H37Rv parental strain, displays no effect on cAMP levels in the ΔRv0805 strain and increase cAMP levels in the ΔCrp_{MT} strain 208

5.8.3 Rv1339 expression leads to significant alterations in the metabolome of *M. tuberculosis* H37Rv parental, ΔRv0805 and ΔCrp_{MT} strains 209

5.9 Rv1339 expression is sufficient to ablate the growth defect of the *M. tuberculosis* H37Rv ΔCrp background 216

Chapter 6 - Final Discussion 218

6.0 Final Discussion 219

Chapter 7 - Annexes 225

7.1.1 Annex 1: 226

cAMP PDE activity assay performed on the second batch of M. bovis BCG lysate, subsequently used for the second CaptoQ and first SEC purifications. 226

7.1.2 Annex 2: 227

The CaptoQ purification of the second batch of *M. bovis* BCG Δ*mb0826* (Rv0805 homologue) clear soluble lysate. A: Chromatogram of the purification. B: Coomassie stained SDS-PAGE of the fractions, correlated to the LC-MS readout of the cAMP PDE activity assay performed on the indicated fractions. C: Thin Layer Chromatography (TLC) visualisation of the cAMP PDE activity assay performed on the indicated fractions from the CaptoQ purification. 227

7.1.3 Annex 3: 228

7.1.4 Annex 4: 229

7.1.5 Annex 5: 230
7.1.6 Annex 6 231
8.0 References 233

Abbreviations

Acronym	Meaning
ABC	ATP-Binding-Cassette
AC	Adenylate Cyclase
ACS	Acetyl-CoA-Synthase
ACT	Adenylate Cyclase Toxin
ADP	Adenosine Diphosphate
AI-1	Autoinducer-1
AMP	Adenosine monophosphate
AMR	Antimicrobial Resistance
AMT	Antimicrobial Tolerance
ATC	Anhydrotetracycline
ATP	Adenosine-Tri-Phosphate
BCA	Bicinchoninic Acid
BCG	Bacillus Calmette–Guérin
BSA	Bovine-Serum-Albumin
BWA	Burrows-Wheeler Transform
CA	Carbonic anhydrase
CAI-1	Cholera autoinducer-1
cAMP	Cyclic Adenosine Monophosphate
CCR	Carbon Catabolite Repression
c-di-AMP	Cyclic-di-Adenosine Monophosphate
c-di-GMP	Cyclic-di-Guanosine Monophosphate
CF	Cystic Fibrosis
CFE	Cell-Free-Expression
CFU	Colony-Forming-Units
cNMP	Cyclic Nucleotide Monophosphate
CREB	cAMP Response Element Binding protein
Crp	cAMP Receptor Binding Protein
CT	Cholera Toxin
CTFR	Cystic Fibrosis Transmembrane Conductance Regulator
DAG	Diacylglycerol
DNA	Deoxyribonucleic acid
DsbC	Thiol:disulfide interchange protein
DTT	Dithiothreitol
ECAR	Extracellular Acidification Rate
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EF	Edema Factor
EPAC	Exchange protein activated by cAMP
EPEC	Enteropathogenic E. coli

ET	Edema Toxin
ETC	Electron Transport Chain
ETZ	Ethoxzolamide
FAD	Flavin adenine dinucleotide
FDR	False-Discovery-Rate
FT	Flow Through
FX cloning	Fragment Exchange cloning
GDP	Gross Domestic Product
GDP	Guanosine Diphosphate
GFAT	Glutamine-Fructose-6-phosphate Aminotransferase
GFP	Green-Fluorescence Protein
GMP	Guanosine Monophosphate
GSH	Glutathione
GTP	Guanosine Triphosphate
HCL	Hydrochloric Acid
HGT	Horizontal Gene Transfer
HIC	Hydrophobic Interaction Chromatography
HIV	Human Immunodeficiency Virus
HRP	Horse-Radish-Peroxidase
HSP-60	Heat-Shock-Protein 60
HSP70	Heat-Shock-Protein 70
ICL	Isocitrate Lysase
IFN	Interferon
IGRA	Interferon Gamma Release Assay
IMP	Imidazopyridines
IPTG	Isopropyl β-d-1-thiogalactopyranoside
K_d	Dissociation Constant
kDa	Kilo-Dalton
LC-MS	Liquid Chromatography - Mass Spectrometry
LT	Labile Toxin
Luria Broth	Luria Broth
MALDI	Matrix -Associated-Laser-Desorption/Ionisation
MBP	Mannose-Binding-Protein
m-DAP	Meso-Diaminopimelate
MES	2-ethanesulfonic acid
MHC-Class II	Major Histocompatibility Complex-Class-II
MIC	Minimum Inhibitory Concentration
MK	Menaquinone
MOA	Mechanism Of Action
MRC	Medical Research Council
MRSA	Methicillin-resistant Staphylococcus aureus

MW	Molecular Weight
NAD	Nicotinamide Adenine Dinucleotide
NCBI	National Centre for Biotechnology Information
NIMR	National Institute for Medical Research
NO	Nitric Oxide
NTA	nickel- nitrilotriacetic acid
NTM	Non-tuberculous mycobacteria
OCR	Oxygen Consumption Rate
OD	Optical Density
ORF	Open-Reading-Frame
ORI	Origin
PA	Protective Antigen
PBS	Phospho-Buffered-Saline
PCR	Polymerase-Chain-Reaction
PDE	Phosphodiesterase
PDL	Poly-D-Lysin
PKA	Protein Kinase A
PKC	Protein Kinase C
PKG	Protein Kinase G
PLC	Phospholipase C
PMF	Proton-Motive-Force
PT	Pertussis Toxin
PTM	Post-Translational-Modification
PTS	Phospho-Transfer-System
Q-TOF	Quad-Time-of-Flight-Mass-Spectrometry
REMA	Resazurin Microtiter Assay
RIN	RNA integrity number
RNA	Ribonucleic acid
RNAP	RNA polymerase
RPM	Rotations Per Minute
RSH	RelA-SpoT homologues
RTX	Repeat-in-Toxin
sdaA	L-serine dehydratase
sdh	Succinate Dehydrogenase
SDS	Sodium-Dodecyl-Sulphate
SDS-PAGE	Sodium-Dodecyl-Sulphate Poly-Acrylamide-Gel-Electrophoresis
SEC	Size Exclusion Chromatography
serC	Phosphoserine aminotransferase
TB	Terrific broth
TB	Tuberculosis
TCA	Tricarboxylic Acid Cycle
TLC	Thin-Layer-Chromatography

TNF	Tumour Necrosis factor
TST	Tuberculin-Skin-Test
TTK	Time-To-Killing
UHPLC	Ultra-High-performance liquid chromatography
UK	United Kingdom
UPEC	Uropathogenic E. coli
USA	United States of America
USP	Universal Stress Protein
UV	Ultraviolet
WHO	World Health Organisation
WT	Wild Type

Chapter 1 - Introduction

1.1 Diseases caused by mycobacteria, their major impact on global health and the dangers of Antimicrobial Resistance/Antimicrobial Tolerance

1.1.1 Tuberculosis

Tuberculosis (TB) is a disease that has plagued mankind since ancient times. Evidence of TB infection has even been found in ancient Egyptian[1] and pre-Columbian Peruvian mummies[2,3]. Tuberculosis is caused by the *Mycobacterium tuberculosis* bacterium (*M. tuberculosis*) and is an obligate human pathogen[4]. *M. tuberculosis* appears to be over 70,000 years old and shows remarkable co-evolution/co-divergence with human mitochondrial DNA[4]. When the genomes of the *M. tuberculosis* isolates and human mitochondrial genomes are phylogenetically plotted, both group into lineages and haplotypes; geographically and with a similar pattern[4]. With no known infectious reservoir besides humans, and the similar evolution patterns to human mitochondrial DNA, this clearly reinforces that *M. tuberculosis* is an obligate human pathogen that has affected humans for centuries[4].

In 2018 there were 10 million new cases of TB and 1.5 million deaths[5] (Figure 1.1). TB has the highest death toll of any infectious disease and can be part of a severe co-infection with HIV (an additional 251,000 deaths in 2018). 67% of new TB cases occurred in India, China, Indonesia, the Philippines, Pakistan, Nigeria, Bangladesh and South Africa[5]. However, TB is also a problem in certain areas of major cities like London[6], parts of Eastern Europe and Russia[7].

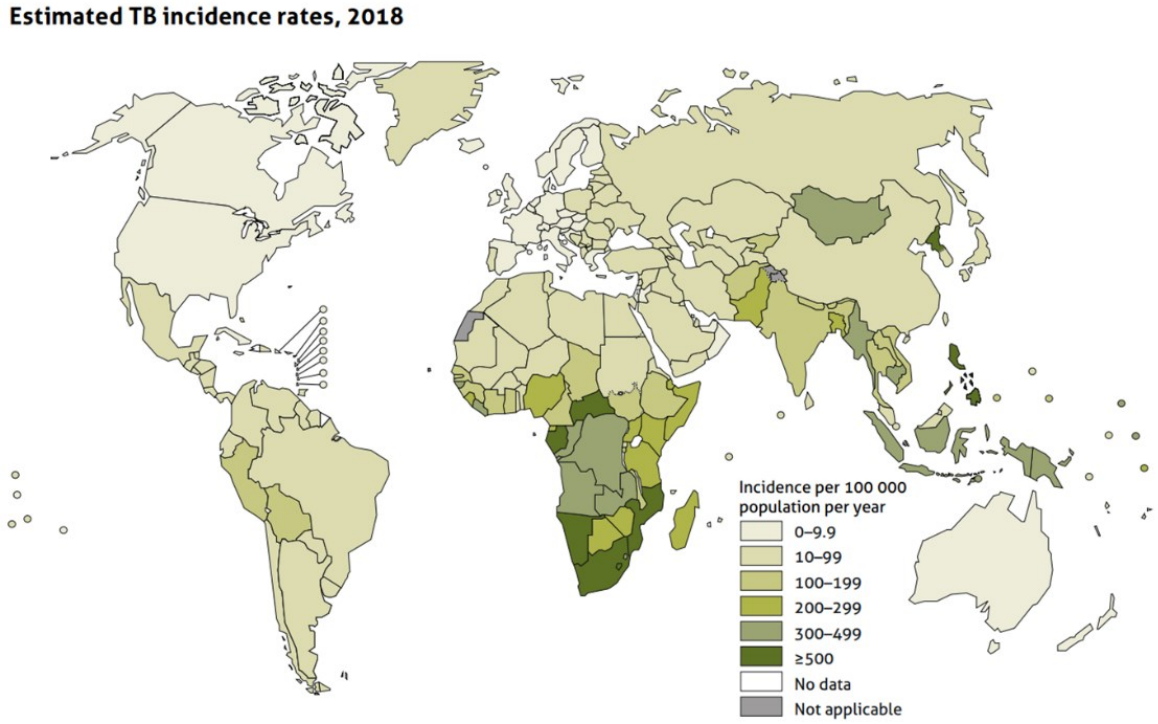


Figure 1.1: Map showing the global incidence of M. tuberculosis per 100,000 of the population, per year. Adapted from WHO Tuberculosis report, 2018.

Traditionally, TB infection is characterised by latent or active disease. Estimates suggest that 2-3 billion people are latently infected with TB. The lifetime risk of developing active disease is 5-15% and this can be much higher in immunocompromised individuals (i.e. with HIV, or malignancies)[5]. A more accurate representation than this dichotomous system is a multistate gradient of infection ranging from subclinical latent infection to active disease[8].

The placement on the spectrum of a TB patient's disease and whether they are contagious or not depends on the resolution of the initial infection. TB has no known environmental reservoir and is an airborne pathogen spread via aerosolised droplets expelled from individuals with active pulmonary disease[5,9]. The bacilli enter the lungs, where they are phagocytosed by alveolar macrophages. Inside the macrophage the bacteria can resist acidification of the phagosome and escape killing[9,10]. In this niche, TB bacilli can replicate and acquire nutrients such as cholesterol[11], iron[12] and nitrogen[13]. Monocytes or dendritic cells that take up the bacilli can travel to a pulmonary lymph node, prime T cells and recruit immune cells to form a granuloma around the site of the bacterial infection. This consists of macrophages with phagocytosed bacilli surrounded by neutrophils, T and B cells. The granuloma can be sterilised and resolved or exist in a balance between containment of the bacterial load and pathology. Alternatively, the infection can be resolved before an adaptive immune response. Various diagnostic tests such as Tuberculin Skin Tests (TST) and Interferon Gamma Release Assays (IGRA), cultures, sputum smears, infectiousness and symptoms all depend on how the infection develops or is resolved[9]. This can also be influenced by coinfection with other pathogens such as Helminths, which modulate the host immune response and supress the Th1 response shown to be effective against TB[14]. The immune response to TB is complex and not fully understood. As a result, effective correlates of immune protection are not definitively known. This poses a significant problem in the development of an efficacious vaccine against TB infection or disease[15].

Currently, the only available vaccine for TB is *Mycobacterium bovis* Bacille Calmette-Guerin (BCG) which shows markedly variable efficacy. This ranges from an overall 0% in Chingleput South India to 80% in the UK[16]. BCG was first administered to humans in 1921[17] but In the ensuing 99 years, the search for a more effective vaccine is still ongoing – with more than 14 vaccines in the development pipeline[18]. Until a new vaccine is found, the variable efficacy of BCG, treatment with antimicrobials and improvements in public health and sanitation are the only weapons against TB. However, resistance to antimicrobials in pathogenic mycobacteria such as *M. tuberculosis* was estimated to account for over 480,000 deaths in 2016[5].

1.2 Antimicrobial resistance and tolerance

The World Health organisation (WHO) and World Economic Forum both recognise antimicrobial resistance (AMR) as one of the most severe global threats we currently face[19,20]. A report for the UK government and Wellcome trust (Tackling Drug-Resistant Infections Globally – O'Neill, 2015) suggests that by 2050, there will be 10 million deaths a year from AMR and the cost to world GDP could be $100 trillion.

AMR is the acquired or developed resistance of pathogens to antimicrobial agents, by modification of the genome or acquisition of antimicrobial resistance genes[21,22]. Effector mechanisms of AMR are shown in Figure 1.2.

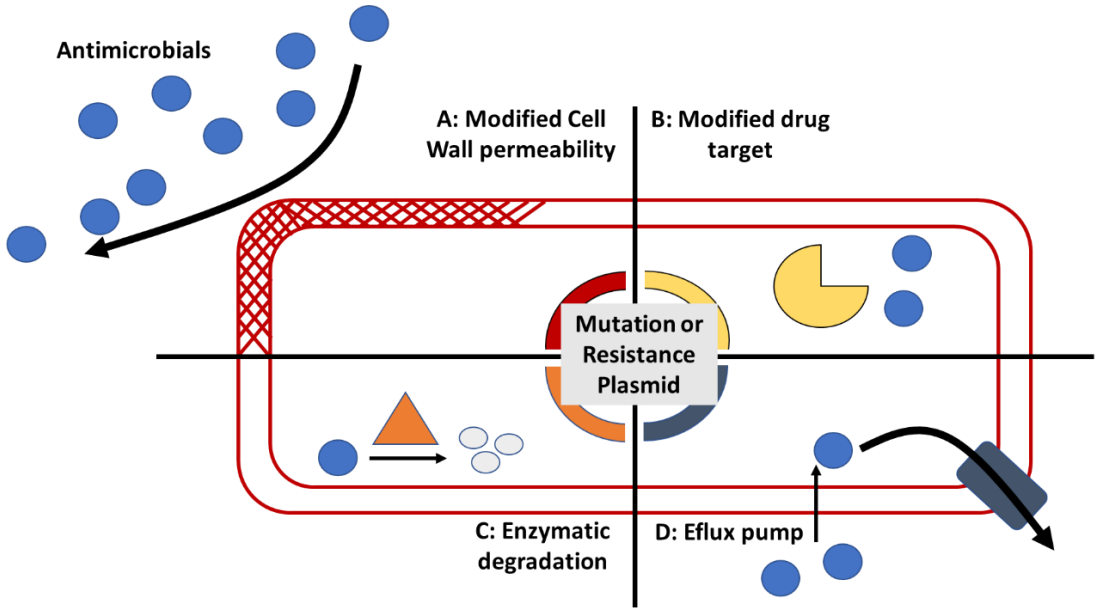


Figure 1.2: The different mechanisms of antimicrobial resistance. A: Bacteria can modify their cell wall composition or downregulate transporters/facilitator proteins that allow antimicrobials entry. B: By changing the structure of the antimicrobials' target, binding or inhibition of function can be prevented. C: Enzymes can degrade or modify antimicrobials to inactivate them. D: Efflux pumps can actively remove antimicrobials that manage to enter the bacteria, decreasing the amount of exposure time and potentially interfering with the antimicrobial mechanism. All of these mechanisms can be the result of mutations in the bacterial chromosome or by acquisition of a mobile genetic element – such as a plasmid.

AMR is best characterised by an increase in the Minimum Inhibitory Concentration (MIC) of an antibiotic that is required to inhibit growth or kill the bacteria (Figure 1.3A). This means that the resistant bacteria can survive exposure to higher levels of the antimicrobial it has developed resistance for. Some bacterial pathogens can transfer genes encoding for resistance to antimicrobials via horizontal/lateral gene transfer (HGT) This is a process utilising plasmids, transposons and bacteriophages to transfer genes or segments of the genome between bacteria[23]. Other pathogens such as *M. tuberculosis* do not appear to use HGT and resistance occurs from mutations in the genome[21].

From an evolutionary perspective, the lack of HGT in *M. tuberculosis* makes sense – it is an intracellular pathogen likely to be isolated from genetically heterogenous bacteria in the host niche. Finally, pathogens such as *Vibrio cholerae* (*V. cholerae*) can be naturally well suited to acquiring DNA from the environment[24]. It is also important to note that there is evidence that AMR pre-dates modern antimicrobials[25]. Ancient bacterial DNA isolated from 30,000-year old Beringian permafrost shows the presence of beta-lactamase and tetracycline resistance genes[25]. In addition, genetic sequences that cluster with the highly specific resistance operons for vancomycin (invariant arrangement of the 3 gene operon *van-H-van-A-van-X*) have also been found. This *vanHAX* operon had been thought to have originated in response to vancomycin treatment in the 1980s and can be found in current day streptomycetes strains – albeit with minor differences in the predicted residues of the Mg^{2+} and ATP binding pockets. These findings were confirmed by expression of the proteins, purification and X-ray crystallography. This indicates that this operon and the encoded proteins, likely predates vancomycin treatment in the 1980s[25].

AMR cannot just be thought of in terms of the isolated pathogens during an infection or under drug treatment – currently prevailing concepts of AMR additionally view it as a culmination of "One health" and "Global health" concepts[26]. "The One health" concept takes in to account that AMR is a problem resulting from health in humans, animals and in the environments in which they intersect. For example, use of antibiotics in the food chain can select for antibiotic resistant pathogens that can then infect humans[26]. Likewise, heavy metals and biocides in the environment can select for antimicrobial resistant bacteria[27]. For example, traditionally volcanic soil dwelling bacteria such as *Ralstonia metallidurans* chromosomally encode genes mediating heavy-metal resistance and that may incidentally provide resistance to antimicrobials. However, strong selection pressures and the prevalence of chemical pollution has led bacteria like *R. metallidurans* to encode these resistance genes on plasmids, that can be spread throughout the bacterial population more easily[28]. Alternatively "antibiotic pollution" in to the environment can alter the make-up of heterogeneous bacterial populations and lead to enrichment of resistance mechanism expressing bacteria[27]. Antimicrobial resistance genes (ARGs) can be spread through the environment as well as from pathogen to pathogen. Unlike chemical contaminants, these ARGs can be concentrated as they move up the food chain due to the high-replication rates of bacteria and other pathogens[27].

The "Global health" concept can be clearly illustrated by the huge network of rapidly interconnected countries and areas of the world. This can enable the spread of pathogens far beyond the environment where they initially developed resistance. International trade, business travel, refugee displacement and even tourism can be driving factors in the spread of antimicrobial resistance[26].

AMR is becoming increasingly prevalent; ranging from nosocomial pathogens in hospitals such as *Klebsiella pneumoniae*[29] or multidrug-resistant *Staphylococcus aureus* (MRSA) strains[30], all the way to

opportunistic pathogens like *Pseudomonas aeruginosa*[31] and even pathogenic *Escherichia Coli* (*E. coli*) strains[32]. They have even evolved resistance over time even to antibiotics of last resort[33].

In contrast to AMR, antimicrobial tolerance (AMT) can be defined as an extension of the period of time that bacteria can survive in the presence of lethal concentrations of an antibiotic, before succumbing to its effects[34]. This decreased susceptibility to antimicrobials is mediated by transient phenotypic adaptions and is non-heritable[34,35]. Tolerant bacteria are still killed by the same dose of antimicrobial as susceptible bacteria, but over a longer time period. For example, the MIC for tolerant bacteria is not increased; just the amount of time the bacteria can survive exposure to this concentration of antibiotic

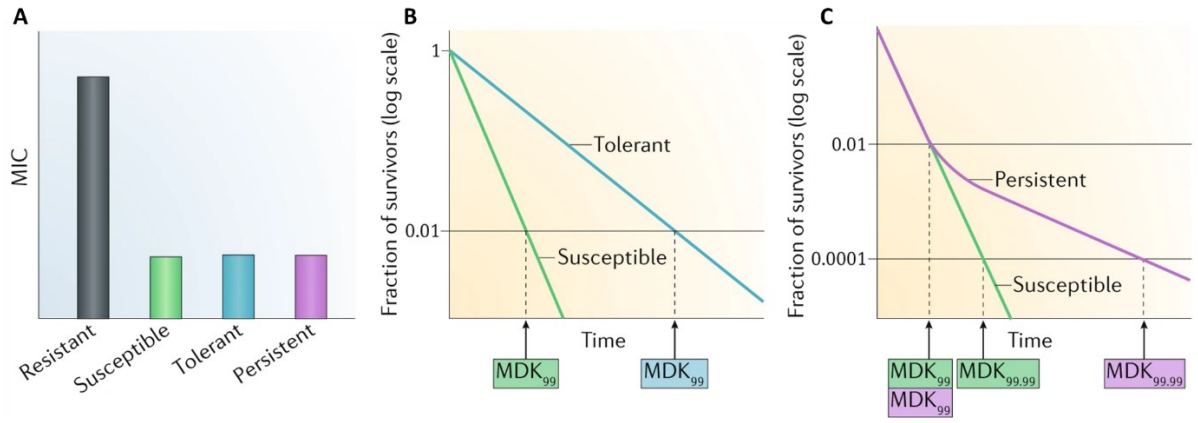


Figure 1.3: The difference between antimicrobial resistance (A), tolerance (B) and persistence (C). Modified from Balaban et al, 2019[35].

AMT (Figure 1.3A) receives far less attention than AMR, and this is because AMT is less well defined clinically[34] due to the transient nature of these and the broad spectrum of tolerance vs persistence. AMT represents just as big a problem and offers some potential solutions. To begin, it is important to distinguish between persistence (Figure 1.3C) and phenotypic antimicrobial tolerance (Figure 1.3B). The best way to illustrate AMT is with an unchanged MIC and the time-to-kill experimental metric, MDK_{99} (Minimal Duration of Killing is the amount of time taken to reduce the numbers of the bacteria in the population by 99%). If colony forming units (CFU) of a bacterial population are plotted over the course of antimicrobial treatment and the bacteria is not genetically resistant to the antibiotic, there will be 3 different kinetics of killing. The 1st kinetic will be the susceptible bacteria in the population (denoted "Susceptible" in Figure 1.3B). These susceptible bacteria will steadily decline over treatment, to 0 if the treatment is sterilising. However, if there are tolerant bacteria in the population, that may have enacted transient mechanisms to be more tolerant to the antimicrobial, there will be a biphasic killing curve (a 2nd kinetic – denoted "Tolerant" in Figure 1.3B). This will also be represented by a linear decline but with a reduced gradient, where it takes longer to sterilise. Finally, there are the persisters (Figure 1.3C), which may only account for a small proportion of the bacterial population, perhaps up to 1%[36,37]. These bacteria can display a variety of phenotypes including: rewiring of their

metabolism[38] slowing of growth[39] and increasing efflux activity towards antibiotics[40]. This phenotypically altered state or ensures the bacteria can survive exposure to the antibiotic for a far greater amount of time. It may even appear that the population is sterilised, if the persisters exist below the limit of detection (the 3rd kinetic – denoted in Figure 1.3C as "Persistent"). Once the treatment is ablated, this population can quickly repopulate[35].

To summarise, AMT is the development of phenotypic adaptions that do not involve heritable genetic changes. These adaptions allow bacteria to survive longer in the presence of antibiotic[34]. In addition, when treated with antimicrobials, there may also be a small proportion of the bacterial population that are "persisters". These bacteria can alter growth[39], metabolism[38] and antibiotic efflux activity[40] thereby making them more tolerant to the killing mechanism of the antimicrobial. Persister populations are often capable of outlasting the duration of the antimicrobial treatment. Once the treatment ends, these persister populations can emerge from their dormant or metabolically slowed state and begin replicating again[35,41].

1.2.1 Mycobacteria and Drug tolerance

The less well-defined end of the tolerance spectrum is phenotypic tolerance that can be observed in many mycobacteria, for example: *Mycobacterium abscessus*, an opportunistic pathogen that increasingly effects Cystic Fibrosis (CF) patients. In the CF lung, the bacteria can form biofilms that increase their tolerance to antimicrobials. This has been shown to be mediated by changes in their outer glycolipid composition in response to infection conditions [42]. Despite being considered opportunistic and distinct from pathogenic tuberculosis mycobacteria, they are capable of infecting the host at any body site, and can cause severe pulmonary disease. Disease can develop even in healthy patients[43]. In addition, *M. abscessus* infections can also display heterogeneous populations with multiple different isolates[44]. Alarmingly, these different isolates can also display differential susceptibility to antimicrobials[44].

Because of the variations between infections and the nature of *M. abscessus* bacteria as rapidly growing (around 4-hours doubling time vs the 24-hour doubling time of *M. tuberculosis*) clinical outcomes are often unfavourable[44,45]. The clinical diseases caused can be highly similar to Tuberculosis infection, and without adequate checks, a non-tuberculosis-mycobacteria (NTM) infection could be treated with a standard TB regimen – with the innate tolerance adaptions providing time for an isolate from the patient to develop resistance[46].

Likewise, *M. tuberculosis* bacterial populations in an infectious setting can be heterogeneous, defined as being genetically identical but with different metabolic states and differing susceptibilities to antimicrobials[47,48]. In mycobacteria, these phenotypic tolerances are mediated by adaptions in bacterial physiology that have been evolved over thousands of years in response to environmental stimuli[47,49,50]. When these adaptions are combined with AMR, *M. tuberculosis* and other disease-causing mycobacteria remain a severe scourge to global health. They can cause severe disease and are responsible for significant morbidity, despite the development of antibiotics[21,51].

Treatment of TB with antimicrobials is a long process and infection with antimicrobial resistant strains can be very problematic. A typical treatment course for non-drug resistant TB consists of 4 antibiotics taken daily for 2 months – Isoniazid, Rifampicin, Pyrazinamide and Ethambutol; followed by 4 months of Isoniazid and Rifampicin treatment[52]. Without Rifampicin and Pyrazinamide (due to resistance), treatment can take 9-12 months[53]. Most of the standard antibiotics used to treat *M. tuberculosis* are effective against continuously growing bacteria but only a few, such as Rifampicin[54] and Pyrazinamide[55] can target intermittently active, or seemingly dormant persister populations respectively (Figure 1.4). This stems from their distinct mechanisms of action – Rifampicin targets RNA polymerase and pyrazinamide is capable of reaching bacteria inside the acidified phagosome of macrophages[55]. Combination therapy is necessary due to the heterogeneous bacterial populations that can be found during *M. tuberculosis* infection[47,48]. A key example of this is during *M. tuberculosis* infection in the

lungs, where improper resolution of the initial infection leads to the deposition of Caseum, the central necrotic material from the *M. tuberculosis* induced lesions[56]. When this niche was extracted from *M. tuberculosis* HN878 infected rabbits, and treated with first and second line TB antimicrobials, the bacteria displayed extreme drug tolerance to several antimicrobials[57]. This was postulated to be due to both phenotypic adaption and the poor pharmacokinetic parameters many drugs display in Caseum[57]. Pyrazinamide has been shown to retain bactericidal activity in this environment, and therefore represents a cornerstone in shortening TB drug treatment regimens[57]. *M. tuberculosis* H37Rv Infection models have investigated some of the physiological differences the bacteria have been shown to display in the infectious environment, that are believed to lead to the increased tolerance. One of these studies found that both transient drug tolerance and permanent drug resistance can be mediated by a shift in bacterial metabolism that feeds trehalose (a cell wall component) into central carbon metabolism when bacteria enter the persister state[38]. Trehalose in the cell wall can be channelled into central carbon metabolism in order to maintain bioenergetics and ATP levels, while bacterial metabolism is otherwise slowed[38].

In addition, seminal work investigating the development of bacterial streptomycin resistance during monotherapy highlighted that over 5 weeks of infection the number of resistant bacteria can increase from 1 in 88,750 CFU to 1 in 267 CFU[58]. As there is currently no credible evidence of HGT in mycobacteria, it is likely that the resistance observed in the study stems from the spontaneous mutation rate – which quite rightly led the author/clinicians to endorse combination therapy[59]. The distinct bacterial populations found during *M. tuberculosis* infection display different metabolic and physiologic properties (Figure 1.4)[48]. For example, persister *M. tuberculosis* bacteria are more tolerant to treatment with many antimicrobials. This can allow time for the bacteria to develop mutations and develop resistance to the antimycobacterial drugs[21]. In such a case, treatment becomes far more problematic, requiring 2nd line drugs which have greater risk of severe side effects[60]. As a result of this, antimicrobial resistance (AMR) is a major concern for TB research.

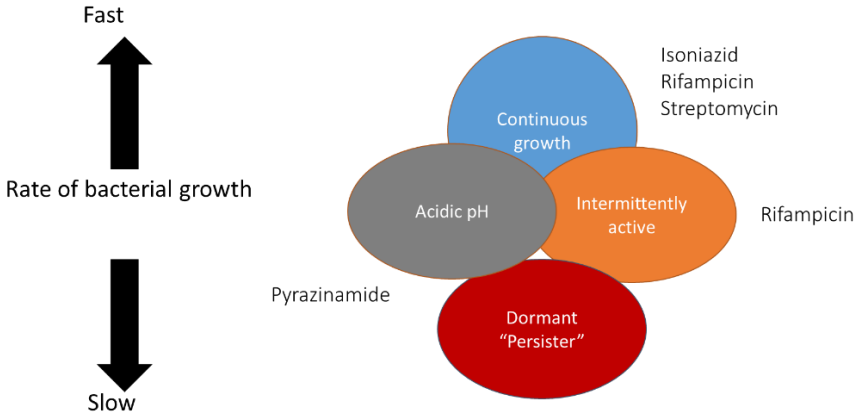


Figure 1.4: Heterogeneous populations of M. tuberculosis bacteria during infection display different metabolic phenotypes. As a result, multiple drugs are required, in order to target each population.

There has been no definitive observation of HGT[21] in *Mycobacterium tuberculosis*. The interaction of bacilli with the treatment, the host response and the robustness with which the treatment regimen is followed are of greater importance. This can be illustrated during treatment with the first line antimicrobials during the initial 2 months. A biphasic decline is observed in the amounts of viable bacteria in sputum. This is representative of at least two distinct populations, one population replicating at a faster rate, that is susceptible to the antibiotic and another population potentially showing altered growth that are more tolerant[39]. The latter population are known as persisters and can be seen in various other bacterial infections[36,52]. TB bacilli are capable of persisting in a dormant non-replicating state that evades the host immune response and antibiotic treatment[61,62]. In addition, there is mounting evidence that there are niches antibiotics cannot always reach e.g. the epicentre of a caseous granuloma[52]. In this resistant state or sheltering environment, TB can mutate and develop resistance or remain dormant until the treatment ends.

The ability of all life forms to detect environmental conditions/stresses, process this information and respond is evolutionarily conserved and essential[63–67]. In the previous examples which involve various pathogens this is especially important for ensuring successful infection and survival[68]. Stresses that pathogens like *M. tuberculosis* would encounter during infection include extreme pH[69], NaCl[47]or potentially antimicrobials[21]. Intracellular, replicating *M. tuberculosis* bacteria have been observed to be less susceptible to antimicrobials; a well characterised example is by increased function of efflux pumps[40]. In this example, although the MIC was unchanged, the bacteria were able to survive exposure to the same dose of antibiotic for longer. Subsequent studies have shown that physiologically relevant levels of NaCl, such as those found within the macrophage (250mM)[47], are capable of reducing bacterial susceptibility to antibiotics *in-vitro*[47]. The physiological changes that were postulated to lead to this increased phenotypic tolerance were also reversed when salt levels were returned to standard media conditions (14.5mM)[47]. The effect of environmental conditions, such as macrophage internalisation, has even been shown to have a greater effect on the development of tolerance than exposure to antimicrobials[49]. Detecting, responding and adapting to environmental conditions is mediated by signalling[70–72].

1.3 Cell signalling and signal transduction

Signalling, is a concept that is integral to all forms of life. At the most simplistic level, signalling can be defined as "communication over distance"[73]. It has been the basis on which life forms evolved to co-operate[63] and even more fundamentally, biological signalling at a cellular level is essential for development and proper function[74]. From early concepts of specific targets and receptive substances (Paul Erhlich – the "magic bullet" in 1862[75] and John Newport Langley in 1905[76]) to phosphorylation/dephosphorylation of enzymes[77] and allosteric regulation[73], signalling now represents a major focus for infectious disease, cancer, developmental and autoimmunity research[78–82]. Our understanding of receptor signalling and signal transduction (the "communication over distance" to elicit a biological response[73]) has evolved over time from reductively simple paradigms e.g. *E. coli* and the lac operon[83], to more complicated systems with multiple signals and spatio-temporal considerations[84]. As such, the earliest example to be elucidated is also one of the best characterised – cyclic nucleotide signalling.

1.4 Cyclic AMP, the first second messenger to be discovered

In 1957, Earl Sutherland and T. W. Rall, published 2 papers. In the first, they identified production of a molecule with the properties of cyclic adenosine monophosphate (cAMP) as a response to hormone stimulation in liver mammalian tissue extracts. In the second paper, they purified and confirmed that this molecule was cyclic AMP (cAMP)[85,86]. By 1971, when Earl Sutherland was awarded the Nobel Prize in medicine for his work on cAMP and the mechanism of hormone action, he had been able to identify that cAMP was an intracellular "second messenger". He had discovered that cAMP levels would increase in response to hormonal activation of an adenylate cyclase (AC – cAMP producing) enzyme, and that the produced cAMP regulated activation of a variety of enzymes in different mammalian tissues e.g. glycogen breakdown in rat muscle tissue[87]. Other work in the field also identified regulation of cAMP levels by production from an AC and "inactivation" (breakdown) by phosphodiesterase enzymes (Figure 1.5)[88,89].

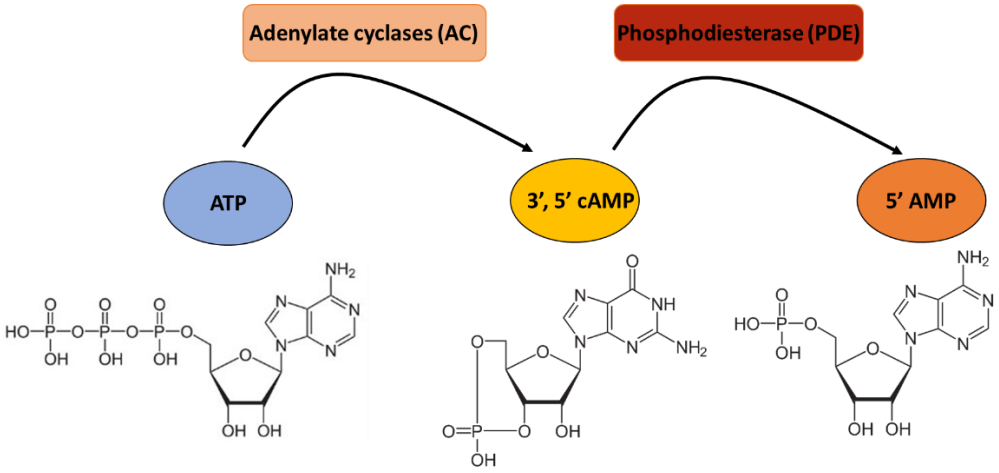


Figure 1.5: cAMP production and degradation. Adenylate cyclase enzymes (AC) produce cAMP from ATP. cAMP can then be hydrolysed by a phosphodiesterase enzyme in the inactive 5' AMP.

1.5 Current understanding of mammalian cAMP signalling

Earl Sutherland's pioneering works would synergise with the discovery that enzyme activity could be modulated by phosphorylation or dephosphorylation[77] and that cAMP mediated many of its effects through the protein serine threonine kinase (PKA)[90,91]. It was also found that cAMP could regulate gene transcription via the cAMP response element binding factor (CREB) in a rat cell line[92].

In order to understand a key aspect of the essentiality of cAMP signalling to the virulence of pathogens, it is important to consider why pathogens would not only evolve to use cAMP for regulating their own physiology, but also why they would evolve virulence mechanisms to exploit cAMP signalling in their hosts. With the widespread presence of cAMP signalling processes throughout mammalian physiology, including a variety of immune functions[93], Sutherland predicted that cAMP may be involved in human disease and could be exploited by pathogens. The example he cited was the bacteria *V. cholerae,* which causes excessive cAMP production in the host via adenylate cyclase activation[89].

Much of the formative work investigating cAMP was undertaken in animal models[87–89], however, our current understanding has confirmed these findings in humans. For example, cAMP has been shown to be involved in neuronal signalling and memory formation at several stages[94], innate and adaptive immunity[95,96], cardiovascular system regulation[97] and many more functions[80].

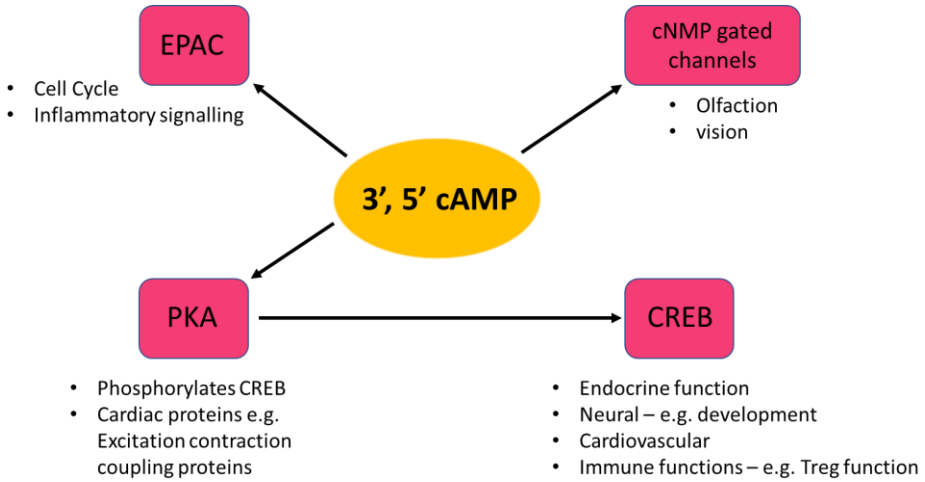


Figure 1.6: Non-exhaustive summary of 3', 5' - cAMP functions in humans. 3', 5' - cAMP mediate various functions by binding to effector proteins that then modulate physiological functions. In humans 3', 5' cAMP is responsible for these functions and more, in almost all known cell types. EPAC – Exchange Protein Activated by Cyclic AMP. PKA – Protein Kinase A. CREB – Cyclic AMP Response Element Binding protein.

The majority of these functions (Figure 1.6) are primarily mediated by PKA and protein de/phosphorylation, with a role for CREB and gene regulation[98]. Additional effectors proteins were later discovered including: Cyclic nucleotide gated ion channels (CNGs) and "exchange proteins directly activated by cAMP" (EPACs). CNGs have been shown to be key components of vision and

olfactory signalling processes[99] and EPACs have been shown to regulate cell cycle and inflammatory signalling in vascular endothelial cells[100]. As such, there has been considerable interest in developing therapies modulating this signalling system. Earl Sutherland postulated that cAMP signalling is involved in a variety of human disorders e.g. the development of type-II diabetes[101] and there is already evidence that therapies targeting this signalling system could be of clinical benefit[98,102].

It is also worth noting that cAMP signalling (Figure 1.7A) is currently the best characterised of 4 distinct second messenger systems identified in humans and animals. Additionally, there is cyclic Guanosine monophosphate (cGMP) (Figure 1.7B), which is produced from guanyl triphosphate by guanylate cyclases and broken down by PDE enzymes. There are also Lipid messengers such a Diacylglycerol (Figure 1.7D), ion messengers (such as calcium or magnesium) (Figure 1.7C) and gaseous messengers such as Nitric oxide, NO_2 (Figure 1.7B)[98,103].

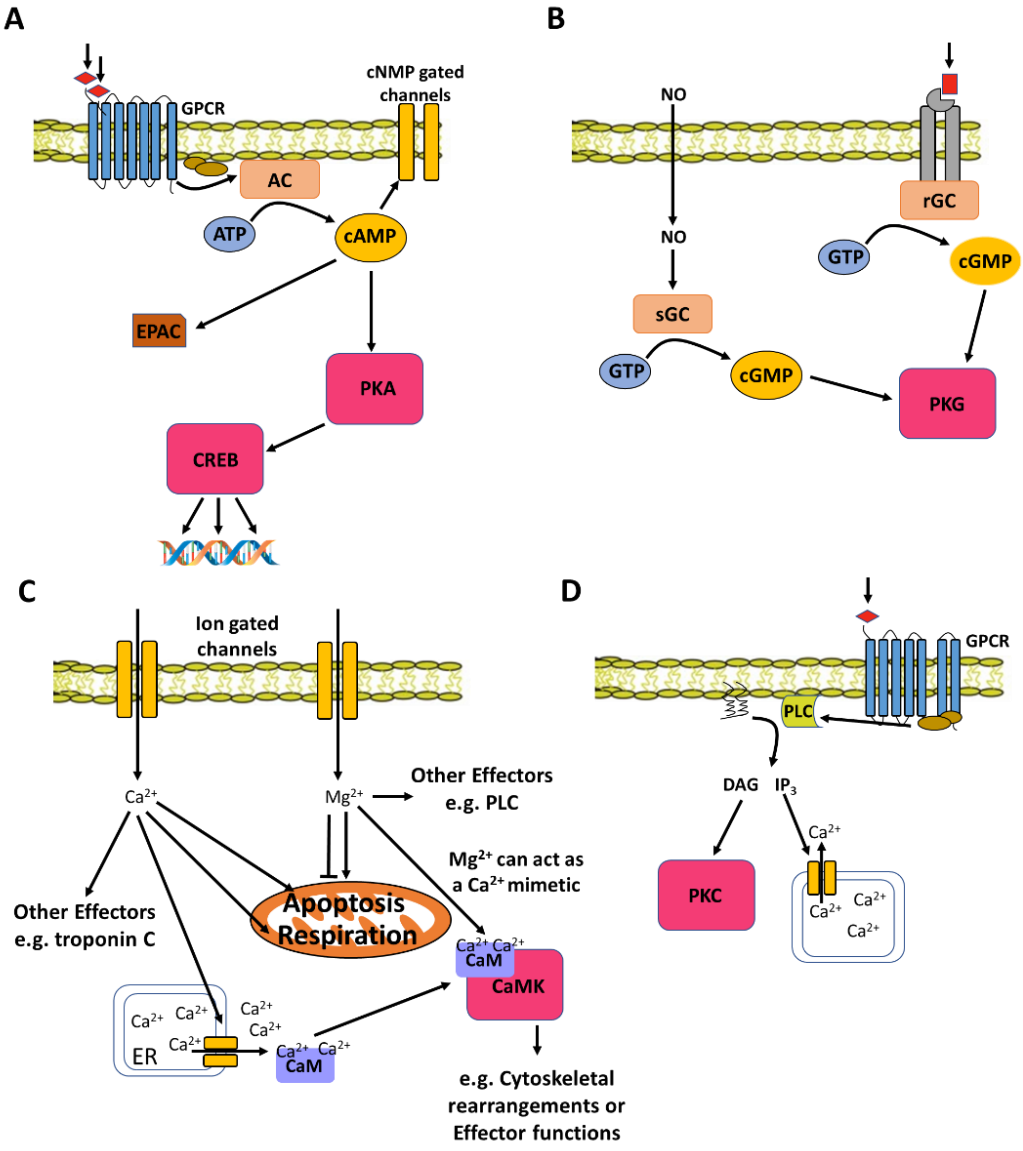


Figure 1.7: Simplified summary of the different second messenger systems in mammals. A: Typical cAMP messenger function. Exogenous stimuli (such as a hormone) binds and activates G-protein coupled receptors (blue). The activated receptor is attached to small G protein subunits (brown) which stimulate membrane bound adenylate cyclase enzymes to produce cAMP from ATP. cAMP then binds effector proteins such as Exchange protein directly activated by cAMP (EPAC), PKA or cAMP responsive exchange binding protein (CREB). These effectors then mediate gene expression or protein conformational changes that lead to physiological adaption. B: Typical cGMP messenger signalling. Exogenous Nitric oxide (NO) can diffuse into the cell and activate soluble guanylate cyclases to produce cGMP. Alternatively, a stimulus can activate cell membrane bound receptor guanylate cyclases to produce cGMP. cGMP can then bind to effector proteins like PKG to mediate physiological adaptions. C: Exogenous Ions or other stimuli can activate cell membrane bound ion gated channels, causing and influx of the specific Ion (Ca^{2+} or Mg^{2+} in this example). These ions can then bind effector proteins in the cell e.g. calmodulin (CaM – purple), troponin C, intracellular ion gated channels harbouring ion stores or mediating changes in mitochondrial function (induce or prevent apoptosis or increase respiration.) D: An example of Lipid second messenger signalling. An exogenous stimulus activates GPCRs which in turn stimulate membrane bound Phospholipase C (PLC). PLC then cleaves Phosphatidylinositol 4,5-bisphosphate (PIP_2) into diacylglycerol (DAG) and inositol triphosphate (IP_3). PIP_2 and IP_3 messengers then activate effector proteins such as PKC and ion gated channels.

A common theme of the cyclic nucleotide and lipid signalling systems is that an initial stimulus e.g. a hormone, growth factor or lipid binds to a receptor on the outside of the cell. This then leads to production of a second messenger or mobilisation of intracellular or extracellular signalling ions to eventually activate a protein kinase (PKA and PKG or PKC respectively, in these examples). These effector proteins can then mediate the required function e.g. by activation of transcription factors or by modulation of metabolic enzymes such as glucose-6-phosphatase[101]. In the case of ion messenger systems (Figure 1.7C), exogenous ion primary messengers (calcium or Magnesium) can activate cell membrane bound voltage gated ion channels, that then source calcium signalling ions from intracellular stores or from extracellular sources. These signalling ions can then diffuse through the cell and directly modulate activity of enzymes or indirectly by binding downstream to intermediate effectors such as calmodulin[98,103]. Finally, in the case of nitrous oxide (Figure 1.7B), exogenous NO_2 can diffuse into cells, activate soluble guanylate cyclases to produce cGMP from GTP, which then binds to the effector protein PKG[104] The downstream mechanism(s) by which PKG mediates the effects of NO_2 are not well understood[105].

In summary, cAMP signalling regulates a broad variety of functions in human physiology[98] The effects of cAMP are mediated by cAMP binding effector proteins. For example, these proteins can be transcription factors in the case of CREB, or protein kinases involved in PTM – in the case of PKA[98]. It is also worth noting that in human physiology, cAMP is not the only cyclic nucleotide messenger. For example, cGMP also mediates specialised roles, with similar machinery to cAMP signalling (e.g. Protein Kinase G)[103].

The involvement of cAMP in immune functions provides the opportunity for manipulation by pathogens[65,89,93]. Several examples of how pathogens manipulate host signalling will be reviewed subsequently, but first, alternate cyclic nucleotide messengers identified in bacteria will be briefly reviewed.

1.6 Pathogens, cyclic nucleotide signalling and virulence

1.6.1 Cyclic nucleotides involved in pathogen signaling and virulence

Cyclic nucleotide signalling and cAMP in particular, is the most broadly conserved signalling system across many different taxa of life[65]. Cyclic AMP signalling has been found in diverse pathogens such as bacteria, fungi and parasites, where it often underpins processes ranging from virulence mechanisms to regulation of physiology [65,66,106]. For example, this can be by subversion of host signalling or immune responses[107], as well as by regulation of the pathogen's own physiology[68]. However, since the discovery of cAMP, other cyclic nucleotides have been discovered in bacteria. Briefly, these alternate cyclic nucleotide messengers will be reviewed.

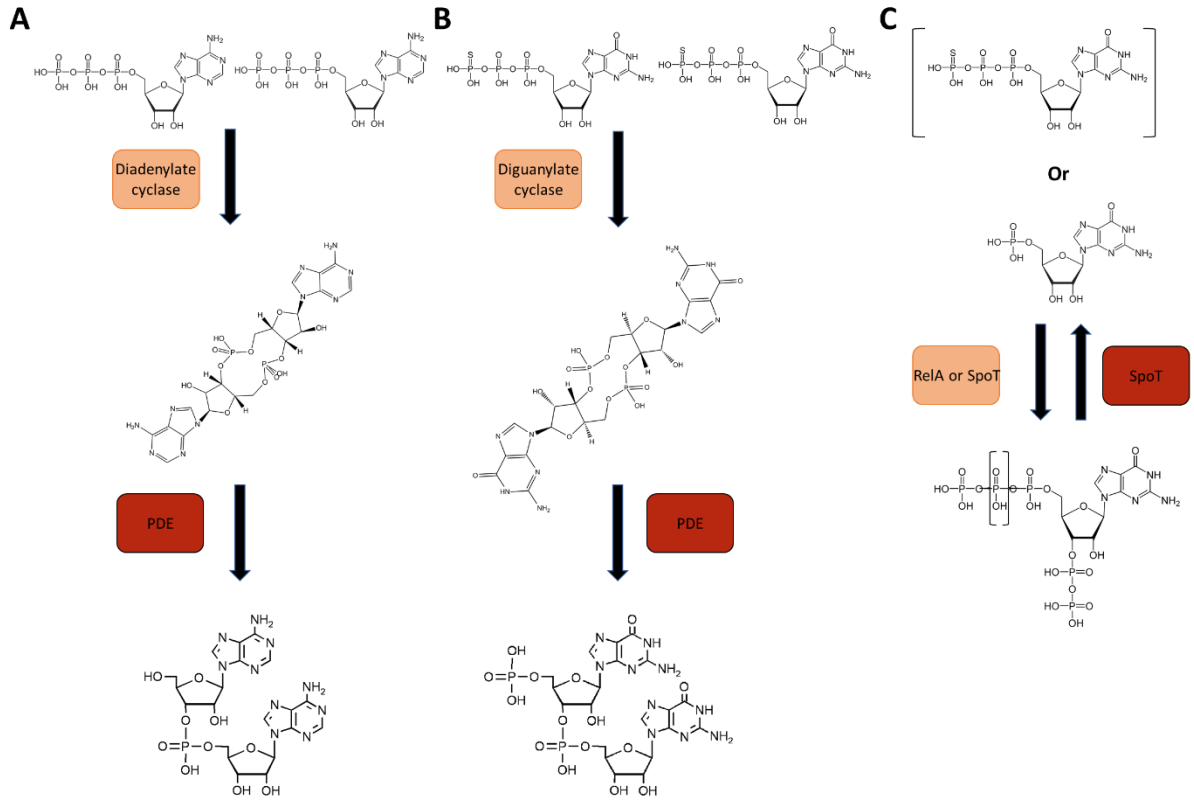


Figure 1.8: Summary of the production and degradation of alternate cyclic nucleotide messengers. A: Cyclic-di-AMP is produced by diadenylate cyclase enzymes from 2 molecules of ATP. Subsequently, c-di-AMP can be linearized by phosphodiesterase enzymes (PDE) to 5'-phosphadenylate-adenosine (pApA). B: Cyclic-di-GMP is produced by diguanylate cyclases from 2 molecules of GTP. Like c-di-AMP, it can be degraded by PDE enzymes via linearization to 5'-phosphguanylyl-adenosine (pGpG). C: pppGpp is produced by addition of pyrophosphate to GTP (substrate and additional phosphate group of the product in brackets). ppGpp is produced by addition of pyrophosphate to GDP). Both molecules can be produced by RelA or SpoT enzymes. SpoT enzymes can also possess PDE activity and hydrolyse (p)ppGpp into molecules of GTP or GDP respectively.

1.6.2 Cyclic Di-nucleotides

In addition to cAMP, several other cyclic nucleotide messengers have since been characterised in bacteria. These include the cyclic-di-nucleotides e.g. cyclic-di-adenosine monophosphate (c-di-AMP) (Figure 1.8A)[108] and cyclic-di-guanosine monophosphate (c-di-GMP)[109] (Figure 1.8B). These are formed of 2 units of ATP or GTP respectively[110,111]. The principles underlying the signalling processes of the cyclic dinucleotides fits the classical second messenger paradigm, similar to cAMP. Cyclic-di-guanosine monophosphate is produced in response to environmental stimuli and can then bind to effector proteins. These effector proteins can then mediate their activity by binding to a genomic or protein target. In addition, levels of the cyclic nucleotide are regulated by specific PDE enzymes which linearize c-di-GMP into 5′-phosphoguanylyl-(3′-5′)-guanosine (pGpG) (Figure 1.8B), before additional non-specific PDE enzymes split this molecule into 2 molecules of GMP[109,110] Alternatively, a distinct family of PDE enzymes (HD-GYP) has been discovered which mediates this breakdown in 1 step[112].

Cyclic-di-GMP is used by both Gram-positive and Gram-negative bacteria to regulate production and secretion of various biofilm components, such as adhesins or cellulose[110]. Additionally, an early seminal report suggests that c-di-GMP is involved in regulation of cell cycle, via regulation of proteolysis of a cell cycle regulator in *Caulobacter crescentus*[113]. In addition to positive regulatory effects of biofilm formation processes, c-di-GMP also negatively regulates virulence and motility[110]– likely because these processes could contraindicate biofilm formation. The central role of c-di-GMP in biofilm formation in numerous bacteria[109,110,114] may suggest it is an example of a second messenger specialised to manage biofilm formation. This is apparent when the levels of c-di-GMP are compared in planktonic (non-biofilm forming bacterial cells) vs biofilm forming *P. aeruginosa* bacteria (30pmol vs 75-110pmol[114]), the essentiality of c-di-GMP for effective biofilm formation and the requirement of c-di-GMP PDEs for transition out of a biofilm (dispersion)[115]. Due to the negative regulatory activities of c-di-GMP on non-biofilm forming processes, the interplay of this messenger with other cyclic nucleotides such as cAMP can be highly important e.g. in *P. aeruginosa* infection[116,117], and this will be elaborated upon later.

Cyclic-di-AMP synthesis, effector mechanism and degradation are similar to c-di-GMPs (Figure 1.8A and 1.8B), however with formation from, and eventual degradation into 2 molecules of ATP, via 5′-phosphoguanylyl-(3′-5′)-adenosine (pApA) (Figure 1.8A)[108]. In addition, it has been found to be essential for *in-vitro* growth in all firmicutes were it is present, having roles in cell wall construction, membrane homeostasis, DNA repair and regulation of potassium ion channels[108,111]. A notable example can be seen with the pathogen *Staphylococcus aureus* (*S. aureus*) – where c-di-AMP has been shown to be essential for growth in rich media and for osmotic regulation[118,119]. Excessive accumulation of this messenger has been shown to be toxic, while low levels can lead to non-viability due to interference with the aforementioned roles[108,111]. In Gram-positive bacteria, this is often postulated to be because of the essentiality of peptidoglycan cell wall maintenance. This is due to the circumstantial evidence that

a variety of *Bacillus* phenotypes caused by deleteriously low c-di-AMP levels can be ablated by the addition of magnesium to the culture[108]. Addition of magnesium has previously been seen to rescue mutants with defects in cell wall synthesis[108]. However, broth culture of the Gram-negative pathogen *Mycoplasma pneumoniae* also displays a requirement for c-di-AMP, despite lacking peptidoglycan[120]. This also reinforces the point that *in-vivo*, the essentiality and role of c-di-AMP is called in to question and is still poorly understood[108,111,120].

Both c-di-GMP and c-di-AMP signalling are not currently thought to be present in humans and these messengers are highly immunogenic, activating cytosolic immune surveillance mechanisms in host cells when secreted[121,122]. For example, in a *Listeria monocytogenes* macrophage infection model, c-di-AMP was found to be secreted into the macrophage cytosol and induced a type-I interferon immune response[122]. This is likely a host evolutionary adaption to combat intracellular bacterial infection[107]. Taken together, this implies that more work is required to understand the importance of these specialised messengers *in-vivo*.

1.6.3 Alarmones

Other noteworthy second messengers found in bacteria and plants are the alarmones – e.g. pppGpp and ppGpp[123,124](Figure 1.8C). These messengers have been shown to be involved in nutrient starvation and the stringent response[125]. The stringent response is a widely conserved bacterial stress adaption, that leads to global changes in gene expression. These changes in gene expression patterns can be in response to many different environmental stimuli, to support adaption and provide resilience against the stress e.g. by reducing non-essential metabolomic processes such as macromolecule production or by slowing growth[126,127]. Stresses, such as nutrient starvation, lead to the transient accumulation of these alarmones in bacterial cells. This can be mediated by 2 proteins – RelA and SpoT. The homologues of these proteins in different plants and bacteria form the RSH superfamily (RelA-SpoT homologues). These enzymes produce ppGpp and pppGpp by transferring the pyrophosphate from an ATP molecule, to a GDP or GTP molecule respectively[49]. While RelA enzymes across bacteria tend to possess just ppGpp synthesising activity, SpoT enzymes tend to possess both ppGpp synthesising activity and ppGpp hydrolysing activity. Degradation of (p)ppGpp and ppGpp leads back to GTP and GDP, respectively[124,127]. The exact mechanism as to how ppGpp prolifically activates expression of multiple genes is not entirely understood, however, recently progress has been made. Most of the understanding of the mechanisms of the alarmones and structures of the proteins involved comes from Gram negative bacteria, such as *E. coli*. A recent study of ppGpp in response to nutrient starvation of *E. coli* showed that both *in-vitro* and *in-vivo,* binding of ppGpp to 2 sites on the RNA polymerase enzyme (RNAP), mediates the gene transcription effects. However, binding to the first discovered site is dispensable[128]. The sites are 60Å apart and the most recently discovered binding site is at the interface of RNAP with a transcription factor DksA. Strains with mutations preventing this binding showed severely impaired growth during nutrient limitation when compared to WT strains[128]. The levels of ppGpp can be regulated at the production and hydrolysis stages, either by regulation of RSH gene expression or by regulation of protein activity (SpoT synthesis vs hydrolysis[127]). The RSH enzymes of different organisms have different preferences for substrate (e.g. GDP for *E. coli* RelA vs GTP for *E. coli* SpoT or MTB's RelA homologue) and therefore favour either ppGpp or pppGpp production. It has been postulated that the sensitivities of effectors in different organisms to these alarmones varies and this could explain the differences[127]. Taken together, although the (p)ppGpp system is conserved across many bacteria and plants, it seems specialised to nutrient limitation or the stress response. The complexity of this system also requires more characterisation, especially in Gram-positive bacteria.

The common theme of these nucleotide messengers is that an environmental stimulus leads to initiation of signalling. This signal is then transduced and converted into an appropriate response. An effective response involves the generation of a signalling nucleotide that binds effectors, that can then mediate the response to the initial stimulus (Figure 1.9 – the example of cAMP). The ability to efficiently

respond to stressors in the environment during infection is essential for the virulence of several pathogens. For example, when (p)ppGpp machinery is knocked out of a *Francisella tularensis* strain, the bacteria are sensitive to oxidative stress and are attenuated in both macrophage and mice infection[129].

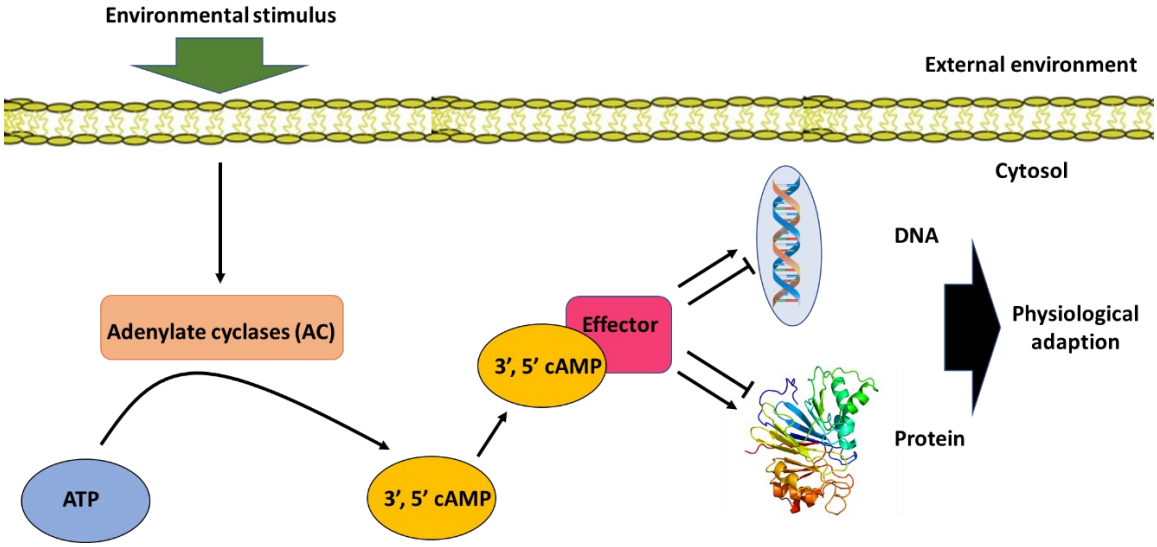


Figure 1.9: Environmental stimuli can lead to activation of adenylate cyclase enzymes (ACs). ACs then produce 3', 5' - cAMP from ATP. 3'5'- cAMP can then bind to effector proteins and mediate changes in gene expression or modulate protein activity to yield physiological adaptions.

While c-di-GMP, c-di-AMP and (p)ppGpp all have important and specialised roles, cAMP signalling is widely conserved in pathogens and mediates a greater variety of important roles. These roles can be in gene regulation, post-translational modification, lifestyle switching, environmental adaption, virulence and host immune evasion[65,66]. The best understood roles of cAMP are for host invasion, gene regulation and virulence. This will be illustrated subsequently with some interesting examples of the uses of cAMP in various disease-causing bacteria.

1.7 Cyclic AMP toxins

Bacterial toxins are virulence factors injected into host cells by a variety of bacterial secretion systems that subvert or take over control of key host processes, thereby creating favourable conditions for the pathogen[130]. Several bacterial pathogens use specialised toxins designed to intoxicate host cells with cAMP. By intoxicating host cells with cAMP, the pathogen can modulate the infectious environment (Figure 1.10), subvert the immune response and/or enable optimal conditions for colonisation – all these processes are required for successful infection of the host[131]. Intoxication of the host cells can either be mediated by activation of host adenylate cyclases or by direct production from a pathogenic adenylate cyclase domain in the toxin[65,131]. Due to the broad range of host immunological processes regulated by cAMP[95,96], it is not surprising that many pathogens have evolved to exploit this system.

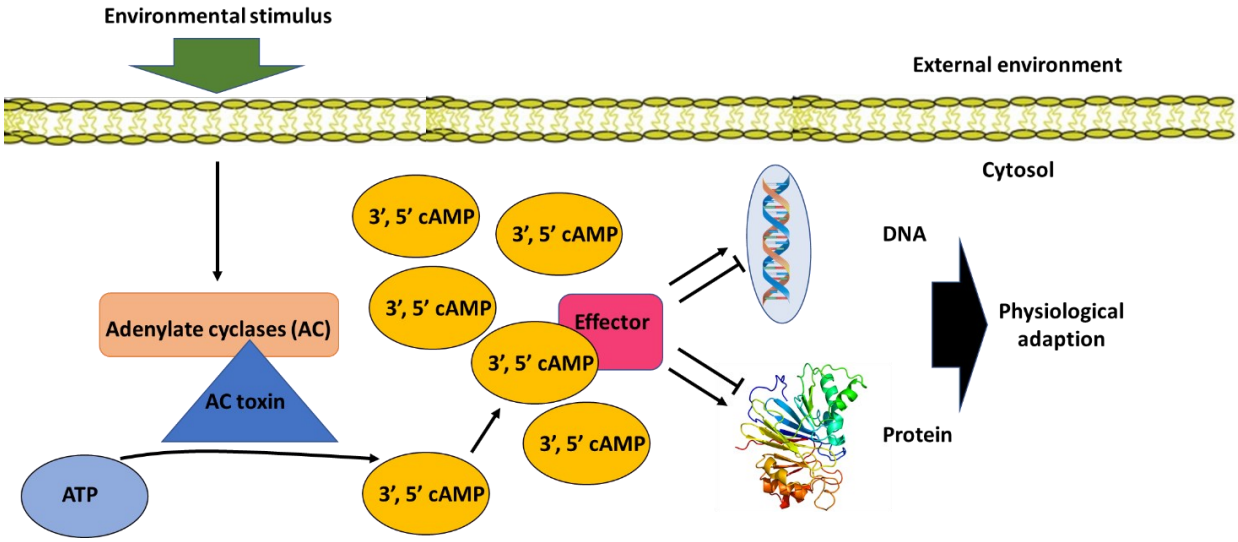


Figure 1.10: Pathogens can subvert host cAMP signalling in the host cell cytosol by perpetually activating host AC enzymes to produce cAMP, or by producing cAMP themselves. This dysregulated signalling at gene expression or protein level can lead to altered host responses, and this can be beneficial for the pathogen.

1.7.1 Cholera toxin of *Vibrio cholerae*

There are many examples of pathogens using cAMP as part of their virulence activities. An early known example is the Gram-negative bacteria *V. cholerae*, as Earl Sutherland eluded to in his Nobel lecture[89]. Cholera is a severe secretory diarrhoeal disease caused by certain serotypes of *V. cholerae* bacteria. Annually it is responsible for 95,000 deaths worldwide and is linked to poor sanitation or unclean drinking water. With 60% of cases occurring in sub-Saharan Africa, wars and regional instability can prevent adequate cholera control[132]. In addition, global warming is potentially causing increased cholera prevalence[133]. It occupies a harsh aquatic environment with varying salinity and protozoan or bacteriophage predation. As a result, humans come into contact with this pathogen via drinking water sources. Pathogenic *V. cholerae* serotypes colonise the surface of intestinal epithelial cells in the host with an essential chitin binding protein (GbpA)[134]. Once bound to the host epithelial cells, *V. cholerae* expresses and secretes cholera toxin (CT), a hexameric subunit protein, where 5 subunits mediate binding to host cells and entry of the 6th subunit into the cytoplasm. Once in the cytoplasm, the CT subunit is allosterically activated by a host ADP-ribosylation factor, which then augments this activity of the toxin[135]. The toxin augmented with the host factor then mediates ADP-ribosylation of a G protein in the membrane, allowing GTP to bind and enabling the G protein to activate the membrane-associated adenylate cyclase enzyme AC6 (Figure 1.11A). This AC is then constitutively active, producing cAMP from host-derived ATP. These increased cAMP levels lead to activation of host PKA, which results in phosphorylation of the cystic fibrosis transmembrane conductance regulator (CFTR). CFTR then allows chloride and bicarbonate secretion into the gut lumen and results in major passive efflux of water – the resulting dehydration is the cause of pathology in Cholera disease[135–137]. This original evolutionary purpose of this mechanism is not entirely understood, but it is essential for successful colonisation in humans. It may potentially be an adaption to fluctuating salinity levels in the pathogen's aquatic environment[137]. This is a key example of a bacterial pathogen exploiting host cell signalling with a bacterial toxin essential for its' successful virulence.

Pathogens can also encode cAMP producing adenylate cyclase toxins that directly increase host cAMP levels. By increasing cAMP to pathogenic levels, the pathogen can subvert the host immune response. This is the case for 3 common bacterial pathogens: *Bordetella pertussis*, *Bacillus anthracis* and *P. aeruginosa*[65,138].

1.7.2 Pertussis toxin and Adenylate cyclase toxin of *Bordetella pertussis*

Bordetella pertussis is a Gram-negative, obligate human pathogen and the causative agent of the highly infectious, chronic respiratory disease "Whooping cough"[139]. While there has been a vaccine since the 1940s, antibiotic resistance and reduced adherence to vaccine programmes mean *B. pertussis* is still a major threat to infants and young children – in which the disease is severe and life threatening[139,140].

The bacteria produce 3 toxins, 2 of which are central to their virulence – Pertussis toxin (PT) and the Adenylate cyclase toxin (CyaA or ACT)[139–141]. PT is transported across the host cell membrane by a type IV secretion system and once inside, catalyses the transfer of a host ADP-ribosylation factor from hydrolysed NAD+ to a host regulatory G-protein. This prevents the inhibition of a host membrane bound AC and causes increase in cAMP levels (Figure 1.11A) – similar to cholera toxin. Additionally, CyaA consists of an adenylate cyclase domain connected to a haemolytic (Hly) toxin domain. Instead of direct secretion into the host cell, Hly binds to the outside and forms cation-selective toxin pores that enable entry of the AC domain to the host cell cytosol (Hly also conducts calcium ions in to, and Potassium ions out of, host cells – leading to dose dependant cytotoxicity[142]). Once inside the host cell, the AC domain is activated by calmodulin (Figure 1.11A), a host Ca^{2+} binding effector protein[139,141,143]. Activation of CyaA within the host cell leads to production of cAMP from host derived ATP and a drastic increase in cAMP levels. CyaA has been shown to be essential for colonisation of the host and impairs T-cell activation/recruitment along with inhibition of T cell receptor – chemokine signalling[144,145]. The effect of PT itself is believed to be responsible for leucocytosis, severity of disease and poor outcome in infants [140]. It has also been shown to have a suppressive effect on MHC-class II presentation on monocytes and dendritic cell surface markers[140,146,147].

1.7.3 Edema toxin of *Bacillus anthracis*

The virulence mechanism of a pathogen releasing an adenylate cyclase toxin into host cells to increase host cAMP levels is also seen in *Bacillus anthracis* Edema Toxin (ET). *Bacillus anthracis* is a Gram-positive spore forming bacterium that is the primary causative agent of Anthrax[148,149]. Anthrax disease can manifest most commonly as a cutaneous infection and rarely as a gastrointestinal or severe pulmonary infection (depending on exposure method)[150]. Normally it infects animals such as ruminants, and as such, farming and consuming infected animals is the most common exposure for humans[149].

Due to a combination of the robustness of *B. antrhracis* spores, a tripartite toxin group and a bacterial capsule with low immunogenicity[149], inhalation mediated systemic infection can be severe – as seen in an instance of bioterrorism in 2001[151]. High levels of bacteria in the systemic circulation leads to immune cell death and sepsis[152]. Edema toxin makes up 1/3 of this toxin group and consists of highly active calmodulin activated adenylate cyclase enzyme subunits (Edema factor – EF), coupled to an adhesion subunit (protective antigen – PA). PA mediates entry into host cells by binding mainly to capillary morphogenesis genes 2 product (CMG2), or to the tumour endothelial marker 8 (TEM8). These integrins are found on the outside of numerous mammalian host cells[149]. When 3 EF subunits are bound to 1 PA subunit, the complex binds the host cell integrins and enters the host cell by endocytosis[153]. For EF to escape the endosome and enter the cytoplasm, PA molecules undergo a conformational change in response to the acidic environment of the late endosome, thereby forming pores into the cytoplasm[149].

Once inside the cytoplasm, EF is activated by host calmodulin (Figure 1.11A)[153]. This leads to production of high levels of cAMP inside the host cell. As seen in previous examples, this leads to pathogenic levels of cAMP. In *B. anthracis* infection, ET has been shown to inhibit efferocytosis – the process by which phagocytic cells engulf dead or dying immune cells[152]. Edema toxin and ACT (from *B. pertussis*), both raise host cell cAMP levels and cause suppression of immune cell activation[144,152]. However, CyaA has a direct suppressive effect on the immunological synapse, not seen in the *B. anthracis* ET[145], despite both toxins raising cAMP to similar levels. This could be an intriguing example of the tight spatiotemporal regulation observed in cAMP signalling[145]. Because of their differential entry mechanisms into the host cell, there is differential localisation of the AC toxins and therefore different locations of cAMP increase. CyaA maintains location near the cell membrane, whereas ET occupies a perinuclear position, after exiting the late endosome. ET cannot elicit the same effect on the immune synapse as CyaA[145]. In addition to localisation differences, these toxins also have structural differences such as the location of the calmodulin binding site[154,155].

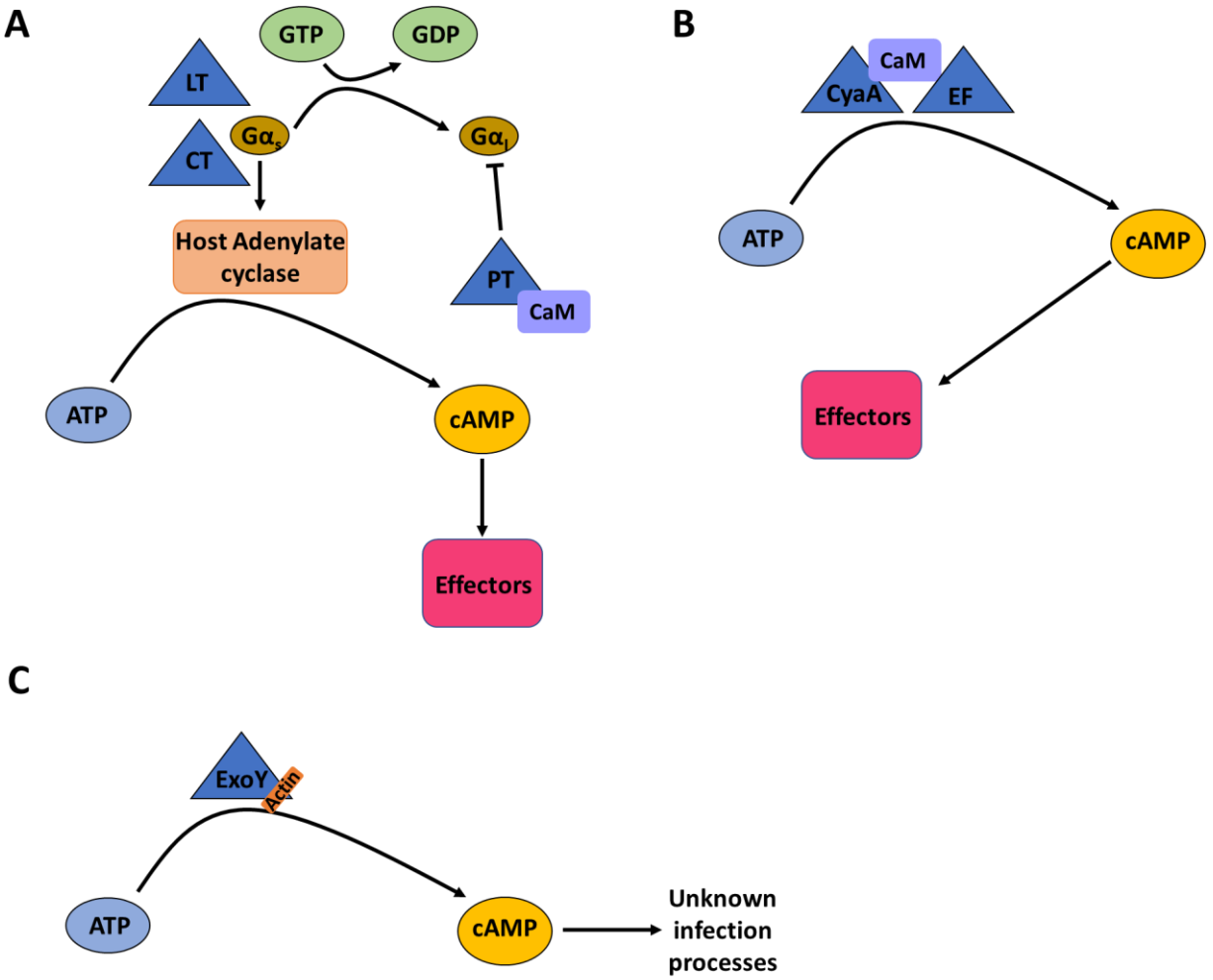


Figure 1.11: Summary of different bacterial AC toxins. A: Labile toxin (LT), Cholera toxin (CT) alter host G protein function and this leads to erroneous cAMP production. B: CyaA and Edema factor produce cAMP directly from host ATP, though at different cytosolic locations. The effect of these toxins, Labile toxin or Cholera toxin is mediated by cAMP binding to host Effector proteins. C: ExoY can produce cAMP directly from host ATP. It is activated by host actin, but the purpose during infection is currently unknown.

1.7.4 ExoY toxin of Pseudomonas aeruginosa

Another secreted AC toxin usually grouped with ET and PT is ExoY from *P. aeruginosa*. *P. aeruginosa* is a Gram-negative, environmental bacterium and opportunistic pathogen – most commonly found in hospitals and the lungs of cystic fibrosis sufferers[156,157]. It displays impressive adaptive capabilities with a broad repertoire of virulence mechanisms and versatile metabolic pathways[158]. Regulation of gene expression, metabolism and virulence is mediated by a complicated signal network that can quickly respond to changing conditions within the host and mediate escape from host immune responses[156]. Of the many virulence determinants that *P. aeruginosa* isolates can accrue and utilise, the gene encoding the AC toxin ExoY, has been shown to present in 90% of clinical isolates, in some studies[159].

ExoY is delivered directly into the host cell cytosol via a type-III secretion system and harbours a promiscuous AC domain. Historically it was thought to produce cAMP primarily, but it is now clear that it can also produce cUMP, cGMP and cCMP. Reports of *in-vivo* infection even suggest that the poorly described cUMP may be the preferred nucleotide produced *in-vivo*. However, the same work shows that the toxin can increase the activity of all 4 nucleotides within host cells[160]. Interestingly, ExoY is completely inactive in the bacteria and is not activated by calmodulin in host cells (unlike ET and PT). Instead it has been found to be activated by localisation and interaction with host F-actin (Figure 1.11B) , a key component of the actin cytoskeleton in all mammalian cells[161]. The role of this promiscuous cyclase, ExoY, during infection still remains to be elucidated. While expression of ExoY in host cells has been found to be toxic and mediate significant pathology and virulence during infection models[162], this may be ascribed to the non-physiological overexpression of ExoY commonly used in these studies[158]. It is clear that binding of ExoY to host F-actin, and subsequent activation, leads to altered turnover/increased absolute levels of actin in host cells[161]. However, the effects of this increase in turnover and actin levels during *P. aeruginosa* infection is unclear.

1.7. 5 Labile toxin of *E. coli*

Cyclic AMP has also been shown to be involved in virulence of enteropathogenic *E. coli* strains (EPEC). These strains of *E. coli* colonise the intestine of the host and can induce diarrhoea (1 potential cause of "travellers' diarrhoea")[163]. Although not normally considered in traditional bacterial toxin paradigms, EPEC *E. coli* strains encode a labile-toxin (LT) (Figure 1.11A). This toxin mediates the binding of *E. coli* to the intestinal epithelial cells, and the subsequent release of ions or fluid leading to diarrhoea. This process was found to require manipulation of host cell cAMP levels by the toxin[163,164]. LT has significant immunogenicity and altered versions of LT, that lack the ability to alter host cAMP levels, are is being considered for the development of a vaccine[164]. Not only does this show a diversity in the roles of cAMP in *E. coli*, but also shows that the human immune system has evolved to respond to the threat of cAMP manipulation.

1.8 cAMP and gene regulation in different pathogenic bacteria

In many organisms, cAMP is involved in gene regulation[98,106,165]. In pathogenic bacteria, this can be especially important due to the variety of environments/stresses bacteria must adapt to and overcome.

1.8.1 The original cAMP bacterial gene regulation paradigm and current understanding in *E. coli*

Cyclic AMP signalling, and the role it plays in gene regulation for *E. coli* has been known for decades. In fact, cAMP signalling in bacteria was first discovered in *E. coli* and this led to the formation of paradigms still in use today. Cyclic AMP signalling as a part of bacterial physiology was discovered during studies of Carbon Catabolite Repression (CCR) in *E. coli*[166,167]. Our understanding of CCR has been iterated upon many times since. Currently it is known that *E. coli* have evolved mechanisms to preferentially utilise the most suitable carbon source available and prevent catabolism processes for other sources, until the preferred source is consumed[65,83,168]. For example, when glucose is present, it is actively transported into the bacteria by a special form of active transport, and phosphorylated. Import and phosphorylation are mediated by the phospho-transfer-system (PTS) (Figure 1.12). In this example, the PTS is specific for glucose import and processing (PTS^{glc}). As the sugar is phosphorylated upon entry, levels of phosphate available to a component of this PTS^{glc}, Enzyme II A ($EIIA^{Glc}$) are reduced. When phosphate availability is low, the non-phosphorylated form of $EIIA^{Glc}$ prevails. This non-phosphorylated form binds enzymes and transporters of secondary carbon source catabolism, thereby inactivating them (Figure 1.12A).

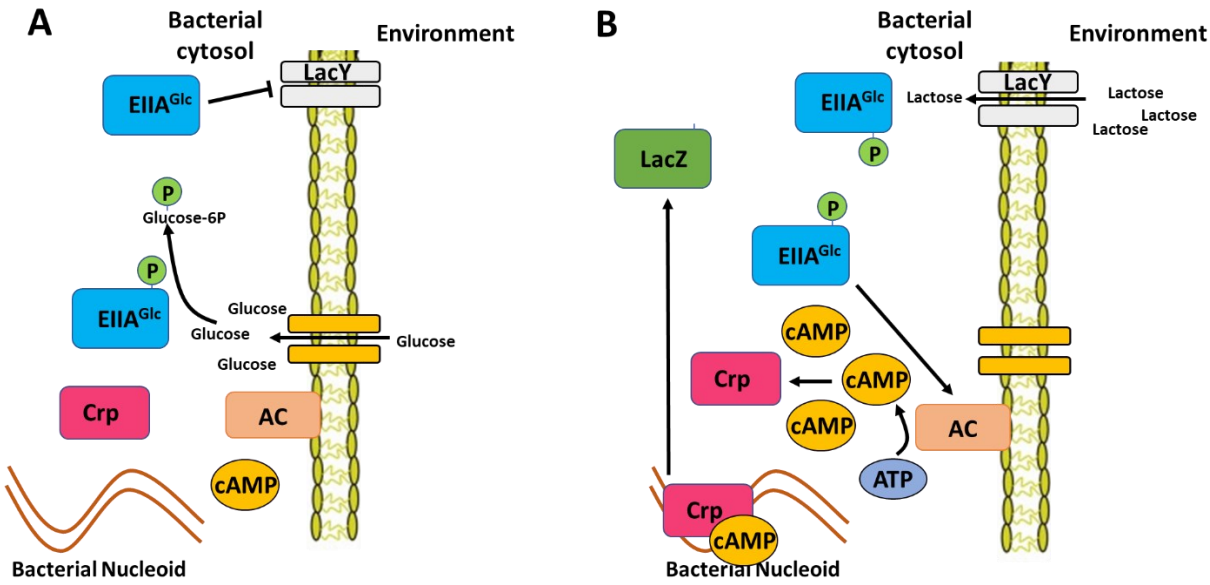


Figure 1.12: Simplified summary of cAMP mediated carbon catabolite repression in E. coli. A: Glucose is available and is transported into the cell. As it is transported, phosphate groups found on the glucose specific (EIIA^{Glc}) proteins, are transferred to incoming glucose molecules to form Glucose-6-phoshphate, by the Phosphotransfer system (PTS). When the EIIA^{Gl} proteins are unphosphorylated, they repress proteins involved in secondary carbon source utilisation, such as the lactose transporting protein (LacY). B: When Glucose levels are low and lactose is available, EIIA^{Glc} proteins are phosphorylated and can allow enzymes of secondary carbon source utilisation to work. In addition, EIIA^{Glc} proteins can activate adenylate cyclase enzymes and this causes the production of cAMP. cAMP then binds to bacterial receptor proteins (such as Crp). Crp can then bind to DNA and promote expression of genes involved in secondary carbon metabolism (such as lactose: LacZ – an enzyme that cleaves lactose into glucose and galactose in this example).

When glucose is readily available, EIIA^{Glc} is non-phosphorylated and works to inhibit catabolism of secondary carbon sources. However, when glucose levels are insufficient, EIIA^{Glc} is phosphorylated (Figure 1.12B). When phosphorylated it can activate the AC enzyme to produce cAMP. When cAMP levels are elevated, cAMP binds a cAMP receptor protein (Crp) and forms the cAMP-Crp complex. This complex can then bind to promotor regions in the genome and initiate transcription of genes from many different catabolic pathways. The literal text book example of this is the lac operon[83]. This paradigm needs constant re-evaluation as recent studies suggest the process of CCR is far more complicated and that understanding of this system in *E. coli* is lagging behind other pathogens[83,168]. For example, it has been recently shown that a combination of carbon and nitrogen source, in the presence of sufficient levels of cAMP is required for optimal function in multiple carbon source setting[169]. Although this study investigates very specific conditions, there is also evidence that this paradigm is not universal for all bacteria. For example, if the CCR system outlined for *E. coli* works by prevention of activation, in firmicutes like *Bacillus subtilis,* the system works by directly repressing transcription of the non-preferred catabolism pathway genes. In other gammaproteobacterial, such as *Pseudomonas putida* (*P. putida*), the cAMP-CRP complex is not even believed to be involved in CCR[170]. It is worth mentioning that although cAMP clearly has a significant contribution to regulation of metabolism and

CCR, there are other major contributors. For example, most recently growth rate has been postulated as a major regulator[171].

The cAMP-Crp complex and the Crp family of transcriptional regulators are conserved across several bacterial species and for a long time, cAMP-Crp was used as a reductive paradigm. Crp contains 2 cAMP binding domains, a DNA binding Helix-turn-Helix domain that recognises and binds surfaces for protein interaction with RNA polymerase[168,172]. Crp in the model organism *E. coli* binds to the palindromic motif TGTGAN6TCACA[173]. Across species, the concentrations of cAMP and the processes it affects varies. This can range from 20 pmol/10^8 bacteria in *M. tuberculosis*[174] to 2 pmol/10^8 bacteria in *E. coli*[175] and although not directly comparable – below limit of detection in *P. putida*[176]. Some explanation for differences in the systems may be explained by differential binding affinities of Crp for cAMP (quantified by K_D. or the dissociation constant). It has been postulated that these differences in affinity are tailored to cAMPs role in the organism's physiology[168]. For example, $Crp_{E.coli}$ has a K_D of 13-16 μM whereas Crp_{MT} has a K_D of 60 μM with 5-fold higher cAMP levels. In order to maintain nuanced control over transcription, even with higher cAMP levels, it would make sense for Crp to have a lower affinity for cAMP.

However, *P. putida* confounds this comparison with a K_D of 23 μM, and sub-picomolar levels[168]. Crp is a homodimeric protein and in complex is asymmetric. Since 2 molecules of cAMP allosterically activate this asymmetric homodimer, seemingly minor alterations in conformation or the bioenergetics of the residues can have a significant effect in changing the binding affinity for cAMP or DNA[172]. The Crp proteins of *E. coli* and *M. tuberculosis* H37Rv have a 32% identity[177] and have been shown to recognise and bind to a similar DNA motif[173,178]. However, some organisms with cAMP signalling have more distantly related members of the Crp transcription regulator family. For example, *P. aeruginosa* which has Virulence factor regulator (Vfr)[179]. This protein is still a member of the Crp superfamily, and has an alteration in the cAMP binding region that accommodates binding, but not activation by cGMP[179]. The processes that cAMP regulates in this pathogen also seem to be functionally distinct, mainly concerning virulence activities and not CCR[179].

Since the *E. coli* CCR paradigm was formed, it has been shown by metabolomic and transcriptomic studies that $Crp_{E.coli}$ can directly and indirectly influence expression of 100s of genes through multiple transcription factors[180]. Previously this was thought to be much lower (Görke and Stülke, 2008[83] vs Shimada et al, 2011[77,181–183]). When taken in the context of the variation of cAMP regulated processes across pathogens and variation within the machinery, this suggests that the picture is much more complicated and the *E. coli* cAMP-Crp complex is likely to be involved in many other roles across different organisms.

In fact, beyond CCR, there are multiple examples where subverting cAMP signalling in *E. coli* can be exploited for biotechnological applications. Some intriguing examples stem from alterations in the

resistance of *E. coli* to low pH or acetate levels. $Crp_{E.\ coli}$ was randomly mutated by error prone PCR and these mutated versions were transformed into ΔCrp *E. coli* strains. Transformed strains which showed increased tolerance to pH as low as 4.12[184] or to normally deleterious levels of acetate[185] were then created. Strains that could survive under these conditions are of great interest for the production of organic acids or for bioengineering processes that involve fermentation[184–186]. The success of this approach is believed to be a result of cAMPs role in managing 100s of genes, including those involved in the stress response, which will be elaborated upon later.

The concept of cAMP managed stress response is more universal than CCR. Indeed, there is even evidence of cAMPs involvement with the formation of non-culturable but viable *E. coli* bacteria. By overexpressing the PDE that breaks down cAMP, or by knocking out Crp / the AC that produces cAMP, this non-culturable phenotype can be induced. Bacteria in this phenotype can maintain virulence potential and become resistant to environmental or host stresses[187].

Furthermore, in uropathogenic *E. coli* strains (UPEC) affecting the urinary tract, the regulation of biofilm formation has been attributed to cAMP and Crp. This has been seen both by the production of curli and cellulose in lab strains (like *E. coli* strain 15) and in clinically relevant UPEC strains[188,189]. In the clinically relevant strain example, cAMP-Crp has been shown to regulate expression of a central regulator of biofilm formation CsgD, with 3 potential binding sites in the CsgD-CsgB intragenic region. Expression of this regulator protein leads to production of curli and cellulose, the 2 main components of the biofilm extracellular matrix (ECM). When Crp or the AC CyaA are knocked out, the bacteria were incapable of producing these ECM components[188]. Again, in other organisms there is evidence the process of biofilm formation is repressed by cAMP or even activated/repressed by other cyclic nucleotides. Several examples will be presented subsequently.

1.8.2 cAMP mediated virulence gene regulation in *Pseudomonas aeruginosa*

There are many other examples of pathogens that utilise cAMP signalling to sense and adapt to their environment. This can be essential for their virulence and life cycle. Although the role during infection of the cyclic nucleotide toxin ExoY remains to be elucidated, it is clear that cAMP and other cyclic nucleotides orchestrate and manage *P. aeruginosa* adaptions to the environment[116,190,191].

A central regulator critical to *P. aeruginosa virulence* is Vfr, which controls key virulence determinants like the Type-III secretion system – the delivery mechanism for many of the bacteria's encoded toxins. Vfr mediates control over the Type-III secretion system by binding to a promotor upstream of the last gene in the exsCEBA polycistronic mRNA sequence, that codes for the system. As previously mentioned, Vfr is a cAMP binding transcription regulating protein and is a member of the Crp family of cAMP regulator proteins. It binds to numerous genomic sites and regulates transcription[192]. It is clear from the *P. aeruginosa* genome and physiology that cAMP signalling is important as it codes for 2 AC enzymes – CyaA and CyaB[193], as well as a PDE (CpdA)[179]. To put this into perspective, the *E. coli* system has just 1 bona-fide AC and PDE[194]. Recently, evidence has emerged that suggests Vfr is itself regulated by glutathione (GSH), as the protein is vulnerable to oxidation. A knock-out mutant for bacterial GSH metabolism shows that under these circumstances, there is no Vfr mediated expression of the Type-III secretion system and decreased virulence in a mouse model[195].

There are multiple susceptible niches in the human host that *P. aeruginosa* can infect. In fact, it has been shown that Type-III secretion systems are only active in a portion of the bacterial population[192]. Another potential phenotype is the formation of biofilms e.g. in the lung of an infected cystic fibrosis patient – where it is indicative of a poor infectious outcome for the patient. Many isolates of *P. aeruginosa* from cystic fibrosis patients have mutations in 2 regulator proteins – YfiR and WspF. These proteins regulate cyclic-di-GMP (c-di-GMP) production[190,196,197]. Biofilm formation is regulated by another cyclic nucleotide, c-di-GMP. However, this involves a complicated hierarchal signalling network that begins with the Pilin protein (Pily1)[116,117]. Expression of this protein begins the process of biofilm formation via production of c-di-GMP and subsequent signalling. There is a negative correlation between c-di-GMP levels and expression of Vfr, as c-di-GMP levels rise, expression of the Type-III secretion system is prevented[116,198]. It was also shown that there is a concomitant decrease in cAMP levels, without inhibition of the AC or activation of the PDE[190]. Likewise, there is a positive correlation between cAMP levels, inhibition of biofilm formation and c-di-GMP signalling. Instead, cAMP induces type-III secretion system expression[191]. This complicated balance of cNMP is crucial to the success of *P. aeruginosa* in the host, and reinforces the importance of cAMP signalling to this pathogen. However, it also offers an example of how cNMP levels are modulated in response to changing environmental stimuli and how this translates into physiological adaptions.

1.8.3 cAMP mediated gene regulation of virulence and lifestyle switching in *Vibrio cholerae*

As previously mentioned, *V. cholerae* utilises a cAMP generating AC toxin to pathogenically raise cAMP levels within host cells – and this mediates the severe diarrhoeal pathology of infection[132,135,136]. *V. cholerae* also displays CCR, with a system similar to *E. coli*[199] and it has been shown experimentally that in the absence of glucose, cAMP is produced from an AC enzyme and then acts as a central regulator of gene expression in *V. cholerae*[199]. For example, cAMP regulates biofilm formation, toxin expression and quorum sensing[65,165]. This is mediated by cAMP binding to Crp_{VC}, which in turn binds at promotor sites in the genome to regulate gene expression. Through a combination of transcriptional profiling and validation studies, it was been shown that in a ΔCrp_{VC} strain, there was altered expression of "Repeats-in-Toxin" genes (a cytotoxicity associated toxin for host infection – RTX). The RTX operon (rtxBDE) was associated with the change in physiology required to infect zebra fish, a natural host for *V. cholerae.* Without Crp_{VC} control, RTX expression became uncoupled from host colonisation but was no longer induced upon infection of the zebra fish midgut[165].

Furthermore, Crp_{VC} represses biofilm formation by both direct and indirect mechanisms. In *V. cholerae*, biofilm formation requires 3 main activators; vibrio polysaccharide proteins (encoded by vpsR and vpsT) and the AphA protein. VpsR expression is directly repressed by Crp_{VC}[72,196]. Crp_{VC} mediates indirect control over biofilm formation via activation of the central repressor of VpsT and AphA – HapR. In fact, Crp_{VC} control over HapR is also direct and indirect – HapR expression requires expression of 2 autoinducers: Cholera autoinducer-1 and autoinducer-1 (CAI-1 and AI-1) and Crp_{VC} positively regulates CAI-1 levels[200,201]. Crp_{VC} also represses Fis, a small protein that promotes quorum sensing (Qrr) sRNAs – which in-turn repress HapR expression. Repressing Fis therefore prevents the Qrr sRNAs from repressing HapR[200,201].

Cyclic AMP and Crp_{VC} has also been seen to regulate metabolism and virulence at a post-translational level. A drosophila *V. cholerae* infection study with an acetate carbon source, showed that triglyceride accumulation in the fly gut leads to mortality[202]. A bacterial deacetylase (CobB) and an acetyltransferase (YfiQ), were required for accumulation of triglycerides in the fly gut. However, in a ΔCrp_{VC} mutant, YfiQ was significantly downregulated – this led the authors to conclude that post-translational lysine modification, and by extension cAMP regulation, could play a major role in bacterial acetate metabolism[202].

Although Cholera is preventable with sufficient sanitation, and usually treatable with oral rehydration or antibiotics, *V. cholerae* has a high level of natural competency. This means it can readily acquire resistance plasmids from the environment or surrounding microbial environment[203]. Competency and genetic plasticity are in part regulated by cAMP, although the exact mechanism still requires further

1.9 Mycobacteria and cAMP

In the earlier examples of pathogens and cAMP signalling (sections 1.7 and 1.8), it is clear there is heterogeneity in the functions entrusted to this system. For example, CCR in *E. coli*, virulence gene regulation in *V. cholerae* and *P. aeruginosa,* biofilm formation and adaption to the environment. The mycobacteria are a broad genus of bacterium, consisting of major pathogens like *M. tuberculosis*, *M. leprae* or *M. abscessus*, as well as the non-pathogenic *M. smegmatis* and the opportunistic *M. avium.* A common theme in this genus is relatively high levels of cAMP as compared to *E. coli* (20 pmol/10^8 [174] vs 2 pmol/10^8 [175]), the presence of multiple ACs, cyclic nucleotide PDEs and cAMP binding effector proteins. Furthermore, cAMP levels in the non-pathogenic *Mycobacterium smegmatis* are 10-fold higher than *M. tuberculosis*[174,204]

The *M. tuberculosis* laboratory strain H37Rv encodes 16 adenylate cyclase enzymes which produce cAMP, 10 cAMP binding effector proteins and thus far only 1 phosphodiesterase enzyme that breaks down cAMP[70,78,177,205–207]. In addition, *M. avium* encodes 12 ACs, *M. marinum* encodes 31 ACs and *M. smegmatis* mc^2155 encodes 7 ACs[208]. Furthermore, the homologue of $Crp_{E.\ coli}$ in *M. tuberculosis*, Crp_{MT} (Rv3676) has been shown to be required for full *M. tuberculosis* virulence in a mouse infection model[177]. The *in-vivo* growth of ΔCrp_{MT} bacteria in the lungs and spleens of BALB-C mice was severely impaired, when compared to WT and Crp_{MT} Complemented strains[177]. In fact, this protein has been shown to be involved with adaption to NaCl levels – such as those found in the macrophage[70], dormancy and resuscitation regulation [207,209], starvation[210] and energy metabolism[211,212]. An explanation of this can be found in the fact that Crp_{MT} has been shown to bind to around 900 genomic locations - mediating gene regulatory activities over 100s of genes[213]. The functions of this protein and other cAMP binding proteins in mycobacteria will be reviewed subsequently, and will be discussed in the context of virulence, environmental adaption, gene regulation and metabolomic adaption..

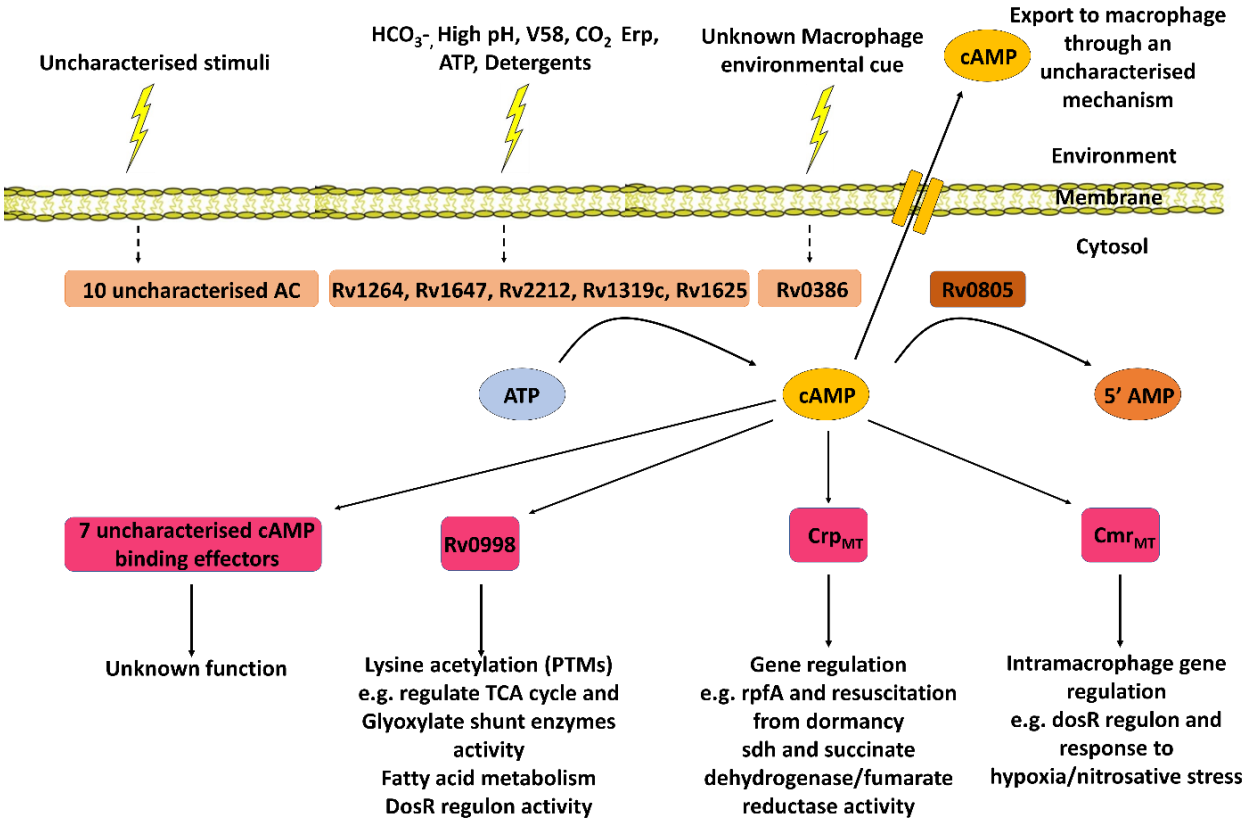


Figure 1.13: Summary of the characterised components of the cAMP signalling system in Mycobacterium tuberculosis. External stimuli lead to activation of mycobacterial AC enzymes, which then produce cAMP from bacterial ATP. Only 5 AC enzymes have had their activating signals characterised[68]. Of the 5 characterised enzymes, Rv1264, Rv1647 and Rv2212 are cytoplasmic whereas Rv1319c and Rv1625 are membrane bound. Rv0386 activation is uncharacterised but it has displayed a striking function, in that cAMP produced from this AC is exported into the macrophage and modulates the host immune response, thereby increasing bacterial survival[78]. Once cAMP is produced, it can bind various effector proteins. 3 of the 7 proteins have been characterised: Rv0998 – a lysine acetyltransferase, Crp_{MT} – a Crp family transcriptional regulator and Cmr_{MT} – another transcriptional regulator. To regulate cAMP levels, currently the only known PDE enzyme in mycobacteria is Rv0805. This enzyme can hydrolyse cAMP into 5' AMP. Off the remaining 10 uncharacterised AC, some have displayed activity in-vitro, but since they are poorly characterised – they will not be discussed further (membrane bound: Rv1320c, Rv3645, Rv1318c, 1900c, Rv2435 and Rv2488c. Cytoplasmic: Rv0891c, Rv1120c, Rv1358 and Rv1359). Likewise, there are 7 uncharacterised cAMP binding proteins beyond all containing a cAMP binding domain – Rv3728, Rv3239c, Rv2564 and Rv0073 are contain transporter domains. Rv3728 and Rv2565 contain phospholipase domains. Rv2434c is fused with a mechanosensitive ion channel. In addition, a universal stress protein (Rv1636) has also been found to bind cAMP with a non-canonical mechanism[214]. It is abundantly expressed in the log-phase and may play a role in modulating central carbon metabolism during the exponential growth phase (awaiting publication)[213,215,216].

1.10 cAMP intoxication of macrophages and *M. tuberculosis* subsequent evasion of killing

The importance of cAMP in infection is further highlighted by work showing that levels of cAMP increase 50-fold in *M. tuberculosis* H37Rv and *M. bovis* BCG bacteria, during macrophage infection[174]. Further studies have shown that a bacterial AC (Rv0386) produces cAMP which is secreted into the macrophage, thereby interfering with host cell signalling and preventing acidification of the phagosome[78]. The effect of this cAMP intoxication is that the bacteria escapes killing, can acquire nutrients, replicate inside the host macrophage and alter the immune response. The mechanism of alteration to host signalling has recently been further elucidated by Chung et al,[79]. When compared to uninfected donors, peripheral blood mononuclear cells (PBMCs) taken from latently infected tuberculosis patient show higher cyclic AMP levels, and reduced levels of the components of the IFN-γ production pathway[79]. The authors went on to show that when cells were exposed to analogues of cAMP (thereby mimicking the effects of the high levels of cAMP *M. tuberculosis* secretes into the macrophage) that CREB was downregulated. This subsequently impaired downstream signalling events and prevented PKA pathway-I activation of IFN-γ secretion. IFN-γ production and secretion has been seen to be required for bacterial killing during infection models and has long been seen as a correlate of protection[78,79]. Macrophages that do not receive IFN-γ stimulation or produce TNF-α have a reduced *M. tuberculosis* killing capacity[78,79]. Levels of TNF-α secretion from macrophages have previously been shown to be reduced[78] or increased[215] after bacterial AC cAMP production; depending on which bacterial AC was producing the cAMP (Rv0386 vs Rv1625). This is interesting as it again potentially illustrates the spatio-temporal component of cAMP signalling, that is widely observed in human cAMP signalling[84,145].

1.11 cAMP and mycobacterial response to the host environmental conditions

Many of the 16 AC enzymes in *M. tuberculosis* H37Rv remain functionally uncharacterised, and some have only been characterised *in-vitro*[68] (Figure 1.13 and Figure 1.14). All 16 of these enzymes are dimeric Class-III adenylate cyclases, a widely distributed class across eukaryotes and bacteria. These proteins typically have adenylate cyclase domains fused to different sensory input domains, thereby enabling them to produce cAMP in response to a specific stimulus[216,217]. The common feature in all the *M. tuberculosis* H37Rv ACs is the presence of a divalent metal binding site, substrate binding site (ATP) and transition state stabilising residues[216]. Beyond the aforementioned conserved aspects, the cyclases display wide diversity in their potential functions, with different N-terminally located sensory domains that can provide clues as to their activating signal. In addition, some of these proteins also harbour additional domains (Summarised in Figure 1.14).

Additional Domains	*M. tuberculosis* H37Rv	Activity	Gene expression	Activating signal
	Rv0891c	ND	NO/hypoxia	ND
	Rv1120c	Pseudogene	ND	Pseudogene
	Rv1264	Yes	Hypoxia	Low pH
	Rv1359	ND	H2O2	ND
	Rv1647	Yes	ND	High pH
αβ-Hydrolayse	Rv1900c	Yes	ND	ND
	Rv2212	Yes	Starvation	Fatty acid
HAMP	Rv1318c	Yes	Starvation	ND
HAMP	Rv1319c	Yes	Starvation	HCO3/CO2
HAMP	Rv1320c	Yes	ND	ND
	Rv1625c	Yes	ND	HCO3/CO2
HAMP	Rv2435c	ND	Hypoxia	ND
HAMP	Rv3645	Yes	ND	ND
Helix-turn-helix, ATPase	Rv0386	Yes	ND	ND
Helix-turn-helix, ATPase	Rv1358	ND	ND	ND
Helix-turn-helix, ATPase	Rv2488c	ND	Starvation	ND

Figure 1.14: Table detailing the 16 AC enzymes identified in M. tuberculosis H37Rv, any additional domains the proteins may have, whether they have been shown to be active and any further characterisation from the literature. Blue indicates that the enzyme is found in the soluble bacterial cell fraction and orange represents that it is found in the membrane fraction. ND = not detected. Additional domain identification was carried out BLAST-PSI and HMMM analysis, and further details can be found in the original publication[216]. The gene expression column describes conditions in which genome changes were measured in response to different conditions, and are not cAMP focussed studies. The original data sets are referenced in[218]. Modified and compiled from[216,218].

There are 5 well characterised examples of different sensory domains possessed by *M. tuberculosis* H37Rv AC enzymes[68]. These sensory domains have been shown to be stimulated by different conditions that can be found during host infection[218]. For example, a well characterised AC (Rv1264) contains an auto-inhibitory N-terminal domain that prevents enzyme activation at pH 7.5. However, at pH 6.0, the enzyme was shown to be 30-fold more active. Removal of this domain likewise boosted activity 300-fold[219]. The intra-macrophage pH in the phagosome is highly acidic and is thought to between pH 5.5 and 6.4[220–222]. This suggests that the role of this particular AC is to adapt to the inhospitable environment of the macrophage.

Another adenylate cyclase, Rv2122, has been shown to be activated by pH, fatty acids, CO_2, ATP and a secreted bacterial protein (Erp)[218,223] . The role of this protein in infection is unknown but this interaction was validated in a macrophage infection model. It shows that beyond adaption to the environment (e.g. pH and fatty acids), cAMP is regulated by multiple different processes, and that it requires fine-tuning.

The AC activity of Rv1319c has been shown experimentally to be responsive to different concentrations of bicarbonate ions (HCO_3^-)[224]. Across a range of HCO_3- concentrations, Rv1319c specific activity was upregulated 3-fold. The role of HCO_3- during mycobacterial infection is not clear, however HCO_3^- levels have been shown to be increased in the serum of TB patients[225]. *M. tuberculosis* H37Rv encodes 3 enzymes that degrade HCO_3^- to CO_2 and carbonic acid. Two of these genes are essential *in-vivo,* the Carbonic anhydrase genes (CA)[226]. This reaction has been postulated to act as a pH balance[227]. An inhibitor of these enzymes (Ethoxzolamide - ETZ) has also been shown to impair expression of genes regulated by the PhoPR 2 component regulatory system regulon[228] is induced by the low pH conditions found within the host macrophage[222]. Mutant strains of several of the genes in the PhoPR regulon show impaired growth during macrophage and mice infection[229] – most notably genes involved in complex lipid synthesis and virulence[229]. The authors of the ETZ study suggest that this carbonic acid hydratase inhibitor interferes with the regulatory cues for the PhoPR regulon, however, it did not appear that this was due to direct modulation of bacterial cytosolic pH[228]. By extension, this work suggests that bicarbonate may be involved in the bacterial adaption to the conditions within the macrophage. With the importance of cAMP production within the macrophage[78] this may explain the presence of multiple HCO_3^- sensing AC.

The *M. tuberculosis* H37Rv AC Rv1625 shows structural similarity to a mammalian AC. In fact, it has displayed robust AC activity when expressed in both mammalian and bacterial cells[230]. It consists of 6 transmembrane domains with a single class II AC domain, thus resembling half of a 12 transmembrane mammalian AC enzyme[216]. While it has also been shown to be responsive to HCO_3^-[224], it has, in addition, been shown to be responsive to CO_2[231] and V58 (a novel orphan cholesterol utilisation inhibitor)[11]. The responsiveness to CO_2 was again indicated by increased activity in the presence of

physiologically relevant CO_2 levels[231], and may serve as a similar environmental cue to HCO_3^-, as previously mentioned. However, when stimulated by V58, Rv1625 increased cAMP levels 20-fold when bacteria were grown in 7H12 media supplemented with cholesterol and acetate. During the macrophage infection model, V58 treatment appeared to intoxicate bacteria and led to drastic bacterial growth defects during macrophage infection[11]. Follow up work showed that this perturbation of normal bacterial cAMP signalling was interfering with numerous genes involved in cholesterol metabolism, although the exact mechanism was not determined[11]. This work was validated by a subsequent study[215] that also further highlighted the essentiality of cholesterol metabolism and detoxification of the downstream products of these pathways, for *M. tuberculosis* bacteria during infection[215]. Indeed, targeting this aspect of *M. tuberculosis* physiology represents a highly promising avenue for future therapeutics. It has been demonstrated that Rv1625, an adenylate cyclase, may have a key role during macrophage infection. The exact mechanism of action (MOA) of V58, and the role that cAMP regulation plays in cholesterol metabolism has not yet been elucidated, however, it makes a compelling argument for the development of therapeutics targeting bacterial cAMP signalling.

Rv1647 is considered a "divergent" adenylate cyclase, as the AC domain appears to be more akin to fungal and protist AC enzymes (30% similarity over 160/328 amino acid residues)[232]. It was highly susceptible to proteolysis in its full length form, but the AC domain was expressed, purified and displayed AC activity[232]. This catalytic domain was found to be responsive to high pH or 5mM concentrations of CHAPS zwitterionic surfactant (displaying 2-2.5fold increased activity) and interestingly still maintained a dimeric state with robust activity in the presence of up to 500mM NaCl. A concentration of 100mM NaCl was enough to impede Rv1625 dimer formation, but not activity[232]. This may be relevant to the intramacrophage infection setting, where the concentration of NaCl in the acidified phagosome has been shown to be 250mM[69]. The fact that this AC is conserved even in *M. leprae*[232] is also intriguing as this bacterium has undergone reductive evolution to reduce its genome (in terms of base pairs) to 75% of the size of H37Rv, with more than half encoding pseudogenes[233–235]. Conservation of genes from other mycobacteria in *M. leprae* is often thought to indicate some level of importance to the bacterium[232,235], however, without proper characterisation of a protein – this remains unproven.

Beyond the 6 AC enzymes discussed here (Rv0386, Rv2212, Rv1264, Rv1319c, Rv1625 and Rv1647), the rest of the AC enzymes remain poorly characterised[68]. The activating signal for Rv0386 also remains unknown, though it is clear that it is a condition found within the macrophage during *M. tuberculosis* infection[78,79]. When taken together, it is likely that *M. tuberculosis* retains so many AC enzymes to enable a nuanced response to a wide range of environmental conditions. As seen with pathogens described earlier, such as *V. cholerae*[199] and *P. aeruginosa*[190], the ability to respond and adapt to environmental cue is crucial for effective virulence.

Beyond adenylate cyclases, the cAMP binding effector protein Crp_{MT} has been implicated in the response to osmotic NaCl stress in *M. tuberculosis*[70]. When *M. tuberculosis*, *M. marinum* and *M. smegmatis* bacteria were exposed to physiologically relevant levels of NaCl (e.g. 250mM – the concentration found in macrophages[69]), *Mycobacterium tuberculosis*, but not *M. marinum* or *M. smegmatis* shows a statistically significant 3-fold increase in intracellular cAMP. When Crp_{MT} is deleted, this increase is no longer observed. This shows that the osmotic stress of 250mM NaCl induces a significant increase in intra-bacterial cAMP levels, and suggests that this is orchestrated by Crp_{MT}. However, the mechanism still requires elucidation[70].

Taken together, the large range of ACs, regulated by different stimuli certainly implies a role for cAMP in adaption to different environmental stimuli. For mycobacteria this is of the utmost importance, as *M. tuberculosis* is a highly successful human pathogen and must adapt to the inhospitable conditions of the macrophage. Figure 1.15 summarises the principles of cAMP signalling in mycobacteria.

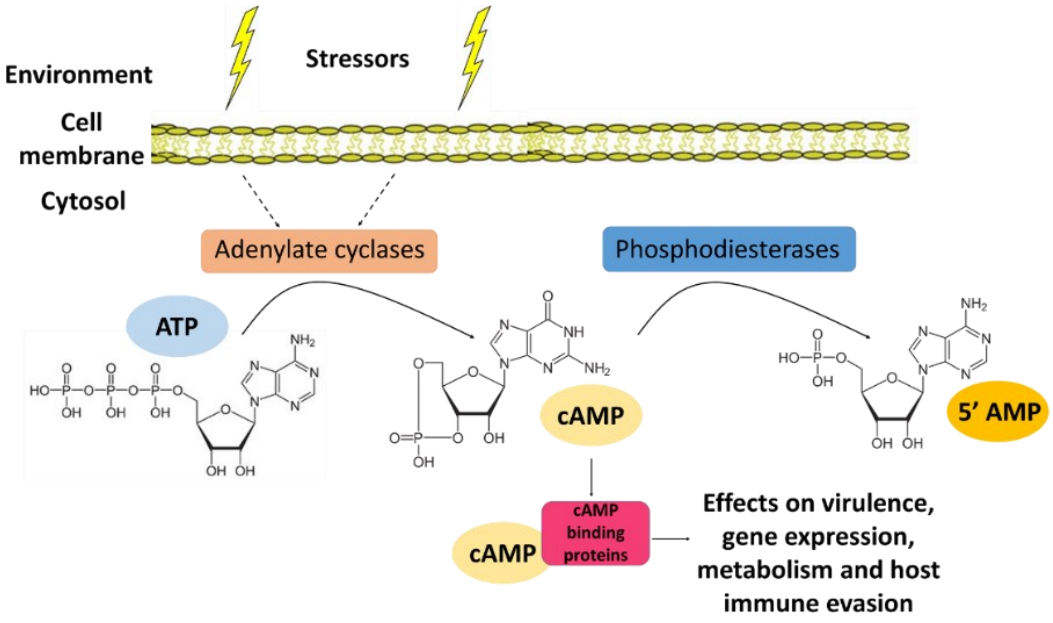


Figure 1.15: Simplified summary of current knowledge of mycobacterial cAMP signalling.

1.12 cAMP mediated gene regulation in mycobacteria

Cyclic AMP occupies a prominent role in mycobacterial gene regulation, for example, studies from Shleeva et al[207,209] have shown that cAMP can regulate "dormancy" in *M. smegmatis* and *M. tuberculosis*. It is worth noting that the authors of this work suggest "dormancy" to be a viable but non-replicative state[209]. In response to several free fatty acids (oleic, linoleic and arachidonic acids), the *M. smegmatis* homologue of Rv2122 (MSMEG_4279) was found to be activated and cAMP levels increased leading to the expression of resuscitation promoting factor A (RpfA). This muralytic enzyme was found to prevent bacteria entering a viable but non-replicative state under acid stress conditions and reverse this process in already viable but non-replicating bacteria. The RpfA-E proteins have been shown to collectively be dispensable for *M. tuberculosis* in-vitro, but are required for virulence and long-term persistence *in-vivo*[236].

The gene regulation activities of cAMP in *M. tuberculosis* H37Rv are ascribed to Crp_{MT} and Cmr_{MT}, 2 potent cAMP responsive transcription factors. In the case of Crp_{MT}, it had originally been thought to regulate around 100 genes[210] including *rpfA*[177], however, in multiple ChiP-seq studies, Crp_{MT} has been shown to bind to 191 points and over 900 points in the genome respectively[173,213]. While the earlier study found roughly equal intragenic and intergenic binding points, the subsequent study[237] suggests that the number of binding points in the genome is much higher, and with a majority (700) intragenic binding points. This study therefore suggests a much broader role for cAMP in regulation of gene expression, metabolism and physiology – with several of the top 10 regions for highest binding being of unknown function. The authors of this study went on to show that Crp_{MT} can functionally regulate a variety of essential genes in *M. tuberculosis* including those coding for fumarate reductase and succinate dehydrogenase[238]. In the case of fumarate reductase/succinate dehydrogenase gene regulons in MTB, 2 out of 3 sets are under the influence of Crp_{MT} – *sdh1* and *frdABCD* (Rv0247-0249 and Rv1552–1555 respectively). Fumarate reductase/Succinate dehydrogenase enzymes are involved in both TCA cycle and electron transport chain (ETC) activities. They mediate the conversion of succinate to fumarate, coupled to the reduction of FAD and also the reverse direction, which is required for the reductive portion of the TCA cycle[239]. It is worth noting that the aforementioned study did not show Crp_{MT} binding to the final succinate dehydrogenase/fumarate reductase regulon (*sdhCD*, Rv3316/Rv3317)[238]. While in *M. tuberculosis* infection these regulons have been shown to be very important in enabling persistence[238,240], there variation in prevalence of these regulons in other mycobacteria. Outside of *M. tuberculosis*, this system is well characterised in the non-pathogenic mycobacterium, *M. smegmatis*. For example, the *sdhCD* and *frdABCD* regulons appear to be absent from *M. smegmatis* mc^2155. However, in addition to orthologues of *M. tuberculosis* H37Rv Crp_{MT} and *sdh1* (MSMEG_0416-MSMEG_0420), *M. smegmatis* mc^2155 harbours a second Crp (Crp2) and *sdh2* (MSMEG_1672-MSMEG_1669). This suggests that cAMP gene regulation and regulation of bioenergetics is more

complex in *M. smegmatis.* While *sdh2* has been shown to be essential for *in-vitro* growth in *M. smegmatis*, *sdh1* was not. Also, only *sdh1* is predicted to be regulated by cAMP[212]. The variation in these systems likely reflects the different functional demands that pathogenic and non-pathogenic mycobacteria may encounter. The fact that both *M. tuberculosis* and *M. smegmatis* harbour multiple *sdh* and/or an *frdABCD* regulon(s) likely indicates the essentiality of their activities. With *sdh1* upregulated in energy limited settings in multiple species of mycobacteria, and with its critical position between the TCA cycle and ETC activities, it is highly likely to have an essential role in *M. tuberculosis infection*. What this example and the example of *rpfA* show is that, even outside of virulence settings, cAMP has been shown to be a regulator of highly important processes.

Cmr has also been shown to bind to around 368 sites in the *M. bovis* BCG genome at 200 points (99% homology to Cmr_{MT})[241]. It appears to autoregulate with binding sites directly upstream of the Cmr_{MT} promotor[241]. It had originally been shown to regulate expression of 4 genes that are known to be cAMP regulated[206] and several of these genes had been shown to be regulated by cAMP under low oxygen conditions, before this was ascribed to Cmr_{MT}[242]. This aspect was particularly interesting, as low oxygen is an environmental cue that would be prevalent in chronic *M. tuberculosis* infection[10]. However, more recent work has shown that 1 of these genes (*groEL2*) is regulated in a cAMP independent manner[243]. In addition, it appears that Cmr_{MT} has a role repressing the DosR regulon[241]. The DosR regulon is a group of 48 co-regulated genes that have been shown to be expressed in mycobacteria in response to hypoxia[244,245]. As previously mentioned, during chronic *M. tuberculosis* infection and due to the actions of the host immune system, *M. tuberculosis* bacteria are likely to be exposed to such low oxygen conditions[10]. The DosR regulon enables the bacteria to enter a viable but non-replicative state, where non-essential metabolism is slowed and energy state is maintained by anaerobic respiration processes[212,244,245]. This state is advantageous for the bacterium, allowing long term persistence in the host despite the immune response or drug treatment[244]. In fact, a *M. tuberculosis* strain that shows increased "hyper-virulence" when compared to the H37Rv lab strain – *M. tuberculosis* Beijing lineage has evolved to uncouple the DosR regulon from regulatory pathways, instead constitutively expressing it during laboratory culutre[246]. It has been postulated that this may give the Beijing lineage a competitive advantage during host infection, as it is already preconditioned with adaptions to aid resistance to the host immune response[246,247]. Indeed, when Cmr_{MT} was deleted, several DosR regulon genes were upregulated, as seen by transcriptomic analysis[243]. This study went on to show that binding of Cmr_{MT} and subsequent gene repression, was impaired by nitrosative stress – an environmental cue *M. tuberculosis* would encounter as part of the host immune response[10]. This suggests that during infection Cmr_{MT} mediated repression of the DosR regulon may be reduced, and that this could enable the bacteria to enter an advantageous persistent state. A Cmr_{MT} deletion mutant showed resistance to nitrosative stress and this is likely due to the increased expression of the DosR regulon genes[243]. However, this Cmr_{MT} deletion mutant showed transient attenuation in an aerosol mouse infection model[243]. When

taken together with the finding that Cmr_{MT} binds at 200 different points in the genome, this implies that the functions of Cmr_{MT} are not limited to DosR regulation.

Studies of Cmr_{MT} suggest that the cAMP binding pocket is occluded by an N-terminal domain that extends through the protein[248]. However, Cmr_{MT} has shown to be essential for cAMP induced changes to the *M. bovis* BCG proteome, observed in earlier studies[206]. This suggests that during *M. tuberculosis* H37Rv infection, the role of Cmr_{MT} and cAMP in regulating dormancy and responding to stress is complex and potentially indirect. Alternatively, it could be via a novel binding mechanism, but there is currently insufficient evidence for this hypothesis. It is clear that cAMP has an important role in gene regulation in *M. tuberculosis* and other mycobacteria. This is evidenced by the number of essential genes regulated by cAMP and the impaired performance of cAMP signalling mutants during infection.

1.13 cAMP and mycobacterial metabolism

Rv0998 is a cyclic AMP binding protein with lysine-acetyltransferase activity – and is therefore capable of post-translationally modifying lysine residues of proteins to regulate their activity[249]. This was initially shown[250] with a study of the *M. smegmatis* homologue (MS_5458)[249]. The *M. smegmatis* homologue of Rv0998 displayed the ability to acetylate a mycobacterial Universal Stress Protein (USP), in a cAMP dependant manner[249]. This was confirmed with bacterial expression a GST-tagged MS_5458 and pull-down of bound/interacting proteins, that were subsequently identified by proteomic mass spectrometry. This was the first report showing lysine acetylation in mycobacteria and the authors finished the study with an *in-vitro* validation of the ability of purified Rv0998 to acetylate the lysine of a USP. The authors showed that in the presence of cAMP, the activity of Rv9998 was increased 10-fold[249]. Additionally, a follow-up study from this group showed that 18 proteins in *M. bovis BCG* were acetylated *in-vitro* and *in-vivo* proteomic mass spectrometry assays[251]. Several of these proteins were from the Acyl-CoA FadD family of proteins, which are responsible for activating fatty-acids for subsequent metabolism[251].

In *E. coli* Acyl-CoA FadD was shown to be essential for utilisation of endogenous fatty acids released from membrane lipids[252]. *M. tuberculosis* has been shown to utilise host cholesterol[253] and other lipids for energy metabolism and biosynthesis of cell wall components[254]. *M. tuberculosis* has a complex cell wall structure that is essential for successful infection, modulating the host environment and resisting antimicrobials[254]. This means acquisition, utilisation and metabolism of available lipids, and therefore the FadD enzymes, are highly important[251]. In addition to identifying these lysine-acetylated enzymes, the authors also characterised several of the lysine residues that were acetylated in 4 of these FadD enzymes (FadD 13, FadD 22, FadD 2 and FadD 5). In the first example, FadD 13, K487 was acetylated, and this led to increased activity of the enzyme *in-vitro*[251]. The lysine acetylation sites in the other 3 FadD enzymes correlated to the positions of K487 – suggesting this mechanism of regulation was conserved across these enzymes. Importantly, the authors also show that acetylation of FadD 13 was cAMP dependent, with a combination of knockout and complemented Rv0998 *M. bovis* BCG homologues and modulation of bacterial cAMP levels – thereby confirming the role of Rv0998 in these modulations *in-vivo*[251].

A subsequent study validated the involvement of Rv0998 in carbon metabolism by showing that the *M. smegmatis* homologue of Rv0998 acetylates Acetyl-CoA synthetase (ACS), in a cAMP dependant manner[250]. Interestingly, this study provides an example of acetylation inhibiting an enzymes activity, when lysine residue (K) 617 is acetylated, activity is lost and is restored by deacetylation (with sirtuin – Rv1151c). ACS is involved in acetate metabolism, suggesting that beyond fatty acid metabolism, Rv0998 also has a role in regulating carbon metabolism.

All the previous findings are validated in an *M. tuberculosis* H37Ra study which shows 137 proteins acetylated by Rv0998[255]. As expected, this involves enzymes from fatty acid metabolism and confirms studies into FadD acetylation. Notably, 8 of the acetylated enzymes were from the TCA cycle, as well as isocitrate lysase (Icl) and malate dehydrogenase G[255]. Isocitrate lyase and malate dehydrogenase are components of the glyoxylate shunt, which has recently been shown to be essential for *M. tuberculosis* utilisation of host cell lactate and pyruvate during infection[256]. Lysine acetylation of ICL by Rv0998 was subsequently characterised in a different study[257]. Of the 2 lysine acetylation sites in ICL, acetylation of K392 increased activity of the enzyme and acetylation of K322 was found to decrease activity[257]. Although this study did not investigate cAMP directly, it noted that the level of acetylation and subsequent activity varied with carbon source. For example, both in WT and a knock-out of the only known sirtuin family deacetylase (ΔRv1511c), the strain showed increased enzyme activity when oleic acid and propionate were the sole carbon sources[257].

Several studies have demonstrated that Rv0998 or its homologues in multiple mycobacterial species can acetylate lysine residues of proteins in a cAMP dependent manner. It is clear that a numerous TCA cycle, glyoxylate shunt and fatty acid metabolism genes can be regulated in this way in infection relevant settings. In summary, this suggests that Rv0998 occupies an important role in regulating metabolism during *M. tuberculosis* infection. To reinforce this point, a recent study has demonstrated that Rv0998 acetylates 269 *M. tuberculosis* H37Rv proteins under hypoxic conditions[258]. Of these proteins, K182 of DosR was acetylated by Rv0998, a key regulator of the DosR regulon that enables the bacteria to survive hypoxic conditions. Indeed, a ΔDosR mutant, complemented with a mutated DosR (K182R) lacking the crucial lysine residue, showed attenuated growth under hypoxic conditions[258]. The authors also carried out a peritoneal mouse infection model with WT, ΔDosR, ΔDosR comp DosR and ΔDosR comp DosR (K182R). The ΔDosR strain complemented with the mutated lysine residue showed the lowest CFU counts in both lung and spleen (with robust statistical significance) – as well as the least histopathology[258]. Taken together, this suggests that Rv0998 mediated lysine acetylation of DosR is important during infection, and even plays a role in the crucial hypoxia response in *M. tuberculosis*.

Rv0998 is conserved across all mycobacterial genomes, so it may mediate an important and conserved role in mycobacterial physiology beyond infections of pathogenic bacteria. In *M. smegmatis* and *M. tuberculosis*, it has been shown to be autoinhibitory in the absence of cAMP[68,251,259]. Propionate and Acetyl-CoA metabolism is highly important for *M. tuberculosis* bacteria, as cholesterol catabolism is a key source of carbon and cell wall precursors during host infection[11,50,215]. Catabolism of cholesterol in *M. tuberculosis* yields acetyl-CoA, propionyl-CoA, and pyruvate – which can feed into central carbon metabolism or fatty acid production[11,215]. However, accumulation of propionyl-CoA is highly toxic to the bacterium and so detoxification of this metabolite is essential to *M. tuberculosis* survival[215]. Propionoyl-CoA can be metabolised into Acetyl-CoA and propionate and cyclic AMP has been

implicated in regulation of the genes mediating this process in *M. bovis* BCG[251]. Furthermore, as previously mentioned, Crp_{MT} regulates key metabolic enzymes that bridge the ETC and TCA cycle, such as fumarate reductase and succinate dehydrogenase[238,239]. Crp_{MT} also regulates rpfA, a resuscitation promoting factor[207] and both Cmr_{MT}[243] and Rv0998[258] regulate the DosR dormancy regulon. Dormancy, resuscitation, ETC/TCA cycle activities, fatty acid metabolism and central carbon metabolism are all essential processes for effective survival in *M. tuberculosis* infection settings and cAMP directly or indirectly has a role in regulating each of these processes. While cholesterol catabolism is a key carbon source during host infection[50], *M. tuberculosis* displays compartmentalised carbon co-catabolism – allowing multiple carbon sources to be metabolised in paralell[253]. This suggests that even the CCR paradigm from *E. coli*, is not the main role of cAMP in *M. tuberculosis*.

Taken together, this evidence further supports the broad importance of cAMP, to metabolism, gene expression and virulence. This highlights that further research is still required to understand the mechanisms underlying these complex roles and that doing so could provide clues on how to better treat *M. tuberculosis* infection.

1.14 Modulating cAMP and antimicrobial tolerance in bacterial pathogens and mycobacteria

As previously described *M. tuberculosis* and other mycobacteria are innately highly resistant to treatment with antimicrobials[42,47,48,245,254,260]. While the reasons for this are multifactorial, briefly, the complex mycobacterial cell wall provides a thick waxy coat decreasing permeability to many antimicrobials[254]. In addition, the host niche of *M. tuberculosis* infection – the macrophage[50] and/or the hypoxic, lipid rich granuloma[61], decrease the effectiveness of antimicrobials and causes the bacteria to activate phenotypic adaptions that incidentally decrease their susceptibility to many antimicrobials[47,48]. Underpinning these physiological mechanisms, is the ability to respond effectively to environmental conditions – a process cAMP has been shown to be intimately involved with in a myriad of studies[65,68]. It is therefore highly likely that cAMP is involved directly or indirectly, with regulating the susceptibility of mycobacteria to antimicrobials. The evidence for this will be reviewed subsequently.

With the importance of cAMP to mycobacteria for sensing and adapting to environmental conditions during infection, coupled with the urgent need for solutions to the rising problems of AMT and AMR, we sought to investigate the link between cAMP and the development of antimicrobials tolerance.

In order to sense and respond to antimicrobials treatment or environmental stimuli, bacteria need a rapid, exquisitely tuneable and metabolically favourable signalling system. A good example is a cyclic nucleotide messenger like cAMP. As previously shown, cAMP and the systems it regulates, are conserved across a broad spectrum of pathogens and organisms. It mediates processes ranging from virulence, host immune evasion, carbon metabolism, gene expression and post translational modifications[65]. Beyond these well reported roles, a lesser known role has been evidenced in *Salmonella typhpimurium*, initially in the work of Alper and Ames (1978)[261]. In their study, they use genetic knockouts of cAMP signalling machinery. For example, the AC enzyme producing cAMP from ATP, the phosphodiesterase enzyme which breaks down cAMP into the inactive 5' AMP, and the cyclic AMP receptor binding protein CRP_{ST} – that mediates the functional consequences of these signals. They show that cAMP signalling regulates permeability to antimicrobials, by positively regulating transporters of carbon or nutrient influx – as well as transporting antibiotics.

Following on from this work, many years later, Kary et al[262] show that the CRP_{ST} protein regulates the susceptibility of *S. typhimurium* to fluoroquinolones, a chemical class of antibiotics. For example, the ΔCrp_{ST} mutant showed far higher growth than WT, when both strains were treated with ciprofloxacin and other antimicrobials[262]. Subsequently, the authors found that in the ΔCrp_{ST} mutant, levels of the gene encoding a drug efflux channel (*acrB*) were increased and that genes encoding porin (drug influx) genes (*ompF* and *ompA*) were reduced[262]. Similar work in *E. coli* has shown that modification of the $Crp_{E.\ Coli}$ and $CyaA_{E.coli}$ genes via transposon insertion leads to decreased susceptibility to several

aminoglycoside antibiotics[263]. The authors of this study postulate that the transposon insertions which led to increased resistance to these antibiotics partially impairedreduced the formation of the cAMP-Crp complex – which was thought to regulate aminoglycoside transporters. Further evidence for the involvement of Crp_{ST} and import/export comes from competition experiments of WT *S. typhimurium* and ΔCrp_{ST}, co-cultured with *S. aureus*, *V. cholerae* or *E. coli*[264]. In rich media, the ΔCrp_{ST} strain could not outcompete the other pathogens. However, in nutrient-poor media, survival was unaffected[264]. Taking this work further, the authors also went on to show that under low temperature conditions or during treatment with antibiotics (such as chloramphenicol), that the ΔCrp_{ST} showed enhanced survival vs. WT. The authors postulate that these observations suggest Crp_{ST} is involved in nutrient up-take (as seen by Alper and Ames[261]) and that it may be involved in regulating metabolism[264]. When these works are taken together, it is highly likely that cAMP could have a role in antibiotic susceptibility in bacteria.

We therefore hypothesised that cAMP may also have a role in regulation of antimicrobial tolerance in mycobacteria. AMT in mycobacteria is of great concern as *M. tuberculosis* and *M. abscessus* are innately highly tolerant to antimicrobials[46,265]. The longer bacteria survive exposure to an antimicrobial, the longer they will have to mutate and evolve resistance[35,266]. If we can better understand cAMP signalling and how it might regulate antimicrobial susceptibility in mycobacteria, perhaps we can find novel resistance breaking strategies for many of the aforementioned pathogens. Mycobacteria represent a perfect model for cAMP as they have innately high levels of this second messenger. In particular, *M. smegmatis,* a commonly used, fast growing, non-pathogenic model for TB that has high cAMP levels (More than 100 times higher than *E.coli* when both organisms are in the exponential phase, and grown on a glucose carbon source[175,204]). *M. smegmatis* is sometimes used as a model for *M. tuberculosis* due to its safety profile, rapid growth, conserved house-keeping processes and related genes[235,267]. However, although a variety of *M. tuberculosis* virulence-associated genes can be found in *M. smegmatis,* the functions of some of these genes have been shown to differ between the organisms[268]. Taken together with the presence of unique sets of genes associated with virulence in *M. tuberculosis* and the differing environmental niches, validation of findings from *M. smegmatis* involving *M. tuberculosis* virulence proteins in *M. tuberculosis* is advisable[269]. Interestingly, in terms of antimicrobial susceptibility *M. smegmatis* and *M. tuberculosis* are similar across a range of relevant antibiotics. For example, under nutrient limited conditions *M. smegmatis* can enter a non-replicating state that mirrors the susceptibility profile of *M. tuberculosis* towards amikacin, capreomycin, rifampicin, ethambutol and isoniazid (it is worth noting that the MIC values may differ)[270]. Furthermore, in terms of cAMP signalling, *M. smegmatis* also possess multiple AC enzymes and two such enzymes have shown conserved activities between *M. smegmatis* and *M. tuberculosis*[204,207]. This supports the suitability of *M. smegmatis* to be used a model to investigate a link between cAMP signalling and antimicrobial susceptibility in mycobacteria.

Cyclic AMP signalling in mycobacteria is highly complex, nuanced and poorly understood (Figure 1.13). It has been difficult to investigate this signalling system because of the redundancy of AC enzymes – 16 in TB comparing with just 1 AC and 1 PDE in *E. coli* [65]. This makes it highly difficult to modulate cAMP levels in the bacteria.

Until now, there has been only limited means of modulating mycobacterial intracellular cAMP levels. These have been with chemical adenylate cyclase activators – such as forskolin[271]. Forskolin is a diterpene of the labdane family, produced from the by the Indian Coleus plant (*Plectranthus barbatus*)[272]. It has been used in a variety of eukaryotic cell types to activate AC enzymes and induce cAMP production[271]. Use of forskolin has not been widely reported in bacterial studies, with some notable exceptions[11]. This may be because the exact MOA is unknown, with debate as to whether it directly activates the catalytic site of an adenylate cyclase[271], or by binding to elsewhere on the protein[273]. Forskolin can also activate other targets within cells[274], which unfortunately reduces the usefulness of Forskolin for *in-vivo* or infection studies. Although forskolin has been shown to activate the *M. tuberculosis* AC Rv1625, it has been postulated that this is due to the similarity of Rv1625 to eukaryotic adenylate cyclases – which can readily be activated by Forskolin[11,232].

Another method involves a cAMP analogue – dibutyryl-cAMP. This analogue is less hydrophobic than cAMP and is therefore able to pass through cell membranes, where it accumulates intracellularly as mono-butyryl-cAMP[275]. Furthermore, dibutyryl-cAMP acts as a competitive PDE inhibitor due to its resistance to PDE hydrolysis[276]. It has been added exogenously to bacterial cultures to mimic the effects of cAMP[207,277]. However, due to its resistance to PDE hydrolysis and the fact that it is added exogenously, it is difficult to specifically modulate the intracellular concentration. This acts as a limitation and normally sees a range of concentrations used to provoke an effect, rather than physiologically relevant modulation of cAMP levels[209].

In order to find more elegant solutions to modulating intra-bacterial cAMP levels, our lab sought to approach this problem by targeting the breakdown of cAMP. Currently there is only 1 cAMP PDE annotated in MTB – Rv0805[205]. This enzyme is a class-III PDE enzyme and is grouped with most of the known bacterial cNMP PDE enzymes[278]. Rv0805 is a 318 amino acid residue homodimer ($Rv0805^{1318}$)[205]. The dimerization was shown to be metal dependant and the protein was seen to be localised to the membrane[205]. Although Rv0805 can act as a cAMP PDE, it is promiscuous in that it shows comparable activity as a cGMP PDE (K_m ~150μM in both cases) and shows some activity towards the general phospho-diester substrate bis(p-nitrophenyl phosphate)(K_m 1.3mM)[205]. When Rv0805 was being crystallised initially, the first 278 residues (excluding the C-terminal) crystallised successfully ($Rv0805^{1-278}$), more readily than the full length[279]. While eventually a full-length protein was crystallised, this led to investigation of the differences between the 2 versions of Rv0805 (summarised in Figure 1.16).

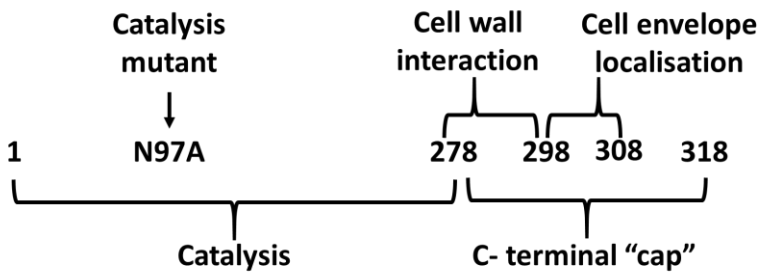


Figure 1.16: Schematic of the Rv0805 amino acid sequence and the functions of each portion of the protein.

The C-terminal portion was found to be dispensable for PDE catalysis activity[279], however, it was conducive for the localisation of Rv0805 to the cell membrane[279,280]. Furthermore multiple studies have utilised an overexpression of Rv0805 N97A activity mutants (cNMP catalysis-deficient) to highlight "moonlighting" or catalysis-independent protein activities[281,282]. Via crystal violet and malachite green permeability assays, and in the absence of Rv0805 mediated catalysis, there was an increase in cell-permeability – the cause of which was not fully elucidated[281]. In addition, a subsequent study found that Rv0805 N97A displayed cAMP catalysis-independent gene regulation. This included alterations in *icl1*, other methyl-citrate pathway enzyme encoding genes, the transcription factor whiB7 and its target *eis*[282]. A ΔEis *M. tuberculosis* mutant has previously been found to have compromised growth on cholesterol[283]. Likewise, the methyl-citrate pathway is required for *M. tuberculosis* utilisation of host cholesterol during infection[256,284]. When taken together, it is clear that Rv0805 possesses other activities and functions beyond 3', 5' - cAMP hydrolysis[285].

Additionally, Rv0805 is only present in pathogenic mycobacteria (*M. tuberculo*sis*, M. bovis* BCG*, M. leprae and M. marinum*), however, a study by Lee[286] shows that cAMP PDE activity also exists in non-pathogenic mycobacteria – such as *M. smegmatis.* When expressed in *M. smegmatis*, Rv0805 only mediates up to a 50% decrease in 3', 5' - cAMP levels; while also displaying higher activity for other cyclic nucleotide messengers, such as the non-second messenger acting 2', 3'- cAMP (a product of DNA degradation that has low physiological levels[278,282,287]) . When overexpressed under the native promotor of Rv0805, in *M. smegmati*s, cAMP hydrolysis can be increased. However, under such conditions, Rv0805 still only mediates a 50% drop in intracellular cAMP levels (in both $Rv0805^{1-318}$ and $Rv0805^{1-278}$)[281].

With comparatively low hydrolysis activity when compared to *E. coli* ($Rv0805^{205,279}$ V_{max} of 30 nmol/sec/mg vs. *E. coli*[194] CpdA 2µmol/sec/mg) – coupled to the 10-100 fold higher cAMP levels found in mycobacteria vs. *E. coli*[174,175,204,288], the fact that Rv0805 mediates various other functions[281,282], and that it is not conserved across all mycobacteria[205,286] – likely indicates that Rv0805 is not the primary cAMP PDE in mycobacteria. It is therefore highly likely that there is another, 3',5' - cAMP degrading "workhorse" enzyme ubiquitously expressed across mycobacteria. This hypothesis appears more likely

when the conserved cAMP PDE in many bacteria – CpdA (from E. *coli*) has no reported homologue in mycobacteria[278]. By identifying a new cAMP PDE with higher activity, it may yield a more physiologically relevant and tuneable method to modulate cAMP levels.

1.15 Aims

1 To identify the ubiquitous cAMP PDE in mycobacteria, express it, and use it as a tool to modulate the intracellular cAMP levels in the model organism *M. smegmatis* mc^{2}155

This will involve utilising an unbiased activity-led approach to identify a novel cAMP PDE enzyme – as traditional homology-based approaches have so far been unsuccessful. Once identified, the enzyme is intended to be expressed first in *M. smegmatis* mc^{2}155 and potentially *M. tuberculosis* H37Rv (time permitting). Overexpression of the enzyme in the mycobacterial species is intended to characterise the phenotypic effects of enzyme expression and altered intrabacterial cAMP levels.

2 To investigate the effect of decreased cAMP levels on bioenergetics, transcription and metabolism of *M. smegmatis* mc^{2}155

By utilising the lab group's specialities of Agilent 6545 Q-TOF LC-MS-led targeted and untargeted metabolomics workflows, bioenergetic investigation with the Seahorse XFP analyser and RNA sequencing techniques, the phenotypic effects of lowered cAMP levels will be investigated. This will allow insight into a previously difficult to modulate signalling pathway – known to involved in regulating virulence activities, metabolism and gene expression. Any findings from the 3 aforementioned screens will be validated with traditional microbiology techniques.

3 To test the hypothesis that cAMP is involved in drug resistance or tolerance in in the model organism *M. smegmatis* mc^{2}155

The focus of this project is to test this hypothesis – which has previously not been investigated in mycobacteria. Due to the conserved nature of cAMP signalling across pathogens, any findings may be applicable to a broad spectrum of fields. Most notably, in devising therapies to block the development of antimicrobial resistance.

4 To express and purify this enzyme for *in-vitro* characterisation and enzyme kinetics

Discovery of any novel cAMP PDE enzyme will be of great interest to the filed, as cAMP signalling is so widely conserved across taxa. A key component of fully determining the novelty or even the best conditions to modulate cAMP levels will require investigation and characterisation of the kinetic parameters of the enzyme. This is intended to be done by utilising a targeted metabolomic work flow, combined with optimised buffer, substrate and time parameters.

Chapter 2 - Materials and Methods

2.0 Materials

All reagents were purchased from Sigma-Aldrich, unless otherwise stated. Metabolomics reagents were purchased from Agilent Technologies (United States).

2.1 Methods

2.1.1 Bacterial strains and growth conditions

All mycobacteria cultured as part of the work of phenotypic characterisation (*M. smegmatis* m^{c}2155 and *M. tuberculosis* H37Rv) were cultured in 7H9 liquid medium or on 7H10 Agar. Cultures were maintained in shaking incubators (180rpm) or static incubators at 37°C. These strains are described in Tables 1, 2 and 3.

All bacteria that were used as part of the purification work (*E. coli*: BL21, MC1061, DH5-α, PlysS, C41, C43 and mycobacteria: *M. smegmatis* mc^{2}155 and *M. smegmatis* mc^{2}155 GroEL1ΔC) were cultured in either Luria-broth, terrific broth or 7H9 (only for mycobacteria). Detailed in Table 1-3. Supplements for the aforementioned medias will be described in their Materials and Methods section, where applicable. Any alterations to these standard media conditions will be detailed in the experimental results section, where applicable.

Table 1: Pathogenic mycobacteria strains used in this project

Strain	Origin	Resistance
***M. tuberculosis* H37Rv**	Laboratory collection	None
***M. tuberculosis* H37Rv pVV16 Empty Vector Control**	Produced as part of this Project from our Laboratory H37Rv parental stock	Kanamycin 25 µg/mL and Hygromycin 50 µg/mL
M. tuberculosis* H37Rv pVV16::*rv1339L	Produced as part of this Project from our Laboratory H37Rv parental stock	Kanamycin 25 µg/mL and Hygromycin 50 µg/mL
M. tuberculosis* H37Rv *Δrv0805	Lab collection	None
***M. tuberculosis* H37Rv *Δrv0805* pVV16 Empty Vector Control**	Lab collection	Kanamycin 25 µg/mL and Hygromycin 50 µg/mL
M. tuberculosis* H37Rv *Δrv0805* pVV16::*rv1339	Lab collection	Kanamycin 25 µg/mL and Hygromycin 50 µg/mL
M. tuberculosis* H37Rv *Δcrp	Dr Roger Buxton, NIMR, Mill Hill, London, UK	None
***M. tuberculosis H37Rv Δcrp* pVV16 Empty Vector Control**	Produced as part of this project from the parental strain originating from the lab of Dr Roger Buxton, NIMR, Mill Hill, London, UKDr Roger Buxton, NIMR, Mill Hill, London, UK	Kanamycin 25 µg/mL and Hygromycin 50 µg/mL
M. tuberculosis* H37Rv *Δcrp pVV16 Empty Vector Control	Produced as part of this project from the parental strain originating from the lab of Dr Roger Buxton, NIMR, Mill Hill, London, UKDr Roger Buxton, NIMR, Mill Hill, London, UK	Kanamycin 25 µg/mL and Hygromycin 50 µg/mL
***M. bovis* BCG Δm*b0828* (*rv0805* homologue)**	Dr Roger Buxton, NIMR, Mill Hill, London, UK Dr Roger Buxton, NIMR, Mill Hill, London, UK	Hygromycin 50 µg/mL

Table 2: Non-pathogenic mycobacterial strains used in this project

Strain	Origin	Resistance
M. smegmatis mc²155	Dr Brian Robertson, MRC CMBI, Imperial College, South Kensington, London UK	None
***M. smegmatis* mc²155 pVV16 Empty Vector Control**	Produced as part of this project from the parental strain originating from the lab of Dr Brian Robertson, MRC CMBI, Imperial College, South Kensington, London UK	Kanamycin 25 µg/mL and Hygromycin 50 µg/mL
M. smegmatis* mc²155 pVV16::*rv1339 Full-length	Produced as part of this project from the parental strain originating from the lab of Dr Brian Robertson, MRC CMBI, Imperial College, South Kensington, London UK	Kanamycin 25 µg/mL and Hygromycin 50 µg/mL
M. smegmatis* mc²155 pVV16::*rv1339S	Produced as part of this project from the parental strain originating from the lab of Dr Brian Robertson, MRC CMBI, Imperial College, South Kensington, London UK	Kanamycin 25 µg/mL and Hygromycin 50 µg/mL
***M. smegmatis* mc²155 GroEL1 ΔC**	Dr Stephane Canaan, CNRS, Marseille, France	None

Table 3*: E. coli* Strains used in this project

Strain	Origin	Resistance
BL21 (DE3) E. coli	Laboratory collection	None
PlysS *E. coli*	Laboratory collection	None
C41 *E. coli*	Laboratory collection	None
C43 *E. coli*	Laboratory collection	None
DH5-α *E. coli*	Laboratory collection	None
MC1061	Profesor Thomas Meier, Imperial College, South Kensington, London	Streptomycin 50µg/mL

2.2 Culture Medium

2.2.1 7H9 liquid medium

Middlebrook 7H9 broth was prepared according to the manufacturer's instructions. 4.7 gram (g) 7H9 powder was added to 875 mL Milli-Q water (ddH_2O). If the osmolarity of the medium was required to be 25 millimolar (mM) sodium chloride (NaCl), 50 mL of NaCl solution 5 M in H_2O (Sigma) were added per litre of final culture medium. The pH of the solution was adjusted using the Mettler Seven Easy S20 pH meter (Sigma). Then, the solution was autoclaved at 121 °C for 20 minutes (min) and permitted to cool. 100 mL of supplement were added to the cooled 7H9 solution under sterile conditions. The 7H9 medium was stored at 4 °C.

2.2.2 Un-buffered 7H9 liquid medium for Seahorse XFP assays

From a recipe obtained from the Lab of Professor Adrien Steyn, the standard Middlebrook 7H9 recipe was prepared, minus the buffering components: Sodium Citrate and Glutamic acid. The pH of this media was adjusted to pH 7.4 by addition of diluted hydrochloric acid. The resulting solution was then syringe filtered in a laminar flow hood. The carbon source supplement (Glycerol and Dextrose) was prepared at a final concentration of 2% in ddH_2O and syringe filtered.

2.2.3 7H10 agar

Middlebrook 7H10 agar was prepared according to the manufacturer's instructions as follows: 21 g of 7H10-agar powder was dissolved in 900 mL ddH2O and mixed well. If the osmolarity of the medium was required to be 250 mM NaCl, 50 mL of NaCl solution 5 M in H2O (Sigma) were added per litre of final culture medium. Then, the pH of the solution was adjusted using the Mettler Seven Easy S20 pH meter (Sigma), the solution was autoclaved at 121 °C for 20 min and permitted to cool in a 55 °C incubator for 1 hour. The preparation was further cooled by stirring at 4 °C for 5 min before adding 100 mL supplement, and any other supplementation required (such as antibiotics) under sterile conditions. Approximately 20 mL were poured into 90-millimetre (mm) Petri dishes and permitted to set and dry in a laminar flow hood for 30 min, then stored at 4 °C.

2.2.4 7H9 and 7H10 Medium supplement

Medium was supplemented with the following: 0.2% of dextrose and glycerol; bovine serum albumin (0.05%), NaCl (0.08%); and tyloxapol (0.05%) to final concentration. Tyloxapol stock was prepared by weighing out tyloxapol solution (Sigma) and adding an appropriate volume of ddH2O in a conical flask to produce a 20% (v/v) stock solution. This is determined by the assumption that pure tyloxapol has a density of 1.1 g/mL. The solution was stirred for approximately 15 min, sterile filtered, and stored at room temperature.

2.2.5 Terrific broth for *E. coli* FX cloning construct purifications

Terrific broth powder (Sigma) was resuspended as per the manufacturer's instructions. The solution was autoclaved at 121 °C for 20 minutes (min) and permitted to cool. Once cooled, the media was supplemented with KH_2PO_4, K_2HPO_4 and 0.4% glycerol (unless otherwise stated) and the relevant antibiotic (to a final concentration of: ampicillin 100 µg/mL or apramycin 50 µg/mL)

2.2.6 Luria broth (LB)

LB powder (Sigma) Was resuspended in ddH_20 and the solution was autoclaved at 121 °C for 20 minutes (min) and permitted to cool.

2.2.7 Growth curves and Optical density (OD) monitoring

10 mL square bottles were prepared by adding 10 mL 7H9 medium supplemented with the carbon source of choice at a pH of 6.8 and incubate in a static incubator to allow media to reach 37 °C. A pre-culture was prepared by adding 100 µL of bacteria stored at -80 °C to 10 mL of culture medium. Once the pre-culture reached mid-log phase, it was used to start the growth curve. Square bottles were inoculated with mycobacteria (from the pre-culture) with an OD_{600} of 0.001 (for *M. smegmatis* mc^2155) or 0.01 (for *M. tuberculosis* H37Rv strains). These were then incubated at 37 °C in a shaking incubator (180rpm). OD_{600} measurements were taken every two to three days for M. tuberculosis H37Rv strains and three times a day for *M. smegmatis* mc^2155 strains to produce growth curves until the stationary phase.

2.2.8 Bicinchoninic acid assay (BCA)

The total protein concentration of the samples analysed via LC-MS was measured using the BCA assay (ThermoFisher), carried out according to the manufacturer's instructions.

2.3 SDS-PAGE and Western blot analysis

2.3.1 His-tagged protein analysis in mycobacteria

To confirm the expression of Rv1339, all strains of *M. smegmatis* were grown in Middlebrook 7H9 broth, as previously described, for three days in a shaking incubator at 37 °C, 180 rpm. The bacterial cultures were centrifuged at 3000 xg, 4 °C for 10 minutes to harvest the bacterial pellets. The pellets were washed with washing buffer (20 nM Tris, 50 mM NaCl, pH 7.5) and then centrifuged at 3000 xg, 4 °C for 10 minutes to remove the BSA protein. This step was repeated three times. For the last wash, the supernatant was poured off, and 1.5 ml lysis buffer, a washing buffer with protease inhibitor (Sigma-Aldrich, USA), was added to the sample before re-suspending and transferring it to an o-ring tube with acid-washed glass beads. All samples were ribolysed with an MP FastPrep®-24 homogenizer (MP pharmaceuticals) at 6 m/s for 60 seconds, three times. Then, the samples were centrifuged at 11,000 xg for 10 minutes, and the supernatant was transferred to an Eppendorf tube and measured for protein concentration with a Nanodrop Lite Spectrophotometer (Thermo Fisher Scientific) or BCA assay. All the samples were diluted with ddH2O to ensure equal protein concentration and then added with a 5X SDS loading buffer (0.25 M Tris-HCl with pH 6.8, 15% SDS, 50% glycerol, 25% β-mercaptoethanol, and 0.01% bromophenol blue), before being boiled at 100 °C for five minutes. After that, 20 µl of each sample was loaded onto the SDS-PAGE gel to run at 150 V, 40 mA for one hour in a tank filled with SDS running buffer (3% Tris, 14.4% glycine, 1% SDS).

After running the SDS-PAGE, the gel was transferred onto a nitrocellulose membrane (Novex, USA) by running at 20 V, 400 mA for 1 hour. The tank was filled with a transfer buffer containing 10X SDS (3%Tris, 14.4% glycine, 1% SDS), methanol, and ddH2O in a 1:2:7 ratio. The non-specific sites on the nitrocellulose membrane were blocked overnight with 10 ml of 5% skimmed milk (Marvell, UK) dissolved in Tris saline buffer, with Tween 20 pH 7.5 composed of 0.06% Tris base, 0.88% NaCl, 0.1% Tween 20. After being washed with the buffer, the membrane was incubated for one hour with mouse HRP α-His Tag antibody (Biolegend, USA) and developed. It was then washed identically to the previous steps, stripped with ReBlot Mild plus (Sigma) and incubated with an anti - mtbGroEL2 antibody (BEI resources). Washed as previously described and incubated with a further goat anti-mouse antibody conjugated to horseradish peroxidase (1:5000 in 5% milk). This served as the loading control. Rv1339 protein was visualized by FujiFilm LAS-3000 image reader and the Amersham™ ECL™ western blotting analysis system (GE Healthcare, UK).

To confirm expression of the Flag-tagged pKW::rv1339 in *M. smegmatis* mc^2155, the above protocol was used, but instead of the α-His-tag antibody, a α-Flag tag antibody (Sigma, produced in mouse) was used as the primary antibody. Subsequent probing of the mouse antibody was performed with an α-mouse HRP conjugated antibody that was used to develop the blot, with the same protocol as above.

To confirm expression of the step-II tagged Rv1339 from the Cell-Free-Expression Assay (CFE), the above protocol was used with some alterations. A streptactin conjugated HRP construct (IBA Biosciences) was used to probe the presence of the Strep-II tag. This construct was suspended in 5% skimmed milk and ddH_2O. Washes were performed with the Tris-HCL-NaCl-Tween 20 pH 7.6 buffer to remove excess detection construct. Development of the blot was carried out as described above.

To confirm expression of His-tagged Rv1339 constructs in *M. tuberculosis* H37Rv strains, the SDS-PAGE running, transfer and development was as described above. However, the acquisition and preparation of the clear soluble lysate was different. Briefly, *M. tuberculosis* H37Rv cultures were grown to the mid-log phase, transferred to 50ml falcon tubes (Corning) and centrifuged in air-tight sealed buckets at 3000xg for 10 minutes. The bacterial pellet was then washed 3 times in 20mM Tris-HCL, 50mM NaCl buffer, by centrifugation at 3000xg for 10 mins/wash. Bacteria were then resuspended in 1.5ml of 20mM Tris-HCL, 50mM NaCl buffer and mechanically lysed with 0.1 mm acid-washed zirconia beads for 1 min using a FastPrep (MPBio®) set at 6.0 m/s. These lysates were filtered twice through 0.22 μm Spin-X column filters (Costar®) to ensure no surviving bacteria were contained within the samples, and to extract the clear soluble lysate. Protein levels were determined by BCA assay for Western blot normalisation.

2.4 Chromatography

All column purifications of *M. bovis BCG* lysate were performed on ÄKTA systems at 4 ºC. Systems were operated by Unicorn manager software. CaptoQ is an anion exchange column that separates protein across an NaCl gradient. Low concentration of 30mM and a high concentration of 600mM NaCl, both with 20mM Tris-HCL pH 7.5, were used as buffer A and B respectively. Size exclusion chromatography was performed on a Superdex 200 10/300 24ml column (GE Healthcare) with 20mM Tris pH 7.5 and 30mM NaCl. Hydrophobic interaction chromatography was performed with the a HiTrap Phenyl FF (high sub) 5ml column (GE Healthcare). All fractions were kept at 4 ºC with added protease inhibitor (Sigma fast) when in use, or stored at -80˚C. HisTrap FF crude 1ml or 5ml (GE Healthcare) was used for ÄKTA Nickel affinity purifications. Low concentration (Buffer A) consisted of 20mM Tris-HCL pH 7.5, 200mM NaCl (unless otherwise stated) and supplemented with 20mM Imidazole (Sigma). Buffer B was the same as Buffer A except the Imidazole concentration was 250mM. The recommended protocol for each column was used to generate the ÄKTA method for the particular purification. FX cloning and some mycobacterial Nickel affinity purifications were performed with gravity flow columns consisting of 1-2ml Nickel affinity resin and a filter to prevent the column drying out.

2.4.1 Thin Layer Chromatography

Silica gel 60 F254 TLC plates were cut to required size of always less than 15 x 10cm. 50 ml mobile phase consisted of 70:30 Ethanol / H2O 0.2M Ammonium bicarbonate. All revelation of the products was performed on the dried post run plate with 5% phosphomolybdic acid dissolved in Ethanol. A Master pro heat gun was used to reveal the products on the plate.

2.4.2 PDE activity assay

10 µl of cell lysate/fractionated lysate/purified protein was incubated at 37 ºC for 16 hours with 2 µl water, 2 µl 10x PDE buffer (20 mM $MgCl_2$, 20mM Tris-HCL pH 9.0 and 100mM NaCl) and either 6 µl of 100mM (for TLC) or 25mM (for LC-MS) of cyclic AMP or water (as a control). For 5’ - AMP controls, 6 µl of 5’ - AMP was used, at the same concentration of cAMP just described. All PDE reactions were performed in Eppendorf tubes.

2.4.3 Cell lysis

M. bovis BCG – cultured cells were aliquoted into 50 ml falcon tubes under a class II hood and sealed. Cells were then spun down at 3,000 x g. All centrifugations were performed at 4 ºC for 5 minutes. The supernatant was discarded and the pellet resuspended in 20mM Tris (pH 7.5) with protease inhibitor (sigma fast) in preparation to be lysed. Cells were washed 3 times. After the final wash, cells were

aliquoted into 2ml micro tubes with 200 µ of acid washed glass beads and loaded into the Fastprep 24 Ribolyser. Cells were lysed at 6m/s for 30 seconds, 3 times and with 5 minutes rest on ice in between runs. The lysed cells were spun down at 11,000 to 17,000 x g for 10 minutes and the clear cell lysate was removed and aliquoted as 200µLs into Eppendorf tubes. These aliquots were then frozen in liquid nitrogen and kept at -80ºC.

M. smegmatis mc^2155 alternate PDE candidate cultures were either lysed with the protocol outlined in "SDS-PAGE and Western blot analysis" or via sonication on ice at a frequency of 50% over 10 seconds, with 30 seconds break (at 4˚C) in between sonications. Depending on the volume, 5 to 7 rounds of sonication were performed.

All large scale (2.4L+) preparations were lysed via cell disruption with the Constant Systems Cell Disruptor Model T5, to the manufacturer's instructions.

2.5 Cloning and generation of mutants

Isothermal assembly (Gibson) assembly was primarily used for mycobacterial constructs. pVV16 (kanamycin and hygromycin resistance cassettes. ORI in *E. coli* and mycobacteria) and pNIC (kanamycin resistance cassette and ORI in *E. Coli*). Primers were designed for Rv1339 and to open vectors. All insert and vector PCR products were cleaned with Qiagen PCR clean up DNA spin columns and template DNA was degraded by DpnI (NEB) restriction digest. Fragments were joined at 50 ºC for 15 minutes using Hifi builder assembly master mix. DH5-α chemically competent cells were used as a plasmid library strain and plated on LB 1.5 % agar kanamycin (60 µg per ml) plates and incubated at 37 ºC overnight. The remaining cells were spun down and had most of the supernatant removed before being plated again at a high concentration.

Colonies were screened using KAPA2G Robust hotstart ready mix and colony PCR protocol.

Constructs for expression in *E. coli* for purification (Fragment exchange cloning) were generated with the methods described in Geertsma et al (2014)[289]. Primers were generated with the FX cloning primer generator[290].

All validations were performed on 1.5% agarose with a Gel Red^{tm} (Biotium) 1:10000 dilution. They were imaged on a Biodoc-It2 315 imaging system (LM-26 Transilluminator) or LAS-3000 Fuji Imager under UV.

2.5.1 Proteomic analysis

The fraction displaying PDE activity as seen by TLC, after CaptoQ and SEC column purifications, and 1 fraction either side were run on an SDS gel as described previously and stained with Instant blue^{tm} Coomassie stain. These gel lanes were cut, reduced, alkylated and digested with trypsin using the protease max surfactant protocol. Samples were then loaded and run on a SYNAPT Q-tof mass spec. Proteins seen were correlated to protein IDs and potential annotations in Uniprot database. The proteins in each fraction were then compared using Venny software. Intact protein mass spec was also performed with this MS and utilised desalting of the samples with Valerian Ziptips then loading on the MALDI target with sinapinic acid matrix suspended in 50% ACN: ddH_2O.

2.5.2 Fx cloning expression trials and preparations

MC1061 Mutants of Fx cloning generated constructs were sub-cultured from cryostock in 10ml volume for at least 6 hours. Cultures were inoculated at 0.001 and left for 2 hours at 37°C in a shaking incubator (160 RPM). The cultures were then induced and either left for a further 4 hours at 37°C or overnight at 25°C. Cultures were then spun down and suspended in 20mM Tris pH 7.5, 50mM NaCl, 20mM MgCl2

and EDTA-free protease inhibitor for lysis. Once lysed, soluble fractions were isolated by centrifugation and loaded onto SDS-PAGE gels for Western blot analysis, as previously described.

2.5.3: Mycobacterial constructs generated as part of this work

Table 4: Mycobacterial cloning constructs

Mycobacterial constructs	kDa
pVV16::rv1339 Full length C-H6	30218.18
pVV16::rv1339 Short C-H6	28532.15
pVV16::rv1339 D180A C-H6	30174.17
pKW::rv1339 Full length C-Flag tag	30148.97
pMINT::rv1339 Full length C-H10 GFP	58526.85
pMINT::rv1339S C-H10 GFP	56654.61
pMINT::ms4902 C-H10 GFP	56352.18
pMex::rv1339 Full length C-H10 GFP	58526.85
pMex::rv1339S C-H10 GFP	56654.61
pMex::ms4902 C-H10 GFP	56352.18
pACE::rv1339 Full length C-H10 GFP	58526.85
pACE::rv1339S C-H10 GFP	56654.61
pACE::ms4902 C-H10 GFP	56352.18
pYUB::rv1339 Full length C-StrepII	30.72867

2.5.4: *E. coli* constructs generated as part of this work

Table 5: *E. coli* cloning constructs

***E. coli* constructs**	**Da**
pNIC::rv1339 Full length N-H6	30218.18
pBXC3H::rv1339 Full length CH-10	31567.61
pBXNH3:: rv1339 Full length NH-10	31567.61
pBXC3GH Full length CH-10	58658.04
pBXPHM3 Full length CH-10 PelB MBP	75625.80
pBXC3H::rv1339 Short CH-10	29695.37
pBXC3H::ms4902 CH-10	29392.95
pBXNH3:: rv1339 Short NH-10	29695.37
pBXNH3:: ms4902 NH-10	29392.95
pBXC3GH:: rv1339 Short C-H10	73451.13
pBXC3GH:: ms4902 C-H10	56483.38
pBXPHM3:: rv1339 Short C-H10 PelB MBP	73753.56
pBXPHM3:: ms4902 C-H10 PelB MBP	73451.13

***E. coli* constructs**	**kDa**
pNIC::rv1339 Full length N-H6	**30218.18**
pBXC3H::rv1339 Full length CH-10	**31567.61**
pBXNH3:: rv1339 Full length NH-10	**31567.61**
pBXC3GH Full length CH-10	**58658.04**
pBXPHM3 Full length CH-10 PelB MBP	**75625.8**
pBXC3H::rv1339 Short CH-10	**29695.37**
pBXC3H::ms4902 CH-10	**29392.95**
pBXNH3:: rv1339 Short NH-10	**29695.37**
pBXNH3:: ms4902 NH-10	**29392.95**
pBXC3GH:: rv1339 Short C-H10	**73451.13**
pBXC3GH:: ms4902 C-H10	**56483.38**
pBXPHM3:: rv1339 Short C-H10 PelB MBP	**73753.56**
pBXPHM3:: ms4902 C-H10 PelB MBP	**73451.13**

2.5.5 Primers used during this work

Gibson assembly	
open_pVV16_for	AAGCTTCACCACCACCACCACCACTGACAG
open_pVV16_rev	CAT ATG GAA GTG ATT CCT CCG GAT CGG GGA TG
Rv1339L_pVV16_f	CCGATCCGGAGGAATCACTTCCATATGCGTCGATGTATTCCGCATCGTT
Rv1339S_pVV16_f	CCGATCCGGAGGAATCACTTCCATATGTCCGTGCGGATAACCGTGCTC
pYUBNTH_pVV16 _f	CCGATCCGGAGGAATCACTTCCATATGGGCAGCAGCCATCATCATCATC
open_pYUB_CT_rev	CATGGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAG
Rv1339S-pYUBCT_for	AACTTTAAGAAGGAGATATACCATG TCCGTGCGGATAACCGTGCTCG
Rv1339L-pYUBCT_for	AACTTTAAGAAGGAGATATACCATG CGT CGA TGT ATT CCG CAT CGT TGT ATC
Rv1339_pVV16_rev	GTCAGTGGTGGTGGTGGTGGTGAAGCTTGCCGGCTCGCCGGACTTCG
open_pYUB_for	GAA TTC GAG CTC CGT CGA CAA GC
open_pYUB_rev	GCT GTG ATG ATG ATG ATG ATG GCT GC
Rv1339L_pVV16-NTF_f	AAAGACGACGACGACAAGGACGATGATGACAAGATGCGTCGATGTATTCC GCATCGTT
Rv1339S_pVV16-NTF_f	TACAAAGACGACGACGACAAGGACGATGATGACAAGATGTCCGTGCGGAT AACCGTGCTC
Rv1339_stop_pVV16_r	GCTTCGAATTCTGCAGCTGGATCCCTAGCCGGCTCGCCGGAC
pVV16_CT_HISFLAG_for	CGAAGCTTCACCACCACCACCACCACGATTACAAAGACGACGACGACAAG
pVV16_CT_HISFLAG_rev	GTTAACTACGTCGACATCGATAAGCTGTCACTTGTCGTCGTCGTCTTTGTAAT CG
open_pKW_rev	AACATTTCCTCCGGATCCTGTCAGGATTC
Rv1339L-pKW_for	CAGGATCCGGAGGAAATGTTATGCGTCGATGTATTCCGCATCGTT
insert_Strep_pYUB_for	GGTCTCACCCGCAGTTCGAAAAATAATCTAAGAATTCGAGCTCCGTCGACAA GCTTG
insert_Strep_pYUB_rev	TTCGAACTGCGGGTGAGACCAGGAACCAGAACCAGAGCCGGCTCGCCGGA CTTCGAAC
Alternate PDE candidates	
Rv0574for	GCCATATGGCTGGCAATCCTGATGT
Rv0574rev	GCAAGCTTCTCCTTGCTCGTTAGGT
Rv1277for	GCCATATGAGTCCGCGCCCGGGACC
Rv1277rev	GCAAGCTTCGCCGCTCCCCGGTCAG
Rv2577for	GCCATATGGGCGCCGATCTGAAGCA
Rv2577rev	GCAAGCTTTCCGCCGCGCGGCTTGG
Rv2795for	GCCATATGACCTGGAAAGGATCGGG
Rv2795rev	GCAAGCTTTCGAGACTGCCGTTGCC
Rv3683 for1	GCTAGCATGGAGGTGACCATGCCAGT
Rv3683 for2	GCTAGCATGCCAGTCTTGACTCCGGG
Rv3683 rev	GCAAGCTTACGCACCGACACTGTCG

2.6 Phenotypic characterisation techniques

2.6.1 Minimal Inhibitory Concentration by Resazurin Microtiter Assay (REMA)

Minimal inhibitory concentration (MIC_{90}) was carried out by the REMA assay, according to the NCCLS guidelines and Palomino et al. (2002)[291,292]. Antibiotic stock solutions of the tested compounds were prepared to yield target concentrations for testing as seen in Table 1. Microdilution assay was performed in 96-well plates. The 10-fold serial dilutions were used to obtain the final drug concentration, ranging from 0.78-200 µg/ml (rifampicin), 0.016-4 µg/ml (ethambutol), 0.78-200 µg/ml (isoniazid), and 0.008-2 µg/ml (streptomycin). For bacterial inoculum, wild-type and Rv1339 strain of *M. smegmatis* were grown to the mid-log phase (OD600 ~ 0.5-1) and diluted 1:1000 with 7H9 broth. 50 µl of the standardized bacterial inoculum and the 50µl 7H9 broth was added to each well of a 96-well plate. The plates were incubated for 48 hours at 37°C with 5% CO_2. Then 30 µl of 0.01% resazurin was added. Wells were read after 24 hours for colour development. The MIC_{90} was defined as the lowest concentration that inhibited the growth of 90% of bacteria, and two independent experiments were performed for this experiment.

Table 6: List of antibiotics used and their preparation

Antibiotics	**Potency (%)**	**Solubilisation solvents**	**Concentration Range (µg/mL)**
Ethambutol	73.0	ddH_2O	16-0.063
Streptomycin	73.3	ddH_2O	2-0.008
Rifampicin	97	Methanol	400-1.563

2.6.2 Time-to-kill Curves

The wild-type and Rv1339 strains of *M. smegmatis* were grown to the mid-log phase (OD_{600} ~ 0.5) and diluted 1:100, corresponding to ~10^5-10^6 CFU/ml, in the Middlebrook 7H9 medium. Antibiotics were added to each sample at defined concentrations and the bacterial samples were collected, before and after adding the antibiotic, at 6, 24, 30, 48 and 72 hours after antibiotic challenge. To perform colony forming unit (CFU), serial 10-fold dilutions were made using 20 µl culture and 180 µl of the 7H9 broth. 20 µl of each dilution was plated onto the Middlebrook 7H10 agar (Sigma-Aldrich, Germany). All plates were incubated at 37 °C and counted on day 4.

2.6.3 Hoechst Staining

Mycobacterial strains were grown in complete 7H9 medium until exponential phase and diluted to final OD_{600} ~ 0.8 with the same medium to a total volume of 5 ml. Bacteria suspension was centrifuged for 10 minutes at 4000 g, 4°C and supernatant was removed. The pellet was washed two times with sterile cold PBS by resuspending and centrifuge at 4,000 x g, 4°C for 25 minutes and removing supernatant. The pellet was resuspended in 5 mL complete 7H9 medium. White-bottomed 96-well plates were prepared firstly by adding 200 µl 7H9 medium to the perimeter to minimise medium evaporation. Then 180 µL of control and ion-treated bacteria suspension was loaded into separate columns of the plate. Blank columns were prepared by adding the same volume of corresponding media. Hoechst 33342 stain solution was prepared to a final concentration of 25 µM with sterile deionised H_2O. 20 µL of stain solution was loaded into each well of the plate, with loading order switching between two groups to avoid loading bias. The changes in fluorescence of each well was monitored in a microwell plate reader (Hidex Sense) at 37°C with 600 rpm shaking, and fluorescence was measured with excitation and emission filters of 355 nm and 460 nm every 10 minutes over 180 minutes, representing a doubling cycle *M. smegmatis* mc^2155 Fluorescence data was exported into Microsoft Excel to quantify changes in fluorescence levels and compare between control and ion-treated groups.

2.6.4 Crystal violet spotting assay

To evaluate the cell membrane permeability of bacteria, the spotted assay is performed according to these papers [293–295]. *Mycobacterium smegmatis mc^2155 - (1) MSMEG pVV16, (2) MSMEG pVV16::rv1339* were grown in Middlebrook 7H9 broth (Sigma-Aldrich, Germany) supplemented with 0.2% W/V dextrose (Thermo Fisher Scientific, USA), 0.2% W/V glycerol (VWR, Belgium) to mid-log phase to obtain OD_{600} around 0.5 and serial dilution was performed (undiluted to 10^{-7}). 5 µl of each dilution was spotted on Middlebrook 7H10 agar (Sigma-Aldrich, Germany) containing 1**0**µg/ml crystal violet, 15 µg/ml malachite green, or 0**.03%** SDS. The agar plates were incubated at 37 °C for **3** days and then were photographed.

2.6.5 Intracellular ATP Measurement

To measure intracellular ATP content, the procedure was performed according to Koul et al. [3] and Parrish et al. [4]. Bacterial samples from different carbon sources, dextrose, and glycerol were taken at different growth phases (the lag phase, log phase, and the beginning of the stationary phase). Bacterial cells were harvested by collecting 1.5 ml of bacterial culture, followed by centrifugation at 17,000 x g for 15 minutes. After that, the pellet samples were kept in -80 °C until the analysis. Each bacterial pellet was mixed with 1.5 ml boiling Tris-EDTA (100 mM Tris, 4 mM EDTA, pH 7.75) and glass beads. All samples were lysed using the MP FastPrep®-24 homogenizer (MP pharmaceuticals) at 6 m/s for 60 seconds, two times. Then, the samples were heated at 100 °C for five minutes and immediately cooled

down on ice for 10 minutes. Cell debris was removed by microcentrifugation at 13,000 x g for 15 minutes, and the supernatant was transferred to new sterile tubes for ATP assay. The ATP content was determined using the ATP Bioluminescence Assay Kit CLS II (Roche), and 50 µL of luciferase reagent was added to 50 µL of the supernatant sample. The reactions were measured in a microwell plate-reading luminometer (Hidex Sense). To obtain the ATP amount per CFU, the bacterial culture was serially diluted and plated on Middlebrook 7H10 agar. The agar plates were incubated at 37 °C and colonies counted on day 4 after plating.

2.6.7 RNA extraction

M. smegmatis mc^{2}155 or *M. tuberculosis* H37Rv strains were grown in 7H9 liquid medium at 37°C as described above. Bacteria were to mid-log phase (light transmittance at 600 nm ~0.6), washed twice with cold PBS at +4°C. Total RNA was extracted according to the FastRNA Pro Blue kit manual (MP). Total RNA was extracted using a fast RNA pro bluekit (MP Biomedicals) 2 times (30 sec/6.0). To remove genomic DNA, the samples were treated once with RNase-free DNase (Promega) for 1 h at 37°C and purified using RNAeasy columns (Qiagen) according to the manufacturer's instructions.

2.6.8 RNA-seq data analysis

Sequence quality was checked using FastQC (v0.11.8; [296]). All sequences passed quality control and paired-end sequences were aligned to the mc^2-155 *M. smegmatis* reference genome (Release 3, Mycobrowser[297]; using Burrows-Wheeler Transform (BWA) sequence aligner (v0.7.17-r1188[298]). Read counts were calculated at the gene level using feature Counts (v1.6.0[299]) using the mc^2-155 *M. smegmatis* annotation (Release 3, Mycobrowser). Sequences were also aligned to the H37Rv *M. tuberculosis* reference genome (GCF_000195955.2_ASM19595v2 assembly, NCBI[300]) and annotated using the H37Rv *M. tuberculosis* annotation (GCF_000195955.2_ASM19595v2 assembly, NCBI[300]) to determine levels of *rv1339* in all samples. Genes with low counts across all samples were removed before sample quality was evaluated through principle component analysis (PCA) using the prcomp function in base R (v3.6.2, R Core Team, 2017[301]). Differential gene expression analysis was performed using the DESeq2 Bioconductor package (v1.26.0[302]), comparing the Rv1339 Expressor with the Empty Vector Control. Genes were corrected for multiple testing using the Benjamini-Hochberg (BH) method and significant genes were identified using a False-Discovery Rate (FDR) cut-off of < 0.05 and a log2 fold change of absolute 1.5. Pathway analysis was performed on differentially expressed genes using DAVID (v6.8[303,304]).

2.6.9 Seahorse XFP analysis of *M. smegmatis* mc^{2}155

Seahorse XFP analysis was performed similarly to previous studies[305–307]. All reagents were purchased from Agilent Technologies and all work was performed in a laminar flow hood. Bacterial cell plates or XFP cartridges were transported to the machine or incubator in airtight containers. The day before an assay was to be run, the Seahorse XFP cartridge probes were hydrated by filling the utility plate with 200 µL of sterile ddH_2O per well and in all border wells. This utility plate-cartridge unit was then incubated in an airtight container overnight.

A minimum of 2 hours before the assay was run, the ddH_2O was removed and replaced with XFP calibrant solution and the cartridge was returned to the 37˚C incubator. 20µL of each substance to be injected was then placed in the relevant injection port, by taking care to ensure none was stuck on the sides of the injection port. In the case of the assays presented here, the only injection was dextrose-glycerol to a final concentration of 0.2%. Separately, each well of the bacterial cell plate was coated with 90 µl of sterile Poly-D-Lysin (PDL), and left to air dry in a sealed laminar flow hood overnight. On the day of the assay, the residual PDL was removed and the wells washed with 90µL of ddH_20. The water was then removed and the plate left to airdry in the laminar flow hood with the lid of. Bacteria were cultured in normal 7H9 media (with BSA, NaCl, glycerol and dextrose supplement. to mid-log phase and then the cultures were diluted to the equivalent of 1ml of OD_{600} 0.51 (this quantity of bacteria was determined to be within the reliable working range of the Seahorse XFP instrument by extensive optimisation). These samples were then centrifuged for 10 minutes at 15,000 x g to pellet the bacteria before the supernatant was discarded. The bacteria were then resuspended in unbuffered 7H9 and centrifuged for 7 minutes. The wash was discarded and the remaining bacteria were resuspended in 1 ml of new unbuffered 7H9. 90µL of well mixed bacterial solution was deposited in each well of the Seahorse XFP bacteria cell plate, the plate was then centrifuged at 2, 000 x g for 10 minutes. 90µL of unbuffered 7H9 was then gently added dropwise to increase the volume to 180µL.

After the minimum 2 hours of calibrant incubation in the utility plate-cartridge unit, the unit was then ready for calibration in the Seahorse XFP instrument. Once calibrated, the utility plate was ejected and could be replaced with the bacterial cell plate. The instrument would then return to 37˚C and then begin the assay. Each measurement cycle requires 4 minutes of mixing and 2 minutes of sensor measurement. The assay consisted of 3 measurement cycles, in the absence of carbon source. Then 20µL of dextrose-glycerol was injected to a final concentration of 0.2% and 12 further measurements were taken. Bacteria counts were normalised by CFU counts of the different bacterial strains used to inoculate each well. Data were normalised to 10^5 CFU.

2.7 LC-MS and metabolomics

Metabolite extraction experiments

For targeted metabolomic profiling studies, mycobacteria were initially grown in 7H9 liquid medium containing the carbon sources of interest until the OD_{600} reached ~0.8-1. Bacteria were then inoculated onto 0.22 μm nitrocellulose filters under vacuum filtration. Mycobacterial-laden filters were then placed on top of chemically equivalent agar media (described above) and allowed to grow at 37°C for 5 doubling times to generate enough biomass for targeted metabolomics studies. Filters were then transferred into 7H10 plates supplemented with 0.5 g/l fraction V bovine serum albumin, 0.2% dextrose and 0.2% glycerol, 10 mM NaCl. Bacteria were metabolically quenched by plunging the filters into the extraction solution composed of acetonitrile/methanol/H_2O (2:2:1) pre-cooled to 4°C. Small molecules were extracted by mechanical lysis of the entire bacteria-containing solution with 0.1 mm acid-washed zirconia beads for 1 min using a FastPrep (MPBio®) set at 6.0 m/s. Lysates were filtered twice through 0.22 μm Spin-X column filters (Costar®). Bacterial biomass of individual samples was determined by measuring the residual protein content of the metabolite extracts using the BCA assay kit (Thermo®) [9, 27]. A 100 μL aliquot of the metabolite solution was then mixed with 100 μL of acetonitrile with 0.2 % acetic acid at -20°C, and centrifuged for 10 mins at 17,000xg at 4°C. The final concentration of 70% acetonitrile was compatible with the starting conditions of the HILIC chromatography. The supernatant was then transferred into an LC/MS V-shape vials (Agilent 5188-2788) and 4 μL was injected into the LC/MS.

2.7.1 Liquid-chromatography-mass spectrometry for targeted and untargeted metabolomics

Aqueous normal phase liquid chromatography was performed using an Agilent 1290 Infinity II LC system equipped with a binary pump, temperature-controlled auto-sampler (set at 4°C) and temperature-controlled column compartment (set at 25°C) containing a Cogent Diamond Hydride Type C silica column (150 mm × 2.1 mm; dead volume 315 μL). A flow rate of 0.4 mL/min was used. Elution of polar metabolites was carried out using solvent A consisting of deionized water (resistivity ~18 MΩ cm) and 0.2% acetic acid and solvent B consisting of 0.2% acetic acid in acetonitrile. The following gradient was used: 0 min 85% B; 0-2 min 85% B; 3-5 min to 80% B; 6-7 min 75% B; 8-9 min 70% B; 10-11 min 50% B; 11.1-14 min 20% B; 14.1-25 min hold 20% B followed by a 5 min re-equilibration period at 85% B at a flow rate of 0.4 ml/min. Accurate mass spectrometry was carried out using an Agilent Accurate Mass 6545 QTOF apparatus. Dynamic mass axis calibration was achieved by continuous infusion, post-chromatography, of a reference mass solution using an isocratic pump connected to an ESI ionization source operated in the positive-ion mode. The nozzle voltage and fragmentor voltage were set at 2,000 V and 100 V, respectively. The nebulizer pressure was set at 50 psig, and the nitrogen drying gas flow rate was set at 5 l/min. The drying gas temperature was

maintained at 300°C. The MS acquisition rate was 1.5 spectra/sec, and *m*/*z* data ranging from 50-1,200 were stored. This instrument enabled accurate mass spectral measurements with an error of less than 5 parts-per-million (ppm), mass resolution ranging from 10,000-45,000 over the *m*/*z* range of 121-955 atomic mass units, and a 100,000-fold dynamic range with picomolar sensitivity. The data were collected in the centroid 4 GHz (extended dynamic range) mode. Detected *m*/*z* were deemed to be identified metabolites based on unique accurate mass-retention time and MS/MS fragmentation identifiers for masses exhibiting the expected distribution of accompanying isotopomers. Typical variation in abundance for most of the metabolites remained between 5 and 10% under these experimental conditions.

2.7.2 LC/Q-TOF data acquisition using HILIC-Z chromatography

Data was acquired on an Agilent 1290 Infinity II UHPLC coupled to a 6546 LC/Q-TOF system. Chromatographic separation was performed on an Agilent InfinityLab Poroshell 120 HILIC-Z, 2.1 × 100 mm, 2.7 µm (p/n 675775-924) column. The HILIC methodology was optimized for polar acidic metabolites (details in Table 1). For easy and consistent mobile phase preparation, a concentrated 10x solution consisting of 100 mM ammonium acetate pH 9.0 in water was prepared to produce mobile phase A and B.

Table 7: LC-MS parameters

LC Conditions			
Column	InfinityLab Poroshell 120 HILIC-Z, 2.1 × 100 mm, 2.7 µm (p/n 675775-924)		
Mobile Phase	A) 10 mM ammonium acetate in water, pH 9 with 5 µm InfinityLab Deactivator Additive (p/n 5191-4506) B) 10 mM ammonium acetate, pH 9 in 10:90 (v:v)water/acetonitrile		
Flow Rate	0.5 mL/min		
Gradient	Time (minutes)	%A	%B
	0	0	100
	11.5	30	70
	12	0	100
	15	0	100
	20	Post Time	
Column Temperature	30°C		
Injection Volume	1 µL		
Multisampler Temperature	6°C		
MS Conditions			
MS System	6545 LC/Q-TOF		
SWARM autotune selection	50-750 *m/z* fragile ion mode		
Ionization Source	Agilent Jet Stream		
Polarity	Negative		
Gas Temperature	200°C		
Drying Gas Flow	10 L/min		
Nebulizer Pressure	40 psig		
Sheath Gas Temperature	300°C		

Sheath Gas Flow	12 L/min
Capillary Voltage	3000 V
Nozzle Voltage	0 V
Fragmentor	115
Acquisition Range	40-1000 *m/z*
Reference Mass	68.995758 & 980.016375

U-^{13}C-Labelling analysis

Under the experimental conditions described above using [U-$^{13}C_3$] glycerol (99%) and [U-$^{13}C_6$] glucose (99%), the extent of ^{13}C labelling for each metabolite was determined by dividing the summed peak height ion intensities of all ^{13}C-labelled species by the ion intensity of both labelled and unlabeled species using the software Agilent Profinder version B.8.0.00 service pack 3.

2.8 Statistical analysis

Metabolomics, Membrane Potential, ATP and NAD/NADH data are presented as the mean ± SD from at least 2 biological replicates and at least 3-9 technical replicates per condition. Unpaired two-tailed Student's *t*-tests were used to compare values, with $p < 0.05$ considered significant.

Growth Curve, Time-to-Kill and Seahorse XFP data are presented as mean ± standard error of the mean from at least 2 biological replicates and at least 2-4 technical replicates. 2-way ANOVA with Bonferroni multiple comparison test used and $p< 0.05$ considered significant. Prism statistics analysis software was used.

Chapter 3 - Identification of Rv1339, a conserved actinobacteria enzyme with putative cAMP PDE hydrolase activity

3.0 Identification of Rv1339, a conserved actinobacteria enzyme with putative cAMP PDE hydrolase activity

3.1 Preliminary work on the project

Prior to the start of my PhD, tools were developed to investigate and validate the presence of an alternate cAMP PDE in *M. tuberculosis.* This work was undertaken by our research group, will be reviewed in brief, and the respective researchers credited in the figure legends.

3.2 *M. bovis* BCG displays cAMP PDE activity even in the absence of Rv0805, the only currently annotated cAMP PDE enzyme

Rv0805 was discovered to be an *M.* tuberculosis H37Rv cNMP PDE enzyme, 15 years ago[205,279]. When expressed in *M. smegmatis* mc^2155 under its native *M. bovis* BCG promotor, Rv0805 is capable of mediating a 50% decrease in cAMP levels[281]. When overexpressed in *M. tuberculosis* H37Rv, it only elicits a 30% decrease in cAMP levels[308]. Rv0805 has also been shown to display catalysis independent functions via mutational studies, moonlighting as a regulator of cell wall permeability[281] and gene expression[308].

In addition, it is only expressed in pathogenic mycobacteria[205], and as mentioned previously in the introduction Chapter, a study by Lee[286] has shown that non-pathogenic mycobacteria also display cAMP PDE activity[286]. Beyond these constraints, *in-vitro* characterisation of Rv0805 suggests that it has a 150-fold higher affinity for the DNA degradation product 2', 3'-cAMP[287]. When compared to 3', 5'- cAMP PDE of *E. coli* – Rv0805 has a surprisingly low V_{max} (30 nmol/sec/mg vs. *E. coli*[194] CpdA 2µmol/sec/mg[279]) considering mycobacteria have 10-100 fold higher cAMP levels[174,175,204].

In sum, this suggests that there is at least 1 other cAMP PDE enzyme in mycobacteria, and that it is likely to be conserved across pathogenic and non-pathogenic mycobacteria.

Identification of an enzyme with cAMP PDE activity would require an effective *in-vitro* read-out of cAMP hydrolysis. An approach found in the literature involved visualisation and semi-quantification 5' AMP levels with Thin-layer chromatography (TLC), when a candidate protein or bacterial lysate was incubated with a cAMP PDE activity buffer and cAMP substrate[309]. To identify cAMP activity other than that of Rv0805, a series of buffers similar to those that were described in the paper that identified cAMP PDE activity in *M. smegmatis* mc^2155[286] were tested. The buffer that consists of 20mM Tris-HCL pH 9.0, 10mM $MgCl_2$, and 100mM NaCl appeared to mediate the most activity when used in the assay with bacterial lysate (Data not shown). The activity assays were run at 37˚C as this would be the standard temperature that *M. tuberculosis* bacilli would be exposed to during human host infection. 16-hours was chosen by default as the duration the assay was allowed to run. (Data not shown, credit: Ipsita Mahapatra).

In order to confirm the presence of an alternative cAMP PDE in mycobacteria, parental *M. bovis* BCG, a clean knock-out mutant of the *M. bovis* BCG Rv0805 homologue (ΔMb0828)

and complemented strains were assayed for cAMP PDE activity (strain provided by Dr Roger Buxton, National Institute for Medical Research, Figure 3.1) (the *M. tuberculosis* H37Rv gene name will be used for clarity). To confirm that this activity was indeed an active process mediated by an enzyme and not just due to degradation of the substrate, the assay was also run with the heat inactivated lysate of the 3 strains. In all strains, when cAMP was present, 5' AMP production was detected, and this activity was not seen when the lysates were heat inactivated. This indicates that even in the absence of Rv0805, *M. bovis* BCG lysate still retains cAMP PDE activity, thereby confirming the presence of another cAMP PDE enzyme (Figure 3.1).

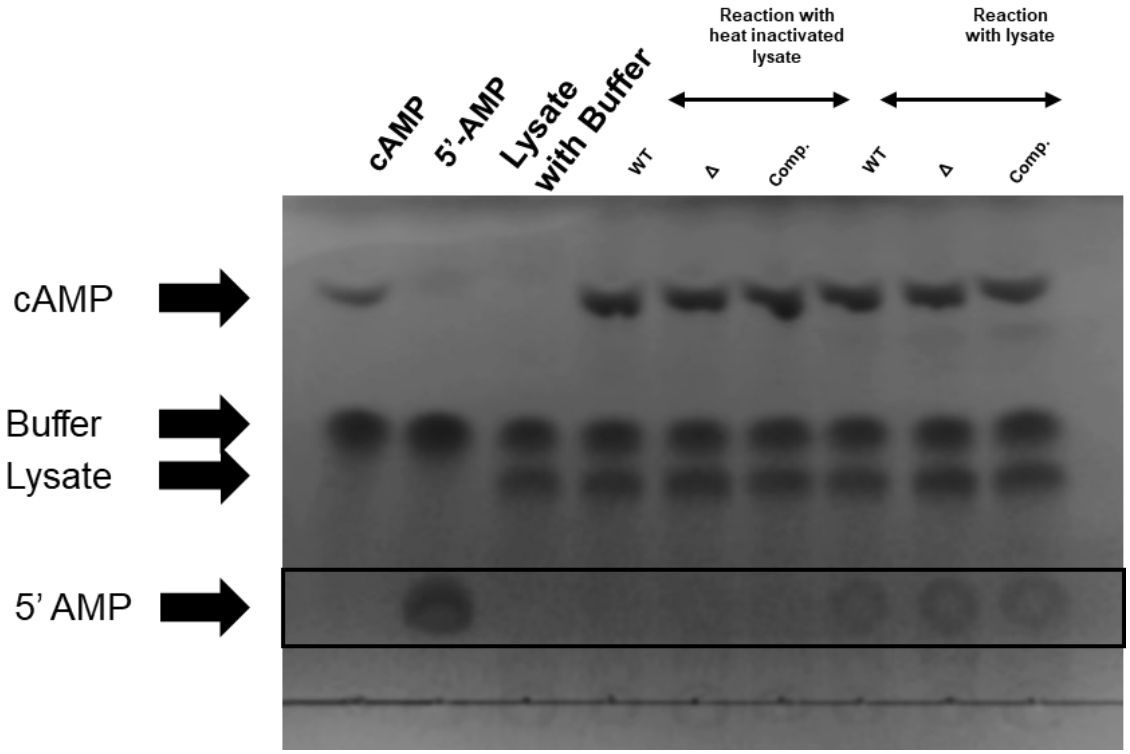
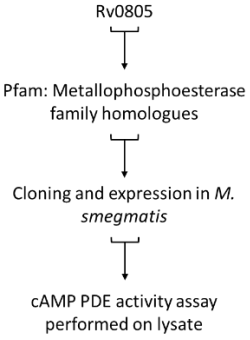


Figure 3.1: A Thin layer chromatography visualisation of the cAMP PDE assay with M. bovis BCG lysates. Heat-inactivated and native lysates were taken from M. bovis BCG parental, ΔRv0805 and Complemented stains and incubated with PDE buffer and cAMP. Lysates with cAMP PDE activity will be able to produce 5' AMP (the hydrolysis product of cAMP). Controls consist of cAMP, 5' AMP and WT lysate and buffer. It is clear that even in the absence of Rv0805, the M. tuberculosis lysate still displays cAMP PDE activity. This activity is ablated in all strains under heat inactivation and therefore is an active process mediated by enzymes. Credit: Dr. Gerald Larrouy-Maumus.

3.3 Screening for other cAMP PDE enzymes via homology to Rv0805 in *M. bovis* BCG

In order to identify the alternative cAMP PDE enzyme to Rv0805 in the *M. tuberculosis* H37Rv genome, the primary amino acid sequence of Rv0805 was used as a starting point. First, a BLAST*p* search was performed against the *M. tuberculosis* H37Rv genome, using Rv0805 as the template. No hits were identified, suggesting that no homologues of Rv0805 could be found based on primary amino acid sequence. Nevertheless, as previously mentioned, Rv0805 contains a metallophospoesterase domain (MPE; Pfam accession ID PF00149). By investigating the presence of this conserved domain in *M. tuberculosis* H37Rv genome instead of primary amino acid sequences similarity, it was found that *M. tuberculosis* H37Rv encodes 6 proteins with a metallophospoesterase domain (Figure 3.2). Sequence alignment of the 6 proteins did not display any sequence similarities (Figure 3.2C). This demonstrates that even if the primary amino acid sequence is not conserved, based on Pfam, the domain is conserved and that a BLAST*p* search alone could not have been used to identify these PDE candidates (Figure 3.2C).

A



B

H37Rv Gene number	Evalue
Rv0574	7.59e-32
Rv0805	2.19e-55
Rv1277	8.2e-38
Rv2577	1.05e-44
Rv2795	8.33e-25
Rv3683	8.07e-27

C

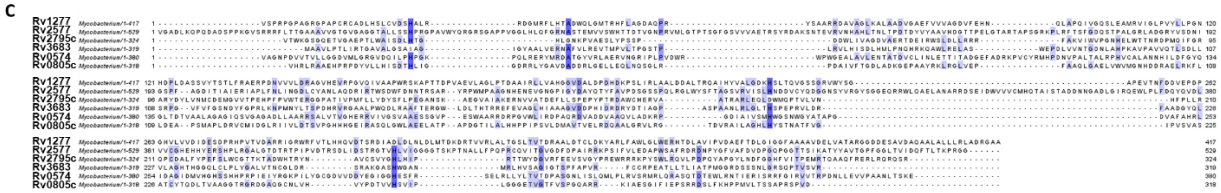


Figure 3.2: Metallophosphoesterase enzymes in M. tuberculosis H37Rv. 2A: Workflow for identification of Rv0805 homologues. 2B: Table of the Evalues of the different homologues for similarity to Rv0805 from Pfam. 2C: Sequence alignment of the Rv0805 homologues. Modified from the work of Dr. Gerald Larrouy-Maumus.

Our research group was able to clone and express 4 of these homologues, and Rv0805 successfully in *M. smegmatis* mc^{2}155 (Figure 3.3A and 3.3B). It is worth noting that Rv2795, Rv3683 and Rv1277 have homologues in *M. smegmatis.* The Rv3683 protein was predicted to consist of a soluble portion and a transmembrane domain. It was therefore decided that the full-length and soluble portions would be expressed and purified separately. The different annotations (Table 3A) were based on similarity to proteins that had been already characterised at that time. For example, Rv0805 was found to be structurally similar to the metallophosphoestarase family protein: glycerolphosphodiesterase from *Enterobacter aerogenes* GpdQ[310]. Rv02795 was also predicted in the database to be similar to GpdQ – but was uncharacterised in *M. tuberculosis*. Rv0574 was also poorly characterised, beyond the fact that it was predicted to be in the DosR regulon[311] and appeared to immunogenic in multiple *M. tuberculosis* antigen studies[311,312]. Rv3683 displayed similarity to YfcE, an enzyme with phosphodiesterase activity characterised in *E. coli*[313]. However, Rv3683 was completely uncharacterised in *M. tuberculosis*.

The expression vector backbone used, contains an HSP-60 constitutively active promotor and a C-terminal 6 Histidine tag (Vector: pVV16 – BEI resources) (Figure 3.3B) in the non-pathogenic model mycobacteria, *M. smegmatis* mc^{2}155. To assess the expression of these homologues, SDS-PAGE and Western blot analysis using an α-His antibody for the His-tag were performed (Figure 3.3B). It appeared that there was an additional band indicating degradation of the expressed full-length Rv0574 (Figure 3.3B). This protein may be cleaved *in-vivo* or alternatively the lysis process could have led to the degradation.

The cAMP PDE activity assay was performed with the *M. smegmatis* mc^{2}155 lysates expressing the Rv0805 homologues or Rv0805. LC-MS analysis was used to quantify the ion counts of intracellular cAMP levels and normalised to the amount of proteins in the sample. As seen in figure 3.3C Rv0805 displays significant cAMP PDE activity ($p<0.05$) in this model (Figure 3.3C). The 30% decrease in cAMP levels observed in the Rv0805 expressing lysate was comparable to that found in the literature[279]. However, none of the Rv0805 homologues showed cAMP PDE activity. This may indicate that they are not cAMP PDE enzymes. Alternatively, the proteins may have been expressed but incorrectly folded or the conditions in the *M. smegmatis* mc^{2}155 lysate/the PDE activity assay buffer may not have been optimal for these proteins (e.g. pH or a different divalent cation).

A

ID	Length Amino acids	Molecular Weight (kDa)	E-value	Region	Family	Function
Rv0574c	380	41.5	$7.59e^{-32}$	8-19,60-312	5'-nucleotidase (syn. UDP-sugar hydrolase), N-terminal domaine	Unknown
Rv0805	318	34.2	$2.19e^{-55}$	12-264	GpdQ-like	cAMP-PDE
Rv2795c	324	37.5	$8.33e^{-25}$	14-268	GpdQ-like	Unknown
Rv3683	319	34.3	$8.07e^{-27}$	47-290	YfcE-like	Unknown
Rv3683tr	273	30.1	$8.07e^{-27}$	47-290	YfcE-like	Unknown

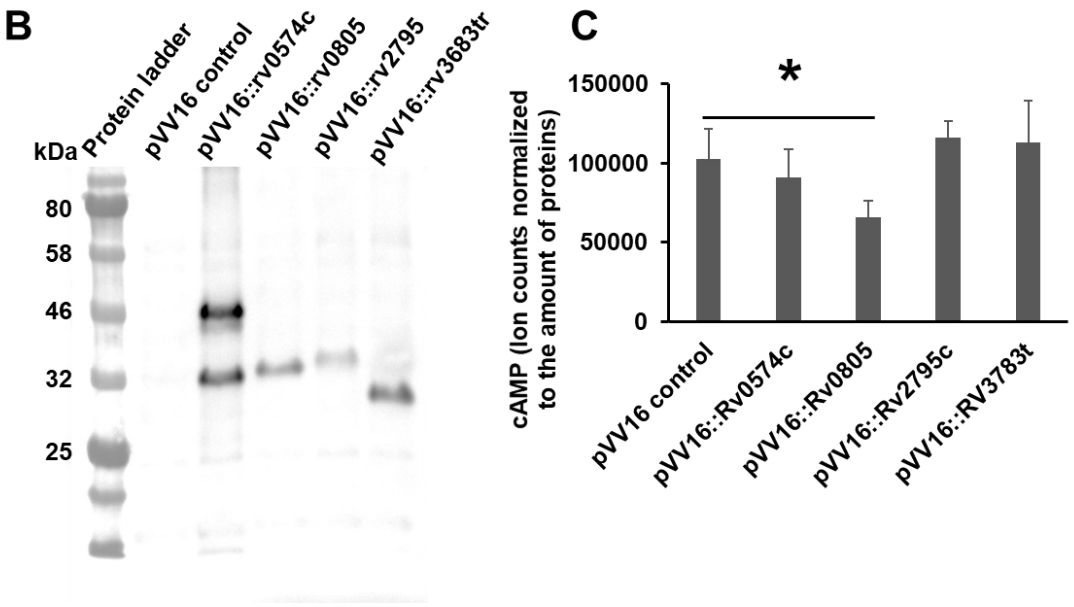


Figure 3.3: A: Table detailing the expressed alternate cAMP PDE candidates derived from Rv0805 homology. B: Western blot of clear soluble lysate of M. smegmatis mc²155 strains expressing the different potential cAMP PDE candidates. C: LC-MS quantification of the cAMP PDE assay performed on M. smegmatis mc²155 strains expressing the alternate cAMP PDE enzyme candidates. Credit: Maeva Denis and Dr Gerald Larrouy-Maumus.

3.4 Optimisation of Rv0805 homologue expression, lysis and further screening for activity

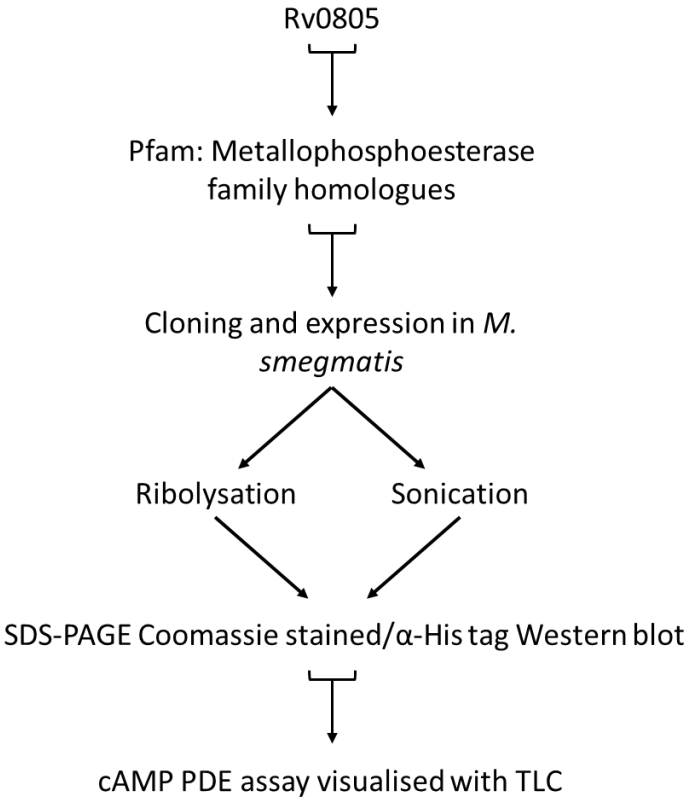


Figure 3.4: Workflow used to investigate if optimisation of Rv0805 homologues expression would allow them to display activity in the cAMP PDE activity assay.

The first aim of my PhD was to identify another cAMP PDE, that is likely conserved across pathogenic and non-pathogenic mycobacteria. I initially sought to build on the data and utilise the tools that the lab had already developed. As seen previously, when the Rv0805 homologues were expressed in *M. smegmatis* mc^2155, and the activity assay was run, none of the homologues showed significant cAMP PDE activity. So far, only Rv0805 had shown activity and only by LC-MS analysis of decreased cAMP levels. In the *M. bovis* BCG ΔRv0805 lysate, cAMP PDE activity could be detected, even with the less sensitive TLC visualisation. I therefore decided to investigate if optimising the purification of the Rv0805 homologues would elucidate whether they had activity or not. The work flow I used is summarised in Figure 3.4, and the data is available in Figure 3.5.

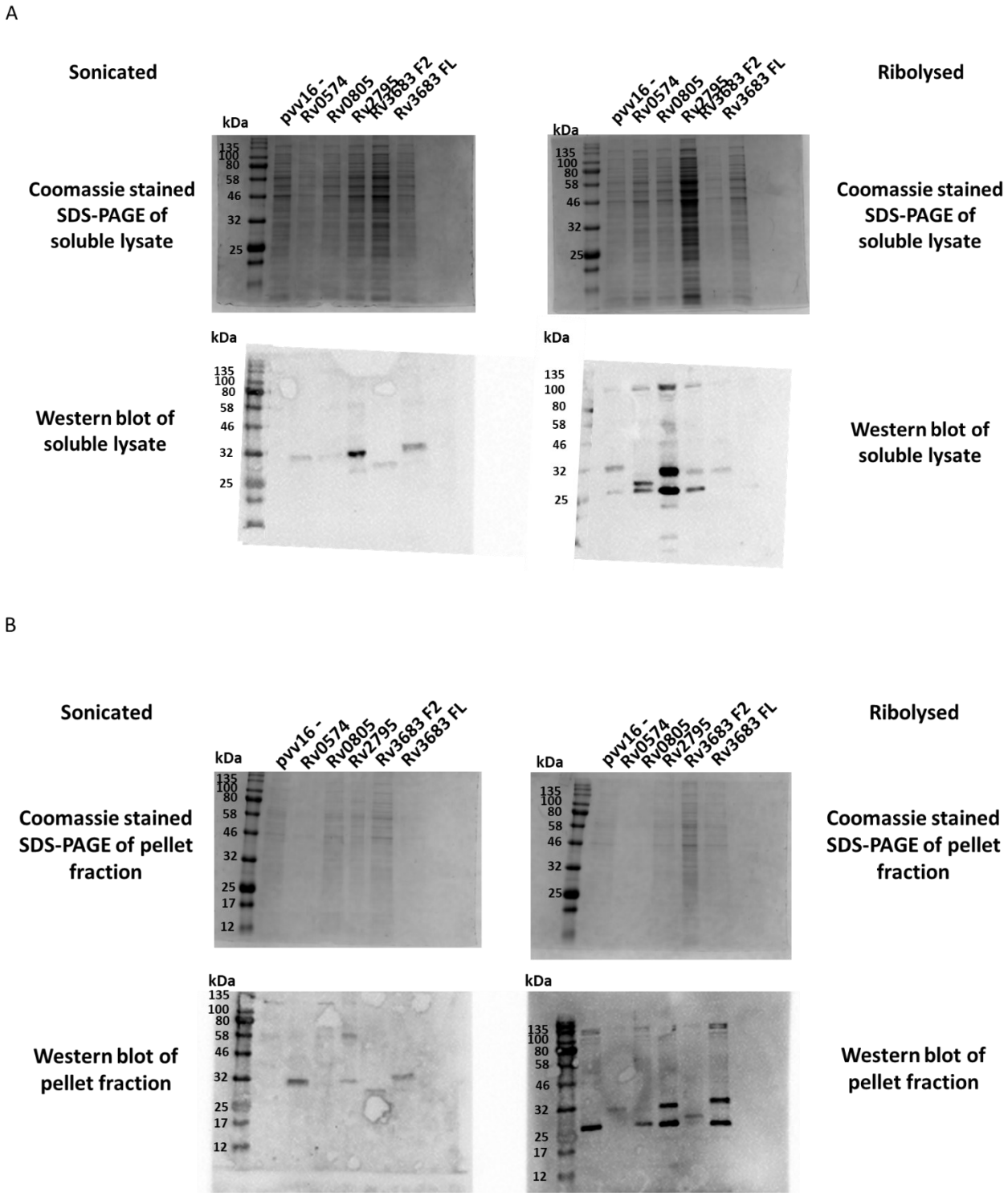


Figure 3.5: SDS-PAGE and Western blot of soluble lysate (A) and solubilised pellet fractions (B) of the M. smegmatis mc²155 strains expressing the His-tagged alternate PDE enzyme candidates. 50µg of protein was loaded from each sample and the negative control is the Empty Vector pVV16. Expected molecular weights of the proteins can be found in Figure 3.3A. An α-His antibody was used for western blot analysis.

To begin, *M. smegmatis* mc^{2}155 cultures expressing the different Rv0805 homologues were cultured to the mid-log phase in conventional 7H9 medium. Clear soluble lysates were obtained via either ribosylation or sonication. This was to investigate which lysis method provided the highest quantity of intact protein. Likewise, insoluble membrane fractions of the lysate were diluted 10-fold and solubilised. 50 µg of each lysate was denatured by boiling with Laemmli and β-Mercapto-ethanol for 5 minutes, and run on SDS-PAGE. Lysate protein composition was assessed with Coomassie staining. From this data, it appears that in the case of the soluble lysates, ribosylation led to the presence of more bands, and this may be due to increased degradation of proteins or more efficient lysis. The same trend can broadly be observed in the solubilised membrane fraction. From Western blot analysis, it seems that sonication tended to yield higher proportions of the target protein and with less bands resembling degradation. This suggests that sonication yielded more full-length proteins of Rv0805, Rv2795, Rv3683 and the truncated version of Rv3683 – possibly because more heat was generated or more small cellular debris is generated from mechanical lysis vs. sonication[314] . However, a band at the predicted mass Rv0574 (41.5 kDa) was not observed, suggesting that due to degradation, the full-length protein was not found in the lysate. Although both blots were developed at the same time, differences in the local ECL abundance can also not be ruled out.

After confirming expression of 4 out of 5 of these proteins in *M. smegmatis* mc^{2}155 clear soluble lysate, I next assayed these lysates for cAMP PDE activity using the cAMP activity assay and TLC for visualisation, that were described earlier (Figure 3.6). Due to contamination during sample spotting onto the TLC solid phase plate, cAMP was also detectable in the 5' AMP control. Unfortunately, none of the *M. smegmatis* mc^{2}155 lysates containing alternate cAMP PDE candidates displayed detectable cAMP PDE activity. Rv0805, which had cAMP PDE activity confirmed by LC-MS analysis, and previous studies, also did not display activity in this assay. It is therefore possible that the TLC visualisation method is not sensitive enough to detect increased 5' AMP in the *M. smegmatis* mc^{2}155 lysate. In the previous data from the lab, LC-MS analysis was focused on the decrease in cAMP and not the production of 5' AMP. It is therefore also likely that other proteins in *M. smegmatis* mc^{2}155 lysate may degrade 5' AMP into adenine, adenosine or inosine monophosphate too rapidly for sufficient accumulation to form a spot on the TLC. This could be assessed with LC-MS analysis of the products of the reaction. Alternatively, active proteins may have been retained in the insoluble pellets which were not assayed.

Although this data did not categorically rule out the possibility that the alternate candidates may possess cAMP PDE activity, it was clear that a different approach would be required to identify another cAMP PDE enzyme. The previous approach left several unknown variables that would require significant optimisation. For example, optimisation of expression, purification of each of the candidates, a broad exploratory screen of cAMP PDE buffers and a more sensitive detection method for analysing the sample metabolome would be required.

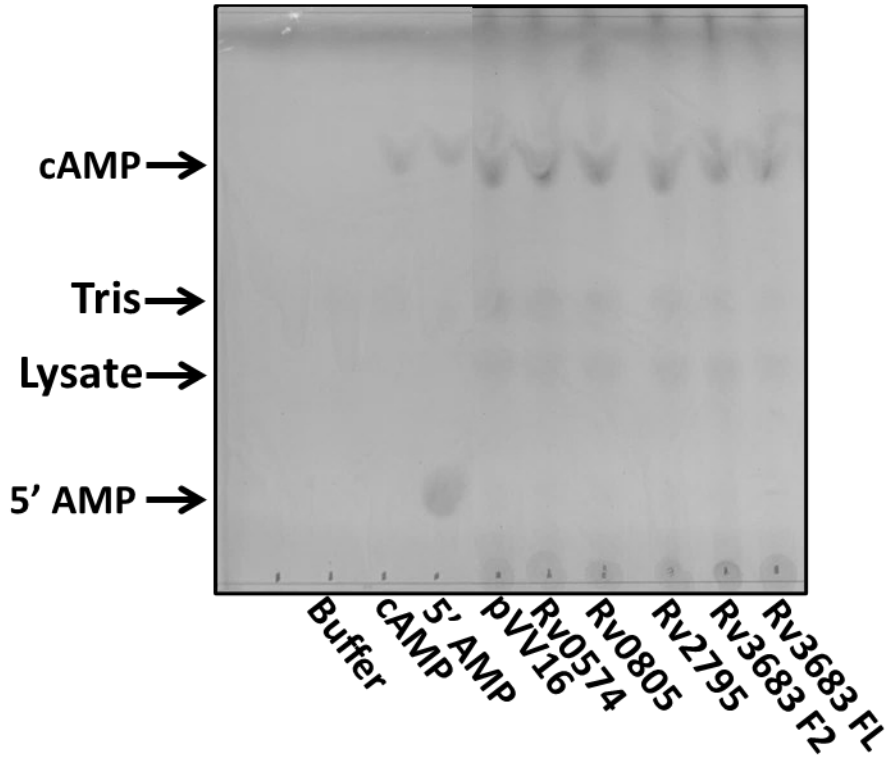


Figure 3.6: cAMP PDE activity assay visualised with TLC to investigate the cAMP PDE activity of the alternate cAMP PDE candidates expressed in M. smegmatis mc^{2}155 lysate. Rv3683 F2 is the truncated version and Rv3683 is the full-length protein.

3.5 Identification of a novel cAMP PDE enzyme via an unbiased activity led, biochemical approach

With no detectable cAMP PDE activity from the alternate metallophosphatase enzymes when expressed in *M. smegmatis* mc^2155, and numerous variables that would require optimisation to make this approach viable, it was decided that an unbiased biochemical approach in *M. bovis* BCG would be used instead. This approach seemed more promising, as it had recently been successfully employed in the discovery of an alternative fructose biphosphatase enzyme in *M. tuberculosis* $H37Rv^{315}$. In *M. bovis* BCG, the cAMP PDE assay detected 5' AMP production with TLC visualisation (Figure 3.1). I repeated this assay, with *M. bovis* BCG ΔRv0805 bacteria cultured without bovine serum albumin (BSA) nor detergent (e.g. Tween-80 or Tyloxopol), and in a static incubator (to ensure surface pellicle formation). This was to prevent detergent from interfering with protein activity and to prevent contamination with non-specific BSA proteins, while also enabling maximal culture density. I was able to observe cAMP PDE activity with TLC visualisation in the *M. bovis* BCG ΔRv0805 lysate (Figure 3.7).

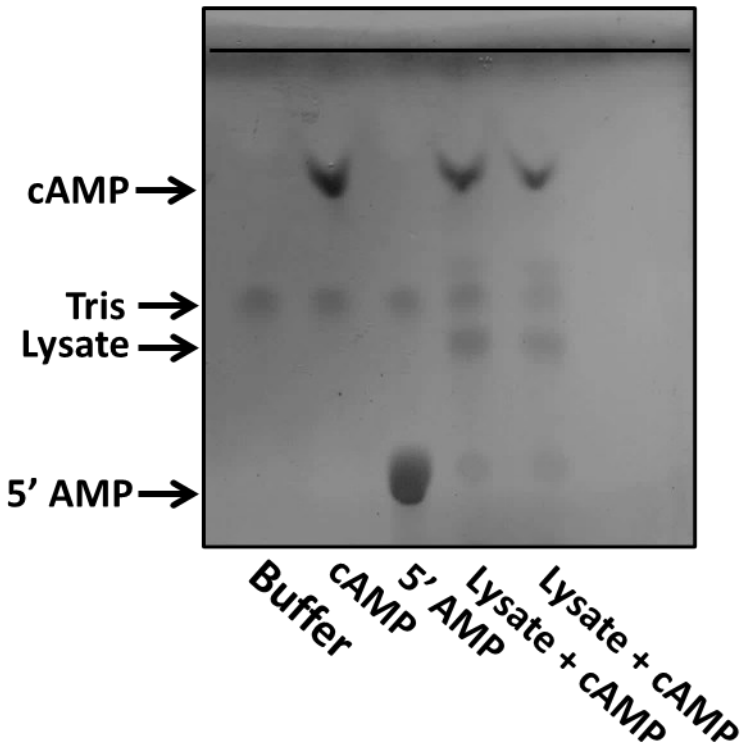


Figure 3.7: cAMP PDE activity assay visualised with TLC to investigate the cAMP PDE activity of BCG ΔRv0805 lysate. 10µg of Lysate was combined with cAMP PDE buffer (as described earlier) and 100mM 3', 5' cAMP as substrate. The assay was run at 37°C for 16 hours (2 technical replicates).

A spot corresponding to the 5' AMP control could be observed in the BCG ΔRv0805 lysate. In order to identify the enzyme with this activity, an activity-led, unbiased biochemical approach was used[315,316]. 500ml cultures of the BCG ΔRv0805 mutant were grown in a static incubator in the defined growth media 7H9, in the absence of detergent and BSA for 8 weeks. Once cultures reached the late stationary phase, they were ribolysed. Although sonication was shown in the *M. smegmatis* mc^2155 data to yield reliably more full-length proteins, due to CL2 safety policies, we were required to ribolyse the bacteria in sealed O-ring screw-cap tubes. The lysates were then centrifuged to acquire the clear soluble protein lysate.

The unbiased activity-led, biochemical approach devised involved fractionating CSL by utilising different chromatographic approaches. Briefly, these approaches and the rationale behind them will be reviewed.

The first approach was to separate the proteins with a CaptoQ Sepharose column. This column separates proteins based on their affinity to an anion. Proteins with higher affinity will bind the column more tightly than proteins with lower affinity. To elute the proteins from the column, an increasing gradient of NaCl (0mM to 600mM) was used. In this way, proteins would be displaced by the stronger affinity of a particular NaCl concentration, and would then elute from the column at the time that the particular NaCl concentration was reached. The CaptoQ Sepharose column was chosen due to the combination of quaternary amine (Q) group with dextran surface extender linkage to the agarose matrix. This was purported to allow fast mass transfer and mediate a high dynamic binding capacity. The further benefits were that this allowed a high flow rate with low back pressure – which was important as we were running crude-lysate through the column (Amersham, GE healthcare).

Pending the success of fractionation of the proteins, and retention of the cAMP PDE activity, the next chromatography approach would be to run the active fractions from the CaptoQ through hydrophobic interaction chromatography (HIC). This approach utilises a Phenyl Sepharose column and an HIC medium that separates proteins based on their surface hydrophobicity. This is mediated by reversible binding to a hydrophobic surface medium in the column. High salt concentrations enhance the interaction of the protein to the column. If all proteins run through the column bind to the surface of the column, depending on the strength of their interactions, they will be eluted along a gradient of decreasing ionic strength. Hydrophobicity interaction chromatography is often used as an intermediate chromatography step between anion exchange and size exclusion chromatography.

The final intended chromatography approach involves size exclusion chromatography (SEC). This approach is mediated by a column that consists of a series of cross-linked agarose and dextran beads of varying sizes. In SEC, samples do not bind to the column, and the buffer they are contained within will not directly affect their purification. This is because what determines when a protein is eluted depends on its size. For instance, smaller proteins will pass through a high number of beads, and will therefore

elute later than larger proteins which must bypass the beads entirely. For this I chose a Superdex 200 column that purportedly separates proteins with a M_r of 10,000 to 600,000 Daltons.

3.5.1 Summary of the unbiased activity-led biochemical approach

During the utilisation of this approach, CaptoQ and SEC approaches worked well to separate the proteins and retain cAMP activity. Unfortunately, the intermediary purification step HIC did not adequately separate the proteins, and it appeared likely that the proteins run on the column had been lost in the waste. Additionally, a collaborator advised that HIC purifications work best with a larger volume of starting sample than we could produce due to CL2 safety regulations. The work-flows undertaken are summarised in Figure 3.8 and experimental results from Figure 3.9 onwards.

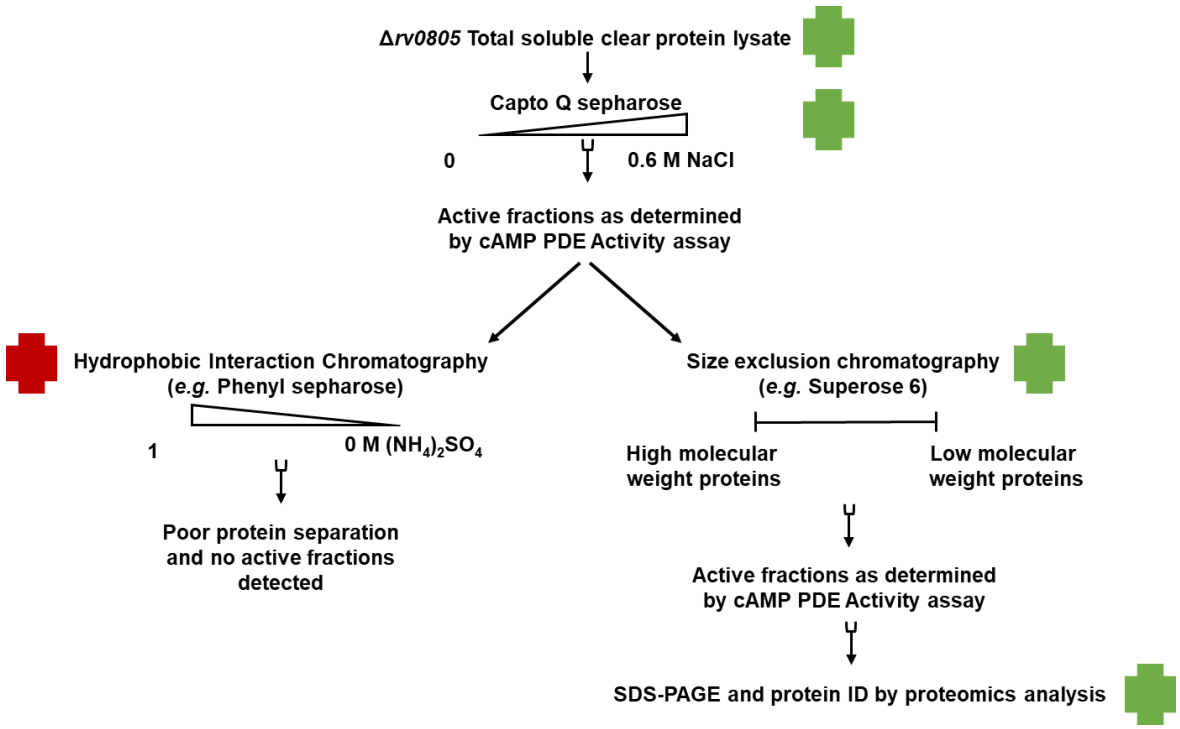


Figure 3.8 Summary of unbiased activity-led, biochemical purification approach. Successful purification protocols are indicated with green crosses. The unsuccessful HIC inclusive work flow is indicated with a red cross.

3.5.2 CaptoQ purification of BCG ΔRv0805 lysate identifies 2 fractions with cAMP PDE activity

As seen in Figure 3.9, when 50 mL clear soluble lysate was run on the CaptoQ column, the proteins in the lysate were separated along the NaCl gradient. The presence of protein in a fraction was inferred by UV absorption. Previous pilot studies of this approach on the *M. bovis* BCG lysate suggested that proteins of interest eluted towards the middle of the NaCl gradient (data not shown; credit: Ipsita Mahapatra).

From fraction A11 onwards, the protein content and activity of the desalted fractions (buffer exchanged with a Vivaspin 2 spin column) was assayed by SDS-PAGE stained with Coomassie (Figure 3.9B) and

by the same activity assay/TLC validation respectively (Figure 3.9B). This assay is the same as the cAMP PDE assay that previously identified the cAMP PDE activity in *M. bovis* BCG. From the activity assay, it was clear that CaptoQ fractions B11 and B10 contained protein(s) with cAMP PDE activity, as indicated by the spot on the TLC corresponding to the 5'AMP control (Figure 3.9C). The fractions eluted at 228mM NaCl on the gradient. This was important to identify, as when this purification was repeated, active protein was again eluted at 232mM NaCl. This provided confidence that purification of 2 fractions with cAMP PDE activity was possible with CaptoQ separation.

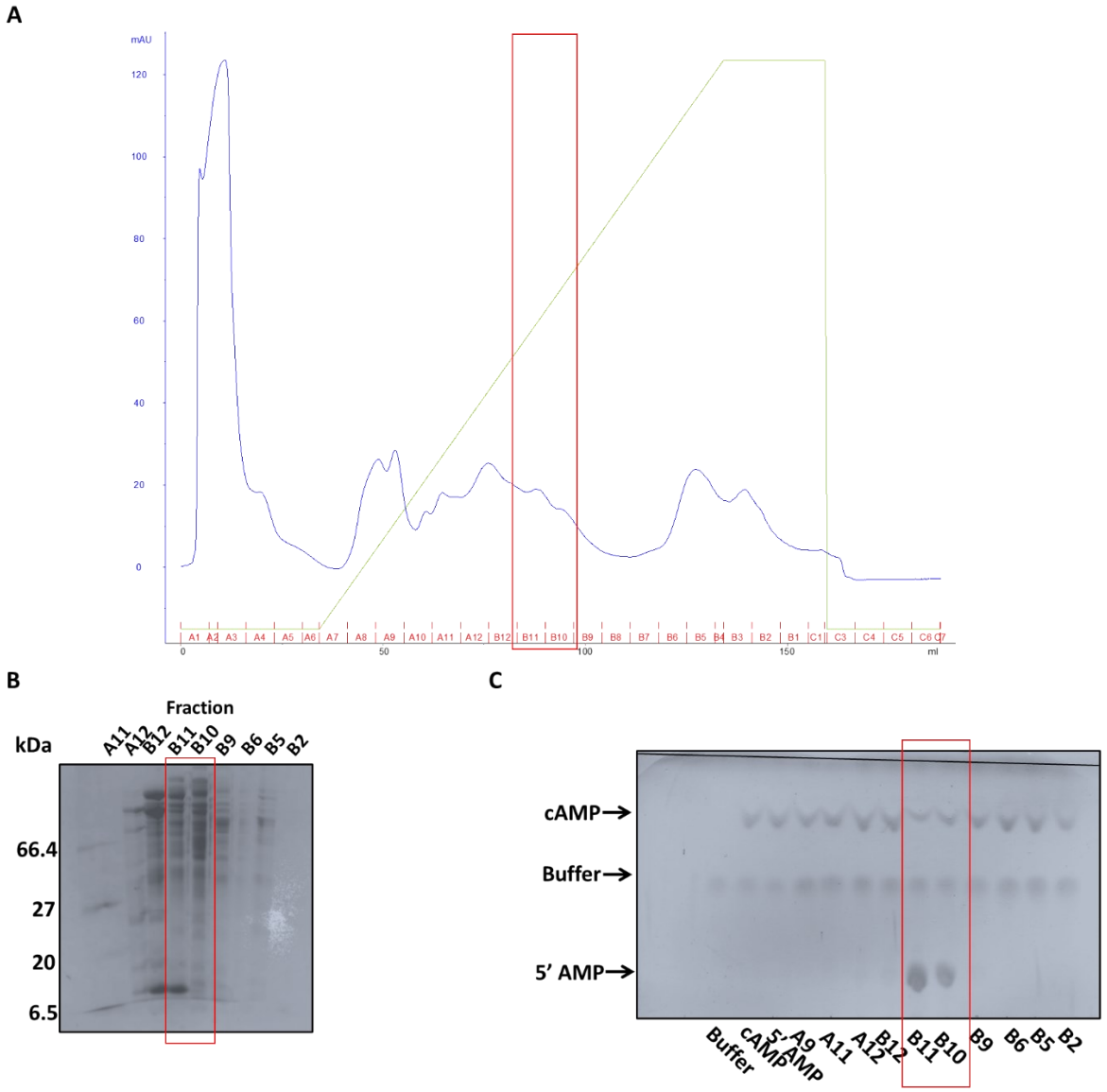


Figure 3.9: CaptoQ purification of BCG ΔRv0805 lysate. A: Chromatogram of the purification. The Y-axis is mAU of UV absorption of material passing the detector (located post column). The X-axis corresponds to the volume that flowed through the column, (including lysate, wash and gradient elution), as well as the fractions that were eluted at that point in the run. B: Coomassie stained SDS-PAGE of a selection of fractions from the CaptoQ run 50µg protein concentration of each fraction was loaded. C: TLC visualisation of the cAMP PDE assay with 50µg of lysate.

3.5.3 Phenyl Sepharose purification of active CaptoQ fractions B11 and B10

To further separate the proteins within fractions B11 and B10 from the CaptoQ, hydrophobic interaction chromatography was attempted. The results of this purification can be seen in Figure 3.10. The chromatogram did not appear to show separation of the proteins along the gradient (Figure 3.10A), with UV readings only beginning to decline at around 40% of the gradient. This decline continued beyond the maximum gradient concentration and this may explain why Coomassie stained SDS-PAGE analysis only found proteins in the final fractions (Figure 3.10B). It is also possible that the majority of the protein load did not bind to the column, perhaps due to a manufacturing defect or improper storage. None of these fractions displayed activity in the cAMP PDE activity assay (after desalting) (Figure 3.10C). Taken together, this data suggested that hydrophobic interaction chromatography would not be a viable intermediary step in the purification protocol.

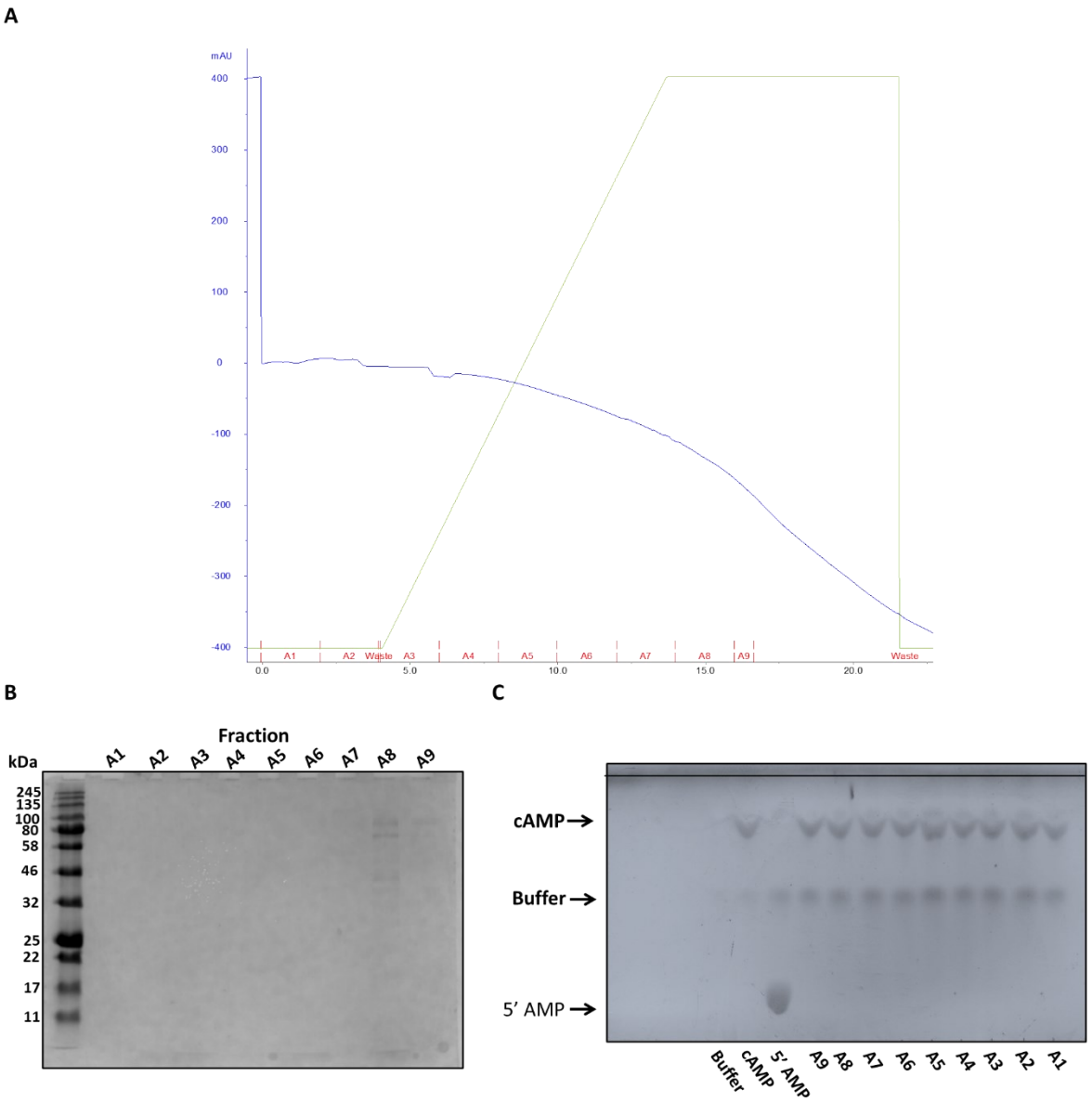


Figure 3.10: Phenyl Sepharose purification of CaptoQ active fractions B11 and B10. A: Chromatogram of the purification. The Y-axis is mAU of UV absorption of material passing the detector (located post column). The X-axis corresponds to the volume that flowed through the column, (including sample, wash and gradient elution), as well as the fractions that were eluted at that point in the run. B: Coomassie stained SDS PAGE of the fractions eluted from the run. 50µg protein concentration of each fraction was loaded. C: TLC visualisation of of the cAMP PDE assay performed on these fractions with 50µg of the eluted fraction .

3.5.4 Size exclusion of the CaptoQ active fraction B11

Instead of attempting to repeat the hydrophobic interaction chromatography step, after reproducing the active *M. bovis* BCG ΔRv0805 lysate activity assay (Annex 1), this lysate was again purified with CaptoQ anion exchange chromatography (Annex 2). cAMP PDE activity was again identified in fractions B11 and B10, which eluted at 232mM NaCl. (Chromatogram, SDS-PAGE, LC-MS and TLC linked activity assay (Annex 2).

The active CaptoQ fraction B11 was further fractionated with size exclusion chromatography (Figure 3.11).

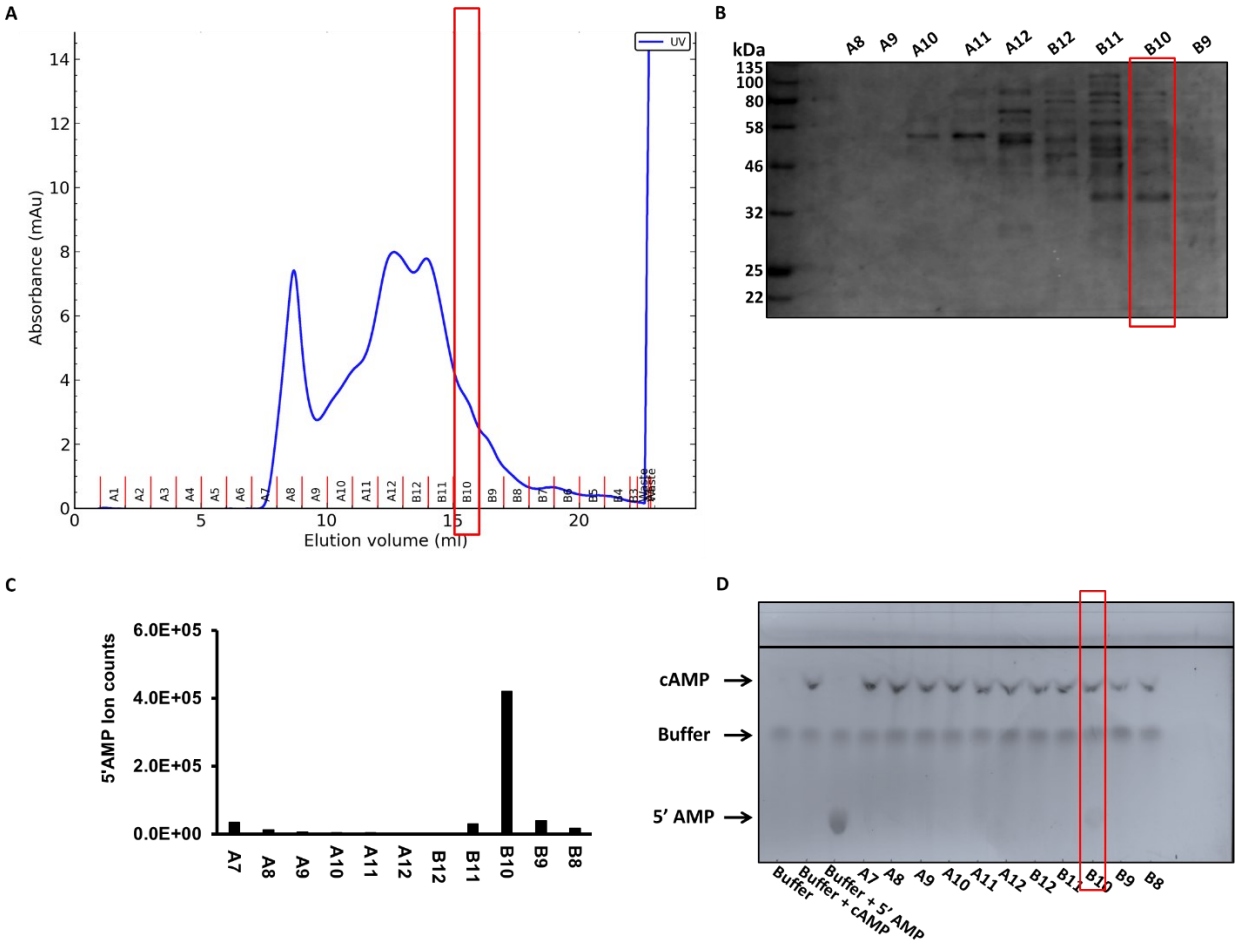


Figure 3.11: SEC purification of the CaptoQ active fraction B11. A: Chromatogram of the run. The Y-axis is mAU of UV absorption of material passing the detector (located post column). The X-axis corresponds to the volume that flowed through the column. B: Coomassie stained SDS PAGE of the fractions. 50µg of the selected eluted fractions were loaded. C: LC-MS quantification of the cAMP PDE activity of each fraction, by measuring production of 5' AMP from 50µg of each selected fraction when combined with cAMP. D: TLC visualisation of the cAMP PDE activity assay.

The chromatogram for the SEC (figure 3.11A) showed a greatly decreased UV intensity when compared to the initial CaptoQ (Figure 3.9A). This is likely due to the fact that only 1 CaptoQ fraction (B11) was loaded instead of the whole lysate. The first peak at A7 and A8 likely corresponds to the "void volume" of the column, and is protein that will pass straight through, without separation by the SEC beads. The protein content of the fractions was assayed with SDS-PAGE and Coomassie stain (Figure 3.11B). It was apparent from this data that the proteins were being fractionated as expected, based on molecular weight. For example, fraction A10 contains proteins around 60 kDa whereas a later fraction such as B9 appears to contain proteins around 32 kDa.

Fraction B10 displayed cAMP PDE activity with the vast majority of activity found in SEC fraction B10 exclusively (either 10 or 15 times enriched compared to SEC B11 and B9 respectively – as measured by LC-MS analysis - Figure 3.11C). TLC visualisation of the assay also corroborated the presence of cAMP PDE activity – and indicated that the levels of 5' AMP seen from the 16-hour incubation of the fractions was sufficient to form a faint spot on the TLC (3.11D).

3.5.5 Proteomic analysis

In order to identify the protein(s) in fraction B10 that possessed cAMP activity, the final step of the purification approach was to run the active fraction on SDS-PAGE, Silver nitrate stain the gel to investigate the complexity of the fractions, and digest the lanes for proteomic mass-spectrometry analysis. A Silver nitrate stain is far more sensitive than Coomassie staining (limit of detection: 0.05-0.2ng of protein vs 5ng) and would allow visualisation of lower abundance proteins.

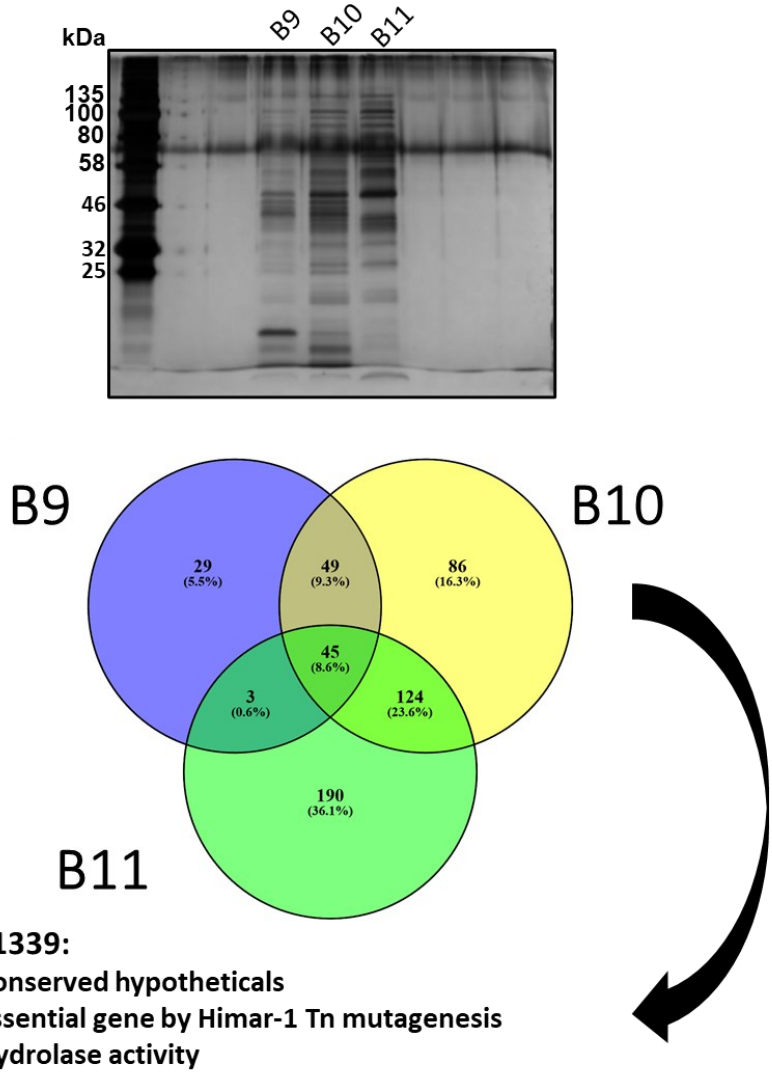


Figure 3.12: Venny analysis of the proteins detected in each of the 3 SEC fractions B9, B10 and B11 (10µg load per fraction was run on an SDS-PAGE gel and fraction composition was visualised with a silver nitrate stain). The fractions were prepared proteomic analysis by tryptic digest and proteins were identified from peptides compared to the Swiss-Prot database. Rv1339 was identified exclusively in fraction B10 - the only fraction with significant cAMP PDE activity.

The proteomic work was performed by the proteomics facility at Imperial College London. Briefly, the gel was de-stained and each lane was excised and trypsin digested. Cysteine residues were reduced and then alkylated to improve peptide yield and identification of disulphide bond containing proteins. The peptides were then loaded into the Waters G2SI-synapt instrument coupled to nano UHPLC[317]. The peptides identified in each fraction were compared to the Swiss-prot database in order to identify the proteins they were likely derived from. These proteins were identified by their accession number and any functional annotations in the Swiss-prot database (all *M. tuberculosis* H37Rv proteins present in fraction B10 can be found in Annex 6). This data was compiled into spreadsheets and manually inspected to remove redundant annotations e.g. the same protein from multiple different organisms. After this process, the protein composition of the fractions were visualised with Venny software[318] (Figure 3.12). Fractions B9, B10 and B11 had 29, 86 and 196 unique proteins respectively. As expected, neighbouring fractions (B9 and B10 and B10 and B11) had greater amounts of common proteins than non-neighbouring fractions (B9 and B11). In addition, 45 proteins were found to be common to all 3 fractions.

From proteins unique to fraction B10, only 8 proteins were found to be present in *M. tuberculosis* H37Rv/*M. bovis* BCG. Only 3 of these were uncharacterised proteins: Rv0250, Rv2568 and Rv1339. Rv0250 is a very small protein (10.8781 kDa), completely uncharacterised and found to be non-essential by Hinmar-1 transposon mutagenesis studies. Rv2568 has a mass of 38.00026 kDa but was also uncharacterised and non-essential. Finally Rv1339 was found to have a mass of 29.154 kDa, and had been shown in some Hinmar-1 transposon mutagenesis studies to be essential[319]. However, a later study utilising a fully saturated transposon approach found it to be non-essential for *in-vitro* growth[320]. More importantly, it was predicted to have hydrolase activity – which would enable cleavage of the phosphodiester bond of cAMP. Furthermore, as will be seen in Chapter 4, Rv1339 is conserved across both pathogenic and non-pathogenic mycobacteria. When taken together, this made it a strong candidate to be a conserved mycobacterial cAMP PDE enzyme.

3.6 Discussion

In this Chapter, my research and that of previous members of the lab was intended to identify a new cAMP PDE enzyme in mycobacteria. This would be a valuable finding for the field and could be used as a tool to alter cAMP levels and test the hypothesis that cAMP is involved with antimicrobials susceptibility in mycobacteria. The only currently identified cAMP PDE, Rv0805, is only found in certain pathogenic mycobacteria (e.g. *M. tuberculosis*, *M. bovis* BCG, *M. leprae* and *M. marinum*). It displays low activity when compared to PDE enzymes found in other organisms e.g. *E. coli – Rv0805* V_{max} of *30* nmol/sec/mg *vs E. coli* CpdA V_{max} of 2µmol/sec/mg [194,205,279]. This is unusual as the levels of cAMP in mycobacteria are at least 10-fold higher than those found in *E. coli* [174,175,204].

It is known that *M. tuberculosis* bacilli secrete cAMP into the host macrophage during infection to intoxicate host cells and prevent host killing mechanisms[78,79]. However, like any second messenger, tight regulation is required. Deleterious levels of cAMP (even a 20-fold increase[11]) in *M. tuberculosis* have been shown to lead to inhibition of growth with cholesterol as a carbon source[11,215]. Cholesterol catabolism is a key carbon source for *M. tuberculosis* H37Rv during infection[253,283,284] and therefore accumulation of a signalling molecule that impairs this process would be highly disadvantageous. It would therefore be important to have an enzyme that can rapidly manage levels of cAMP – unlike Rv0805. Rv0805 has been shown to possess cAMP catalysis-independent gene regulation functions[308] and to have a higher affinity for alternate substrates[281]. For example, transcriptional studies have shown that Rv0805 expression dysregulates genes involved in the methyl citrate cycle and cholesterol catabolism[308]. Beyond these pathogenic mycobacteria, cAMP PDE activity has been identified in non-pathogenic mycobacteria[286], so it was likely that this alternate enzyme was also ubiquitous across mycobacterial species.

Initial homology-based attempts at finding an enzyme with the same domain as Rv0805 that was also a cAMP PDE were unsuccessful. This was possibly due to a combination of investigating these proteins outside of their normal *in-vivo* conditions, by expressing them in the fast-growing *M. smegmatis* mc^2155 or due to differences in the biochemistry precluding our ability to assess production of 5'AMP via TLC. It is also worth noting that of the proteins sharing the same domain as Rv0805 in *M. tuberculosis* H37Rv, that our group was able to express and investigate, only Rv2795 and Rv3683 were conserved across mycobacteria. Of the remaining 2 proteins that could not be expressed, Rv1277 and Rv2577, the latter was also not conserved across mycobacteria. While we cannot rule out that these proteins may have cAMP PDE activity, they did not show up subsequently in the unbiased activity-led approach.

By using an unbiased activity-led biochemical approach, I was able to identify a promising candidate for this alternate cAMP PDE. This approach was powerful and had the advantage of being able to identify proteins with different domains and potentially non-existent homology to known enzymes[316].

The key to this approach was in having a rapid and sensitive readout of the target activity. While the TLC visualisation of the cAMP PDE activity assay worked well in the pathogenic mycobacteria *M. tuberculosis* H37Rv and *M. bovis* BCG, it did not appear sensitive enough to identify cAMP PDE activity in *M. smegmatis* mc^2155 lysate. This could be evidenced by the ability to detect Rv0805 activity in *M. smegmatis* mc^2155 by LC-MS quantification, but not by TLC visualisation (Figure 3.3 vs Figure 3.6). By coupling the cAMP activity assay to LC-MS metabolite quantification, it enabled discrimination between different levels of cAMP activity. Additionally, I was able to utilise quantification of activity to discriminate between protein fractions that had an active protein, and those that may only have trace amounts of the protein or contamination with exogenous cAMP.

Rv1339 was identified as a promising candidate for the alternate cAMP PDE. This protein is conserved ubiquitously across mycobacteria, is potentially essential and was predicted to possess hydrolase activity. Likewise, a homologue of this protein was identified as a cAMP PDE in *Corynebacterium glutamicum* – CpdA (45.71% identity)[321]. From the characterisation work on this *C. glutamicum* CpdA enzyme, it appears to have 1000 times more cAMP PDE activity than Rv0805 (V_{max} 33.6µmol/min/mg vs Rv0805 V_{max} of 30 nmol/sec/mg)[321]. This implies that this enzyme would be better able to control the high cAMP levels in mycobacteria (relative to model organisms such as *E. coli* [174,194]). Indeed, the authors of the paper that identified CpdA in *C. glutamicum* characterised it as a regulatory component of cAMP levels and by extension, cAMP binding transcription factor mediated gene expression[321]. Rv1339 may mediate a similar role in *M. tuberculosis* H37Rv. After environmental stimulation of an AC enzyme to produce cAMP[11,68,216], when the environmental conditions change, cAMP levels would need to be returned to normal levels for effective homeostasis. Beyond secretion of cAMP under specific circumstances (as seen with Rv0386 during *M. tuberculosis* H37Rv infection[78]), enzymatic degradation would be the most likely mechanism for lowering cAMP levels.

In sum, these findings make a strong case for characterising Rv1339 and investigating potential cAMP PDE activity. Therefore, the next phase of my project would be to express, purify and crystallise this enzyme. This would allow enzymatic studies that would enable comparisons with Rv0805, CpdA from *C. glutamicum* or *E. coli*. Furthermore, enzymatic studies could elucidate the substrate specificity of this enzyme or determine whether it possesses other "moonlighting" activities. A tool that can modulate cAMP levels in mycobacteria, without other catalysis independent effects[281,308], is sorely needed in order to understand this complex signalling pathway. Identification of Rv1339 could aid investigation of this signalling system in mycobacteria. Identifying a novel class of cAMP PDE enzyme would also be of major benefit to fields beyond mycobacteria as this superfamily of enzymes is poorly characterised[278], despite cAMP being conserved across many taxa of life[65].

Chapter 4 - Expression, purification and enzymatic characterisation of Rv1339

4.0 Rv1339 is a putative atypical class-II PDE enzymes, from the metallobetalactamase superfamily

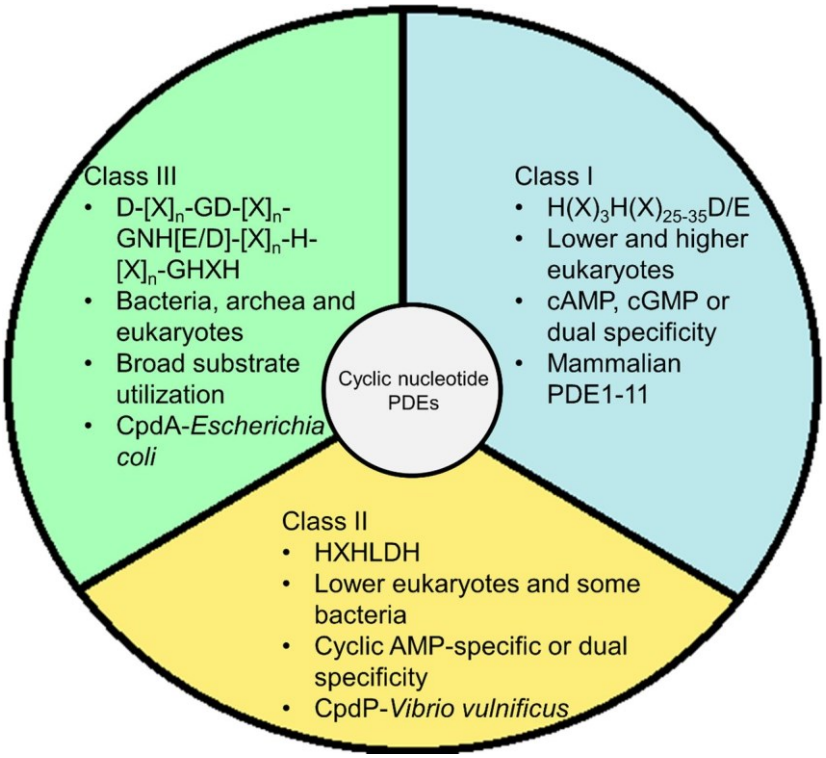


Figure 4.1: Summary of the 3 current classes of cAMP PDE enzymes. Modified from[278]

Three evolutionary unrelated classes of cyclic nucleotide phosphodiesterase (PDE) enzymes have been described[278]. The first class (Class-I) contains all the mammalian PDE enzymes. There are 11 families of mammalian PDE enzymes and these have either specificity to 3', 5' - cAMP, 3', 5' – cyclic GMP (cGMP) or both [322]. This class of enzyme is also found in higher and lower eukaryotes e.g. protozoa[278]. The enzymes in this class and other classes are usually classified based on a common motif e.g. class-I - $H(X)_3H(X)_{25-35}D/E$ motif (Figure 4.1).

Class-II consists of enzymes that can also be found in lower eukaryotes like protozoa and yeasts e.g. *Candida albicans*[323,324], or in some bacteria e.g. *Yersinia pestis*[325]. These enzymes contain an HXHLDH motif and a β-lactamase fold [278](Figure 4.1). A sequence resembling this motif was identified in Rv1339 – HLHADH. The significance of this will be reviewed shortly.

Class-III PDE enzymes have a wider phylogenetic distribution. They are found in bacteria, archaea and eukaryotes, and belong to the calcineurin like metallophosphoesterse (Pfam: PF00149) family of proteins with a D-(X)n-GD-(X)n-GNH[E/D]-(X)n-H-(X)n-GHXH motif [326] (Figure 4.1). A feature of enzymes in this class is the presence of divalent metal ions contained in the active site. For example, in *P. aeruginosa* CpdA[179]and in *M. tuberculosis* H37Rv Rv0805[205] these ions are iron and manganese. It has been postulated that these metal ions are important for substrate/product binding and that the metal

ions support a water molecule that could potentially act as a nucleophile for the hydrolysis reaction. The class-III enzymes are mainly characterised in bacteria and the original member of this class, CpdA, was identified in *E.coli* by Imamura et al[194]. The authors showed that *in vitro*, the presence of CpdA prevented Crp dependant transcription from the lac promotor. Crp is a cyclic AMP receptor binding protein and it stands to reason that modulating levels of a PDE enzyme would interfere with 3', 5' - cAMP levels, then subsequently the functions of Crp. Rv0805, a class-III PDE identified in *M. tuberculosis* H37Rv, was originally found to be a 3', 5' - cAMP PDE[205]. Subsequently it emerged that the ability of Rv0805 to hydrolyse 2', 3'- cAMP or 2', 3' - cGMP was 150-fold higher than for 3', 5' - cAMP[287]. This illustrates the diverse specificity these enzymes can display for cyclic nucleotides. Rv0805 is membrane bound and in addition, has recently been linked to cholesterol utilisation in mycobacterial growth – via regulation of the methyl citrate cycle genes and propionate metabolism. However, this is through cAMP hydrolysis independent moonlighting activities[308].

As previously reviewed in the introduction, Rv0805 is currently the only annotated cAMP PDE enzyme in *M. tuberculosis* H37Rv[68], despite evidence of other cAMP PDE enzymes in mycobacteria since 1979[286], and non-cAMP hydrolysis related moonlighting activies[281]. This means that in order to convincingly identify a novel cAMP PDE enzyme in *M. tuberculosis* H37Rv, in-depth phenotypic, bioinformatic and biochemical characterisation would be required. Subsequently, bioinformatic characterisation of Rv1339 will be investigated. Then, attempts to purify this enzyme, and some eventual preliminary enzyme kinetics will be outlined.

4.1 Bioinformatic and evolutionary analysis of Rv1339, and atypical Class-II PDE family enzyme

According to Pfam, predicted structural analysis (Figure 4.2) and blast*p* sequence alignment coupled to evolutionary analysis algorithms (Figure 4.3): Rv1339 is a metallo-β-lactamase 2 family protein (Pfam: PF12706).

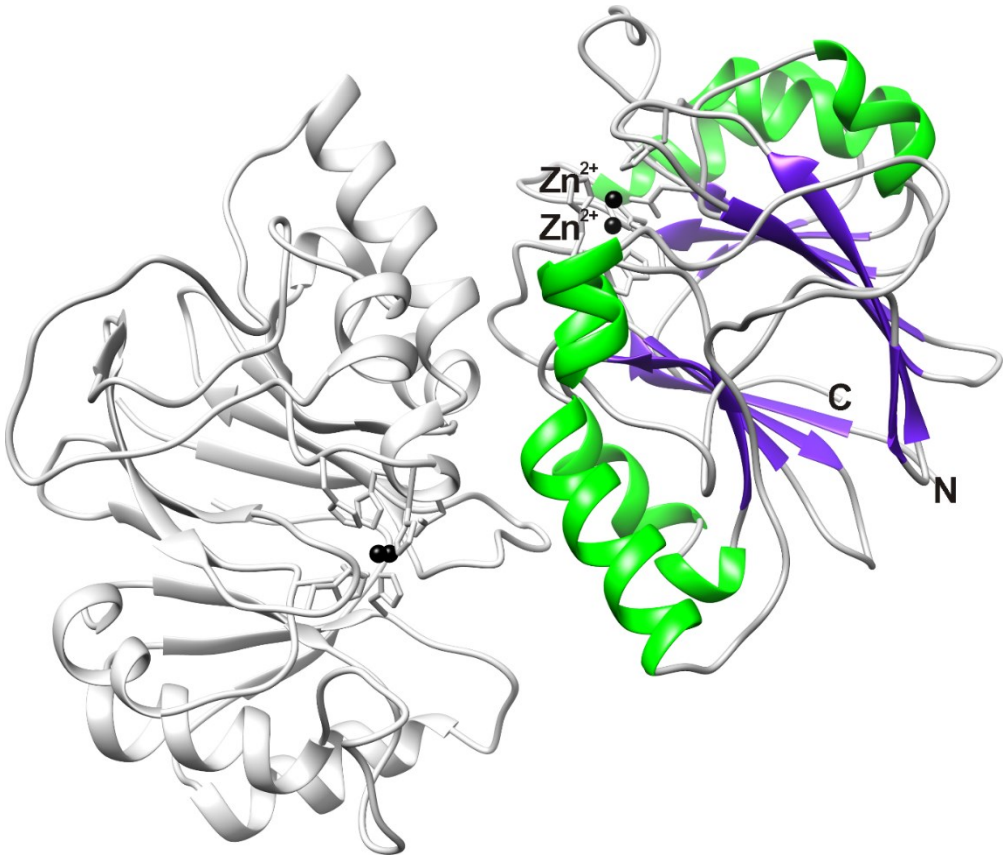


***Figure 4.2**: Ribbon representation of a homology model of Rv1339 calculated using Modeller 9.23*[327] *using as template the crystal structure of the metallo-beta-lactamase fold protein YhfI from Bacillus subtilis (PDB: 6KNS – 27.02% idenitiy). YhfI was found to be a homodimer in solution by size exclusion chromatography*[328]*. The structure forms the αβ-sandwich configuration characteristic of the metallo-β-lactamase fold, consisting of 2-β sheets of 7 β-strands each (purple), with the α helices (green) capping the β-sheet. The two zinc cations in the active sites are depicted in black. The figure was created using UCSF Chimera*[329]

From the evolutionary analysis, Rv1339 was identified as an orthologue of CpdA in *C. glutamicum*[321]. The authors of the study identifying CpdA in *C. glutamicum* as a cAMP PDE proposed that this enzyme was part of the class-II PDEs as it contains the motif HXHTDH, similar to the HXHLDH, observed in class-II enzymes[321]. Characteristics of class-II PDEs include the presence of the metal binding site motif HXHLDH , and the ability to also catalyse the breakdown of cGMP to 5' GMP[278].

Bona fide class-II PDEs belong to the PFAM family of cAMP phosphodiesterase class-II (PF02112), but Rv1339 and *C. glutamicum* CpdA are members of family metallo-β-lactamase-2 (PF12706). Both families are evolutionary related, as they form part of the Metallo-hydrolase/oxidoreductase superfamily (SSF56281). It is also noteworthy that the Pfam class-II PDE family (PF02112) is restricted to Fungi, Bacteroidetes and Proteobacteria, while PF12706 has a much broader distribution that includes Archea, Bacteria and Eukaryota. I will hereafter refer to PDEs belonging to PF12706 as "atypical class-II PDEs" and suggest the use of the motif [T/S]HXHXDH as signature of all type-II PDEs, where X tends to be a hydrophobic, small or very small amino acid residue (as opposed to the current, reductive characterising motif of HXHLDH). This classification rule would allow for the alterations observed in the metal binding motif that can be observed in Rv1339, CpdA from *C. glutamicum* and YfhI from *B. subtilis* (Figure 4.3).

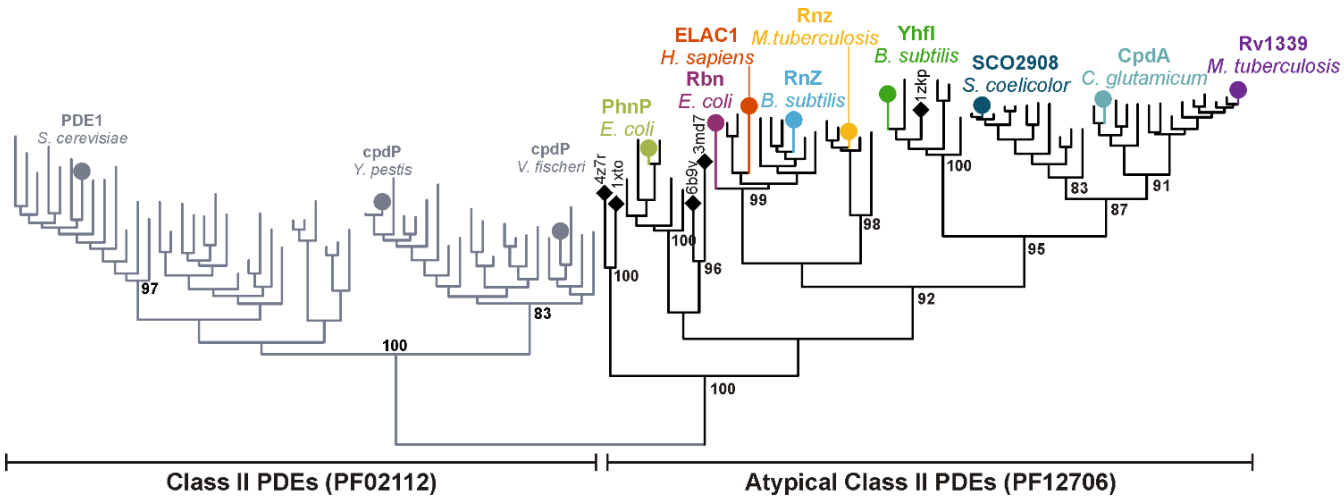


Figure 4.3: Maximum-likelihood phylogenetic consensus tree of selected members of typical (PF02112) and atypical (PF12706) phosphodiesterase (PDE) class-II families. Sequences were aligned with MUSCLE[330]. The best-fit amino acid substitution model for the alignment was LG+R6, found using ModelFinder[331]. Trees were calculated in IQ-Tree (1.6.11[332]), using 100 bootstrap replicates. Trees were visualised using Dendroscope[333]. Selected solved 3D-structures are shown with a diamond and labelled with their PDB accession code. Characterised members of the families are named and have the following Uniprot and PDB accession codes in brackets: Aliivibrio fischeri cpdP (Q56686), Yersinia pestis cpdP (Q8ZD92), Saccharomyces cerevisiae PDE1 (P22434; 4OJV), Escherichia coli Phosphotriesterase homology protein PhP (P45548;1BF6), Homo sapiens zinc phosphodiesterase ELAC protein 1 (Q9H777;3ZWF), Escherichia coli Ribonuclease BN Rbn (P0A8V0; 2CBN), Mycobacterium tuberculosis Ribonuclease Z (P9WGZ5), Bacillus subtilis Ribonuclease Z (P54548; 1Y44), Bacillus subtilis Metallo-hydrolase YhfI (O07607; 6KNS), Streptomyces coelicolor (SCO2908), Corynebacterium glutamicum CpdA a.k.a. Cgl2508 (Q8NMQ7), Mycobacterium tuberculosis H37Rv Rv1339 (P9WGC1).

The potential orthologue of Rv1339 in *C. glutamicum* CpdA has been characterised as a cAMP PDE enzyme[321]. Although the structural fold defining the protein family (PFAM12706) is named after the beta-lactamase enzymes known for their role in antibiotic resistance[334–336], other members are also known to have thioester hydrolase activity instead – such as hydroxyacylglutathione hydrolases[337] and the phosphodiester hydrolases discussed in this thesis[278,321].

All members of PF12706 characterised so far display phosphodiesterase activity (the ability to hydrolyse a phosphodiester bond) but substrates vary, including RNA, tRNA, 3', 5' -cNMP, 2', 3' cNMPs and FAD. For example, a 3', 5' – cAMP or cGMP PDE[321], a ribonuclease (that assists in tRNA maturation by cleaving extra 3' nucleotides from the tRNA precursor)[328], *E. coli* PhnP which displays activity towards 2', 3' – CNMPs[338]as well as hydrolysis activity of photoreactive flavin adenine dinucleotides (FAD) in *Koribacter versatalis*[339] .

Besides CpdA from *C. glutamicum*, YfhI from *B. subtilis* is the closest related orthologue that has been characterised. YfhI displays the characteristic metallo-β-lactamase metal binding site motif of histidine and aspartate residues stabilising 2 divalent metal ions[328,340]. Unfortunately, the authors did not characterise the substrate specificity beyond the general phosphodiester bond containing substrate Bis(para-nitrophenyl)-phosphate. Activity towards this non-physiological substrate is common to 3', 5' – cAMP PDE enzymes[279,321] but also 2', 3' – CNMP PDEs[338]. As noted in the paper characterising YfhI, the broad functional roles and activities mediated by metallo-β-lactamase enzymes makes it difficult to predict function or substrate specificity from sequence analysis/predicted structures[328]

This broad substrate specificity is very interesting but unsurprising – considering the hydrolase activity of this protein family is also believed to have evolved at multiple different points concurrently; making ancestral analysis difficult[341]. As a result, with just a bioinformatic approach we cannot predict the substate specificity of Rv1339. Instead, protein purification for determination of Michaelis-Menten kinetic parameters with a variety of phosphodiester bond containing substrates will be required.

4.1.1 The predicted active site of Rv1339

1ZKP in *Bacillus anthracis* has a 29% identity with Rv1339 and unknown activity, however, it has been crystallised (1.5Å resolution) (Figure 4.4). 1ZKP is also a member of the metallo-β-lactamase superfamily of proteins. The proteins of this family have a pair of Zinc ions in their active site and a Metallo-β-Lactamase fold (4-layered β-sandwich with 2 mixed β-sheets flanked by α-helices). These ions have been shown to be involved in substrate binding and catalysis[340]. This metal ion containing site found in 1ZKP (with two annotated zinc ions) is conserved in the predicted structure of Rv1339 (Figure 4.2).

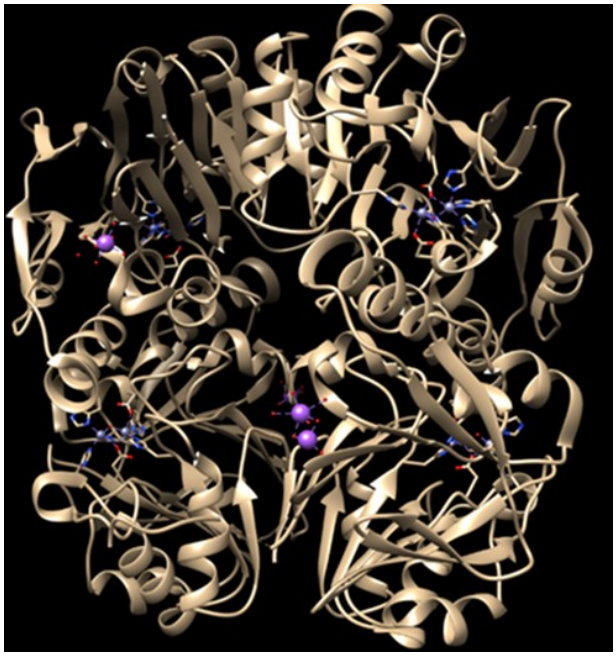


Figure 4.4: Crystal structure of 1ZKP, a metallo-β-lactamase protein from Bacillus anthracis with unknown function.

A common approach in inhibiting enzymes in this superfamily involves destabilising the metal ion stabilising interactions[340] (Figure 4.5). In order to generate an activity mutant of Rv1339, the key residue (D180 in Rv1339) was substituted with an alanine, via site directed mutagenesis[342].

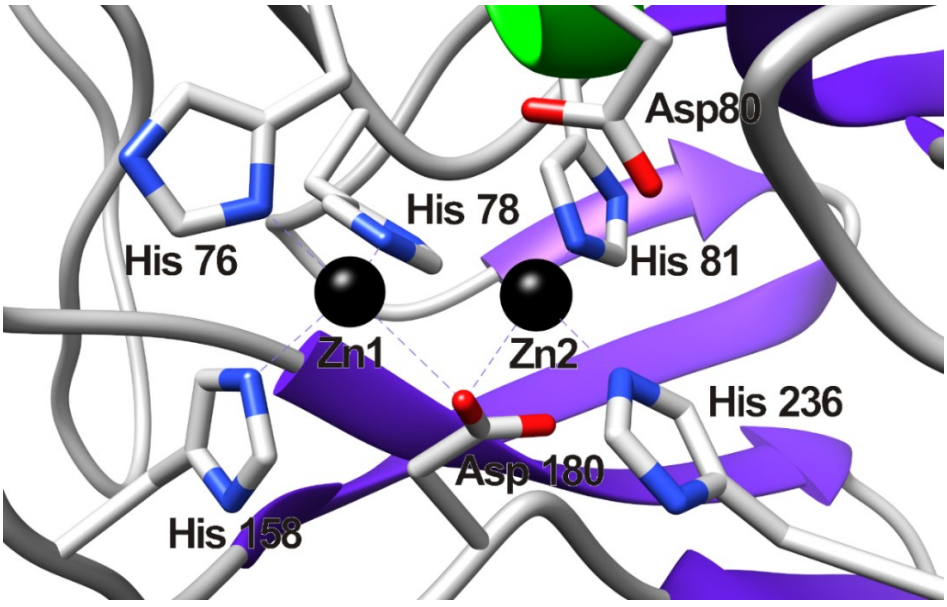


Figure 4.5: Zoom in of one of the active sites of the homology model of Rv1339 based on Bacillus subtilis YfhI (PDB: 6KNS – 27.02% identity). From the metal ion coordinating histidine and aspartate residues, Asp180 was chosen for mutagenesis because it is expected to be involved in the coordination of both zinc cations and is conserved across the metallo-β-lactamase family (PF12706)[340].

4.2 Expression of Rv1339 in *M. smegmatis* mc^{2}155

When the sequences of several Rv1339 homologues from mycobacteria were aligned (in MEGA7 with ClustalW and Muscle analysis), it became clear that there was an additional sequence of 17 amino acid residues at the start of the protein in *M. tuberculosis* and *M. bovis* BCG, but that was not found in the other aligned species of mycobacteria (for example, Figure 4.6 compares the sequence observed in *M. tuberculosis* H37Rv and *M. bovis* BCG with *M. smegmatis* mc^{2}155 and *M. marinum*).

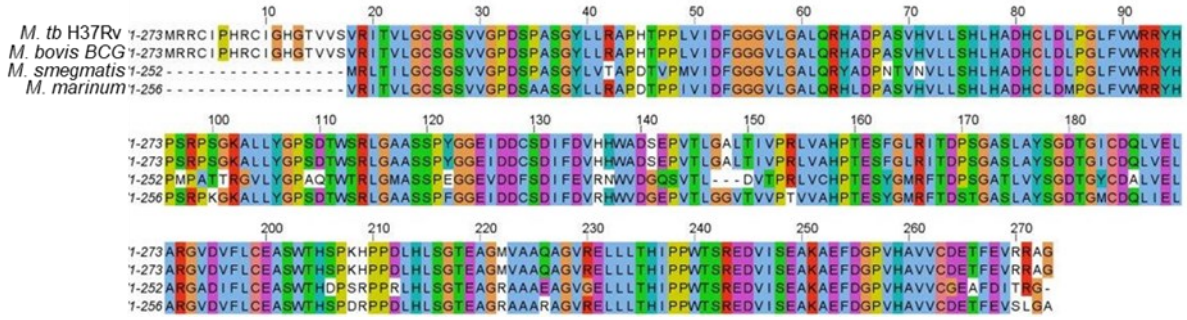


Figure 4.6: DNA nucleotide alignments of Rv1339 and some of its key homologues in mycobacteria.

A potential explanation for the additional 17 residues was that they represented a signal peptide sequence, involved in determining the localisation or secretion of the protein[343,344]. In order to investigate this, the FASTA sequence of the full-length Rv1339 protein was analysed by SignalP 5.0[345] (an online tool for analysis of protein sequences and identification of potential signal peptide presence -http://www.cbs.dtu.dk/services/SignalP/). This analysis identified a lipoprotein signal peptide at the start of Rv1339. This lipoprotein sequence would address this protein to the membrane or periplasm[346], if not for eventual secretion[347]. Lipoproteins are transported across the bacterial cytoplasmic membrane by the SEC translocon system, and cleaved by signal peptidase II (LspA)[348,349]. This system, and lipoprotein processing, has been shown to be dispensable *in-vitro* but is essential for virulence in *M. tuberculosis* H37Rv[347,348]. No signal peptides were detected when the FASTA sequence of Rv1339 lacking the first 17 residues was used as the SignalP input. As a result, in all of the mycobacterial species investigated, the Rv1339 homologue would begin transcription on a valine. It is not unusual for mycobacterial genes to begin translation on a valine (GTG) codon. This phenomenon has been predicted to occur with 38% of *M. tuberculosis* H37Rv open reading frames (ORF)[320]. Initiation of translation on a non-ATG start codon is postulated to be a gene regulation mechanism (as non ATG start codons show a lower efficiency of binding to translation initiation factors, like f-MET-tRNA[320]). This may explain why the Rv1339 homologue in other mycobacterial species would be predicted to begin on a valine.

As a result of the alignments and signal peptide prediction, it was decided that both full length (Rv1339L) and the shorter sequence lacking the "MRRCIPHERCIGHGTVVS" portion (Rv1339S) would be investigated – as cleavage of the signal peptide may alter the structure or function of the protein.

In order to determine if Rv1339 possessed 3', 5' – cAMP PDE activity, C-terminal H6 tagged *rv1339L* and *rv1339S* gene sequences were cloned into the mycobacterial shuttle vector pVV16 with Gibson isothermal assembly[350] and expressed in *M. smegmatis* mc^{2}155 bacteria. The vector contained both hygromycin and kanamycin resistance cassettes, a HSP-60 (GroELp) strong constitutively active promotor, as well as origins of replication for *E. coli* and mycobacteria (Figure 4.7). Subsequently, a point mutant was generated to ablate cAMP catalysis – this mutant will be characterised in the next Chapter but briefly, based on bioinformatic analysis, residue D180 was mutated to alanine by site-directed mutagenesis. This residue is responsible for stabilising both metal ions in the substrate binding and the active site of a mettallo-β-lactamase protein[340].

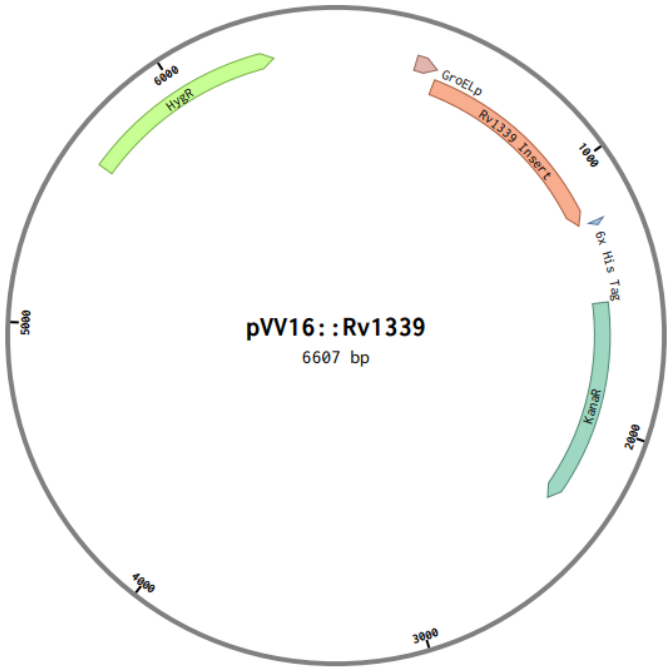


Figure 4.7: Plasmid map for the pVV16::rv1339 construct. pVV16 utilises the promotor from HSP60 (GroELp) to constitutively express the insert. Hygromycin and Kanamycin resistance casettes enable antibiotic selection in E. coli and mycobacteria.

Expression of all 3 of these constructs was confirmed by Western blot of clear soluble *M. smegmatis* mc^{2}155 lysate for the C-terminal H6 tag (Figure 4. 8).

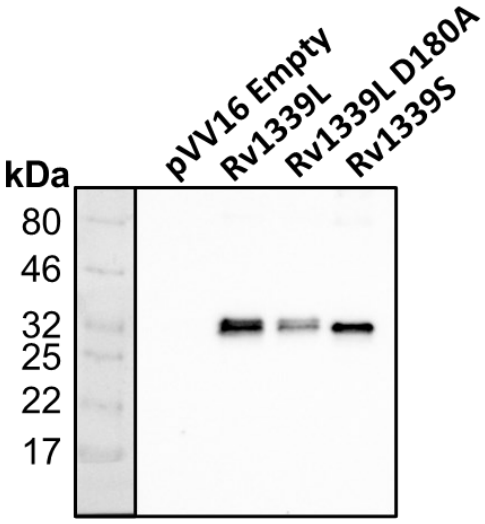


Figure 4.8: Western blot of pVV16 Empty Vector Control, Rv1339L, Rv1339L D180A and Rv1339S proteins in M. smegmatis mc^{2}155 soluble lysate. All constructs possessed a C-terminal His-6 tag and the western blot utilised an α-His antibody and was imaged by chemiluminescence.

The 1st lane of the Western blot corresponds to the Empty Vector Control with no insert – no band expected (Figure 4.8). Bands were detected in the 2nd, 3rd and 4th lanes and corresponded to the expected size of full-length Rv1339 (30.21818 kDa) in the Rv1339L and D180A (30.17417 kDa) expressing lysates and the smaller Rv1339S construct (28.53215 kDa, Lane 4, Figure 4.8). It appeared that in Rv1339L and Rv1339L D180A expressing lysate (lanes 2 and 3, Figure 4.8), there was a second band below the main band. This suggests that the protein is cleaved *in-vivo* or degraded as part of bacterial lysis. It may also indicate that this cleavage or degradation removes the 17 amino acid residues at the start of the protein – corresponding to the difference between the Rv1339L and Rv1339S construct. The presence of only a single band in the Rv1339S expressing lysate that corresponds to the Rv1339S construct size (28532.15 Da), would support this theory. Intact protein mass-spectrometry was originally planned for confirmation, however, logistical problems with the Imperial Proteomics Department delayed this un-workably and I proceeded to investigate the differences between the constructs phenotypically.

4.2.1 Rv1339L and Rv1339S display cAMP PDE activity and Rv1339L, but not Rv1339S alters the metabolome significantly

In order to understand the potential phenotypic impact of the 17 additional residues, targeted metabolomics analysis was first carried out on *M. smegmatis* bacteria expressing the Empty Vector Control (pVV16), Rv1339 or Rv1339S (Figure 4.9). This analysis showed that the levels of cAMP were significantly lowered when Rv1339L and Rv1339S were expressed in bacterial lysate. However, Rv1339L expressing lysate displayed a 10-fold decrease in cAMP levels, relative to Empty Vector Control lysate, when Rv1339S only mediated a 2-fold decrease. Ion counts were normalised to protein levels in each sample. This suggests that the Rv1339L construct was a more active cAMP PDE. This was an interesting observation, as it may imply that the first 17 N-terminal residues in full-length Rv1339 are involved in activity. From the predicted model, (Figure 4.2) it appears that this N-terminal section extends down the middle of the protein, and may potentially affect binding of substrate to the metal binding site.

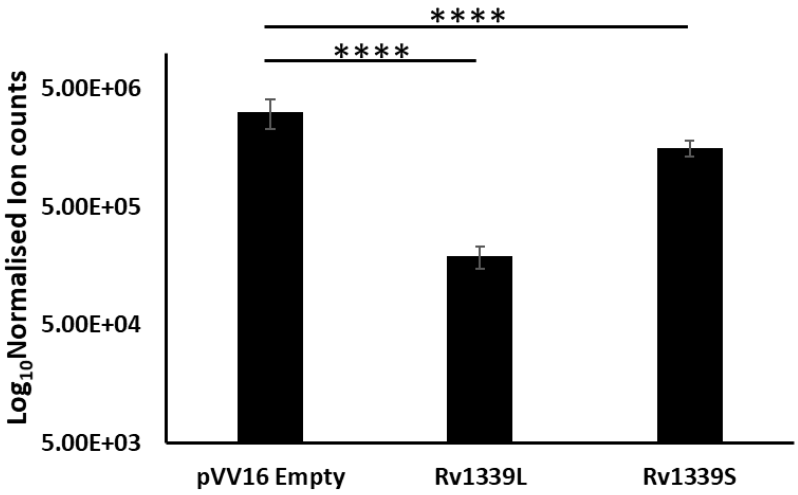


Figure 4.9: Targeted metabolomics analysis of cAMP levels in M. smegmatis bacteria cultured in 7H10 solid medium containing glycerol and dextrose as carbon sources. The bacteria expressed either pVV16 (Empty Vector Control), Rv1339L or Rv1339S. Ion counts are normalised to protein levels of each sample (from BCA quantification) and expressed on the log_{10} scale. 9 technical replicates were used of each strain. p< 0.0001 as measured by unpaired 2-tailed Student t-test.

Subsequently, untargeted metabolomic analysis was used to quantify the effect of expression of each on the metabolomes of the pVV16 Empty Vector Control, Rv1339L and Rv1339S bacteria (gated for 2-fold differential metabolite abundance, with $p<0.05$). From the PCA (Figure 4.10), it was clear that all 9 samples of Rv1339L could be separated from the pVV16 Empty Vector Control samples by multiple statistical components of the principle component analysis (PCA) (Figure 4.10). In contrast, the Rv1339S clustered closer to the pVV16 Empty Vector Control samples, but were still distinct. When

taken with the targeted data (Figure 4.9) this suggests that Rv1339S still has some cAMP PDE activity, but a reduced amount. Additionally, it suggests that Rv1339S expression has a lesser impact on the bacterial metabolome than Rv1339L expression – and this may be due to the decreased effect on the cAMP pool. A potential mechanism that will be investigated in Chapter 5 involves the cAMP binding transcription factors (Crp) effects on transcription when cAMP levels are altered. As a result, Rv1339L, the native full-length TB protein was chosen for purification (This Chapter) and in-depth phenotypic characterisation (Chapter 5). This was to ensure that the native protein was characterised, and to fulfil the 1st and 2nd objectives of my PhD – to characterise the cAMP PDE and use it as an effective tool to lower bacterial cAMP levels.

Figure 4.10: PCA representing the metabolomes of 9 samples of pVV16 (Empty Vector Control), Rv1339L or Rv1339S expressing M. smegmatis bacteria – grown on solid 7H10 media with glycerol and dextrose as carbon sources. Each dot corresponds to a technical replicate of either pVV16 (Empty Vector Control), Rv1339L or Rv1339S expressing bacteria. Rv1339L can be separated on 3 axis/components whereas pVV16 Empty Vector control and Rv1339S can only appear to be separated on 2.

4.3 Full length Rv1339 expression and purifications

In order to fulfil my 1st and 2nd PhD objectives, to characterise the enzyme and potentially acquire the crystal structure, my efforts to purify the enzyme had several requirements. Bioinformatic and evolutionary analysis coupled to literature searches revealed that no close homologues of Rv1339L (henceforth referred to as Rv1339 Full-length or Rv1339) had been crystallised at the time. Likewise, the Rv1339 branch of atypical class-II PDE enzymes was poorly characterised and may display functional heterogeneity to the prototypical class-II PDE members. As a result, it was considered important to obtain a crystal structure of the protein, and combine this with in-depth enzyme kinetics. Such analyses would require purified protein with a high homogeneity. For determining enzyme kinetic parameters, mathematical equations incorporating reaction substrate and product levels, temperature time and protein abundance are used. For example, the V_{max} represents the maximum amount of substrate that the enzyme can convert to product over time, at pre-set conditions[351,352]. In order to ensure accurate values from these calculations, a high homogeneity of purified protein would be required. Likewise, in order to ensure the greatest chances of successful crystallisation of Rv1339, a very high homogeneity of purified protein would be required[353]. As such, a target purity of 80% for enzyme kinetics and over 90% for X-ray crystallography was decided as the minimum required purity for successful purification of the protein. These values formed the 1st criteria for judging the success of my purification protocols.

Additionally, for enzyme kinetic and phenotypic characterisation, it was of paramount importance that the protein retained activity[316], was not denatured and was as close to the *in-vivo M. tuberculosis* H37Rv enzyme as possible. Retaining activity therefore formed the 2nd criteria for judging the success of a purification protocol. Furthermore, the final criteria were that the protein was soluble and that the purification was reproducible. Although techniques have been developed to reconstitute proteins *in-vitro*[354] or recover insoluble proteins from inclusion bodies[355], such approaches were considered to be unfeasible in the timeframe of my PhD. Additionally, CpdA from *Corynebacterium glutamicum*, was reported to be soluble – though with no specification of the threshold of purity reached[321].

The criteria that were decided to judge the success of my purification can be summarised as:

1. A minimum purity of 80% for enzyme kinetics, and/or over 90% for crystallography
2. The protein must retain activity
3. The protein must be expressed in the soluble fraction
4. The preparation that leads the purification needs to be reproducible

As such, the success of a particular purification approach was assessed against these criteria. Over the course of 3 years, I pursued many different approaches to fulfil these criteria. Any approach or purification that did not meet all of these criteria was discarded and a new approach was made the focus.

In the interest of clarity, I have included data that was used to decide the success or failure of each approach, or that potentially provided insight into the Rv1339 protein.

Of the many protein expression systems available[356], *E. coli* is likely the most widely used for bacterial protein expression and purification[357,358]. Purification of a recombinant protein in *E. coli* has the capability to efficiently produce high quantities of protein, with feasibility to purify this protein to a very high homogeneity[358]. This system was initially chosen for purification of Rv1339.

4.3 Rv1339 purification approaches in *E. coli*

4.3.1 Rv1339 expression in BL21 (DE3) STAR *E. coli* leads to toxicity and bacterial death

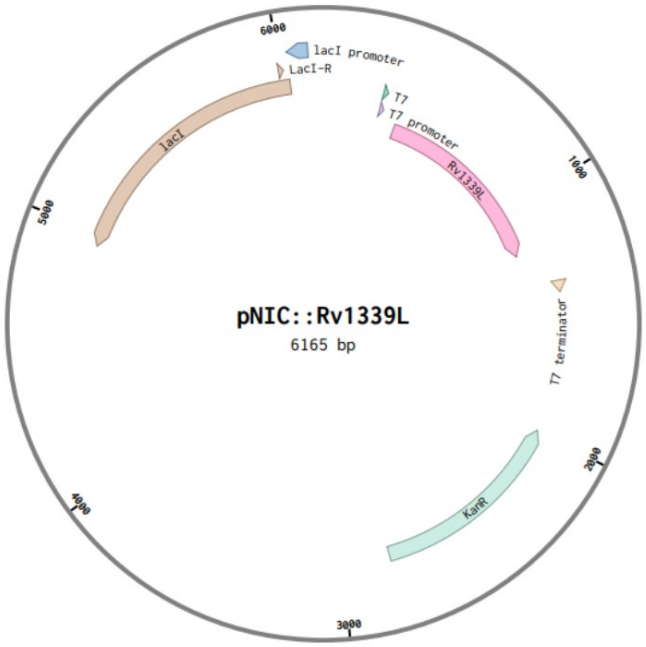


Figure 4.11: Plasmid map of pNIC::Rv1339L construct. An IPTG inducible T7 promotor regulated expression vector. Kanamycin resistance cassette and an origin of replication in E. coli allow for expression and purification of the encoded insert in E. coli.

In order to express Rv1339 in *E. coli*, pNIC::rv1339 N-terminal His-6 tagged constructs were generated (Figure 4.11). This IPTG inducible T7 promotor system *E. coli* shuttle vector was chosen based on its ability to exert tight repression of gene expression in the absence of inducer[359], a leaky promotor is often cited as a problem for inefficient recombinant protein expression[358].

These constructs were initially expressed in BL21 (DE3) STAR *E. coli* at both 16 and 37°C, in LB medium. When the expression cultures reached OD_{600} 0.6, Rv1339 expression was induced with addition of 0.1mM IPTG (a concentration validated in previous studies[359]). This led to a major drop in OD_{600} that stabilised at 6 hours (at OD_{600} 1) but that never recovered to the uninduced control levels – OD_{600} 2.4 (data not shown). The lower temperature of 16°C was used as a means of slowing down *E. coli* metabolism to try and prevent overexpression of Rv1339 causing toxicity.

When the soluble(S) and insoluble (M) cellular lysate were isolated and run on SDS-PAGE gels stained with Coomassie, it is clear that Rv1339 is expressed in the insoluble fractions. In addition, the construct may be leaky as Rv1339 appears to be detectable at 0 hours (Figure 4.12A and B).

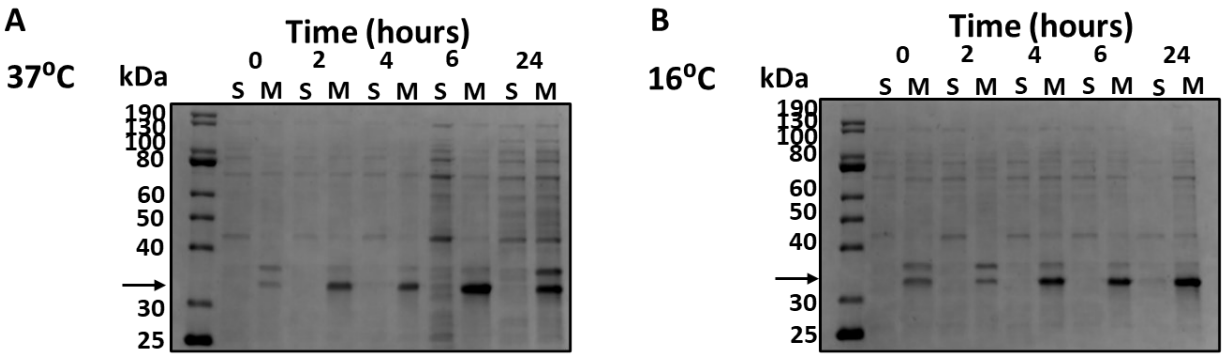


Figure 4.12: Coomassie stained SDS-PAGE gels of pNIC::rv1339 transformed bacteria lysates (Supernatant – S, Insoluble– M) at 0, 2, 4, 6 and 24 hours post IPTG induction of Rv1339 expression. The bacteria were induced at 37 (A) and 16°C (B).

3', 5' - cAMP is essential for *E. coli* to utilise non-glucose carbon sources[360–362]. Rv1339 may be decreasing 3', 5' – cAMP levels to the point where it becomes toxic to the bacteria. In an attempt to allow for this, three specialised expression strains of *E. coli* were used – C41, C43 and PlysS (BL21 (DE3) (Figure 4.13)

E. coli	Genetic Background	Source
DH5α	F–recA1 endA1 hsdR17 supE44 λ– thi-1 gyrA96 relA1 Δ(lacZYA-argF) U169 [Φ80lacZΔM15]	Lab collection
BL21(DE3)	fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5	Lab collection
BL21(DE3) PlysS	F– ompT hsdSB(rB–, mB–) gal dcm (DE3) pLysS (CamR)	Lab collection
BL21(DE3) C41	Δgor Δ(malF)3F´ lac, pro, lacIQ / Δ(ara-leu)7697 araD139 fhuA2 lacZ::T7 gene1 Δ(phoA)PvuII phoR ahpC* galE (or U) galK Δgor Δ(malF)3	Lab collection
BL21(DE3) C43	As C41 but additionally: ΔdcuS fur (GTC insertion) lacI G-T I Ion IS4 excission yjcO A-G cydA (IS1 insertion) ccmF ~ ompC* (IS1 deletion) yjiV ~ yjjN* (IS1 deletion)	Lab collection

Figure 4.13: E. coli strains that were used and their genetic backgrounds. DH5-α used exclusively for plasmid library generation

C41 and C43 are "Walker" strains of *E. coli* utilising a T7 RNA polymerase mediated expression system. They show resistance to protein overexpression toxicity when compared to standard BL21 *E. coli* strains and allow accumulation of overexpressed proteins in cytoplasmic aggregates[363].

Likewise, PlysS is a T7 RNAP expression system encoding *E. coli* strain. While expression of the T7 RNAP is under the control of the LacUV5 promotor in all 3 strains, PlysS features a modification of this promotor that leads to decreased dependence on carbon source and cAMP binding proteins for

initiation of transcription[364,365]. The intention of using these strains was to prevent the drop in OD_{600} associated with Rv1339 expression and to potentially express the protein in the soluble fraction.

Expression trials were repeated in these 3 expressers with and without 0.2% glucose (final concentration) in the medium. Again, 0.1mM IPTG was used as inducer and 16˚C was chosen as culture temperature. Unfortunately, although the OD_{600} of the cultures did not display any decrease upon Rv1339 induction, the protein remained insoluble. PlysS had the highest levels of expression and stable levels of growth. The results were similar between glucose (Figure 4.14A) and non-glucose media (Figure 4.14B).

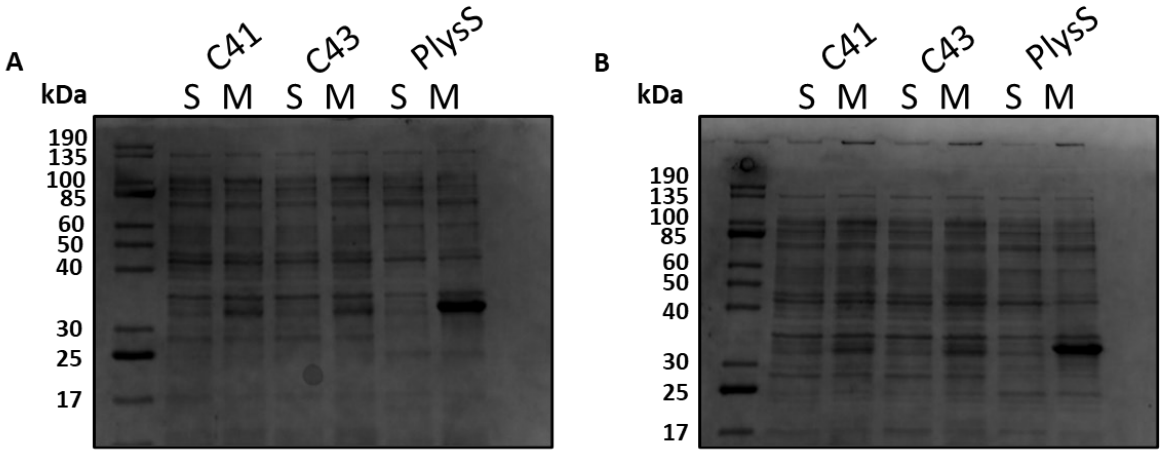
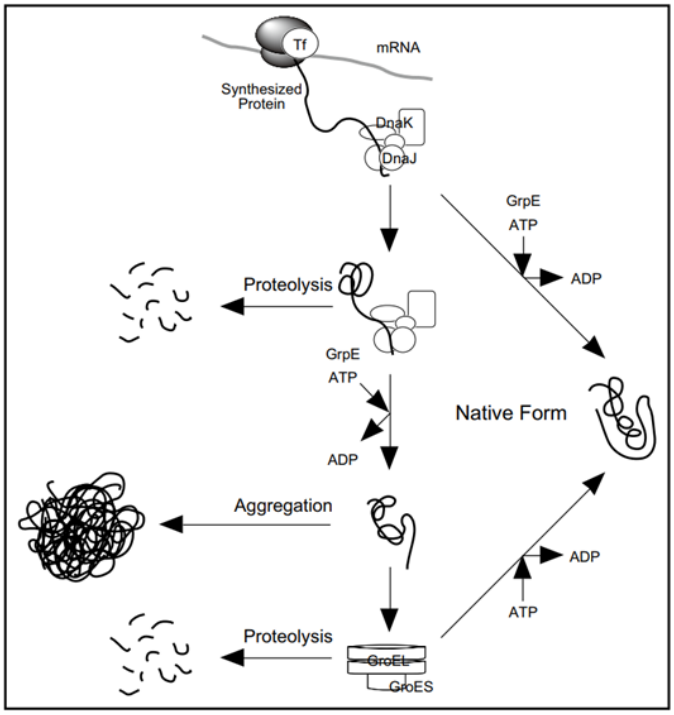


Figure 4.14: Coomassie stained SDS-PAGE gels of the lysate of three pNIC:rv1339 transformed toxic protein expression E. coli strains. Expression was carried out in non-glucose (A) or glucose containing media (B). S = soluble lysate, M = insoluble.

As the levels of expression and solubility were the same in glucose and non-glucose media, the hypothesis that Rv1339 was toxic to *E. coli* became more likely. The Walker strains were previously found to be more capable of expressing toxic proteins, but studies have noted that protein or plasmid overexpression related toxicity, is a key factor in successful expression[366]. To investigate this, 5 Takara protein expression plasmids containing different combinations of GroES, GroEL, DnaK, DnaJ-GrpE and Tig chaperones from *E. coli* were cloned into PlysS cells with Rv1339 and co-induced using the Takara protocol (Figure 4.15 – Chaperone plasmid set, Takara)

A



B

No.	Plasmid	Chaperone	Promoter	Inducer	Resistant Marker
1	pG-KJE8	*dna*K-*dna*J-*grp*E *gro*ES-*groEL*	*araB* *Pzt-1*	L-Arabinose Tetracycline	Cm
2	pGro7	*gro*ES-*gro*EL	*araB*	L-Arabinose	Cm
3	pKJE7	*dna*K-*dna*J-*grp*E	*araB*	L-Arabinose	Cm
4	pG-Tf2	*gro*ES-*gro*EL-tig	*Pzt-1*	Tetracycline	Cm
5	pTf16	tig	*araB*	L-Arabinose	Cm

Figure 4.15: Takara chaperone sets composition. A: Schematic indicating the different mechanisms by which each set of chaperone combinations can potentially intervene to prevent different causes of: proteolysis, aggregation or improper folding. B: Tabulated summary of the different chaperone combinations, their promotors system, induction and resistance marker.

Depending on which of these "chaperone teams" would enable soluble Rv1339 expression, it would give an indication of whether the problem was proteolysis, aggregation or improper folding[367,368]. The soluble (S) and insoluble (M) fractions from expression trials with these chaperone sets and Rv1339 are shown in Figure 4.16. Briefly, the PlysS *E. coli* strains were transformed with the different "chaperone team" encoding plasmids and subsequently transformed with pNIC::Rv1339 plasmids. Expression of the chaperone teams was performed at OD_{600} 0.2 with specified inducers and expression of pNIC::Rv1339 was induced at OD_{600} 0.8 with 0.1mM IPTG. All bacteria were cultured at 16˚C.

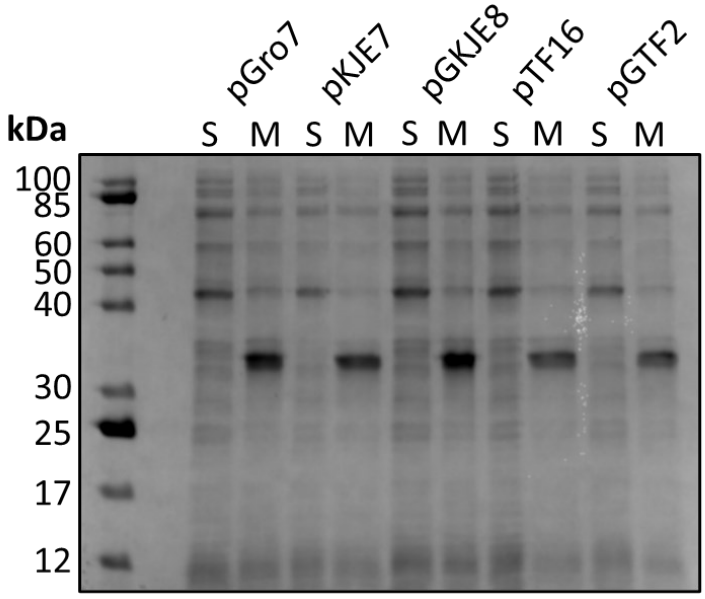


Figure 4.16: Coomassie stained SDS-PAGE gel of soluble and insoluble fractions from PlysS E. coli transformed with addition chaperones and expressing Rv1339. Specified inducer for each chaperone team encoding plasmid can be found in figure 4.15B. 0.1mM IPTG was used to induce pNIC::rv1339. All cultures performed at 16˚C. S = soluble lysate, M = insoluble fraction.

Similarly, to previous attempts, the protein remained insoluble even with the additional chaperones (Figure 4.16). Instead of focussing on optimising expression in this particular vector (e.g. by attempting a C-terminal His-tag), and to ensure the greatest chance of eventual success, it was decided that a high-throughput cloning approach was required – in order to generate as many different constructs as possible for expression trials. This new approach might identify a construct/condition that minimised any potential toxicity of Rv1339 to *E. coli*. For example, a C-terminal His-6 tag version could conceivably show better expression because the position of the N-terminal tag could be lowering protein thermal stability, or affecting other thermodynamic parameters that determine folding[369]. Alternatively, a fusion with GFP may help solubilise the protein, if poor solubility is the problem[370].

4.32 Fragment exchange (FX) cloning of Rv1339 for expression in *E. coli*

In order to enable the efficient production of different constructs of Rv1339 for expression, the high-throughput Fragment exchange cloning system was used. This approach involves the class-IIS restriction enzymes that are capable of binding to palindromic sequences and cleaving any 3 nucleotides next to their recognition site[371]. Fx cloning also makes use of subcloning into an initial vector (pINIT). This vector is not intended for expression but is instead used as the template to subclone the insert into subsequent vectors via a 2-step PCR mediated protocol[289]. The advantage of this technique is that if an insert is designed with the Class-IIS restriction site and 3 chosen nucleotides, the restriction enzyme will be cleaved away, leaving the overhang. The vector can be simultaneously digested in the same reaction, again leaving a complementary 3 base overhang – the 1st step. The 2nd step involves ligation

of the insert and vector, which are ligated at the complementary 3 base overhangs. Furthermore, the vector pINIT contains a *ccdB* gene[372], lethal to *E. coli,* that is removed only by successful digestion and ligation of the insert into the vector. This ensures that only a vector containing the desired insert is successfully transformed and selected. Fx cloning can be used for high throughput production of different constructs via the straight-forward PCR reactions, the restriction enzyme chosen was SapI, which allows expression of constructs with just 1 additional amino acid residue[371]. This method would enable a high-throughput screen of different constructs, as pINIT containing the desired insert can be simultaneously digested with any FX cloning expression vector (e.g. N or C terminal His-tag, GFP fusion or signal binding peptides).

It was decided that in addition to cloning Rv1339 (full length), that the Rv1339S (lacking the first 17 residues) and MS_4902 (the *M. smegmatis* homologue of Rv1339 – with a 97% identity to the Rv1339S) would be generated in each approach. This was to provide as many options as possible for eventual purifications, although, Rv1339 full-length was always the focus.

To begin, Rv1339, Rv1339S and MS_4902 were cloned into N or C terminal His-tagged and/or C terminal His-tagged GFP fusion constructs. Due to the robustness of the Class-IIS digestion and ligation system, construction of the desired vector and insert was confirmed by comparing BSA-1 digests of 2 mini-prepped colonies that had previously been transformed with the construct, and comparing this to a virtual digest (Data not shown). In the following *E. coli* Fx cloning approaches, the strain used Is MC1061[373] and the inducer is L-arabinose. The benefit of this strain was the presence of genes coding for stress response systems (such as RelA and SpoT) that can enable some tolerance to toxic constructs and due to the lack of interference with L-arabinose based induction systems[373].

With the high-throughput generation of different constructs, a high-throughput system to investigate expression was also required. The workflow used is summarised in Figure 4.17.

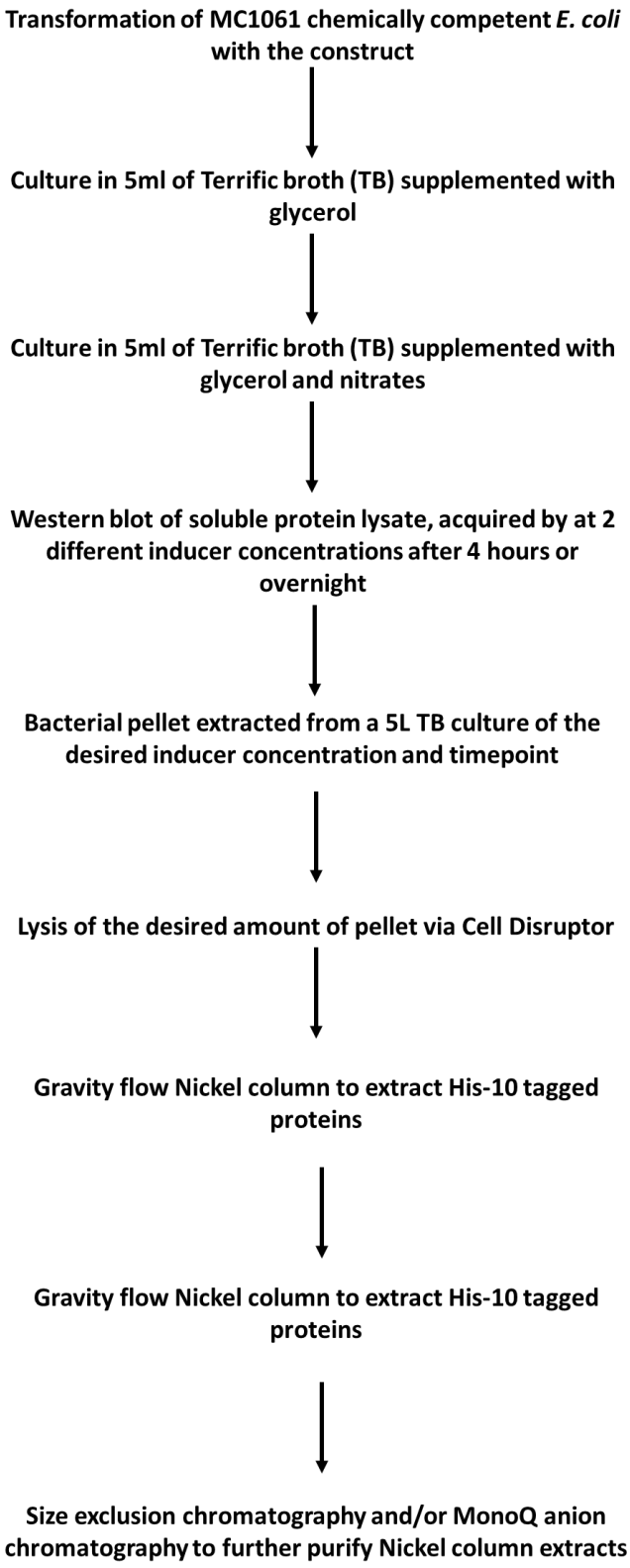


Figure 4.17: The standard workflow used to express and purify Fx cloning constructs. Terrific broth (TB) was chosen as media over LB as it has been shown to extend the E. coli log phase[374] and increase cell density/the quantity of protein extracted from plasmid mediated overexpression.

The first constructs tested were Rv1339, Rv1339S and MS_4902 in pBXNH3, pBXC3H or pBX3GH vectors. The intention was to provide the broadest range of options to enable eventual successful purification, although native Rv1339 was prioritised significantly (as it would be the most physiologically relevant). The vectors pBXN3H and pBXC3H were identical apart from the placement of a His-10 tag at the N or C terminal respectively. The vector pBXCGH fused the insert to a C-terminal His-10 tag and GFP (and therefore increased the expected molecular weight – Materials and Methods table 4). Between the insert and the His-10 tag or His-10 GFP, there was a C3 PreScission protease site to enable cleavage of these elements from the protein. These constructs were run through small scale (5mL) expression trials with a Western blot of the soluble cellular lysate as a readout. The two different methods involve expression at 25 or 37°C, overnight or over the course of 4 hours respectively. This was the standard suggested Fx cloning protocol and expression at 25˚C was intended to ease any potential toxicity from protein overexpression, by slowing metabolism[375]. The constructs were induced with L-arabinose at 0.001% or 0.01% final concentration (Figure 4.18).

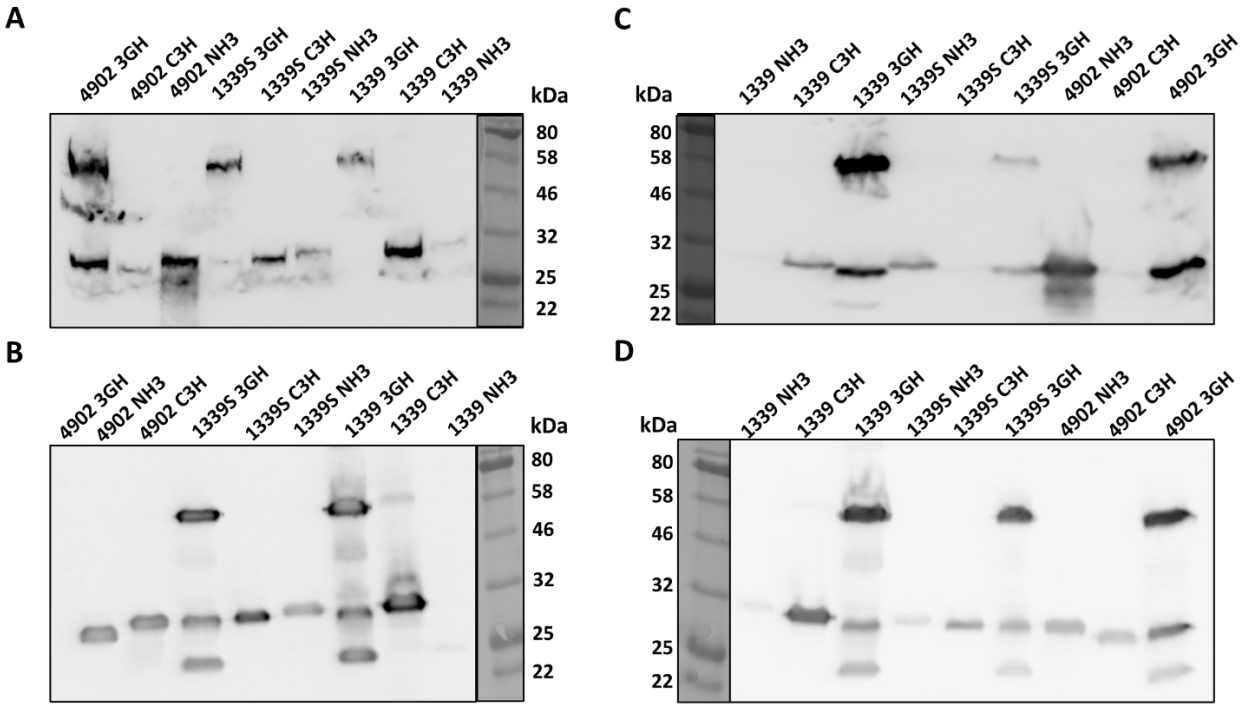


Figure 4.18: His-tag Western blots performed on soluble lysate of MC1061 E. coli transformed with Fx cloning constructs and induced with 0.01% (A and B) or 0.001% (C and D) L-arabinose. The upper blots (A and C) are from 4 hours post induction at 37°C. The lower blots (B and D) are from overnight expression at 25°C.

Interestingly, while there appeared to be expression of the Rv1339 C-terminal (1339 C3H) His-10 and C-terminal His-10 GFP (3GH) at all concentrations of inducer and at both time points, the Rv1339 N-terminal His-10 construct could only be seen at comparable levels at 4 hours post induction at the higher inducer concentration (Figure 4.18). This suggests that this construct may be degraded or misfolded,

and this could explain why expression of pNIC::Rv1339 N-terminal His-6 was only found in the insoluble fraction (Figure 4.12). Improperly folded protein might aggregate or display decreased solubility[369]. While small scale (5mL) expression trials displayed specific expression of each construct in soluble lysate, in order to yield enough protein for crystallography and enzyme kinetic studies, it is integral that the chosen construct can be scaled up. Rv1339 C-terminal His-10 (1339 C3H) and overnight expression at 25˚C with 0.01% arabinose yielded a blot with a good relative abundance when compared to other constructs. This construct was therefore chosen for larger scale expression trials.

From a 5L culture induced and cultured in Terrific Broth overnight. The soluble cellular lysate (extracted from 15g out of 56g extracted from an OD_{600} 3.66 culture) was run through a 1 mL Nickel column (Figure 4.19) and then subsequently purified by either MonoQ anion exchange (Figure 4.20) or Size exclusion chromatography (Figure 4.21), in order to find the protocol that would produce the highest purity of overexpressed Rv1339. If necessary, to ensure purity, a sequential approach of both would be used. Cell disruption was used for lysis, as it allows efficient lysis of high volumes of bacteria at 4˚C. The mechanism involves forcing the bacteria through a small orifice at high-pressure (30 Kpsi), thereby compressing them and then allowing them to rapidly expand when they have passed through. This rapid change in pressure causes lysis as a result of the shear forces breaking apart the bacterial cell wall[314]

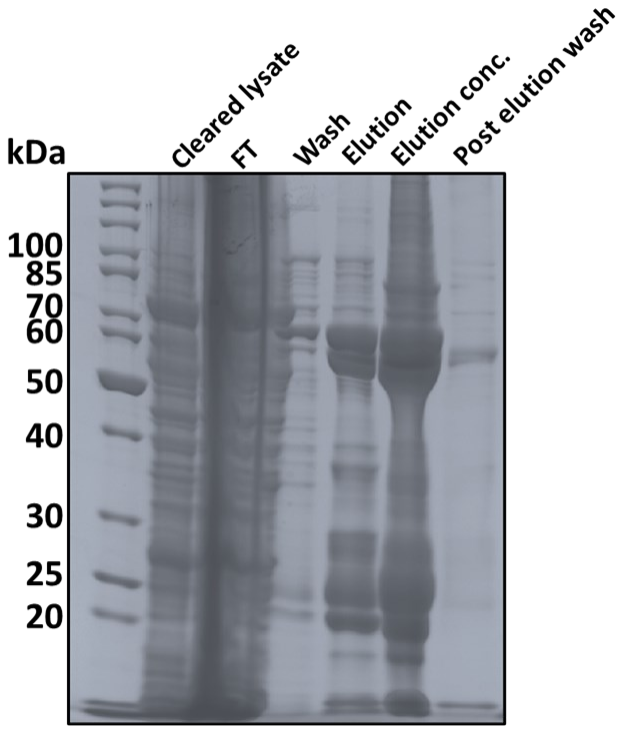


Figure 4.19: Nickel Affinity Column Purification of pBXC3H::Rv1339L in MC1061 E. coli. Cleared lysate, Nickel column flow through, Nickel column wash, Nickel column elution, concentrated nickel column elution and a post elution wash of the nickel column. FT: Flow through, Eluction conc: Ni column elution desalted and concentrated (Collaboration with Dr Michael Hohl, credit for the gel).

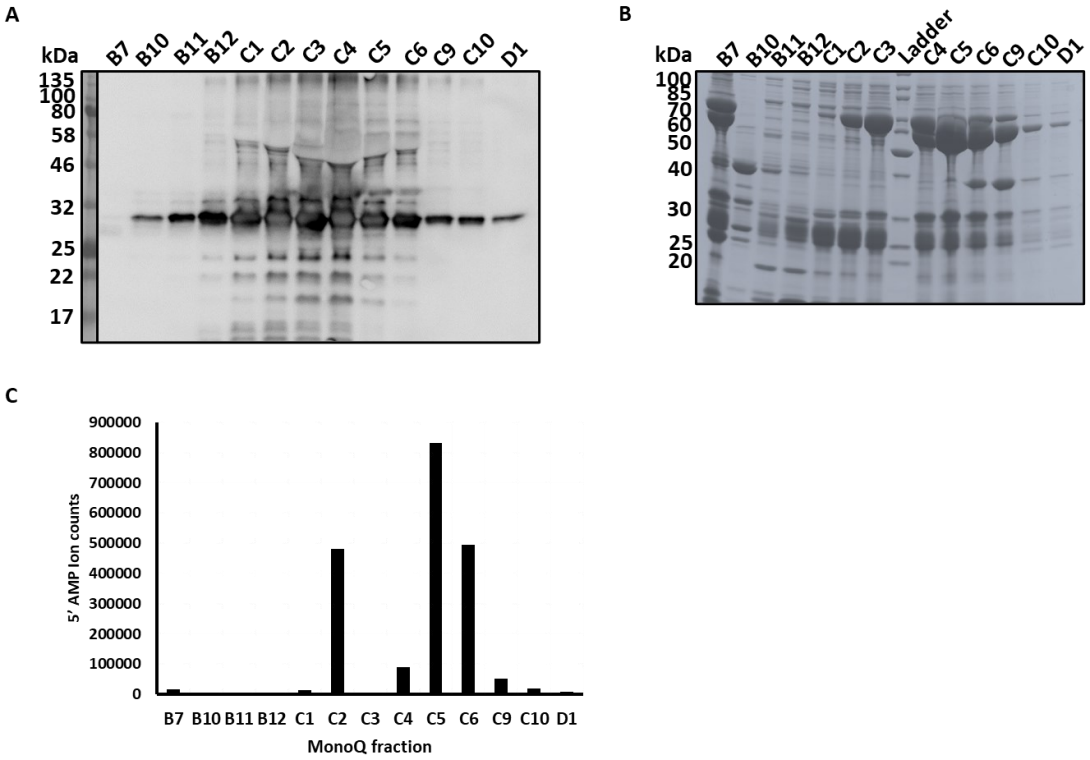


Figure 4.20: MonoQ purification fractions of the pBXC3H::Rv1339L construct. A: Western blot of MonoQ anion exchange fractions. B: Coomassie stained SDS-PAGE of MonoQ anion exchange chromatography fractions (Credit: Dr Michael Hohl). C: cAMP PDE activity assay of the fractions 15g of bacterial pellet was lysed via cell disruption, clear soluble lysate extracted by centrifugation.

A MonoQ step after the nickel column proved insufficient to purify Rv1339 to any significant degree; the western blot and SDS-PAGE of this purification (Figure 4.20A and 20B) showed a band approximately corresponding to His-10-Rv1339 in all fractions assayed with a bell-curve-like profile and with myriad additional bands. Additionally, cAMP PDE activity was found in at least 5 fractions via LC-MS quantification of 5' AMP levels after the activity assay (Figure 4.20C). Furthermore, in addition to thick bands corresponding to the size of the Rv1339 C3H construct, there were thicker bands present at around 60 kDa. Initially I postulated that the bands at 60 kDa were the Rv1339 C3H construct in a dimeric form (MW 63.13522 kDa). This could be explained by the presence of multiple His-10 tags, binding with higher affinity to the nickel column. Regardless, the MonoQ step was insufficient to purify Rv1339.

A Size exclusion step (Figure 4.21) (with an S200 increase for broad mass separation) following the nickel column, was attempted next instead of another the anion exchange. 15g of bacterial pellet from the same 5L culture as the MonoQ was again fractionated with the 1mL gravity flow nickel column. The SEC yielded a higher level of purity, with fewer unexpected bands. In addition, the levels of 5' -AMP production from the fractions were investigated using the same PDE activity as previously described (Figure 4.21).

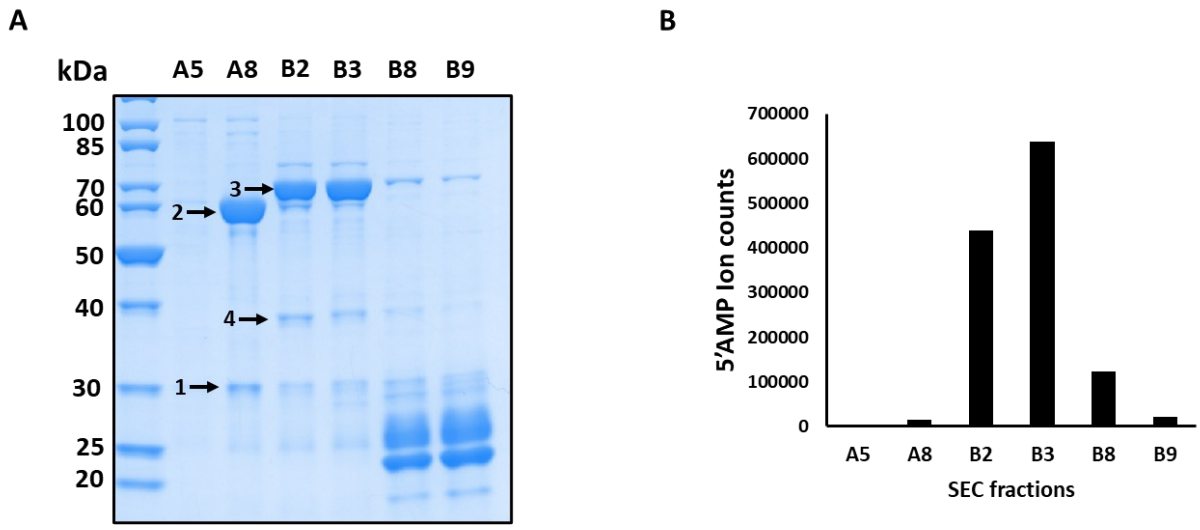


Figure 4.21: Coomassie stained SDS-PAGE gel of protein containing fractions eluted from the SEC chromatography column. The levels of 5' - AMP formation in the PDE activity assay performed on each fraction is shown below. Numbers 1-4 indicate bands that would be identified by proteomic mass-spectrometry.

5' - AMP production was found to be highest in fractions B2 and B3, however, this seemed unusual as these fractions displayed more heterogeneity, and had a less intense band corresponding to the size of monomeric Rv1339 C3H (MW 31.56761 kDa) than the band displayed in A8. B2 and B3 displayed additional bands at around 40 kDa and 70 kDa, and a band at the size of the potential dimeric form of Rv1339 C3H (MW 63.13522 kDa – Figure 4.21A).

To investigate the strong bands seen in the earlier preparations, and the new bands identified in this preparation (indicated in Figure 4.21 with arrows), these samples were re-run on an SDS-PAGE gel, cut and sent for proteomic analysis at the St Andrew's University Proteomics facility.

The proteomic analysis revealed the identity of each of the 4 bands.

1. pBXC3H Rv1339 L construct
2. GroEL (HSP65) chaperone – *E. coli*
3. Glutamine-fructose-6-phosphate aminotransferase – *E. coli*
4. Glycerol dehydrogenase – *E. coli*

From the proteomic analysis it is clear that the most pronounced bands (1 and 3) as well as band 4 are proteins up-regulated in response to stress or glycerol metabolism in *E. coli*[376,377]. The presence of such bands might indicate that induction of Rv1339 expression is highly stressful for the bacteria. This may be due to the protein being misfolded and toxic or to metabolic changes induced by reduction of cAMP levels. In this theory, higher levels of soluble Rv1339 expression could conceivably be deleterious, this would explain why a 10-fold increase in inducer concentration did not lead to a noticeable increase in

expression in earlier trials. Alternatively, 3', 5'-cAMP could be involved in the mechanisms of these stress enzymes – with depletion thereby leading to altered expression. Due to the low abundance of Rv1339 in the fractions (around 5% - when Imagej was used to quantify the relative intensity of the bands), it is likely that these proteins out-compete Rv1339 for binding sites in the nickel column, thereby leading to non-specific binding[378]

To further confirm the presence of Rv1339, Western blots of both preparations were performed. In both preparations, C-terminal His10-tagged Rv1339 was only observed at the monomeric size (SEC S200 "Increase" – Figure 4.22). Interestingly, monomeric Rv1339 also appeared to be present in fractions where no cAMP PDE activity had been detected (e.g. A5) and it appeared to be expressed at comparable levels across fractions A8-B8, despite differing amounts of activity being observed (Figure 4.21B). This might indicate aggregation or improper folding of the protein, precluding activity.

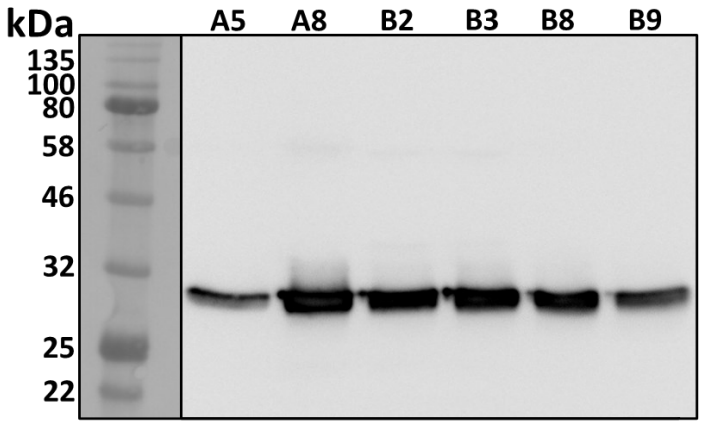


Figure 4.22: Western blot of Size exclusion fractions probed with a α-His tag antibody. Collaboration with Dr. Hohl.

With the high relative abundance of stress proteins as compared to Rv1339 C3H, it was clear that this approach would not be sufficient to express large quantities of homogeneous Rv1339. In order to overcome this toxicity or lack of solubility, a new approach was required. A potential solution was to express Rv1339 in the periplasm of *E. coli*. The oxidising environment of the periplasm can benefit proteins with cysteine residues (8 in Rv1339 at position C4, 9, 25, 82, 129, 154, 169 and 263) by enabling the formation of disulphide bonds[379]. In the periplasm, Rv1339 might also be less able to affect the metabolome or physiology of *E. coli*.

4.3.3 Expression of Rv1339 in the periplasm of *E. coli*

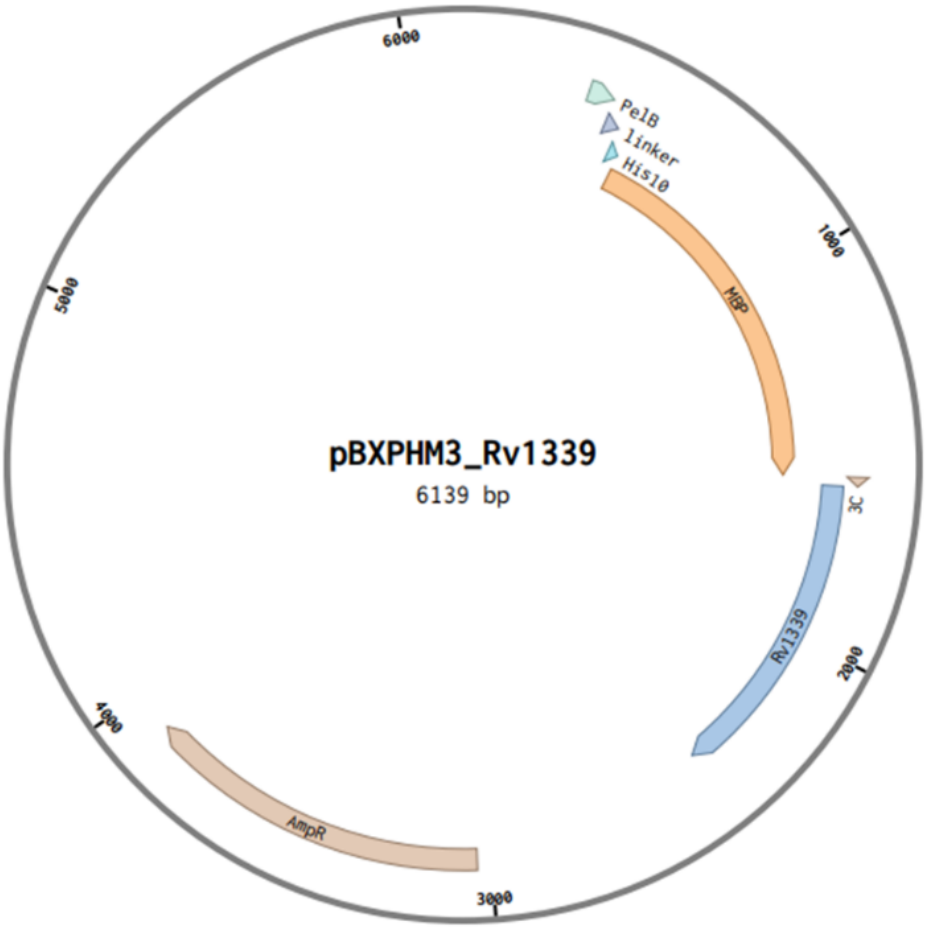


Figure 4.23: Plasmid map of pBXPHM3::Rv1339L. This plasmid utilises a mannose binding protein (MBP), PelB signal sequence and a linker region to express a His10 tag insert in the E. coli periplasm. Between these adjuncts and the insert is a C3 PreScission protease cleavage site.

Rv1339 was predicted to be expressed in the insoluble membrane fraction during high-throughput TB protein expression studies[344,380]. Rv1339 also contains 8 cysteine residues, potentially allowing the formation of 4 covalent disulphide bonds (formed by oxidation of the thiol groups on 2 cysteine residues). If these bonds form between inappropriate cysteine residues e.g. in a non-native configuration, it can lead to improper folding and orientation of the protein[381]. This was considered as a potential explanation for the co-purification of Rv1339 with chaperones or stress enzymes and expression in the insoluble fraction. Taken together with the putative positioning of Rv1339 in the insoluble membrane fraction in *M. tuberculosis* H37Rv studies, a potential solution was to express the protein in the *E. coli* periplasm. The periplasm contains machinery to assist in the formation of disulphide bonds, proper folding and isomerisation of such proteins[382,383]. For example, the enzyme Disulphide Bond Isomerase (DsbC) which is expressed in the *E. coli* periplasm and can rearrange

disulphide bonds in order to ensure the correct/native bonds are formed[384,385]. Expression in the periplasm can therefore increase the yield of correctly formed disulphide bonds and can enable more efficient purifcation[381]. It has also been shown to enable expression of otherwise toxic proteins[383]. It was hoped that this approach would lead to increased purity.

To begin, small scale (5mL) expression trials with a Western blot readout were again performed, as before with the C3H, NH3 and CGH constructs. However, this time the analysis was performed on the insoluble fraction (without extracting the periplasm). The vector construct chosen was pBPHM3, and again Rv1339L, Rv 1339S and MS_4902 were chosen as inserts (to provide the broadest range of future purification options). This vector contained the N-terminal fusion: mannose binding protein (MBP), PelB signal sequence, flexible 14 amino acid linker sequence (in combination, these fusions would enable export to the periplasm) and a His-10 tag (Figure 4.23). These elements were separated from the insert by a C3 PreScission protease cleavage site. Western blot analysis (Figure 4.24) showed expression of the constructs after both 4 hours at 37˚C (Figure 4.24A) and overnight at 16˚C (Figure 4.24B). All constructs displayed a band at the predicted size (Rv1339 C3H PelB MBP: 75.62580 kDa). However, Rv1339 constructs displayed markedly less non-specific/degradation bands than Rv1339S and MS_4902 constructs (Rv1339S C3H PelB MBP: 73753.56 and MS_4902 C3H PelB MBP: 73.45113 kDa). It is possible that this is because the Rv1339S and MS_4902 constructs lack the first 17 amino acid residues that code for the Rv1339 signal peptide, targeting them to the cell wall – this may mean that when these constructs are found in the insoluble fraction, that it reflects enrichment of their degradation products. Alternatively, loss of this sequence may make the protein more vulnerable to cleavage or degradation. It was therefore decided that the Rv1339 full-length (1339L) construct would be taken forward. The bands from the overnight cultures also appeared more defined, and so this time point was chosen.

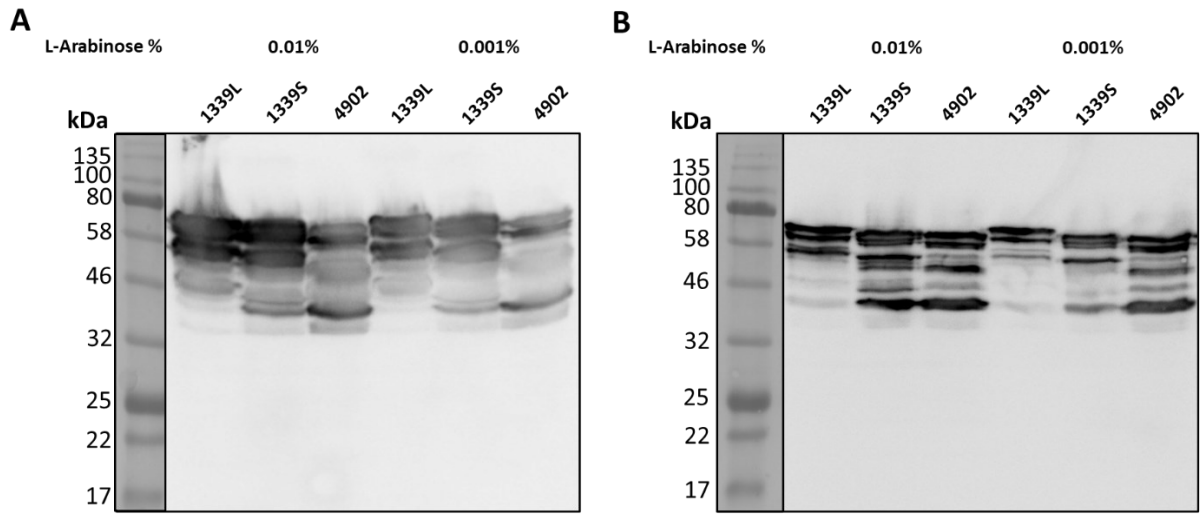


Figure 4.24: Western blot analysis of the membrane fraction from MC1061 E. coli expression pBXPHM3::Rv1339L, Rv1339S or MS_4902 constructs.

Next, the optimal extraction method for extracting the periplasm was investigated (Figure 4.25). In Figure 4.25A, expression in the total lysate was compared to the periplasm (obtained via periplasmic extraction osmotic shock buffer supplemented with 20% sucrose and 0.25g/l of lysozyme[386]) on cultures induced with different concentrations of L-arabinose overnight. This approach seemed to yield only degradation products or non-specific bands (Figure 4.25A).

However, a protocol for extracting the periplasm that did not involve lysozyme, instead using a combination of hypertonic and hypotonic solutions, appeared to yield full length protein at the mass expected[387] (MW 75.62580 kDa) (Figure 4.25B). There appeared to be slightly greater levels of expression at 0.01% of inducer, and so this extraction method, overnight culture and inducer concentration were taken forward to a 5L preparation.

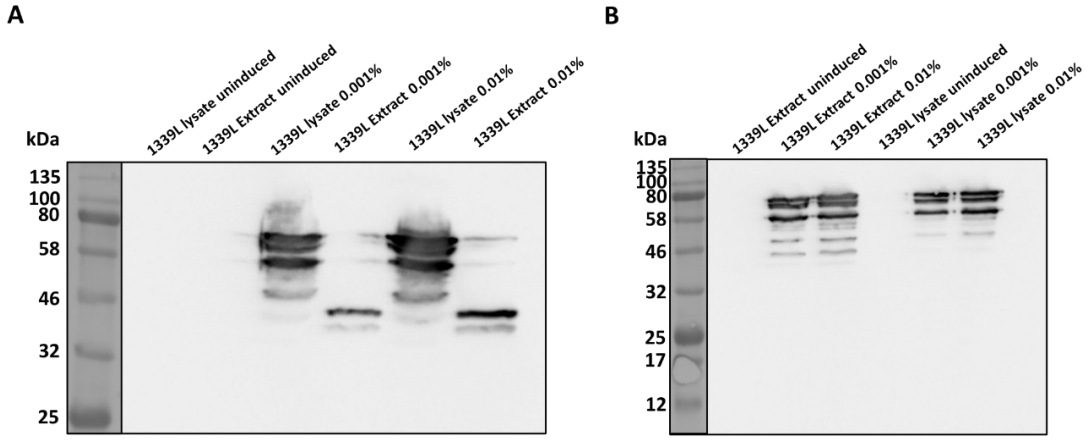


Figure 4.25: Western blot analysis of lysate and extracted periplasm of pBXPHM3::Rv1339L expressing MC1061 E. coli cells, at different inducer concentrations and using different extraction methods. A: Western blot with an α-His tag antibody of the lysozyme (0.25g/L) and 20% sucrose periplasm extraction method. Whole lysate is compared to periplasm extracts and at different concentrations of inducer. B: Western blot with an α-His tag antibody of the extraction method involving hypotonic and hypertonic solutions to extract the periplasm. Whole lysate is again compared to the periplasm extracted with this method.

From this 5L culture, the periplasm was extracted from 15g of bacterial pellet, and the remaining bacterial fraction was lysed via cell disruption. The periplasmic extract was fractionated with a Nickel affinity gravity column and desalted. To increase purity, N-terminal linker, MBP, PelB and His-10 tag were cleaved off overnight (with the C3 PreCission protease cleavage site) and then this solution was "reverse-purified" through a new 1mL nickel affinity gravity column (this time the desired protein should be in the flow through). Unfortunately, it seemed that the resulting flow through displayed low protein abundance and did not appear to be sufficiently purified (Figure 4.26, red box). The presence of a band that may correspond to Rv1339 in the flow through of the 2nd Nickel column and in the subsequent strip of this column may indicate that the cleavage was not successful.

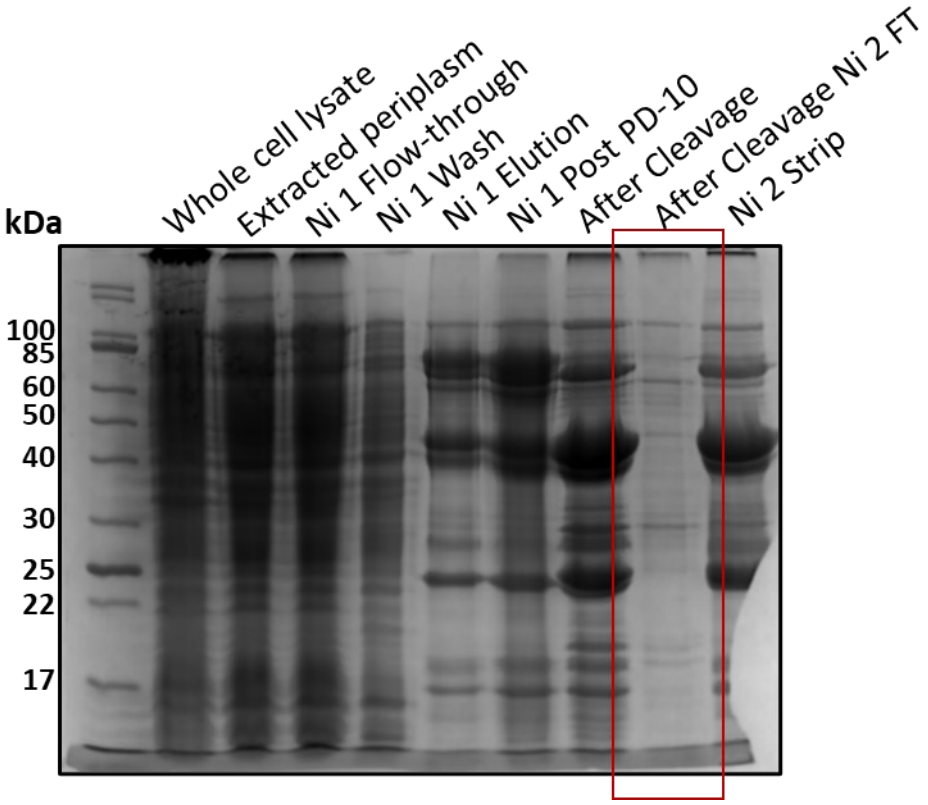


Figure 4.26: Periplasm purification from 15g of bacterial pellet expressing pBXPHM3::Rv1339L. Ni 1 – the first Nickel affinity purification. PD-10 is a desalting column. A second nickel column purification (Ni 2) was performed after cleaving the MBP-PelB-Linker-H10 tag. Ni 1 Post PD-10: Ni 1 elution after desalting column, Ni 2 strip: high imidazole concentration elution of the remaining bound proteins.

With the lack of success purifying native Rv1339 sufficiently from the *E. coli* periplasm, and concerns that the activity or folding of Rv1339 may be toxic to *E. coli,* it was decided that purification of a catalytically inactive mutant would be attempted.

4.3.4 Expression of an activity mutant – Rv1339 D180A in *E. coli*

E. coli have been shown to require cAMP signalling to metabolise non-glucose carbon sources[360,361]. As Rv1339 was seen to decrease cAMP levels in mycobacteria earlier, this activity in *E. coli* may lead to toxicity. Therefore, purification of the pBXC3H::Rv1339 D180 cAMP catalysis mutant (characterised fully in Chapter 3) from MC1061 *E. coli* could potentially ablate any toxicity associated with decreasing cAMP levels.

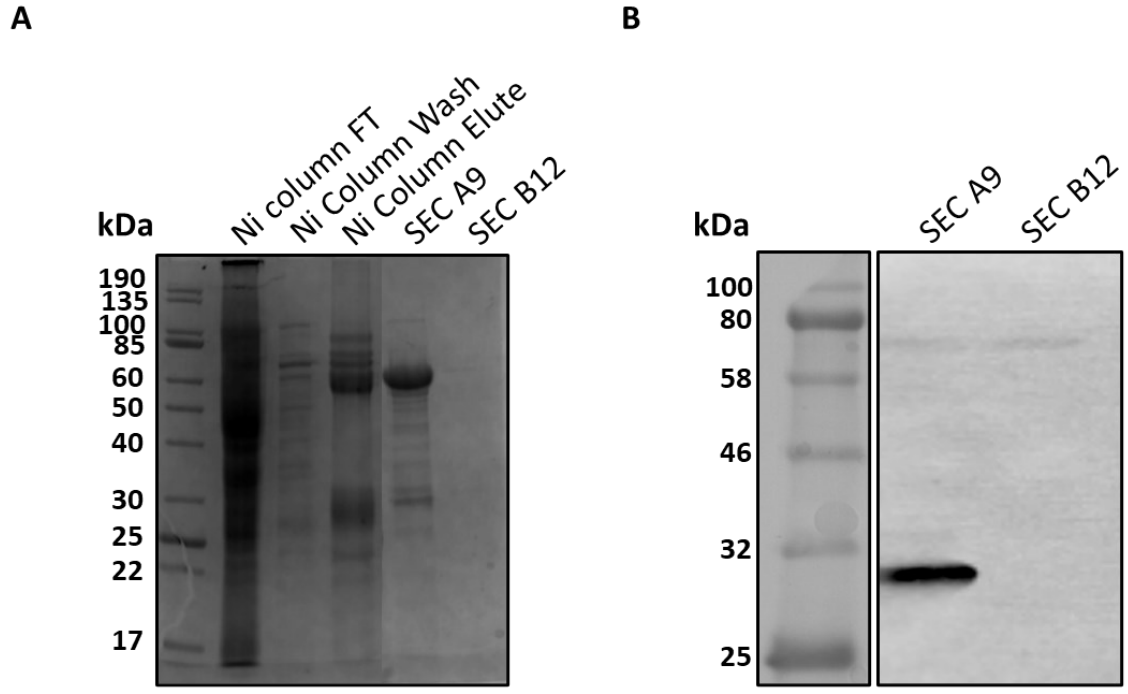


Figure 4.27: Coomassie stained SDS-PAGE and Western blot analysis of the pBXC3H::Rv1339L D180A purification. A: Coomassie stained SDS-PAGE of the gravity flow Nickel column – consisting of samples taken from the key steps and 2 fractions of the subsequent Size Exclusion Chromatography step. B: Western blot of the Size Exclusion Chromatography step with an α-His tag antibody.

The pBXC3H::Rv1339L D180A construct was cultured, the cell pellet was extracted and 15g was lysed via cell disruption with the same protocol as the earlier pBXC3H::Rv1339L construct purification (in order to allow comparison between the active and inactive constructs). The clear soluble lysate was fractionated with a gravity flow nickel affinity column and subsequent size exclusion chromatography (S200 Increase) (Figure 4.27A). Western blot analysis confirmed the presence of Rv1339 D180A in SEC fraction A9 (Figue 4.27B). However, the Coomassie stained SDS-PAGE suggested that although the purity of Rv1339 D180A was higher than when Rv1339 was tentatively purified with a similar protocol (10% via imageJ analysis of band intensity), that it is still not sufficiently homogeneous for enzymology or crystallisation studies (80% or 90% purity required, respectively). Likewise, the most prominent contaminating band once again appeared to be at the size of a chaperone, such as GroEL.

Interestingly, when the SEC chromatograms from the pBXC3H::Rv1339 D180A purification (Figure 4.28B) and pBXC3H::Rv1339 (Figure 4.28A) are compared, it appears as though there is less heterogeneity in protein species eluted – as indicated by UV readings (despite similar UV abundance). This may indicate that the cAMP hydrolysis activity of active Rv1339 induces a stress-response that leads to the upregulation of stress related enzymes (such as those identified in the earlier proteomic analysis). However, in order to properly compare the effects of active vs inactive Rv1339 expression in *E. coli,* a more in-depth proteomic MS analysis would have been required, but this was considered beyond the scope of my PhD.

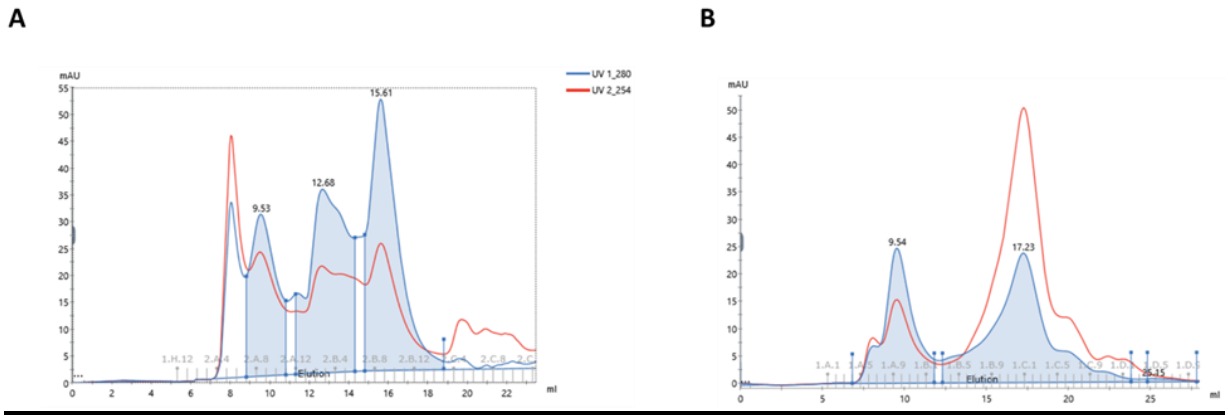


Figure 4.28: Chromatograms from SEC purification of Nickel affinity gravity flow column of purified pBXC3H::Rv1339 (A) and pBXC3H::Rv1339 D180 (B) constructs.

To summarise the *E. coli* purifications, it appeared after the various approaches that Rv1339 would prove difficult to purify in *E. coli*. This may be due to toxicity, or just the act of purifying a mycobacterial protein in another species. Although the FX cloning system was initially developed in *E. coli,* mycobacterial shuttle vectors were also recently developed for the system[375]. With the lack of success in sufficiently purifying Rv1339 in *E. coli*, expression and purification of these constructs was investigated in *M. smegmatis mc²155* – with the newly described mycobacterial FX cloning constructs[375]. The suitability of *M. smegmatis* mc²155 for recombinant mycobacterial protein expression is reviewed in the discussion.

4.4 Purifications of Rv1339 in mycobacteria

4.4.1 Mycobacterial Fx cloning expression vectors

To begin, 3 mycobacterial FX cloning vectors were considered. pMEX-, pMyc- and pMINT-C3GH. All 3 contained a C-terminal GFP and His-10 tag separated from the insert via a C3 Prescission protease cleavage site. Expression was controlled by an anhydrotetracycline inducible system in pMEX and pMINT. However, pMyc was induced by acetamide[375] (Figure 4.29).

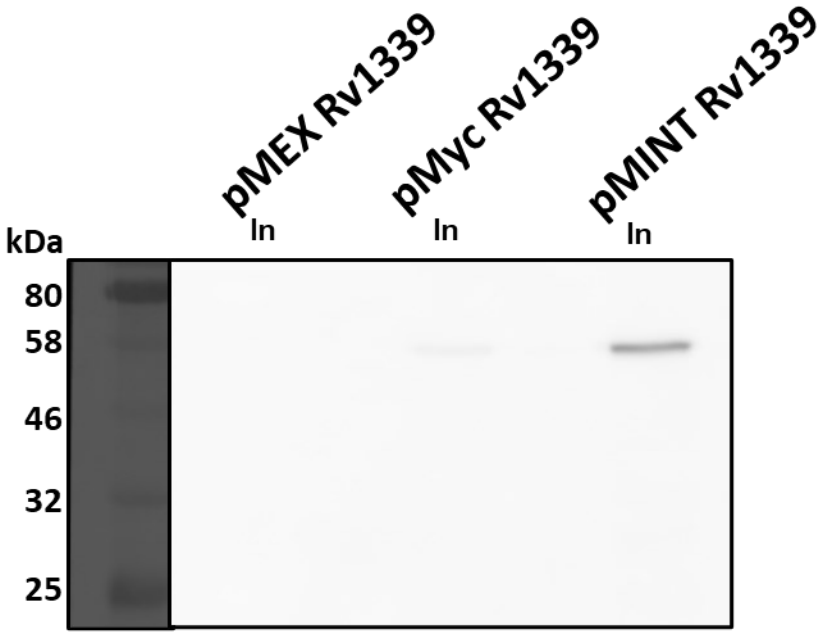


Figure 4.29: Western blot of different FX cloning mycobacterial constructs of Rv1339L. All 3 vectors add a GFP-H10 sequence to the insert separated by a C3 cleavage site. pMEX and pMINT are induced by Anhydrotetracycline, whereas pMYC is acetamide induced (all indicted with "In" labelled lanes).

Since the purifications were being attempted in *M. smegmatis* mc^{2}155, only Rv1339 full-length was initially considered – again, as characterising the native protein was the priority. Western blot analysis of the small-scale expression cultures only showed bands for the proteins expressed in pMINT and pMyc (Figure 4.29). pMINT::Rv1339 appeared to display the highest level of expression (MW of Rv1339-H10-GFP: 59.04945 kDa) and was therefore selected for further purification attempts.

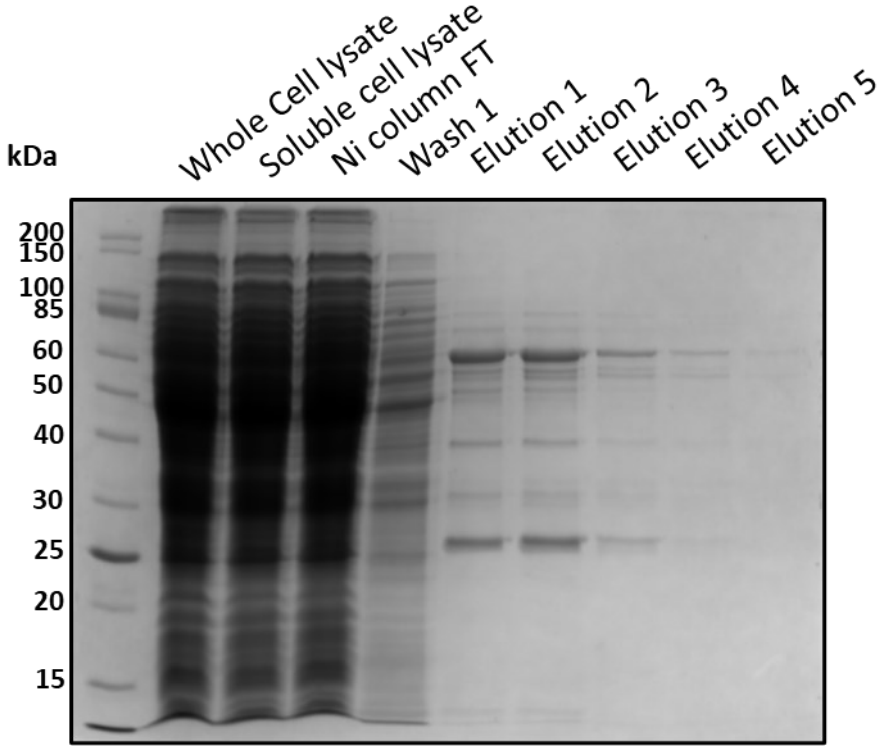


Figure 4.30: 1mL Nickel affinity gravity flow column purification of pMINT::Rv1339. gradient (Elution 1 = 100mM, 2 = 200mM, 3 = 300mM, 4 = 400mM and Elution 5 = 500mM Imidazole).

M smegmatis mc^2155 bacteria transformed with the pMINT::Rv1339L construct were cultured and induced at OD_{600} 0.6. Clear soluble lysate was obtained after overnight expression via cell disruption. This lysate was purified via a gravity flow nickel affinity column, with incremental elution's along an imidazole gradient (Elution 1 = 100mM, 2 = 200mM, 3 = 300mM, 4 = 400mM and 5 = 500mM Imidazole) (Figure 4.30). It appeared that at elution 1 and 2, there was a thick band corresponding to the expected size of the full length Rv1339 GFP C-terminal His10 tagged construct – 59.04945 kDa. In addition, it appeared that there were bands corresponding to the size of Rv1339, though this may have been due to degradation from boiling and denaturing prior to SDS-PAGE analysis, or from during initial lysis.

Unfortunately, none of the Nickel column purified fractions or the soluble cell lysate displayed cAMP PDE activity by LC-MS analysis. It was considered likely that this was a result of fusion to GFP. It is possible the bulky GFP protein occluded substrate binding – although it would not appear likely from the predicted model of Rv1339 (Figure 4.2). To test this, a further expression and purification was carried out – this time with cleavage of the C-terminal GFP and H10 tag (Figure 4.31).

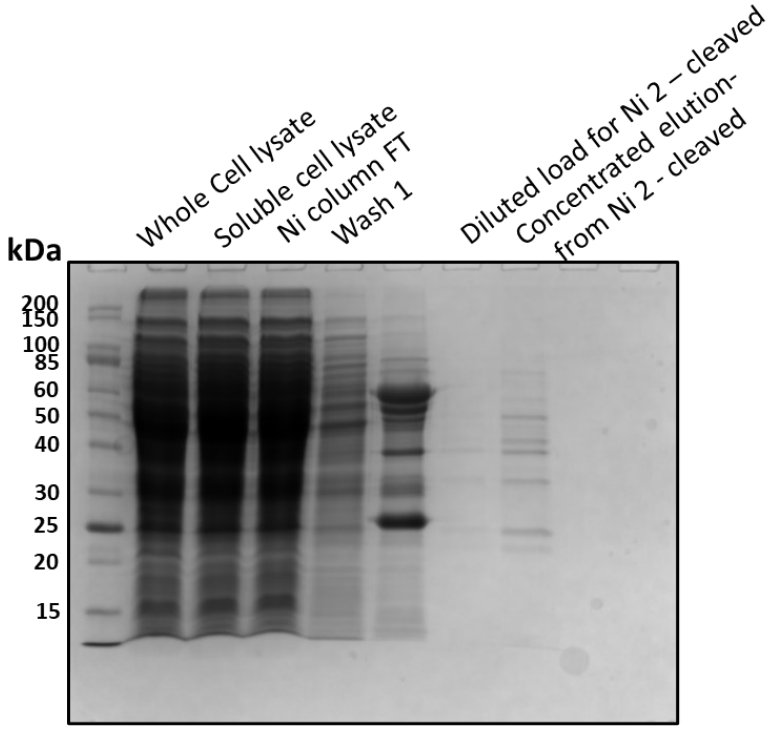


Figure 4.31: A second purification of pMINT::Rv1339L.

M smegmatis mc^{2}155 bacteria transformed with the pMINT::Rv1339L construct were again cultured and induced at OD_{600} 0.6. Clear soluble lysate was obtained after overnight expression, by cell disruption and centrifugation. The clear soluble lysate was fractionated with a 1mL gravity flow nickel column. Fractions Ni 1 and Ni 2 were concentrated and buffer exchanged to remove the Imidazole and allow optimum conditions for C3 PreCission protease activity. The GFP His-10 tag portion was then cleaved overnight at 4˚C with C3 PreScission protease. Unfortunately, neither fraction displayed cAMP PDE activity – and the band corresponding to the predicted size of Rv1339 did not appear be sufficiently purified. This is likely due to improper cleavage, or degradation of the protein when it was no longer fused to GFP.

To summarise the mycobacterial FX cloning purification attempts, it appeared that purity could be increased when compared to *E. coli* (25% by imageJ analysis in Figure 4.31). However, expression of Rv1339 full-length with GFP seemed to ablate activity, before and after cleavage. This may indicate that the C3 PreCission protease either displays poor activity or in some way degrades Rv1339 – although no recognition site was found within the sequence. Much of the Fx cloning work was carried out in collaboration with Dr Michael Hohl from the Meier group at Imperial (credit indicated in relevant figures). Unfortunately, neither of our groups possessed a positive control for the C3 PreCission protease and therefore we did not investigate this further. Instead, I moved on to try purifying full-length Rv1339 with the constitutive expression vector, pVV16. This vector had previously shown successful expression of Rv1339 (Figure 4.7) and was concurrently being used for phenotypic characterisation of Rv1339 (Chapter 5).

4.5 Purification of Rv1339L in mycobacteria with the constitutive expression vector pVV16

Purification of Rv1339 would be an essential step in characterising this enzyme via kinetics and X-ray crystallographic studies. Furthermore it is a necessary step in confirming the presence of PDE activity *in-vitro*. With the observed soluble expression of Rv1339 in *M. smegmatis* mc^2155 (Figure 4.7), and accompanying phenotypes (Chapter 5), it was decided that purification would be attempted from a pVV16::Rv1339 construct.

To begin, a 2.4L culture of pVV16::Rv1339 transformed *M. smegmatis* mc^2155 was cultured in LB medium, lysed via sonication and the soluble fraction was run through a gravity flow Nickel affinity column (5.7g of bacterial pellet). The fractions collected from this column were then run on an SDS-PAGE gel and stained with Coomassie (Figure 4.32).

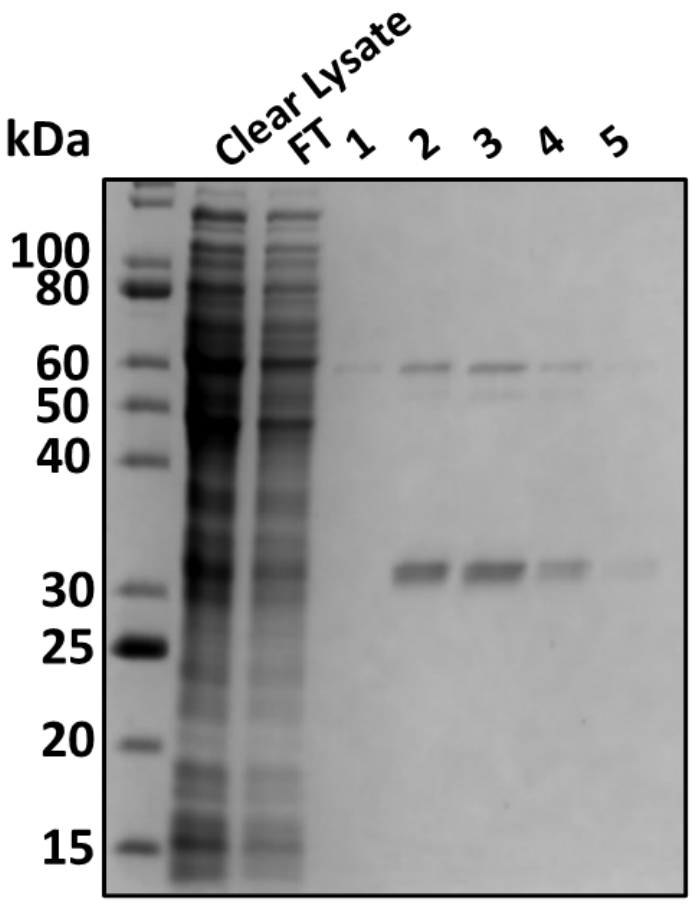


Figure 4.32: Coomassie stained SDS-PAGE of clear soluble lysate from MSMEG expressing Rv1339. Although relatively pure, there appears to be evidence of degradation. 1-5 indicate the 1.5 mL fractions collected from elution of the nickel column, with 250mM imidazole.

The yield seemed highly pure with a band for the Rv1339 monomer and potentially a dimeric form of Rv1339. Even if the 60 kDa band was GroEL (as with the *E. coli* purifications), a band at the size of the Rv1339 monomer (30.21816 kDa) was still the most prevalent in the fraction. While this homogeneity was promising (~80% by imageJ analysis), the yield of protein was not (200 µl of 0.24mg of protein per mL). This preparation would need to be scaled up in order to extract enough protein for crystallography and enzymology studies. To this end, 50L of the Rv1339 expressing *M. smegmatis* mc^2155 mutant was cultured in a bioreactor (BR). 134.68g of cell paste was collected and kept at -80°C.

~30-40g of cell pellet was lysed via a cell disruptor (volume 220mL) and the soluble fraction was loaded onto a 1mL Nickel column connected to an ÄKTA system (Figure 4.33). The bound C-terminally his-tagged protein was then eluted from the column. An SDS-PAGE gel was run and stained with Coomassie to visualise proteins in each of the fractions (Figure 4.34).

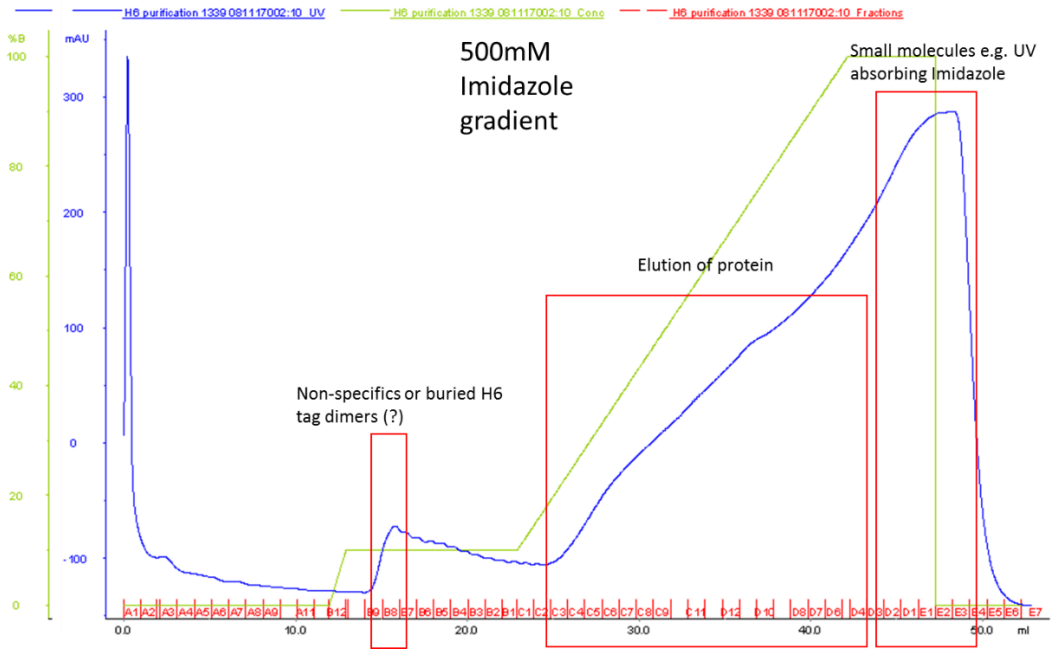


Figure 4.33: Chromatogram of the Rv1339 expressing MSMEG lysate being run on a 1mL nickel column.

The profile of the chromatogram was mainly due to the UV reacting Imidazole that was used. This explains why the profile of protein wavelength is very similar to the Imidazole gradient.

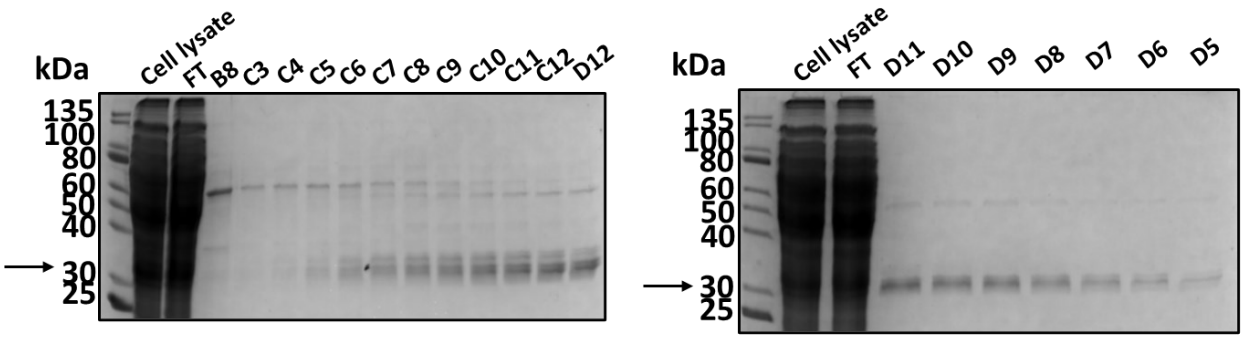


Figure 4.34: Coomassie stained SDS-PAGE gels visualising the proteins in each of the Nickel column fractions.

In this preparation, the protein eluted with a high homogeneity in later fractions. Fractions D11 to D7 were pooled, concentrated to 500 µl and run through an S200 size exclusion chromatography step. The SEC purification Chromatogram (Figure 4.35), SDS-PAGE and PDE activity assays of the fractions (Figure 4.36A and 4.36B) are shown overleaf.

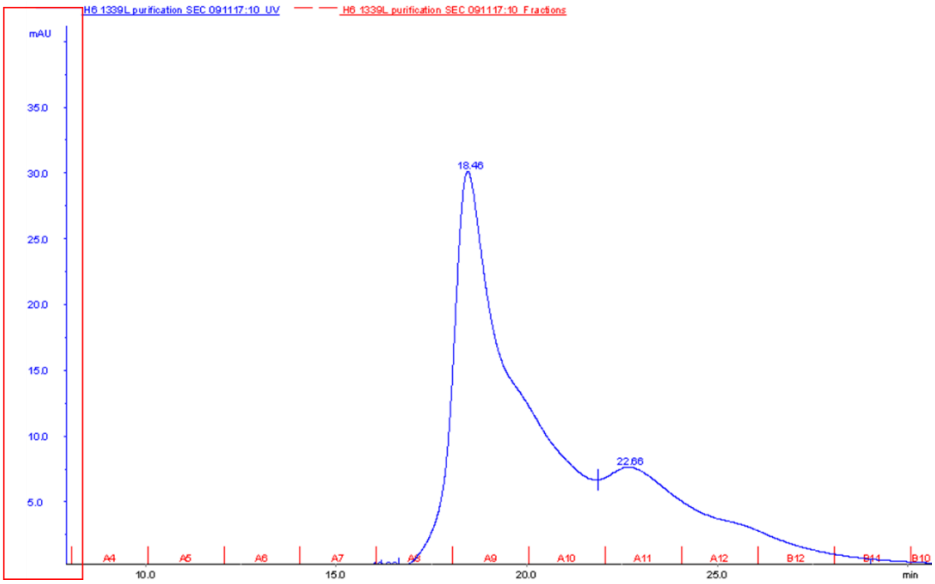


Figure 4.35: Chromatogram of the nickel purified lysate run on an S200 size exclusion chromatography column.

The levels of UV were very low (Figure 4.35, red box along the Y-axis), this likely indicates that there were low amounts of protein after the size exclusion. In order to visualise the likely proteins that eluted in fractions A8 to A11, Coomassie stained SDS-PAGE was used (Figure 4.36A). The gel corroborates the profile of the chromatogram (Figure 4.35) well with bands at the size of the monomer and potentially a dimer or GroEL chaperone. In addition, the fractions were screened for PDE activity – the product of the reaction (5' - AMP) being monitored by LC-MS (Figure 4.36).

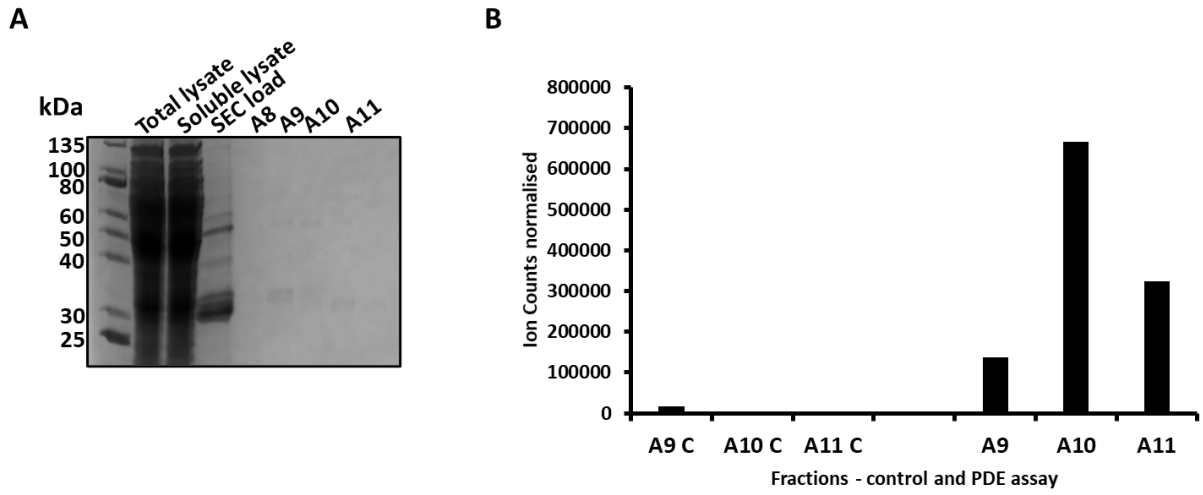


Figure 4.36: Coomassie stained SDS-PAGE gel visualising the proteins in the fractions eluted from the SEC column (left). Normalised ion counts of 5' - AMP detected in these fractions run in the PDE activity assay. Activity assay fractions labelled with "C" correspond to controls which were not provided with 3', 5' - cAMP in the PDE activity assay.

Although the preparation was pure (around 80%), the yield was far too low for crystallography and enzymology purposes (after concentrating to 200µl – 0.24 mg/mL). The PDE activity of Rv1339 was well conserved and intact protein mass spectrometry showed that the actual mass of protein bands observed corresponded to the theoretical mass of monomeric (30.21816 kDA) and potentially dimeric Rv1339 (60.43632 kDa). However, only the monomer reading was within the range of error for the apparatus. In contrast, peaks at the dimeric size were outside the acceptable range of error –and could potentially be a chaperone instead of dimeric Rv1339, as seen in the *E. coli* purifications (data not shown). The difficulty with this analysis was that an Rv1339 dimer would be at a comparable size to a variety of chaperones (e.g. mycobacterial GroEL – 56.726 kDa and DnaK –66.8309 kDa), that would likely be co-purified to some extent with any mycobacterial protein. Due to the insufficient yield however, the purification was repeated with 56g of bacterial cell pellet.

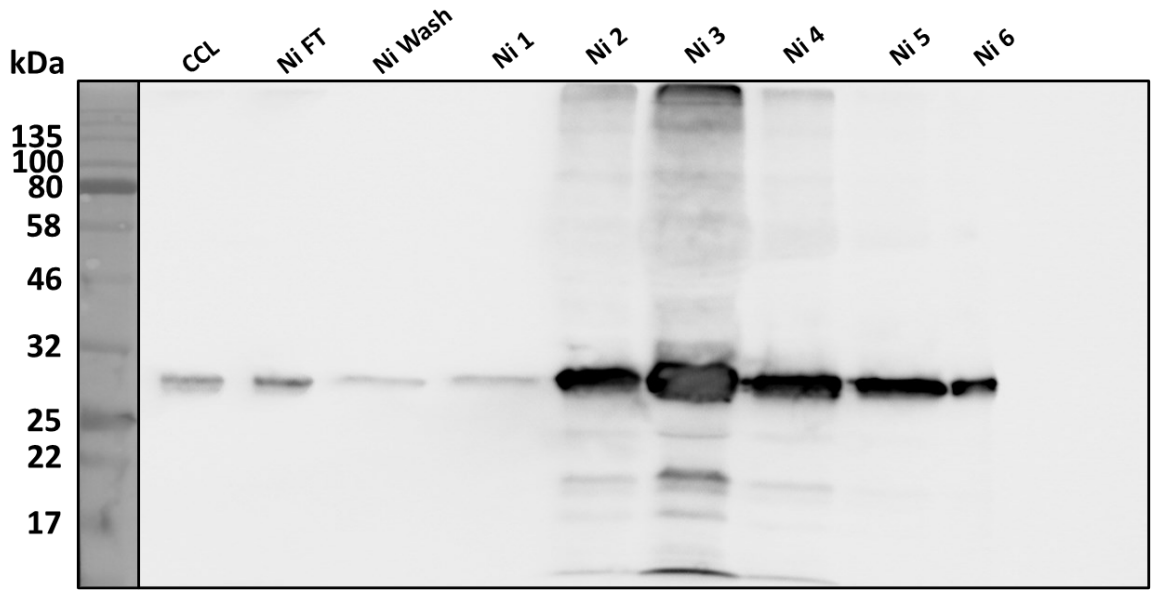


Figure 4.37: Western blot of Nickel affinity gravity column purification of pVV16::Rv1339L

With the increased starting mass of bacterial pellet (56g vs 30-40g previously) a 3 mL gravity flow nickel affinity column purification was performed. The results of this purification were analysed by Western blot analysis (Figure 4.37). This blot showed that fractions 4 and 5 displayed relatively high intensity of the band corresponding to the predicted size of Rv1339 (32.21816 kDa), but with fewer non-specific bands than fractions Ni 2 and Ni 3.

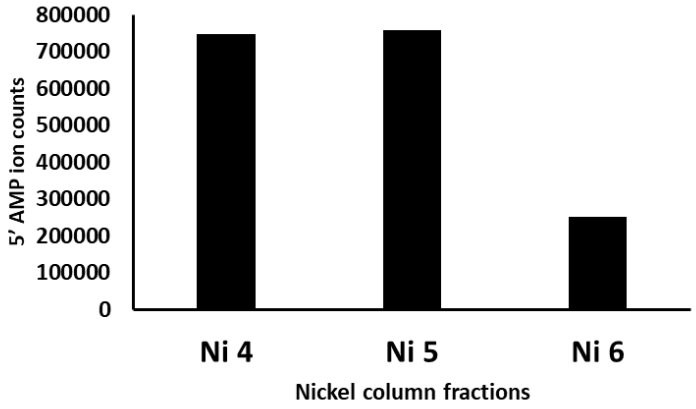


Figure 4.38: LC-MS quantification of the cAMP PDE activity assay for the Nickel affinity gravity flow column purified fractions.

Furthermore, after the Ni fractions were buffer exchanged, the cAMP PDE activity assay indicated that Ni fractions 4 and 5 displayed significant production of 5' AMP (Figure 4.38). This cAMP PDE activity was not seen in the heat denatured controls (data not shown).

So far, only SEC purification steps had yielded relatively pure preparations of protein and therefore a protocol that only involved nickel purification and SEC steps was considered.

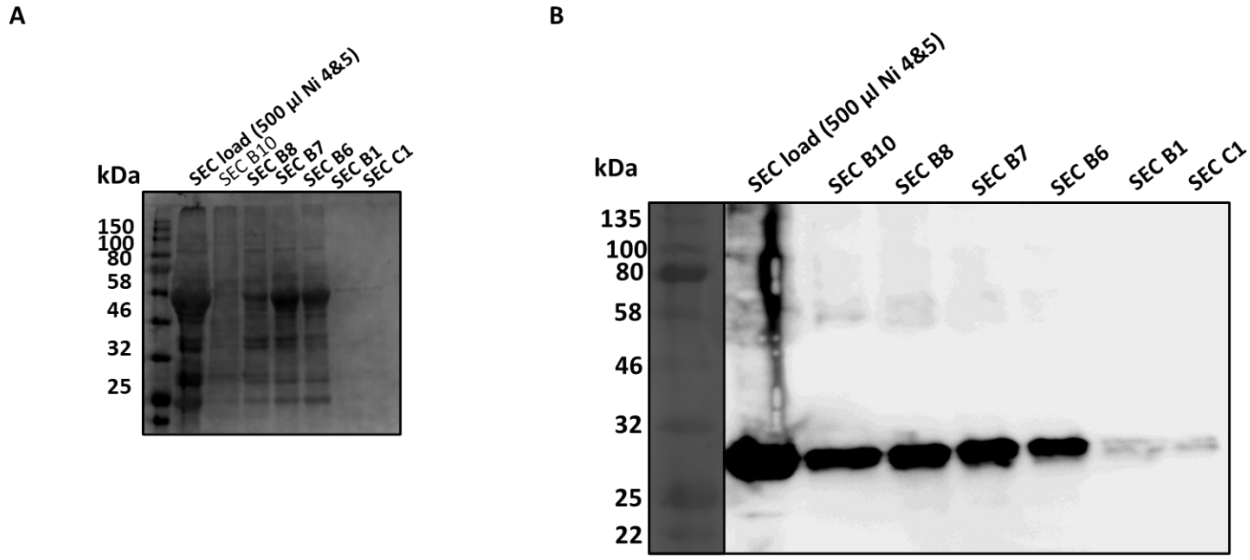


Figure 4.39: Coomassie stained SDS-PAGE (A) and Western blot analysis (B) of S200 SEC purified Nickel column fractions (Ni fraction 4 and Ni fraction 5).

Nickel column fractions 4 and 5 were therefore subsequently purified by 2 consecutive SEC chromatography steps. In the first step, an S200 SEC column was first used to purify fractions from a nickel affinity gravity flow column, with the intention of separating the proteins over a broad range of molecular weights (Figure 4.39A). Again, it appeared that as well as bands within the right size for Rv1339, there was also bands corresponding to the size of the chaperones (such as GroEL2). This hypothesis appeared more likely as western blot analysis confirmed the presence of a band at the predicted monomeric size of Rv1339, but only very faint bands around the size of an Rv1339 dimer (MW 60.43636 kDa) (Figure 4.39B). The presence of a dimer was considered unlikely as the pre-requisite boiling of SDS-PAGE samples in Laemmli (SDS) and β-mercaptoethanol containing buffer for 5 minutes before running should linearise the polypeptides and break down any quaternary interactions between protein subunits.

Subsequently, SEC S200 fractions B8, B7 and B6 were concentrated and loaded into an S75 SEC column, with the intention of hopefully coaxing apart any potential Rv1339-GroEL complex or to at least remove other protein contaminants.

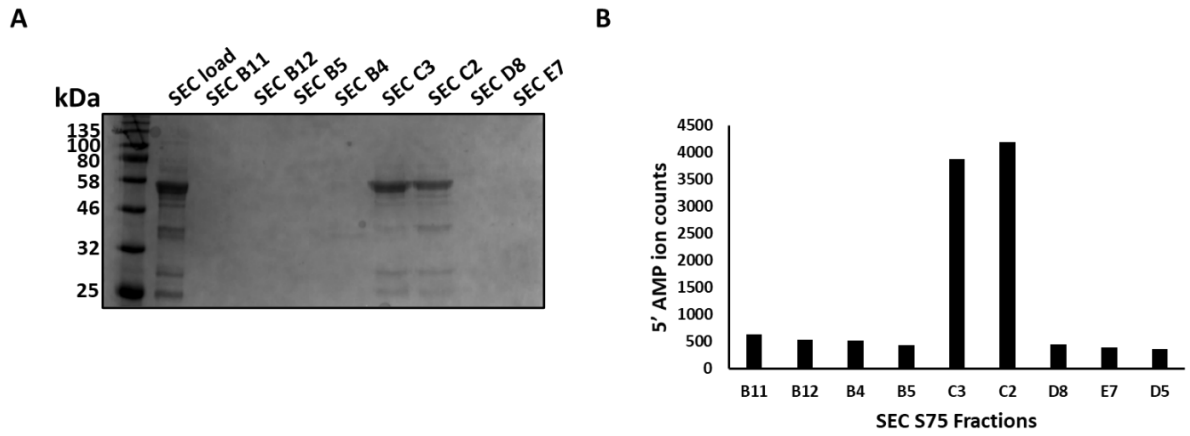


Figure 4.40: Coomassie stained SDS-PAGE (A) and cAMP PDE activity assay quantified with LC-MS (B), of the second, S75 SEC purification of pVV16::Rv1339.

After the second SEC purification (Figure 4.40A), there seemed to be fewer unknown bands – this suggests that the second SEC step did in fact increase homogeneity. However, proteomic mass-spectrometry analysis (performed by the proteomics facility at the University of St Andrews), showed that the bands in fractions C3 and C2, at around 60 kDa, were indeed GroEL and DnaK chaperones. The band expected to be Rv1339 was indeed Rv1339 (data not shown). The cAMP PDE activity assay (4.40B) showed activity above background levels in fractions C3 and C2, however, ion counts were significantly lower than had been seen in previous purification activity assays. While not directly compatible as the samples were run at different times, through different columns and with potentially different protein concentrations, the 5' AMP ion counts expected from a successful cAMP PDE activity assay are in the range of 10^4-10^6 ion counts. This suggested that there were low amounts of active protein left after the multiple purification steps. With purity far below the required 80%, low activity and contamination with non-specific proteins, this purification did not meet the required criteria for enzymology or crystallography.

4.5.1 Approaches to dissociate the potential chaperone-Rv1339 complex in *M. smegmatis* purifications

After several purification attempts with pVV16::Rv1339 proved insufficient to yield active and pure protein, and the proteomic data which showed the likely consistent co-elution of Rv1339 with GroEL, my efforts became focussed on disrupting this complex.

Co-elution with the GroEL1 chaperone is common in mycobacterial protein purifications – when a His-tag is used. This is due to a combination of the histidine rich C-terminus binding to a Nickel affinity column and eluting similarly to a His-tag, and also the potential reliance of certain mycobacterial proteins on chaperones for correct folding. This also means that within 1 GroEL1 super complex, there can be heterogenous populations of protein, thereby decreasing purity[388,389].

Firstly, to address this problem, a chromosomal GroEL1ΔC strain of *M. smegmatis* was gratefully received from the group of Stephane Canaan[390]. This strain features a mutated C-terminus for GroEL1, which removed the histidine rich region (7 out of 11 residues) that interacts with Nickel affinity columns[389]. With the reduced affinity of this chaperone for Nickel affinity columns, I hoped to disrupt the GroEL-Rv1339 complex. Expression of Rv1339 and Rv1339 D180A were confirmed in The *M. smegmatis* mc^{2}155GroEL1ΔC mutant by Western blot analysis (Figure 4.41).

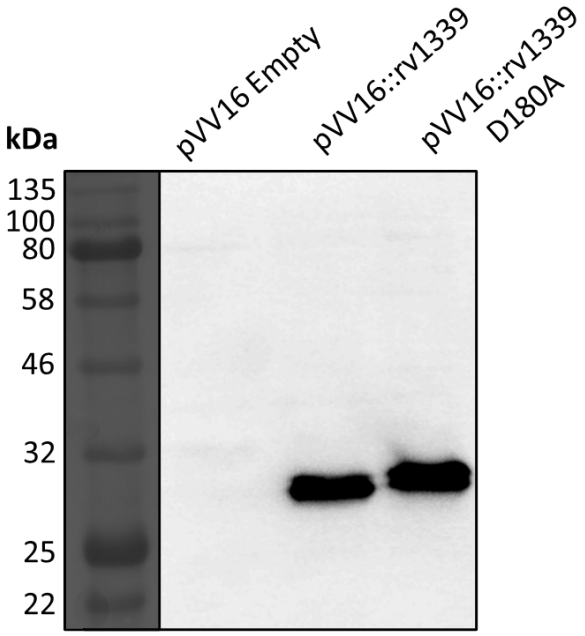


Figure 4.41: Western blot of Rv1339 and Rv1339 D180A expressed in GroEL1ΔC M. smegmatis bacteria.

GroEL forms a reversible complex with its targeted proteins. This interaction is dependent on ATP/ADP levels, chemistry and kinetics[377,391]. With this in mind, I attempted to alter the lysis buffer conditions, with the intention to discourage or weaken the interaction between Rv1339 and GroEL1ΔC. Initially,

in order to form a baseline for the purity of expression of Rv1339 in the GroEL1ΔC *M. smegmatis* mutant, I repeated the method of gravity flow Nickel affinity column, and SEC that earlier yielded relatively pure Rv1339, but with low protein abundance (Figure 4.36). This involved a gravity flow nickel affinity column purification and subsequent S200 SEC purification steps.

The results of this purification method in the GroEL1ΔC strain can be seen in Figure 4.42. Coomassie stained SDS-PAGE analysis (Figure 4.42A and 4.42C) showed drastically decreased intensity of bands matching the size of the chaperones GroEL and DnaK – this suggests that mutating the GroEL1 C-terminus does in fact reduce binding. Unfortunately, the intensity of the band at the size of Rv1339 also appeared to be of a lower intensity, and potentially with significant degradation (bands found at around 25 kDa). Western blot analysis (Figure 4.42B and 4.42D) displayed a band for Rv1339 in the clear soluble lysate, diminished in the Ni column flow through and presence of a Rv1339 band in the SEC column load, as well as fractions A4 to C8. This suggests that Rv1339 was eluted over many fractions, decreasing purity.

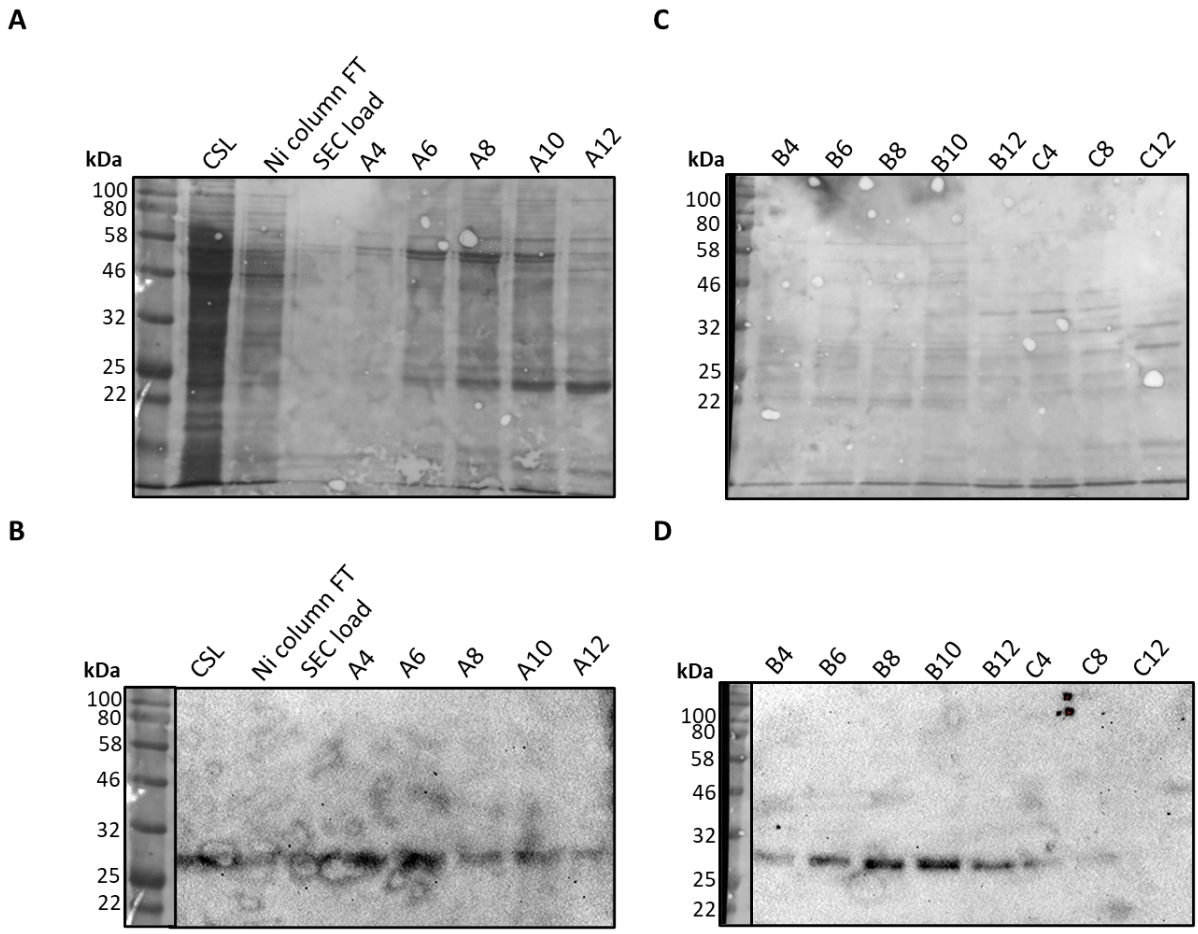


Figure 4.42: A and C: Coomassie stained SDS-PAGE analysis of SEC purified Rv1339 expressing GroEL1ΔC M. smegmatis. B and D: Western blot analysis of SEC purified expressing GroEL1ΔC M. smegmatis.

From this purification (Figure 4.42) the proportion of Rv1339 in SEC fractions as compared to bands at the size of chaperones, appeared to be increased. However, there appeared to be degradation of Rv1339, potentially from the C-terminal portion, as this is where the His-6 tag was contained and there appeared to be only 1 full length band on the western blot. Although promising, increased purity was required.

Altering the pH of the lysis buffer may affect the interaction of GroEL and Rv1339. Chaperones were initially thought to bind to their target proteins by hydrophobic interactions[392]. Hydrophobic interactions with protein residues are affected by the conformation and structure of the protein e.g. the number of hydrophobic residues exposed on the protein surface. Alteration of pH and salt concentration can change the conformation of a protein[393,394], thereby potentially affecting hydrophobic interactions.

Likewise, low salt levels (50mM) were considered preferable as a previous study of chaperone activity at various salt concentrations (e.g. 171mM), favoured chaperone-protein interactions via increased mediation of electrostatic interactions between chaperone and protein[395].

Additionally, reports in the literature indicated that chaperones were found to bind to denatured proteins[396,397]. By combining denatured *E. coli* lysate (from pBXC3H::Rv1339 expressing MC1061 cells) with the Rv1339 expressing GroEL1ΔC *M. smegmatis* mc^2155 lysate (to a final concentration of 10% based on protein concentration) it was hoped that GroEL or other chaperones would bind to the denatured proteins instead of Rv1339.

With these criteria in mind, I chose 4 lysis buffer conditions to test. A fresh 5L culture of Rv1339 expressing GroEL1ΔC *M. smegmatis* mc^2155 was cultured, and the bacterial cell pellet extracted by centrifugation. 12.5g of bacterial pellet was used for each condition to be tested. The pellets were lysed via cell disruption in the different lysis buffers, and combined with a proportion equal to 10% of the denatured *E. coli* lysate – by protein concentration.

The first 2 conditions attempted utilised a 20mM solution of Tris buffer at pH 9. This was selected because the buffer for the cAMP PDE activity assay was pH 9.0, and Rv1339 had previously shown robust activity as this pH. Additionally, 50mM NaCl was added to both lysis buffers. One of the lysis buffers was further supplemented with 2mM of the reducing agent Dithiothreitol (DTT). This reducing agent was intended to potentially aid formation of disulphide bonds, and enable proper folding of Rv1339. The results of these purifications can be seen in Figure 4.43. When lysed in the buffer containing 20mM tris pH 9.0 and 50mM NaCl (Figure 4.43A), it appeared that nickel affinity purification under these conditions yielded bands at both the size of predicted GroEL and Rv1339, particularly in fractions A2 and A3. However, the bands corresponding to the size of Rv1339 appeared to have a low relative abundance to the heterogeneity of bands within these fractions.

When DTT was added to the lysis buffer, the colour of the nickel resin darkened over time, likely reflecting the change in charge state caused by DTT. The concentration of DTT used was still below the maximum recommended limit of the nickel resin. Regardless of the decolouration, the Coomassie stained SDS-PAGE (Figure 4.43B) still indicated that proteins were binding to the column. The discernible effect of DTT addition appeared to be that the highest intensity bands appeared to be in NTA fraction 3, whereas in buffer without DTT, proteins showed a wider distribution over fractions A2 and A3. This may indicate that DTT still had an effect on the interactions of proteins with the Ni ion containing resin – despite being within the recommended concentration range from the manufacturer. cAMP PDE activity was confirmed via the activity assay in fractions A2 and A3 in Figure 4.43A but no activity was observed in fractions from the DTT supplemented purification (4.43B) (Data not shown).

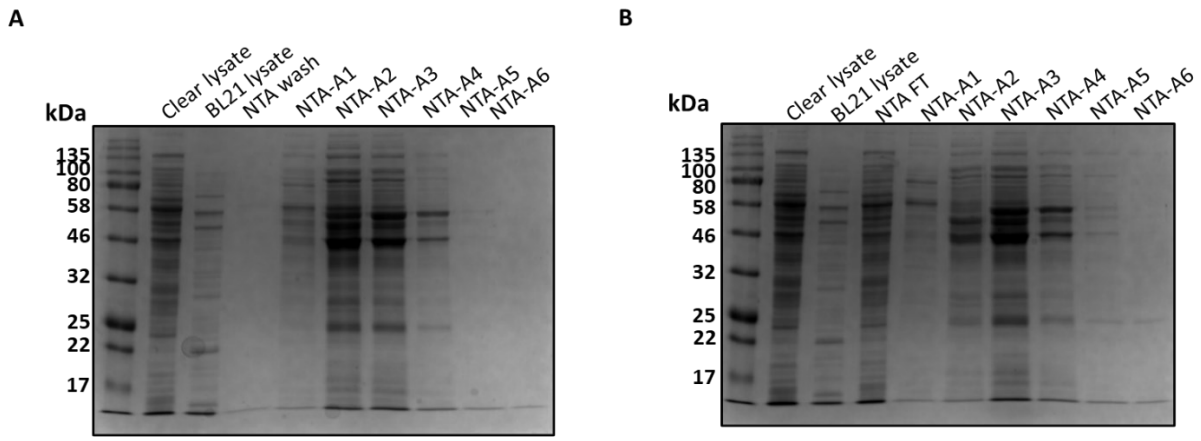


Figure 4.43: A: Nickel affinity purification of Rv1339 GroEL1ΔC M. smegmatis lysed with 20mM Tris pH 9.0 and 50mM NaCl (A) or 20mM Tris pH 9.0, 50mM NaCl and supplemented 2mM DTT (B).

The next 2 conditions investigated were the standard 20mM pH 7.5 and 50mM NaCl that proved successful in initial pVV16::Rv1339 NTA and SEC S200 column purification of Rv1339 (Figure 4.44A) or 20mM MES buffer (pH 5.0) (Figure 4.44B). This pH was chosen as it was predicted to be lower than most bacteria would normally encounter, however, *M. tuberculosis* has been shown to be exposed to environmental pH as low as 5 during host infection[220–222] . Likewise, the isoelectric point (pI) of Rv1339 is 5.8, the pH at which the protein has a net charge of 0 – a key factor in determining the strength of electrostatic interactions[398,399]. Together, this indicated that a lysis buffer at pH 5.0 may be appropriate[399]. The results of these purification can be found in Figures 4.44A (Tris pH 7.5) or 44B (MES pH 5.0).

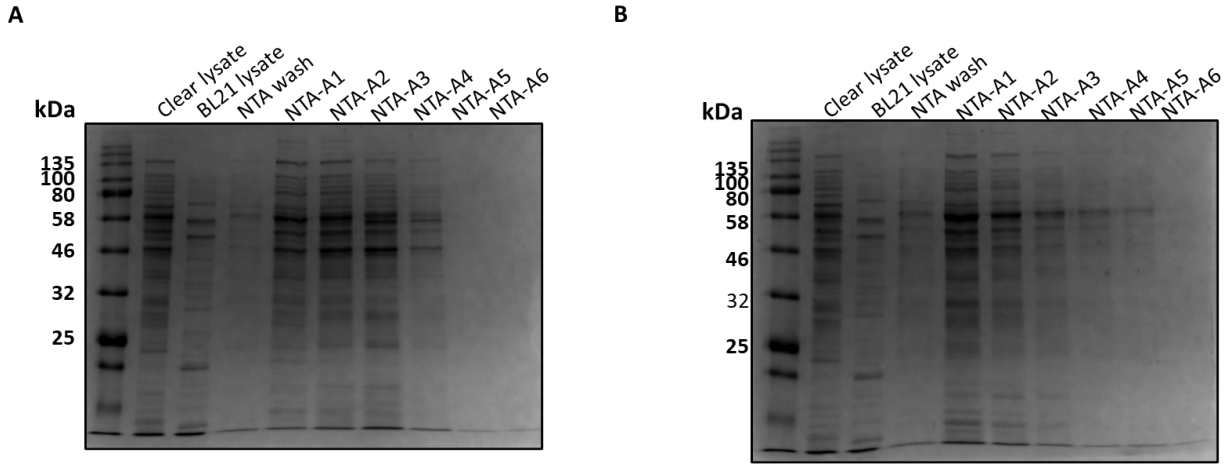


Figure 4.44: A: Nickel affinity purification of Rv1339 GroEL1ΔC M. smegmatis lysed with 20mM Tris pH 7.5 and 50mM NaCl (A) or 20mM MES pH 5.0, 50mM NaCl.

In both purifications, proteins were still able to bind to the NTA column, though in pH 5.0 condition (Figure 4.44B), proteins were concentrated in the initial 3 Nickel column fractions and it appeared as though overall they were less abundant than the Tris pH 7.5 condition (Figure 4.44A). Purification at pH 5.0 appeared to yield less heterogeneity in fraction composition overall, but still did not yield sufficiently pure preparations of Rv1339. Likewise, the pH 7.5 condition did not show any increased level of purity over previous purifications.

Taken together, this suggests the interaction between Rv1339 and the chaperones is strong. The lack of homogeneity in the later mycobacterial purifications also suggests that purification in this system would consequently not be straight forward. The impact of Rv1339 on mycobacterial physiology may also play a role in prevented a high purity e.g. lowered cAMP may compromise key processes that in turn lead to stress. If the problem with Rv1339 purification was due to its cAMP PDE activity, cAMP signalling is universal and broadly conserved across many different taxa[65] and it would be difficult to find any organism suited to protein purification that does not rely on this signalling pathway in some way. Unfortunately, with the time and resources left of my PhD, it was not possible to find other feasible *in-vivo* purification systems. A potential solution was to attempt cell-free expression.

4.6 Cell free expression

With the lack of success in obtaining sufficiently pure Rv1339 using a multitude of approaches and organisms involving *in-vivo* expression, I considered cell free expression (CFE).

The approach is mediated by continuous-exchange cell-free protein synthesis systems, that consist of *in-vitro* reaction buffers containing transcription and translational machinery e.g. RNA polymerase, ribosomal components and proprietary reaction buffers[400]. These components can extracted from *E. coli*[401] or even insect cells[402]. By providing this machinery with DNA nucleotide bases (dNTPs), amino acids and important co-factors, the insert of an expression plasmid can be transcribed and translated in the well of a microplate, under optimal conditions.

Our lab purchased a CFE kit from Biotech Rabbit who used a secretive and proprietary system similar to previous studies[400]. The draw-back of CFE is the low-yield and issues with scalability[400]. However, Biotech Rabbit products appeared to accommodate scaling up to preparative quantities.

Because of the unknown composition of the lysate, and the low quantities of protein, I selected the Strep-II tag system for purification. This system utilises an 8 amino acid long synthetic peptide tag (Trp-Ser-His-Pro-Gln-Phe-Glu-Lys) that binds with high affinity to streptactin – a bioengineered version of streptavidin. Streptavidin was initially isolated from *Streptomyces avidinii* and was noteworthy due to its very high affinity to biotin (dissociation constant of 10^{-15}M) [403]. The strep-II tag can therefore bind a streptacin matrix with high affinity and be eluted under mild buffer conditions, by competition with biotin or its derivatives[404]. It is purported to offer a robust one-step purification protocol with high purity[404].

The CFE system used a T7 promotor system so I utilised an IPTG inducible mycobacterial shuttle vector (pYUB) that fit the CFE reaction criteria exactly. This vector natively possessed a C-terminal His-6 tag, but I replaced this tag with a C-terminal strep-II tag using site directed mutagenesis[342].

Another drawback of the explorative small-scale CFE reactions was that they were a maximum of 50µl reaction volume. This meant purification would require Strep-II binding spin columns. It also meant that there was limited material for stained SDS-PAGE, western blot and activity assay analysis – therefore I compromised and ran only Western blot and activity assays. I performed the cell free expression assay and passed the lysate through Strep-II tag binding spin columns (Figure 4.45).

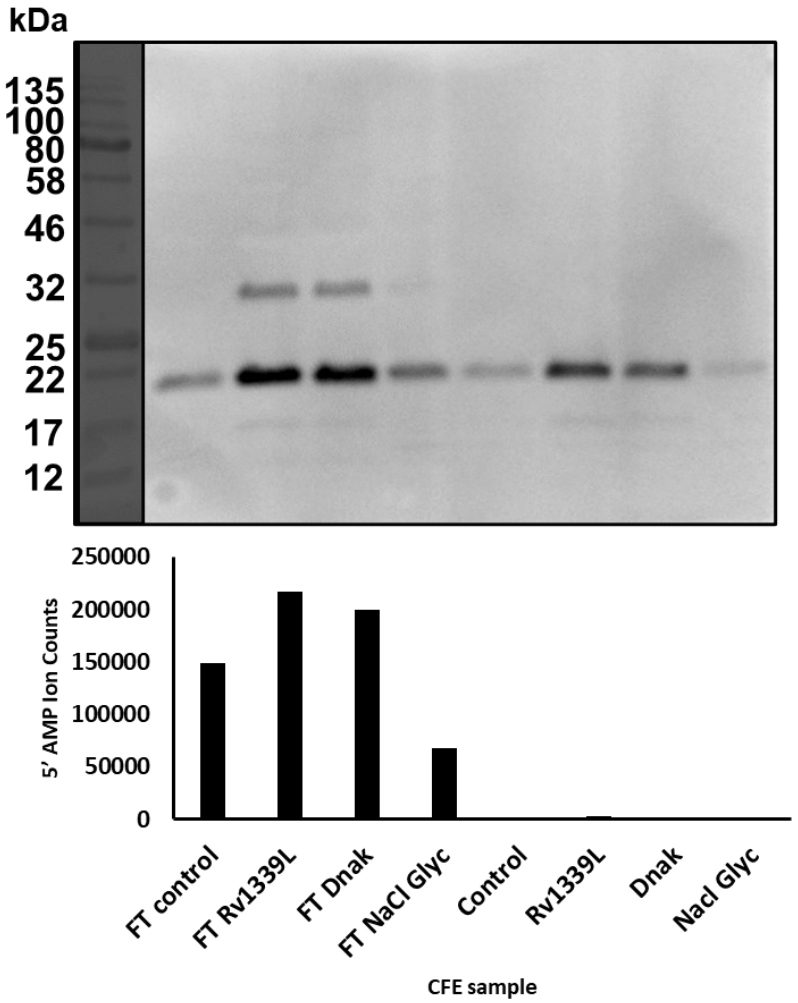


Figure 4.45: Western blot analysis of the CFE reactions. Control represents a non-strep-II tagged urokinase. Rv1339 corresponded to pYUB::Rv1339L Strep-II. DnaK denotes supplementation of pYUB::Rv1339L Strep-II with purified chaperone DnaK. NaCl Glyc denoted pYUB::Rv1339L Strep-II supplemented with NaCl and Glycerol.

Samples denoted with FT indicate that they are the proteins that did not bind the column. From western blot analysis combined with LC-MS quantification of the cAMP PDE activity assay, it appeared that Rv1339 was produced and that this exhibited 31% higher activity than the control reaction (a non-strep-II tagged urokinase). Samples denoted with "DnaK" and NaCl indicate CFE reactions with the addition of purified mycobacterial chaperone DnaK or addition of 250mM of salt and 10% glycerol – 2 suggestions from the CFE kits troubleshooting guide. However, the column could not bind and purify Rv1339 (MW: 30.698.65 kDa), which only appeared to be expressed well in the Rv1339 and DnaK conditions (for purification of this chaperone, see Annex 4) supplemented reactions. This may be because of unspecified high biotin levels in the *E. coli* CFE mixture, out-competing the strep-II tag. Biotin has been shown to bind with high affinity to streptavidin (a capture matrix for the Strep-II tag)[405]. Indeed, it appeared that within the CFE reaction lysates, there was a biotynolated protein that could bind and elute from the column (22 kDa band).

The size of the kit did not allow this experiment to be repeated, but if time and resources allowed, it would have been prudent to repeat this purification, but after adding avidin to the samples – which has been shown to minimise the interference of biotin[406].

4.7 Investigation of Rv1339 substrates specificity from expression of an Rv1339 and an Rv1339 D180A pVV16 construct in *M. smegmatis mc²155* clear soluble lysate

Purification of Rv1339 to the 80% target required for accurate purified protein enzyme kinetic studies was either not reached, or not found to be scabale from all attempted purification attempts. In order to still characterise this enzyme to some extent, *M. smegmatis* mc^{2}155 lysate expressing the Empty Vector Control (pVV16), Rv1339 or Rv1339 D180A was obtained via sonication and combined with the PDE acitivty assay buffer, and various substrates over a 3 hour time period at 37˚C (Figure 4.46A-E). The temperature, time-point and substarte concentration (0.01mM) were optimised intensively and use of the Agilent Bravo Liquid Handling System was also optimised. This invovlved optimising a reaction volume (50 µL), a manual pipetting method and an automated dilution protocol for loading samples into the LC-MS.

The resulting product or substrate metabolites of the reaction quantified by LC-MS analysis after 3-hours and plotted as a percentage of Control levels. The Control used for this experiment was the lysate of the Empty Vector expressing strain, heat denatured. This was an attempt to set a baseline for metabolite levels with the biological matrix effect, and without degradation by active proteins. A similar approach has previously been used to characterise an enzyme that proved difficult to purify in *Staphlaccocus aureus*[118].

The substrates investigated were: 3', 5' – cAMP – the conserved second messenger (4.46B), 2', 3' – cAMP – a DNA degradation product (4.46C), 3', 5' - cGMP – a mammalian second messenger that can be broken down by Class-II PDE eznymes (4.46D) and 5' AMP – the degradation product of cAMP (4.46E). The biological relevance of cGMP to *M. smegmatis* mc^{2}155 currently appears to be neglidible – but this will be reviewed briefly in the discussion. 5' AMP was inteded as an additional control, to ensure that Rv1339 was not directly altering 5' AMP levels.

A

Compound Name	Chemical Formula	Mass	Theoretical m/z (pos)	Experimental m/z (pos)	Δppm Pos	Retention Time (min)	CAS	METLIN	KEGG
2',3'-Cyclic adenosine monophosphate	C10H12N5O6P	329.0525	330.0598	330.06	-0.6059508	3	634-01-5	3459	C02353
Adenosine5'-monophosphate	C10H14N5O7P	347.0631	348.0704	348.0708	-1.149192807	5.5	61-19-8	34478	C00020
2',3'-Cyclic GMP	C10H12N5O7P	345.0474	346.0547	346.0551	-1.155886627	1.75	634-02-6	3487	C06194

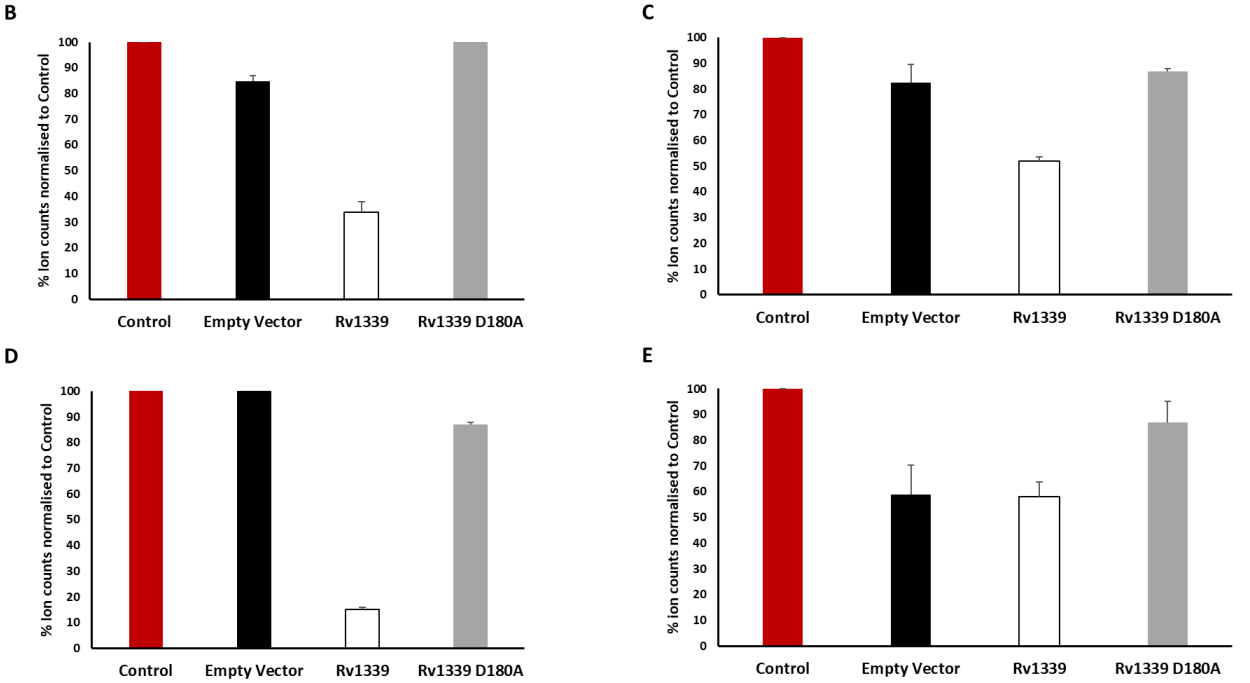


Figure 4.46: In-vitro activity assays of (pVV16) Empty Vector Control, Rv1339 and Rv1339 D180 when incubated with cAMP PDE activity assay buffer and different substrates, for 3h. A: Table of substrates used in this assay. 2', 3' - cAMP and 3', 5' - cAMP are identical except for bond position. B: 3', 5' - cAMP. C: 2', 3' - cAMP. D: 3', 5' - cGMP. E: 5' AMP.

When 3', 5' - cAMP (cAMP) was used as substrate in this assay (4.46B), there was a 66% decrease in cAMP levels at the 3 hour timepoint with lysate expressing Rv1339. Such a drastic decrease was not seen in the Empty Vector or Rv1339 D180A expressing lysate. However, there was 15% decrease observed in the Empty Vector expressing lysate. This likely corresponds to degradation or potentially the activity of the Rv1339 *M. smegmatis* mc^{2}155 homolgue, MS_4902. If this was the case however, it would not explain the relative lack of cAMP degradation observed in the Rv1339 D180A expressing lysate. As a result, these differences may just reflect variation in the levels of cAMP measured in each sample. When taken together, it suggests that Rv1339 does infact display cAMP PDE activity *in-vitro*.

2', 3' – cAMP was also investigated (Figure 4.46C), as the only currently annotated 3', 5' - cAMP PDE enzyme in mycobacteria, Rv0805, displayed 150-fold increased activity towards this nucleotide[281,308]. In fact, this was a criteria used to suggest that there must be another 3', 5' - cAMP PDE. When the lysates were assayed with this substrate, indeed Rv1339 expressing lysate appeared to show some PDE activity (mediating a 48% decrease). A decrease of this scale was not seen in the Empty Vector or Rv1339 D180A expressing lysates, thereby suggesting that this decrease was the result of Rv1339

expression. While it is not unusal for a cAMP PDE enzyme to display promiscuous activity on both 3', 5' - cAMP and 2', 3' - cAMP, as discussed earlier in the bioinofrmatics section[278], it appeared that Rv1339 displayed greater activity towards 3', 5' - cAMP (mediating a 66% vs 48% decrease).

As a probable atypical class-II PDE enzyme (discussed in the bioinformatic section), Rv1339 would be expected to display PDE activty to 3', 5' - cGMP [278]. In order to test this, and provide more evidence to support the identification of Rv1339 as an atypical class-II PDE enzyme, this nucleotide was used in the assay. In fact, Rv1339 expressing lysate displayed a 85% decrease in 3', 5' - cGMP levels. A decrease of this scale was not observed in the Empty Vector or Rv1339 D180A expressing lysates. The problem with this particular data was that due to likely differences in ionisation efficiency of this particular metabolite, ion counts were very low. In fact, the Denatured Empty Vector Control samples displayed 500 ion counts and the Rv1339 expressing lysate displayed 69 ion counts (below the background ion counts threshold). This is contrasted with 8000 and 2000 ion counts found for the 3', 5' - cAMP containing samples. While the cGMP samples may begin at just above background ion count levels, they decrease below the reliable threshold. To improve this experiment, increasing the quanitity of substrate used to boost ion counts would help. In light of the limitations, it is not possible to draw many conclusions from the reactions containing this substrate. It is however likely that Rv1339 possess some PDE activity towards cGMP, as expected[278].

Finally, 5' AMP was assayed as a substrate (Figure 4.46E). In the control reactions, the starting point was 1000 ion counts and this decreased to a minimum of 500 ion counts in Empty Vector, Rv1339 and Rv1339 D180A. The same scale of decrease (around 45%) was seen in Empty Vector and Rv1339 expressing lysates, with D180A displaying a minor decrease (15%). While this may indicate that within the *M. smegmatis* mc^2155 lysate that there are enzymes that break down 5' AMP (into its degradation products e.g. ionosine-monophosphate), the data does not suggest this is mediated by Rv1339. The variation between replicates was also high for this substrate, with an irregular, asymetrical chromatogram that proved difficult to integrate and extract ion counts from. This is reflected in the larger error bars than are observed in the other reactions.

The limitation of this lysate enzyme kinetics approach is that there may be other enzymes contained in the lysate that effect the results. In addition, the singular time point of 3 hours did not allow proper measurements of the kinetics of these substrate level decreases. A far greater sampling would be required to investigate the rate at which Rv1339 cleaves each substrate. What these assays do show, is that Rv1339 possess cAMP PDE activity, that is not seen in the Rv1339 D180A point mutant. This provides some evidence to support the hypothesis that Rv1339 is an atypical-class II PDE enzyme (with characteristic activity towards both cAMP and cGMP). However, phenotypic characterisation would prove to be a better indicator of this (as seen in Chapter 3). Finally, it might also suggest that the preference of Rv1339 is towards the second messenger 3', 5' - cAMP as opposed to the DNA

degradation product 2', 3' - cAMP. This would be in direct contrast to Rv0805 – and therefore support the hypothesis that Rv1339 is a cAMP PDE enzyme in mycobacteria. However it must be noted that the assay did not measure affinity (K_d) or other kinetic parameters that would be required to confirm such a preference.

4.8 Discussion

Expression and purification of mycobacterial proteins in *E. coli* can be difficult e.g. differential codon usage and physiology[382]. A review of this topic suggests that more than 60% of mycobacterial proteins do not show soluble expression in *E. coli*[388]. This is postulated to be a result of the differences in physiology between the 2 organisms. For example, mycobacteria such as *M. leprae* encode 2 GroEL chaperones[407] and *M. smegmatis* encodes 3[408] – in contrast to just 1 GroEL chaperone in *E. coli*. This is in addition to other chaperones such as GroES and DnaK[409]. Additionally, essential roles for some of these chaperones in native protein folding have been identified in mycobacteria[410]. The chaperones GroEL and DnaK have also been found to be essential in *M. tuberculosis* H37Rv via saturated transposon mutagenesis studies[320]. Taken together, this emphasises the importance of chaperones for successful mycobacterial protein folding. These systems in *E. coli* may not be adapted or comprehensive enough to enable correct folding of mycobacterial proteins. Likewise, the doubling times of these organisms is very different e.g. 28-56 hours for *M. tuberculosis*[411] vs. 50-100 minutes for *E. coli*[412].With the propensity of chaperones like GroEL1 to co-elute with His-tagged proteins and even out-compete them for binding to Nickel affinity columns[389], this adds another layer of difficulty.

From my attempts to purify Rv1339, it is clear that it was not an easy protein to express. This may be due to the presence of 8 cysteine residues, that can form disulphide bridges. It may also be due to the bulky signal sequence attached to the N-terminus or even because of the effects on physiology of overexpressing an enzyme capable of reducing cAMP levels. As reviewed in the Introduction chapter, cAMP is broadly conserved and essential under a wide range of circumstances. Perhaps decreasing cAMP levels leads to up-regulation of the stress response (as seen with stress-related proteins found in the proteomic mass spectrometry data derived from pBXC3H::Rv1339 purifications).

Another study involving purification of murine HSP-70 in *E. coli* identified co-purification with glutamine-fructose-6-phosphate aminotransferase (GFAT)[376]. This was 1 of the key contaminating enzymes I observed during pBXC3H::Rv1339 purification in *E. coli* (Figure 4.21). The aforementioned study was able to subsequently prevent contamination with GFAT by altering the media conditions and metabolism of the bacteria. For example, GFAT catalyses the conversion of N-acetylglucosamine to D-glucosamine-6-phosphate and is part of the hexosamine metabolic pathway[376]. By supplementing the media with D-glucosamine-6-phosphate, the levels of GFAT contaminant were markedly reduced. Furthermore, this had the effect of boosting overexpressed murine HSP70 production in *E. coli* by 5-fold – and therefore allowed efficient purification[376]. While it is likely that to some extent this may have been an effect on the expression of this particular protein, the idea of modulating bacterial metabolism to improve protein purification is very interesting. With more time, it would have been interesting to attempt purification of pBXC3H::Rv1339 with media supplemented with D-glucosamine-6-phosphate

to see if this approach could universally be applied to removing GFAT contamination. Even if this succeeded in reducing contamination of GFAT in my purification, it would still leave the significant contaminant, glycerol dehydrogenase and the other unknown bands.

Given enough time and resources, it may also have been possible to find another appropriate carbon source, as Terrific broth utilised 0.2% glycerol as carbon source, this may have led to increased levels of glycerol dehydrogenase contaminants. I believe manipulating the bacterial metabolome would be an exciting avenue to investigate, to improve protein purification but was not feasible in the timeframe of my PhD. By combining proteomics to identify contaminating proteins, and metabolome pathway analysis, with supplementation of the media with relevant metabolites, it may be possible to modulate bacterial metabolism to improve purity or protein yield in this way. Likewise, perhaps genetic knockouts of key metabolic enzymes may offer ways to decrease contamination – this was how the Walker strains of E. coli (C41, C43 and PlysS) were originally created[363]. From the aforementioned example, this may not be possible as GFAT deficient strains of *E. coli* have been generated but display significant growth defects when grown in non-glucosamine supplemented media[413].

With the relative ease of gene knockout and CRISPR/Cas 9 multigene knockout approaches in *E. coli*[414,415], high-throughput metabolomic analysis[416] and fluorescence mediated quantification of protein expression[370] it may be possible to generate a high-throughput screening system of different genetic mutants to find a strain of *E. coli* most suited to purification of a particular protein. This could involve microplate-based culture systems linked to microscopy quantification of GFP levels. Such fluorescence microscopy approaches are already being used to screen drug efficacy in high-throughput screens[417]. By increasing the sensitivity of these detection systems, it is conceivable that protein overexpression and even localisation[370,418] could be effectively quantified in a microplate format. A literature search reveals that the tools to mediate such analyses already exist with electron microscopy and different fluorophore approaches (e.g. with binding partners that enhance fluorescence signal)[419]. The challenge would be to optimise these tools in bacteria, and in small volume microplates.

The homologue of Rv1339 in *Corynebacterium glutamicum* CpdA[321] was purified to supposedly sufficient purity for enzyme kinetic studies. In kinetic studies, it was found to display significantly higher activity than *E. coli* CpdA and Rv0805 (33.63 µmol vs. 2 µmol vs 30 nmol cAMP converted/sec per mg protein, respectively)[194,279,321]. Rv1339 and *Corynebacterium glutamicum* CpdA share a 45.71% identity at the protein level and CpdA does not contain any signal peptide sequences. This may be a key difference in altering the solubility of the protein. Unfortunately, the authors did not crystallise CpdA and also do not discuss the final purity achieved for their enzyme kinetics. Combined with only 45.71 % identity, it is difficult to draw any conclusions.

A potential homologue of Rv1339 in *Bacillus anthracis* (1ZKP) that was used to identify the active site in Rv1339 was crystallised (1.5Å resolution) by a high-throughput crystallography consortium. This

enzyme displays a 29% identity with Rv1339, however, correspondence with the now retired project leader for the purification and crystallisation revealed that without the N-terminal His-tag and unspecified linker, 1ZKP was insoluble. He also noted great difficulty in purifying and crystallising it. Taken together, this suggests that Rv1339 may be difficult to purify in any biological system. The success of CpdA purification in Corynebacterium[321] suggests that this may be an interesting system to try and express Rv1339. However, due to the limited time and resources left of my PhD, this approach was deemed unfeasible. As previously mentioned, without in-depth biochemical characterisation, it is not possible to predict the true substrate specificity of a metallo-β-lactamase protein from bioinformatics alone – due to the broad nature of the family and their functional heterogenity[328].

Rv1339 appears to possess activity to both 3', 5' – cAMP and 2', 3' – cAMP nucleotides, as well as 3', 5' –cGMP. A key limitation of existing PDE enzyme studies, such as CpdA of *C. glutamicum*, is that the authors of these studies tend to assay only the non-physiologically relevant substrate Bis(para-nitrophenoyl)-phosphate[321,328,338] and/or a highly limited subset of potential substrates. In the case of the CpdA from *C. glutamicum* study, only 3', 5' – cAMP is assayed[321] – this drastically limits the potential scope and comparability of the findings of these studies. In my approach, I attempted to compensate for this by assaying both 3', 5' and 2' 3', - nucleotides, as well as an alternate cNMP – cGMP. The significance of cGMP to mycobacteria is questionable, as although an isolated report from 1984 described the presence of cGMP in *M. smegmatis* mc^2155[420], an extensive literature search did not appear to reveal any subsequent confirmation of these findings in mycobacteria.

Regardless of the difficulties encountered during attempts to purify Rv1339, I was able to express Rv1339 successfully in mycobacteria and quantify some level of substrate specificity[118]. Additionally, I was able to show that Rv1339 likely displays cAMP PDE activity *in-vitro* and that mutagenesis of a residue (D180A) in the predicted metal binding site seems to ablate this activity.

Bioinformatic and evolutionary analysis of Rv1339 showed that it was conserved across actinobacteria and that it represents an expansion of class-II PDE enzymes to include both the current bona fide class-II PDEs and a novel classification of atypical class-II PDE. The class-II PDE family classification motif requires significant overhauling, and I recommend the new defining motif for class-II PDE to be [T/S]HXHXDH. All characterised proteins from this family possess the ability to hydrolyse a phosphodiester bond, but the key difference in function appears to be the substrate specificity

With these findings, I would proceed to focus on characterising Rv1339 phenotypically and using it as a tool to decrease cAMP levels in mycobacteria.

Chapter 5 - Phenotypic characterisation of Rv1339 expression and decreased cAMP levels in *M. smegmatis* mc^{2}155 and *M. tuberculosis* H37Rv

5.0 Phenotypic characterisation of Rv1339 expression and decreased cAMP levels in *M. smegmatis* mc^{2}155 and *M. tuberculosis* H37Rv

As previously reviewed in the introduction, cAMP signalling is broadly conserved across many different taxa of life[65,106,324]. In pathogenic bacteria, cAMP signalling has been found to regulate a diverse range of processes including carbon source utlisation[360], biofilm formation[188], gene expression[165,210], metabolism[181,251] and virulence[78,145,164,177,190].

In mycobacteria, cAMP signalling has shown to mediate response to host environmental conditions such as the NaCl concentration found in the phagosome[47] – which leads to a significant 6-fold increase of intrabacterial cAMP levels[70]. Additionally, it has been shown that a cAMP binding transcription factor regulates bacterial virulence during infection[177]. *M. tuberculosis* bacilli even use cAMP to intoxicate host cells, preventing phagosome acidification and allowing evasion of host immune killing in the macrophage[78,79]. A broad range of mycobacterial processes are regulated by cAMP. For example, *M. tuberculosis* H37Rv bacteria can enter a dormant persister state, with significantly slowed growth, metabolism[245] and rewired bioenergetics to maintain membrane potential and energy production under inhospitable infection conditions (such as hypoxia[307] or redox-stress[421]). Many of these changes can be mediated by induction of the ~40 gene DosR regulon[244] – which has been shown to be regulated by a cAMP binding transcription factor, Cmr[241,243]. Likewise, cAMP binding transcription factors have been shown to regulate expression of key components of bioenergetics that provide mycobacteria with the flexibility to maintain essential ATP production while in the dormant state e.g. Succinate dehydrogenase, cytochromes and NADH dehydrogenases[211,237]. Furthermore, components of cAMP signalling have been shown to regulate metabolism of cholesterol[308] (a key carbon source during infection[284]), as well as detoxification of downstream metabolites like propionyl-CoA[251].

While it is clear from the aforementioned examples that cAMP regulates and contributes to many essential processes during infection, the role of cAMP signalling in modulating susceptibility to antimicrobials has not been investigated in mycobacteria. Several recent and historical reports have provided evidence that cAMP signalling is involved in antimicrobial susceptibility – potentially by regulation of drug or metabolite transporters in *Salmonella typhimurium*[261,262,264] and *E. coli*[263]. As a result, and to fulfil the 3rd aim of my PhD, I sought to investigate the link between cAMP signalling and antimicrobial susceptibility in mycobacteria.

As previously mentioned, cAMP signalling in mycobacteria is complex, with *M. tuberculosis* H37Rv encoding 16 enzymes that can produce cAMP[68,422]. Because of this functional redundancy, it has been difficult to modulate bacterial cAMP levels and measure the effects. Approaches with chemical activation of adenylate cyclases[11] or exogenous modulation of cAMP levels[207] demonstrate limited utility in an infection setting. To overcome this, my PhD project was designed to target mycobacterial cAMP signalling from the direction of cAMP degradation. Currently, Rv0805 is the only enzyme annotated in mycobacteria with cAMP degradation activity (a phosphodiesterase enzyme – PDE). Rv0805 possesses 150-fold increased specificity for other substrates[281], moonlights with non-cAMP catalysis related activities[308] and displays relatively poor cAMP hydrolysis activity[279] compared to *E. coli*[194], (which has 10-100 fold lower cAMP levels than mycobacteria[174,175,204,288]) and *Corynebacterium glutamicum*[321].

In experimental Chapter 3, I identified a novel potential cAMP PDE enzyme, Rv1339 – with the goal of finding a tool to effectively modulate mycobacterial cAMP levels. In Chapter 3, I classified this enzyme as an atypical member of the Class-II PDE family and showed that it displayed cAMP PDE activity *in-vitro* and when expressed in *Mycobacterium smegmatis* mc^{2}155. Furthermore, I generated an activity mutant by mutating[342] the key residue (D180A) that stabilises the metal ions required for substrate binding and catalysis[340]. I also investigated its effectiveness of this mutation at ablating cAMP PDE activity *in-vitro*.

In this Chapter (5), I investigate the phenotypic effects of lowering cAMP in the model mycobacteria *M. smegmatis* mc^{2}155. I also begin some preliminary work required for investigating the effects of Rv1339 overexpression in *M. tuberculosis* H37Rv. Briefly, I use a combination of metabolomics, antimicrobial susceptibility assays, RNA sequencing, bioenergetics and permeability assays to elucidate and characterise a link between cAMP signalling and antimicrobial susceptibility.

5.1 Expression of Rv1339 and Rv1339 D180

The metal binding sites of metallo-β-lactamase proteins are well characterised[423] and are important for stabilising substate binding interactions[340].

From bioinformatic analysis described in the previous Chapter, I was able to identify the metal binding motif involved in substrate binding. In order to investigate modulation of cAMP signalling, and any potential catalysis independent effects, an activity mutant was devised. The key residue, D180, responsible for stabilising both divalent metal ions[340] was altered by using a site directed mutagenesis protocol previously described[342].

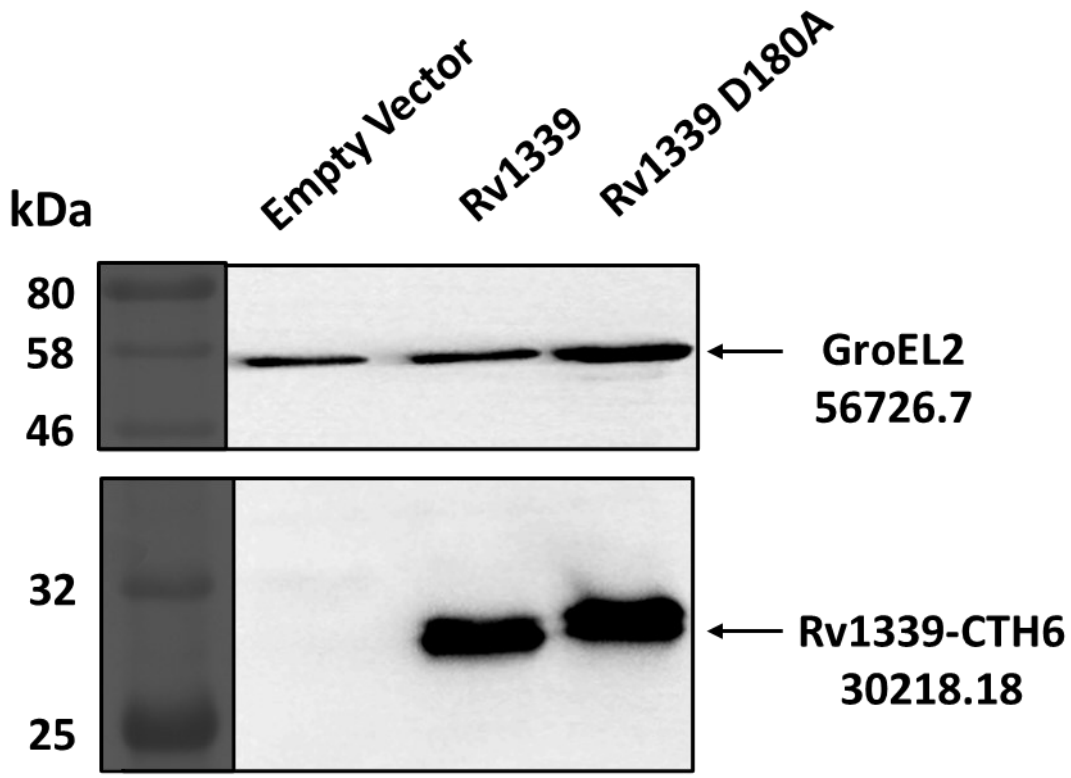


Figure 5.1: Western blot of Empty Vector Control, Rv1339 and Rv1339 D180 expressing M. smegmatis mc²155 strains. Proteins were probed with an α-His and α-GroEL2 antibodies, and each lane was loaded with 50µg of protein to normalise.

In order to confirm expression of the Rv1339 D180A mutant, western blot analysis with an α-His was performed. From this blot (Figure 5.1), it was clear that Rv1339 was expressed. However, it appears that the point mutation in Rv1339 caused the protein to migrate differently through the gel. This may be due to alterations in the binding to SDS or the conformation of the protein. Altered protein confirmation is most likely, as the GroEL2 (a mycobacterial HSP) loading control appears to have migrated at comparable levels to the Empty Vector Control and Rv1339.

5.1.1 Rv1339 expression leads to significantly decreased cAMP levels, increased turnover and a growth defect

In order to confirm that Rv1339 possessed cAMP hydrolysis activity *in-vivo*, I next performed metabolomic analysis of the *M. smegmatis* mc^{2}155 strains expressing the Empty Vector Control, Rv1339 and D180A (Figure 5.2). This analysis was performed by culturing the bacteria to the mid-log phase, transferring the bacteria to a nitrocellulose filter, incubating this filter on an agar plate, lysing the bacteria after 6-doubling times and extracting metabolites for running on the LC-MS (Figure 5.2A). Both the liquid culture and agar steps were performed with the defined medias 7H9 and 7H10, supplemented with Glycerol and Dextrose carbon sources. In order to study the "turnover" or rate of production of metabolites after transfer to the nitrocellulose filter, the bacterial containing filters were transferred to agar plates containing universally labelled [U-$^{13}C_3$]-glycerol and [$^{13}C_6$]-dextrose. The filters and labelled plates were then incubated for the specified time point, (0, 1.5 and 3.5 hours – corresponding to 1/2, and 1 doubling time respectively) before the bacteria were lysed and metabolites extracted for running on the LC-MS (this time for quantification of the ion counts and percentage label incorporation).

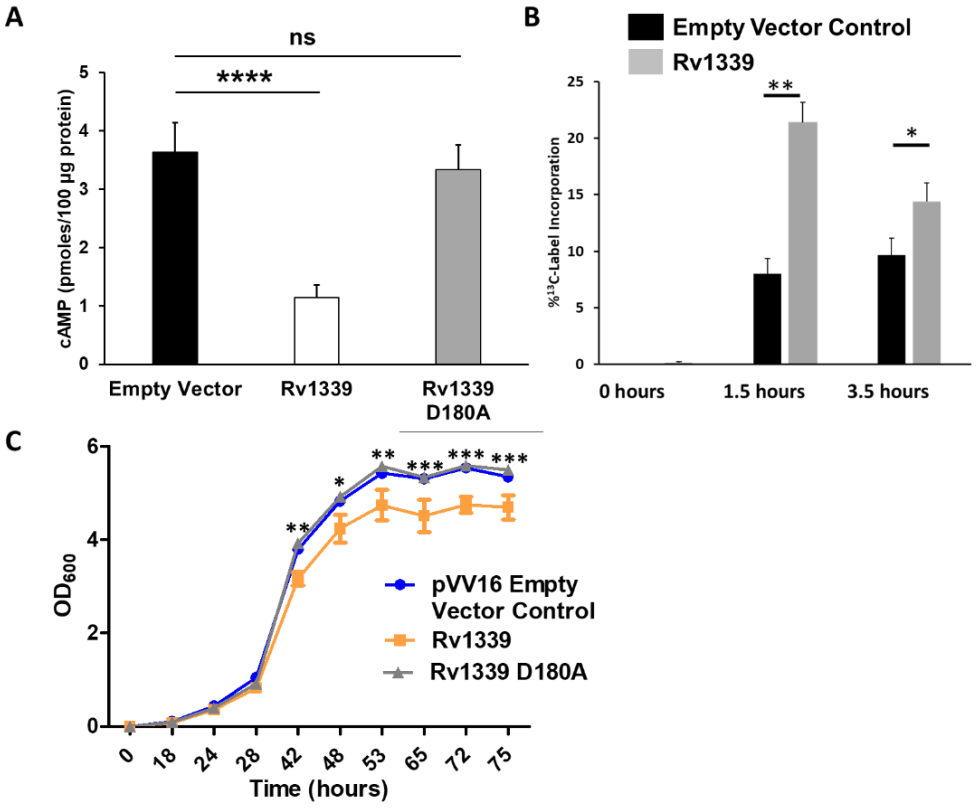


Figure 5.2: The effects of Rv1339 expression on cAMP levels, turnover and growth. A: LC-MS metabolomics analysis of cAMP levels within mid-log phase of Empty Vector Control, Rv1339 and Rv1339 D180A expressing M. smegmatis mc^{2}155. B: Turnover of cAMP in Empty Vector Control and Rv1339 expressing M. smegmatis mc^{2}155. 1C: Growth of M. smegmatis mc^{2}155 strains expressing the Empty Vector Control, Rv1339 and Rv1339 D180A. A and B: 2-tailed Student's's T-test *p<0.05, **p<0.01 or ***p<0.0001. C: Two-way ANOVA with Bonferroni multiple comparison tests. All panels representative of 2 independent biological replicates.

Expression of Rv1339 led to a significant 3.2-fold decrease in cAMP levels (Figure 5.2A). Concomitantly, the turnover of cAMP (as measured by ^{13}C-label incorporation) was also significantly increased by 3-fold (Figure 5.2B). These data indicate that both the absolute levels of cAMP are decreased and this is despite the fact that rate of production is increased (as increased label incorporation over time indicates increased *de-novo* production of this metabolite). When taken together, this strongly suggests that Rv1339 possess cAMP PDE activity *in-vivo*. Notably, the reduction in cAMP levels was seen to be ablated in the Rv1339 D180A mutant (Figure 5.2A), although this data was not available for metabolite turnover studies. This suggests that by destabilising the metal binding site, cAMP PDE catalysis activity is ablated. This is likely due to the disruption of the metal co-factors in the protein, which are predicted to be required for the nucleophilic attack of the phosphodiester bond carboxyl residue and the binding of substrate respectively. Both metal ions have been shown in mutational studies to be essential for full activity of a metallo-β-lactamase protein from *Bacillus cereus*[424].

Additionally, Rv1339 but not Rv1339 D180A expressing bacteria displayed a significant growth defect of around 30% when cultured in 7H9 defined media supplemented with Dextrose and Glycerol carbon sources (Figure 5.2C) (from time-points 42-75 hours, $p<0.05$ – $p<0.001$, as measured by two-way ANOVA with Bonferroni multiple comparison tests). This suggests that Rv1339 expression and reduction of cAMP levels has a significant effect on the bacterial physiology.

5.1.2 Increased Rv1339 expression positively correlates to increased growth defect in an anhydrotetracycline (ATC) inducible vector

To investigate the growth defect and develop a way to modulate the levels of Rv1339 expression, an Rv1339 N-terminal Flag tagged construct into the ATC inducible vector pKW. The expression system of pKW relies on the tetracyline repressor (TetR) and operator system (TetO). TetO acts as a promotor of downstream gene expression, but is repressed by TetR. When tetracycline (or its derivatives ATC or doxycycline) bind TetR, the conformation changes, and TetR repression of TetO is released – thereby leading to expression of the gene[425]. The amount of tetracycline or derivatives added therefore positively correlates with the level of expression. The gene *rv1339* was cloned down-stream of the TetO promotor, thereby positively linking induction of *rv1339* expression with ATC levels (Figure 5.3). ATC was chosen as the inducer because it was found to be non-toxic to the bacteria and in the studies where pKW was developed, showed the tightest regulation of expression[426]

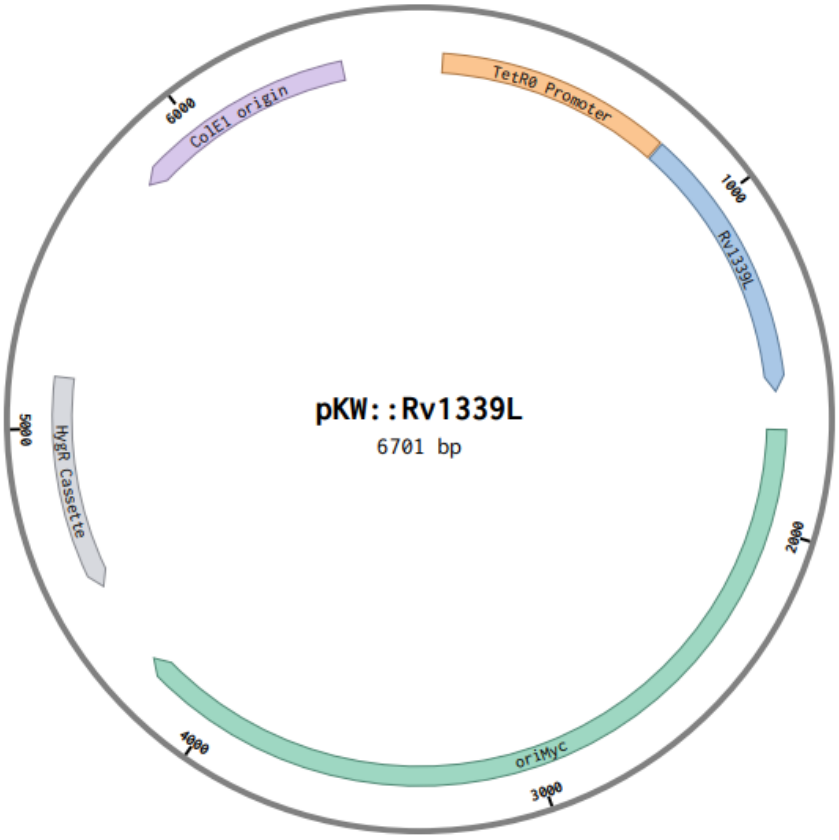


Figure 5.3: Vector map of the anhydrotetracycline (ATC) inducible Rv1339 construct for mycobacterial expression - pKW.

During the placement of an MSc student under my supervision, *M. smegmatis mc²155* bacteria were transformed with the pKW Empty Vector Control, or Rv1339 (Figure 5.4). Expression was tested by culturing the transformants with different concentrations of ATC (0 ng/ml, 200ng/ml and 500 ng/ml), lysing the bacteria, loading 50µg of clear soluble lysate on an SDS-PAGE gel and assaying with Coomassie stain/Western blot analysis (α-Flag tag antibody) (Figure 5.4A and 5.4B). A band for Rv1339 was detected at 0, 200 and 500ng of ATC (Figure 5.4B). The intensity of the bands, as measured by Fiji image analysis software, at each concentration was also comparable (Figure 5.4C) and did not show significant increase in response to higher concentrations of ATC. Furthermore, when the intrabacterial cAMP levels were measured by LC-MS quantification of ion counts normalised to sample protein levels, pKW::Rv1339 bacteria displayed 10-fold significantly decreased cAMP levels ($p<0.001$) (4D). However, this significant decrease in cAMP levels did not display any dose dependence on ATC levels. Taken together this suggested that the vector displayed "leaky" expression of Rv1339, even in the absence of inducer. Additionally, it indicated that at increasing ATC concentrations, expression of the protein in clear soluble lysate was similar to uninduced levels.

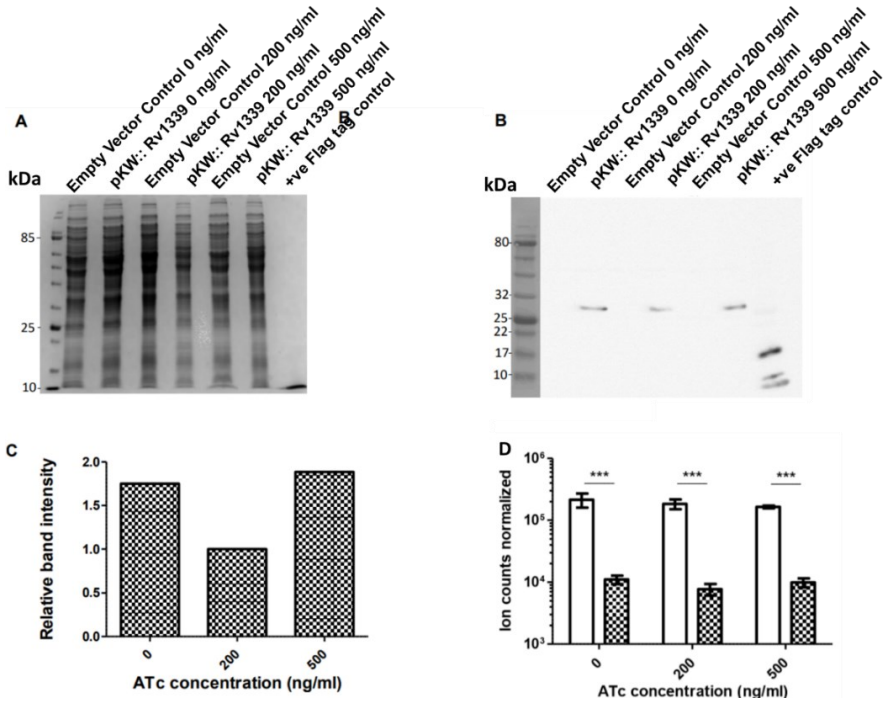


Figure 5.4: The effect of increasing ATC concentrations on pKW::Rv1339 expression. A: Coomassie stained SDS-PAGE of 50µg M. smegmatis soluble lysate from bacteria expressing pKW (Empty Vector Control) or Rv1339L at increasing concentrations of ATC. B: Western blot with α-Flag tag antibody for the Rv1339L C-terminal Flag tagged construct. C: Quantification of α-Flag tag Western blot band intensity at different concentrations of ATC as compared to uninduced levels. D: cAMP levels in pKW Empty Vector Control and Rv1339 expressing M. smegmatis mc²155, at different concentrations of ATC. p<0.001, where indicated – as measured by two-way ANOVA and Bonferroni multiple comparison tests. (Credit: Krista Grimes).

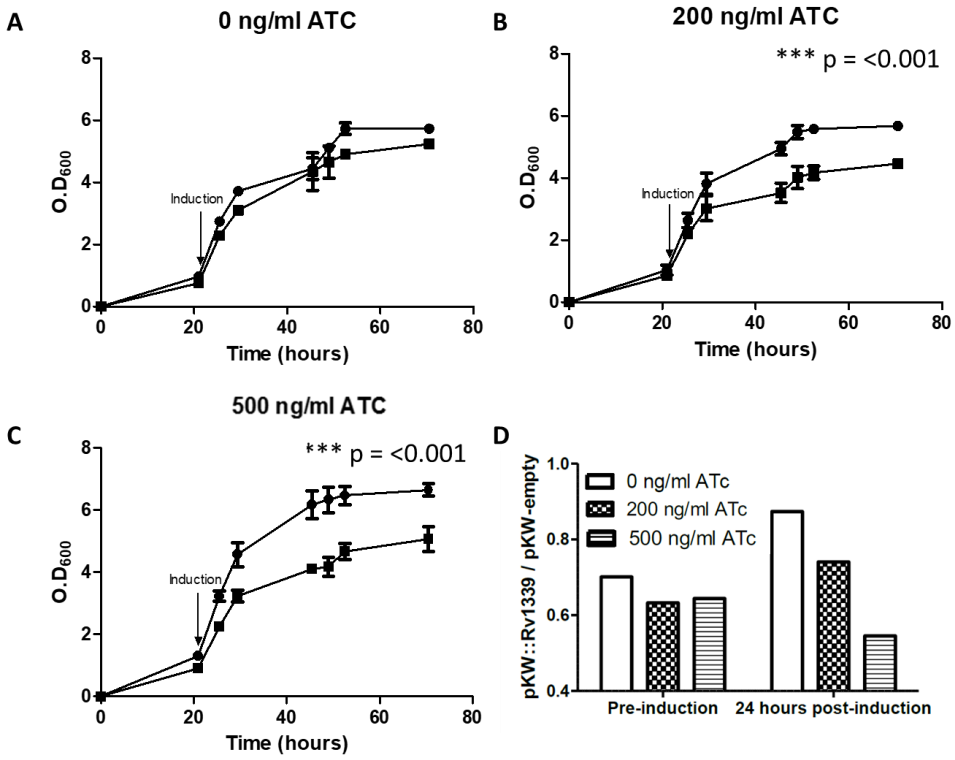


Figure 5.5: Growth curves of pKW Empty Vector Control and Rv1339 expressing M. smegmatis mc²155. A-C: Induction of pKW::Rv1339 expression with 0, 200 and 500ng/ml ATC. D: Normalisation of pKW::Rv1339 ODs as compared to Empty Vector Control. p<0.001, as measured by two-way ANOVA with Bonferroni multiple comparison tests.

However, when the growth of the strains was assayed, increased concentrations of ATC led to dose dependant growth defects (Figure 5.5) in Rv1339 expressing bacteria (Figure 5.5). Concentrations of ATC up to 1000ng/ml were shown to have no effect on bacterial growth, in the Empty Vector expressing strain (data not shown). This therefore suggested 2 possibilities. Firstly, that expression of the protein at the levels induced by ATC was toxic. It is conceivable that build-up of the protein in the insoluble fraction could cause the toxicity, without adding to cAMP catalysis. Secondly, Rv1339 may be expressed and then secreted. The latter hypothesis seemed more likely, as the signal peptide present at the *N*-terminal of full-length Rv1339 (as described in Chapter 4) would target the protein to the membrane and potentially for secretion[346,347]. My parallel focus on purification of the protein and further phenotypic characterisation precluded my continuation of this work. The next step would have been to optimise the inducibility of this construct and enrich the supernatant of bacterial cultures, for western blot analysis of the presence of Rv1339. Henceforth, all data reported utilises the constitutive expression plasmid, pVV16 – which had already shown excellent expression in previous metabolomic characterisation assays (Chapter 4).

5.2 Rv1339 expression leads to increased antimicrobial susceptibility in *M. smegmatis mc²155*, and the D180A catalysis mutant can ablate this increase

In recent years, there has been increasing evidence to support the hypothesis that cAMP signalling is involved in regulating the susceptibility of bacteria to antimicrobials[71,262]. In light of the previous data showing that cAMP levels could be effectively decreased within the bacteria, I therefore decided to measure the antimicrobial susceptibility of Empty Vector Control and Rv1339 expressing *M. smegmatis* mc²155 strains. To begin, I chose a cross-section of frontline TB antimicrobials with varied MOAs. This was to investigate whether Rv1339 expression and decreased cAMP levels led to a broad alteration in antimicrobials susceptibility, or if the response was specific to particular bacterial processes. For example, Rifampicin targets bacterial RNA polymerase and inhibits effective transcription[54,427.] . It is highly effective and has been shown in *M. smegmatis* mc²155 to inhibit phenylalanine incorporation within 6 minutes of administration[428]. Streptomycin targets protein production by binding the 16s subunit of the bacterial ribosome and inhibiting effective translation[429]. It was the first anti-TB antibiotic and studies into spontaneous development of resistance in *M. tuberculosis* isolates during monotherapy, prompted the development of combination therapy for TB[58]. Rifampicin and streptomycin are both bactericidal antibiotics at clinically relevant doses[430]. Ethambutol targets the synthesis of arabingalactan[431–433], a key mycobacterial cell wall component[434].

Firstly, the minimum inhibitory concentration (MIC_{90}) of the chosen antimycobacterial drugs was measured for the *M. smegmatis* mc²155 strains expressing Rv1339 or the Empty Vector Control (during supervision of an additional MSc student – Kanokkan Nunta). No significant alterations in the MIC_{90} of rifampicin and ethambutol were observed; with only a minor difference in streptomycin susceptibility. This indicates that resistance levels of the bacteria were not significantly affected by Rv1339 expression (Figure 5.6A). The MIC_{90} is an essential parameter for running TTK assays – as it enables relevant concentrations of antibiotic to be used. For example, in addition to measuring bacterial CFU counts during treatment with 0x MIC_{90} (no antibiotic) and 1x the MIC_{90} with each antibiotic, we also quantified the TTK at the C_{max} concentration of each respective antibiotic. This value is the peak concentration of antibiotic found in patient serum during treatment[435] – and is a clinically relevant measure of antimicrobial susceptibility. In this way, we aimed to see the effects on the susceptibility profiles of the bacteria both in the standard lab setting (1x MIC) and in the clinic (C_{max}).

When the time-to-killing was measured in Rv1339, Rv1339 D180A and Empty Vector control expressing *M. smegmatis* mc²155 strains – a trend indicating increased streptomycin (figure 5.6B) and rifampicin (Figure 5.6C) susceptibility could be observed in Rv1339 expressing bacteria. The concentrations of antibiotic used ranged from 0x, 1x and the relevant C_{max} for each individual antibiotic (20x MIC for streptomycin and 4x MIC for Rifampicin). When cultured without antibiotic (0 x MIC_{90}

– light blue lines), bacterial CFU counts increase over time, from an initial 6 $\log_{10}$ to a maximum of between 9.5 and 10 $\log_{10}$ CFU/ml in all strains (Figure 5.6B streptomycin and 6C rifampicin).

Conversely, when the bacteria are treated with 1x MIC_{90} (purple line), bacterial CFU counts decrease over time from the initial 6 $\log_{10}$ to below the limit of detection of our assay – 2 $\log_{10}$ CFU/ml (5.6B streptomycin and 5.6C rifampicin). This indicates that treatment with antibiotic at the MIC_{90} effectively reduces CFU over time.

Furthermore, when treated with the C_{max} concentration of each antibiotic (20x MIC_{90} for streptomycin and 4x MIC_{90} for rifampicin – orange line) bacterial CFU counts were reduced below the limit of detection more rapidly in all strains treated with streptomycin (Figure 5.6B) and rifampicin (Figure 5.6C).

When differences between Empty Vector Control (solid lines), Rv1339 (dashed lines) and Rv1339 D180A (dotted lines) expressing bacterial are investigated, Rv1339 expressing bacteria appear more susceptible to streptomycin and rifampicin treatment at both 1x MIC and the C_{max} – when compared to both Empty Vector Control and Rv1339 D180A expressing bacteria (figure 5.6B and Figure 5.6C, respectively).

A

Antibiotics	MIC (µg/mL) *M. smegmatis* pVV16	MIC (µg/mL) *M. smegmatis* pVV16::*r1339*
Ethambutol	2.0 ± 0.0	2.0 ± 0.0
Rifampicin	50.0 ± 0.0	50.0 ± 0.0
Streptomycin	0.50 ± 0.0	0.25 ± 0.0

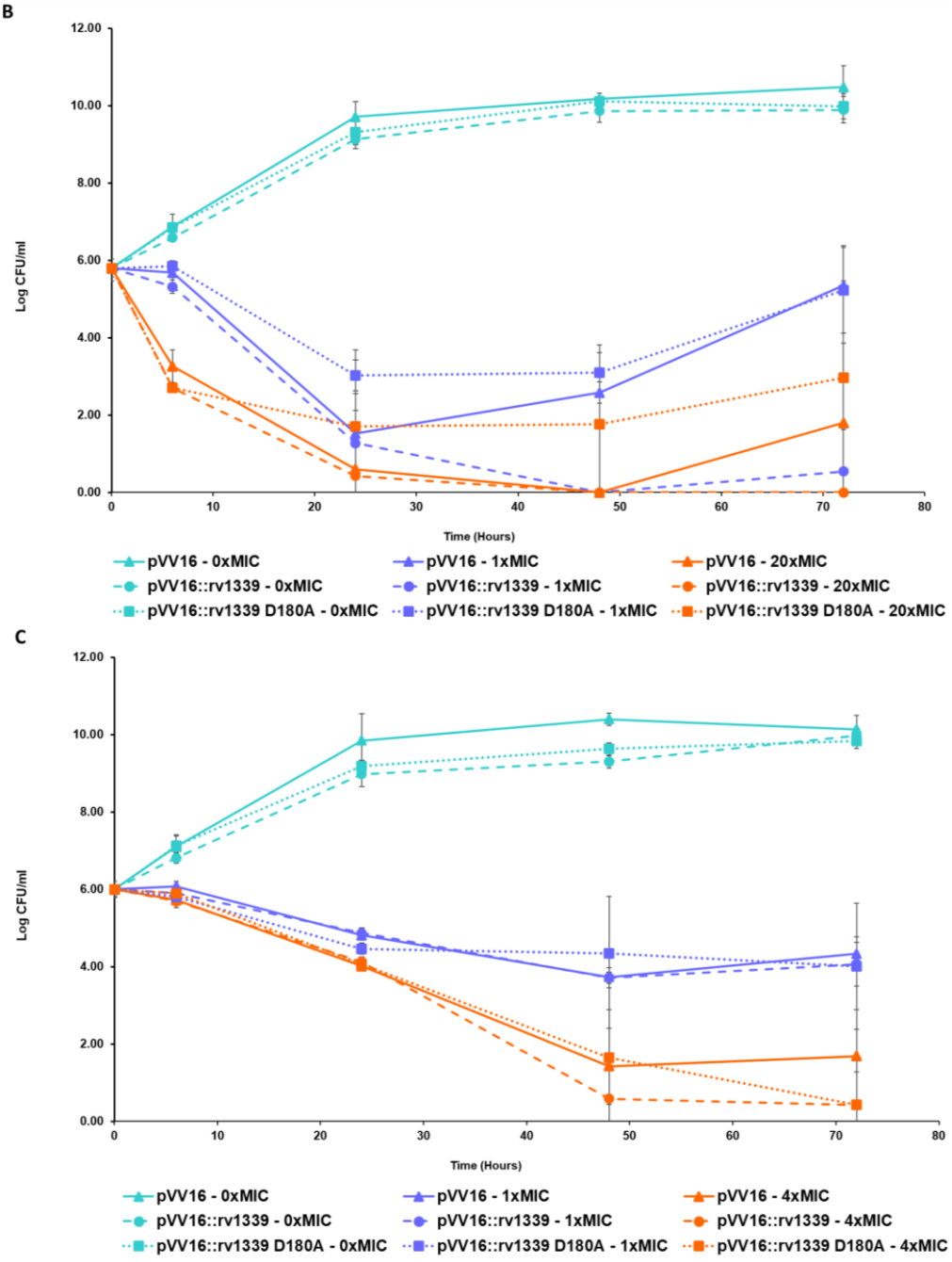


Figure 5.6:. A: MIC_{90} of various frontline TB antimicrobials. B: The time-to-killing of Streptomycin treated Empty Vector Control, Rv1339 or Rv1339 D180A expressing M. smegmatis mc^2155 strains at 0, 1x and 20x MIC. C: The time-to-killing of Rifampicin treated Empty Vector Control, Rv1339 or Rv1339 D180A expressing M. smegmatis mc^2155 strains at 0, 1x and 4x MIC. All panels representative of 2 independent biological replicates. Collaboration with Kanokkan Nunta.

However, statistical analysis indicated that there was only significant difference between Rv1339 and Empty Vector Control/ D180A strains at 1x MIC_{90} treatment with streptomycin (Bonferroni tests $p < 0.001$) (Figure 5.6B). This may be a result of rapid bactericidal activity at the clinically relevant levels of antibiotic, combined with the limit of detection of our assay (2 $\log_{10}$ CFU). Optimisation for the time points of the TTK assay used or the concentration of antibiotic may be able to mitigate these limitations

Intriguingly, although regrowth was observed in 1x MIC streptomycin treated bacteria at around 50 hours in Empty Vector Control and Rv1339 D180A expressing strains, Rv1339 expression appeared to ablate this regrowth (Figure 5.5B). This may suggest that Rv1339 expression not only increases the susceptibility to streptomycin, but also potentially prevented the development of phenotypic tolerance development in streptomycin treated *M. smegmatis* mc^2155. An alternative, less likely hypothesis, could be that after 24 hours the streptomycin antibiotic degraded in the Empty Vector Control and Rv1339 D180A conditions. As previously mentioned, both streptomycin and rifampicin are bactericidal antibiotics. In the case of the potential regrowth under streptomycin treatment, the antibiotic may select for a resistant sub-population of bacteria (e.g. with increased efflux activity or an alteration of target mutation), and this could explain the regrowth. As this is not observed in Rv1339 expressing bacteria, this would suggest that Rv1339 expression somehow mitigates the difference of this subpopulation, or potentially prevents the development of this population. In order to investigate this, whole genome sequencing of each strain, at each concentration of streptomycin, should be undertaken at 0, 24- and 72-hour time points. This would elucidate whether spontaneous mutations were responsible for allowing bacterial regrowth and how this process might be altered in Rv1339 expressing bacteria. Another crucial investigation would be to quantify antibiotic levels and degradation over time, via LC-MS analysis of streptomycin, and its degradation products. However, the time remaining of my PhD did not allow me to investigate this further.

Interestingly, when the time-to-kill assay was repeated with the bacteriostatic antibiotic ethambutol (bacteriostatic even at clinically relevant doses), the bacteria were also more susceptible (Figure 5.7). In fact, when used to treat Rv1339 expressing bacteria with decreased cAMP levels, ethambutol became bactericidal, sterilising the culture. Taken together, these data clearly indicate that Rv1339 expression, and potentially decreased cAMP levels, renders the bacteria more susceptible to antimicrobials. This suggests that lowering cAMP levels may be enough to turn a bacteriostatic antibiotic into a bactericidal antibiotic.

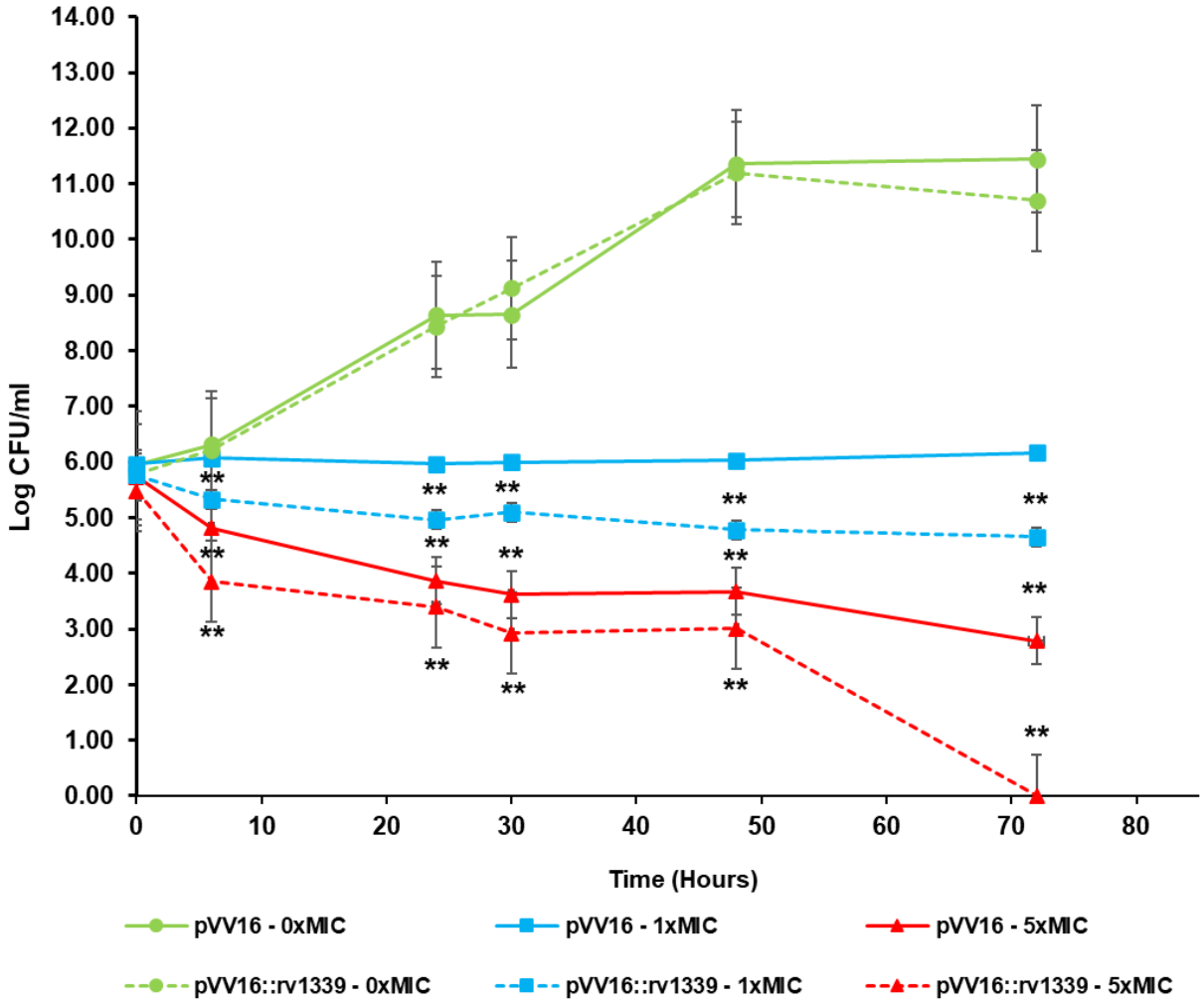


Figure 5.7: TTK for the bacteriostatic antibiotic, ethambutol when treating Empty Vector Control and Rv1339 expressing M. smegmatis mc^{2}155 strains. $p<0.01$, as measured by two-way ANOVA with Bonferroni multiple comparison tests. Representative of 2 independent biological replicates of technical quadruplets. In collaboration with Kanokkan Nunta.

All 3 of the antibiotics used for TTK assays, possessed different mechanisms of action. This suggests that it is unlikely that the adaptions mediated by Rv1339 expression and decreased cAMP levels are specific compensations to a particular antibiotic. This therefore implies that the difference(s) observed when Rv1339 is expression and cAMP levels decreased, are a result of generalised changes in bacterial physiology. In order to investigate any generalised changes in bacterial physiology RNA sequencing was performed in collaboration with the Imperial College Genomics facility.

5.3 Expression of Rv1339 leads to an altered transcriptome, with decreased expression of genes known to be regulated by cAMP

From the literature, the best characterised mycobacterial cAMP binding effector proteins are Crp and Cmr[243]. These transcription factors have been shown to regulate processes ranging from virulence[177,206,241] to carbon metabolism[251] and dormancy[207]. It was therefore likely that decreased cAMP levels would alter the transcriptome – and so RNA sequencing was performed on Empty Vector Control and Rv1339 expressing *M. smegmatis* mc^2155 bacteria. The bacteria were grown to the mid-log phase, in 7H9 media with Glycerol and Dextrose as the carbon source. Differential gene analysis was then performed, with the help of another PhD student from our lab – Ashleigh Cheyne.

4 replicates of each strain were cultured, and the RNA extracted. DNA digests were performed on the samples and their quality was assessed with a bioanalyzer. The 3 samples with the best RNA integrity (RIN[436] – a value out of 10, 10 being the best RNA quality and with a threshold of RIN 7 deemed acceptable for RNA-seq) values were then sent to the Imperial College Genomics Facility, for RNA sequencing (detailed in Table 5.1)

Sample Description	Volume [µl]	Diluent	Concentration [ng/µl]	Quantification Method	RIN
Empty Vector Control 1	50	Water	2454.8	Bionalyzer 2100	9.6
Empty Vector Control 2	50	Water	1435.2	Bionalyzer 2100	9.7
Empty Vector Control 3	50	Water	2537.2	Bionalyzer 2100	9.7
Rv1339 expressor 1	50	Water	2322.7	Bionalyzer 2100	9.9
Rv1339 expressor 2	50	Water	2151	Bionalyzer 2100	9.8
Rv1339 expressor 3	50	Water	3507.1	Bionalyzer 2100	9.9

Table 5.1: RNA concentrations and RIN values for each sample sent to RNA sequencing.

M. smegmatis mc^{2}155 harbours two Crp genes (Crp 1 and Crp 2[437]). A previous study investigated the effect on the transcriptome of knocking out Crp 1 or overexpressing Crp 2[211]. The reason for overexpressing Crp 2 was that attempts to knock it out were repeatedly unsuccessful, suggesting it was essential for *M. smegmatis* mc^{2}155[211]. When Crp 1 was knocked out, the authors saw significantly increased expression of 54 genes. Many were involved in bioenergetic processes, such as ATP production, succinate dehydrogenases and a variety of ABC transporters[211]. This suggests that Crp 1 mediates repression of these genes. Furthermore, in the Crp 2 overexpressor mutant, 20 genes displayed significantly increased expression of between 1.5 and 6-times[211]. This suggests that Crp 2 promotes expression of these genes, which also included several succinate dehydrogenase subunits and the resuscitation promoting factor *rpfE*. Another study also suggests differential roles of Crp 1 and Crp 2 with negative and positive gene regulation respectively[437]. This study functionally validated this theory with a single *M. smegmatis* mc^{2}155 gene.

In our approach, we were not knocking out or overexpressing the different Crps. Instead, we were expressing Rv1339, decreasing cAMP levels and measuring the effect of this on the transcriptome. The data from our analysis are presented in Figure 5.8.

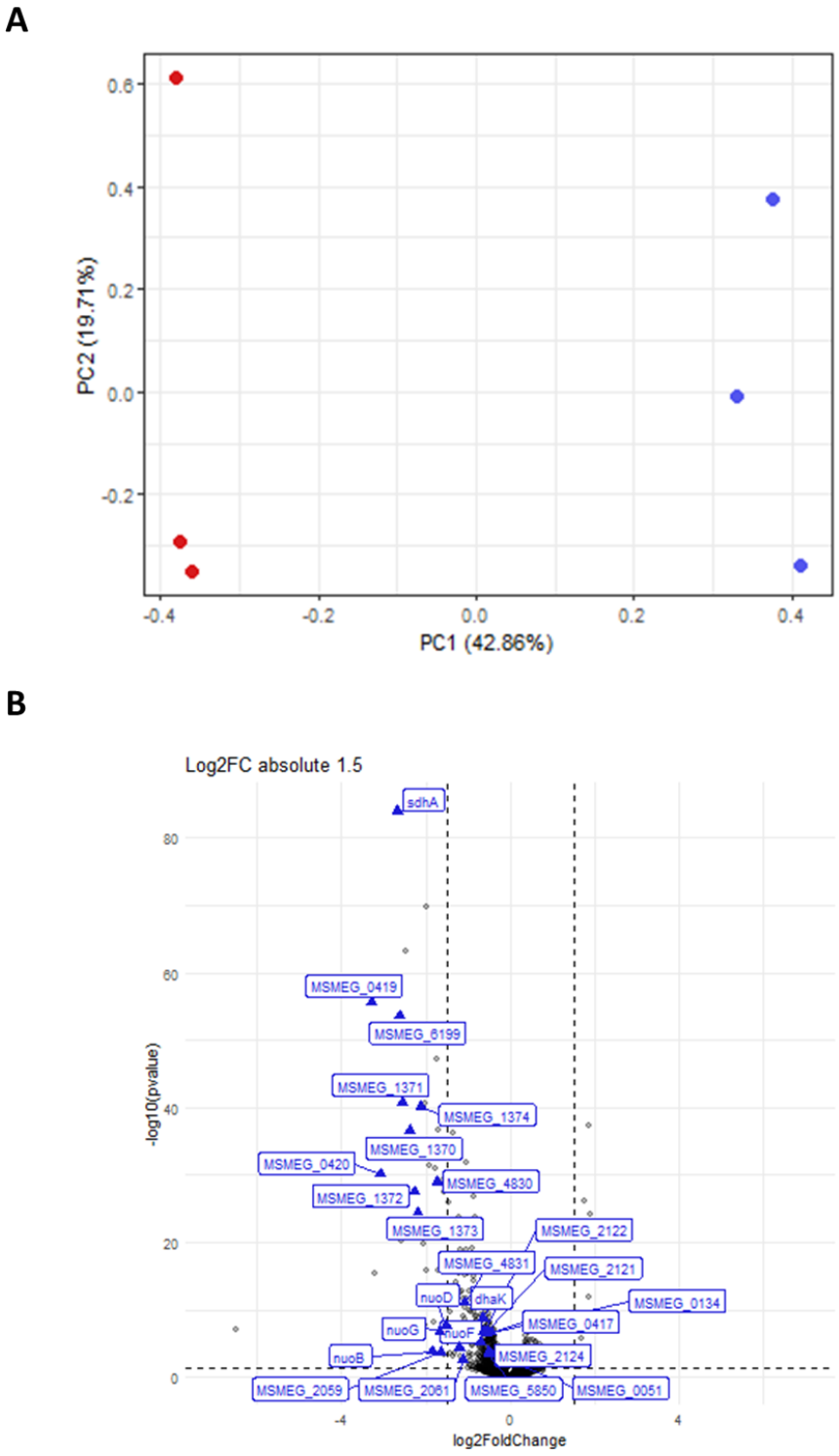


Figure 5.8: RNA sequencing of Empty Vector Control and Rv1339 expressing M. smegmatis mc^{2}155 strains in the mid-log phase of growth when grown in 7H9 media supplemented with glycerol and dextrose. A: PCA of the 2 bacterial strains. Red corresponds to Rv1339 and Blue to Empty Vector Control samples. B: Volcano plot of significantly differentially expressed genes ($p<0.05$) in M. smegmatis mc^{2}155 Empty Vector Control and Rv1339 expressing M. smegmatis mc^{2}155. Labelled genes have previously been shown to be regulated by cAMP. The Y-axis displays the $\log_{10}$ p-value plotted against the $\log_2$ fold change on the X-axis. Data representative of 3 technical replicates.

The result from our RNA-seq and differential gene analysis showed that when Rv1339 is expressed and cAMP levels significantly decreased, 387 genes were significantly differentially expressed (Annex 2). Of the 387 genes, the majority showed decreased expression (312) and 40 genes were significantly downregulated by 1.5-fold or more ($p< 0.05$). Conversely, only 5 genes were upregulated by 1.5-fold or more ($p< 0.05$). The 5 upregulated genes were MSMEG_2013-2017 (Figure 5.8B). The 2 genes that displayed the highest downregulation were MSMEG_5321 (-3.3-fold) and MSMEG_5322 (-6.6-fold). These genes together likely form an *M. smegmatis* mc^2155 homologue of Rv1651c, a PE/PPE channel protein. This family of proteins has recently been elucidated to mediate nutrient transport across the outer membrane of *M. tuberculosis* [438], but a previous study expressing Rv1651c in *M. smegmatis* mc^2155 showed that it led to a delay in the lag-phase of bacterial growth[439]

3 putative WhiB family proteins were significantly downregulated. Some members of this family have been identified as putative cAMP binding transcription factors (though direct cAMP binding has never been proved)[421,440,441]. Unfortunately, these genes are uncharacterised in *M. smegmatis* mc^2155. WhiB3 has been predicted to negatively regulate expression of molybdate transporters and regulate redox homeostasis[442]. This is interesting as in our data, the 5 genes identified as having 1.5-fold or more increased expression appear to be in a molybdate transporter operon or adjacent to it (MSMEG_2013-2017). The increased expression of these transporters when WhiB family transcription factor expression is decreased supports the hypothesis that these transcription factors negatively regulate molybdate transport.

In terms of the genes regulated, our data was in agreement with observations from the Crp 1 and Crp 2 microarray study[211]. However, the authors observed that when Crp 1 was deleted, there were significant increases in the expression ratio of 54 genes. This suggests that Crp 1 acts as a transcriptional repressor and this could be mediated by promotor exclusion. This mechanism involves binding of Crp 1 at a point upstream of the gene, that prevents adequate binding of RNA polymerase. It has previously been postulated that Crp proteins are capable of both positive and negative regulation, and that negative regulation is mediated by promotor exclusion[177,437,443]. The aforementioned microarray study showed that both the putative Crp 1 and Crp 2 regulons had 11 genes in common. They further postulated that Crp 2 was a positive regulator of expression of *sdh1*, whereas Crp 1 was a negative regulator, based on observations. The binding of Crp 1 to identified Crp binding sites was shown to be cAMP independent during the microarray study, whereas Crp 2 binding was thought to be cAMP dependent[211]. However, another study suggests that binding of Crp 1 to a site regulating the MSMEG_3781 adenylate cyclase gene was cAMP dependent[437]. Furthermore, the study investigating Crp regulation of MSMEG_3781 suggested that Crp 2 was repressive, and binding was cAMP independent. The 2 studies appear to contradict each other. However, in other studies of Crp binding to genes within the *M. tuberculosis* H37Rv, multiple Crp binding sites were found within and outside genes[173,237]. Likewise, in order for a transcription factor to mediate positive or negative regulation (in this case by promotor exclusion),

binding at different sites would be required e.g. further upstream within the previous gene, or within an intergenic region. We therefore propose the model, that depending on the site of binding, the different Crps can be positive or negative regulators of expression. The disagreement in the literature suggests that neither Crp 1 or Crp 2 can be considered definitively positive or negative regulators. Our data is more supportive of the microarray study, and in our model (Figure 5.9), we suggest that Crp 1 mediates repression of gene expression by binding to upstream sites and acting by promotor exclusion. This binding is likely cAMP independent, and in the presence of low levels of cAMP either: release of Crp binding and repression could be mediated by binding of cAMP or that positive regulation requires Crp 1 binding at a different site, that is mediated by the conformational change of cAMP binding.

In our data, genes in the Crp 1 regulon showed decreased expression. It would therefore make sense that in the presence of reduced cAMP levels, that Crp 1 mediated gene repression is impaired. Likewise, in the microarray study, it was postulated that Crp 2 was a positive regulator, significantly increasing expression of 18 genes found in the Crp regulon that decrease in our data. This further supports the hypothesis that Crp 2 is a positive regulator. Overall, the presence of 2 Crps, and disagreement in the literature, means that it is difficult to conclude anything further. Our hypothesis is summarised in Figure 5.9.

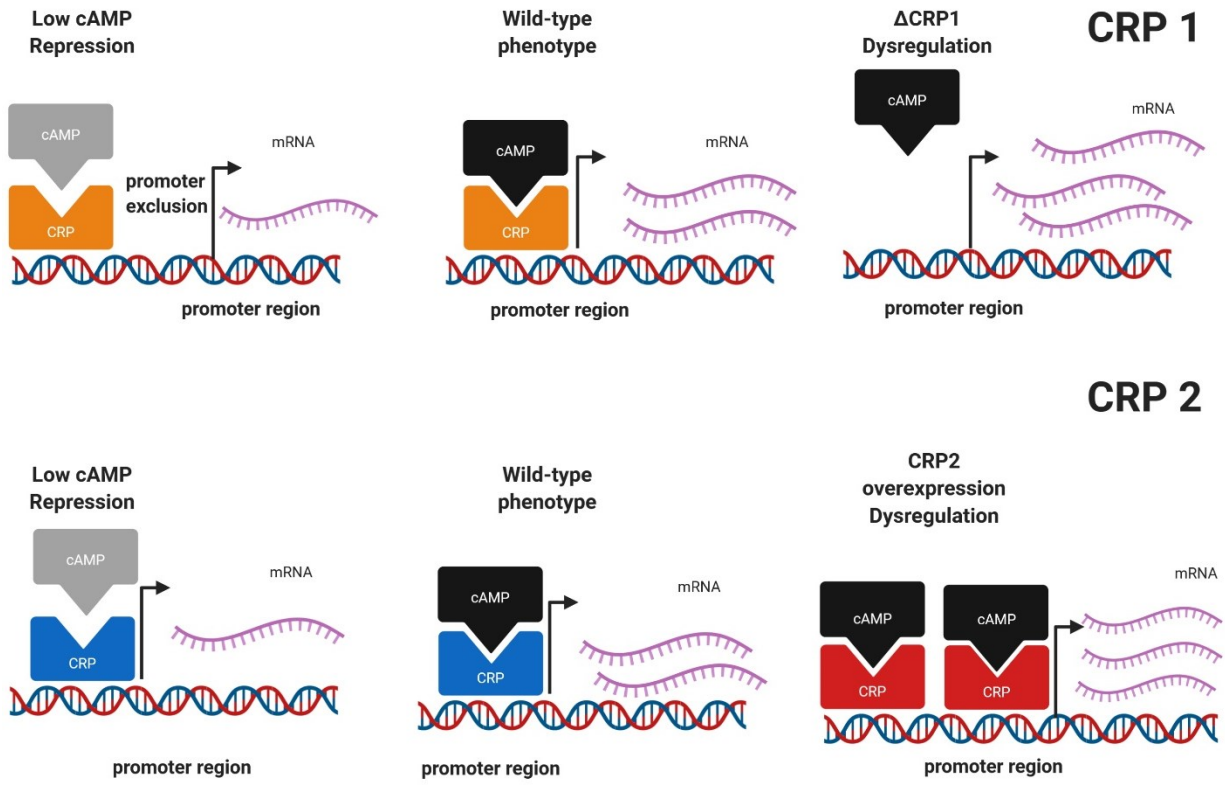


Figure 5.9: Our proposed model for the respective repression and activation of gene expression by Crp 1 and Crp 2, under low cAMP (this study), Crp 1 knockout or Crp 2 overexpression conditions[211]. In the low cAMP condition, Crp 1 and Crp 2 appear to repress expression of genes. Crp 1 via cAMP independent binding to a promotor exclusion point and Crp 2 via insufficient promotor binding when cAMP levels are reduced. When Crp 1 is knocked out, there can be no repression, thereby leading to dysregulation. Similarly, when Crp 2 is overexpressed, there are more Crp 2 cAMP complexes formed, and consequently greater levels of promoted gene expression.

14 out of 40 of the significantly downregulated genes in our data were previously seen to be regulated by Crp 1 and Crp 2. This included the 3 out of 4 genes encoding the Sdh1 Succinate Dehydrogenase (MSMEG_0417-0420) operon[444] – which were decreased by between 3 and 4-fold. It also included 6 out of 14 genes encoding the Nuo NADH dehydrogenase operon (all by around 2.5-fold). Together, these genes form part of complex II and I of the electron transport chain respectively, and in *M. tuberculosis,* the activity of Sdh1 has been shown to be essential for survival and maintenance of proton motive force (PMF) under hypoxic infection conditions[238,445]. Taken together, this suggests that the bacterial bioenergetics of the Rv1339 expressing bacteria may be compromised.

5.4 Rv1339 expression decreases succinate dehydrogenase activity, decreases membrane potential and leads to compromised bioenergetics

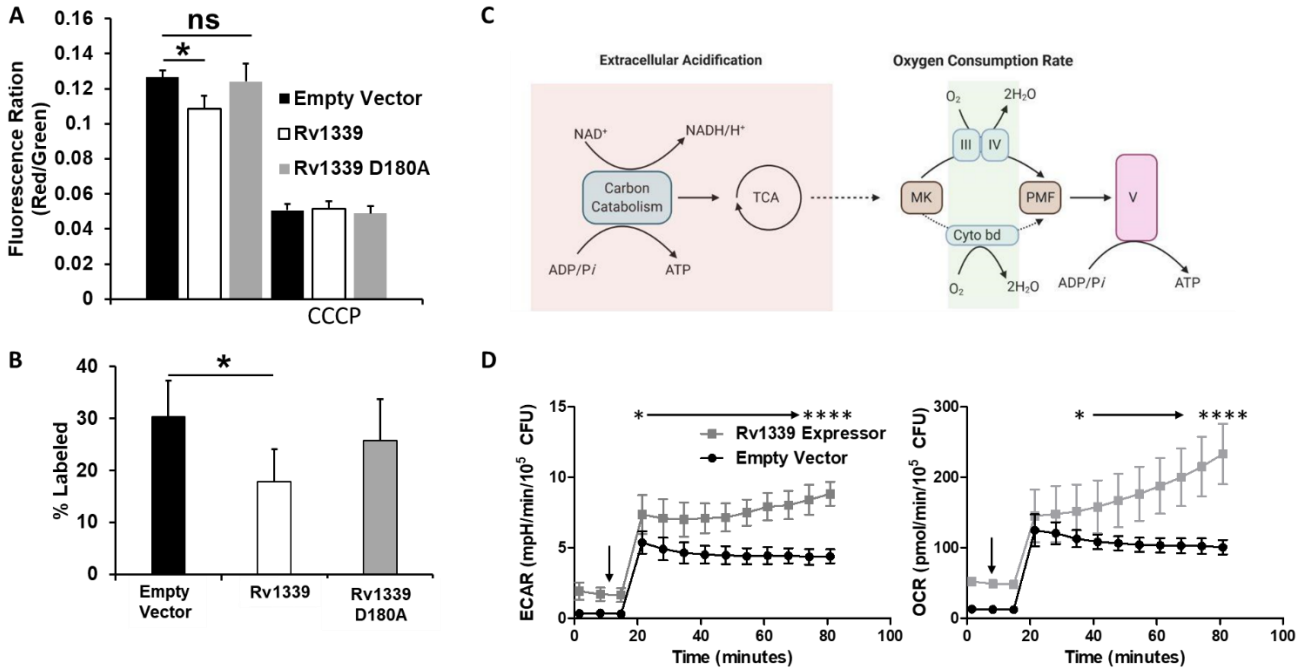


Figure 5.10: Investigation of the bioenergetics of Empty Vector Control and Rv1339 expressing M. smegmatis mc²155 strains. A: Membrane potential measurements of Empty Vector Control, Rv1339 and Rv1339 D180A expressing bacteria. B: Succinate turnover as measured by ^{13}C-label incorporation. C: Schematic of the different aspects of bioenergetics that are measured by the Seahorse XFP analyser – MK = menaquinone pool, PMF = protonmotive force. D: Extracellular acidification rate and oxygen consumption rate of Empty Vector Control and Rv1339 expressing M. smegmatis mc²155 strains. Statistical analysis: two-way ANOVA and bonferoni multiple comparison tests - (Cycle 4 from the OCR and onwards is p<0.05 – p<0.0001, Cycle 6 onwards from ECAR p< 0.05 – to p<0.0001). Panels A, B and D are representative of 2, 2 and 3 independent biological replicates.

To investigate the effect of Rv1339 expression and decreased cAMP levels on bacterial bioenergetics, a multipronged approach was used. Firstly, the membrane potential in Empty Vector Control, Rv1339 and Rv1339 D180A expressing M. smegmatis mc²155strains was quantified. Rv1339 expressing bacteria showed a significant ($p<0.05$) 15% decrease in membrane potential when compared to the Empty Vector Control (Figure 5.10A). This decrease was not observed in Rv1339 D180 catalysis mutant expressing bacteria (Figure 5.10A). The membrane potential is a key determinant of the protonmotive force (PMF). PMF drives the proton gradient across the bacterial membrane required for ATP synthesis[212,446]. A decrease in membrane potential would impact the protonmotive force (PMF) and so could compromise bacterial bioenergetics and viability (Figure 5.10A).

Secondly, the decrease in Sdh1 expression previously observed was investigated by analysing the turnover of succinate with ^{13}C-isotopologue analysis (Figure 5.10B). As previously mentioned, Sdh1 catalyses the conversion of succinate to fumarate[444], therefore a read-out of decreased Sdh1 activity would be decreased turnover of succinate. The rate of succinate turnover was indeed significantly decreased by 33% ($p<0.05$) in Rv1339 expressing bacteria as compared to Empty Vector Control expressing bacteria. This significant decrease was not observed in The Rv1339 D180 cAMP catalysis mutant strain (Figure 5.10B).

The activities of key sections of bacterial bioenergetics can be investigated by measurement of the extracellular acidification rate (ECAR) and oxygen consumption rate (OCR)[305,306,442] (Figure 5.10C). Activities measured by ECAR include carbon catabolism and the tricarboxylic acid cycle (TCA) – this is due to production of H^+ as a result of glycolysis, or NADH/H^+. NADH is a reducing equivalent/electron donor, that feeds electrons into the menaquinone pool (MK) and subsequently, the electron transport chain (ETC). Succinate is also a reducing equivalent/electron donor, but succinate dehydrogenase activity is not measured by the ECAR. Mycobacteria, such as *M. tuberculosis* and *M. smegmatis*, have a branched respiratory chain that allows electrons from the MK to be routed through either complex III and IV of the ETC (cytochrome bc) or cytochrome bd[447]. This allows bioenergetic flexibility that can compensate for treatment of antimicrobials that target portions of the electron transport chain[447,448].

Either branch of the respiratory chain leads to the acceptance of electrons from the Menaquinone (MK) pool, the consumption of oxygen and production of H_2O (Figure 5.10C). Therefore, the OCR is a readout of electron transport chain activity. The Seahorse XFP analyser utilises a closed system where bacterial oxygen consumption rate (OCR – a readout of oxidative phosphorylation/the electron transport chain activity) and the extracellular acidification rate (ECAR – a readout of TCA cycle and carbon catabolism activities) are measured over time[305,306]. This is mediated by probes that detect changes in pH (H^+ concentration) or O_2 consumption, as well as a mixing system which ensure oxygen and nutrient availability[305]. Increased OCR and ECAR are therefore readouts of bioenergetic activities. Over time, as these processes continue, oxygen will be consumed and the H^+ of the media will increase. Alterations in bacterial bioenergetics can lead to alterations in the level or rate of these increases. Furthermore, the closed system enables injection of compounds, carbon sources or even antibiotics at any point, and the response to this injection can be measured[305,306].

In order to gain further insight into potential alterations of bacterial bioenergetics Seahorse XFP analysis was performed on the 3 *M. smegmatis* mc^2155 *strains*. A protocol was designed and informed by previous studies in the literature[305,449] (Figure 5.10D). Briefly, bacteria were cultured to the mid-log phase in media containing Glycerol and Dextrose. The bacteria were centrifuged, washed with unbuffered media without carbon source, and then suspended in 1 mL of unbuffered, carbon source free

media. These bacteria were then loaded into the Seahorse XFP analyser plate and entered the closed system (temperature 37˚C). Bacterial numbers between wells and strains were normalised by CFU counts.

After 3 cycles of measurements over the course of 20 minutes (1 cycle = 4 minutes of mixing and 2 minutes of reading), the basal OCR and ECAR in the absence of carbon source were measured, in order to gain an understanding of basal respiratory capacity. It was clear from the data that Rv1339 expressing bacteria had a significantly higher basal OCR and ECAR over the course of the initial 3 measurements (~20 minutes). Already this indicates that the bioenergetic machinery is working at increased levels in these bacteria, even in the absence of carbon source. After 20 minutes Glycerol and Dextrose were injected to a final concentration of 0.2% - in line with the concentration of carbon source in the media the bacterial culture media used in all other experiments. Once carbon source was injected, all strains of bacteria assayed saw an increase of more than 10-fold in OCR and ECAR. However, in Rv1339 expressing bacteria, the levels of OCR and ECAR were maintained at significantly higher levels over the remaining course of the run (~60 mins/10 cycles – p-values in figure legend) This indicates that even in the presence of carbon source, the bioenergetic machinery is again forced to work much harder than in the Empty Vector Control strain.

Rv1339 D180 catalysis mutant expressing bacteria displayed similar OCR and ECAR levels with WT throughout the assay with no statistically significant differences (Annex 5).

Increased OCR and ECAR activity should correlate to increased production of ATP (from carbon catabolism and/or TCA cycle and/or ETC activities), the precursor of cAMP and the critically important energy source for a multitude of bacterial activities. Likewise, increased ECAR should correlate to increased NADH levels, as NADH is produced from carbon catabolism and TCA cycle activity[305].

To test this, ATP levels in all 3 strains were measured (Figure 5.11A). Surprisingly, despite the significant increase in bioenergetic rates, no significant difference in ATP levels was observed across the 3 strains. Although Aung et al[211] found regulation of bd cytochromes and ATP synthase components by Crp 1, there were no significant increases in these genes in our data.

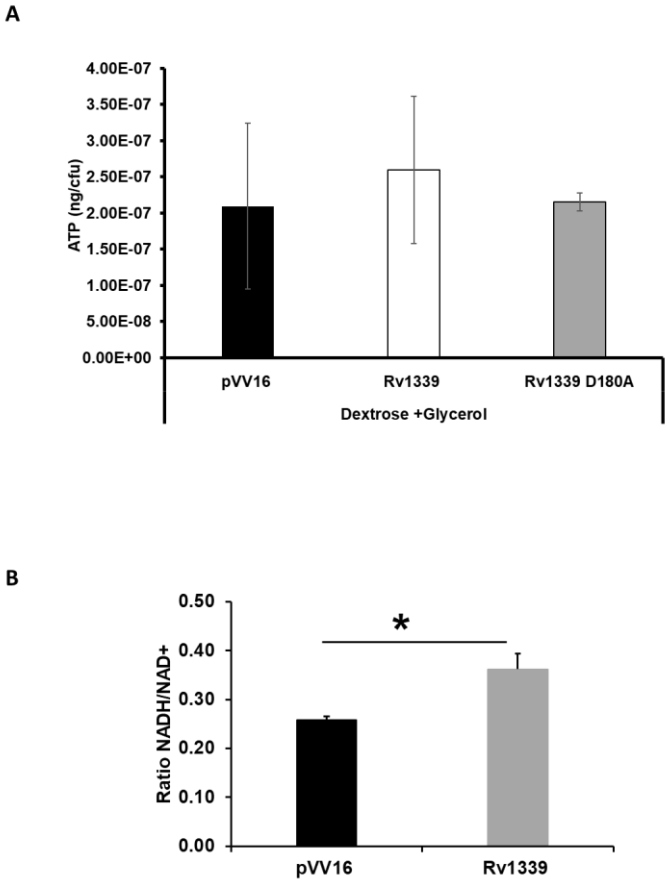


Figure 5.11: ATP levels and the NADH/ NAD+ ratio of Empty Vector Control and Rv1339 expressing bacteria. 9A: ATP levels measured at Mid-log phase of bacterial growth, when bacteria were cultured with Glycerol and Dextrose as carbon sources. 9B: The ratio of NADH/NAD+ levels in Empty Vector Control and Rv1339 expressing M. smegmatis mc^{2}155 strains. p<0.05 as measured by 2-tailed Student's's T-test. Each plot representative of 2 biological replicates in technical quadruplets (A) and triplicates (B).

Furthermore, the NADH to NAD+ ratio was determined (Figure 5.11B). This ratio is an indication of carbon catabolism or the TCA cycle activity. NADH, produced from NAD+, feeds electrons in to the electron transport chain[305,450,451]. The ratio of NADH/NAD+ was significantly increased by 28% ($p<0.05$). This may be indicative of an increase in carbon catabolism. This would reinforce the increased bioenergetics observed in the seahorse data as increased ECAR may be indicative of increased carbon catabolism or increased TCA cycle activities. Considering the increase in OCR, it is unlikely that the increase in the NADH/NAD+ ratio is due to reduced influx of electrons from NAD+/NADH into the ETC.

To investigate carbon catabolism activities, U-^{13}C labelling (of glycerol and dextrose) was again used to measure the turnover of serine and glycine (Figure 5.12). From this analysis, it was clear that Rv1339 expressing bacteria show significantly increased percentage incorporation of label and a different

isotopologue profile (e.g m+3 is increased by 10% in Rv1339 vs. Empty Vector Control expressing bacteria). This means that there is increased turnover of serine. This metabolite has been demonstrated by systems modelling and *in-vitro* validation to be a potential carbon source in *M. tuberculosis* H37Rv as well as *M. bovis* BCG and its parental strain[452]. A further study showed that during *M. tuberculosis* H37Rv macrophage infection, serine production was essential and that it could act as a carbon source to replenish the pyruvate pool, or as a nitrogen backbone for glycine production[453]. The isotopologue distribution of glycine (Figure 5.12B) was not significantly altered, however, the error bars were prohibitively large – potentially due to the low ion counts detected and the resulting variability between replicates. Additionally, the gene *sdaA* showed significant 0.35-fold decreased expression. This gene codes for L-serine ammonia lyase which is an enzyme that can convert serine to pyruvate. Taken together, this suggests that serine turnover is increased, but that glycine is not significantly increased. Concurrently, with the decreased *sdaA* expression, this may indicate a compensation for the reduced enzyme activity of this pathway – by increasing the production of substrate. This suggests that it may be replenishing the pyruvate pool – an alteration that could produce more NADH and ATP to contribute to the increased bioenergetic activity that has already been observed. To follow this up, optimising our chromatography and sample method may allow better detection of other metabolites involved in carbon catabolism, a key contributor to NADH levels and bioenergetic activity (Figure 5.10C).

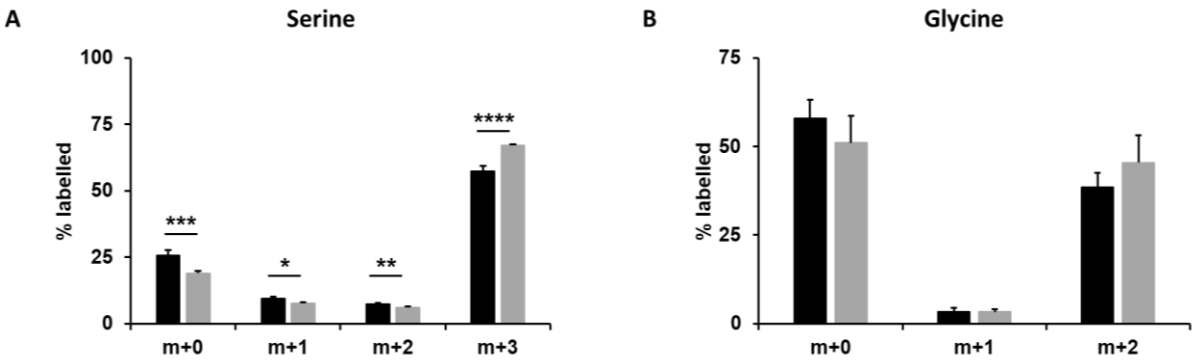


Figure 5.12: U-^{13}C of M. smegmatis mc^{2}155 Empty Vector Control (black bars) and Rv1339 expressing bacteria (grey bars), at a half doubling time after growth with labelled glycerol and dextrose as carbon sources – on 7H10 agar. Each plot is indicative of 2 biological replicates in technical triplicate. Statistical analysis performed with 2-tailed Student's T-test. A: p<.0.05, p<0.01, p<0.001 and p<0.001. B: not significant.

5.5 Untargeted metabolomics analysis reveals that Rv1339 expression alters the metabolome, leading to increased plasticity in peptidoglycan synthesis metabolites

In order to better understand the functional consequences of Rv1339 expression, the subsequent decreased cAMP levels and compromised bioenergetics, LC-MS untargeted metabolomic analysis was performed (Figure 5.13A and 5.13B). This metabolomic workflow would enable an unbiased investigation of altered metabolite levels, and could give an insight in to which processes are affected. The chromatography used for this analysis was the hydrophilic interaction column HILIC-Z. This column and the associated protocols were intended to provide good detection of moderate to highly polar molecules, including carbon metabolism metabolites (such as pentose phosphate pathway metabolites), sugars and ATP.

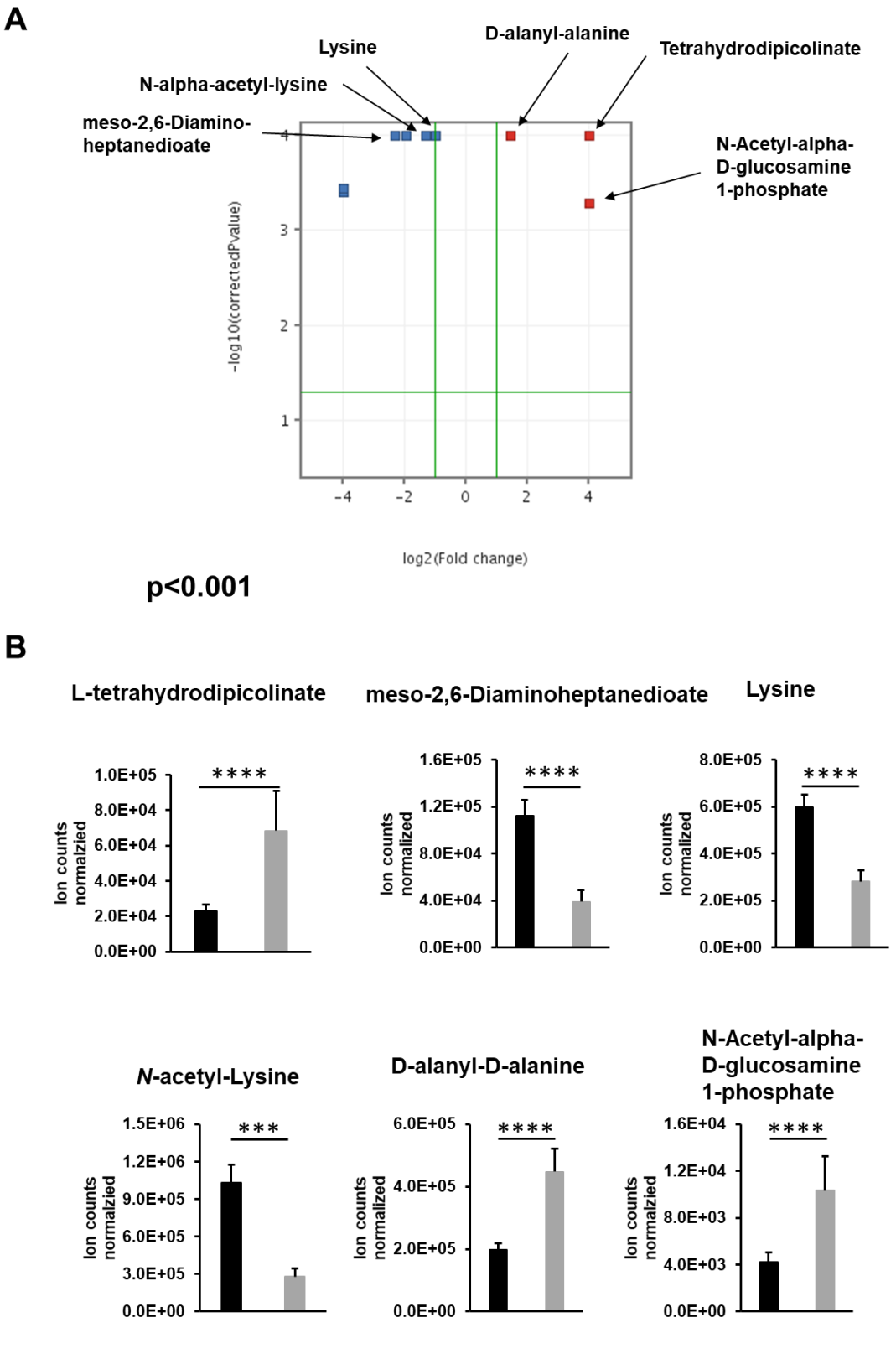


*Figure 5.13: Untargeted HILIC-Z metabolomics analysis of Empty Vector Control (Black) and Rv1339 expressing M. smegmatis mc²155 strains (Grey) – grown in 7H9 supplemented with glycerol and dextrose carbon sources. 10A: Volcano plot of significantly differentially expressed metabolites. 10B: Absolute quantification of several key peptidoglycan synthesis pathway metabolites. p<0.001 = ***, p< 0.0001 = ****, as measured 2-tailed Student's's T-test.*

From this untargeted analysis of the metabolome (Figure 5.13), there were significant alterations in the levels of compounds later confirmed to be meso-2,6-Diaminoheptanedioate (m-DAP – 2.2-fold), lysine (1.2-fold), D-alanyl-D-alanine (1.5-fold), N-acetyl lysine (1-fold), glucosamine-6-phosphate (4-fold) and tetrahydrodipicolinate (4-fold). All of these metabolites are involved in peptidoglycan biosynthesis – an ATP intensive process[434].

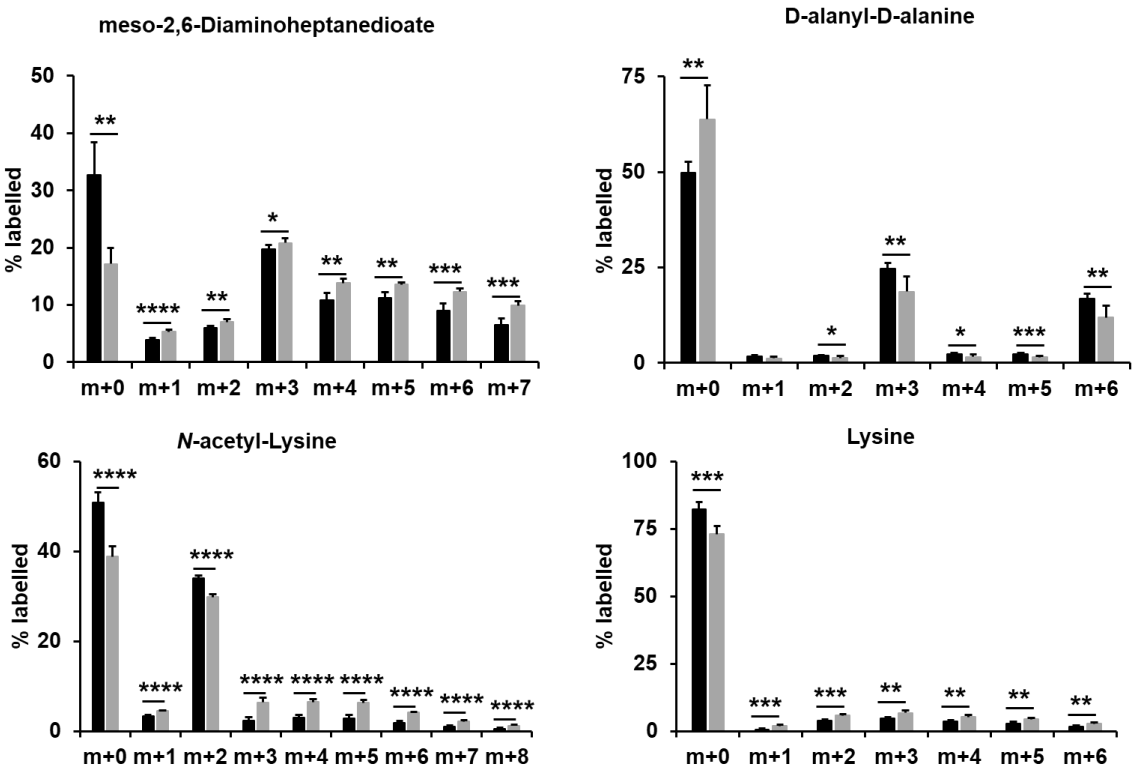


Figure 5.14: Isotopologue distribution of a subset of peptidoglycan synthesis metabolites from ^{13}C label incorporation studies of Empty Vector Control (Black) and Rv1339 expressing M. smegmatis mc^{2}155 (grey) strains. All bacteria were cultured in 7H9 with glycerol and dextrose as carbon source. Each plot is representitve of 2 biological replicates with 3 technical replicates. A Student's t-test was used to determine significance with *p<0.05, **p<0.01, ***p<0.001 and ****p<0.0001.

The metabolome alterations observed in the absolute levels of these metabolites was validated by measuring their turnover via the incorporation of ^{13}C over time (Figure 5.14). From this data there was significant alteration in the turnover m-DAP, D-ala-D-ala, N-acetyl-lysine and lysine. The altered turnover observed for these metabolites and changes in the metabolite pools likely indicates greater plasticity in bacterial peptidoglycan. For example, both Lysine, and m-DAP displayed decreased metabolite pool sizes, despite increased turnover (Figure 5.13 and Figure 5.14). m-DAP is a key component of peptidoglycan and the fact that the absolute levels of m-DAP are decreased, with an increase in turnover suggests that more m-DAP is being produced and then consumed. The fact the lysine metabolite pool size is also decreased, despite increased turnover may suggest that while more

m-DAP is being produced, the muralytic enzyme LysA[454] may be breaking down more m-DAP from peptidoglycan into lysine. This would indicate increased production and degradation peptidoglycan, and therefore increased plasticity. DAP crosslinks provide stability to the cell wall[454], and so increased plasticity may indicate increased instability. Tetrahydrodipicolinate is the precursor of m-DAP, and displayed a 4-fold increase in metabolite pool size – this may be to compensate for the significantly increased turnover of m-DAP. Production of m-DAP has been shown to be essential in mycobacteria, as disruption leads to decreased stability of peptidoglycan[454,455]. Furthermore, D-alanyl-D-alanine (D-Ala-D-Ala) and N-acetyl-glucosamine (GlcNAc) are also key components of peptidoglycan, with GlcNAc representing 1 half of the glycan strand. Both of these metabolites saw increases in their pool sizes of 1.5 and 4-fold respectively. Additionally, the turnover of D-ala-D-ala was decreased, suggesting that the increase in the metabolite pool may not be the result of increased synthesis, but perhaps degradation. Taken together, this further supports the theory that there is increased production and degradation of peptidoglycan. Increased plasticity of peptidoglycan, and decreased stability, would likely lead to consequences for bacterial permeability.

5.6 Rv1339 expression leads to increased permeability

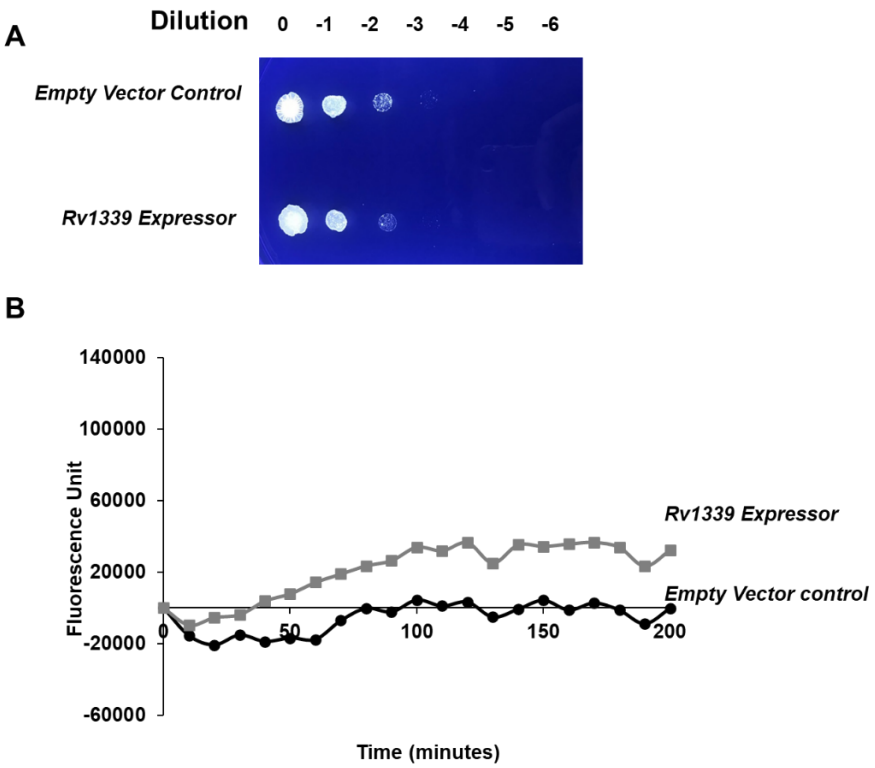


Figure 5.15: Permeability of Empty Vector Control and Rv1339 expressing M. smegmatis mc²155 strains. A: Different dilutions of Empty Vector Control or Rv1339 expressing bacteria spotted on crystal violet containing 7H10 agar (supplemented with glycerol and dextrose carbon sources). B: Hoechst 33342 dye incorporation assay performed on Empty Vector Control and Rv1339 expressing M. smegmatis mc²155 strains in 7H9 liquid culture containing glycerol and dextrose as carbon sources.

The functional consequences of the increased peptidoglycan plasticity were investigated with Crystal violet and Hoechst dye 33342 uptake assays[456] (Figure 5.15A and B). Crystal violet is a toxic dye, and bacteria with higher permeability see greater uptake of this dye[437]. This therefore leads to reduced growth when the equivalent of OD 0.5 mid-log phase bacteria are spotted at 10^{-1} – 10^{-7} dilutions on agar containing 10µg/ml of crystal violet and incubated at 37˚C for 72 hours. Rv1339 expressing bacteria showed decreased growth on this agar when compared to the Empty Vector Control (Figure 5.15A).

This increase in permeability was confirmed by measuring the uptake of a DNA binding dye, Hoechst 33342 in bacteria grown in liquid culture. When this dye binds to DNA intracellularly, its fluorescence wavelength is altered (from a maximum emission of 510-540nm when unbound to 350-461nm when bound). This dye has a good cell permeability profile with a lower toxicity than other nucleic acid staining dyes, like DapI[456]. The Rv1339 expressing bacteria showed increased incorporation of the dye, thereby indicating increased permeability.

5.7 Discussion

Several species of pathogenic mycobacteria have evolved to be intrinsically resistant to antimicrobials, with among other mechanisms, a complex peptidoglycan, arabinogalactan and mycolic acid containing cell wall representing a significant barrier to antimicrobials entry[254]. Mycobacteria have also been shown to display different phenotypic and metabolomic states during infection, which can alter their susceptibility to antimicrobials[42,47,48]. Likewise, mycobacteria like *M. tuberculosis* are able to rewire their bioenergetics and drastically slow growth to enable persistence in response to infection conditions[445,457].

Phenotypic antimicrobial tolerance represents a gateway to resistance as it allows time for pathogens to develop resistance mutations[266]. The development of tolerance or persistence requires the detection of stresses or signals in the environment and phenotypic adaptions to respond. It is already known that response to environmental stimuli in pathogenic bacteria requires rapid, tuneable and metabolically favourable signalling systems – such as cAMP[116,165,243]. The host environment induced a greater effect on the antimicrobial susceptibility transcriptome in *M. tuberculosis* bacteria during an *in-vivo* infection model, than treatment with antimicrobials[49]. In the host niche, *M. tuberculosis* shows decreased antimicrobial susceptibility[47,49]. With the link between cAMP and adaption to environmental cues in the host niche[68,70,216], and the significant effect the environment has been shown to have on the antimicrobial susceptibility transcriptome, it is likely that cAMP is involved in antimicrobial susceptibility.

While the involvement of cAMP with antimicrobials susceptibility in pathogenic bacteria has been previously elucidated in several studies[71,262], our data represents the first time this has been characterised by reducing cAMP levels in mycobacteria – without knocking out any of the synthesis machinery. Decreasing cAMP levels in this way can prove to be a useful tool in investigating this signalling system, with less background than previous approaches in the field (such as forskolin – which does not work on all adenylate cyclases or db-cAMP – which can also act as a PDE inhibitor, as reviewed in the introduction Chapter)[271,308].

Rv1339 expression decreased cAMP levels and increased bacterial susceptibility to antimicrobials – with widely different MOAs. Concomitantly with decreased cAMP levels there was decreased expression of key components of mycobacterial bioenergetics, decreased PMF and increased permeability – probably due to increased peptidoglycan plasticity. I therefore hypothesise that the increased permeability leads to increased levels of the antimicrobial inside the bacteria – and as a result increases their efficacy. Optimising a targeted metabolomics workflow for the antibiotics and their degradation products would enable quantification of antibiotic levels inside the bacteria, and this may be a way to prove this hypothesis. Alternatively, quantification of fluorescence within the bacteria, when they are treated with fluorophore-labelled antibiotics or their derivatives would also confirm or

refute this hypothesis. There has already been success with this approach for MOA studies with antibiotics like vancomycin and linezolid[458]. Alternatively, antibiotics with distinctive structural components can be imaged using a combination of mass-spectrometry and microscopy approaches – and this could enable quantification of intracellular antibiotic concentrations (e.g. bedaquiline with $^{79}Br_{35}$[459]). Bedaquiline targets mycobacterial ATP synthase, causing progressive depletion of ATP that eventually compromises bacterial viability[240,460]. Initially, bedaquiline is bacteriostatic until bacterial metabolic adaptions are overcome, at which point it displays bactericidal action[240,448,460,461]. This would make it especially interesting to use in this case, as my data shows an impact on bacterial bioenergetics. However, the time constraints of my PhD did not allow me to optimise the TTK assays with bedaquiline and coupled with the limited access to appropriate microscopy facilities, the approach of using bedaquiline and antibiotic imaging was was not deemed feasible.

Despite observing significant increases in the OCR and ECAR of Rv1339 expressing bacteria, there was no significant increase in ATP levels. While this was unexpected, other studies have shown that even without significant changes in absolute ATP levels, bacterial metabolism and bioenergetics can still be significantly altered – as was observed here[462]. Previous studies in *M. smegmatis* mc^2155 have shown that cAMP levels vary over the course of bacterial growth, peaking in the lag phase and rapidly declining in the log and stationery phases[204]. At the mid-log phase, cAMP levels were already predicted to be at half the maximal level seen in the lag phase[204]. If these experiments were repeated in the lag phase, it is likely that there would be greater phenotype – as there would be more demand for cAMP and subsequently a greater effect on the ATP pool. Furthermore, I observed significantly decreased expression of 33 genes encoding ATP binding proteins – including numerous ABC transporters, a 30% decrease in growth and increased plasticity in peptidoglycan synthesis. The proteins encoded by these genes, growth and peptidoglycan biosynthesis are ATP dependent[434,463,464]. With increased ETC activity, TCA and/or carbon catabolism, I would expect to see increased ATP levels – but they were not. In sum, it therefore appears that by forcibly decreasing cAMP levels, even despite increased synthesis, that this is enough to put stress on the ATP pool. I hypothesise that in order to maintain the essential membrane potential and bioenergetic processes e.g. with Sdh1, that the bacteria are forced to commit a deleterious quantity of resources towards maintenance of cAMP levels. There is already a precedence for this hypothesis as during macrophage infection, *M. tuberculosis* and *M. bovis* BCG bacteria increase their cAMP levels 50-fold[174]. With cAMP levels already innately high in *M. tuberculosis* (more than 10-fold increased vs. E. coli[174,175,288]), it is likely that such a sudden increase in cAMP levels would increase demand on the ATP pool. This hypothesis is summarised in Figure 5.16.

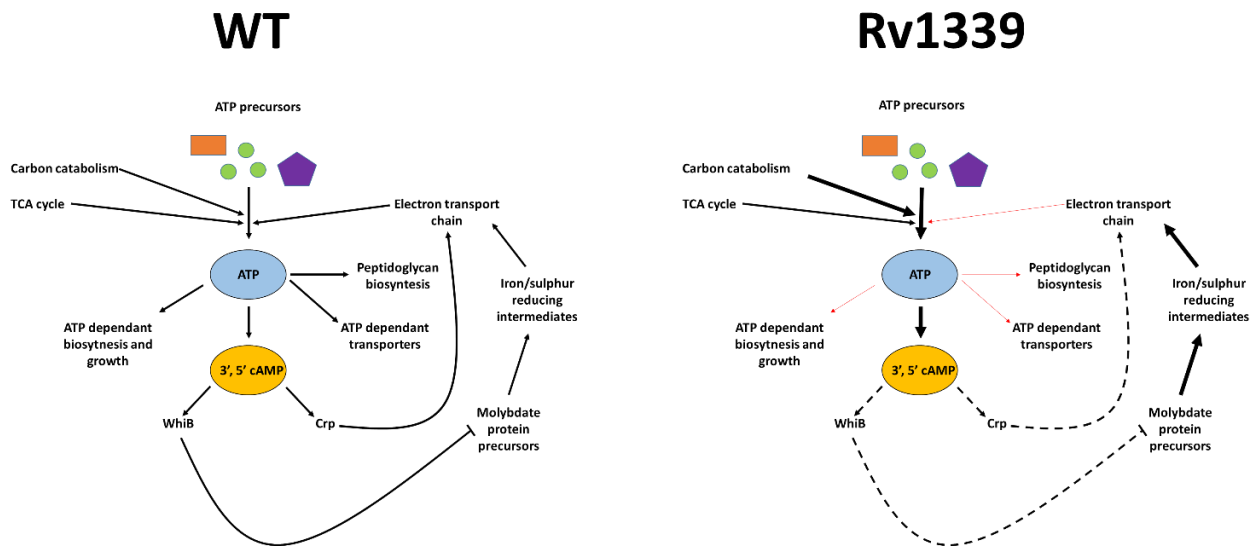


Figure 5.16: Proposed model of Rv1339 expression, and the effect of decreased cAMP levels. Forcibly decreasing cAMP levels leads to compensatory increases in the production of cAMP. These compensations are insufficient to restore cAMP to WT levels, but instead maintain cAMP levels at around 25%. Decreased cAMP alters the expression of key bioenergetic genes such as sdh1 and NADH dehydrogenases. This leads to compromised PMF and electron transport chain activities, despite significant increases in ECAR and OCR. The strain this places on the ATP pool leads to downregulation of ATP dependant processes, such as adequate peptidoglycan biosynthesis. This leads to increased permeability and influx of antimicrobials. Red lines indicate decreased input. Thicker lines indicate increased input. Dashed lines ended with arrows indicate processes that are inhibited to some extent. Dashed lines ended with a perpendicular dash indicate removal of repression.

Interestingly, there is more circumstantial evidence for the hypothesis that cAMP levels impact upon the ATP pool in the literature. Single point mutations in Rv1339 have been found to confer resistance to the entire class of Imidazopyridines (IMP) – ATP depleting compounds that are structurally related to Q203[465]. Characterisation of the IMPs suggests that they target QcrB, a proton pumping component of the terminal cytochrome oxidase (complex III of the ETC). *M. tuberculosis* bacteria treated with the IMPs saw dose dependent decreases in ATP levels and disruption of intrabacterial pH. Mutations in QcrB also led to resistance[465]. Of the single point mutations reported in Rv1339, all cluster around the metal binding site in Rv1339 or are predicted to change the structure[465]. If these mutations ablated the activity of Rv1339 and decreased demand on the ATP pool, this could compensate for the effects of ATP depletion.

Maintenance of sufficient ATP levels is essential for *M. tuberculosis* during infection[457,466] and requires a carbon source[284]. Cholesterol is a key carbon source of *M. tuberculosis* isolates during host infection[283,284]. Successful infection requires the ability to metabolise cholesterol and detoxify the

downstream products, such as propinyl-CoA[250,251]. Detoxification of the cholesterol catabolism product propinoyl-CoA requires expression of genes known to be regulated by cAMP[251]. Additionally, a screen of mycobacterial cholesterol targeting small molecules identified several promising compounds with significant bactericidal efficacy against intracellular *M. tuberculosis* bacteria during macrophage infection [11]. While the MOA was not fully characterised, treatment with the compounds led to drastically increased cAMP levels via activation of a bacterial AC, Rv1625 [11,215]. Transposon insertions in the open reading frame of Rv1625 ablated this increase and conferred resistance to the compounds[215]. The fact that cAMP appears important to both cholesterol metabolism and downstream detoxification, makes cAMP signalling a tempting target, as it would be potentially more difficult for the bacteria to evolve away from.

The link between cAMP and permeability has previously been postulated as an explanation of why perturbation of this signalling pathway leads to altered antimicrobial susceptibility in *Salmonella spp*[261,262] and *E. coli*[263]. In the aforementioned studies, perturbation of cAMP signalling led to decreased expression of nutrient transporters that were also thought to uptake antibiotics. The benefit of cAMP regulated nutrient transporters was observed when *Salmonella typhimurium* bacteria were cocultured with other bacterial species (to mimic the host intestinal environment)[264]. In this highly competitive environment, *S. typhimurium* bacteria expressing Crp_{ST} thrived whereas growth of ΔCrp_{ST} strains was significantly compromised[264]. This competitive advantage was not observed when the bacteria were cultured in nutrient poor media – thereby suggesting that in nutrient rich conditions, cAMP positively regulates uptake[264]. Unfortunately, these studies were performed in *gram-negative* bacteria, which have decreased prominence of peptidoglycan when compared to mycobacteria[434]. The link between cAMP, permeability and peptidoglycan may therefore be more complex in mycobacteria. Repeating the MIC_{90} and TTK experiments with peptidoglycan cell wall component targeting antibiotics (e.g. D-cycloserine or vancomycin) may help with further characterisation of this link. I would expect that treatment with these antibiotics would show increased efficacy, due to the potential alterations in peptidoglycan crosslinking suggested in this study. My data supports a link between cAMP and permeability in *M. smegma*tis mc^2155 and therefore this may represent a unifying mechanism conserved across other bacterial species. This definitely warrants investigation or re-investigation of the importance of cAMP in other bacterial species – to further elucidate the role cAMP plays in permeability and antimicrobial susceptibility.

Taken together, the data in this Chapter suggests that targeting cAMP signalling could lead to impairment of bacterial bioenergetics and viability during infection through multiple mechanisms. It also shows circumstantial evidence that decreasing cAMP levels can interfere with the development of resistance to streptomycin (though genome sequencing and further investigation are required for confirmation). It also shows that decreasing cAMP increases the efficacy of ethambutol and potentially rifampicin. There is a strong case for further investigation of the validity of targeting this signalling system and doing so may provide solutions to increase susceptibility of bacteria to antimicrobials treatment – or to block the development of resistance.

5.8 Preliminary work to validate the findings from *M. smegmatis* mc^{2}155 in *M. tuberculosis* H37Rv

An original goal of my project, time-permitting, was to validate any findings from *M. smegmatis* mc^{2}155 in *M. tuberculosis* H37Rv. Although the timeframe of my PhD proved too short to fully complete this objective, I was able to begin preliminary work in *M. tuberculosis* H37Rv. While there is still a large volume of work to be done to take these findings forward, I will briefly describe them here, and discuss any potential future avenues.

5.8.1 Transformation and potential expression in *M. tuberculosis* H37Rv parental, ΔCrp$_{MT}$ and ΔRv0805 with pVV16::*rv1339* full-length

3 *M. tuberculosis* H37Rv strains: Parental, ΔCrp$_{MT}$ and ΔRv0805 were transformed via electroporation with pVV16 Empty Vector Control and Rv1339 Full length containing constructs. The bacteria were then plated on 7H10 agar supplemented with Bovine serum albumin, glycerol, dextrose and with kanamycin/hygromycin for selection of successful transformants. The plates were incubated at 37˚C for 30 days to allow for colony formation. These strains were chosen to fulfil several goals:

1. To investigate if Rv1339 could complement the role of Rv0805 in the ΔRv0805 strain
2. To investigate the effect of lowering cAMP in the ΔCrp$_{MT}$ strains
3. To investigate whether Rv1339 was capable of reducing 3', 5' – cAMP levels in *M. tuberculosis* H37Rv

To fulfil these goals, confirmation of expression by Western blot analysis, quantification of metabolome alterations and growth phenotypes was performed.

To begin, when the successfully transformed (colony forming) bacteria were cultured to the mid-log phase and lysed to extract the clear soluble lysate, Western blot analysis with α-His tag antibody was performed. Unfortunately, no bands were found in all conditions, except in the ΔRv0805 Rv1339 condition (Data not shown). This was surprising, as concurrently with the preparation for Western blot analysis, there appeared to be an interesting growth phenotype in the ΔCrp$_{MT}$ and ΔCrp$_{MT}$ Rv1339 conditions, although there appeared to be no difference in growth in the ΔRv0805 conditions (Figure 5.17). The ΔCrp$_{MT}$, strain displays a major growth defect when compared to the parental strain[70,177,467]. Indeed, the ΔCrp$_{MT}$ pVV16 Control displayed a growth defect (Figure 5.17B). Strikingly, when Rv1339 was expressed in the ΔCrp$_{MT}$ background, it appeared to complement growth back to H37Rv parental strain levels (5.17B as compared to 5.17A). In contrast, ΔRv0805 did not display alterations in growth when compared to the H37Rv parental strain, or when Rv1339 was expressed in this background (Figure 5.17C). With the knowledge that Rv1339 displayed cAMP PDE activity in *M. smegmatis* mc^{2}155, my next step was to quantify cAMP levels in these strains. Additionally, I wanted to confirm that Rv1339 was being expressed before forming a hypothesis.

Even though there were no bands by Western blot of the clear soluble lysate, the positive transformation control (pVV16 Empty Vector) was successfully transformed in all conditions, and the negative control (double distilled H_2O) was not. With these results in mind, I started cultures of all strains in order to extract RNA for qRT-PCR analysis of Rv1339 gene expression (using the MP FASTRNA pro Blue kit). In tandem, I transformed *M. tuberculosis* H37Rv parental, ΔRv0805 and ΔCrp_{MT} strains with the pVV16::Rv1339L D180A activity mutant and awaited colonies. This was to ensure adequate controls for further investigation of the interesting growth phenotype.

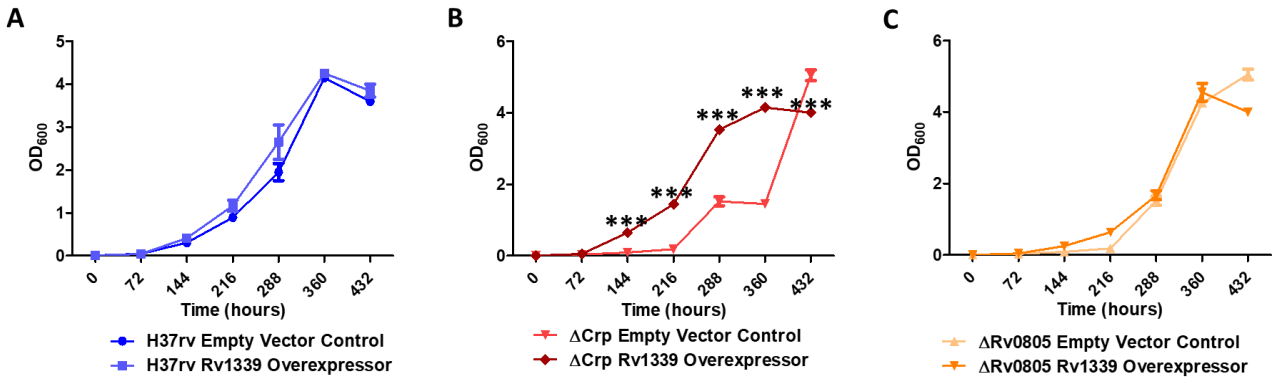


Figure 5.17: Growth curve of H37Rv (A), ΔCrp_{MT} (B) and ΔRv0805 strains expression the Empty Vector Control (pVV16) or Rv1339. Each point is representative of 2 technical replicates. Two-way ANOVA with Bonferroni multiple comparison tests determined p< 0.001 for ΔCrp_{MT} strains. All strains were produced from the H37Rv background and measurements were taken as part of the same experiment.

5.8.2 Rv1339 expression decreases cAMP levels in the *M. tuberculosis* H37Rv parental strain, displays no effect on cAMP levels in the ΔRv0805 strain and increase cAMP levels in the ΔCrp_{MT} strain

While awaiting, the Rv1339 D180A colonies and the growth of cultures to extract RNA from all current conditions, I performed targeted metabolomics analysis on the Empty Vector Control and Rv1339L expressor strains. This was in order to investigate the effect of Rv1339 expression on cAMP levels (Figure 5.18).

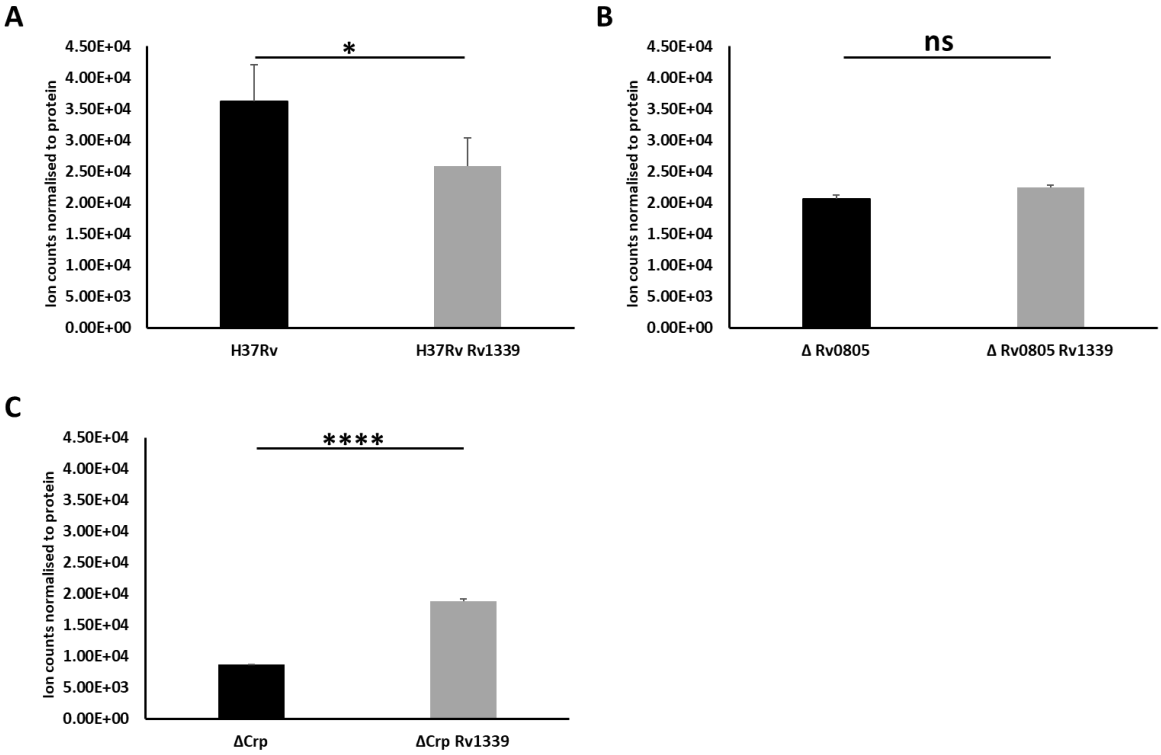


Figure 5.18: Targeted LC-MS analysis of cAMP levels in the H37Rv (A), ΔRv0805 (B) and ΔCrp_{MT} (C) strains expressing the Empty Vector Control (pVV16) or Rv1339. 2-tailed Student's T-test was used to investigate significance of the altered metabolite levels. A: p<0.05. C: p<0.0001.

From this data, it was clear that Rv1339 expression in the *M. tuberculosis* H37Rv parental strain led to a 30% decrease in cAMP levels (Figure 5.18A). This decrease was not observed in the ΔRv0805 strain, which appeared to display similar cAMP levels, even when Rv1339 was expressed in this background (Figure 5.18B). Surprisingly, the ΔCrp_{MT} strain showed 100% increase in cAMP levels, upon Rv1339 expression (5.18C). It is also worth noting that cAMP levels in the ΔRv0805 and ΔCrp_{MT} pVV16 Empty Vector Control expressing strains were 70% or 25% of the *M. tuberculosis* H37Rv pVV16 Empty Vector Control respectively. (Figures 5.18B and 5.18C compared to 5.18A). This combination of findings seemed unusual, as the knocking-out of Rv0805 (a cAMP degrading enzyme) or Crp, lowered cAMP levels. However, the decrease in cAMP levels observed in H37Rv (Figure 5.18A) was in-keeping with the literature reports on Rv0805 overexpression mediating a similar decrease[308]. In sum, it appeared perturbation of cAMP signalling components Rv0805 or Crp_{MT} had the effect of decreasing cAMP.

Meanwhile, overexpression of Rv1339 in the parental strain (H37Rv) led to a 30% decrease in cAMP levels. The contrast between the Rv1339 overexpression, and the knockout backgrounds seems to imply that disruption of the signalling pathway can have profound effects on cAMP levels. It also implies that Rv1339 cannot complement all the activities of Rv0805, although to test this, Rv1339 must be expressed under the native Rv0805 promotor (as opposed to the HSP-60 promotor in pVV16). It is possible that there is a high degree of co-regulation, to maintain cAMP homeostasis. This would most likely be mediated at the synthesis level, as *M. tuberculosis* H37Rv possesses 16 cAMP producing AC enzymes[68]. Although there have been no reports of AC enzymes being regulated by Crp, Rv0805 has itself been found to possess a strong Crp_{MT} binding recognition site and this may indicate it is regulated by Crp_{MT} (although no functional validation was performed)[210].

5.8.3 Rv1339 expression leads to significant alterations in the metabolome of *M. tuberculosis* H37Rv parental, ΔRv0805 and ΔCrp_{MT} strains

In order to gain an insight into the complicated picture of cAMP regulation and the effect of modulating cAMP on metabolism, all 3 strains expressing the Empty Vector Control (pVV16) and Rv1339 were prepared for LC-MS metabolomic analysis. *M. tuberculosis* H37Rv transformed with Rv1339 vs. Empty Vector Control showed the most similarity to my data from *M. smegmatis* mc^2155. However, the magnitude of the reduction in cAMP was already seen to be less pronounced (Figure 5.17A). This may be due to different localisation of the protein, different cAMP production capabilities of the different mycobacteria or due to expression under a non-native promotor (expressing Rv0805 under its native promotor increased activity in overexpression studies[308]).

A key difference between the *M. tuberculosis* H37Rv Empty Vector Control and Rv1339 expressing bacteria was the significant 20% decrease in pyruvate and methyl-citrate metabolite pool levels (Figure 5.19A and 5.19B). Pyruvate is a key component of central carbon metabolism, representing a carbon catabolism product that feeds into the TCA cycle[256]. Likewise, methyl-citrate is also a key central carbon metabolite, from the eponymous methyl-citrate cycle[256]. This cycle takes catabolism products from cholesterol or odd-chain fatty acids and converts them in to pyruvate, or succinate (via methyl-citrate)[256]. By decreasing cAMP levels with Rv1339 expression, it appears as though activity of the methyl-citrate cycle is impaired to some extent. This could be explained by the cAMP dependant regulation of Rv0998 – a lysine acetyltransferase[255,257]. Rv0998 that has been shown to acetylate lysine residues on isocitrate lysase (ICL) and modulate its activity[257]. ICL converts methyl-isocitrate into pyruvate, as well as isocitrate to glyoxylate (in the glyoxylate shunt – although we could not detect these metabolites well with the chosen chromatography)[256,468]. It is therefore possible that decreasing cAMP levels decreases the positive regulation activity by Rv0998 acetylation of ICL. In order to confirm this, quantitative proteomic profiling of lysine acetylation would be required. This would

involve extraction of bacterial proteins, trypsin digestion and preparation for MS analysis (similar to approaches in the literature[255,258].

In the ΔCrp_{MT} background, expression of Rv1339 led to a significant 20% and 100% increase in the metabolite pool levels of pyruvate and methyl-citrate respectively (Figure 5.19C and 5.19D). This may suggest that in the absence of Crp_{MT}, Rv1339 expression has the effect of increasing methyl-citrate cycle activity. When the ΔCrp_{MT} pVV16 Control and ΔCrp_{MT} Rv1339 were compared to the H37Rv parental strain, pyruvate and methyl-citrate levels were already 34% and 44% decreased (Figure 5.19E and 19F). This implies that expression of Rv1339 in an already dysregulated system can reverse some of the effects of Crp_{MT} knockout. It is worth noting that Rv1339 expression in the ΔCrp_{MT} background returns pyruvate and methyl-citrate levels to H37Rv parental Empty Vector Control levels. This may form part of the explanation behind the ΔCrp_{MT} Rv1339 strain displaying comparable growth to the H37Rv parental strain (Empty Vector Control). However, the underlying mechanisms will be difficult to disentangle as I observe a 100% increase in cAMP levels when Rv1339 is expressed in the ΔCrp_{MT} background – as opposed to the expected decrease of overexpressing a cAMP degrading enzyme. A transcriptome analysis of each of these strains similar to the approach I used in *M. smegmatis* mc^2155 may provide some insights, although cAMP signalling is likely more complex in the pathogenic *M. tuberculosis* H37Rv. This can be evidenced by multiple signalling components with transcriptional regulatory activity (e.g. Cmr[206], Crp[237], Rv0805[308(p20)] and WhiB transcription factors[441]).

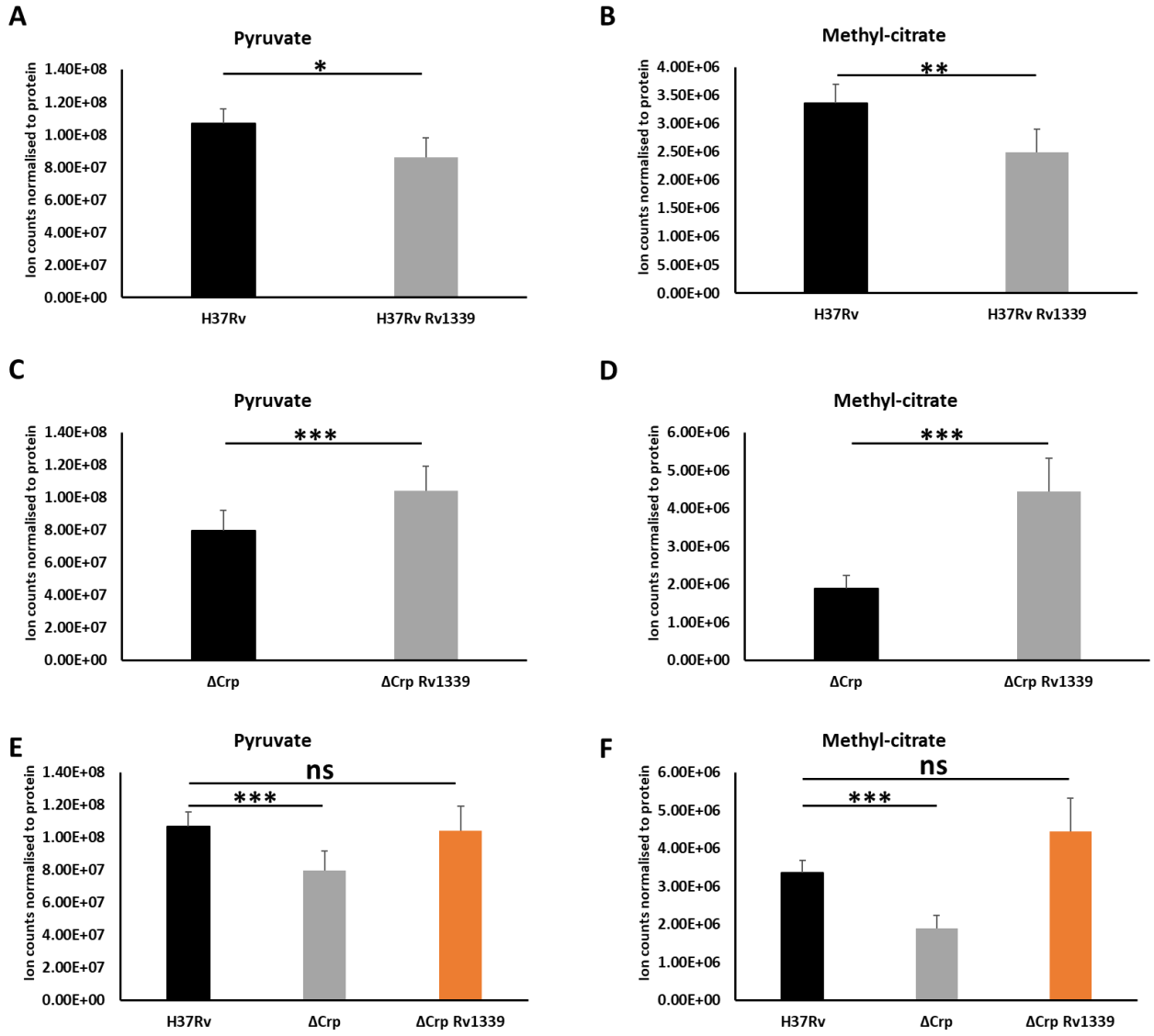


Figure 5.19: Targeted LC-MS analysis of key central carbon metabolites in H37Rv (A and B) or ΔCrp strains (C and D) expressing Rv1339 or Empty Vector Control (pVV16). E and F: comparison of pyruvate (E) and methyl-citrate levels in H37Rv and ΔCrp$_{MT}$ strains expressing Rv1339 or the Empty Vector Control (pVV16). two-tailed Student's's T-test used for statistical analysis p<0.05 (A), p<0.01 (B) and p<0.001 (C, D, E and F).

Rv1339 expression in the ΔRv0805 background led to relatively few changes when compared to H37rv parental and ΔCrp$_{MT}$. However, the levels of succinate, diaminopimelate and D-ala-D-ala were all significantly decreased (by 25%, 50% and 50% respectively) (Figure 5.20). Although the levels of cAMP remain the same in the ΔRv0805 background despite Rv1339 expression, they are still decreased by 50% relative to the H37Rv parental (Empty Vector Control) strain (Figure 5.18A and 5.18B). It is possible that cAMP levels cannot be lowered any further in the presence of Crp$_{MT}$ (which is still present in the ΔRv0805 strain). This may explain why there appears to be no additional decrease of Rv1339 expression. Alternatively, Rv0805 has been shown to be membrane bound and have a cAMP hydrolysis-independent role in regulating permeability[281,308]. Although this role was not elucidated in the original study[281], it is conceivable this could be mediated by alterations in peptidoglycan – and this could explain

the alterations I observe in diaminopimelate acid and D-ala-D-ala (2 key components of peptidoglycan). Unfortunately, there was no evidence in any of the Rv0805 studies showing a connection to peptidoglycan synthesis, and the authors of the permeability study do not speculate[281]. When compared to the H37Rv parental (Empty Vector Control) stain, the levels of succinate are not significantly altered in the ΔRv0805 background (Empty Vector Control) (Figure 5.20C). However, the levels of diaminopimelate and D-ala-D-ala were increased by around 40% when compared to the H37Rv parental (Empty Vector Control) strain. This indicates that deletion of Rv0805 may compromise peptidoglycan biosynthesis and this could lead to the increased permeability previously observed[281]. Interestingly, when Rv1339 is expressed in the ΔRv0805 background (Empty Vector Control) strain, the levels of diaminopimelate and D-ala-D-ala decrease by 60% and 50% respectively. This suggests that the activity of Rv1339 further alters peptidoglycan biosynthesis. When taken with my *M. smegmatis* mc^{2}155 data, this may indicate a role for cAMP dependent and cAMP independent regulation mechanisms over peptidoglycan synthesis or permeability. Again, the complicating factor is that I do not observe further decrease of cAMP levels in the ΔRv0805 background expressing Rv1339. In order to investigate whether cAMP activity is still being altered despite no change in the absolute levels in the metabolite pool, I would need to carry out stable isotope labelling and metabolite turnover studies. If there was increased turnover of cAMP, it would support the hypothesis that Rv1339 is still working to degrade cAMP, but that in the dysregulated system there were mechanisms acting to compensate for this activity e.g. increased cAMP production from adenylate cyclase enzymes.

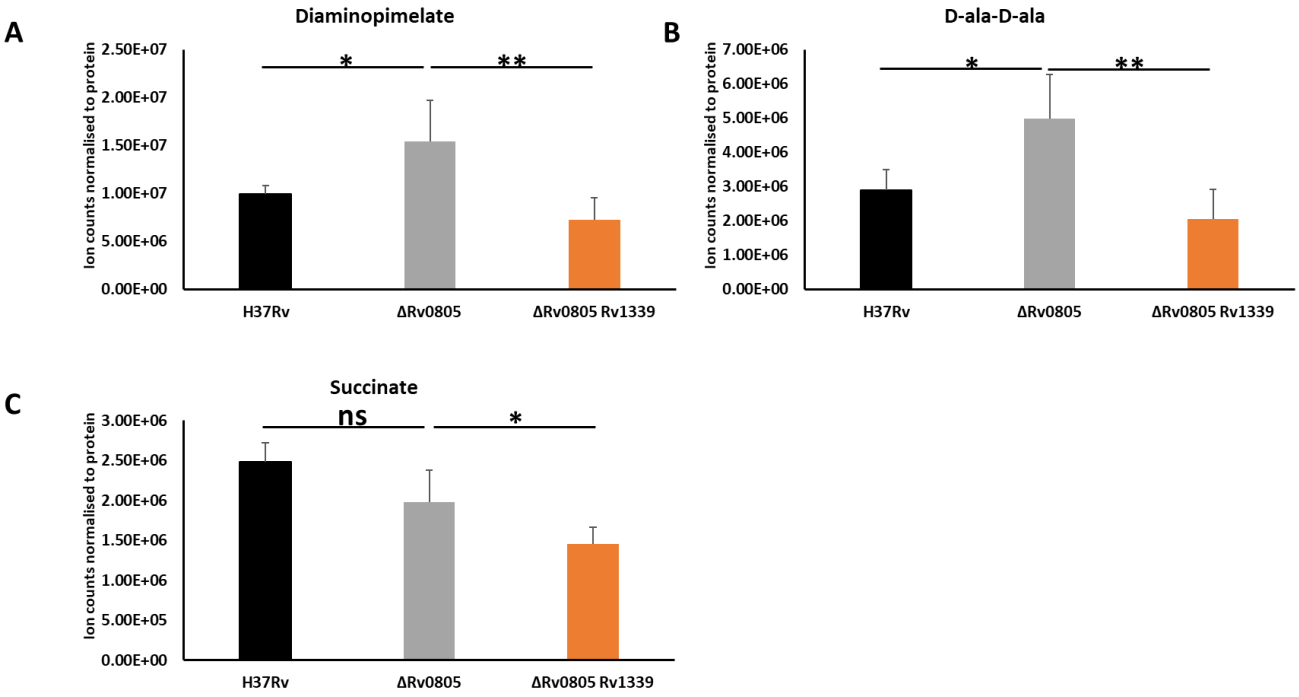


Figure 5.20: Targeted LC-MS analysis of the levels of diaminopimelate (A), D-ala-D-ala (B) and succinate in H37Rv parental and ΔRv0805 strains expressing the Empty Vector Control or Rv1339. Statistical analysis was performed with 2-tailed Student's's T-Test. p<0.05 and p<0.01.

When Rv1339 was expressed in the H37Rv parental strain (Empty Vector Control), there were also significant decreases in the levels of glutamate (22%) ornithine and citrulline (38% for both). Additionally, the levels of arginine were increased by 82% (Figure 5.21A-D). Such alterations were not observed in the ΔRv0805 or ΔCrp_{MT} conditions. A recent study showed that arginine starvation led to rapid sterilisation of *M. tuberculosis* H37Rv in an infection model with antimicrobials treatment[469]. Although not elaborated upon, the study also noted that the levels of cAMP were increased by 150-fold under arginine starvation[469]. When taken together with my data, this suggests there may be a link between cAMP signalling an arginine metabolism. Glutamate, the first metabolite in the de-novo arginine synthesis pathway[470], can also enter the TCA cycle via Rv3722c mediated conversion to α-ketogluterate[471]. However, from my data, the levels of α-ketoglutarate were not significantly altered upon Rv1339 expression (Figure 5.21F). This likely suggests an alteration in the production of arginine. However, without performing metabolite turnover analysis using stable isotope, it is not possible to confirm this.

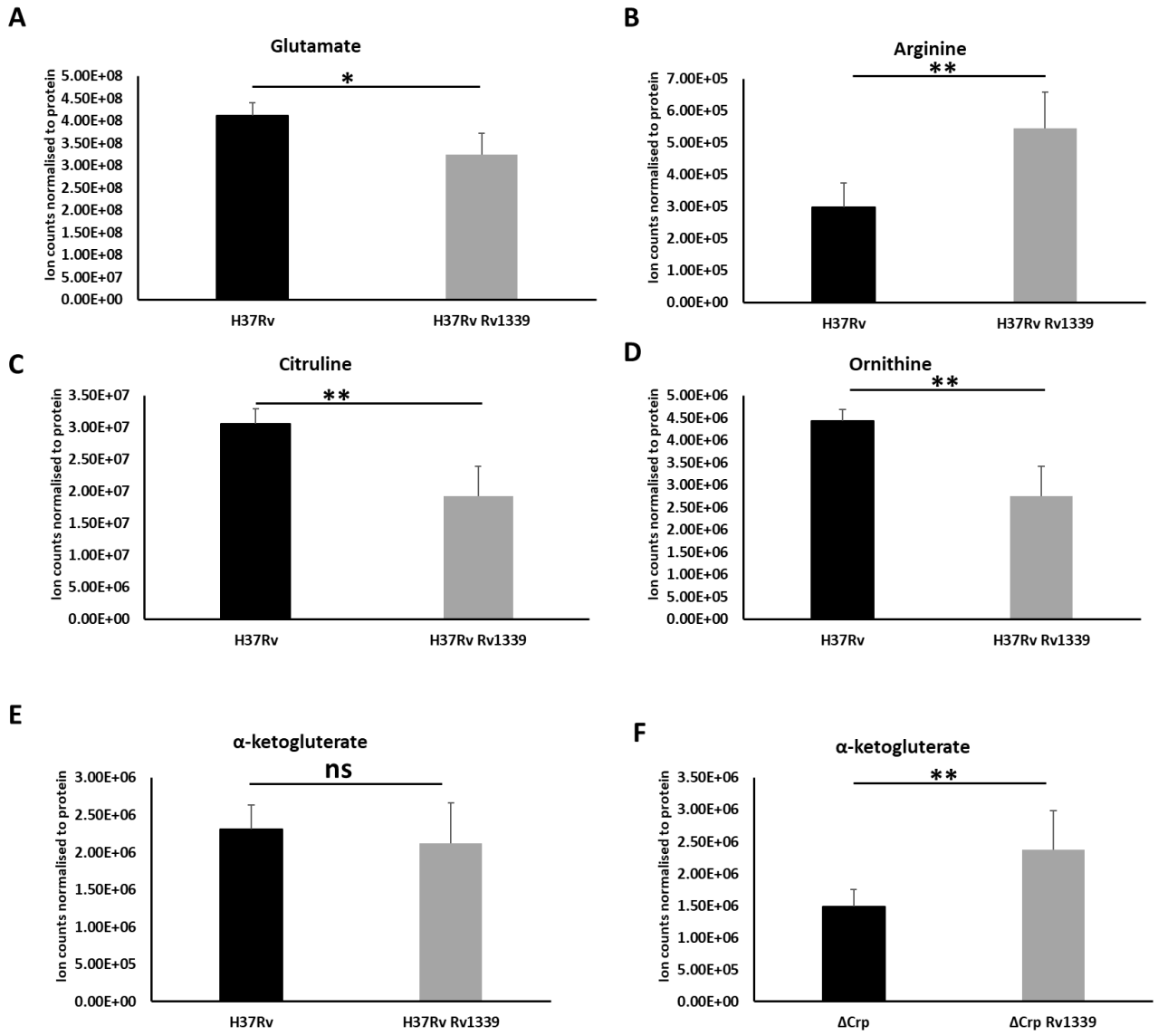


Figure 5.21: Targeted LC-MS analysis of the levels of metabolites related to arginine synthesis in M. tuberculosis H37Rv or ΔCrp$_{MT}$ strains expressing the Empty Vector Control (pVV16) or Rv1339. Statistical analysis performed with 2-tailed Student's's T-test. p<0.05 and p<0.01.

As previously mentioned, the M. *tuberculosis* H37Rv ΔCrp$_{MT}$ mutant displays a growth defect that was attributed to impaired expression of *serC*, the gene encoding a phosphoserine aminotransferase[467]. The data I acquired supports the hypothesis that serine production is impaired in the ΔCrp$_{MT}$ strain when compared to the H37Rv parental (Empty Vector Control strains) – there is a significant 39% decrease in serine levels in ΔCrp$_{MT}$ background (Figure 5.22). Serine production has been shown to be essential during infection conditions and can act as both a nitrogen or carbon source[453]. The attenuated virulence of the ΔCrp$_{MT}$ strain[177] might also be attributable to impairment of serine synthesis, as *M. tuberculosis* H37Rv bacteria cannot attain serine from the host phagosome[452,453].

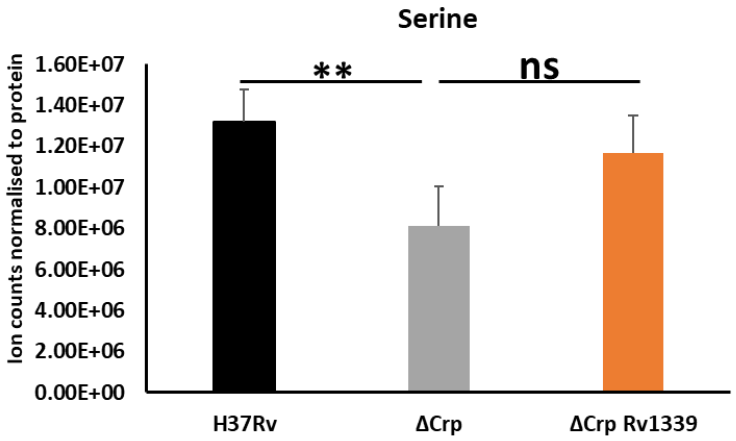


Figure 5.22: Targeted LC-MS analysis of Serine levels in H37Rv parental and ΔCrp_{MT} Empty Vector Control or ΔCrp_{MT} Rv1339 expressing strains. Statistical analysis performed with two-tailed Student's's T-test. p<0.01.

5.9 Rv1339 expression is sufficient to ablate the growth defect of the *M. tuberculosis* H37Rv ΔCrp background

Finally, to confirm the ability of Rv1339 expression to restore the ΔCrp_{MT} strain to the *M. tuberculosis* H37Rv parental level of growth, the growth curve of the H37Rv parental and ΔCrp_{MT} strains expressing the Empty Vector Control (pVV16) and Rv1339 was repeated. In this instance, Rv1339 D180A was also used to confirm if the restoration of growth was due to cAMP PDE activity (Figure 5.23).

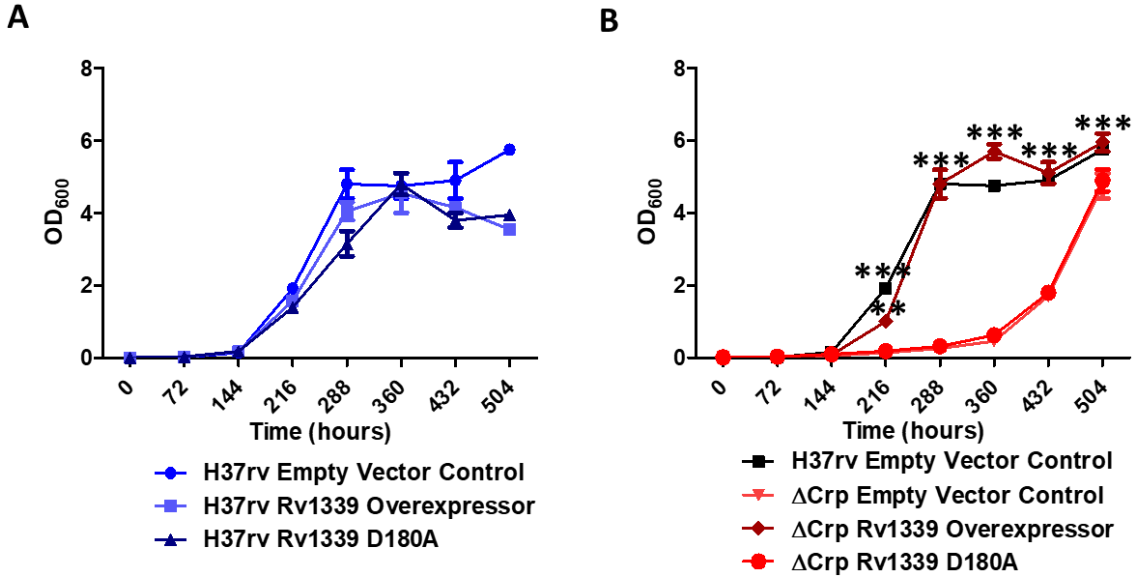


Figure 5.23: Growth curve of H37Rv (A) or ΔCrp (B) strains expressing the Empty Vector Control (pVV16), Rv1339 or Rv1339 D180A. Each point is representative of 2 technical replicates. Two-way ANOVA with Bonferroni multiple comparison tests determined p< 0.01 and p<0.001 when all other strains are compared to ΔCrp.

It was clear from the growth curve that expression of Rv1339 in the ΔCrp_{MT} strain was sufficient to restore growth to the H37Rv parental (Empty Vector Control) strain (Figure 5.23B). This restoration was not seen when Rv1339 D180A was expressed in the ΔCrp_{MT} background (Figure 5.23B) and no differences were observed in the H37Rv parental strains under any condition (Figure 5.23A). This suggests that although the absolute levels of cAMP are not decreased in the ΔCrp_{MT} strain, that the restoration-of-growth phenotype depends on the atypical class-II PDE domain in Rv1339. *In-vitro* and *in-vivo* I have previously confirmed that the D180A mutation in Rv1339 ablates cAMP PDE activity. This supports the hypothesis that the cAMP PDE activity of Rv1339 is having an effect in the bacteria, but that other mechanisms are compensating for the reduction of cAMP levels. Again, in order to confirm this, stable isotope tracing mediated investigation of cAMP turnover would be essential. Furthermore, qRT-PCR of *serC* and other genes involved in serine metabolism would be essential (e.g. *sdaA* - L-serine ammonia lyayse).

Unfortunately, due to the time constraints of my PhD, I was unable to continue the *M. tuberculosis* H37Rv validation any further. The most essential step to following up this work is qRT-PCR to confirm Rv1339 is expressed in the bacterial strains. However, the data I was able to acquire is highly interesting and suggests investigation of the effect of Rv1339 expression in these strains during macrophage infection would be very valuable. For example, succinate is essential for maintaining the membrane potential under hypoxic conditions during *M. tuberculosis* H37Rv infection[238] as it can act as an energy store to maintain electron transport chain activity[238]. Likewise, my *M. smegmatis* mc^{2}155 data suggests that even if we do not observe a decrease in the absolute levels of cAMP in the ΔRv0805 strain in response to Rv1339 expression, there may still be an increase in turnover. This may place greater strain on the ATP pool, and lead to the compromised bioenergetics observed in the *M. smegmatis* mc^{2}155 data. I have already seen that Rv1339 expression in the ΔRv0805 background leads to further decrease in the levels of the peptidoglycan components m-DAP and D-ala-D-ala, during an infection setting this is likely to be deleterious.

Furthermore, it would be very interesting to see if Rv1339 mediated rescue of the ΔCrp_{MT} growth defect, serine, methyl-citrate and pyruvate levels correlates to restoration of virulence during macrophage infection. In order to understand the dysregulation in the cAMP signalling system in this background, transcriptome and metabolite turnover analyses would also be essential.

In the *M. smegmatis* mc^{2}155 RNA-seq data (Figure 5.8) it was clear that there was significantly altered expression of many genes encoding transporters in response to Rv1339 expression and the associated decrease in cAMP levels. It would therefore be very interesting to see if the H37Rv parental strain expressing Rv1339 displayed attenuated virulence in a macrophage infection model. Following this up with dual-RNA sequencing approaches[50,472] could also elucidate the role of cAMP signalling during *M. tuberculosis* H37Rv infection – in both pathogen and host[78]. Additionally, validating my *M. smegmatis* mc^{2}155 antimicrobial susceptibility data in *M. tuberculosis* H37Rv would be essential to understanding the link between cAMP, environmental adaption and antimicrobial susceptibility. As seen previously, macrophage internalisation of *M. tuberculosis* H37Rv during infection has a greater effect on antimicrobial susceptibility related genes than treatment with specific antimicrobials[49]. With the alterations I observed in permeability and peptidoglycan *in M. smegmatis* mc^{2}155, it would be interesting to see if *M. tuberculosis* H37R expressing Rv1339 can still utilise trehalose for maintenance of carbon metabolism and tolerance to antimicrobials during infection[38]. Although quantifying the MIC_{90} and TTK data on bacteria during macrophage infection would be laborious by traditional methods, models involving fluorescently labelled *M. tuberculosis* H37Rv are already being used to quantify antimicrobial susceptibility during macrophage infection in a microplate format[417]. This approach could also be used in a high-throughput manner to screen the orphan cholesterol inhibitors used in previous studies[11], when cAMP levels are decreased in the bacteria during macrophage infection.

Chapter 6 - Final Discussion

6.0 Final Discussion

M. tuberculosis is an obligate human pathogen that has co-evolved with mankind over 1000s of years[1,2,4]. It is the causative agent of Tuberculosis (TB), which in 2018 was responsible for 10 million new infections and 1.5 million deaths – the most of any infectious disease worldwide[5]. Furthermore, with the emergence of the Covid-19 pandemic, modelling analysis suggests the 5-year Tuberculosis death toll could increase by 20% - as a result of the public health interventions required to prevent Covid-19 spread[473].

A key problem in dealing with eradicating *M. tuberculosis* is the emergence of antimicrobial resistance[21] and the innate tolerance[474] of the bacteria towards many antimicrobials. For example, *M. tuberculosis* bacilli populations in the infected host can display physiologically and metabolically heterogenous phenotypes[48]. The different populations of bacilli during infection display different susceptibility profiles to antimicrobials[47,48]. Without treatment with a combination of antimicrobials targeting the bacteria in each of these populations, resistance can rapidly develop from spontaneous mutations in the bacterial genome[58,266]. Additionally, the thick, waxy and complex cell walls of mycobacteria[254,475], flexible bioenergetic pathways[305] and propensity for formation of dormant persister cells[247], mean that *M. tuberculosis* bacilli are tolerant to a broad set of antimicrobials[474]. Compounding this tolerance, is the infectious niche that develops during prolonged *M. tuberculosis* infection. For example, deposition of Caseum in necrotic Tuberculosis granulomas interferes with the pharmacokinetic parameters of many drugs, thereby decreasing their efficacy even further[56].

The problems of innate tolerance and this allowing time for the development of full-blown resistance stem, in part, from the physiological adaptions the bacteria undergo in order to mediate successful infection in their host environment[247,476,477] . The primary infectious niche of *M. tuberculosis* bacilli is within the host alveolar macrophage[50]. For a pathogen in the macrophage, it is a harsh environment with the constant threat of suppression by the immune response[50,78,478]. In order to survive in this niche, acquire nutrients and propagate their existence, *M. tuberculosis* bacilli have evolved to adapt their physiology to meet these challenges[4,50].

Adaption to environmental conditions is mediated by signalling in a diverse range of life forms[63,84,217]. In bacteria, cAMP signalling is broadly conserved across many different taxa[65]. In *M. tuberculosis*[70], *P. aeruginosa*[179], *V. cholerae*[165], *E. coli*[263] and *Salmonella spp*[261], cAMP signalling has been shown to mediate adaption to the host environmental conditions. *M. tuberculosis* is also a perfect example of how the physiological adaptions to the environment, mediated by cAMP, have the effect of modulating susceptibility to antimicrobials treatment. When exposed to levels of NaCl found within the alveolar macrophage (250mM), *M. tuberculosis* H37Rv bacteria displayed drastically increased minimum inhibitory concentration (MIC_{90}) values to a variety of 1st and 2nd line TB antimicrobials[47]. In a

subsequent study, *M. tuberculosis* H37Rv bacilli exposed to 250mM NaCl displayed a 2.5 fold increase in cAMP levels in response[70]. When the cAMP binding transcription factor Crp_{MT} was knocked out, this same increase in cAMP levels in response to 250mM NaCl was not observed[70]. This suggests that not only is cAMP involved in the response to physiological NaCl levels, but that the cAMP binding transcription factor, Crp_{MT} orchestrates this response[70].

Beyond this example, *M. tuberculosis* H37Rv encodes 16 adenylate cyclase (AC) enzymes that produce cAMP[68]. The *M. tuberculosis* H37Rv AC enzymes are classed as Class-III adenylate cyclase enzymes – this denotes that they typically have a cAMP producing domain linked to a sensory domain[216,217]. *E. coli* in contrast only possesses 1 bona fide cAMP adenylate cyclase – Cya[194]. The sheer number of AC enzymes reinforces the importance of cAMP; and the fusion of AC domains to sensory domains implies that they are stimulated in response to environmental conditions[68]. *M. tuberculosis* H37Rv encodes 10 cAMP binding effector proteins, although not all are characterised[68]. Of the characterised cAMP effector binding proteins, Crp_{MT} has been shown to be required for virulence in a mouse infection model[177]. Additionally, it regulates key enzymes required to maintain essential bioenergetics and ATP production during hypoxic infection conditions e.g. succinate dehydrogenase 1 (Sdh1) – mediating an essential process to maintain membrane potential under anaerobic conditions[173,213].

Regulation of persistence, where *M. tuberculosis* H37Rv bacilli drastically slow their growth and shut down non-essential metabolism are also essential processes for successful infection[209,247]. Cmr, another cAMP binding transcription factor has been shown to repress the DosR, dormancy regulon[248] and is responsive to host immune mediated redox stress[243]. Furthermore, Rv0998, a cAMP binding lysine acetyltransferase has shown 10-fold increased activity in the presence of cAMP and is autoinhibitory in the absence of cAMP[251(p201)]. Rv0998 mediated post translational modifications (PTMs) have been shown to regulate enzymes required for fatty acid utilisation – this is highly important as host derived cholesterol is a key carbon source of *M. tuberculosis* H37Rv bacteria during macrophage infection[284]. The toxic downstream product of cholesterol metabolism is propionyl CoA – the detoxification of which is essential for *M. tuberculosis* H37Rv survivial[215]. Rv0998 has also been shown to regulate Acetyl-CoA synthetase[250], as well as enzymes from the glyoxylate shunt[255] and the methyl citrate pathway[256] – all required for detoxifying the downstream products of cholesterol metabolism.

Adaption to the host environment is essential for full virulence[50], but the transcriptional response to the host immune pressures during *M. tuberculosis* H37Rv infection has been shown to have a greater effect on decreasing antimicrobial susceptibility, than treatment with the actual antimicrobials[49]. Combined with evidence of the involvement of cAMP in modulation of antimicrobials tolerance in *Salmonella spp*[261,262] and *E. coli*[263] (via regulation of transporters and nutrient uptake[264]), this implies that cAMP could also regulate antimicrobial susceptibility in mycobacteria.

It has proven difficult to investigate *M. tuberculosis* H37Rv cAMP signalling, due to the functional redundancy of enzymes that produce it and the numerous effector proteins[68]. To approach this problem, our lab decided to target the degradation of cAMP. Prior to this project, Rv0805 was the only annotated cAMP degrading enzyme (a phosphodiesterase – PDE) in mycobacteria[68,205]. A study from 1979 which showed cAMP PDE activity in the non-pathogenic mycobacteria *M. smegmatis*. However, Rv0805 is only present in pathogenic mycobacteria[279]. Rv0805 displays relatively poor cAMP PDE activity relative to *E. coli* (Rv0805 V_{max} of 30 nmol/sec/mg vs. *E. coli*[194] CpdA 2µmol/sec/mg), despite mycobacterial cAMP levels being 10-100 fold higher than *E. coli*[174,175,204]. Additionally, Rv0805 displays 150-fold increased activity to alternate substrates[281] and cAMP hydrolysis-independent moonlighting activities[282]. Taken together, it was clear that there must be another cAMP PDE enzyme.

Identifying this cAMP PDE enzyme was not possible by standard homology-based approaches. However, by utilising an unbiased, activity-led, biochemical approach, I was able to identify this enzyme: Rv1339, a metallo-β-lactamase 2 enzyme (Chapter 3). In Chapter 4, I used evolutionary bioinformatic analysis to characterise this enzyme as an atypical Class-II PDE enzyme. In the process, I was able to identify the need for expansion of this family to other metallo-β-lactamase proteins and a reclassification of the defining motif as[T/S]HXHXDH to accommodate Rv1339 and other recently discovered cAMP PDEs (e.g. CpdA in *C. glutamicum*[321] and YfhI in *Bacillus subtillis*[328]). I now recommend the Class-II PDE family be subdivided into Bona fide (Pfam: PF02112) and Atypical Class-II PDEs (Pfam: PF12706). These enzymes likely all share the ability to hydrolyse phosphodiester bonds, and diverge based on their substrates[328].

Attempts to purify this enzyme for enzyme kinetic studies or crystallography did not prove completely successful (Chapter 4). In order to eventually purify this enzyme, I believe the best approach would be by optimising the culture media and lysis method of any preparation. In this way, the amount of chaperone binding may be reduced. Furthermore, expression in the M. *smegmatis* mc^2155 GroEL1ΔC appeared to decrease chaperone contamination. A combination of culture, lysis conditions and this strain could yield relatively pure protein (~80%) for enzymatic characterisation. In fact, this was what was seen in my early mycobacterial pVV16::Rv1339 purification, the levels of probable chaperones were less than Rv1339. However, I was still able to do some preliminary substrate studies on the enzyme and show that it possesses cAMP PDE activity *in-vitro* and *in-vivo*.

Taking this work forward, in Chapter 5, I characterised the effect of decreased cAMP levels on *M. smegmatis* mc^2155. I was able to confirm the hypothesis that modulation of cAMP levels altered bacterial tolerance to various antimicrobials with different mechanisms of action. This appeared to stem from a generalised change in physiology that I found to be potentially related to increased permeability and compromised bioenergetics. With more time, I would validate this work in *M. tuberculosis* H37Rv. However, I was still able to do some preliminary work in this area. I found that by expressing Rv1339

in *M. tuberculosis* H37Rv parental, ΔCrp_{MT} and ΔRv0805 backgrounds, I was able to identify different alterations to the metabolome. Expression of Rv1339 in the H37Rv parental strain mediated a 30% decrease in cAMP (similar to overexpression of Rv0805[282]). This was accompanied by significantly decreased abundance of methyl-citrate and pyruvate that may be a result of altering the cAMP dependant lysine acetyltransferase activity of Rv0998 (known to regulate ICL – the enzyme that produces pyruvate from methyl-isocitrate[256]). Beyond this, I saw an 80% increase in arginine levels and decreases in several upstream metabolites from this synthesis pathway. This might indicate a role for cAMP in arginine biosynthesis, which by extension, could mean a role in protecting against reactive oxygen species (ROS) damage to DNA or cell envelope damage from oxidative stress[470]. Reinforcing this hypothesis, Tiwari et al (2018) recently published a study indicating that cAMP levels were increased 150-fold in *M. tuberculosis* H37Rv under arginine starvation[469]. Alternatively, this could indicate that cAMP levels are raised in response to oxidative cell envelope stress or ROS mediated damage to DNA or cellular components – as these were also seen in arginine starvation. There has been some evidence of a link between the response to redox stress and cAMP signalling – with the transcription factor Cmr[243] or the Crp_{MT} regulated transcription factors WhiB1-7[479] being responsive to the bacterial redox state.

When Rv1339 was expressed in an *M. tuberculosis* H37Rv ΔRv0805 strain, the levels of diaminopimelate, D-ala-D-ala and succinate were all significantly decreased. This may support my finding from *M. smegmatis* mc^2155 that cAMP has a role in directly or indirectly modulating peptidoglycan biosynthesis. While I previously suggested this may be through stress placed on the ATP pool, it is possible Rv1339 possesses additional moonlighting activity similar to Rv0805. This seems unlikely as there was no detectable decrease in D-ala-D-ala or diaminopimelate levels when Rv1339 was expressed in the H37Rv parental strain. There may have been a 26% increase in D-ala-D-ala levels (similar to my *M. smegmatis* mc^2155 data), though the p-value for this change was just above significance (0.0788). I hypothesise that this more likely reflects decreased cAMP levels, and the decrease observed in the ΔRv0805 background when compared to the H37Rv parental strain is indicative of the permeability regulating effects previously reported for Rv0805[281,282]. Likewise, the further decrease in diaminopimelate and D-ala-D-ala may be mediated by Rv1339 expression in the ΔRv0805 background may be due to a compounding effect on the pre-existing dysregulation. Although cAMP levels were not significantly altered, they were already 50% of the levels found in the H37Rv parental strain. This may indicate that the ability to decrease cAMP from modulation of degradation was at its limit and it is likely that increased synthesis compensated for the additional activity of Rv1339. Conversely, a 100% increase in cAMP levels was observed when Rv1339 was expressed in the ΔCrp_{MT} background. Again, it is worth noting that cAMP levels were already decreased by 75% compared to the H37Rv parental strain. It is therefore likely the increased cAMP PDE activity mediated

by Rv1339 led to compensation by increased synthesis. Presuming this is the case, U-^{13}C studies would be required to confirm this hypothesis.

Deletion of Crp_{MT} from *M. tuberculosis* H37Rv leads to impaired growth and compromised virulence[177] and this has been seen to be at least in part mediated by compromising expression of *serC*, the Crp_{MT}-regulated essential gene coding for phosphoserine aminotransferase[467]. This enzyme is involved in serine biosynthesis and was required for growth in the absence of serine supplementation[467]. More recent studies have shown that *M. tuberculosis* H37Rv produces serine as it cannot acquire serine during host infection, thereby suggesting that this pathway represents a promising therapeutic target[453]. It is likely there are other such pathways that are in part regulated by cAMP signalling. From my *M. tuberculosis* H37Rv data, it was clear that expression of Rv1339 complemented the growth of ΔCrp_{MT} bacteria back to H37Rv parental strain levels. Concomitantly, the levels of serine, pyruvate and methyl-citrate were increased, it is therefore possible that this explains why growth levels were restored to comparable levels to the H37Rv parental strain – and this suggests a role for cAMP signalling in central carbon metabolism.

In order to finish the preliminary *M. tuberculosis* H37Rv studies, it is imperative to confirm Rv1339 expression by qRT-PCR and the cAMP PDE activity in all strains via U-^{13}C labelling metabolite turnover analysis. It will then be vital to develop a fluorescent *M. tuberculosis* H37Rv system to quantify the TTK and MIC_{90} of various antimicrobials, in order to confirm my data from *M. smegmatis* mc^{2}155. These tools already exist: this approach is utilised by pharmaceutical companies for high-throughput screening of compounds and by academic researchers for virulence or antimicrobial efficacy studies[11,50,417].

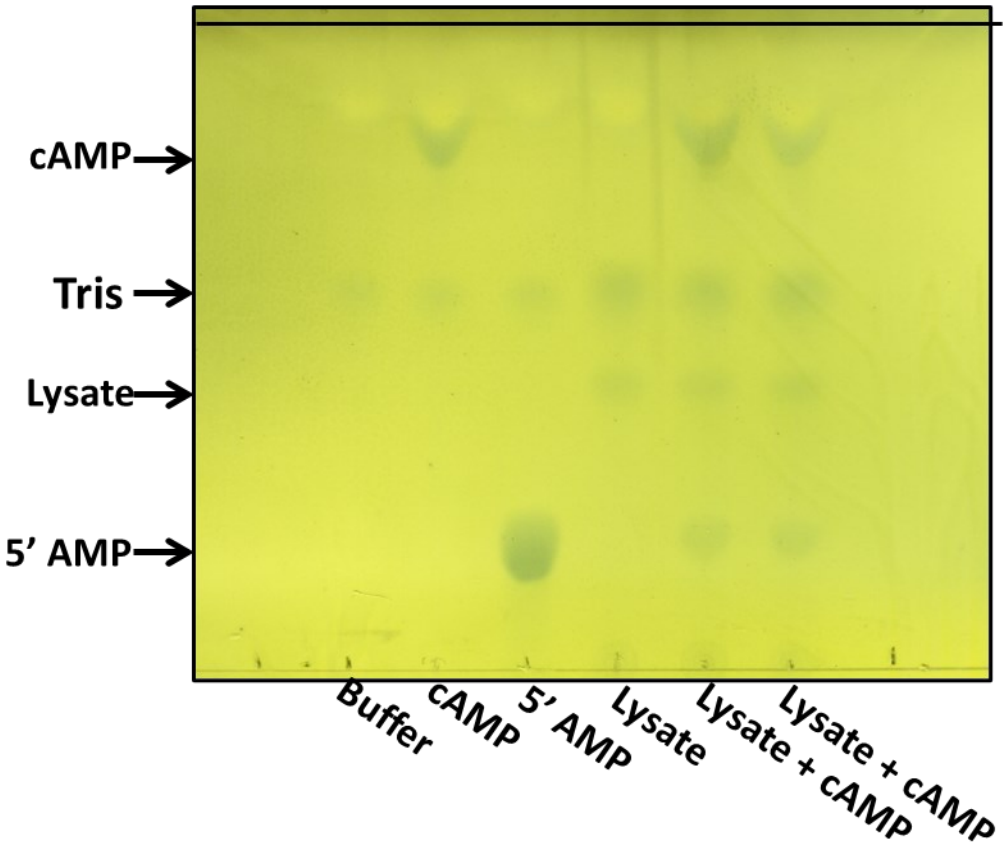
The implications of these findings are numerous. To begin, the integral functions of cAMP signalling in the regulation of mycobacterial virulence[78,177,243], essential gene expression[173,206,211,213] and carbon metabolsim[250,251,255,308] should already make it a tempting target for the development of new antimicrobials. A screen of novel cholesterol orphan inhibitors identified that chemical activation of the cAMP producing adenylate cyclase Rv1625 led to bacterial killing when cholesterol was used as carbon source[11]. The same inhibitor was also shown to compromise propionate metabolism in a subsequent study and lead to increased cAMP levels[215]. Although the exact mechanism was not elucidated, if we can increase our understanding of cAMP signalling in mycobacteria – it may be possible to identify a component of cAMP signalling to target. Identifying a means to modulate cAMP levels is a significant step forward in enabling such discoveries. It is also worth noting that by targeting a signalling pathway involved in both cholesterol utilisation and propionate metabolism/detoxification of toxic cholesterol metabolism products – that it would be difficult for the bacteria to evolve away from[215]. Finding essential bacterial processes where the bacteria can not readily evolve tolerance or resistance is already recognised as a key consideration for antimicrobial development[266].

Further to my findings that modulating cAMP levels potentially leads to compromised peptidoglycan crosslinking and increased permeability to antimicrobials, measuring the effects of D-cylcloserine (DCS) and vancomycin when cAMP levels are decreased may confirm this hypothesis. DCS targets 2 enzymes in the D-ala-D-ala branch of peptidoglycan production (alanine racemase and D-ala-D-ala ligase[480] and vancomycin binds to the extending D-ala-D-ala peptides[481], thereby compromising cell wall construction. If decreased cAMP levels are compromising peptidoglycan integrity, these antibiotics should display significantly decreased MIC values or reduced time to killing. It could also open the possibility of combing a novel therapeutic targeting cAMP signalling, with cell wall targeting antibiotics – as this may lead to synergy and increased bactericidal activity.

Given my findings and the available literature, I suggest combining promising compounds from the novel cholesterol orphan inhibitor screen[11], with cAMP signalling system and cholesterol utilisation enzyme genetic mutants. This could potentially highlight the points of these pathways that can be targeted with novel therapeutic approaches. Cholesterol metabolism represents a significant vulnerability for mycobacteria, as although they are capable of co-catabolising different carbon[253] sources or nitrogen sources[453], cholesterol remains a major carbon source, essential for successful virulence[50,284]. I also suggest investigating the mechanisms linking cAMP to antimicrobial susceptibility in other species of bacteria. As previously mentioned, cAMP signalling is ubiquitous across bacterial pathogens[65] and implicated in modulating antimicrobial susceptibility[262,263]. The ubiquitous nature of cAMP signalling means that these findings, or any incidental findings are likely universal across multiple pathogens.

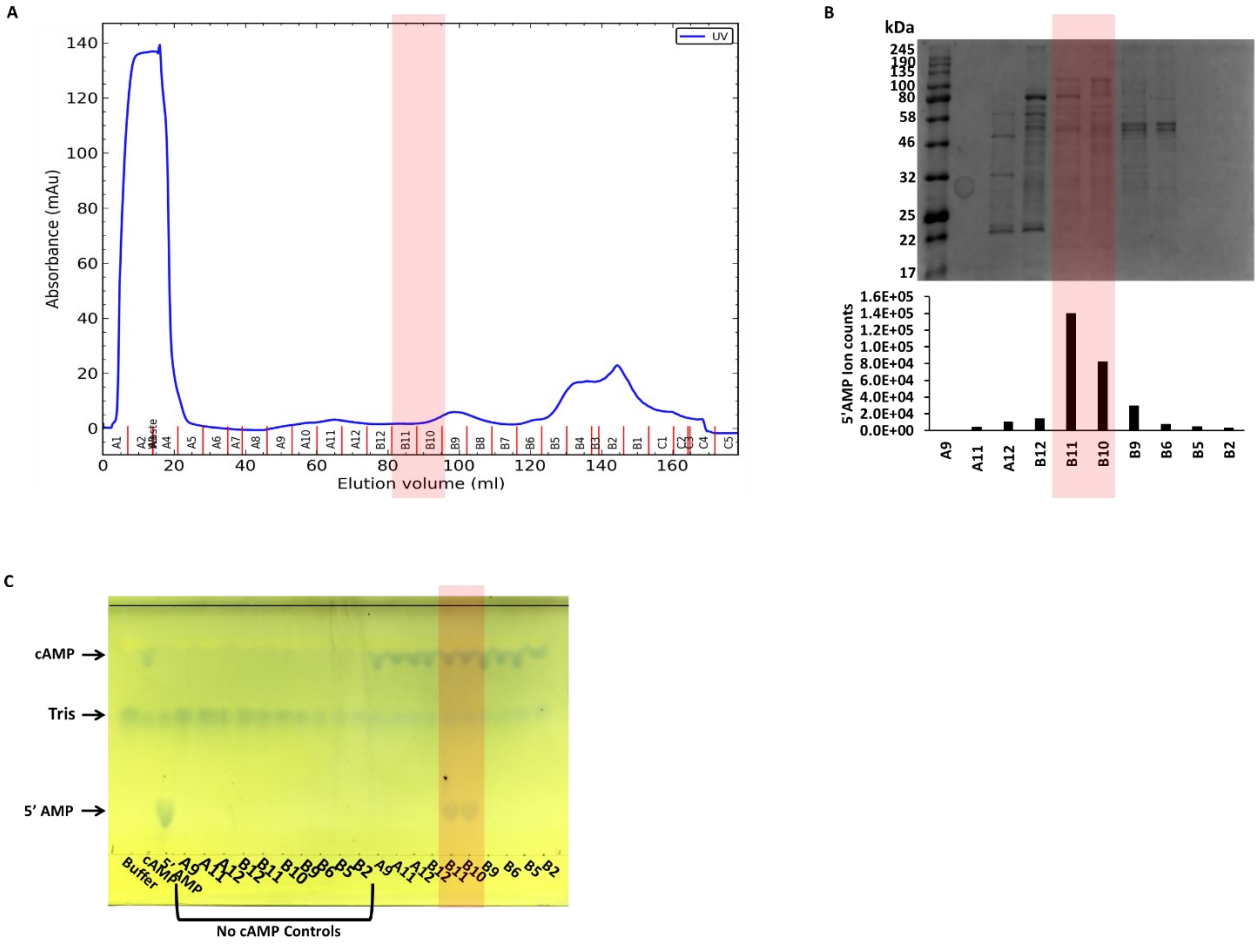
Targeting cAMP signalling may represent a novel avenue to prevent the adaption of bacteria to the host infection environment, and therefore prevent development of the tolerance mechanisms associated with these adaptions[47]. In any case, cAMP signalling represents an area of pathogen research that requires further investigation. With the tool identified in this work, Rv1339, cAMP levels can be decreased in bacteria, and this can allow previously unattainable information about this signalling pathway in mycobacteria and beyond. I hope that the most convincing aspect of this work is that cAMP signalling in pathogenic bacteria deserves more attention, as the evidence clearly suggests great potential for novel therapeutic approaches – potentially to block the development of antimicrobial tolerance, and resistance.

Chapter 7 - Annexes



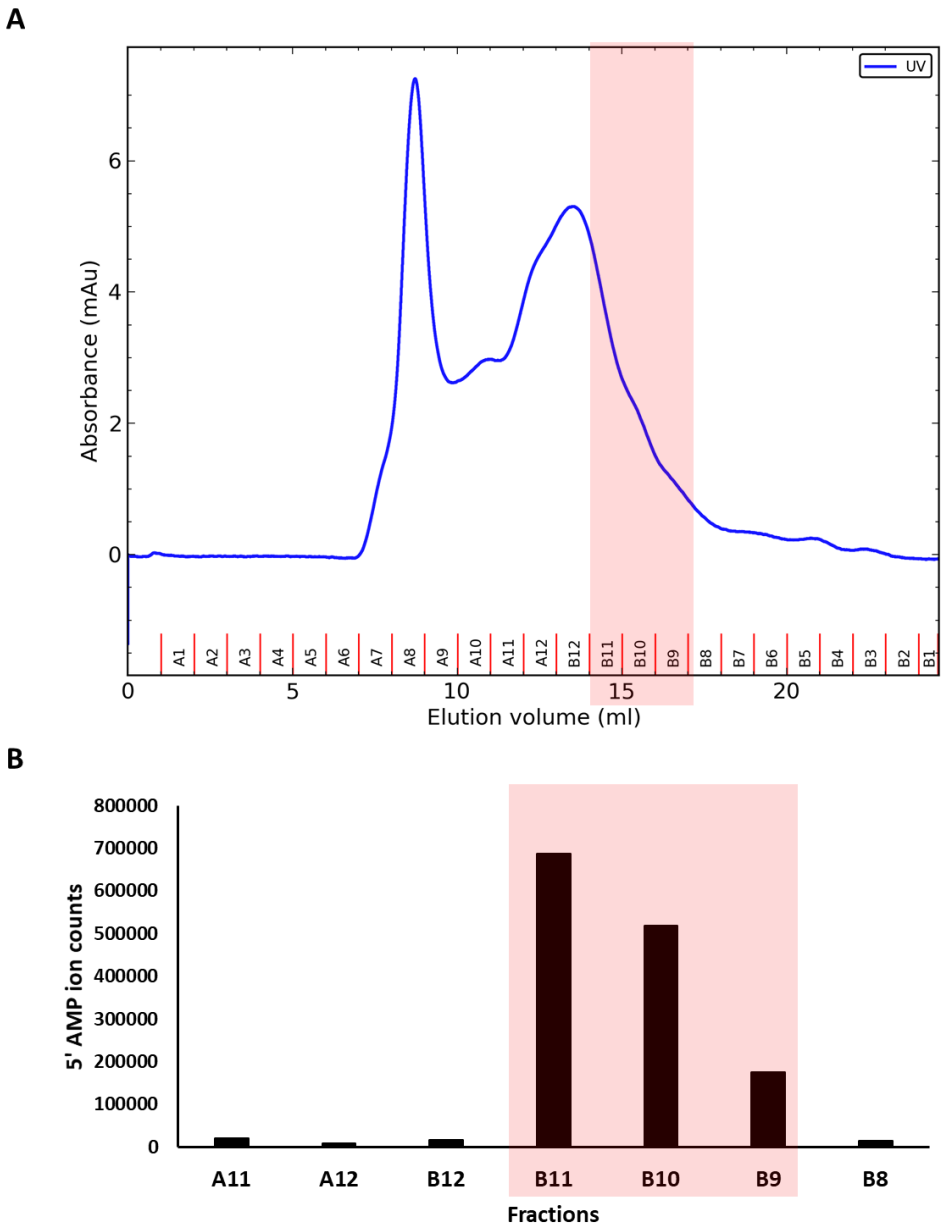
7.1.1 Annex 1:

cAMP PDE activity assay performed on the second batch of M. bovis BCG lysate, subsequently used for the second CaptoQ and first SEC purifications.



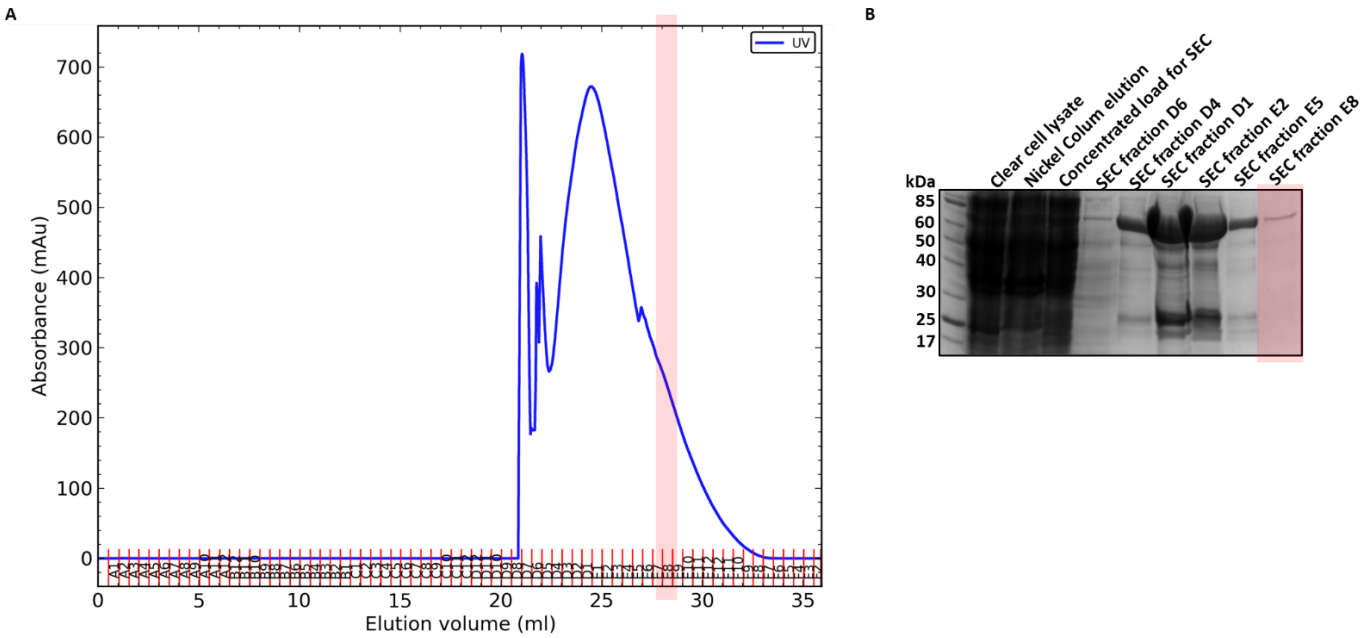
7.1.2 Annex 2:

The CaptoQ purification of the second batch of *M. bovis* BCG Δ*mb0826* (Rv0805 homologue) clear soluble lysate. A: Chromatogram of the purification. B: Coomassie stained SDS-PAGE of the fractions, correlated to the LC-MS readout of the cAMP PDE activity assay performed on the indicated fractions. C: Thin Layer Chromatography (TLC) visualisation of the cAMP PDE activity assay performed on the indicated fractions from the CaptoQ purification.



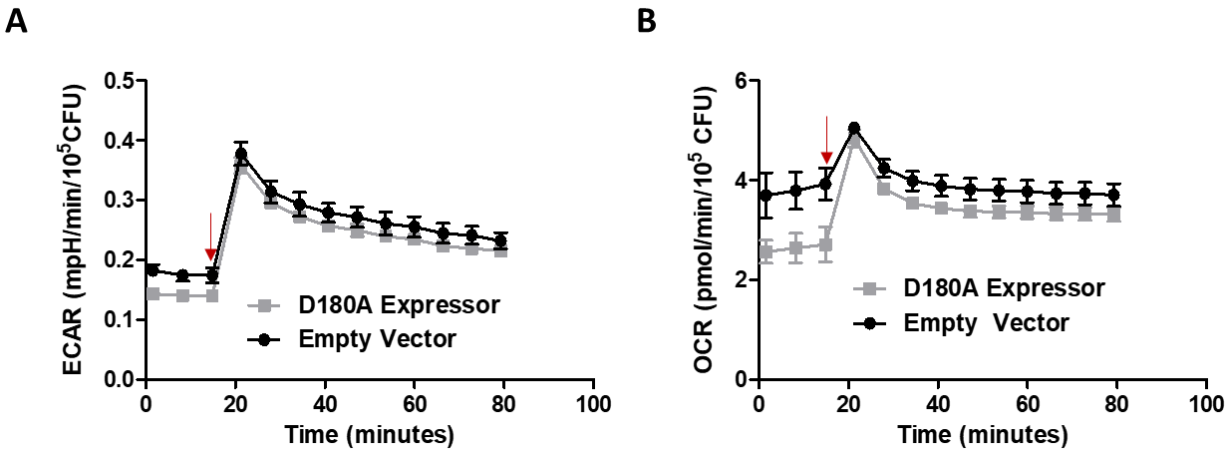
7.1.3 Annex 3:

The second Size Exclusion Chromatography (SEC 2) purification of CaptoQ 2 purified *M. bovis* BCG Δ*mb0828* (rv0805 homologue) clear soluble lysate. A: Chromatogram of the purification. B: LC-MS data from the PDE activity assay performed on the indicated fractions. Significant activity was observed in multiple fractions (B11-B9), in contrast to SEC 1. If the proteomic analysis of SEC 1 had been unsuccessful, my approach would have been to look for a protein present in all 3 fractions.



7.1.4 Annex 4:

Purification of the *M. tuberculosis* H37Rv chaperone DnaK overexpressed in *E. coli* PlysS for the CFE expression assay. A: Chromatogram of the Size Exclusion Chromatography (SEC) purification from the Nickel affinity column elution. B: Coomassie stained SDS-PAGE of each step of the Dnak purification. Fraction E8 was concentrated and stored at -80˚C until use as a supplement in the CFE assay.



7.1.5 Annex 5:

Seahorse XFP assay of Rv1339 D180A and pVV16 Empty Vector Control expressing *M. smegmatis* mc^{2}155. A: Extracellular Acidification Rate (ECAR), no significant differences between the strains after cabron source (glycerol and dextrose to final concentration 0.2%) is injected (red arrow). B: Oxygen Consumption Rate (OCR),), no significant differences observed between the strains after cabron source (glycerol and dextrose to final concentration 0.2%) is injected (red arrow).

Swiss prot entry	Accession n	Annotated description	Average Mass Protein (Da)	Protein matched Peptides	Protein seq coverage (%)
GARA_MYCTU	P9WJA9	Glycogen accumulation regulator GarA OS=	17250.924	4	62.35
GCSH_MYCTU	P9WN55	Glycine cleavage system H protein OS=Myc	14237.6098	1	12.69
Y1339_MYCTU	P9WGC1	Uncharacterized protein Rv1339 OS=Mycob	29154.0739	4	24.54
Y2568_MYCTU	P9WL95	Uncharacterized protein Rv2568c OS=Mycol	38002.7613	6	28.15
Y250_MYCTU	O53672	Uncharacterized protein Rv0250c OS=Mycol	10896.1485	4	51.55
HEM3_MYCTU	P9WMP3	Porphobilinogen deaminase OS=Mycobacte	31938.4536	8	42.39
SERC_MYCTU	P9WQ73	Phosphoserine aminotransferase OS=Myco	40233.4299	6	39.36
RL25_MYCTU	P9WHB5	50S ribosomal protein L25 OS=Mycobacteriu	22441.2422	1	15.81
PCKG_MYCTU	P9WIH3	Phosphoenolpyruvate carboxykinase [GTP]	67253.2672	15	40.43
FADA4_MYCTU	P9WG69	Probable acetyl-CoA acetyltransferase OS=	40080.6563	7	31.11
SAHH_MYCTU	P9WGV3	Adenosylhomocysteinase OS=Mycobacteri	54323.7648	16	42.22
ETFB_MYCTU	P9WNG7	Electron transfer flavoprotein subunit beta	28080.828	13	76.69
ADHB_MYCTU	P9WQC7	Alcohol dehydrogenase B OS=Mycobacteriu	39747.824	10	48.8
ACPM_MYCTU	P9WQF3	Meromycolate extension acyl carrier protei	12523.9751	10	93.91
HRP1_MYCTU	P9WJA3	Hypoxic response protein 1 OS=Mycobacte	15517.7834	12	89.51
Y2971_MYCTU	P9WQA5	Uncharacterized oxidoreductase Rv2971 OS	30364.1541	12	63.12
MDH_MYCTU	P9WK13	Malate dehydrogenase OS=Mycobacterium	34321.8606	13	64.44
AHPC_MYCTU	P9WQB7	Alkyl hydroperoxide reductase subunit C O	21566.2985	12	78.46
ADOK_MYCTU	P9WID5	Adenosine kinase OS=Mycobacterium tube	34472.1404	10	50
GARA_MYCTU	P9WJA9	Glycogen accumulation regulator GarA OS=	17250.924	4	62.35
ADHB_MYCTU	P9WQC7	Alcohol dehydrogenase B OS=Mycobacteriu	39747.824	10	48.8
ETFB_MYCTU	P9WNG7	Electron transfer flavoprotein subunit beta	28080.828	13	76.69
ACEA_MYCTU	P9WKK7	Isocitrate lyase OS=Mycobacterium tubercu	47086.6918	14	55.61
PCKG_MYCTU	P9WIH3	Phosphoenolpyruvate carboxykinase [GTP]	67253.2672	15	40.43
CISY1_MYCTU	P9WPD5	Citrate synthase 1 OS=Mycobacterium tube	47978.5273	9	22.04
DESA2_MYCTU	P9WNZ5	Putative acyl-[acyl-carrier-protein] desatur	31359.5587	10	53.82
ADHC_MYCTU	P9WQC5	NADP-dependent alcohol dehydrogenase C	37075.3997	4	19.36
GPMA_MYCTU	P9WIC9	2_3-bisphosphoglycerate-dependent phosp	27215.8253	5	31.33
RF2_MYCTU	P9WHG1	Peptide chain release factor 2 OS=Mycobac	41473.255	4	12.67
DLDH_MYCTU	P9WHH9	Dihydrolipoyl dehydrogenase OS=Mycobac	49239.2275	5	15.52
RIR2B_MYCTU	P9WH71	Ribonucleoside-diphosphate reductase sub	36991.5873	6	14.81
MASZ_MYCTU	P9WK17	Malate synthase G OS=Mycobacterium tube	80403.1902	4	6.07

7.1.6 Annex 6:

Table containing all *M. tuberculosis* H37Rv proteins unique to the proteomic fraction B10.

Chapter 8 - References

8.0 References

1. Nerlich AG, Haas CJ, Zink A, Szeimies U, Hagedorn HG. Molecular evidence for tuberculosis in an ancient Egyptian mummy. *The Lancet*. 1997;350(9088):1404. doi:10.1016/S0140-6736(05)65185-9

2. Salo WL, Aufderheide AC, Buikstra J, Holcomb TA. Identification of Mycobacterium tuberculosis DNA in a pre-Columbian Peruvian mummy. *Proc Natl Acad Sci U S A*. 1994;91(6):2091-2094. doi:10.1073/pnas.91.6.2091

3. Daniel TM. The history of tuberculosis. *Respir Med*. 2006;100(11):1862-1870. doi:10.1016/j.rmed.2006.08.006

4. Comas I, Coscolla M, Luo T, et al. Out-of-Africa migration and Neolithic co-expansion of Mycobacterium tuberculosis with modern humans. *Nat Genet*. 2013;45(10):1176-1182. doi:10.1038/ng.2744

5. WHO fact sheet Tuberculosis (TB) 2018. Accessed March 29, 2020. https://www.who.int/news-room/fact-sheets/detail/tuberculosis

6. Casali N, Broda A, Harris SR, Parkhill J, Brown T, Drobniewski F. Whole Genome Sequence Analysis of a Large Isoniazid-Resistant Tuberculosis Outbreak in London: A Retrospective Observational Study. *PLOS Med*. 2016;13(10):e1002137. doi:10.1371/journal.pmed.1002137

7. Kontsevaya I, Nikolayevskyy V, Kovalyov A, et al. Tuberculosis cases caused by heterogeneous infection in Eastern Europe and their influence on outcomes. *Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis*. 2017;48:76-82. doi:10.1016/j.meegid.2016.12.016

8. Rangaka MX, Cavalcante SC, Marais BJ, et al. Controlling the Seedbeds of Tuberculosis: Diagnosis and Treatment of Tuberculosis Infection. *Lancet Lond Engl*. 2015;386(10010):2344-2353. doi:10.1016/S0140-6736(15)00323-2

9. Pai M, Behr MA, Dowdy D, et al. Tuberculosis. *Nat Rev Dis Primer*. 2016;2(1):1-23. doi:10.1038/nrdp.2016.76

10. Russell DG. Mycobacterium tuberculosis and the intimate discourse of a chronic infection. *Immunol Rev*. 2011;240(1):252-268. doi:10.1111/j.1600-065X.2010.00984.x

11. VanderVen BC, Fahey RJ, Lee W, et al. Novel Inhibitors of Cholesterol Degradation in Mycobacterium tuberculosis Reveal How the Bacterium's Metabolism Is Constrained by the Intracellular Environment. *PLOS Pathog*. 2015;11(2):e1004679. doi:10.1371/journal.ppat.1004679

12. Fang Z, Sampson SL, Warren RM, Gey van Pittius NC, Newton-Foot M. Iron acquisition strategies in mycobacteria. *Tuberc Edinb Scotl*. 2015;95(2):123-130. doi:10.1016/j.tube.2015.01.004

13. Gouzy A, Larrouy-Maumus G, Bottai D, et al. Mycobacterium tuberculosis Exploits Asparagine to Assimilate Nitrogen and Resist Acid Stress during Infection. *PLoS Pathog*. 2014;10(2). doi:10.1371/journal.ppat.1003928

14. Babu S, Nutman TB. Helminth-Tuberculosis Co-infection: An Immunologic Perspective. *Trends Immunol*. 2016;37(9):597-607. doi:10.1016/j.it.2016.07.005

15. Bhatt K, Verma S, Ellner JJ, Salgame P. Quest for correlates of protection against tuberculosis. *Clin Vaccine Immunol CVI*. 2015;22(3):258-266. doi:10.1128/CVI.00721-14

16. Mangtani P, Abubakar I, Ariti C, et al. Protection by BCG vaccine against tuberculosis: a systematic review of randomized controlled trials. *Clin Infect Dis Off Publ Infect Dis Soc Am*. 2014;58(4):470-480. doi:10.1093/cid/cit790

17. LUCA S, MIHAESCU T. History of BCG Vaccine. *Mædica*. 2013;8(1):53-58.

18. Kaufmann SHE, Weiner J, von Reyn CF. Novel approaches to tuberculosis vaccine development. *Int J Infect Dis IJID Off Publ Int Soc Infect Dis*. 2017;56:263-267. doi:10.1016/j.ijid.2016.10.018

19. Weltwirtschaftsforum, Zurich Insurance Group. *Global Risks 2019: Insight Report*.; 2019. Accessed March 29, 2020. http://www3.weforum.org/docs/WEF_Global_Risks_Report_2019.pdf

20. WHO Antimicrobial resistance Fact sheet (2018). Accessed March 29, 2020. https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance

21. Gygli SM, Borrell S, Trauner A, Gagneux S. Antimicrobial resistance in Mycobacterium tuberculosis: mechanistic and evolutionary perspectives. *FEMS Microbiol Rev*. 2017;41(3):354-373. doi:10.1093/femsre/fux011

22. Jorge P, Magalhães AP, Grainha T, et al. Antimicrobial resistance three ways: healthcare crisis, major concepts and the relevance of biofilms. *FEMS Microbiol Ecol*. 2019;95(8). doi:10.1093/femsec/fiz115

23. Juhas M. Horizontal gene transfer in human pathogens. *Crit Rev Microbiol*. 2015;41(1):101-108. doi:10.3109/1040841X.2013.804031

24. Antonova ES, Hammer BK. Genetics of Natural Competence in Vibrio cholerae and other Vibrios. *Microbiol Spectr*. 2015;3(3). doi:10.1128/microbiolspec.VE-0010-2014

25. D'Costa VM, King CE, Kalan L, et al. Antibiotic resistance is ancient. *Nature*. 2011;477(7365):457-461. doi:10.1038/nature10388

26. Hernando-Amado S, Coque TM, Baquero F, Martínez JL. Defining and combating antibiotic resistance from One Health and Global Health perspectives. *Nat Microbiol*. 2019;4(9):1432-1442. doi:10.1038/s41564-019-0503-9

27. Berendonk TU, Manaia CM, Merlin C, et al. Tackling antibiotic resistance: the environmental framework. *Nat Rev Microbiol*. 2015;13(5):310-317. doi:10.1038/nrmicro3439

28. Martinez JL. Environmental pollution by antibiotics and by antibiotic resistance determinants. *Environ Pollut*. 2009;157(11):2893-2902. doi:10.1016/j.envpol.2009.05.051

29. Ruiz J, Gordon M, Villarreal E, et al. Influence of antibiotic pressure on multi-drug resistant Klebsiella pneumoniae colonisation in critically ill patients. *Antimicrob Resist Infect Control*. 2019;8(1):38. doi:10.1186/s13756-019-0484-8

30. Vicetti Miguel CP, Mejias A, Leber A, Sanchez PJ. A decade of antimicrobial resistance in Staphylococcus aureus: A single center experience. *PLoS ONE*. 2019;14(2). doi:10.1371/journal.pone.0212029

31. Pang Z, Raudonis R, Glick BR, Lin T-J, Cheng Z. Antibiotic resistance in Pseudomonas aeruginosa: mechanisms and alternative therapeutic strategies. *Biotechnol Adv*. 2019;37(1):177-192. doi:10.1016/j.biotechadv.2018.11.013

32. Haberecht HB, Nealon NJ, Gilliland JR, et al. Antimicrobial-Resistant Escherichia coli from Environmental Waters in Northern Colorado. Journal of Environmental and Public Health. doi:https://doi.org/10.1155/2019/3862949

33. Havenga B, Ndlovu T, Clements T, Reyneke B, Waso M, Khan W. Exploring the antimicrobial resistance profiles of WHO critical priority list bacterial strains. *BMC Microbiol*. 2019;19(1):303. doi:10.1186/s12866-019-1687-0

34. Brauner A, Shoresh N, Fridman O, Balaban NQ. An Experimental Framework for Quantifying Bacterial Tolerance. *Biophys J*. 2017;112(12):2664-2671. doi:10.1016/j.bpj.2017.05.014

35. Balaban NQ, Helaine S, Lewis K, et al. Definitions and guidelines for research on antibiotic persistence. *Nat Rev Microbiol*. 2019;17(7):441-448. doi:10.1038/s41579-019-0196-3

36. Lewis K. Persister cells, dormancy and infectious disease. *Nat Rev Microbiol*. 2007;5(1):48-56. doi:10.1038/nrmicro1557

37. Grassi L, Di Luca M, Maisetta G, et al. Generation of Persister Cells of Pseudomonas aeruginosa and Staphylococcus aureus by Chemical Treatment and Evaluation of Their Susceptibility to Membrane-Targeting Agents. *Front Microbiol*. 2017;8. doi:10.3389/fmicb.2017.01917

38. Lee JJ, Lee S-K, Song N, et al. Transient drug-tolerance and permanent drug-resistance rely on the trehalose-catalytic shift in Mycobacterium tuberculosis. *Nat Commun*. 2019;10(1):2928. doi:10.1038/s41467-019-10975-7

39. Pontes MH, Groisman EA. Slow growth determines nonheritable antibiotic resistance in Salmonella enterica. *Sci Signal*. 2019;12(592). doi:10.1126/scisignal.aax3938

40. Adams KN, Takaki K, Connolly LE, et al. Drug Tolerance in Replicating Mycobacteria Mediated by a Macrophage-Induced Efflux Mechanism. *Cell*. 2011;145(1):39-53. doi:10.1016/j.cell.2011.02.022

41. Fisher RA, Gollan B, Helaine S. Persistent bacterial infections and persister cells. *Nat Rev Microbiol*. 2017;15(8):453-464. doi:10.1038/nrmicro.2017.42

42. Hunt-Serracin AC, Parks BJ, Boll J, Boutte CC. Mycobacterium abscessus Cells Have Altered Antibiotic Tolerance and Surface Glycolipids in Artificial Cystic Fibrosis Sputum Medium. *Antimicrob Agents Chemother*. 2019;63(7). doi:10.1128/AAC.02488-18

43. Cowman S, Burns K, Benson S, Wilson R, Loebinger MR. The antimicrobial susceptibility of non-tuberculous mycobacteria. *J Infect*. 2016;72(3):324-331. doi:10.1016/j.jinf.2015.12.007

44. Shaw LP, Doyle RM, Kavaliunaite E, et al. Children With Cystic Fibrosis Are Infected With Multiple Subpopulations of Mycobacterium abscessus With Different Antimicrobial Resistance Profiles. *Clin Infect Dis*. 2019;69(10):1678-1686. doi:10.1093/cid/ciz069

45. Mougari F, Guglielmetti L, Raskine L, Sermet-Gaudelus I, Veziris N, Cambau E. Infections caused by Mycobacterium abscessus: epidemiology, diagnostic tools and treatment. *Expert Rev Anti Infect Ther*. 2016;14(12):1139-1154. doi:10.1080/14787210.2016.1238304

46. Nasiri MJ, Haeili M, Ghazi M, et al. New Insights in to the Intrinsic and Acquired Drug Resistance Mechanisms in Mycobacteria. *Front Microbiol*. 2017;8. doi:10.3389/fmicb.2017.00681

47. Larrouy-Maumus G, Marino LB, Madduri AVR, et al. Cell-Envelope Remodeling as a Determinant of Phenotypic Antibacterial Tolerance in Mycobacterium tuberculosis. *ACS Infect Dis*. 2016;2(5):352-360. doi:10.1021/acsinfecdis.5b00148

48. Evangelopoulos D, Fonseca JD da, Waddell SJ. Understanding anti-tuberculosis drug efficacy: rethinking bacterial populations and how we model them. *Int J Infect Dis*. 2015;32:76-80. doi:10.1016/j.ijid.2014.11.028

49. Liu Y, Tan S, Huang L, et al. Immune activation of the host cell induces drug tolerance in Mycobacterium tuberculosis both in vitro and in vivo. *J Exp Med*. 2016;213(5):809-825. doi:10.1084/jem.20151248

50. Russell DG, Huang L, VanderVen BC. Immunometabolism at the interface between macrophages and pathogens. *Nat Rev Immunol*. 2019;19(5):291-304. doi:10.1038/s41577-019-0124-9

51. Fonseca JD, Knight GM, McHugh TD. The complex evolution of antibiotic resistance in Mycobacterium tuberculosis. *Int J Infect Dis*. 2015;32:94-100. doi:10.1016/j.ijid.2015.01.014

52. Horsburgh CR, Barry CE, Lange C. Treatment of Tuberculosis. *N Engl J Med*. 2015;373(22):2149-2160. doi:10.1056/NEJMra1413919

53. Falzon D, Schünemann HJ, Harausz E, et al. World Health Organization treatment guidelines for drug-resistant tuberculosis, 2016 update. *Eur Respir J*. 2017;49(3). doi:10.1183/13993003.02308-2016

54. Campbell EA, Korzheva N, Mustaev A, et al. Structural Mechanism for Rifampicin Inhibition of Bacterial RNA Polymerase. *Cell*. 2001;104(6):901-912. doi:10.1016/S0092-8674(01)00286-0

55. Lamont EA, Dillon NA, Baughn AD. The Bewildering Antitubercular Action of Pyrazinamide. *Microbiol Mol Biol Rev*. 2020;84(2). doi:10.1128/MMBR.00070-19

56. Sarathy JP, Dartois V. Caseum: a Niche for Mycobacterium tuberculosis Drug-Tolerant Persisters. *Clin Microbiol Rev*. 2020;33(3). doi:10.1128/CMR.00159-19

57. Sarathy JP, Via LE, Weiner D, et al. Extreme Drug Tolerance of Mycobacterium tuberculosis in Caseum. *Antimicrob Agents Chemother*. 2018;62(2). doi:10.1128/AAC.02266-17

58. Pyle MM. Relative numbers of resistant tubercle bacilli in sputa of patients before and during treatment with streptomycin. *Proc Staff Meet Mayo Clin*. 1947;22(21):465-473.

59. Kerantzas CA, Jacobs WR. Origins of Combination Therapy for Tuberculosis: Lessons for Future Antimicrobial Development and Application. *mBio*. 2017;8(2). doi:10.1128/mBio.01586-16

60. Ettehad D, Schaaf HS, Seddon JA, Cooke GS, Ford N. Treatment outcomes for children with multidrug-resistant tuberculosis: a systematic review and meta-analysis. *Lancet Infect Dis*. 2012;12(6):449-456. doi:10.1016/S1473-3099(12)70033-6

61. Dutta NK, Karakousis PC. Latent tuberculosis infection: myths, models, and molecular mechanisms. *Microbiol Mol Biol Rev MMBR*. 2014;78(3):343-371. doi:10.1128/MMBR.00010-14

62. Alnimr AM. Dormancy models for Mycobacterium tuberculosis: A minireview. *Braz J Microbiol Publ Braz Soc Microbiol*. 2015;46(3):641-647. doi:10.1590/S1517-838246320140507

63. Skyrms B. Evolution of signalling systems with multiple senders and receivers. *Philos Trans R Soc B Biol Sci*. 2009;364(1518):771-779. doi:10.1098/rstb.2008.0258

64. Godfrey-Smith P. Signals: Evolution, Learning, and Information, by Brian Skyrms. *Mind*. 2011;120(480):1288-1297. doi:10.1093/mind/fzs002

65. McDonough KA, Rodriguez A. The myriad roles of cyclic AMP in microbial pathogens, from signal to sword. *Nat Rev Microbiol*. 2011;10(1):27-38. doi:10.1038/nrmicro2688

66. Botsford JL, Harman JG. Cyclic AMP in prokaryotes. *Microbiol Rev*. 1992;56(1):100-122.

67. Rickemberg HV. Cyclic AMP in Prokaryotes. *Annu Rev Microbiol*. 1974;28(1):353-369. doi:10.1146/annurev.mi.28.100174.002033

68. Johnson RM, McDonough KA. Cyclic nucleotide signaling in Mycobacterium tuberculosis: an expanding repertoire. *Pathog Dis*. 2018;76(5). doi:10.1093/femspd/fty048

69. Tan S, Sukumar N, Abramovitch RB, Parish T, Russell DG. Mycobacterium tuberculosis responds to chloride and pH as synergistic cues to the immune status of its host cell. *PLoS Pathog*. 2013;9(4):e1003282. doi:10.1371/journal.ppat.1003282

70. Rebollo-Ramirez S, Larrouy-Maumus G. NaCl triggers the CRP-dependent increase of cAMP in Mycobacterium tuberculosis. *Tuberculosis*. 2019;116:8-16. doi:10.1016/j.tube.2019.03.009

71. Molina-Quiroz RC, Silva-Valenzuela C, Brewster J, Castro-Nallar E, Levy SB, Camilli A. Cyclic AMP Regulates Bacterial Persistence through Repression of the Oxidative Stress Response and SOS-Dependent DNA Repair in Uropathogenic Escherichia coli. *mBio*. 2018;9(1). doi:10.1128/mBio.02144-17

72. Fong JCN, Yildiz FH. Interplay between Cyclic AMP-Cyclic AMP Receptor Protein and Cyclic di-GMP Signaling in Vibrio cholerae Biofilm Formation. *J Bacteriol*. 2008;190(20):6646-6659. doi:10.1128/JB.00466-08

73. Changeux J-P, Christopoulos A. Allosteric Modulation as a Unifying Mechanism for Receptor Function and Regulation. *Cell*. 2016;166(5):1084-1102. doi:10.1016/j.cell.2016.08.015

74. Berridge M. Module 12: Signalling Defects and Disease. *Cell Signal Biol*. 2014;6:csb0001012. doi:10.1042/csb0001012

75. Schwartz RS. Paul Ehrlich's Magic Bullets. *N Engl J Med*. 2004;350(11):1079-1080. doi:10.1056/NEJMp048021

76. Maehle AH. 'Receptive substances' : John Newport Langley (1852-1925) and his path to a receptor theory of drug action. *Med Hist*. 2004;48(2):153-174.

77. Krebs EG, Beavo JA. Phosphorylation-Dephosphorylation of Enzymes. *Annu Rev Biochem*. 1979;48(1):923-959. doi:10.1146/annurev.bi.48.070179.004423

78. Agarwal N, Lamichhane G, Gupta R, Nolan S, Bishai WR. Cyclic AMP intoxication of macrophages by a Mycobacterium tuberculosis adenylate cyclase. *Nature*. 2009;460(7251):98-102. doi:10.1038/nature08123

79. Chung Y, Pasquinelli V, Jurado JO, et al. Elevated Cyclic AMP Inhibits Mycobacterium tuberculosis-Stimulated T-cell IFN-γ Secretion Through Type I Protein Kinase A. *J Infect Dis*. 2018;217(11):1821-1831. doi:10.1093/infdis/jiy079

80. Koga Y, Tsurumaki H, Aoki-Saito H, et al. Roles of Cyclic AMP Response Element Binding Activation in the ERK1/2 and p38 MAPK Signalling Pathway in Central Nervous System, Cardiovascular System, Osteoclast Differentiation and Mucin and Cytokine Production. *Int J Mol Sci*. 2019;20(6):1346. doi:10.3390/ijms20061346

81. Raker VK, Becker C, Steinbrink K. The cAMP Pathway as Therapeutic Target in Autoimmune and Inflammatory Diseases. *Front Immunol*. 2016;7. doi:10.3389/fimmu.2016.00123

82. Bernal-Ulloa SM, Heinzmann J, Herrmann D, et al. Cyclic AMP Affects Oocyte Maturation and Embryo Development in Prepubertal and Adult Cattle. *PLoS ONE*. 2016;11(2). doi:10.1371/journal.pone.0150264

83. Görke B, Stülke J. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nat Rev Microbiol*. 2008;6(8):613-624. doi:10.1038/nrmicro1932

84. Kholodenko BN. CELL SIGNALLING DYNAMICS IN TIME AND SPACE. *Nat Rev Mol Cell Biol*. 2006;7(3):165-176. doi:10.1038/nrm1838

85. Sutherland EW, Rall TW. Fractionation and characterization of a cyclic adenine ribonucleotide formed by tissue particles. *J Biol Chem*. 1958;232(2):1077-1091.

86. Rall TW, Sutherland EW. Formation of a Cyclic Adenine Ribonucleotide by Tissue Particles. *J Biol Chem*. 1958;232(2):1065-1076.

87. Robinson GA, Butcher RW, Sutherland EW. Cyclic AMP. *Annu Rev Biochem*. 1968;37(1):149-174. doi:10.1146/annurev.bi.37.070168.001053

88. Senft G, Schultz G, Munske K, Hoffmann M. Influence of insulin on cyclic 3',5'-AMP phosphodiesterase activity in liver, skeletal muscle, adipose tissue, and kidney. *Diabetologia*. 1968;4(6):322-329. doi:10.1007/BF01211766

89. The Nobel Prize in Physiology or Medicine 1971. NobelPrize.org. Accessed March 28, 2020. https://www.nobelprize.org/prizes/medicine/1971/sutherland/lecture/

90. Kammer GM. The adenylate cyclase-cAMP-protein kinase A pathway and regulation of the immune response. *Immunol Today*. 1988;9(7):222-229. doi:10.1016/0167-5699(88)91220-0

91. Hunter T. Signaling—2000 and Beyond. *Cell*. 2000;100(1):113-127. doi:10.1016/S0092-8674(00)81688-8

92. Gonzalez GA, Montminy MR. Cyclic AMP stimulates somatostatin gene transcription by phosphorylation of CREB at serine 133. *Cell*. 1989;59(4):675-680. doi:10.1016/0092-8674(89)90013-5

93. Bourne HR, Lichtenstein LM, Melmon KL, Henney CS, Weinstein Y, Shearer GM. Modulation of Inflammation and Immunity by Cyclic AMP. *Science*. 1974;184(4132):19-28.

94. Kandel ER. The molecular biology of memory: cAMP, PKA, CRE, CREB-1, CREB-2, and CPEB. *Mol Brain*. 2012;5(1):14. doi:10.1186/1756-6606-5-14

95. Serezani CH, Ballinger MN, Aronoff DM, Peters-Golden M. Cyclic AMP: master regulator of innate immune cell function. *Am J Respir Cell Mol Biol*. 2008;39(2):127-132. doi:10.1165/rcmb.2008-0091TR

96. Klein M, Bopp T. Cyclic AMP Represents a Crucial Component of Treg Cell-Mediated Immune Regulation. *Front Immunol*. 2016;7. doi:10.3389/fimmu.2016.00315

97. Leroy J, Vandecasteele G, Fischmeister R. Cyclic AMP signaling in cardiac myocytes. *Curr Opin Physiol*. 2018;1:161-171. doi:10.1016/j.cophys.2017.11.004

98. YAN K, GAO L-N, CUI Y-L, ZHANG Y, ZHOU X. The cyclic AMP signaling pathway: Exploring targets for successful drug discovery (Review). *Mol Med Rep*. 2016;13(5):3715-3723. doi:10.3892/mmr.2016.5005

99. Biel M, Michalakis S. Cyclic nucleotide-gated channels. *Handb Exp Pharmacol*. 2009;(191):111-136. doi:10.1007/978-3-540-68964-5_7

100. Borland G, Smith BO, Yarwood SJ. EPAC proteins transduce diverse cellular actions of cAMP. *Br J Pharmacol*. 2009;158(1):70-86. doi:10.1111/j.1476-5381.2008.00087.x

101. Soty M, Chilloux J, Delalande F, et al. Post-Translational Regulation of the Glucose-6-Phosphatase Complex by Cyclic Adenosine Monophosphate Is a Crucial Determinant of Endogenous Glucose Production and Is Controlled by the Glucose-6-Phosphate Transporter. *J Proteome Res*. 2016;15(4):1342-1349. doi:10.1021/acs.jproteome.6b00110

102. Blokland A, Heckman P, Vanmierlo T, Schreiber R, Paes D, Prickaerts J. Phosphodiesterase Type 4 Inhibition in CNS Diseases. *Trends Pharmacol Sci*. 2019;40(12):971-985. doi:10.1016/j.tips.2019.10.006

103. Newton AC, Bootman MD, Scott JD. Second Messengers. *Cold Spring Harb Perspect Biol*. 2016;8(8). doi:10.1101/cshperspect.a005926

104. Heldin C-H, Lu B, Evans R, Gutkind JS. Signals and Receptors. *Cold Spring Harb Perspect Biol*. 2016;8(4). doi:10.1101/cshperspect.a005900

105. Li Y, Talotta-Altenburg LM, Silimperi KA, Ciabattoni GO, Lowe-Krentz LJ. Endothelial nitric oxide synthase activation is required for heparin receptor effects on vascular smooth muscle cells. *Am J Physiol-Cell Physiol*. 2019;318(3):C463-C475. doi:10.1152/ajpcell.00284.2018

106. Harcus D, Nantel A, Marcil A, Rigby T, Whiteway M. Transcription profiling of cyclic AMP signaling in Candida albicans. *Mol Biol Cell*. 2004;15(10):4490-4499. doi:10.1091/mbc.e04-02-0144

107. Dey B, Bishai WR. Crosstalk between Mycobacterium tuberculosis and the host cell. *Semin Immunol*. 2014;26(6):486-496. doi:10.1016/j.smim.2014.09.002

108. Commichau FM, Dickmanns A, Gundlach J, Ficner R, Stülke J. A jack of all trades: the multiple roles of the unique essential second messenger cyclic di-AMP. *Mol Microbiol*. 2015;97(2):189-204. doi:10.1111/mmi.13026

109. Römling U, Galperin MY, Gomelsky M. Cyclic di-GMP: the First 25 Years of a Universal Bacterial Second Messenger. *Microbiol Mol Biol Rev MMBR*. 2013;77(1):1-52. doi:10.1128/MMBR.00043-12

110. Hengge R. Principles of c-di-GMP signalling in bacteria. *Nat Rev Microbiol*. 2009;7(4):263-273. doi:10.1038/nrmicro2109

111. Fahmi T, Port GC, Cho KH. c-di-AMP: An Essential Molecule in the Signaling Pathways that Regulate the Viability and Virulence of Gram-Positive Bacteria. *Genes*. 2017;8(8). doi:10.3390/genes8080197

112. Bellini D, Caly DL, McCarthy Y, et al. Crystal structure of an HD-GYP domain cyclic-di-GMP phosphodiesterase reveals an enzyme with a novel trinuclear catalytic iron centre. *Mol Microbiol*. 2014;91(1):26-38. doi:10.1111/mmi.12447

113. Duerig A, Abel S, Folcher M, et al. Second messenger-mediated spatiotemporal control of protein degradation regulates bacterial cell cycle progression. *Genes Dev*. 2009;23(1):93-104. doi:10.1101/gad.502409

114. Valentini M, Filloux A. Biofilms and Cyclic di-GMP (c-di-GMP) Signaling: Lessons from Pseudomonas aeruginosa and Other Bacteria. *J Biol Chem*. 2016;291(24):12547-12555. doi:10.1074/jbc.R115.711507

115. Roy AB, Sauer K. Diguanylate cyclase NicD based signaling mechanism of nutrient-induced dispersion by Pseudomonas aeruginosa. *Mol Microbiol*. 2014;94(4):771-793. doi:10.1111/mmi.12802

116. Luo Y, Zhao K, Baker AE, et al. A Hierarchical Cascade of Second Messengers Regulates Pseudomonas aeruginosa Surface Behaviors. *mBio*. 2015;6(1). doi:10.1128/mBio.02456-14

117. Kazmierczak BI, Schniederberend M, Jain R. Cross-regulation of Pseudomonas motility systems: the intimate relationship between flagella, pili and virulence. *Curr Opin Microbiol*. 2015;28:78-82. doi:10.1016/j.mib.2015.07.017

118. Bowman L, Zeden MS, Schuster CF, Kaever V, Gründling A. New Insights into the Cyclic Di-adenosine Monophosphate (c-di-AMP) Degradation Pathway and the Requirement of the Cyclic Dinucleotide for Acid Stress Resistance in Staphylococcus aureus. *J Biol Chem*. 2016;291(53):26970-26986. doi:10.1074/jbc.M116.747709

119. Zeden MS, Schuster CF, Bowman L, Zhong Q, Williams HD, Gründling A. Cyclic di-adenosine monophosphate (c-di-AMP) is required for osmotic regulation in Staphylococcus aureus but dispensable for viability in anaerobic conditions. *J Biol Chem*. 2018;293(9):3180-3200. doi:10.1074/jbc.M117.818716

120. Blötz C, Treffon K, Kaever V, Schwede F, Hammer E, Stülke J. Identification of the Components Involved in Cyclic Di-AMP Signaling in Mycoplasma pneumoniae. *Front Microbiol*. 2017;8. doi:10.3389/fmicb.2017.01328

121. Abdul-Sater AA, Tattoli I, Jin L, et al. Cyclic-di-GMP and cyclic-di-AMP activate the NLRP3 inflammasome. *EMBO Rep*. 2013;14(10):900-906. doi:10.1038/embor.2013.132

122. Woodward JJ, Iavarone AT, Portnoy DA. c-di-AMP secreted by intracellular Listeria monocytogenes activates a host type I interferon response. *Science*. 2010;328(5986):1703-1705. doi:10.1126/science.1189801

123. Boniecka J, Prusińska J, Dąbrowska GB, Goc A. Within and beyond the stringent response-RSH and (p)ppGpp in plants. *Planta*. 2017;246(5):817-842. doi:10.1007/s00425-017-2780-y

124. Steinchen W, Bange G. The magic dance of the alarmones (p)ppGpp. *Mol Microbiol*. 2016;101(4):531-544. doi:10.1111/mmi.13412

125. Magnusson LU, Farewell A, Nyström T. ppGpp: a global regulator in Escherichia coli. *Trends Microbiol*. 2005;13(5):236-242. doi:10.1016/j.tim.2005.03.008

126. Dalebroux ZD, Svensson SL, Gaynor EC, Swanson MS. ppGpp Conjures Bacterial Virulence. *Microbiol Mol Biol Rev MMBR*. 2010;74(2):171-199. doi:10.1128/MMBR.00046-09

127. Irving SE, Corrigan RM. Triggering the stringent response: signals responsible for activating (p)ppGpp synthesis in bacteria. *Microbiol Read Engl*. 2018;164(3):268-276. doi:10.1099/mic.0.000621

128. Ross W, Sanchez-Vazquez P, Chen AY, Lee J-H, Burgos HL, Gourse RL. ppGpp Binding to a Site at the RNAP-DksA Interface Accounts for Its Dramatic Effects on Transcription Initiation during the Stringent Response. *Mol Cell*. 2016;62(6):811-823. doi:10.1016/j.molcel.2016.04.029

129. Ma Z, King K, Alqahtani M, et al. Stringent response governs the oxidative stress resistance and virulence of Francisella tularensis. *PLoS ONE*. 2019;14(10). doi:10.1371/journal.pone.0224094

130. do Vale A, Cabanes D, Sousa S. Bacterial Toxins as Pathogen Weapons Against Phagocytes. *Front Microbiol*. 2016;7. doi:10.3389/fmicb.2016.00042

131. Ahuja N, Kumar P, Bhatnagar R. The Adenylate Cyclase Toxins. *Crit Rev Microbiol*. 2004;30(3):187-196. doi:10.1080/10408410490468795

132. Rabaan AA. Cholera: an overview with reference to the Yemen epidemic. *Front Med*. 2019;13(2):213-228. doi:10.1007/s11684-018-0631-2

133. Chowdhury FR, Nur Z, Hassan N, von Seidlein L, Dunachie S. Pandemics, pathogenicity and changing molecular epidemiology of cholera in the era of global warming. *Ann Clin Microbiol Antimicrob*. 2017;16(1):10. doi:10.1186/s12941-017-0185-1

134. Kirn TJ, Jude BA, Taylor RK. A colonization factor links Vibrio cholerae environmental survival and human infection. *Nature*. 2005;438(7069):863-866. doi:10.1038/nature04249

135. Thomas A, Ramananda Y, Mun K, Naren AP, Arora K. AC6 is the major adenylate cyclase forming a diarrheagenic protein complex with cystic fibrosis transmembrane conductance regulator in cholera. *J Biol Chem*. 2018;293(33):12949-12959. doi:10.1074/jbc.RA118.003378

136. Randak CO. An elusive adenylate cyclase complicit in cholera is exposed. *J Biol Chem*. 2018;293(33):12960-12961. doi:10.1074/jbc.H118.004669

137. Sakib SN, Reddi G, Almagro-Moreno S. Environmental Role of Pathogenic Traits in Vibrio cholerae. *J Bacteriol*. 2018;200(15). doi:10.1128/JB.00795-17

138. Finley NL. Revealing how an adenylate cyclase toxin uses bait and switch tactics in its activation. *PLOS Biol*. 2018;16(2):e2005356. doi:10.1371/journal.pbio.2005356

139. Dorji D, Mooi F, Yantorno O, Deora R, Graham RM, Mukkur TK. Bordetella Pertussis virulence factors in the continuing evolution of whooping cough vaccines for improved performance. *Med Microbiol Immunol (Berl)*. 2018;207(1):3-26. doi:10.1007/s00430-017-0524-z

140. Carbonetti NH. Pertussis toxin and adenylate cyclase toxin: key virulence factors of Bordetella pertussis and cell biology tools. *Future Microbiol*. 2010;5:455-469. doi:10.2217/fmb.09.133

141. Cannella SE, Ntsogo Enguéné VY, Davi M, et al. Stability, structural and functional properties of a monomeric, calcium–loaded adenylate cyclase toxin, CyaA, from Bordetella pertussis. *Sci Rep*. 2017;7(1):1-17. doi:10.1038/srep42065

142. Fiser R, Masin J, Bumba L, et al. Calcium influx rescues adenylate cyclase-hemolysin from rapid cell membrane removal and enables phagocyte permeabilization by toxin pores. *PLoS Pathog*. 2012;8(4):e1002580. doi:10.1371/journal.ppat.1002580

143. O'Brien DP, Durand D, Voegele A, et al. Calmodulin fishing with a structurally disordered bait triggers CyaA catalysis. *PLOS Biol*. 2017;15(12):e2004486. doi:10.1371/journal.pbio.2004486

144. Paccani SR, Dal Molin F, Benagiano M, et al. Suppression of T-lymphocyte activation and chemotaxis by the adenylate cyclase toxin of Bordetella pertussis. *Infect Immun*. 2008;76(7):2822-2832. doi:10.1128/IAI.00200-08

145. Arumugham VB, Ulivieri C, Onnis A, et al. Compartmentalized Cyclic AMP Production by the Bordetella pertussis and Bacillus anthracis Adenylate Cyclase Toxins Differentially Affects the Immune Synapse in T Lymphocytes. *Front Immunol*. 2018;9. doi:10.3389/fimmu.2018.00919

146. Shumilla JA, Lacaille V, Hornell TMC, et al. Bordetella pertussis infection of primary human monocytes alters HLA-DR expression. *Infect Immun*. 2004;72(3):1450-1462. doi:10.1128/iai.72.3.1450-1462.2004

147. Martino A, Volpe E, Auricchio G, Colizzi V, Baldini PM. Influence of pertussis toxin on CD1a isoform expression in human dendritic cells. *J Clin Immunol*. 2006;26(2):153-159. doi:10.1007/s10875-006-9009-3

148. Orlos Z, Rakoczi E, Misak O, et al. Outbreak of anthrax in adults and adolescents: a review of nine cases in a regional teaching hospital in East Hungary. *Clin Microbiol Infect*. 2017;23(1):52-54. doi:10.1016/j.cmi.2015.11.023

149. Pilo P, Frey J. Pathogenicity, population genetics and dissemination of Bacillus anthracis. *Infect Genet Evol*. 2018;64:115-125. doi:10.1016/j.meegid.2018.06.024

150. Welkos S, Bozue J, Twenhafel N, Cote C. Animal Models for the Pathogenesis, Treatment, and Prevention of Infection by Bacillus anthracis. *Microbiol Spectr*. 2015;3(1). doi:10.1128/microbiolspec.TBS-0001-2012

151. Jernigan DB, Raghunathan PL, Bell BP, et al. Investigation of Bioterrorism-Related Anthrax, United States, 2001: Epidemiologic Findings. *Emerg Infect Dis*. 2002;8(10):1019-1028. doi:10.3201/eid0810.020353

152. Pan Z, Dumas EK, Lawrence C, et al. Bacillus anthracis Edema Toxin Inhibits Efferocytosis in Human Macrophages and Alters Efferocytic Receptor Signaling. *Int J Mol Sci*. 2019;20(5). doi:10.3390/ijms20051167

153. Tang W-J, Guo Q. The adenylyl cyclase activity of anthrax edema factor. *Mol Aspects Med*. 2009;30(6):423-430. doi:10.1016/j.mam.2009.06.001

154. Guo Q, Jureller JE, Warren JT, Solomaha E, Florián J, Tang W-J. Protein-protein docking and analysis reveal that two homologous bacterial adenylyl cyclase toxins interact with calmodulin differently. *J Biol Chem*. 2008;283(35):23836-23845. doi:10.1074/jbc.M802168200

155. Schuler D, Lübker C, Lushington GH, et al. Interactions of Bordetella pertussis adenylyl cyclase toxin CyaA with calmodulin mutants and calmodulin antagonists: Comparison with membranous adenylyl cyclase I. *Biochem Pharmacol*. 2012;83(7):839-848. doi:10.1016/j.bcp.2012.01.005

156. Malhotra S, Hayes D, Wozniak DJ. Cystic Fibrosis and Pseudomonas aeruginosa: the Host-Microbe Interface. *Clin Microbiol Rev*. 2019;32(3). doi:10.1128/CMR.00138-18

157. Klockgether J, Tümmler B. Recent advances in understanding Pseudomonas aeruginosa as a pathogen. *F1000Research*. 2017;6. doi:10.12688/f1000research.10506.1

158. Munder A, Rothschuh J, Schirmer B, et al. The Pseudomonas aeruginosa ExoY phenotype of high-copy-number recombinants is not detectable in natural isolates. *Open Biol*. 8(1):170250. doi:10.1098/rsob.170250

159. Feltman H, Schulert G, Khan S, Jain M, Peterson L, Hauser AR. Prevalence of type III secretion genes in clinical and environmental isolates of Pseudomonas aeruginosa. *Microbiology,*. 2001;147(10):2659-2669. doi:10.1099/00221287-147-10-2659

160. Beckert U, Wolter S, Hartwig C, et al. ExoY from Pseudomonas aeruginosa is a nucleotidyl cyclase with preference for cGMP and cUMP formation. *Biochem Biophys Res Commun*. 2014;450(1):870-874. doi:10.1016/j.bbrc.2014.06.088

161. Belyy A, Mechold U, Renault L, Ladant D. ExoY, an actin-activated nucleotidyl cyclase toxin from P. aeruginosa: A minireview. *Toxicon*. 2018;149:65-71. doi:10.1016/j.toxicon.2017.12.046

162. Hritonenko V, Mun JJ, Tam C, et al. Adenylate cyclase activity of Pseudomonas aeruginosa ExoY can mediate bleb-niche formation in epithelial cells and contributes to virulence. *Microb Pathog*. 2011;51(5):305-312. doi:10.1016/j.micpath.2011.08.001

163. Johnson AM, Kaushik RS, Francis DH, Fleckenstein JM, Hardwidge PR. Heat-Labile Enterotoxin Promotes Escherichia coli Adherence to Intestinal Epithelial Cells. *J Bacteriol*. 2009;191(1):178-186. doi:10.1128/JB.00822-08

164. Huang J, Duan Q, Zhang W. Significance of Enterotoxigenic Escherichia coli (ETEC) Heat-Labile Toxin (LT) Enzymatic Subunit Epitopes in LT Enterotoxicity and Immunogenicity. *Appl Environ Microbiol*. 2018;84(15). doi:10.1128/AEM.00849-18

165. Manneh-Roussel J, Haycocks JRJ, Magán A, et al. cAMP Receptor Protein Controls Vibrio cholerae Gene Expression in Response to Host Colonization. *mBio*. 2018;9(4). doi:10.1128/mBio.00966-18

166. Pastan I, Perlman R. Cyclic Adenosine Monophosphate in Bacteria: In many bacteria the synthesis of inducible enzymes requires this cyclic nucleotide. *Science*. 1970;169(3943):339-344. doi:10.1126/science.169.3943.339

167. Pastan I, Adhya S. Cyclic adenosine 5'-monophosphate in Escherichia coli. *Bacteriol Rev*. 1976;40(3):527-551.

168. Green J, Stapleton MR, Smith LJ, et al. Cyclic-AMP and bacterial cyclic-AMP receptor proteins revisited: adaptation for different ecological niches. *Curr Opin Microbiol*. 2014;18(100):1-7. doi:10.1016/j.mib.2014.01.003

169. Bren A, Park JO, Towbin BD, Dekel E, Rabinowitz JD, Alon U. Glucose becomes one of the worst carbon sources for E.coli on poor nitrogen sources due to suboptimal levels of cAMP. *Sci Rep*. 2016;6(1):1-10. doi:10.1038/srep24834

170. Milanesio P, Arce-Rodríguez A, Muñoz A, Calles B, de Lorenzo V. Regulatory exaptation of the catabolite repression protein (Crp)-cAMP system in Pseudomonas putida. *Environ Microbiol*. 2011;13(2):324-339. doi:10.1111/j.1462-2920.2010.02331.x

171. Zampieri M, Hörl M, Hotz F, Müller NF, Sauer U. Regulatory mechanisms underlying coordination of amino acid and glucose catabolism in Escherichia coli. *Nat Commun*. 2019;10(1):1-13. doi:10.1038/s41467-019-11331-5

172. Evangelista W, Dong A, White MA, Li J, Lee JC. Differential modulation of energy landscapes of cyclic AMP receptor protein (CRP) as a regulatory mechanism for class II CRP-dependent promoters. *J Biol Chem*. 2019;294(42):15544-15556. doi:10.1074/jbc.RA119.009151

173. Kahramanoglou C, Cortes T, Matange N, et al. Genomic mapping of cAMP receptor protein (CRPMt) in Mycobacterium tuberculosis: relation to transcriptional start sites and the role of CRPMt as a transcription factor. *Nucleic Acids Res*. 2014;42(13):8320-8329. doi:10.1093/nar/gku548

174. Bai G, Schaak DD, McDonough KA. cAMP levels within Mycobacterium tuberculosis and M. bovis BCG increase upon infection of macrophages. *FEMS Immunol Med Microbiol*. 2009;55(1):68-73. doi:10.1111/j.1574-695X.2008.00500.x

175. Notley-McRobb L, Death A, Ferenci T. The relationship between external glucose concentration and cAMP levels inside Escherichia coli: implications for models of phosphotransferase-mediated regulation of adenylate cyclase. *Microbiol Read Engl*. 1997;143 (Pt 6):1909-1918. doi:10.1099/00221287-143-6-1909

176. Arce-Rodríguez A, Durante-Rodríguez G, Platero R, Krell T, Calles B, de Lorenzo V. The Crp regulator of Pseudomonas putida: evidence of an unusually high affinity for its physiological effector, cAMP. *Environ Microbiol*. 2012;14(3):702-713. doi:10.1111/j.1462-2920.2011.02622.x

177. Rickman L, Scott C, Hunt DM, et al. A member of the cAMP receptor protein family of transcription regulators in Mycobacterium tuberculosis is required for virulence in mice and controls transcription of the rpfA gene coding for a resuscitation promoting factor. *Mol Microbiol*. 2005;56(5):1274-1286. doi:10.1111/j.1365-2958.2005.04609.x

178. Aung HL, Berney M, Cook GM. Hypoxia-activated cytochrome bd expression in Mycobacterium smegmatis is cyclic AMP receptor protein dependent. *J Bacteriol*. 2014;196(17):3091-3097. doi:10.1128/JB.01771-14

179. Fuchs EL, Brutinel ED, Jones AK, et al. The Pseudomonas aeruginosa Vfr Regulator Controls Global Virulence Factor Expression through Cyclic AMP-Dependent and -Independent Mechanisms. *J Bacteriol*. 2010;192(14):3553-3564. doi:10.1128/JB.00363-10

180. Kochanowski K, Gerosa L, Brunner SF, Christodoulou D, Nikolaev YV, Sauer U. Few regulatory metabolites coordinate expression of central metabolic genes in Escherichia coli. *Mol Syst Biol*. 2017;13(1):903. doi:10.15252/msb.20167402

181. Shimada T, Fujita N, Yamamoto K, Ishihama A. Novel roles of cAMP receptor protein (CRP) in regulation of transport and metabolism of carbon sources. *PloS One*. 2011;6(6):e20081. doi:10.1371/journal.pone.0020081

182. Kremling A, Geiselmann J, Ropers D, de Jong H. Understanding carbon catabolite repression in Escherichia coli using quantitative models. *Trends Microbiol*. 2015;23(2):99-109. doi:10.1016/j.tim.2014.11.002

183. Gayán E, Cambré A, Michiels CW, Aertsen A. RpoS-independent evolution reveals the importance of attenuated cAMP/CRP regulation in high hydrostatic pressure resistance acquisition in E. coli. *Sci Rep*. 2017;7(1):1-11. doi:10.1038/s41598-017-08958-z

184. Basak S, Geng H, Jiang R. Rewiring global regulator cAMP receptor protein (CRP) to improve E. coli tolerance towards low pH. *J Biotechnol*. 2014;173:68-75. doi:10.1016/j.jbiotec.2014.01.015

185. Chong H, Yeow J, Wang I, Song H, Jiang R. Improving Acetate Tolerance of Escherichia coli by Rewiring Its Global Regulator cAMP Receptor Protein (CRP). *PLoS ONE*. 2013;8(10). doi:10.1371/journal.pone.0077422

186. Geng H, Jiang R. cAMP receptor protein (CRP)-mediated resistance/tolerance in bacteria: mechanism and utilization in biotechnology. *Appl Microbiol Biotechnol*. 2015;99(11):4533-4543. doi:10.1007/s00253-015-6587-0

187. Nosho K, Fukushima H, Asai T, et al. cAMP-CRP acts as a key regulator for the viable but non-culturable state in Escherichia coli. *Microbiology,*. 2018;164(3):410-419. doi:10.1099/mic.0.000618

188. Hufnagel DA, Evans ML, Greene SE, Pinkner JS, Hultgren SJ, Chapman MR. The Catabolite Repressor Protein-Cyclic AMP Complex Regulates csgD and Biofilm Formation in Uropathogenic Escherichia coli. *J Bacteriol*. 2016;198(24):3329-3334. doi:10.1128/JB.00652-16

189. Sutrina SL, Daniel K, Lewis M, et al. Biofilm Growth of Escherichia coli Is Subject to cAMP-Dependent and cAMP-Independent Inhibition. *J Mol Microbiol Biotechnol*. 2015;25(2-3):209-225. doi:10.1159/000375498

190. Almblad H, Harrison JJ, Rybtke M, et al. The Cyclic AMP-Vfr Signaling Pathway in Pseudomonas aeruginosa Is Inhibited by Cyclic Di-GMP. *J Bacteriol*. 2015;197(13):2190-2200. doi:10.1128/JB.00193-15

191. Almblad H, Rybtke M, Hendiani S, Andersen JB, Givskov M, Tolker-Nielsen T. High levels of cAMP inhibit Pseudomonas aeruginosa biofilm formation through reduction of the c-di-GMP content. *Microbiology,*. 2019;165(3):324-333. doi:10.1099/mic.0.000772

192. Marsden AE, Intile PJ, Schulmeyer KH, et al. Vfr Directly Activates exsA Transcription To Regulate Expression of the Pseudomonas aeruginosa Type III Secretion System. *J Bacteriol*. 2016;198(9):1442-1450. doi:10.1128/JB.00049-16

193. Wolfgang MC, Lee VT, Gilmore ME, Lory S. Coordinate Regulation of Bacterial Virulence Genes by a Novel Adenylate Cyclase-Dependent Signaling Pathway. *Dev Cell*. 2003;4(2):253-263. doi:10.1016/S1534-5807(03)00019-4

194. Imamura R, Yamanaka K, Ogura T, et al. Identification of the cpdA Gene Encoding Cyclic 3',5'-Adenosine Monophosphate Phosphodiesterase in Escherichia coli. *J Biol Chem*. 1996;271(41):25423-25429. doi:10.1074/jbc.271.41.25423

195. Zhang Y, Zhang C, Du X, et al. Glutathione Activates Type III Secretion System Through Vfr in Pseudomonas aeruginosa. *Front Cell Infect Microbiol*. 2019;9. doi:10.3389/fcimb.2019.00164

196. Yang X, Yang X-A, Xu M, Zhou L, Fan Z, Jiang T. Crystal structures of YfiR from Pseudomonas aeruginosa in two redox states. *Biochem Biophys Res Commun*. 2015;461(1):14-20. doi:10.1016/j.bbrc.2015.03.160

197. Hickman JW, Tifrea DF, Harwood CS. A chemosensory system that regulates biofilm formation through modulation of cyclic diguanylate levels. *Proc Natl Acad Sci U S A*. 2005;102(40):14422-14427. doi:10.1073/pnas.0507170102

198. Borlee BR, Goldman AD, Murakami K, Samudrala R, Wozniak DJ, Parsek MR. Pseudomonas aeruginosa uses a cyclic-di-GMP-regulated adhesin to reinforce the biofilm extracellular matrix. *Mol Microbiol*. 2010;75(4):827-842. doi:10.1111/j.1365-2958.2009.06991.x

199. Poncet S, Milohanic E, Mazé A, et al. Correlations between Carbon Metabolism and Virulence in Bacteria. *Contrib Microbiol*. 2009;16:88-102. doi:10.1159/000219374

200. Liang W, Sultan SZ, Silva AJ, Benitez JA. Cyclic AMP post-transcriptionally regulates the biosynthesis of a major bacterial autoinducer to modulate the cell density required to activate quorum sensing. *FEBS Lett*. 2008;582(27):3744-3750. doi:10.1016/j.febslet.2008.10.008

201. Teschler JK, Zamorano-Sánchez D, Utada AS, et al. Living in the matrix: assembly and control of Vibrio cholerae biofilms. *Nat Rev Microbiol*. 2015;13(5):255-268. doi:10.1038/nrmicro3433

202. Liimatta K, Flaherty E, Ro G, Nguyen DK, Prado C, Purdy AE. A Putative Acetylation System in Vibrio cholerae Modulates Virulence in Arthropod Hosts. *Appl Environ Microbiol*. 2018;84(21). doi:10.1128/AEM.01113-18

203. Verma J, Bag DrS, Saha B, et al. Genomic plasticity associated with antimicrobial resistance in Vibrio cholerae. *Proc Natl Acad Sci*. 2019;116. doi:10.1073/pnas.1900141116

204. Dass BKM, Sharma R, Shenoy AR, Mattoo R, Visweswariah SS. Cyclic AMP in mycobacteria: characterization and functional role of the Rv1647 ortholog in Mycobacterium smegmatis. *J Bacteriol*. 2008;190(11):3824-3834. doi:10.1128/JB.00138-08

205. Shenoy AR, Sreenath N, Podobnik M, Kovacevic M, Visweswariah SS. The Rv0805 gene from Mycobacterium tuberculosis encodes a 3',5'-cyclic nucleotide phosphodiesterase: biochemical and mutational analysis. *Biochemistry*. 2005;44(48):15695-15704. doi:10.1021/bi0512391

206. Gazdik MA, Bai G, Wu Y, McDonough KA. Rv1675c (cmr) regulates intramacrophage and cyclic AMP-induced gene expression in Mycobacterium tuberculosis-complex mycobacteria. *Mol Microbiol*. 2009;71(2):434-448. doi:10.1111/j.1365-2958.2008.06541.x

207. Shleeva M, Goncharenko A, Kudykina Y, Young D, Young M, Kaprelyants A. Cyclic Amp-Dependent Resuscitation of Dormant Mycobacteria by Exogenous Free Fatty Acids. *PLoS ONE*. 2013;8(12). doi:10.1371/journal.pone.0082914

208. Agarwal N, Bishai WR. cAMP signaling in Mycobacterium tuberculosis. *INDIAN J EXP BIOL*. Published online 2009:8.

209. Shleeva MO, Kondratieva TK, Demina GR, et al. Overexpression of Adenylyl Cyclase Encoded by the Mycobacterium tuberculosis Rv2212 Gene Confers Improved Fitness, Accelerated Recovery

from Dormancy and Enhanced Virulence in Mice. *Front Cell Infect Microbiol*. 2017;7:370. doi:10.3389/fcimb.2017.00370

210. Bai G, McCue LA, McDonough KA. Characterization of Mycobacterium tuberculosis Rv3676 (CRPMt), a Cyclic AMP Receptor Protein-Like DNA Binding Protein. *J Bacteriol*. 2005;187(22):7795-7804. doi:10.1128/JB.187.22.7795-7804.2005

211. Aung HL, Dixon LL, Smith LJ, et al. Novel regulatory roles of cAMP receptor proteins in fast-growing environmental mycobacteria. *Microbiol Read Engl*. 2015;161(Pt 3):648-661. doi:10.1099/mic.0.000015

212. Pecsi I, Hards K, Ekanayaka N, et al. Essentiality of Succinate Dehydrogenase in Mycobacterium smegmatis and Its Role in the Generation of the Membrane Potential Under Hypoxia. *mBio*. 2014;5(4). doi:10.1128/mBio.01093-14

213. Knapp GS, McDonough KA. Cyclic AMP Signaling in Mycobacteria. In: *Molecular Genetics of Mycobacteria*. John Wiley & Sons, Ltd; 2015:281-295. doi:10.1128/9781555818845.ch14

214. Banerjee A, Adolph RS, Gopalakrishnapai J, et al. A Universal Stress Protein (USP) in Mycobacteria Binds cAMP. *J Biol Chem*. 2015;290(20):12731-12743. doi:10.1074/jbc.M115.644856

215. Johnson RM, Bai G, DeMott CM, et al. Chemical activation of adenylyl cyclase Rv1625c inhibits growth of Mycobacterium tuberculosis on cholesterol and modulates intramacrophage signaling. *Mol Microbiol*. 2017;105(2):294-308. doi:10.1111/mmi.13701

216. Shenoy AR, Visweswariah SS. Mycobacterial adenylyl cyclases: Biochemical diversity and structural plasticity. *FEBS Lett*. 2006;580(14):3344-3352. doi:10.1016/j.febslet.2006.05.034

217. Bassler J, Schultz JE, Lupas AN. Adenylate cyclases: Receivers, transducers, and generators of signals. *Cell Signal*. 2018;46:135-144. doi:10.1016/j.cellsig.2018.03.002

218. Knapp GS, McDonough KA. Cyclic AMP Signaling in Mycobacteria. *Microbiol Spectr*. 2014;2(2). doi:10.1128/microbiolspec.MGM2-0011-2013

219. Tews I, Findeisen F, Sinning I, Schultz A, Schultz JE, Linder JU. The structure of a pH-sensing mycobacterial adenylyl cyclase holoenzyme. *Science*. 2005;308(5724):1020-1023. doi:10.1126/science.1107642

220. Mehta M, Rajmani RS, Singh A. Mycobacterium tuberculosis WhiB3 responds to vacuolar pH-induced changes in mycothiol redox potential to modulate phagosomal maturation and virulence. *J Biol Chem*. Published online December 4, 2015:jbc.M115.684597. doi:10.1074/jbc.M115.684597

221. Schump MD, Fox DM, Bertozzi CR, Riley LW. Subcellular Partitioning and Intramacrophage Selectivity of Antimicrobial Compounds against Mycobacterium tuberculosis. *Antimicrob Agents Chemother*. 2017;61(3). doi:10.1128/AAC.01639-16

222. Rohde KH, Abramovitch RB, Russell DG. Mycobacterium tuberculosis invasion of macrophages: linking bacterial gene expression to environmental cues. *Cell Host Microbe*. 2007;2(5):352-364. doi:10.1016/j.chom.2007.09.006

223. Ganaie AA, Trivedi G, Kaur A, et al. Interaction of Erp Protein of Mycobacterium tuberculosis with Rv2212 Enhances Intracellular Survival of Mycobacterium smegmatis. *J Bacteriol*. 2016;198(20):2841-2852. doi:10.1128/JB.00120-16

224. Cann MJ, Hammer A, Zhou J, Kanacher T. A defined subset of adenylyl cyclases is regulated by bicarbonate ion. *J Biol Chem*. 2003;278(37):35033-35038. doi:http://dx.doi.org/10.1074/jbc.M303025200

225. Olalekan AW, Oluwaseun FA, Oladele HA-W, Akeem AD. Evaluation of electrolyte imbalance among tuberculosis patients receiving treatments in Southwestern Nigeria. *Alex J Med*. 2015;51(3):255-260. doi:10.1016/j.ajme.2014.10.003

226. J R, Br B, Ej R. Genome-wide requirements for Mycobacterium tuberculosis adaptation and survival in macrophages. *Proc Natl Acad Sci U S A*. 2005;102(23):8327-8332. doi:10.1073/pnas.0503272102

227. Supuran CT. Bacterial carbonic anhydrases as drug targets: toward novel antibiotics? *Front Pharmacol*. 2011;2:34. doi:10.3389/fphar.2011.00034

228. Johnson BK, Colvin CJ, Needle DB, Medie FM, Champion PAD, Abramovitch RB. The Carbonic Anhydrase Inhibitor Ethoxzolamide Inhibits the Mycobacterium tuberculosis PhoPR Regulon and Esx-1 Secretion and Attenuates Virulence. *Antimicrob Agents Chemother*. 2015;59(8):4436-4445. doi:10.1128/AAC.00719-15

229. Walters SB, Dubnau E, Kolesnikova I, Laval F, Daffe M, Smith I. The Mycobacterium tuberculosis PhoPR two-component system regulates genes essential for virulence and complex lipid biosynthesis. *Mol Microbiol*. 2006;60(2):312-330. doi:10.1111/j.1365-2958.2006.05102.x

230. Guo YL, Seebacher T, Kurz U, Linder JU, Schultz JE. Adenylyl cyclase Rv1625c of Mycobacterium tuberculosis: a progenitor of mammalian adenylyl cyclases. *EMBO J*. 2001;20(14):3667-3675. doi:10.1093/emboj/20.14.3667

231. Townsend PD, Holliday PM, Fenyk S, et al. Stimulation of Mammalian G-protein-responsive Adenylyl Cyclases by Carbon Dioxide. *J Biol Chem*. 2009;284(2):784-791. doi:10.1074/jbc.M807239200

232. Shenoy AR, Sreenath NP, Mahalingam M, Visweswariah SS. Characterization of phylogenetically distant members of the adenylate cyclase family from mycobacteria: Rv1647 from Mycobacterium tuberculosis and its orthologue ML1399 from M. leprae. *Biochem J*. 2005;387(Pt 2):541-551. doi:10.1042/BJ20041040

233. Singh P, Cole ST. Mycobacterium leprae: genes, pseudogenes and genetic diversity. *Future Microbiol*. 2011;6(1):57-71. doi:10.2217/fmb.10.153

234. Lamichhane S. Mycobacterium tuberculosis: Gene and Genome analysis. Published online June 12, 2018.

235. Malhotra S, Vedithi SC, Blundell TL. Decoding the similarities and differences among mycobacterial species. *PLoS Negl Trop Dis*. 2017;11(8). doi:10.1371/journal.pntd.0005883

236. The resuscitation-promoting factors of Mycobacterium tuberculosis are required for virulence and resuscitation from dormancy but are collectively dispensable for growth in vitro - Kana - 2008 - Molecular Microbiology - Wiley Online Library. Accessed March 29, 2020. https://onlinelibrary.wiley.com/doi/full/10.1111/j.1365-2958.2007.06078.x

237. Knapp GS, Lyubetskaya A, Peterson MW, et al. Role of intragenic binding of cAMP responsive protein (CRP) in regulation of the succinate dehydrogenase genes Rv0249c-Rv0247c in TB complex mycobacteria. *Nucleic Acids Res*. 2015;43(11):5377-5393. doi:10.1093/nar/gkv420

238. Eoh H, Rhee KY. Multifunctional essentiality of succinate metabolism in adaptation to hypoxia in Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A*. 2013;110(16):6554-6559. doi:10.1073/pnas.1219375110

239. Watanabe S, Zimmermann M, Goodwin MB, Sauer U, Barry CE, Boshoff HI. Fumarate reductase activity maintains an energized membrane in anaerobic Mycobacterium tuberculosis. *PLoS Pathog*. 2011;7(10):e1002287. doi:10.1371/journal.ppat.1002287

240. Bald D, Villellas C, Lu P, Koul A. Targeting Energy Metabolism in Mycobacterium tuberculosis, a New Paradigm in Antimycobacterial Drug Discovery. *mBio*. 2017;8(2). doi:10.1128/mBio.00272-17

241. Ranganathan S, Bai G, Lyubetskaya A, et al. Characterization of a cAMP responsive transcription factor, Cmr (Rv1675c), in TB complex mycobacteria reveals overlap with the DosR (DevR) dormancy regulon. *Nucleic Acids Res*. Published online 2016:134-151.

242. Gazdik MA, McDonough KA. Identification of Cyclic AMP-Regulated Genes in Mycobacterium tuberculosis Complex Bacteria under Low-Oxygen Conditions. *J Bacteriol*. 2005;187(8):2681-2692. doi:10.1128/JB.187.8.2681-2692.2005

243. Smith LJ, Bochkareva A, Rolfe MD, et al. Cmr is a redox-responsive regulator of DosR that contributes to M. tuberculosis virulence. *Nucleic Acids Res*. 2017;45(11):6600-6612. doi:10.1093/nar/gkx406

244. Chen T, He L, Deng W, Xie J. The Mycobacterium DosR regulon structure and diversity revealed by comparative genomic analysis. *J Cell Biochem*. 2013;114(1):1-6. doi:10.1002/jcb.24302

245. Leistikow RL, Morton RA, Bartek IL, Frimpong I, Wagner K, Voskuil MI. The Mycobacterium tuberculosis DosR Regulon Assists in Metabolic Homeostasis and Enables Rapid Recovery from Nonrespiring Dormancy. *J Bacteriol*. 2010;192(6):1662-1670. doi:10.1128/JB.00926-09

246. Fallow A, Domenech P, Reed MB. Strains of the East Asian (W/Beijing) lineage of Mycobacterium tuberculosis are DosS/DosT-DosR two-component regulatory system natural mutants. *J Bacteriol*. 2010;192(8):2228-2238. doi:10.1128/JB.01597-09

247. Domenech P, Zou J, Averback A, et al. Unique Regulation of the DosR Regulon in the Beijing Lineage of Mycobacterium tuberculosis. *J Bacteriol*. 2017;199(2). doi:10.1128/JB.00696-16

248. Ranganathan S, Cheung J, Cassidy M, Ginter C, Pata JD, McDonough KA. Novel structural features drive DNA binding properties of Cmr, a CRP family protein in TB complex mycobacteria. *Nucleic Acids Res*. 2018;46(1):403-420. doi:10.1093/nar/gkx1148

249. Nambi S, Basu N, Visweswariah SS. cAMP-regulated protein lysine acetylases in mycobacteria. *J Biol Chem*. 2010;285(32):24313-24323. doi:10.1074/jbc.M110.118398

250. Xu H, Hegde SS, Blanchard JS. Reversible acetylation and inactivation of Mycobacterium tuberculosis acetyl-CoA synthetase is dependent on cAMP. *Biochemistry*. 2011;50(26):5883-5892. doi:10.1021/bi200156t

251. Nambi S, Gupta K, Bhattacharyya M, et al. Cyclic AMP-dependent Protein Lysine Acylation in Mycobacteria Regulates Fatty Acid and Propionate Metabolism. *J Biol Chem*. 2013;288(20):14114-14124. doi:10.1074/jbc.M113.463992

252. Pech-Canul Á, Nogales J, Miranda-Molina A, et al. FadD Is Required for Utilization of Endogenous Fatty Acids Released from Membrane Lipids ▿ . *J Bacteriol*. 2011;193(22):6295-6304. doi:10.1128/JB.05450-11

253. de Carvalho LPS, Fischer SM, Marrero J, Nathan C, Ehrt S, Rhee KY. Metabolomics of Mycobacterium tuberculosis reveals compartmentalized co-catabolism of carbon substrates. *Chem Biol*. 2010;17(10):1122-1131. doi:10.1016/j.chembiol.2010.08.009

254. Gago G, Diacovich L, Gramajo H. Lipid metabolism and its implication in mycobacteria-host interaction. *Curr Opin Microbiol*. 2018;41:36-42. doi:10.1016/j.mib.2017.11.020

255. Liu F, Yang M, Wang X, et al. Acetylome analysis reveals diverse functions of lysine acetylation in Mycobacterium tuberculosis. *Mol Cell Proteomics MCP*. 2014;13(12):3352-3366. doi:10.1074/mcp.M114.041962

256. Serafini A, Tan L, Horswell S, et al. Mycobacterium tuberculosis requires glyoxylate shunt and reverse methylcitrate cycle for lactate and pyruvate metabolism. *Mol Microbiol*. 2019;112(4):1284-1307. doi:10.1111/mmi.14362

257. Bi J, Wang Y, Yu H, et al. Modulation of Central Carbon Metabolism by Acetylation of Isocitrate Lyase in Mycobacterium tuberculosis. *Sci Rep*. 2017;7(1):1-11. doi:10.1038/srep44826

258. Yang H, Sha W, Liu Z, et al. Lysine acetylation of DosR regulates the hypoxia response of Mycobacterium tuberculosis. *Emerg Microbes Infect*. 2018;7(1):1-14. doi:10.1038/s41426-018-0032-2

259. Lee HJ, Lang PT, Fortune SM, Sassetti CM, Alber T. Cyclic AMP regulation of protein lysine acetylation in Mycobacterium tuberculosis. *Nat Struct Mol Biol*. 2012;19(8):811-818. doi:10.1038/nsmb.2318

260. Nash KA, Brown-Elliott BA, Wallace RJ. A Novel Gene, erm(41), Confers Inducible Macrolide Resistance to Clinical Isolates of Mycobacterium abscessus but Is Absent from Mycobacterium chelonae. *Antimicrob Agents Chemother*. 2009;53(4):1367-1376. doi:10.1128/AAC.01275-08

261. Alper MD, Ames BN. Transport of antibiotics and metabolite analogs by systems under cyclic AMP control: positive selection of Salmonella typhimurium cya and crp mutants. *J Bacteriol*. 1978;133(1):149-157.

262. Kary SC, Yoneda JRK, Olshefsky SC, Stewart LA, West SB, Cameron ADS. The Global Regulatory Cyclic AMP Receptor Protein (CRP) Controls Multifactorial Fluoroquinolone Susceptibility in Salmonella enterica Serovar Typhimurium. *Antimicrob Agents Chemother*. 2017;61(11). doi:10.1128/AAC.01666-17

263. Girgis HS, Hottes AK, Tavazoie S. Genetic architecture of intrinsic antibiotic susceptibility. *PloS One*. 2009;4(5):e5629. doi:10.1371/journal.pone.0005629

264. Sawant K, Shashidhar R. CRP i.e. cAMP receptor protein provides a competitive edge to Salmonella Typhimurium in a microbial community set-up. *Access Microbiol*. 2019;1(1A):423. doi:10.1099/acmi.ac2019.po0244

265. Luthra S, Rominski A, Sander P. The Role of Antibiotic-Target-Modifying and Antibiotic-Modifying Enzymes in Mycobacterium abscessus Drug Resistance. *Front Microbiol*. 2018;9. doi:10.3389/fmicb.2018.02179

266. Evangelopoulos D, Prosser GA, Rodgers A, et al. Comparative fitness analysis of D -cycloserine resistant mutants reveals both fitness-neutral and high-fitness cost genotypes. *Nat Commun*. 2019;10(1):4177. doi:10.1038/s41467-019-12074-z

267. Reyrat J-M, Kahn D. Mycobacterium smegmatis: an absurd model for tuberculosis? *Trends Microbiol*. 2001;9(10):472-473. doi:10.1016/S0966-842X(01)02168-0

268. Barry CE. Mycobacterium smegmatis: an absurd model for tuberculosis? *Trends Microbiol*. 2001;9(10):473-474. doi:10.1016/S0966-842X(01)02169-2

269. Shiloh MU, DiGiuseppe Champion PA. To catch a killer. What can mycobacterial models teach us about Mycobacterium tuberculosis pathogenesis? *Curr Opin Microbiol*. 2010;13(1):86-92. doi:10.1016/j.mib.2009.11.006

270. Lelovic N, Mitachi K, Yang J, Lemieux MR, Ji Y, Kurosu M. Application of Mycobacterium smegmatis as a surrogate to evaluate drug leads against Mycobacterium tuberculosis. *J Antibiot (Tokyo)*. Published online May 29, 2020:1-10. doi:10.1038/s41429-020-0320-7

271. Insel PA, Ostrom RS. Forskolin as a Tool for Examining Adenylyl Cyclase Expression, Regulation, and G Protein Signaling. *Cell Mol Neurobiol*. 2003;23(3):305-314. doi:10.1023/A:1023684503883

272. Seamon KB, Daly JW. Forskolin: a unique diterpene activator of cyclic AMP-generating systems. *J Cyclic Nucleotide Res*. 1981;7(4):201-224.

273. Schneyer CR, Piñeyro MA, Gregerman RI. Mechanism of action of forskolin on adenylate cyclase: effect on bovine sperm complemented with erythrocyte membranes. *Life Sci*. 1983;33(3):275-279. doi:10.1016/0024-3205(83)90387-9

274. Garber SS, Hoshi T, Aldrich RW. Interaction of forskolin with voltage-gated K+ channels in PC12 cells. *J Neurosci Off J Soc Neurosci*. 1990;10(10):3361-3368.

275. Kaukel E, Hilz H. Permeation of dibutyryl cAMP into HeLa cells and its conversion to monobutyryl cAMP. *Biochem Biophys Res Commun*. 1972;46(2):1011-1018. doi:10.1016/S0006-291X(72)80242-0

276. Blecher M. Biological effects and catabolic metabolism of 3', 5'-cyclic nucleotides and derivatives in rat adipose tissue and liver. *Metabolism*. 1971;20(1):63-77. doi:10.1016/0026-0495(71)90060-6

277. Veron W, Lesouhaitier O, Pennanec X, et al. Natriuretic peptides affect Pseudomonas aeruginosa and specifically modify lipopolysaccharide biosynthesis. *FEBS J*. 2007;274(22):5852-5864. doi:10.1111/j.1742-4658.2007.06109.x

278. Matange N. Revisiting bacterial cyclic nucleotide phosphodiesterases: cyclic AMP hydrolysis and beyond. *FEMS Microbiol Lett*. 2015;362(22). doi:10.1093/femsle/fnv183

279. Shenoy AR, Capuder M, Draškovič P, Lamba D, Visweswariah SS, Podobnik M. Structural and Biochemical Analysis of the Rv0805 Cyclic Nucleotide Phosphodiesterase from Mycobacterium tuberculosis. *J Mol Biol*. 2007;365(1):211-225. doi:10.1016/j.jmb.2006.10.005

280. Malhotra N, Chakraborti PK. Eukaryotic-Type Ser/Thr Protein Kinase Mediated Phosphorylation of Mycobacterial Phosphodiesterase Affects its Localization to the Cell Wall. *Front Microbiol*. 2016;7:123. doi:10.3389/fmicb.2016.00123

281. Podobnik M, Tyagi R, Matange N, et al. A mycobacterial cyclic AMP phosphodiesterase that moonlights as a modifier of cell wall permeability. *J Biol Chem*. 2009;284(47):32846-32857. doi:10.1074/jbc.M109.049635

282. Matange N, Hunt DM, Buxton RS, Visweswariah SS. Overexpression of the Rv0805 phosphodiesterase elicits a cAMP-independent transcriptional response. *Tuberc Edinb Scotl*. 2013;93(5):492-500. doi:10.1016/j.tube.2013.05.004

283. Griffin JE, Pandey AK, Gilmore SA, et al. Cholesterol catabolism by Mycobacterium tuberculosis requires transcriptional and metabolic adaptations. *Chem Biol*. 2012;19(2):218-227. doi:10.1016/j.chembiol.2011.12.016

284. Pandey AK, Sassetti CM. Mycobacterial persistence requires the utilization of host cholesterol. *Proc Natl Acad Sci U S A*. 2008;105(11):4376-4380. doi:10.1073/pnas.0711159105

285. Matange N, Podobnik M, Visweswariah SS. The Non-catalytic "Cap Domain" of a Mycobacterial Metallophosphoesterase Regulates Its Expression and Localization in the Cell. *J Biol Chem*. 2014;289(32):22470-22481. doi:10.1074/jbc.M114.578328

286. Lee CH. Metabolism of cyclic AMP in non-pathogenic Mycobacterium smegmatis. *Arch Microbiol*. 1979;120(1):35-37. doi:10.1007/bf00413269

287. Keppetipola N, Shuman S. A phosphate-binding histidine of binuclear metallophosphodiesterase enzymes is a determinant of 2',3'-cyclic nucleotide phosphodiesterase activity. *J Biol Chem*. 2008;283(45):30942-30949. doi:10.1074/jbc.M805064200

288. Padh H, Venkitasubramanian TA. Adenosine 3', 5'-monophosphate in mycobacteria. *Life Sci*. 1977;20(7):1273-1280. doi:10.1016/0024-3205(77)90502-1

289. Geertsma ER, Dutzler R. A versatile and efficient high-throughput cloning tool for structural biology. *Biochemistry*. 2011;50(15):3272-3278. doi:10.1021/bi200178z

290. FX cloning. Accessed June 28, 2020. https://www.fxcloning.org/

291. Ginocchio CC. Role of NCCLS in antimicrobial susceptibility testing and monitoring. *Am J Health-Syst Pharm AJHP Off J Am Soc Health-Syst Pharm*. 2002;59(8 Suppl 3):S7-11. doi:10.1093/ajhp/59.suppl_3.S7

292. Palomino J-C, Martin A, Camacho M, Guerra H, Swings J, Portaels F. Resazurin microtiter assay plate: simple and inexpensive method for detection of drug resistance in Mycobacterium tuberculosis. *Antimicrob Agents Chemother*. 2002;46(8):2720-2722. doi:10.1128/aac.46.8.2720-2722.2002

293. Podobnik M, Tyagi R, Matange N, et al. A mycobacterial cyclic AMP phosphodiesterase that moonlights as a modifier of cell wall permeability. *J Biol Chem*. 2009;284(47):32846-32857. doi:10.1074/jbc.M109.049635

294. Wang C, Mao Y, Yu J, et al. PhoY2 of Mycobacteria Is Required for Metabolic Homeostasis and. 2013;195(2):243-252. doi:10.1128/JB.01556-12

295. Karinou E, Schuster CF, Pazos M, Vollmer W. *Inactivation of the Monofunctional Peptidoglycan Glycosyltransferase SgtB Allows Staphylococcus Aureus to Survive in the Absence of Lipoteichoic Acid*. Vol 44.; 2018. doi:10.1128/JB.00574-18

296. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. Accessed June 27, 2020. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

297. Kapopoulou A, Lew JM, Cole ST. The MycoBrowser portal: a comprehensive and manually annotated resource for mycobacterial genomes. *Tuberc Edinb Scotl*. 2011;91(1):8-13. doi:10.1016/j.tube.2010.09.006

298. Li H, Durbin R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics*. 2010;26(5):589-595. doi:10.1093/bioinformatics/btp698

299. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinforma Oxf Engl*. 2014;30(7):923-930. doi:10.1093/bioinformatics/btt656

300. Home - Assembly - NCBI. Accessed June 27, 2020. https://www.ncbi.nlm.nih.gov/assembly/

301. R: The R Project for Statistical Computing. Accessed June 27, 2020. https://www.r-project.org/

302. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550. doi:10.1186/s13059-014-0550-8

303. Huang DW, Sherman BT, Tan Q, et al. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biol*. 2007;8(9):R183. doi:10.1186/gb-2007-8-9-r183

304. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4(1):44-57. doi:10.1038/nprot.2008.211

305. Lamprecht DA, Finin PM, Rahman MA, et al. Turning the respiratory flexibility of Mycobacterium tuberculosis against itself. *Nat Commun*. 2016;7(1):12393. doi:10.1038/ncomms12393

306. Saini V, Chinta KC, Reddy VP, et al. Hydrogen sulfide stimulates Mycobacterium tuberculosis respiration, growth and pathogenesis. *Nat Commun*. 2020;11(1):1-17. doi:10.1038/s41467-019-14132-y

307. Cook GM, Hards K, Vilchèze C, Hartman T, Berney M. Energetics of Respiration and Oxidative Phosphorylation in Mycobacteria. *Microbiol Spectr*. 2014;2(3). doi:10.1128/microbiolspec.MGM2-0015-2013

308. Matange N, Hunt DM, Buxton RS, Visweswariah SS. Overexpression of the Rv0805 phosphodiesterase elicits a cAMP-independent transcriptional response. *Tuberc Edinb Scotl*. 2013;93(5):492-500. doi:10.1016/j.tube.2013.05.004

309. Higashida H, Hossain KZ, Takahagi H, Noda M. Measurement of adenylyl cyclase by separating cyclic AMP on silica gel thin-layer chromatography. *Anal Biochem*. 2002;308(1):106-111. doi:10.1016/s0003-2697(02)00237-3

310. Jackson CJ, Carr PD, Liu J-W, Watt SJ, Beck JL, Ollis DL. The structure and function of a novel glycerophosphodiesterase from Enterobacter aerogenes. *J Mol Biol*. 2007;367(4):1047-1062. doi:10.1016/j.jmb.2007.01.032

311. Osada-Oka M, Tateishi Y, Hirayama Y, et al. Antigen 85A and mycobacterial DNA-binding protein 1 are targets of immunoglobulin G in individuals with past tuberculosis. *Microbiol Immunol*. 2013;57(1):30-37. doi:10.1111/j.1348-0421.2012.12005.x

312. Kassa D, Ran L, Geberemeskel W, et al. Analysis of Immune Responses against a Wide Range of Mycobacterium tuberculosis Antigens in Patients with Active Pulmonary Tuberculosis. *Clin Vaccine Immunol CVI*. 2012;19(12):1907-1915. doi:10.1128/CVI.00482-12

313. Miller DJ, Shuvalova L, Evdokimova E, Savchenko A, Yakunin AF, Anderson WF. Structural and biochemical characterization of a novel Mn2+-dependent phosphodiesterase encoded by the yfcE gene. *Protein Sci Publ Protein Soc*. 2007;16(7):1338-1348. doi:10.1110/ps.072764907

314. Shehadul Islam M, Aryasomayajula A, Selvaganapathy PR. A Review on Macroscale and Microscale Cell Lysis Methods. *Micromachines*. 2017;8(3). doi:10.3390/mi8030083

315. Ganapathy U, Marrero J, Calhoun S, et al. Two enzymes with redundant fructose bisphosphatase activity sustain gluconeogenesis and virulence in Mycobacterium tuberculosis. *Nat Commun*. 2015;6(1):1-12. doi:10.1038/ncomms8912

316. Prosser GA, Larrouy-Maumus G, de Carvalho LPS. Metabolomic strategies for the identification of new enzyme functions and metabolic pathways. *EMBO Rep*. 2014;15(6):657-669. doi:10.15252/embr.201338283

317. Ryumin P, Brown J, Morris M, Cramer R. Protein identification using a nanoUHPLC-AP-MALDI MS/MS workflow with CID of multiply charged proteolytic peptides. *Int J Mass Spectrom*. 2017;416:20-28. doi:10.1016/j.ijms.2016.12.006

318. Venny 2.1.0. Accessed May 22, 2020. https://bioinfogp.cnb.csic.es/tools/venny/

319. Griffin JE, Gawronski JD, Dejesus MA, Ioerger TR, Akerley BJ, Sassetti CM. High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. *PLoS Pathog*. 2011;7(9):e1002251. doi:10.1371/journal.ppat.1002251

320. DeJesus MA, Gerrick ER, Xu W, et al. Comprehensive Essentiality Analysis of the Mycobacterium tuberculosis Genome via Saturating Transposon Mutagenesis. *mBio*. 2017;8(1). doi:10.1128/mBio.02133-16

321. Schulte J, Baumgart M, Bott M. Identification of the cAMP phosphodiesterase CpdA as novel key player in cAMP-dependent regulation in Corynebacterium glutamicum. *Mol Microbiol*. 2017;103(3):534-552. doi:10.1111/mmi.13574

322. Francis SH, Houslay MD, Conti M. Phosphodiesterase Inhibitors: Factors That Influence Potency, Selectivity, and Action. In: Francis SH, Conti M, Houslay MD, eds. *Phosphodiesterases as Drug Targets*. Handbook of Experimental Pharmacology. Springer; 2011:47-84. doi:10.1007/978-3-642-17969-3_2

323. A Candida albicans cyclic nucleotide phosphodiesterase: cloning and expression in Saccharomyces cerevisiae and biochemical characterization of the recombinant enzyme | Microbiology Society. Accessed June 3, 2020. https://www.microbiologyresearch.org/content/journal/micro/10.1099/13500872-140-7-1533

324. Rascón A, Soderling SH, Schaefer JB, Beavo JA. Cloning and characterization of a cAMP-specific phosphodiesterase (TbPDE2B) from Trypanosoma brucei. *Proc Natl Acad Sci*. 2002;99(7):4714-4719. doi:10.1073/pnas.002031599

325. phosphodiesterase activity of the HmsP EAL domain is required for negative regulation of biofilm formation in Yersinia pestis | FEMS Microbiology Letters | Oxford Academic. Accessed June 3, 2020. https://academic.oup.com/femsle/article/247/2/123/490858

326. Richter W. 3',5'-Cyclic nucleotide phosphodiesterases class III: Members, structure, and catalytic mechanism. *Proteins Struct Funct Bioinforma*. 2002;46(3):278-286. doi:10.1002/prot.10049

327. Webb B, Sali A. Comparative Protein Structure Modeling Using MODELLER. *Curr Protoc Bioinforma*. 2016;54:5.6.1-5.6.37. doi:10.1002/cpbi.3

328. Na H, Namgung B, Song WS, Yoon S. Structural and biochemical analyses of the metallo-β-lactamase fold protein YhfI from Bacillus subtilis. *Biochem Biophys Res Commun*. 2019;519(1):35-40. doi:10.1016/j.bbrc.2019.08.106

329. Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem*. 2004;25(13):1605-1612. doi:10.1002/jcc.20084

330. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 2004;32(5):1792-1797. doi:10.1093/nar/gkh340

331. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods*. 2017;14(6):587-589. doi:10.1038/nmeth.4285

332. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*. 2015;32(1):268-274. doi:10.1093/molbev/msu300

333. Huson DH, Scornavacca C. Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. *Syst Biol*. 2012;61(6):1061-1067. doi:10.1093/sysbio/sys062

334. Walsh TR. The emergence and implications of metallo-β-lactamases in Gram-negative bacteria. *Clin Microbiol Infect*. 2005;11:2-9. doi:10.1111/j.1469-0691.2005.01264.x

335. Walsh TR, Toleman MA, Poirel L, Nordmann P. Metallo-beta-lactamases: the quiet before the storm? *Clin Microbiol Rev*. 2005;18(2):306-325. doi:10.1128/CMR.18.2.306-325.2005

336. Somboro AM, Osei Sekyere J, Amoako DG, Essack SY, Bester LA. Diversity and Proliferation of Metallo-β-Lactamases: a Clarion Call for Clinically Effective Metallo-β-Lactamase Inhibitors. *Appl Environ Microbiol*. 2018;84(18). doi:10.1128/AEM.00698-18

337. Lee C, Park C. Bacterial Responses to Glyoxal and Methylglyoxal: Reactive Electrophilic Species. *Int J Mol Sci*. 2017;18(1). doi:10.3390/ijms18010169

338. Podzelinska K, He S-M, Wathier M, et al. Structure of PhnP, a phosphodiesterase of the carbon-phosphorus lyase pathway for phosphonate degradation. *J Biol Chem*. 2009;284(25):17216-17226. doi:10.1074/jbc.M808392200

339. Beaudoin GAW, Li Q, Bruner SD, Hanson AD. An unusual diphosphatase from the PhnP family cleaves reactive FAD photoproducts. *Biochem J*. 2018;475(1):261-272. doi:10.1042/BCJ20170817

340. Meini M-R, Llarrull LI, Vila AJ. Overcoming differences: The catalytic mechanism of metallo-β-lactamases. *FEBS Lett*. 2015;589(22):3419-3432. doi:10.1016/j.febslet.2015.08.015

341. Alderson RG, Barker D, Mitchell JBO. One origin for metallo-β-lactamase activity, or two? An investigation assessing a diverse set of reconstructed ancestral sequences based on a sample of phylogenetic trees. *J Mol Evol*. 2014;79(3-4):117-129. doi:10.1007/s00239-014-9639-7

342. Liu H, Naismith JH. An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol*. 2008;8(1):91. doi:10.1186/1472-6750-8-91

343. Leversen NA, de Souza GA, Målen H, Prasad S, Jonassen I, Wiker HG. Evaluation of signal peptide prediction algorithms for identification of mycobacterial signal peptides using sequence data from proteomic methods. *Microbiology*. 2009;155(Pt 7):2375-2383. doi:10.1099/mic.0.025270-0

344. de Souza GA, Leversen NA, Målen H, Wiker HG. Bacterial proteins with cleaved or uncleaved signal peptides of the general secretory pathway. *J Proteomics*. 2011;75(2):502-510. doi:10.1016/j.jprot.2011.08.016

345. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol*. 2019;37(4):420-423. doi:10.1038/s41587-019-0036-z

346. Bagos PG, Tsirigos KD, Liakopoulos TD, Hamodrakas SJ. Prediction of Lipoprotein Signal Peptides in Gram-Positive Bacteria with a Hidden Markov Model. *J Proteome Res*. 2008;7(12):5082-5093. doi:10.1021/pr800162c

347. Sander P, Rezwan M, Walker B, et al. Lipoprotein processing is required for virulence of Mycobacterium tuberculosis. *Mol Microbiol*. 2004;52(6):1543-1552. doi:10.1111/j.1365-2958.2004.04041.x

348. Tucci P, Portela M, Chetto CR, González-Sapienza G, Marín M. Integrative proteomic and glycoproteomic profiling of Mycobacterium tuberculosis culture filtrate. *PLoS ONE*. 2020;15(3). doi:10.1371/journal.pone.0221837

349. Feltcher ME, Sullivan JT, Braunstein M. Protein export systems of Mycobacterium tuberculosis: novel targets for drug development? *Future Microbiol*. 2010;5(10):1581-1597. doi:10.2217/fmb.10.112

350. Gibson DG. Enzymatic assembly of overlapping DNA fragments. *Methods Enzymol*. 2011;498:349-361. doi:10.1016/B978-0-12-385120-8.00015-2

351. Robinson PK. Enzymes: principles and biotechnological applications. *Essays Biochem*. 2015;59:1-41. doi:10.1042/bse0590001

352. NickelsJ JT. Purification, Characterization, and Kinetic Analyosfisa 55-kDa Form of Phosphatidylinositol 4-Kinase fromSaccharomyces cereuisiae. :8.

353. Drenth J. Crystallizing a Protein. In: Drenth J, ed. *Principles of Protein X-Ray Crystallography*. Springer; 2007:1-20. doi:10.1007/0-387-33746-6_1

354. Shevket SH, Gonzalez D, Cartwright JL, et al. The CcmC–CcmE interaction during cytochrome c maturation by System I is driven by protein–protein and not protein–heme contacts. *J Biol Chem*. 2018;293(43):16778-16790. doi:10.1074/jbc.RA118.005024

355. Singh A, Upadhyay V, Upadhyay AK, Singh SM, Panda AK. Protein recovery from inclusion bodies of Escherichia coli using mild solubilization process. *Microb Cell Factories*. 2015;14(1):41. doi:10.1186/s12934-015-0222-8

356. Gecchele E, Merlin M, Brozzetti A, Falorni A, Pezzotti M, Avesani L. A Comparative Analysis of Recombinant Protein Expression in Different Biofactories: Bacteria, Insect Cells and Plant Systems. *J Vis Exp JoVE*. 2015;(97). doi:10.3791/52459

357. Jia B, Jeon CO. High-throughput recombinant protein expression in Escherichia coli: current status and future perspectives. *Open Biol*. 6(8):160196. doi:10.1098/rsob.160196

358. Rosano GL, Ceccarelli EA. Recombinant protein expression in Escherichia coli: advances and challenges. *Front Microbiol*. 2014;5. doi:10.3389/fmicb.2014.00172

359. Savitsky P, Bray J, Cooper CDO, et al. High-throughput production of human proteins for crystallization: The SGC experience. *J Struct Biol*. 2010;172(1):3-13. doi:10.1016/j.jsb.2010.06.008

360. Brückner R, Titgemeyer F. Carbon catabolite repression in bacteria: choice of the carbon source and autoregulatory limitation of sugar utilization. *FEMS Microbiol Lett*. 2002;209(2):141-148. doi:10.1111/j.1574-6968.2002.tb11123.x

361. Aidelberg G, Towbin BD, Rothschild D, Dekel E, Bren A, Alon U. Hierarchy of non-glucose sugars in Escherichia coli. *BMC Syst Biol*. 2014;8:133. doi:10.1186/s12918-014-0133-z

362. Utsumi R, Noda M, Kawamukai M, Komano T. Effects of *cya* and *cid* Mutations on Cell Growth in *Escherichia coli*. *Agric Biol Chem*. 1987;51(2):465-469. doi:10.1080/00021369.1987.10868035

363. Wagner S, Klepsch MM, Schlegel S, et al. Tuning Escherichia coli for membrane protein overexpression. *Proc Natl Acad Sci*. 2008;105(38):14371-14376. doi:10.1073/pnas.0804090105

364. Pan SH, Malcolm BA. Reduced background expression and improved plasmid stability with pET vectors in BL21 (DE3). *BioTechniques*. 2000;29(6):1234-1238. doi:10.2144/00296st03

365. Wagner S, Baars L, Ytterberg AJ, et al. Consequences of Membrane Protein Overexpression in Escherichia coli. *Mol Cell Proteomics*. 2007;6(9):1527-1550. doi:10.1074/mcp.M600431-MCP200

366. Dumon-Seignovert L, Cariot G, Vuillard L. The toxicity of recombinant proteins in Escherichia coli: a comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3). *Protein Expr Purif*. 2004;37(1):203-206. doi:10.1016/j.pep.2004.04.025

367. Nishihara K, Kanemori M, Kitagawa M, Yanagi H, Yura T. Chaperone coexpression plasmids: differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in Escherichia coli. *Appl Environ Microbiol*. 1998;64(5):1694-1699.

368. Nishihara K, Kanemori M, Yanagi H, Yura T. Overexpression of Trigger Factor Prevents Aggregation of Recombinant Proteins in Escherichia coli. *Appl Environ Microbiol*. 2000;66(3):884-889.

369. Booth WT, Schlachter CR, Pote S, et al. Impact of an N-terminal Polyhistidine Tag on Protein Thermal Stability. *ACS Omega*. 2018;3(1):760-768. doi:10.1021/acsomega.7b01598

370. Optimization of membrane protein overexpression and purification using GFP fusions | Nature Methods. Accessed June 15, 2020. https://www.nature.com/articles/nmeth0406-303

371. Szybalski W, Kim SC, Hasan N, Podhajska AJ. Class-IIS restriction enzymes — a review. *Gene*. 1991;100:13-26. doi:10.1016/0378-1119(91)90345-C

372. Bernard P, Gabant P, Bahassi EM, Couturier M. Positive-selection vectors using the F plasmid ccdB killer gene. *Gene*. 1994;148(1):71-74. doi:10.1016/0378-1119(94)90235-6

373. Meissner PS, Sisk WP, Berman ML. Bacteriophage lambda cloning system for the construction of directional cDNA libraries. *Proc Natl Acad Sci U S A*. 1987;84(12):4171-4175.

374. Kram KE, Finkel SE. Rich Medium Composition Affects Escherichia coli Survival, Glycation, and Mutation Frequency during Long-Term Batch Culture. *Appl Environ Microbiol*. 2015;81(13):4442-4450. doi:10.1128/AEM.00722-15

375. Arnold FM, Hohl M, Remm S, et al. A uniform cloning platform for mycobacterial genetics and protein production. *Sci Rep*. 2018;8(1):9539. doi:10.1038/s41598-018-27687-5

376. Raska M. Novel Modification of Growth Medium Enables Efficient E. coli Expression and Simple Purification of an Endotoxin-Free Recombinant Murine hsp70 Protein. *J Microbiol Biotechnol*. Published online 2009. doi:10.4014/jmb.0810.570

377. Yan X, Shi Q, Bracher A, et al. GroEL Ring Separation and Exchange in the Chaperonin Reaction. *Cell*. 2018;172(3):605-617.e11. doi:10.1016/j.cell.2017.12.010

378. Bornhorst JA, Falke JJ. [16] Purification of Proteins Using Polyhistidine Affinity Tags. *Methods Enzymol*. 2000;326:245-254.

379. Chang CCH, Li C, Webb GI, Tey B, Song J, Ramanan RN. Periscope: quantitative prediction of soluble protein expression in the periplasm of Escherichia coli. *Sci Rep*. 2016;6(1):21844. doi:10.1038/srep21844

380. Gu S, Chen J, Dobos KM, Bradbury EM, Belisle JT, Chen X. Comprehensive proteomic profiling of the membrane constituents of a Mycobacterium tuberculosis strain. *Mol Cell Proteomics MCP*. 2003;2(12):1284-1296. doi:10.1074/mcp.M300060-MCP200

381. Berkmen M. Production of disulfide-bonded proteins in Escherichia coli. *Protein Expr Purif*. 2012;82(1):240-251. doi:10.1016/j.pep.2011.10.009

382. Gopal GJ, Kumar A. Strategies for the Production of Recombinant Protein in Escherichia coli. *Protein J*. 2013;32(6):419-425. doi:10.1007/s10930-013-9502-5

383. Malik A. Protein fusion tags for efficient expression and purification of recombinant proteins in the periplasmic space of E. coli. *3 Biotech*. 2016;6(1):44. doi:10.1007/s13205-016-0397-7

384. Rietsch A, Bessette P, Georgiou G, Beckwith J. Reduction of the periplasmic disulfide bond isomerase, DsbC, occurs by passage of electrons from cytoplasmic thioredoxin. *J Bacteriol*. 1997;179(21):6602-6608. doi:10.1128/jb.179.21.6602-6608.1997

385. Kadokura H, Beckwith J. Detecting Folding Intermediates of a Protein as It Passes Through the Bacterial Translocation Channel. *Cell*. 2009;138(6):1164-1173. doi:10.1016/j.cell.2009.07.030

386. Latifi AM, Khajeh K, Farnoosh G, Hassanpour K, Khodi S. The Cytoplasmic and Periplasmic Expression Levels and Folding of Organophosphorus Hydrolase Enzyme in Escherichia coli. *Jundishapur J Microbiol*. 2015;8(12). doi:10.5812/jjm.17790

387. Sockolosky JT, Szoka FC. Periplasmic production via the pET expression system of soluble, bioactive human growth hormone. *Protein Expr Purif*. 2013;87(2):129-135. doi:10.1016/j.pep.2012.11.002

388. Bashiri G, Baker EN. Production of recombinant proteins in Mycobacterium smegmatis for structural and functional studies. *Protein Sci Publ Protein Soc*. 2015;24(1):1-10. doi:10.1002/pro.2584

389. Noens EE, Williams C, Anandhakrishnan M, Poulsen C, Ehebauer MT, Wilmanns M. Improved mycobacterial protein production using a Mycobacterium smegmatis groEL1ΔC expression strain. *BMC Biotechnol*. 2011;11:27. doi:10.1186/1472-6750-11-27

390. Santucci P, Point V, Poncin I, et al. LipG a bifunctional phospholipase/thioesterase involved in mycobacterial envelope remodeling. *Biosci Rep*. 2018;38(6). doi:10.1042/BSR20181953

391. Jaworek MW, Möbitz S, Gao M, Winter R. Stability of the chaperonin system GroEL–GroES under extreme environmental conditions. *Phys Chem Chem Phys*. 2020;22(6):3734-3743. doi:10.1039/C9CP06468K

392. Kim YE, Hipp MS, Bracher A, Hayer-Hartl M, Hartl FU. Molecular chaperone functions in protein folding and proteostasis. *Annu Rev Biochem*. 2013;82:323-355. doi:10.1146/annurev-biochem-060208-092442

393. Jiang L, Wang Z, Li Y, et al. Relationship Between Surface Hydrophobicity and Structure of Soy Protein Isolate Subjected to Different Ionic Strength. *Int J Food Prop*. 2015;18(5):1059-1074. doi:10.1080/10942912.2013.865057

394. Xia F, Nagrath D, Cramer SM. Effect of pH changes on water release values in hydrophobic interaction chromatographic systems. *J Chromatogr A*. 2005;1079(1):229-235. doi:10.1016/j.chroma.2005.04.005

395. Lee C, Kim H, Bardwell JCA. Electrostatic interactions are important for chaperone–client interaction in vivo. *Microbiology*. 2018;164(7):992-997. doi:10.1099/mic.0.000676

396. Marchenko NYu, Sikorskaya EV, Marchenkov VV, Kashparov IA, Semisotnov GV. Affinity chromatography of chaperones based on denatured proteins: Analysis of cell lysates of different origin. *Protein Expr Purif*. 2016;119:117-123. doi:10.1016/j.pep.2015.11.020

397. Kampinga HH. Chaperones in Preventing Protein Denaturation in Living Cells and Protecting Against Cellular Stress. In: Starke K, Gaestel M, eds. *Molecular Chaperones in Health and Disease*. Handbook of Experimental Pharmacology. Springer; 2006:1-42. doi:10.1007/3-540-29717-0_1

398. Łapińska U, L. Saar K, V. Yates E, et al. Gradient-free determination of isoelectric points of proteins on chip. *Phys Chem Chem Phys*. 2017;19(34):23060-23067. doi:10.1039/C7CP01503H

399. Widmann M, Trodler P, Pleiss J. The Isoelectric Region of Proteins: A Systematic Analysis. *PLOS ONE*. 2010;5(5):e10546. doi:10.1371/journal.pone.0010546

400. Yang J-P, Cirico T, Katzen F, Peterson TC, Kudlicki W. Cell-free synthesis of a functional G protein-coupled receptor complexed with nanometer scale bilayer discs. *BMC Biotechnol*. 2011;11:57. doi:10.1186/1472-6750-11-57

401. Youn Jun S, Hyeon Kang S, Lee K-H. Continuous-exchange cell-free protein synthesis using PCR-generated DNA and an RNase E-deficient extract. *BioTechniques*. 2008;44(3):387-391. doi:10.2144/000112690

402. A Continuous-Exchange Cell-Free Protein Synthesis System Based on Extracts from Cultured Insect Cells. Accessed June 3, 2020. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0096635

403. Pähler A, Hendrickson WA, Kolks MA, Argaraña CE, Cantor CR. Characterization and crystallization of core streptavidin. *J Biol Chem*. 1987;262(29):13933-13937.

404. Schmidt TG, Skerra A. The Strep -tag system for one-step purification and high-affinity detection or capturing of proteins. *Nat Protoc*. 2007;2(6):1528-1535. doi:10.1038/nprot.2007.209

405. Dundas C, Demonte D, Park S. Streptavidin-biotin technology: Improvements and innovations in chemical and biological applications. *Appl Microbiol Biotechnol*. 2013;97. doi:10.1007/s00253-013-5232-z

406. Bratthauer GL. The avidin-biotin complex (ABC) method and other avidin-biotin binding methods. *Methods Mol Biol Clifton NJ*. 2010;588:257-270. doi:10.1007/978-1-59745-324-0_26

407. Rinke de Wit T, Bekelie S, Osland A, et al. Mycobacteria contain two groEL genes: The second Mycobacterium leprae groEL gene is arranged in an operon with groES. *Mol Microbiol*. 1992;6:1995-2007. doi:10.1111/j.1365-2958.1992.tb01372.x

408. Ojha A, Anand M, Bhatt A, Kremer L, Jacobs WR, Hatfull GF. GroEL1: a dedicated chaperone involved in mycolic acid biosynthesis during biofilm formation in mycobacteria. *Cell*. 2005;123(5):861-873. doi:10.1016/j.cell.2005.09.012

409. Bandyopadhyay B, Gupta TD, Roy D, Gupta SKD. DnaK Dependence of the Mycobacterial Stress-Responsive Regulator HspR Is Mediated through Its Hydrophobic C-Terminal Tail. *J Bacteriol*. 2012;194(17):4688-4697. doi:10.1128/JB.00415-12

410. Fay A, Glickman MS. An Essential Nonredundant Role for Mycobacterial DnaK in Native Protein Folding. *PLoS Genet*. 2014;10(7). doi:10.1371/journal.pgen.1004516

411. Gill WP, Harik NS, Whiddon MR, Liao RP, Mittler JE, Sherman DR. A replication clock for Mycobacterium tuberculosis. *Nat Med*. 2009;15(2):211-214. doi:10.1038/nm.1915

412. Caglar MU, Houser JR, Barnhart CS, et al. The E. coli molecular phenotype under different growth conditions. *Sci Rep*. 2017;7(1):45303. doi:10.1038/srep45303

413. Wu G, Sun Y, Qu W, et al. Application of GFAT as a novel selection marker to mediate gene expression. *PloS One*. 2011;6(2):e17082. doi:10.1371/journal.pone.0017082

414. Wah Tang P, San Chua P, Kee Chong S, et al. A Review of Gene Knockout Strategies for Microbial Cells. *Recent Pat Biotechnol*. 2015;9(3):176-197.

415. Jiang Y, Chen B, Duan C, Sun B, Yang J, Yang S. Multigene Editing in the Escherichia coli Genome via the CRISPR-Cas9 System. *Appl Environ Microbiol*. 2015;81(7):2506-2514. doi:10.1128/AEM.04023-14

416. Sévin DC, Fuhrer T, Zamboni N, Sauer U. Nontargeted in vitro metabolomics for high-throughput identification of novel enzymes in Escherichia coli. *Nat Methods*. 2017;14(2):187-194. doi:10.1038/nmeth.4103

417. Sorrentino F, Rio RG del, Zheng X, et al. Development of an Intracellular Screen for New Compounds Able To Inhibit Mycobacterium tuberculosis Growth in Human Macrophages. *Antimicrob Agents Chemother*. 2016;60(1):640-645. doi:10.1128/AAC.01920-15

418. Radhakrishnan A, M. Furze C, Syed Ahangar M, Fullam E. A GFP-strategy for efficient recombinant protein overexpression and purification in Mycobacterium smegmatis. *RSC Adv*. 2018;8(58):33087-33095. doi:10.1039/C8RA06237D

419. Giepmans BNG, Adams SR, Ellisman MH, Tsien RY. The Fluorescent Toolbox for Assessing Protein Location and Function. *Science*. 2006;312(5771):217-224. doi:10.1126/science.1124618

420. Bhatnagar NB, Bhatnagar R, Venkitasubramanian TA. Characterization and metabolism of cyclic guanosine 3’5’-monophosphate in Mycobacteriumsmegmatis. *Biochem Biophys Res Commun*. 1984;121(2):634-640. doi:10.1016/0006-291X(84)90229-8

421. Mehta M, Singh A. Mycobacterium tuberculosis WhiB3 maintains redox homeostasis and survival in response to reactive oxygen and nitrogen species. *Free Radic Biol Med*. 2019;131:50-58. doi:10.1016/j.freeradbiomed.2018.11.032

422. Bai G, Knapp GS, McDonough KA. Cyclic AMP signaling in mycobacteria: redirecting the conversation with a common currency. *Cell Microbiol*. 2011;13(3):349-358. doi:10.1111/j.1462-5822.2010.01562.x

423. Bush K, Fisher JF. Epidemiological expansion, structural studies, and clinical challenges of new β-lactamases from gram-negative bacteria. *Annu Rev Microbiol*. 2011;65:455-478. doi:10.1146/annurev-micro-090110-102911

424. de SENY D, Prosperi-Meys C, Bebrone C, et al. Mutational analysis of the two zinc-binding sites of the Bacillus cereus 569/H/9 metallo-β-lactamase. *Biochem J*. 2002;363(3):687-696. doi:10.1042/bj3630687

425. Ehrt S, Guo XV, Hickey CM, et al. Controlling gene expression in mycobacteria with anhydrotetracycline and Tet repressor. *Nucleic Acids Res*. 2005;33(2):e21. doi:10.1093/nar/gni013

426. Williams KJ, Joyce G, Robertson BD. Improved Mycobacterial Tetracycline Inducible Vectors. *Plasmid*. 2010;64(2):69-73. doi:10.1016/j.plasmid.2010.04.003

427. Saito K, Warrier T, Somersan-Karakaya S, et al. Rifamycin action on RNA polymerase in antibiotic-tolerant Mycobacterium tuberculosis results in differentially detectable populations. *Proc Natl Acad Sci*. 2017;114(24):E4832-E4840. doi:10.1073/pnas.1705385114

428. White RJ, Lancini GC, Silvestri LG. Mechanism of Action of Rifampin on Mycobacterium smegmatis. *J Bacteriol*. 1971;108(2):737-741.

429. Ruiz P, Rodríguez-Cano F, Zerolo FJ, Casal M. Investigation of the in vitro activity of streptomycin against Mycobacterium tuberculosis. *Microb Drug Resist Larchmt N*. 2002;8(2):147-149. doi:10.1089/107662902760190707

430. Yamori S, Ichiyama S, Shimokata K, Tsukamura M. Bacteriostatic and bactericidal activity of antituberculosis drugs against Mycobacterium tuberculosis, Mycobacterium avium-Mycobacterium

intracellulare complex and Mycobacterium kansasii in different growth phases. *Microbiol Immunol*. 1992;36(4):361-368. doi:10.1111/j.1348-0421.1992.tb02035.x

431. Belanger AE, Besra GS, Ford ME, et al. The embAB genes of Mycobacterium avium encode an arabinosyl transferase involved in cell wall arabinan biosynthesis that is the target for the antimycobacterial drug ethambutol. *Proc Natl Acad Sci U S A*. 1996;93(21):11919-11924. doi:10.1073/pnas.93.21.11919

432. Takayama K, Kilburn JO. Inhibition of synthesis of arabinogalactan by ethambutol in Mycobacterium smegmatis. *Antimicrob Agents Chemother*. 1989;33(9):1493-1499. doi:10.1128/AAC.33.9.1493

433. Telenti A, Philipp WJ, Sreevatsan S, et al. The emb operon, a gene cluster of Mycobacterium tuberculosis involved in resistance to ethambutol. *Nat Med*. 1997;3(5):567-570. doi:10.1038/nm0597-567

434. Maitra A, Munshi T, Healy J, et al. Cell wall peptidoglycan in Mycobacterium tuberculosis: An Achilles' heel for the TB-causing pathogen. *FEMS Microbiol Rev*. 2019;43(5):548-575. doi:10.1093/femsre/fuz016

435. McCallum AD, Sloan DJ. The importance of clinical pharmacokinetic–pharmacodynamic studies in unraveling the determinants of early and late tuberculosis outcomes. *Int J Pharmacokinet*. 2017;2(3):195-212. doi:10.4155/ipk-2017-0004

436. Schroeder A, Mueller O, Stocker S, et al. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol*. 2006;7(1):3. doi:10.1186/1471-2199-7-3

437. Sharma R, Zaveri A, Gopalakrishnapai J, Thiruneelakantan S, Varshney U, Visweswariah SS. Paralogous cAMP Receptor Proteins in Mycobacterium smegmatis Show Biochemical and Functional Divergence. *Biochemistry*. 2014;53(49):7765-7776. doi:10.1021/bi500924v

438. Wang Q, Boshoff HIM, Harrison JR, et al. PE/PPE proteins mediate nutrient transport across the outer membrane of Mycobacterium tuberculosis. *Science*. 2020;367(6482):1147-1151. doi:10.1126/science.aav5912

439. Chatrath S, Gupta VK, Dixit A, Garg LC. The Rv1651c-encoded PE-PGRS30 protein expressed in Mycobacterium smegmatis exhibits polar localization and modulates its growth profile. *FEMS Microbiol Lett*. 2011;322(2):194-199. doi:10.1111/j.1574-6968.2011.02354.x

440. Bush MJ. The actinobacterial WhiB-like (Wbl) family of transcription factors. *Mol Microbiol*. 2018;110(5):663-676. doi:10.1111/mmi.14117

441. Smith LJ, Stapleton MR, Buxton RS, Green J. Structure-function relationships of the Mycobacterium tuberculosis transcription factor WhiB1. *PloS One*. 2012;7(7):e40407. doi:10.1371/journal.pone.0040407

442. Cumming BM, Rahman MA, Lamprecht DA, et al. Mycobacterium tuberculosis arrests host cycle at the G1/S transition to establish long term infection. *PLOS Pathog*. 2017;13(5):e1006389. doi:10.1371/journal.ppat.1006389

443. Stapleton M, Haq I, Hunt DM, et al. Mycobacterium tuberculosis cAMP Receptor Protein (Rv3676) Differs from the Escherichia coli Paradigm in Its cAMP Binding and DNA Binding Properties and Transcription Activation Properties. *J Biol Chem*. 2010;285(10):7016-7027. doi:10.1074/jbc.M109.047720

444. Hards K, Rodriguez SM, Cairns C, Cook GM. Alternate quinone coupling in a new class of succinate dehydrogenase may potentiate mycobacterial respiratory control. *FEBS Lett*. 2019;593(5):475-486. doi:10.1002/1873-3468.13330

445. Hartman T, Weinrick B, Vilchèze C, et al. Succinate dehydrogenase is the regulator of respiration in Mycobacterium tuberculosis. *PLoS Pathog*. 2014;10(11):e1004510. doi:10.1371/journal.ppat.1004510

446. Rao SPS, Alonso S, Rand L, Dick T, Pethe K. The protonmotive force is required for maintaining ATP homeostasis and viability of hypoxic, nonreplicating Mycobacterium tuberculosis. *Proc Natl Acad Sci*. 2008;105(33):11945-11950. doi:10.1073/pnas.0711697105

447. Mascolo L, Bald D. Cytochrome bd in Mycobacterium tuberculosis: A respiratory chain protein involved in the defense against antibacterials. *Prog Biophys Mol Biol*. 2020;152:55-63. doi:10.1016/j.pbiomolbio.2019.11.002

448. Koul A, Vranckx L, Dhar N, et al. Delayed bactericidal response of Mycobacterium tuberculosis to bedaquiline involves remodelling of bacterial metabolism. *Nat Commun*. 2014;5(1):3369. doi:10.1038/ncomms4369

449. Kalia NP, Hasenoehrl EJ, Ab Rahman NB, et al. Exploiting the synthetic lethality between terminal respiratory oxidases to kill Mycobacterium tuberculosis and clear host infection. *Proc Natl Acad Sci U S A*. 2017;114(28):7426-7431. doi:10.1073/pnas.1706139114

450. Vilchèze C, Weinrick B, Leung LW, Jacobs WR. Plasticity of Mycobacterium tuberculosis NADH dehydrogenases and their role in virulence. *Proc Natl Acad Sci*. 2018;115(7):1599-1604. doi:10.1073/pnas.1721545115

451. Miesel L, Weisbrod TR, Marcinkeviciene JA, Bittman R, Jacobs WR. NADH dehydrogenase defects confer isoniazid resistance and conditional lethality in Mycobacterium smegmatis. *J Bacteriol*. 1998;180(9):2459-2467.

452. Lofthouse EK, Wheeler PR, Beste DJV, et al. Systems-based approaches to probing metabolic variation within the Mycobacterium tuberculosis complex. *PloS One*. 2013;8(9):e75913. doi:10.1371/journal.pone.0075913

453. Borah K, Beyß M, Theorell A, et al. Intracellular Mycobacterium tuberculosis Exploits Multiple Host Nitrogen Sources during Growth in Human Macrophages. *Cell Rep*. 2019;29(11):3580-3591.e4. doi:10.1016/j.celrep.2019.11.037

454. Gokulan K, Rupp B, Pavelka MS, Jacobs WR, Sacchettini JC. Crystal Structure of Mycobacterium tuberculosis Diaminopimelate Decarboxylase, an Essential Enzyme in Bacterial Lysine Biosynthesis. *J Biol Chem*. 2003;278(20):18588-18596. doi:10.1074/jbc.M301549200

455. Usha V, Lloyd AJ, Lovering AL, Besra GS. Structure and function of Mycobacterium tuberculosis meso-diaminopimelic acid (DAP) biosynthetic enzymes. *FEMS Microbiol Lett*. 2012;330(1):10-16. doi:10.1111/j.1574-6968.2012.02527.x

456. Coldham NG, Webber M, Woodward MJ, Piddock LJV. A 96-well plate fluorescence assay for assessment of cellular permeability and active efflux in Salmonella enterica serovar Typhimurium and Escherichia coli. *J Antimicrob Chemother*. 2010;65(8):1655-1663. doi:10.1093/jac/dkq169

457. Watanabe S, Zimmermann M, Goodwin MB, Sauer U, Barry CE, Boshoff HI. Fumarate Reductase Activity Maintains an Energized Membrane in Anaerobic Mycobacterium tuberculosis. Deretic V, ed. *PLoS Pathog*. 2011;7(10):e1002287. doi:10.1371/journal.ppat.1002287

458. Stone MRL, Butler MS, Phetsang W, Cooper MA, Blaskovich MAT. Fluorescent Antibiotics: New Research Tools to Fight Antibiotic Resistance. *Trends Biotechnol*. 2018;36(5):523-536. doi:10.1016/j.tibtech.2018.01.004

459. Greenwood DJ, Santos MSD, Huang S, et al. Subcellular antibiotic visualization reveals a dynamic drug reservoir in infected macrophages. *Science*. 2019;364(6447):1279-1282. doi:10.1126/science.aat9689

460. Andries K, Verhasselt P, Guillemont J, et al. A diarylquinoline drug active on the ATP synthase of Mycobacterium tuberculosis. *Science*. 2005;307(5707):223-227. doi:10.1126/science.1106753

461. Cholo MC, Mothiba MT, Fourie B, Anderson R. Mechanisms of action and therapeutic efficacies of the lipophilic antimycobacterial agents clofazimine and bedaquiline. *J Antimicrob Chemother*. 2017;72(2):338-353. doi:10.1093/jac/dkw426

462. Peyrusson F, Varet H, Nguyen TK, et al. Intracellular Staphylococcus aureus persisters upon antibiotic exposure. *Nat Commun*. 2020;11(1):2200. doi:10.1038/s41467-020-15966-7

463. Bald D, Koul A. Respiratory ATP synthesis: the new generation of mycobacterial drug targets? *FEMS Microbiol Lett*. 2010;308(1):1-7. doi:10.1111/j.1574-6968.2010.01959.x

464. Soni DK, Dubey SK, Bhatnagar R. ATP-binding cassette (ABC) import systems of Mycobacterium tuberculosis: target for drug and vaccine development. *Emerg Microbes Infect*. 2020;9(1):207-220. doi:10.1080/22221751.2020.1714488

465. O'Malley T, Alling T, Early JV, et al. Imidazopyridine Compounds Inhibit Mycobacterial Growth by Depleting ATP Levels. *Antimicrob Agents Chemother*. 2018;62(6). doi:10.1128/AAC.02439-17

466. Rao SPS, Alonso S, Rand L, Dick T, Pethe K. The protonmotive force is required for maintaining ATP homeostasis and viability of hypoxic, nonreplicating Mycobacterium tuberculosis. *Proc Natl Acad Sci*. 2008;105(33):11945-11950. doi:10.1073/pnas.0711697105

467. Bai G, Schaak DD, Smith EA, McDonough KA. Dysregulation of serine biosynthesis contributes to the growth defect of a Mycobacterium tuberculosis crp mutant. *Mol Microbiol*. 2011;82(1):180-198. doi:10.1111/j.1365-2958.2011.07806.x

468. Eoh H, Rhee KY. Methylcitrate cycle defines the bactericidal essentiality of isocitrate lyase for survival of Mycobacterium tuberculosis on fatty acids. *Proc Natl Acad Sci U S A*. 2014;111(13):4976-4981. doi:10.1073/pnas.1400390111

469. Tiwari S, van Tonder AJ, Vilchèze C, et al. Arginine-deprivation–induced oxidative damage sterilizes Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A*. 2018;115(39):9779-9784. doi:10.1073/pnas.1808874115

470. Mizrahi V, Warner DF. Death of Mycobacterium tuberculosis by l-arginine starvation. *Proc Natl Acad Sci U S A*. 2018;115(39):9658-9660. doi:10.1073/pnas.1813587115

471. Jansen RS, Mandyoli L, Hughes R, et al. Aspartate aminotransferase Rv3722c governs aspartate-dependent nitrogen metabolism in Mycobacterium tuberculosis. *Nat Commun*. 2020;11(1):1960. doi:10.1038/s41467-020-15876-8

472. Zimmermann M, Kogadeeva M, Gengenbacher M, et al. Integration of Metabolomics and Transcriptomics Reveals a Complex Diet of Mycobacterium tuberculosis during Early Macrophage Infection. *mSystems*. 2017;2(4). doi:10.1128/mSystems.00057-17

473. Hogan A, Jewell B, Sherrard-Smith E, et al. *Report 19: The Potential Impact of the COVID-19 Epidemic on HIV, TB and Malaria in Low- and Middle-Income Countries*. Imperial College London; 2020. doi:10.25561/78670

474. Antonova AV, Gryadunov DA, Zimenkov DV. Molecular Mechanisms of Drug Tolerance in Mycobacterium tuberculosis. *Mol Biol*. 2018;52(3):372-384. doi:10.1134/S0026893318030020

475. Kieser KJ, Rubin EJ. How sisters grow apart: mycobacterial growth and division. *Nat Rev Microbiol*. 2014;12(8):550-562. doi:10.1038/nrmicro3299

476. Baker JJ, Dechow SJ, Abramovitch RB. Acid Fasting: Modulation of Mycobacterium tuberculosis Metabolism at Acidic pH. *Trends Microbiol*. 2019;27(11):942-953. doi:10.1016/j.tim.2019.06.005

477. Beites T, O'Brien K, Tiwari D, et al. Plasticity of the Mycobacterium tuberculosis respiratory chain and its impact on tuberculosis drug development. *Nat Commun*. 2019;10(1):1-12. doi:10.1038/s41467-019-12956-2

478. Rohde KH, Abramovitch RB, Russell DG. Mycobacterium tuberculosis invasion of macrophages: linking bacterial gene expression to environmental cues. *Cell Host Microbe*. 2007;2(5):352-364. doi:10.1016/j.chom.2007.09.006

479. Larsson C, Luna B, Ammerman NC, Maiga M, Agarwal N, Bishai WR. Gene Expression of Mycobacterium tuberculosis Putative Transcription Factors whiB1-7 in Redox Environments. *PLoS ONE*. 2012;7(7). doi:10.1371/journal.pone.0037516

480. Hett EC, Rubin EJ. Bacterial growth and cell division: a mycobacterial perspective. *Microbiol Mol Biol Rev MMBR*. 2008;72(1):126-156, table of contents. doi:10.1128/MMBR.00028-07

481. Catalão MJ, Filipe SR, Pimentel M. Revisiting Anti-tuberculosis Therapeutic Strategies That Target the Peptidoglycan Structure and Synthesis. *Front Microbiol*. 2019;10. doi:10.3389/fmicb.2019.00190