

The Universal Stress Proteins of Bacteria

Dominic Bradley

Department of Life Sciences

*A thesis submitted for the degree of Doctor of Philosophy and the
Diploma of Imperial College London*

Declaration of Originality

I, Dominic Bradley, declare that this thesis is my own work and has not been submitted in any form for another degree or diploma at any university or other institute of tertiary education. Information derived from the published and unpublished work of others has been acknowledged in the text and a list of references is given in the bibliography.

Abstract

Universal stress proteins (USPs) are a widespread and abundant protein family often linked to survival during stress. However, their exact biochemical and cellular roles are incompletely understood. *Mycobacterium tuberculosis* (*Mtb*) has 10 USPs, of which Rv1636 appears to be unique in its domain structure and being the only USP conserved in *M. leprae*. Over-expression of Rv1636 in *M. smegmatis* indicated that this protein does not share the growth arrest phenotype of another *Mtb* USP, Rv2623, suggesting distinct roles for the *Mtb* USPs.

Purified Rv1636 was shown to have novel nucleotide binding capabilities when subjected to UV crosslinking. A range of site-directed mutants of Rv1636 were produced, including mutations within a predicted nucleotide binding motif, with the aim of identifying and characterising key residues within the Rv1636 protein. Further putative biochemical activities, including nucleotide triphosphatase, nucleotidylation and auto-phosphorylation were also investigated *in vitro*; however Rv1636 could not be shown to definitively possess these activities, raising the possibility that addition factors may be present *in vivo*.

Bioinformatic analysis of Rv1636 has provided an in-depth look at the protein. The crystal structure of Rv1636 shows a strand swapped dimer conformation that appears to block the predicted nucleotide binding site, providing a possible reason for the low NTPase activity previously observed. Truncated Rv1636 constructs were successfully generated, in which the strand swapped dimer was disrupted, and subjected to biophysical analysis, including analytical ultracentrifugation and size exclusion chromatography combined with multi-angle light scattering.

Previous *Mtb* single-USP mutants are known to have no phenotype under a range of stress conditions. For this reason the *P. aeruginosa* USP PA3309, which does possess a stress survival phenotype, was also investigated. This provided the opportunity to investigate the role of USPs and their putative nucleotide binding motif *in vivo*. Site-directed mutants of PA3309 were generated to investigate the role of the nucleotide binding motif *in vivo*. It remains to be determined if the survival defect observed for Δ PA3309 strain can be complemented with these mutants as the vector system used in these experiments proved unable to integrate into the attB site of the genome.

Through the analysis of the USPs from mycobacteria and *Pseudomonas*, the aim was to elucidate a greater understanding of the role of Rv1636 in *Mtb* and the role of USPs in general. The bioinformatic and biochemical analyses of USPs, in addition to the site directed mutants generated as a result of this work, will provide a strong foundation for future studies.

Acknowledgements

Firstly I would pass my enormous appreciation to Huw Williams for his guidance and supervision. In the Williams' group I express great thanks to Ben Ryall, Suzie Hingley-Wilson, Javier Levia and Kathryn Loughheed for their continued advice and support. Special mention goes to Kathryn for her meticulous reading of this thesis.

I would like to show gratitude to Steve Smerdon and especially Simon Pennell for assisting with the structural studies undertaken at the National Institute for Medical Research. I am also grateful to Bob Coutts and Martin Buck for their supervision in during my MRes in Biochemical Research, which gave me my first experiences of academic research.

Others to I would like acknowledge within the Imperial College Life Sciences Department include The Flying Ducks, Falcons and Dragons football team and their supporters. Members Mike, Nick, Thorsten, Luis, Rustem and Diogo brought a taste of success to the department. Ellen, Sophie, Maija, Lucy and Lei were the best supporters we could wish for.

Wilson House provided a balanced environment and with four years as a sub-warden, there was always excitement waiting on returning home. From parties to 4am wake-up calls I enjoyed it all. Kudos to Mike Bluck, Laura Nicolaou, Shweta Singh and Alex Mottura, as they were brilliant and we enjoyed the challenge of keeping Wilson an awesome place to live.

Through Diogo and Lara Geraldes I was afforded the opportunity to visit Portugal each year during my PhD and as a result Portugal is a country full of friends. I give credit to my kickboxing instructor Gustarve, the Scuba Diving Club, IC Tennis, IC Squash and Wilson House FC for 4 years of sporting fun and success.

Finally, I would like to dedicate this thesis to my wonderful family.

Table of Contents

Declaration of Originality	1
Abstract.....	2
Acknowledgements.....	3
Table of Contents.....	4
List of Figures.....	11
List of Tables	14
Abbreviations.....	15
Chapter 1 - Introduction	16
1.1 Scope of the Introduction	17
1.2 The <i>Mycobacterium</i> Genus	17
1.3 <i>Mycobacterium tuberculosis</i>	18
1.3.1 The <i>Mtb</i> Bacterium	18
1.3.2 The <i>Mtb</i> Cell Wall.....	19
1.4 Tuberculosis	20
1.4.1 History of TB.....	21
1.4.2 Distribution of TB	21
1.4.3 HIV associated TB.....	23
1.4.4 Diagnosis of TB.....	23
1.4.5 Drug Resistance and its Development in TB	25
1.4.6 Drug discovery	26
1.4.7 TB Vaccine Development	26
1.5 Persistence of <i>Mtb</i> within the Host	27
1.5.1 Survival within Macrophage	28
1.5.2 The Granuloma	29
1.5.3 Latency, and Non-Replicating Persistence.....	32
1.5.4 Reactivation, Resuscitation and Disease	33
1.6 Models to Study <i>Mtb</i>	34
1.6.1 <i>In vivo</i> Models.....	34
1.6.2 <i>In vitro</i> Models	35
1.7 Stress Responses of <i>Mtb</i>	36
1.7.1 Reactive Oxygen Radicals	37
1.7.2 Reactive Nitrogen Intermediates.....	38

The Universal Stress Proteins of Bacteria

1.7.3	pH Stress	38
1.7.4	Nutrient Starvation	39
1.7.5	Hypoxia	39
1.8	The DosR Regulon	40
1.9	<i>E. coli</i> UspA.....	41
1.10	UspA Superfamily	42
1.10.1	Distribution	42
1.10.2	Structural Forms of USPs	44
1.11	USP Classification	45
1.11.1.1	PP Loop Superfamily	45
1.11.1.2	The G-X ₂ -G-X ₉ -G(S/T) Nucleotide Binding Motif	46
1.12	USP Function	48
1.12.1	Function of <i>E. coli</i> USPs.....	48
1.12.2	USP Function in Other Bacteria	49
1.13	The USPs of <i>Mtb</i>	49
1.14	Why Study the <i>Mtb</i> USPs?	51
1.15	The USP Rv1636	52
1.16	Aims and Objectives.....	55
Chapter 2 -	Materials and Methods.....	57
2.1	General.....	58
2.1.1	Media	58
2.1.2	Buffers and Solutions.....	59
2.1.3	Antibiotics	60
2.1.4	Strains	61
2.1.5	Plasmids	62
2.1.6	Oligonucleotides	65
2.1.7	Strain Storage and Culture.....	69
2.1.7.1	Glycerol Stocks.....	69
2.1.7.2	Strain Culture	69
2.2	Protein Methods	70
2.2.1	Protein Purification	70
2.2.2	Size Exclusion Chromatography (SEC).....	71
2.2.3	SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE).....	72

The Universal Stress Proteins of Bacteria

2.2.4	Western Blotting.....	72
2.2.5	<i>In silico</i> Analysis of Proteins.....	73
2.3	Nucleic Acid Methods	73
2.3.1	Transformation of Bacteria.....	73
2.3.1.1	Escherichia coli.....	73
2.3.1.2	Mycobacteria	74
2.3.1.3	Pseudomonas.....	75
2.3.2	DNA Plasmid Purification	76
2.3.3	Agarose Gel Electrophoresis	76
2.3.4	Sequencing DNA	76
2.3.5	DNA Concentration	76
2.3.6	Polymerase Chain Reaction	77
2.3.7	PCR Purification	77
2.3.8	Gel Extraction.....	77
2.3.9	DNA Restriction Digestion.....	77
2.3.10	Genomic DNA extraction for PCR	78
2.3.10.1	Pseudomonas.....	78
2.3.10.2	Mycobacteria	78
2.3.11	Site Directed Mutagenesis (SDM)	79
2.3.11.1	Primer Design.....	79
2.3.11.2	Mutagenesis.....	79
2.3.12	Polyacrylamide Gel Analysis of Oligonucleotides	80
2.4	Biochemical Assays	81
2.4.1	<i>In vitro</i> Kinase Assay	81
2.4.2	Phosphoamino Acid Analysis	82
2.4.3	<i>In vitro</i> Nucleotidylation Assay.....	83
2.4.4	<i>In vitro</i> RNA Triphosphatase Assay	83
2.4.5	<i>In vitro</i> NTPase Assay	84
2.4.6	Nucleotide Binding Experiments	84
2.4.6.1	UV Cross-linking Experiments	84
2.4.6.2	Vacuum Filtration	85
2.5	Biophysical Techniques	85
2.5.1	Multi Angle Laser Light Scattering (MALLS) Analysis	85

The Universal Stress Proteins of Bacteria

2.5.2	Analytical Ultracentrifugation.....	86
2.6	Identification of Protein-Protein Interactions.....	87
2.6.1	Cell Fractionation.....	87
2.6.2	GST Pull-down.....	87
2.6.3	Co-immunoprecipitation.....	88
2.6.4	Macrophage Infection with <i>Mtb</i>	89
Chapter 3 -	Biochemical Characterization of Rv1636	90
3.1	Introduction	91
3.1.1	Biochemical Activities of USPs	91
3.1.2	Biochemical Characterisation of <i>Mtb</i> USP Rv1636	92
3.1.2.1	Autokinase Activity	93
3.1.2.2	Nucleotidylylation Activity.....	93
3.1.2.3	RNA Triphosphatase Activity	94
3.1.3	Aims	95
3.2	Results.....	96
3.2.1	Protein Purification	96
3.2.1.1	Rv1636 _{WT} Purification	96
3.2.1.2	Generation of Site Directed Mutants in Rv1636.....	98
3.2.2	Investigating the Biochemical Activities of Rv1636	101
3.2.2.1	Investigation of the Rv1636 Autokinase Activity	101
3.2.2.2	Investigation of Autonucleotidylylation Activity	107
3.2.2.3	Is Rv1636 an RNA Triphosphatase?	109
3.2.2.1	Rv1636 Binds Adenosine Triphosphate	111
3.3	Discussion.....	116
3.3.1	Purification of Rv1636 Proteins	116
3.3.2	Biochemical Activities of Rv1636	117
3.3.2.1	Post-translational Modifications of Rv1636	117
3.3.2.2	RNA Triphosphatase Activity	120
3.3.2.3	Nucleotide Binding of Rv1636	121
3.3.3	Biochemical Studies: <i>M. smegmatis</i> USPs	121
3.3.3.1	cAMP in <i>Mtb</i>	123
3.3.4	Summary.....	123
Chapter 4 -	Characterisation of the Putative NTPase Activity of Rv1636.....	125

The Universal Stress Proteins of Bacteria

4.1	Introduction	126
4.1.1	Putative NTPase Activity of Rv1636.....	127
4.1.2	Structural Methods Used in this Study	128
4.1.2.1	Multi-Angle Laser Light Scattering.....	128
4.1.2.2	Analytical Ultracentrifugation.....	129
4.1.3	Aims	131
4.2	Results.....	132
4.2.1	USP Ligand Binding	132
4.2.2	Analysis of the Crystal Structure of Rv1636.....	134
4.2.3	Classification of Rv1636	136
4.2.4	Rv1636 is a Strand Swapped Dimer	138
4.2.4.1	Disrupting the Strand Swapped Rv1636	140
4.2.4.2	Purification of Truncated Rv1636 Constructs	144
4.2.5	Size Exclusion Chromatography - Multi-Angle Laser Light Scattering	145
4.2.5.1	SEC-MALLS shows MAKT Exists in a Monomeric Form.....	146
4.2.6	Analytical Ultracentrifugation.....	149
4.2.6.1	Size Exclusion Chromatography.....	149
4.2.6.2	AUC of Rv1636 _{WT}	151
4.2.6.3	AUC of Rv1636 _{MAKT} Construct.....	153
4.2.6.4	Does Rv1636 Hydrolyse NTPs?	155
4.2.6.5	Do the Truncated Constructs Bind Nucleotide Triphosphate?	161
4.3	Discussion.....	163
4.3.1	Rv1636 Crystal Structure	163
4.3.2	Design and Purification of Truncated Rv1636	163
4.3.3	Biophysical Analysis of Rv1636	164
4.3.4	Biochemical Properties of Rv1636	165
4.3.5	Ligands of USPs	166
4.3.6	Structural Insights into USPs.....	167
4.3.7	USP Structure Related to Function	168
4.3.8	Summary	170
Chapter 5 -	Intracellular Studies of Universal Stress Proteins	171
5.1	Introduction	172
5.1.1	Mycobacterial USPs <i>in vivo</i>	172

The Universal Stress Proteins of Bacteria

5.1.2	<i>Pseudomonas aeruginosa</i> USPs	173
5.1.2.1	<i>P. aeruginosa</i> Overview	173
5.1.2.2	Universal Stress Proteins of <i>P. aeruginosa</i>	175
5.1.2.3	The Universal Stress Protein PA3309.....	176
5.1.2.4	PA3309 Function.....	176
5.1.3	Aims	177
5.2	Results.....	178
5.2.1	Rv1636 Homologues	178
5.2.2	Creation of a <i>Mtb</i> Rv1636 Mutant.....	182
5.2.2.1	Macrophage Infection with Putative <i>Mtb</i> H37Rv Rv1636 Mutants Reveals a Possible Survival Defect Phenotype.....	184
5.2.3	Identification of Proteins Interacting with Rv1636.....	185
5.2.3.1	Database Search	186
5.2.3.2	Rv1636 Expression	187
5.2.3.3	GST Pull-down of Rv1636.....	189
5.2.3.4	Co- immunoprecipitation of Rv1636	191
5.2.4	Over-expression of Rv1636.....	195
5.2.4.1	Cloning of Rv1636 into Over-expression Vectors	195
5.2.4.2	Over-expression of Rv1636 with pMV261::Rv1636.....	198
5.2.5	The Universal Stress Protein PA3309 from <i>P. aeruginosa</i>	199
5.2.5.1	PA3309 Structure.....	199
5.2.6	Does the Predicted PA3309 Nucleotide Binding Motif have a Direct Biological Role?.....	201
5.2.6.1	Expressing PA3309 in <i>P. aeruginosa</i> using pUCp18 vector.....	201
5.2.6.2	Expression of PA3309 Using the Integrative Mini-CTX2 Vector	203
5.2.6.3	Construction of Mini-CTX2::PA3309 _{G118A}	203
5.2.6.4	Confirmation of Insertion Site in <i>P. aeruginosa</i>	204
5.3	Discussion.....	207
5.3.1	Genomic context and lineage	207
5.3.2	Creation of a Rv1636 knockout.....	208
5.3.3	Identification of Protein Interactions	209
5.3.3.1	Rv1636 Expression	209
5.3.3.2	GST Pull-down of Rv1636.....	210
5.3.3.3	Co-immunoprecipitation of Rv1636	211

The Universal Stress Proteins of Bacteria

5.3.3.4	Known USP Interactions	212
5.3.4	Over-expression of USPs.....	213
5.3.5	<i>P. aeruginosa</i> Aerobic Survival and the PA3309 Protein	214
5.3.6	Summary	216
Chapter 6 -	General Discussion.....	217
6.1	Introduction	218
6.2	Biochemical Activities	218
6.2.1	Activities of Rv1636	218
6.2.2	PP-loop Members	219
6.3	Biophysical Studies.....	220
6.3.1	Structure of Rv1636.....	220
6.3.2	Functional Oligomerisation.....	221
6.4	Intracellular Role of USPS.....	223
6.4.1	Role of Rv1636	223
6.4.2	USP Rv2623 has a role in Establishing Chronic Infection in Mice	225
6.4.3	Role of PA3309.....	226
6.5	Future work.....	226
6.6	General Role of USPs.....	227
Chapter 7 -	Bibliography	229

List of Figures

Figure 1.1. <i>Mycobacterium tuberculosis</i> scanning electron micrograph.....	19
Figure 1.2. Estimated TB incidence rates 2008.....	22
Figure 1.3. Granulomas of <i>Mtb</i>	31
Figure 1.4. Taxonomic distribution of the universal stress protein domains	43
Figure 1.5. Schematic representation of different UspA family proteins.....	44
Figure 1.6. ATP binding of the <i>Methanococcus jannaschii</i> MJ0577 protein.....	47
Figure 1.7. The USPs of <i>Mtb</i>	50
Figure 1.8. Protein sequence alignment of Rv1636 with ML1390, MJ0577, UspA (<i>E. coli</i>) and UspA (<i>H. influenzae</i>).	54
Figure 3.1. Time dependent modification of His ₁₀ -UspG <i>in vitro</i>	92
Figure 3.2. Apparent <i>in vitro</i> autophosphorylation of Rv1636 _{WT}	93
Figure 3.3. Apparent <i>in vitro</i> nucleotidylylation of Rv1636 _{WT}	94
Figure 3.4. Putative RNA triphosphatase releases a radiolabelled phosphate from the γ -labelled RNA transcript.....	95
Figure 3.5. Purification of Rv1636 _{WT} by immobilized metal-ion affinity chromatography.....	97
Figure 3.6. Purified His ₍₆₎ Rv1636 _{WT} protein.	98
Figure 3.7. Putative nucleotide binding site of Rv1636	99
Figure 3.8. Expression and purification of his-tagged site direct mutant proteins	100
Figure 3.9. Autoradiograph of Rv1636 _{WT} following incubation with radiolabelled GTP.....	101
Figure 3.10. Phosphoamino acid analysis by semi-dry electrophoresis on a thin-layer chromatography electrophoretic plate	102
Figure 3.11. NetPhos Prediction of the serine phosphorylation sites in Rv1636.	103
Figure 3.12. Multiple alignment of Rv1636 from <i>Mtb</i> with homologues from <i>M. bovis</i> (Mb1662), <i>M. smegmatis</i> (MSMEG3811), <i>M. leprae</i> (ML1390), <i>M. avium</i> (MAP1339), <i>M. vanbaalenii</i> (Mvan3342), <i>M. gilvum</i> (Mflv3551) and <i>M. ulcerans</i> (MUL1619).	104
Figure 3.13. Expression and purification of Rv1636 _{S16A-S17A} and Rv1636 _{S118A} proteins.....	105
Figure 3.14. Phosphorylation assays with Rv1636 and MBP	106
Figure 3.15. Analysis of the nucleotidylylation of Rv1636 using piperidine release	108
Figure 3.16. <i>In vitro</i> nucleotidylylation of Rv1636 mutants.....	109
Figure 3.17. RNA triphosphatase control assay. Rv1636 _{WT} and Rv1636 _{G113A} protein	111
Figure 3.18. UV crossing linking of Rv1636 proteins.....	113
Figure 3.19. Nucleotide binding assay using a filter vacuum.....	114

The Universal Stress Proteins of Bacteria

Figure 4.1. The crystal structures of <i>H. influenzae</i> UspA and <i>M. jannaschii</i> MJ0577	126
Figure 4.2. Michaelis-Menten kinetics of Rv1636 _{WT} ATPase activity	128
Figure 4.3. Schematic diagram of the Beckman Optima XL-A absorbance system.	130
Figure 4.4. Deposited crystal structure of Rv1636.....	134
Figure 4.5. The structure of Rv1636.	135
Figure 4.6. Alignment of Rv1636 with the MJ0577, UspA from <i>H. influenzae</i> and the USPs of <i>E. coli</i>	137
Figure 4.7. Strand swap of the Rv1636 protein..	139
Figure 4.8. Sequence alignment of Rv1636 and the truncated version of Rv1636.	141
Figure 4.9. Schematic diagram of the R1636 constructs	143
Figure 4.10. Solubility of truncated proteins expressed in <i>E. coli</i> Rosetta cells.	145
Figure 4.11. SEC-MALLS analysis of the Rv1636 _{MYKT} and Rv1636 _{MFKT}	147
Figure 4.12. SEC-MALLS analysis of the Rv1636 _{MAKT} protein.	148
Figure 4.13. Size exclusion chromatography of Rv1636 _{WT} and Rv1636 _{MAKT}	150
Figure 4.14. Analytical centrifugation data of Rv1636 _{WT}	152
Figure 4.15. Analytical centrifugation of Rv1636 _{MAKT}	154
Figure 4.16. Nucleotide hydrolysis by Rv1636 _{WT} and the site direct mutants.....	155
Figure 4.17. Nucleotide hydrolysis by Rv1636 _{WT} and the truncated constructs.	156
Figure 4.18. Size exclusion chromatography of Rv1636 _{WT} followed by an NTPase assay.....	158
Figure 4.19. NTPase assay on fractions after IMAC purification.....	160
Figure 4.20. Nucleotide binding of Rv1636 _{WT} and truncated constructs.....	162
Figure 5.1. The regulation of the USPs of <i>Mtb</i>	173
Figure 5.2. Pyruvate fermentation pathway model by <i>P. aeruginosa</i>	174
Figure 5.3. The USPs of <i>P. aeruginosa</i>	175
Figure 5.4. Phylogenetic tree view of the fastest minimal evolution of Rv1636.....	179
Figure 5.5. Genomic context of <i>Rv1636</i> homologues.....	181
Figure 5.6. General strategy for mutagenesis of <i>Rv1636</i> gene.....	183
Figure 5.7. Survival curve of intracellular bacilli from a range of <i>Mtb</i> mutants	185
Figure 5.8. String database search for predicted protein-protein interactions of Rv1636 ..	187
Figure 5.9. Growth curves for <i>M. bovis</i> BCG.....	188
Figure 5.10. Western Blot using anti-His ₆ -Rv1636 antibody against samples of cytosolic fractions produced from <i>M. bovis</i> BCG cells.....	189
Figure 5.11. GST-pull-down reaction with GST-Rv1636.....	190
Figure 5.12. GST-pull-down reaction of GST-Rv1636 from cytosolic fractions.....	191

The Universal Stress Proteins of Bacteria

Figure 5.13. Immunoprecipitation of His ₆ -Rv1636 _{WT} protein using protein A agarose	193
Figure 5.14. Co-immunoprecipitation of the Mb1662 protein from <i>M. bovis</i> BCG.....	194
Figure 5.15. PCR analysis of across <i>M. smegmatis</i> attB site.	197
Figure 5.16. Western blot confirming over-expression of Rv1636 in <i>M. smegmatis</i>	198
Figure 5.17. Over-expression of Rv1636 in <i>M. smegmatis</i>	199
Figure 5.18. Sequence comparison of <i>Mtb</i> USP Rv1636 with <i>P. aeruginosa</i> USP PA3309....	200
Figure 5.19. Predicted tertiary structure of <i>P. aeruginosa</i> USP PA3309.	200
Figure 5.20. Anaerobic survival of <i>P. aeruginosa</i> on pyruvate-supplemented media.....	202
Figure 5.21. Double digestion of the mini-CTX2:: <i>PA3309</i> _{G118A} plasmid	204
Figure 5.22. PCR analysis of <i>P. aeruginosa</i> strains.....	206
Figure 6.1. Sequence alignment of crystallised USPs..	222

List of Tables

Table 2.1. Composition of media	58
Table 2.2. Buffers and solutions.....	59
Table 2.3. Antibiotics and solvents	60
Table 2.4. Strains.....	61
Table 2.5. Plasmids and origin	62
Table 2.6. Primers used in site directed mutagenesis of Rv1636.	65
Table 2.7. Primers designed to create the truncated versions of Rv1636.....	66
Table 2.8. Primers designed for amplification of <i>Rv1636</i> and hygromycin cassette.....	66
Table 2.9. Primers used to clone Rv1636 into the over-expression vectors	67
Table 2.10. Primers to amplify <i>M. smegmatis</i> attB site.....	67
Table 2.11. Primers used for confirmation of miniCTX2:: <i>PA3309</i> integration.	68
Table 2.12. Primers used to clone PA3309 gene region.	68
Table 2.13. Composition of SDS-PAGE gels.....	72
Table 2.14. Cycling parameters for site-directed mutagenesis.	80
Table 2.15. 10% Polyacrylamide denaturing urea gel.....	81
Table 2.16. Loading Dye for denaturing urea PAGE.....	81
Table 4.1. Ligand binding and probable quaternary structures of crystallised USPs	133
Table 4.2. Interface statistics of Rv1636 deposited crystal structure.....	135
Table 5.1. Homologues of Rv1636 across <i>Mycobacterium</i> species	178
Table 5.2. Related proteins to Rv1636 within the <i>Corynebacterineae</i> suborder.	179
Table 5.3. PCR primer pairs used for analysis of the <i>P. aeruginosa</i>	205

Abbreviations

AIDS	Acquired immune deficiency syndrome
ATP	Adenosine triphosphate
AUC	Analytical ultracentrifugation
BCG	Bacillus Calmette-Guérin
CAT	Computerized axial tomography
CF	Cystic fibrosis
CFU	Colony forming units
DNA	Deoxyribonucleic acid
DOTS	Directly observed treatment and short-course drug therapy
DTT	Dithiothreitol
FPLC	Fast protein liquid chromatography
GTP	Guanosine triphosphate
HIV	Human immunodeficiency virus
HPLC	High performance liquid chromatography
IMAC	Immobilised metal affinity chromatography
LAM	Lipoarabinomannan
LB	Luria Bertani broth
MALLS	Multi angle laser light scattering
MDR	Multi drug resistant
<i>Mtb</i>	<i>Mycobacterium tuberculosis</i>
NRP	Non replicating persistence
NTPase	Nucleotide triphosphatase
OD₆₀₀	Optical density at 600 nm
PET	Positron emission tomography
RNA	Ribonucleic acid
RNIs	Reactive nitrogen intermediates
ROIs	Reactive oxygen intermediates
rpm	Revolutions per minute
SDM	Site-directed mutagenesis
SDS	Sodium-dodecyl-sulphate
SDS-PAGE	Sodium-dodecyl-sulphate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
TB	Tuberculosis
USP	Universal stress protein
UV_{280nm}	Ultraviolet absorbance at 280nm
V	Volts
w/v	Weight to volume
WHO	World Health Organisation
XDR-TB	Extreme drug resistant TB

Chapter 1 - Introduction

1.1 Scope of the Introduction

Tuberculosis (TB) remains a huge global problem and the current pressing issues relating to the disease, including detection, prevention and treatment, are introduced below. TB kills up to 1.8 million people a year worldwide, yet up to a third of the world's population are thought to harbour an asymptomatic latent infection (WHO, 2010, Dye *et al.*, 1999). The highly specialised bacteria employ an array of survival mechanisms and strategies to persistently evade the host immune system. As a result, the bacteria are able to maintain latent infections over long periods of time, in some cases decades and, when reactivated, the disease often results in a high mortality rate.

Members of the Universal Stress Protein (USP) superfamily are postulated to play a role in the successful survival of the pathogen *Mycobacterium tuberculosis* (*Mtb*) within the human host. The USPs are widespread, found in the three kingdoms of life; however the action of these small proteins continues to be incompletely understood. The study of the USPs aims to determine if this protein family has a role in aiding the survival of the *Mtb* for prolonged latent periods within the host.

The later part of the introduction focuses on the USP family and the roles of these proteins in bacteria in general, with a particular emphasis on the response to diverse stress conditions. The research data presented aims to further characterise the USP Rv1636 from *Mtb* and provide insights into the broad biological roles that USPs themselves play.

Part I: *Mycobacterium Tuberculosis*

1.2 The *Mycobacterium* Genus

The genus *Mycobacterium* is part of the *Actinobacteria*, which also contains the genera *Rhodococcus*, *Gordonaiceae* and *Tsukamurellae*. The genus contains more than 50 species and although most are non-pathogenic the group does contain important human pathogens, such as *Mtb* and *M. leprae*. The latin 'myco' means both *fungus* and *wax*, likely in relation to the waxy nature of the compounds that make up the mycobacterial cell wall and their resulting fungal-like colony morphology.

The Universal Stress Proteins of Bacteria

The bacteria are considered to be gram positive; however they do not stain well due to the unique cell wall properties. The thick cell wall consists of two layers, one of mycolic acids and one of peptidoglycan, held together by the polysaccharide arabinogalactan. The cells stain bright red with dyes such as carbol fuchsin, used in the Ziehl-Neelsen staining technique, and the cell wall leads to retention of the dye when decolourised with acidified organic compounds, giving rise to the common name of 'acid-fast bacteria'. Organisms within the genus are typically widespread being found in water and food sources. Some, however, such as *Mtb* and *M. leprae* are intracellular parasites and are not found as free-living bacteria.

Mycobacteria are split into two groups depending on the growth rate: slow growing and fast growing (Wayne & Kubica, 1986). Of the slow growing bacteria, the most significant group to human health is the *Mycobacterium tuberculosis* complex which contains the causative agents of human and animal tuberculosis. These include *Mtb*, *M. bovis*, *M. bovis* BCG, *M. africanum*, *M. canetti*, *M. caprae*, *M. microti* and *M. pinnipedii*.

1.3 *Mycobacterium tuberculosis*

1.3.1 The *Mtb* Bacterium

Mtb are non-motile rod-shaped bacteria that are 2-4 μm in length and 0.2-0.5 μm in width (Figure 1.1). The tubercle bacillus is a facultative aerobe and has a slow generation time of 15-20 hours. Humans are the only known reservoir for the bacterium and latently infected individuals represent a huge reservoir of infection. The bacteria can grow on a variety of carbon sources, including glycerol, which is phosphorylated before being converted to pyruvate during metabolism (Ratlidge, 1982).

Mtb encodes an inducible nitrate reductase, *narG*, which allows respiration using NO_3 as a final electron acceptor (Weber, 2000). However, it is not clear if it alone supports growth in microaerophilic environments or merely sustains the viability of the bacterium. In the laboratory, *Mtb* is usually grown on two media types either the agar based Middlebrook's medium or an egg based medium, containing albumin, termed Lowenstein-Jensen medium. The bacteria, after 4 to 6 weeks, form small colonies with a characteristic colony morphology and appear to form a 'serpentine-type growth'.

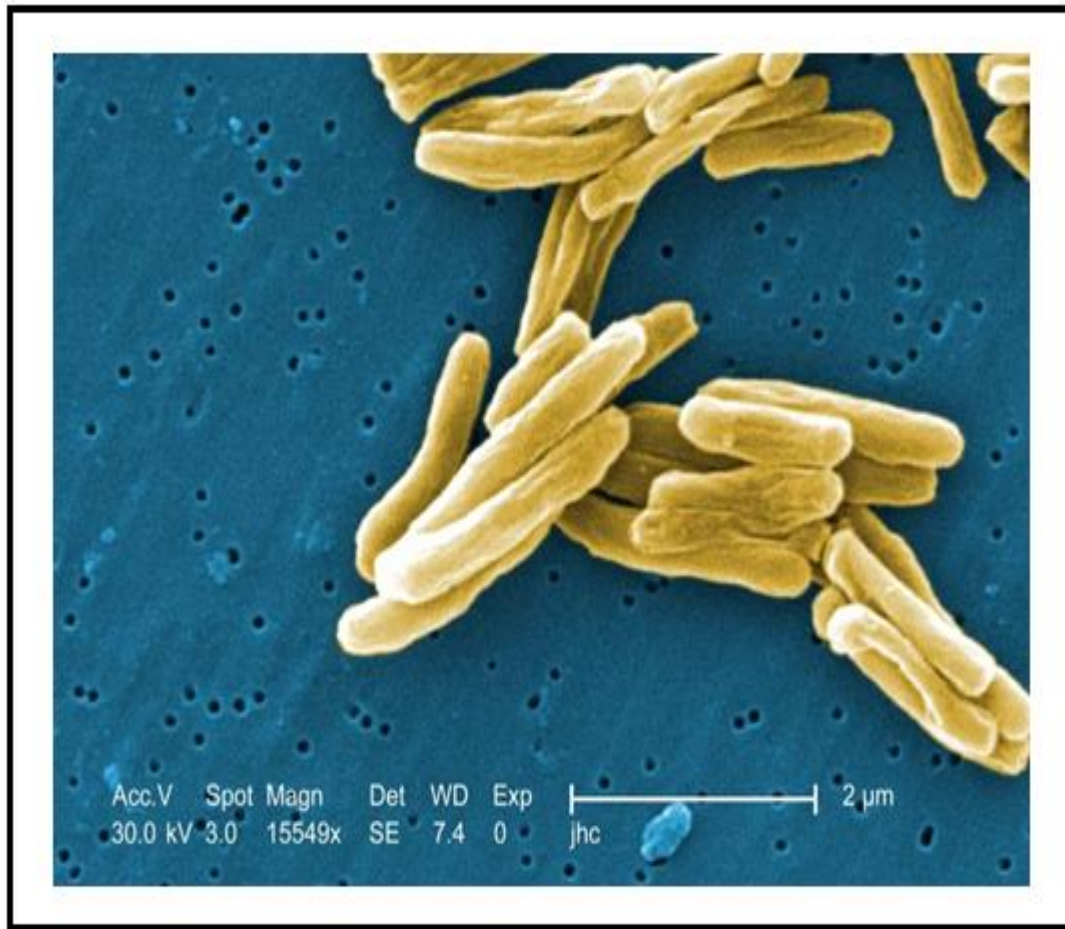


Figure 1.1. *Mycobacterium tuberculosis* scanning electron micrograph at 15549X magnification. Micrograph from the Centers for Disease Control and Prevention (CDC) www.cdc.gov/tb.

1.3.2 The *Mtb* Cell Wall

The cell wall of *Mtb* is extremely complex and is vital for its survival. The inner layer of the cell wall is composed of peptidoglycan to which arabinoglycan is covalently attached. It is to this layer that the unique feature of the mycobacterial cell wall, mycolic acids, are linked. The crosslinking of mycolic acids and long chain fatty acids with a variety of cell wall lipids is the feature of the mycobacterial cell wall that leads to its acid fast properties (Daffe & Draper, 1998). This thick layer of mycolic acids is a key factor in *Mtb*'s resistance to a range of antibacterial agents and provides the bacterium with protection against the host immune system. The importance of the TB cell wall has been exploited in the development of anti-TB drugs and indeed two of the current TB drugs, isoniazid and ethambutol, target the synthesis of mycolic acids.

The Universal Stress Proteins of Bacteria

The cell wall contains a number of factors that promote virulence, including lipoarabinomannan (LAM), which acts as a virulence factor by interfering with the coordinated immune response against the bacteria through inhibiting protein kinase C, scavenging oxygen free radicals and also blocking the transcriptional effects of interferon- γ (Chan *et al.*, 1991). However, LAM also can induce the release of the tumour necrosis factor from macrophages promoting bacterial elimination (Chatterjee, 1992). The cell wall contains cell surface molecules that permit binding and entry into mononuclear phagocytes. For example *Mtb* can enter macrophages using the integrin family CR1 and CR3 complement receptors (Ahearn & Fearon, 1989).

1.4 Tuberculosis

The bacterium *Mtb* is the etiologic agent of the infectious disease tuberculosis (TB). The disease kills more than 35,000 people every week (1.8 million people every year) (WHO, 2010). The illness most commonly affects the lungs; however, other parts of the body, such as brain, bones and lymph glands, can also be compromised. Common symptoms of the disease include a persistent cough that brings up sputum or phlegm, often containing blood. Other symptoms of TB are lung destruction, tiredness, weight loss, chills, fever, and night sweats. The disease is transmitted when someone with active lung disease coughs and produces aerosol droplets that harbour infectious bacilli, and which are subsequently inhaled by a susceptible individual. A large reservoir of the bacteria exists within humans, as the bacterium is able to survive in the human host for long periods of time without causing clinical symptoms associated with TB (Parrish *et al.*, 1998).

Following exposure to *Mtb*, those people who are unable to completely clear the infection either develop active TB or a latent infection. In the majority of cases (~90%) the infection proceeds along the latent route and the immune system is effective at keeping the bacterium contained in the latent form; however the bacterium is often not eliminated from the body (Kaufmann, 2001). The non-replicative persistent bacteria can then be re-activated when immune system function decreases, such as in old age (Manabe and Bishai, 2000) or through infection with HIV. In many developing countries, two thirds of adults carry a latent infection and, as the population ages, the disease is likely to re-activate, causing millions of deaths. The BCG vaccine has been shown to have limited efficacy and is not suitable to protect against re-activation once the initial infection is established. Key areas in current TB research include the HIV-TB co-epidemic, the rising levels of drug resistant TB, the need for better diagnostic assays and the development of new vaccines.

1.4.1 History of TB

Tuberculosis has been a notorious killer of humans over thousands of years. Tubercular decay was shown to be present in the spines of Egyptian mummies from 3000-2400BC (Zink, 2003). Over the course of history, the disease has been known by many names, including consumption, phthisis, scrofula, Pott's disease, and the white plague. The name 'consumption' was due to the disease being said to consume people from within, with a bloody cough, fever and a general wasting away being common symptoms of the disease. The Greek term Phthisis for the disease was described by physician Hippocrates (460BC), who revolutionized medicine in ancient Greece, as the most widespread disease of the time that almost certainly ended in death. Tuberculosis continued to ravage the western world and afflicted the urban poor populations in epidemic proportions. In the 19th and early 20th Century, TB was of widespread public concern and, known as the white plague, it was responsible for a quarter of all deaths in the UK.

Although TB was recognised as a contagious disease in the 1820s, the tuberculoase bacillus was not identified and described until 1882 by Dr Robert Koch. Dr Koch was later awarded the Nobel Prize in physiology or medicine in 1905 for his discovery. It was one year later in 1906 that Albert Calmette and Camille Guérin began developing an attenuated bovine strain of tuberculosis, a line of research that took until 1919 to complete. Their attenuated strain showed success in immunizing against the disease and the vaccine was first used on humans in France in 1921. Its use became widespread after World War II in the USA, Great Britain and Germany and with increased hygiene and sanitation, combined with antibiotics and vaccination, the war against tuberculosis seemed in the ascendency. However, TB has shown a recent resurgence and, in 1993, the World Health Organization (WHO) declared the disease a global health emergency.

1.4.2 Distribution of TB

Tuberculosis remains an important challenge to world health even now in the 21st Century. In 2008 there were 9.4 million new cases of TB, of which 80% were in just 22 countries, and this resulted in 1.8 million deaths. The number of new cases of TB across the world is steadily rising; however, with the world's population growing about 2% per annum, per capita the global TB incidence rate is declining at an estimated 1% per year (WHO report, 2009). TB is classed as a worldwide pandemic with the greatest burden being found in Africa and Asia (Figure 1.2). Thirteen of the 15 countries with the highest TB incidence rates are in Africa, highlighting the seriousness of

The Universal Stress Proteins of Bacteria

the problem for this continent. Of new cases recorded, 55% occur in six Asian countries (Bangladesh, China, India, Indonesia, Pakistan and the Philippines) highlighting the problems in prevention of TB in these countries.

The World Health Organisation and the Stop TB Partnership has set out the Global Plan to Stop TB (2006–2015) with an investment of \$67 billion US dollars. Within this plan, there are certain targets to meet by 2015 including the Millennium Development Goal of reducing the global burden of TB (per capita prevalence and death rates) by 50% relative to levels in 1990. The WHO has a global plan to eliminate TB as a public health problem aiming for the global incidence of active TB to be less than 1 case per million population per year by 2050.

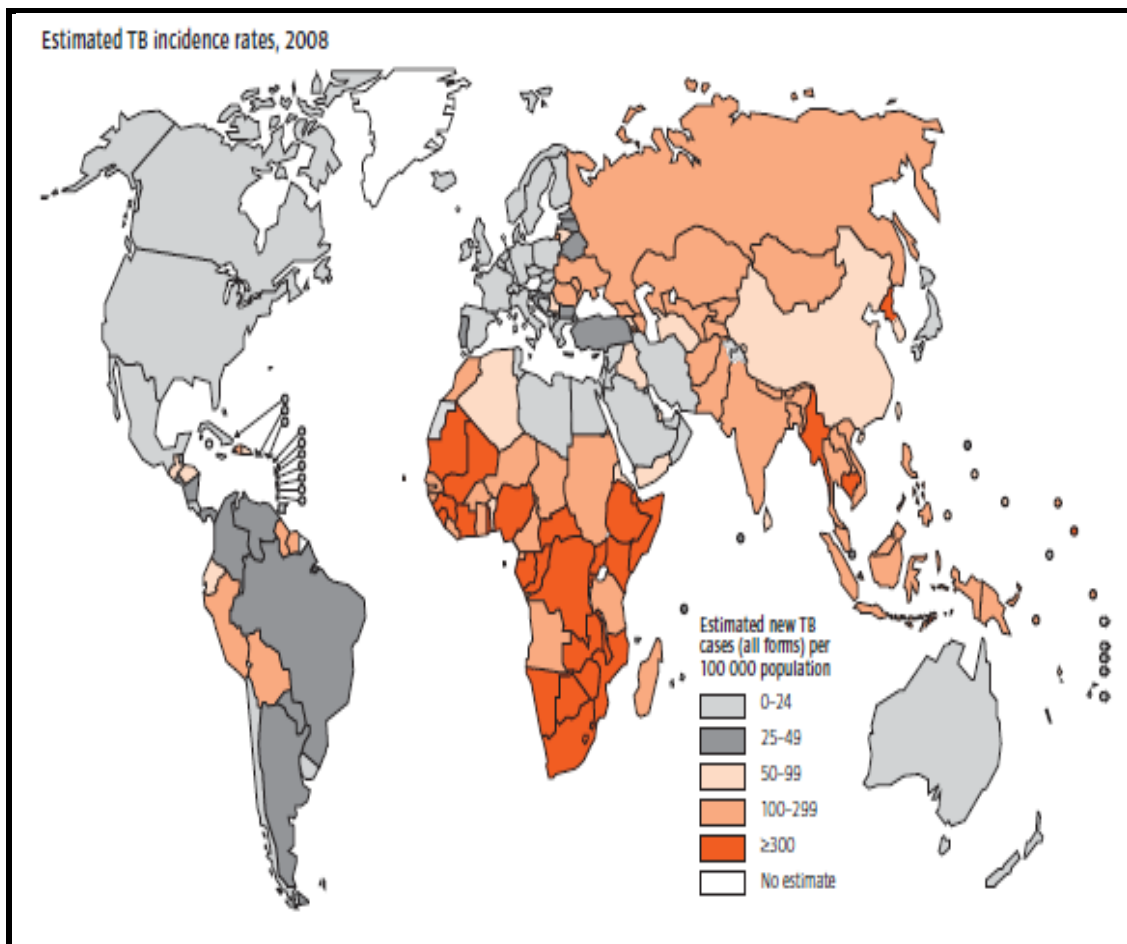


Figure 1.2. Estimated TB incidence rates 2008 (WHO Report 2010).

1.4.3 HIV associated TB

TB is the leading killer among people infected with HIV, with sufferers 20-40 times more likely to develop active TB due to the weakened state of their immune system. In 2008 there were 1.4 million new TB cases associated with HIV positive patients and a total of 0.52 million TB deaths (29% of total TB deaths) among HIV positive cases (WHO, 2009). The HIV-associated TB epidemic represents a major challenge to accomplishing the 2015 Millennium Development Goals for TB control. Nearly 80% of HIV-TB cases are in sub-Saharan Africa, with South Africa alone accounting for 25% of cases despite only having less than 1% of the world's total population (Abdool-Karim *et al.*, 2009).

In England and Wales, cases of TB combined with HIV increased from 3.1% of TB cases in 1999 to 8.3% in 2003 (Ahmed *et al.*, 2007), mirroring problems in Eastern Europe where overlapping HIV and drug-resistant TB epidemics have seen regional TB incidence rates double since 1990 (Lazarus *et al.*, 2008). The United States, on the other hand, has had a threefold decrease between 1993 and 2004 due to improvements in control, diagnosis and treatment (Albalak *et al.*, 2007).

Treatment programmes for HIV-infected individuals include highly active antiretroviral therapy (HAART) which uses expensive antiretroviral drugs. There are problems associated with the therapy including the Immune Restoration Inflammatory Syndrome (IRIS). IRIS is the belated response of a recovering individual's immune system towards to an infection and has been seen in patients with infections such as *Mtb*, *Mycobacterium avium complex* and other HIV-associated infections (French *et al.*, 2000). For *Mtb*-related IRIS, management includes the continuation of HAART therapy and, in some cases, steroids to modulate the immune system. The symptoms of IRIS are often confused with those of other diseases including HIV-related illnesses and accurate diagnosis is therefore of high importance. In addition, HAART treatment has become increasingly complicated with the increase in drug resistant strains of *Mtb*.

1.4.4 Diagnosis of TB

To ensure effective treatment and control of TB, diagnosis needs to be improved in reliability and speed (Abu-Raddad *et al.*, 2009). TB is usually diagnosed through identification of *Mtb* in a clinical sample such as sputum or from the site of disease if extra-pulmonary; however, culturing the organism in the laboratory may take up to three months and the laboratory infrastructure required is usually a major limitation in many developing countries. Acid-fast bacilli or bacilli fluorescently

The Universal Stress Proteins of Bacteria

labelled can be detected rapidly in sputum smears, but the high rate of smear negativity means that culturing of the bacilli always needs to be carried out in association with this. In most cases a full medical evaluation is required including chest X-rays, tuberculin skin test (Mantoux test) and microscopic analysis and culturing of sputum smears or sampling from the site of disease. The tuberculin skin test is not 100% accurate with false negatives/positives occurring as the results are affected by prior vaccination with BCG, previous exposure to mycobacteria and immune-competence levels.

More recent and more specific and sensitive diagnostic tests include the interferon-gamma release assays (IGRAs), which are *in vitro* blood tests which test the immune response against mycobacterial proteins such as the 6-kDa early secretory antigenic target (ESAT-6) and culture filtrate protein 10KDa (CFP-10), proteins which are not present in the vaccine strain BCG and which are therefore not cross-reactive in vaccinated individuals. The IGRA's are either ELISA-based, such as Quantiferon, or based on ELISpot, such as TSPOT-TB and routinely show enhanced specificity and sensitivity compared to the tuberculin skin test.

Molecular biology based assays such as polymerase chain reaction assays have been developed to detect mycobacterial DNA in samples and this technique may provide the key to diagnosing drug resistance in a swifter manner than culturing, which is currently used. Fast and effective diagnosis of drug resistance is key to controlling and treating MDR and XDR-TB and nucleic acid amplification techniques have been developed which amplify drug resistance genes such as *rpo* (rifampicin) and isoniazid (*inh* and *katg*). These include the Genotype *MtbDR* plus (Hain Lifesciences, Nehren, Germany) and GeneXpert (Cephaid) assays.

New TB tests are continuing to be developed with the hope of providing cheaper, faster and more accurate diagnosis that are urgently required. Currently, microscopic diagnosis of acid fast bacilli in sputum/site of disease samples is widely used in developing countries due to the relatively low cost and requirements. Improvements in microscopic diagnosis are being researched such as the modification of light-emitting-diodes (LED) for use in fluorescent microscopy. They are more affordable, have long lasting bulbs and have lower electric-power demand (Hanscheid, 2008). An example of labelling *Mtb* involves using the stain auramine, which binds mycolic acids, followed by the counter stain rhodamine. However, due to the high prevalence of smear-negativity, particularly in HIV+ individuals, and the high levels of drug resistance, the need for improved diagnostics applicable to resource poor settings is urgently required.

1.4.5 Drug Resistance and its Development in TB

TB can be treated with antibiotics; however, drug resistance and economic factors often prevent this method of control being successfully used in the developing countries where the health infrastructure is often poor and lacks resources (Brennan, 2005). Anti-tuberculosis drugs were first isolated and brought into use mid-way through the 20th Century. The first to be discovered, in 1944, was Streptomycin (Schatz *et al.*, 2005), followed by Isoniazid in 1951 (Bernstein *et al.*, 1952; Fox, 1952; Offe & Domagk, 1952), Pyrazinamide (Yeager *et al.*, 1952), Ethambutol (Delaude *et al.*, 1962), and Rifampicin (Furesz *et al.*, 1965). Chemotherapeutic regimens often require more than six months treatment with multiple drugs due to the persistent nature of the *Mtb* infection.

Due to the problems of compliance and the rise in drug resistance, the World Health Organisation launched a control strategy based on Directly Observed Treatment and Short-course drug therapy (DOTS). Multi-drug resistant (MDR) strains of *Mtb* are resistant to isoniazid and rifampicin, the two most powerful anti-TB drugs (WHO, 2010), and MDR-TB accounted for 511, 000 (5.3%) of cases in 2007. Of total MDR-TB cases reported 56% were in China, India and the Russian Federation; however statistical analysis and modelling has suggested that MDR levels in Africa maybe substantially higher than original estimates (Nunn *et al.*, 2007; Ben Amor *et al.*, 2008).

Even more worrying is the fact that extensively drug-resistant strains (XDR-TB) are also increasing in prominence around the world. XDR-TB is defined as resistance to rifampicin, isoniazid, any fluoroquinolone and one of the second-line injectable agents such as kanamycin, capreomycin or amikacin. It is extremely troublesome to treat, sometimes requiring up to 2 years of chemotherapy, and cases have been confirmed in more than 50 countries around the world (WHO, 2009). Diagnosis of MDR-TB and XDR-TB is also difficult and adds complications to the already problematic diagnostic area of TB, with many of the countries carrying the burden of MDR and XDR-TB lacking the diagnostic capacity to test for it. More recently Velayati *et al.* (2009) reported the emergence of totally drug-resistant (TDR) TB strains which are impossible to treat with any of the drugs currently available.

The requirement for new drugs is, therefore, in urgent demand. Most antibiotics are effective against actively dividing bacteria, and therefore when the tuberculosis bacilli enter the non-replicating state within quiescent lesions, they are often resistant to killing (Mitchison, 2000). Even with current treatments, bacilli are sometimes able to remain within infected hosts (Hu *et al.*, 2000) and potentially reactivate at any time, and the development of new drugs, with novel modes of action, is urgently required to target the persistent bacteria.

1.4.6 Drug discovery

The goal of drug discovery research is to improve the treatment of tuberculosis and to ultimately yield new medicines that combat *Mtb*. Key considerations include treatment durations, compatibility with antiretroviral treatments, lower toxic side effects, effectiveness against susceptible and drug resistant strains and action against latent infection. Any new drugs must be capable of gaining entry into the mycobacterial cell by crossing the thick hydrophobic cell wall. Creation and screening of new drugs requires innovative technologies to ensure discovery efforts are applied across multiple therapeutic areas. Identification of novel *Mycobacterium* enzymes/targets aims to provide rational drug targets that can be attacked to inhibit the bacterium.

Andries *et al.* (2005) identified a diarylquinoline compound called R207910 which can target the bacterial ATP synthase. The novel mode of action allows the drug to inhibit both drug-sensitive and drug-resistant strains of *Mtb*. Mice studies show R207910 has a higher bactericidal activity than the common isoniazid and rifampicin drugs. Absorption in mice and humans is rapid and the drug appears safe and tolerable. Ideally, a new drug would cut the treatment time down from six months, thus patients would be more likely to complete the drug course.

Hegymegi-Barakonyi *et al.*, 2008 described a new line of drug discovery focusing on the bacterial or host cell signalling including the serine/threonine protein kinases, PknB and PknG, NAD kinase and the NAD synthetases. Virtual screening with thousands of molecules indicated the potential in rapid screening methods, which aims to aid in the identification of new potent antimycobacterial substances. Pethe *et al.* (2010) emphasise the importance of *in vivo* bacterial metabolism in the drug discovery field as they describe *Mtb* growth inhibitors which have no efficacy *in vivo*. Clearly drug discovery *in vitro* is only part of the key challenge to develop new drugs to add to the current out-dated portfolio of anti-TB drugs.

1.4.7 TB Vaccine Development

Vaccines are likely to have a large impact in reducing the incidence of TB disease. The Bacillus Calmette–Guérin (BCG) vaccine has been widely used with over three billion vaccinations given worldwide since the first introduction in France in 1921. However many questions surrounding the variables that determine efficacy remain. Although offering more than 80% protection against disseminated TB infection in children, the protection offered when used in adults

The Universal Stress Proteins of Bacteria

is not reliable with the protective efficacy for preventing pulmonary TB in adolescents and adults ranging from 0 to 80% (Bannon & Finn, 1999).

There is, therefore, an urgent requirement for a newer, more effective vaccine that offers protection against all forms of TB to people across all ages. Avenues of research to produce such a vaccine include modifying the current BCG strain and attenuating *Mtb* by deleting specific metabolic pathways or virulence factors. Development of a vaccine that can be administered without needles would be also beneficial for distribution. The immune capabilities of people infected with HIV are also important in the design of any vaccine and the prospect of vaccinating latently infected individuals also needs to be considered. In fact in 2007, WHO re-recommended, in their weekly Epidemiology Record, against giving the BCG vaccination to infants known to be HIV positive, including in countries where TB is endemic (WHO, 2007) due to the risk of dissemination of the vaccinating strain.

There are currently new vaccines being trialled in humans such as recombinant BCG (rBCG), rBCG30, which over-expresses antigen 85B and which was showed to be safe and immunogenic in early trials in the USA since 2004 (Horwitz *et al.*, 2006; Hoft, 2008). A key area is developing a heterologous prime-boost strategy, which involves amplifying particular facets of the immune response, coupled with the presence of a TB antigen, to offer enhanced immune protection against TB. The MVA85A vaccine is currently in phase IIb trials in South Africa and is based on a replication-deficient *Vaccinia* virus that expresses antigen 85A (McShane *et al.*, 2004). Four sites have been developed for Phase III trials in South Africa, Mozambique, Uganda and Kenya which represents an important advance in vaccine development (Hanekom *et al.*, 2010). Any new vaccine produced requires the human immune system to provide adequate protection and the different branches of the immune system that need to be activated remain unclear as it is not definitively known what constitutes protective immunity. Vaccine development and the correlation of protective immunity remain important lines of research.

1.5 Persistence of *Mtb* within the Host

Tuberculous bacilli are able to remain viable outside the human host for several hours and, once air-borne, they pose an infectious risk to susceptible individuals. When inhaled, the bacilli reach the pulmonary alveoli within the human lung and are engulfed by alveolar macrophages, where the bacteria can survive and replicate. The initial host-pathogen interaction involves receptor mediated actions. Bacteria can be picked up by antigen presenting cells, such as dendritic cells, and

The Universal Stress Proteins of Bacteria

transported to lymph nodes, where they can disseminate to cause secondary infections in other tissues and organs such as the kidneys and brain and bones.

Different branches of the immune system are known to be important in fighting *Mtb*. Harding and Boom (2010) recently reviewed antigen presentation and the Toll-like receptors role in activating the immune response against *Mtb*. The immune response is responsible for containing or eliminating the bacteria and subjects the bacilli to a variety of stresses (see Chapter 1.7). Macrophages are activated when T lymphocytes secrete cytokines such as interferon gamma. Cytotoxic T cells can also directly kill infected macrophages, by secreting perforin and granulysin. However, *Mtb* has many survival mechanisms and has evolved specialised ways to promote survival within the human host and evade the immune response, as described below.

1.5.1 Survival within Macrophage

Macrophages are the primary defence mechanism employed against microbial invasion within the lungs (Deretic *et al.*, 2009; Nathan & Shiloh, 2000). Wolf *et al.* (2007) inferred that following inhalation of aerosolised droplets containing infectious bacilli, the alveolar macrophages are primarily the first cells infected. These macrophages make up a significant element of the innate immune response. Later they become a central component of granuloma formation, following the activation of the acquired immune response as a result of which the lesions become stratified and surrounded by a lymphocyte-rich marginal zone (Russel *et al.*, 2009; Flynn & Chan, 2005).

Once phagocytosed, *Mtb* is able to prevent killing and can replicate within the macrophage. The bacilli inhibit both phagolysosome formation and phagosomal acidification, thus allowing sustained survival and replication (Hackam *et al.*, 1998; Goren, 1977). Individual *Mtb* avoid lysosomal killing by budding into vacuoles that do not undergo phagosome-lysosomal fusion and hence do not come into contact with the highly toxic lysosomal contents (McDonough *et al.*, 1993). In addition, the production of sulfatides (anionic trehalose glycolipids) has an anti-fusion effect (Goren, 1977) and the production of ammonium chloride is used to increase the pH within the intralysosomal compartment (Hart *et al.*, 1983). van der Wel *et al.* (2007) also report that *Mtb* are able to escape the phagosome and translocate to the cytoplasm of the macrophages to aid survival.

Prevention of phagosomal maturation is achieved through an *Mtb*-induced alteration in expression of phagosomal proteins, such as Rab5, which results in the phagosome remaining in the early endosomal stage (Clemens *et al.*, 2000). Prevention of the trafficking of the bacilli to late

The Universal Stress Proteins of Bacteria

endosomal compartments also allows the bacteria to avoid contact with reactive oxygen and nitrogen species (Deretic & Fratti, 1999).

Mtb is also known to modulate the immune response in its favour, causing a reduction in interferon- γ interleukin-12 (Dao *et al.*, 2008), and, through the action of ESAT-6 and the CPF-10, *Mtb* downregulates the production of lipopolysaccharide (LPS)-induced reactive oxygen species (Ganguly *et al.*, 2007).

In addition to subverting the immune response, *Mtb* possesses numerous protection mechanisms to enhance survival in the non-optimal conditions experienced in the host macrophage. The ability of *Mtb* to survive the various stresses encountered, such as the low pH experienced in the phagosome, is discussed in more detail later in this introduction. It is also worth noting that *Mtb* is well-adapted to survival within the macrophage, importantly altering its metabolism to utilise the carbon sources available. Recently, there has been a large amount of interest in *Mtb*'s use of host cholesterol as a carbon source (Pandey & Sasseti, 2008) and the accompanied changes in lipid metabolism (Nesbitt *et al* 2010; McKinney *et al.*, 2000).

The mechanisms mentioned above allow the bacteria to survive even in the presence of a non-compromised immune response. In the majority of cases, the host immune response is merely able to control the infection, preventing active disease. Rather than being eliminated, *Mtb* can persist within the host, causing a latent infection. The infection remains dormant for many years until such a time when the host immune system can be overcome and the bacteria can reactivate, resulting in active disease. It is hypothesised that during latency the bacteria survive in a state of non-replicating persistence with minimal metabolic activity.

1.5.2 The Granuloma

The formation of granulomas is used as a containment mechanism to prevent the *Mtb* disseminating further. There are numerous types of granulomas with the size, shape and composition of the spherical mass of immune cells varying greatly within infected tissues and between different organisms (see Chapter 1.6 - Models to Study *Mtb*). Macrophages, B lymphocytes, T lymphocytes and fibroblasts are among the cells that aggregate to form a granuloma (Figure 1.3). Granulomas are formed in stages as different immune cells are recruited to the infection site. Macrophages are the key immune cells that are responsible for the formation of granulomas. Inflammation, death of neutrophils and macrophages, and toxic immune products all lead to formation a caseous core. Extracellular bacteria can then utilise the breakdown products.

The Universal Stress Proteins of Bacteria

The necrotic core is then surrounded by a mixture of lymphocytes, activated macrophages, and fibroblasts (Saunders & Britton, 2007), thus creating a granuloma.

The caseous granuloma (Figure 1.3A) is thought to be the conventional type found in humans, although other types can be present. A necrotic centre contains an array of cellular debris which is likely to be highly toxic to both host cells and to *Mtb*. Studies of granulomas from latently infected monkeys show them to be fibrotic and contain a hypoxic, caseous core, characteristics similar to latent lesions in humans (Lin *et al.*, 2009). Human granulomas are thought to be hypoxic as they do not have endothelial and blood vessel markers present (Tsai *et al.*, 2006), however, the conditions within the caseous granuloma are not well studied and understood. Recent evidence suggests within humans the granulomas are indeed hypoxic (Tsai *et al.*, 2006; Via *et al.*, 2008) although it is unclear if the bacteria are surviving within these granulomas. It has even been suggested *Mtb* survives on the periphery or in lymph nodes (Ulrichs *et al.*, 2004.)

Further granulomas described by Barry *et al.* (2009) include a non-necrotising granuloma (Figure 1.3B) which is made up of larger numbers of lymphocytes. The containment of a few bacilli appears controlled and therefore the core is not necrotic. Levels of immune attack are likely to be reduced with less reactive oxygen and nitrogen species and less killing. The third type of granuloma described is a fibrotic granuloma (Figure 1.3C) which consists of a hard spherical unit that contains minimal levels of lymphocytes and bacteria. The range of granuloma types discovered indicate the challenges in studying the *Mtb* infection and that the host–bacillus interaction relies on many factors.

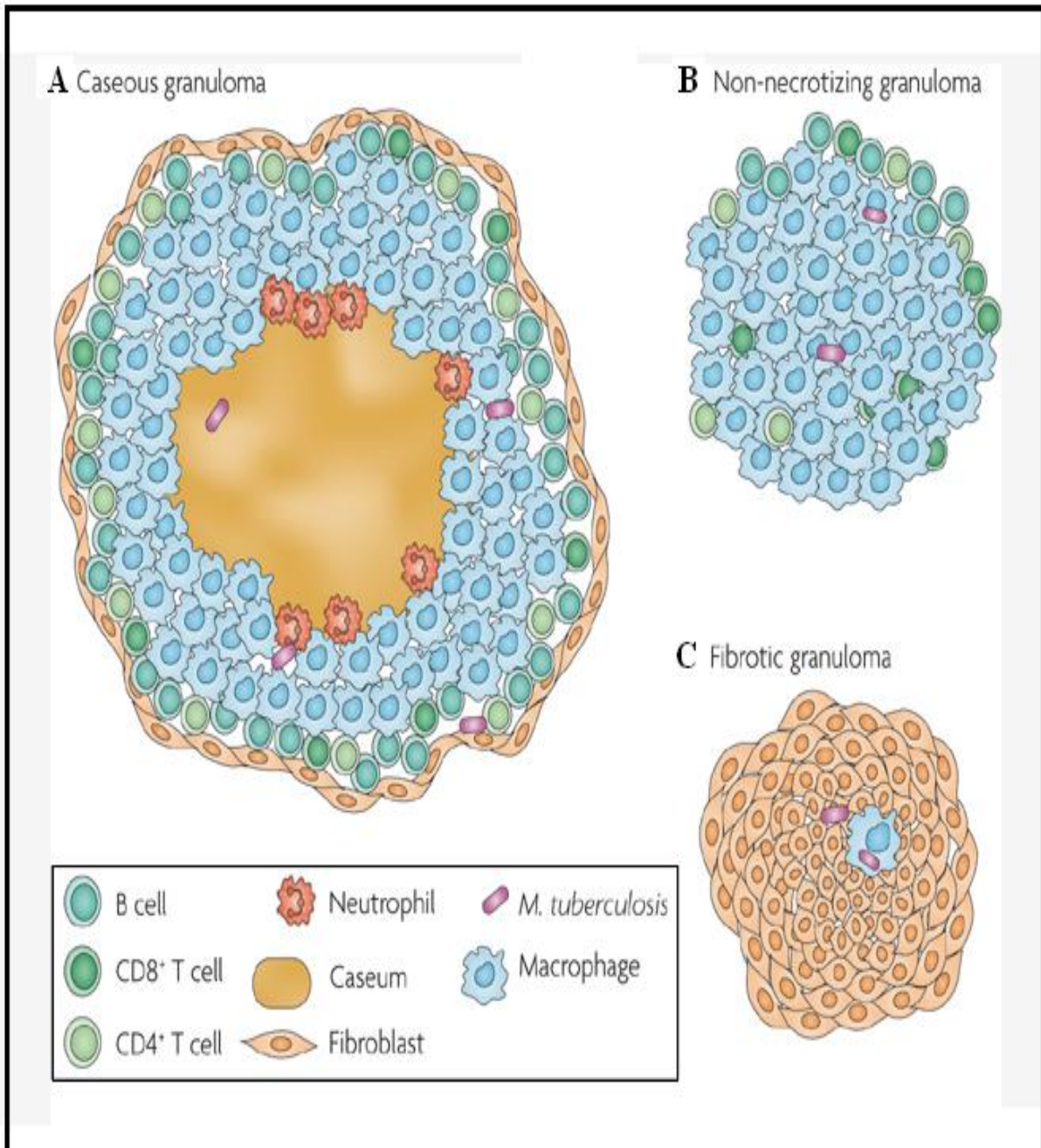


Figure 1.3. Granulomas of *Mtb* with the components of three types of granulomas found in human and non-human primates. (A) Classic caseous granuloma often found in active disease and latency cases. Microenvironments differ within the granuloma and bacilli persist intracellularly and within the hypoxic necrotic core. (B) Non-necrotizing granuloma often found in active disease where lymphocytes and macrophages surround the bacilli with the lesion. (C) Fibrotic granuloma often found in latent infections. Few *Mtb* remain and are surrounded by fibroblasts with few, if any, macrophages present. Diagram from Nature reviews Microbiology (Barry *et al.*, 2009).

1.5.3 Latency, and Non-Replicating Persistence

Latent and chronic infections of *Mtb* are clinical terms used to describe the nature of the infection. Latency is a clinical observation referring to an individual that has an infection without showing any apparent signs of disease, whereas a chronic infection refers to an ongoing infection with clinical symptoms. A key feature of *Mtb* is therefore its ability to evade the immune system and prevent the elimination from the body, thus creating a huge reservoir of infection, until conditions are more suitable for re-activation.

Wayne first coined the term non replicating persistence (NRP) in 1954 when studying the growth of *Mtb* under different oxygen tensions. In a classic study, Wayne and Hayes, 1996, grew *Mtb* in sealed vessels and a defined volume of headspace. As the cultures grew, the oxygen levels were gradually depleted and the bacteria entered the NRP state. Two phases of adaptation to the hypoxic conditions were described in relation to the changes observed in the optical density of the cultures. They were named NRP1, with oxygen saturation at 1%, and NRP2, with oxygen saturation at 0.06%.

The increased ability of the bacteria in stationary phase to survive particular stresses correlates with physiological changes (Kolter *et al.*, 1993, Smeulders *et al.*, 1999). These physiological changes are mediated by differing patterns of gene expression for NRP1 and NRP2 (Wayne & Sohaskey, 2001). In active disease, *Mtb* cells can show signs of being dormant. Garton *et al.* (2008) analysed mycobacterial cells from sputum from patients with active disease and visualised evidence of dormant cells including presence of lipid bodies and transcript profiles similar to NRP cells. Physiological changes include remodelling of the cell wall as shown by Seiler *et al.*, 2003.

The state in which *Mtb* survives during latent infection has been termed dormancy in the past. However, this may not be the most accurate description of the bacteria. It remains to be established whether the bacilli are persisting in a dormant or suppressed state (Parrish *et al.*, 1998). Dormancy analogous to that observed in the formation of spores by species such as *B. subtilis* implies that the bacteria cease all metabolic activity. However, it is likely that, during persistence, TB bacilli maintain a low level of metabolic activity and gene expression.

Aunchin *et al.*, 2009 studied non-replicating forms of *M. smegmatis* and suggested the viable but non-culturable state was associated with a transition towards coccoid and oval cellular shapes, which resembled spore-like morphology. It should be noted that the possibility does exist that mycobacterial may enter a true, spore-like state, however this seems unlikely as mycobacteria do not contain the genes necessary for spore formation as seen in *B. subtilis*. Ghosh *et al.*, 2009 suggested *M. marinum*, a fish pathogen, when grown to stationary phase in long-term cultures can

form spores and Tragg *et al.*, 2010 questioned the observation after setting out certain criteria for the dormant forms to be called spores. Whether the tubercle bacillus can form spores remains a question to be answered, but if so, could represent a key area for drug discovery.

1.5.4 Reactivation, Resuscitation and Disease

Mtb can remain in the host for substantial periods with Lillebaek *et al.*, 2002, reporting molecular evidence of endogenous reactivation of *Mtb* after 33 years of a latent infection, indicating bacterial infection can remain established for decades. The reactivation of the infection promotes transmission of the bacilli within droplet aerosols, allowing the bacteria to spread to new, naive hosts. The reactivation is thought to occur in between 5-10% of individuals who are latently infected. Reactivation of *Mtb* has been investigated by studying the ability of the bacteria to reactivate when grown to stationary phase in *in vitro* models. Under specific conditions, a viable but non-culturable state was observed in *Mtb*. When cultured under nutrient limiting conditions, such as 4 month cultures maintained in stationary phase, the bacteria could not form CFU when plated out on nutrient agar, but could be resuscitated in supplemented liquid media (Shleeva *et al.*, 2004).

The reactivation of *Mtb* requires an upregulation of gene expression and involves the action of resuscitation promoting factors (*Rpfs*). These were identified when the secreted *Micrococcus luteus* Rpf was shown to be able to promote the resuscitation of dormant mycobacteria (Mukamolova *et al.*, 2002). In fact, recombinant mycobacterial Rpf was shown to increase the recovery of *Mtb* from the sputum of active TB patients (Mukamolova *et al.*, 2010). Long-term persistent TB remodels the cell wall with different crosslinks between the peptidoglycan. When it resuscitates, the wall needs to be remodelled to allow growth and the *Rpfs* appear to be involved in this. Crystal structures show Rpf proteins have a C-type lysozyme-like fold which is predicted to allow enzymatic cleavage of the glycosidic bond between *N*-acetyl muramic acid and *N*-acetyl glucosamine in the peptidoglycan layer of the cell wall (Cohen-Gonsaud *et al.*, 2005). There is also some suggestion that the release of muropetides as the result of Rpf cleavage acts as a signal for exit from latency (Kana *et al.*, 2010).

Through understanding the mechanisms of reactivation from the latent state there is the potential for a further understanding of both the host immune and bacterial response and of developing drug targets.

1.6 Models to Study *Mtb*

1.6.1 *In vivo* Models

Common animal models used to study infection include mice, guinea pigs, rabbits and primates. No individual model is completely suitable and the validity of each model system is widely debated. The murine model, although widely used, has a different disease progression to that seen in humans and the bacterial burden remains high throughout infection with the bacteria continuing to replicate during chronic infection (Gill *et al.*, 2009). Granulomas, if formed, are small and poorly differentiated, and recent work by Via *et al.*, 2008 suggests the mouse granulomas appear to lack necrosis, fibrosis or hypoxia. Different versions of murine models are used including SCID (severe combined immunodeficient) mice and the Cornell model, where mice are treated with antibiotics to try to induce a persistence-like infection and where reactivation ensues upon administration of steroids. Human granulomas provide the bacteria with unique environments which are likely to be important when studying chemo- and immunotherapy.

Guinea pigs and rabbits produce mature lesions which show similar caseous necrosis to those in humans (Via *et al.*, 2008). However, these two models also have disadvantages as guinea pigs are very sensitive to TB, with no latency described as yet, and rabbits have to be infected with *M. bovis* as they have a natural resistance to *Mtb* (Flynn, 2006). Primate models provide a model with the most similarities to human disease, including disease progression, the immunological response and latency. However, working with primates has very high cost, high space requirements and time constraints, and there are also ethical considerations.

The use of animal models aims to provide further insight into the progression of the infection within humans, which is otherwise difficult to evaluate without damaging invasive methods. The search for biomarkers, such as metabolites or immunological antigens, within blood/urine to ascertain the status of the infection with humans would be a beneficial research tool (Doherty *et al.*, 2009).

To evaluate disease progression within humans, imaging tools such as positron emission tomography (PET) and computerized axial tomography (CAT) are being developed to allow visualisation of disease progression and the bacterial burden *in vivo*. CAT scans provide structural information of the lung and lymph nodes and can highlight the presence of granulomas. The probes used in PET, such as ¹⁸F- fluorodeoxyglucose (FDG), can then mark tissues and cells of interest. The temporal visualisation of TB infection as disease progresses shows that the populations of bacteria within one individual are not homogenous and there is a dynamic nature within the individual

lesions as reviewed by Barry *et al.*, 2010. Thorough studies that compare the disease in animal models to humans are crucial to move the TB therapy and TB vaccinology fields forward.

1.6.2 *In vitro* Models

Mtb is able to survive in human hosts through employing numerous strategies to promote survival. To study the bacteria *in vitro*, models have been developed that aim to replicate particular stages of infection.

Mtb grown under nutrient limiting conditions is driven towards a persistent state and this nutrient starvation model is used to represent the nutritionally starved mycobacteria that persist within lung lesions. Nyka (1974) showed that *Mtb* could be re-grown even after being starved of nutrients for up to two years. Betts *et al.* (2002) performed microarray and proteome analysis on nutrient starved *Mtb* showing that alterations in metabolism, transcription and biosynthesis were made in order to maintain long-term survival.

Alternative *in vitro* models involve growing *Mtb* in hypoxic conditions. Wayne and Hayes, 1996, suggested the non-replicating persistent state could be mimicked *in vitro* through growth of *Mtb* in an oxygen deprived environment. The Wayne model relies on a self-generated O₂ gradient and therefore the availability of the nutrient oxygen is the initial limiting factor. The bacteria within a sealed vessel deplete the oxygen present, with the culture becoming progressively more hypoxic, until a shift in the *Mtb* physiology occurs (Wayne 1977, Wayne & Lin, 1983). Rodrigue *et al.* (2006) showed the alternative sigma factors *sigB*, *sigE* and *sigF* were all upregulated in the model and are thought to drive the bacteria towards the non-replicating persistent state observed.

A further hypoxic model termed the defined hypoxia model is thought to be more reproducible and relies on growing *Mtb* in defined hypoxic atmospheres that are controlled and remain constant throughout culture. Using an oxygen tension of less than 1% inhibits growth; however the bacteria remain viable (Yuan *et al.*, 1998). These models, although not proven to directly represent the conditions within the lung, aim to provide a basis for studying the bacteria. Other *Mycobacterium* species including *M. bovis* BCG and *M. smegmatis* are also used to study the non-replicative persistent state *in vitro* through promoting entry to stationary phase by growing the cells in hypoxic conditions.

1.7 Stress Responses of *Mtb*

The persistence associated with latent infection and expansion associated with active disease of the *Mtb* bacterial population within a granuloma invariably involves a shift in the balance of the physiological state of the bacteria. Investigating the infection cycle in mice, Barry (2009) showed bacterial strains with gene disruptions in particular metabolic pathways exhibited different defects over the course of the infection cycle, thus suggesting that the ability of the bacteria to differentially regulate its metabolism during the progression of disease is important. These metabolic shifts are likely to be driven by the varying conditions experienced within the host over the course of the lifecycle of *Mtb* and the infection stage. The availability of certain nutrients and levels of oxygen to *Mtb*, and the amount of stress induced by infected host cells and the immune response would all be transiently changing in a temporal manner.

Mtb has developed ways to deal with these types of stresses, whether biotic or abiotic in nature, through biological processes that involve responding to certain stimuli. A response to stress can be defined as 'a change in state or activity of a cell or organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a disturbance in the organismal or cellular homeostasis' (EMBL database). Over the life cycle of *Mtb* there are key stages where the bacteria may encounter different stresses. *Mtb* has a complex set of regulatory proteins which includes thirteen sigma factors (SigA-M) which all fall into the Sigma-70 family (Cole *et al.*, 1998). When combined with specific promoter regions on the genome the factors can combine with RNA polymerase in order to allow the modulation of transcription in response to external and internal factors. SigA is the primary transcription factor and is essential for growth due to its importance in general housekeeping gene control (Doukhan *et al.*, 1995). The alternative sigma factors are then up-regulated in response to particular signals.

Extracellular survival is important to ensure transmission between susceptible hosts and possible stresses encountered here include temperature shock such as cold shock, which up-regulates *sigI* (Manganelli *et al.*, 2004), as the outside temperature is likely to be lower than the optimal 37 °C encountered when inside the human host. Hu and Coates (1999a) postulated that the surface of the alveoli within the lung may induce a detergent-type response due to the array of phospholipids present, and it has since been found that *sigB* and *sigE* are up-regulated in response to SDS-detergent (Manganelli *et al.*, 2004). *sigD* is up-regulated in response to nutrient starvation (Ramam *et al.*, 2004), *sigB* is up-regulated in response to oxygen depletion and heat stress (Manganelli *et al.*, 1999) and *sigG* is up-regulated during macrophage infection (Capelli *et al.*, 2006).

The Universal Stress Proteins of Bacteria

An *Mtb* knockout in *sigF* was shown to have reduced virulence in mice (Chen *et al.*, 2000) and the gene was shown to be upregulated in stationary-phase and after exposure to several stress agents in *M. bovis* BCG (DeMaio *et al.*, 1996).

Other regulatory mechanisms, including two component systems, are employed by *Mtb* to promote survival in challenging conditions. An important two component sensor response system with significance to this project is the DosS/DosR system, which is described in Chapter 1.8.

A further level of regulation occurs predominantly at the protein level via protein phosphorylation, a detailed discussion of which is beyond the scope of this introduction. It is worth noting that it has been suggested that *Mtb* is able to modify host proteins in addition to its own. The tyrosine phosphatase PtpB has no known partner in *Mtb* and likely acts upon host proteins, leading to the suggestion the protein could be a possible drug target (Zhou *et al.*, 2010). An interesting area of contention is whether the serine/threonine protein kinase PknG is able to phosphorylate protein kinase C and thus promote intracellular survival (Walburger *et al.*, 2004).

The purpose of the multiple levels of regulation is to allow the bacteria to survive stresses including hypoxia, oxygen free radicals, nitric acid, pH, temperature shocks, nutrient starvation and other stresses caused by the host immune response.

1.7.1 Reactive Oxygen Radicals

The hostile environment within macrophages is characterised by the abundance of toxic chemicals including oxygen and nitrogen radicals. Cytotoxic reactive oxygen intermediates (ROIs) cause damage as they are highly oxidising and can react with bacterial cell walls and particularly the polyunsaturated phospholipids of the cell membrane. Gordon and Hart (1994) showed that macrophages produce ROIs in response to intracellular *Mtb*. The ability of *Mtb* to survive and replicate within the toxic environment of the host macrophage is enhanced by enzymes such as superoxide dismutase (SOD), which is an antioxidant that can detoxify superoxide into oxygen and hydrogen peroxide. Methods to deal with peroxide complement this pathway, with *Mtb* producing a protein that cyclopropanates mycolic acid double bonds, making them ten-fold less susceptible to peroxide (Yuan *et al.*, 1992).

1.7.2 Reactive Nitrogen Intermediates

Reactive nitrogen intermediates (RNIs) produced by activated macrophages are also a key element of the immune response. When murine macrophages are subjected to interferon- γ they are able to produce nitric oxide through the action of the enzyme nitric oxide synthase 2 (NOS2). RNIs can also be produced when nitric acid is auto-oxidised and results in the accumulation of nitrite (NO^{2-}), which can then generate and re-generate the radicals NO and NO_2 when at a low pH (Stuehr & Nathan, 1989). The free radicals then damage proteins, lipids and DNA and are used as antimicrobial agents. Exposure to NO results in a similar pattern of gene expression as when the bacilli are subjected to hypoxia, as both conditions inhibit normal respiration (Voskuil *et al.*, 2004).

Using mice models to study the role of RNI has indicated the importance of RNI. Mice deficient in producing RNIs have shown an increased susceptibility to *Mtb* infection (Flynn & Chan 2003; Chan & Flynn 2004). NOS2 is therefore important for containing acute and chronic infection but not involved directly in the complete eradication of the bacilli. To avoid damage, *Mtb* has several factors that contribute to the survival, including the *noxR1* and *noxR3* genes that are known to have a key role in allowing the bacteria to survive the toxic effect of RNIs, however the exact mechanism of action of these genes that confer resistance remains unknown (Ruan *et al.*, 1999).

1.7.3 pH Stress

As macrophages mature, they undergo an acidification process that can reach a pH of 4.5. When Metchnikoff (1905) discovered phagocytosis he speculated that the waxy cell wall of *Mtb* is likely to allow the bacteria to survive in these challenging conditions. *Mtb* is able to prevent the acidification and halt the pH drop at a less damaging pH 6.3 (Sturgill-Koszycki *et al.*, 1994). This low pH still requires some gene up-regulation including genes that have role in metabolism, lipid metabolism, virulence, cell wall proteins and regulatory proteins (Fisher *et al.*, 2002). Vandal *et al.* (2008) showed that *Mtb* survived in simple buffer at pH 4.5 and maintenance of the intrabacterial pH was dependent on the membrane associated protein Rv3671c. Identified through transposon mutagenesis, Rv3671c was shown to be necessary to maintain intrabacterial pH *in vitro* and in activated macrophages.

1.7.4 Nutrient Starvation

The non-replicating state of the bacilli is suggested to be a result of the lack of availability of nutrients including oxygen, carbon and amino acids. The stringent response to these conditions is initiated by the accumulation of hyper phosphorylated guanosine (p)ppGpp and results in a variety of protective measures to ensure persistence. Steps taken include the induction of stress response genes and the shutting down of ribosomal RNA transcription machinery, which has a high metabolic cost.

Mycobacteria alter their gene expression using different sigma factors such as SigD and SigM, and begin protein degradation processes, stabilise RNA species including rRNA and tRNA, and inhibit DNA replication. The *Rel* gene can hydrolyse and synthesise ppGpp and knockouts were shown to be unable to establish long term infection in the Cornell mouse model (Dahl *et al.*, 2003).

1.7.5 Hypoxia

Hypoxia is widely thought to be a key trigger for the *Mtb* bacilli to enter the non-replicating persistent state. *Mtb* infections are preferentially associated with the most oxygen rich sites in the body, indicating the importance of oxygen to the reactivation and continued success of the pathogen. The host may induce hypoxic stress on the bacterium to limit the growth of the pathogen *in vivo*. With the adaptive immune response mounted against the bacteria a granuloma forms when the immune cells surround and entrap the bacteria, and oxygen is thought to be limiting within this environment. In addition, if the granuloma proceeds to a calcified state, as seen in humans, the formation of the calcified surrounding is thought to prevent gaseous exchange resulting in the oxygen levels slowly depleting, thus leaving the encased bacteria in a hypoxic environment. The altered metabolic state persists until the opportunity to reactivate arises.

The survival under challenging hypoxic conditions involves two main responses that have been named the DosR transcriptional response and the enduring hypoxic response. In the short term (<8 hours), the DosR response involves the upregulation of approximately 50 genes (Roberts *et al.*, 2004); however, over increased time more than 400 genes are shown to be upregulated by day 4 of oxygen limitation. After one week of hypoxic conditions, up to 730 genes are upregulated and the subset of genes induced between day 4 and day 7 have been named the enduring hypoxic response (Rustad *et al.*, 2008). Sigma factors SigE and SigH are two such genes that are upregulated within this period. The enduring hypoxic response genes are upregulated several days after the *dos*

regulon but then return to basal levels (Rustad *et al.*, 2009). It is clear the long term survival of the bacteria and entry into a non-replicating persistent state is multifactorial and hypoxia alone is only one of the contributing factors promoting latency within the host.

1.8 The DosR Regulon

Studying the transcriptomic response to hypoxia, using microarrays, Sherman *et al.*, 2001, identified roughly 100 genes that underwent altered expression. Metabolic activity was notably reduced in genes relating to lipid, protein and amino acid synthesis and DNA replication. Among the upregulated genes was the two component sensor-regulator DosR/DosS system. Tetrameric DosR is able to interact with DNA and through interactions with the sensors DosS/T, which are both histidine kinases, DosR acts as a transcriptional regulator. The DosS protein is able to indicate the net redox state of the cell as it can bind ferrous iron, whereas the DosT protein is able to bind haem bound to O₂ and therefore directly senses the oxygenic state of the cell.

Subsequently, studies using a knockout *dosR* strain, determined that the gene was required to allow induction of nearly all the initial genes induced by hypoxia (Park *et al.*, 2003). Further studies have since linked the DosR response to an array of stress conditions including the response to nitric oxide (Voskuil *et al.*, 2003), in macrophages (Schnappinger *et al.*, 2003) and in early (Karakousis *et al.*, 2004) and late (Shi *et al.*, 2003) mouse models of TB infection, to carbon monoxide (Kumar *et al.*, 2008), SDS (Pang *et al.*, 2007) and low pH (Rohde *et al.*, 2007). However, the *DosR* knockout strain did not have the strong growth phenotype that was predicted and could enter bacteriostasis when grown *in vitro* in the Wayne hypoxia model. The mutant strain had no phenotype until week 3, where a 1-2 log survival defect was present.

The DosR transcriptional response is, therefore, not synonymous with the dormancy response and is now thought to play a more focused regulatory role in the response to nitrosative/redox stresses and in response to other respiratory challenges. In a recent study Leistikow *et al.* (2010) showed a mutant unable to upregulate the *Dos* regulon had defects in maintaining ATP and NADH levels. The mutant showed delayed resuscitation from the exit of hypoxia in growth experiments.

Of the approximately 50 genes that were highlighted by Park *et al.* (2003) as being induced during the DosR response, only a small number have a known function. Upregulated genes for which a function is known include the *narX*, *nark2*, *fdxA* genes, involved in nitrate accumulation and alternative electron transport, and the *nrd3* and *tgs1* genes, involved in dNTP and triglyceride

The Universal Stress Proteins of Bacteria

synthesis respectively. The majority of genes, however, were found to be conserved hypothetical proteins. Interestingly, a number of the conserved hypothetical proteins were classed as putative proteins that belonged to the Universal Stress Protein (USP) superfamily. In *Mtb*, the USPs Rv2623, Rv2005c, Rv2028c, Rv2624c, and Rv3134c were all shown to be upregulated when oxygen is limiting via the DosR-mediated response (Sherman *et al.*, 2001). The orthologous two-component response regulator in *M. smegmatis*, *DevR*, has also been shown to control the expression of some of the USPs present in this bacterium (O'Toole *et al.*, 2003).

These USPs, therefore, may help the *Mycobacterium* adapt and survive when under hypoxic conditions. As some USPs were discovered as members of the DosR regulon, it is possible that understanding USP function could lead to further understanding of the DosR hypoxic adaptation. This is supported by the fact that nitric oxide also induces the DosR regulon described and the discovery that Rv2623 is expressed after phagocytosis by the host macrophage (Mohan *et al.*, 2001).

Part II: Universal Stress Proteins

1.9 *E. coli* UspA

When *E. coli* cells undergo the transition from exponential growth phase to stationary growth phase, there are a number of physiological changes undertaken to promote survival. A variety of cellular processes including cell wall structural organisation, gene expression, DNA organisation and metabolism are modified (Kolter, 1995). The changes result in the cells becoming resistant to a variety of stresses including heat shock, UV light, changes in pH, osmotic shock and oxidation (Jenkins *et al.*, 1988, Lange & Hengge-Aronis, 1991, Small *et al.*, 1994).

Nystrom, & Neidhardt, 1992, identified the key protein Universal Stress Protein A in *E. coli* and described the protein as having unique responsiveness to diverse stress conditions. The gene was found to be upregulated in response to a wide variety of environmental stresses such as heat, ethanol, osmotic, nutrient starvation (including stationary phase) and damaging agents such as antibiotics, oxidants and metals. A cytoplasmic protein, UspA was shown to be one of the most abundant proteins in growth arrested cells. *E. coli* with an inactivated *uspA* gene was found to be more sensitive to the wide array of stresses. The name 'Universal Stress Protein' has continued to be used and subsequently related genes have been placed in the UspA superfamily.

The Universal Stress Proteins of Bacteria

UspA was found to be a serine and threonine phosphoprotein and it was shown that the phosphorylation *in vivo* is dependent on the TypA phosphotyrosine protein (Freestone *et al.*, 1997, Freestone *et al.*, 1998). The transcriptional regulation of *uspA* was shown to occur through a σ 70-dependant promoter (Nystrom and Neidhardt, 1992); however, differential expression suggests there must be other control levels of transcription initiation. Indeed, Nystrom and Neidhardt, 1994, showed that transcription of *uspA* is rapidly turned off when the growth arrested bacteria are subjected to a nutritional up-shift. *uspA* is part of the *FadR* regulon, which is involved in fatty acid and membrane lipid metabolism (Farewell *et al.*, 1996).

Further analysis of the *E. coli* genome (Blattner *et al.*, 1997) revealed the presence of six *usp* family genes, *uspA*, *uspC*, *uspD*, *uspE*, *uspF* and *uspG*. After the name was widely accepted, the UspA gene was shown to be repressed in some stress conditions including exposure to the antibiotic tetracycline and cold shock (Nyström & Neidhardt, 1993).

1.10 UspA Superfamily

1.10.1 Distribution

The Universal stress protein A (UspA) superfamily is a conserved group of proteins present in Eukaryotes, Eubacteria and Archae (Figure 1.4). USPs all contain at least one USP domain consisting of ~150 amino acids. The role of the domain is, however, unknown. The abundance of USP domains varies depending on the organism. In Archaea, Cyanobacteria and some plants, USP domains have been shown to be present in proteins that contain other functional domains such as protein kinases and amino acid permeases (Kvint *et al.*, 2003).

The Universal Stress Proteins of Bacteria

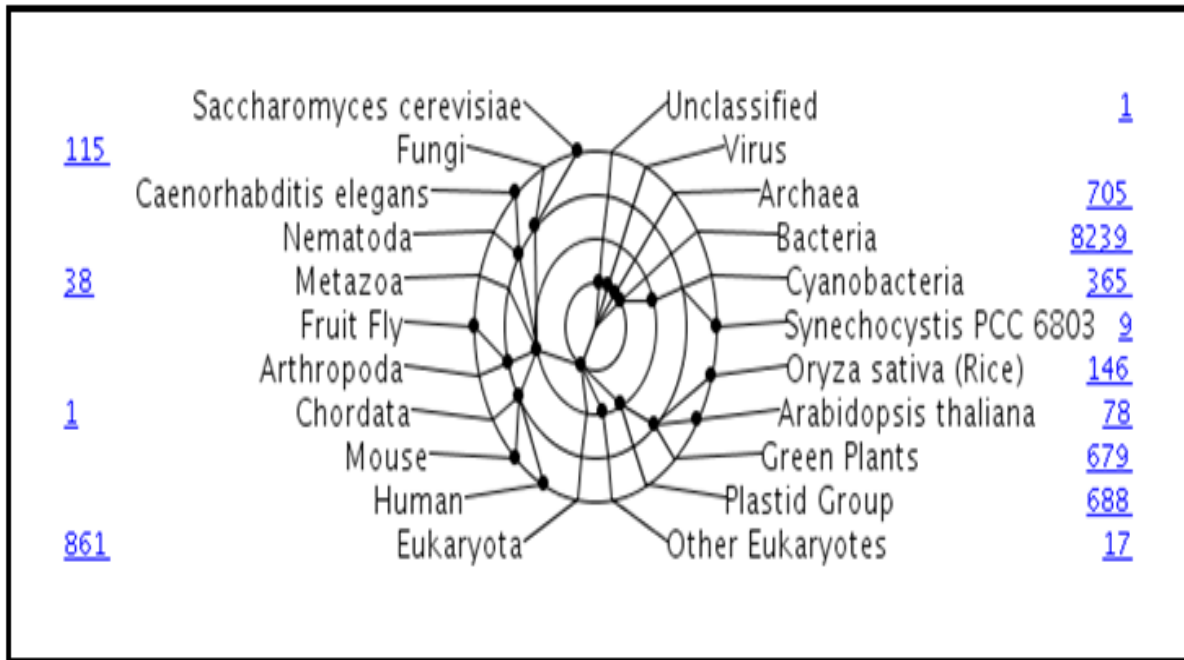


Figure 1.4. Taxonomic distribution of the universal stress protein domains across the kingdoms of life. The numbers refer to the number of USP domain containing proteins within the classification groups listed. The distribution web contains USPs that have been identified by genome sequencing and analysis as of December 2010 (EMBL Database).

USPs are found in 106 species of Eukaryotes including 4 species of metazoan, 43 species of plants including common crops such as barley, rice and wheat, and 49 species of fungi, including industrial important yeasts and penicillin species. The 38 domains found in metazoans include a single USP in the chordate *Salmo salar* (Atlantic salmon); 18 USPs in two species of Blood Fluke in *Platyhelminthes* and finally 19 USPs found in *Nematostella vectensis* (Starlet Sea anemone). The number of USP domains in higher eukaryotes is low suggesting they do not require this family of proteins or that other proteins have replaced the biological roles of USPs.

USPs in the plant kingdom have been shown to be involved in stress adaptation when rice is submerged in water (Sauter *et al.*, 2002), response to pathogens in barley (Steiner-Lange *et al.*, 2003) and fruit ripening in tomatoes (Zegzouti *et al.*, 1997). In rice and tomatoes, the expression of USPs has been linked to the plant hormone ethylene and it was shown changes in metabolic activity leads to an increase in expression of the USPs.

The Universal Stress Proteins of Bacteria

1.10.2 Structural Forms of USPs

USPs appear in three main structural forms; single domain USPs, tandem domain USPs and USP domains fused with other catalytic domains. Bacterial *usp* genes normally encode either a small USP that contains one USP domain or a larger USP consisting of two tandem USP domains (Figure 1.5). Single and tandem protein domains have been shown to make up approximately 80% of USPs in Eubacteria and Archaea, but only 50% in Eukarya (Siegele, 2005). Siegele (2005) describes that the biochemical and biological functions of USPs largely remain elusive. The ubiquitous nature of USPs in prokaryotes and limited distribution in eukaryotes means the biological role of the superfamily remains an intriguing unresolved enigma.

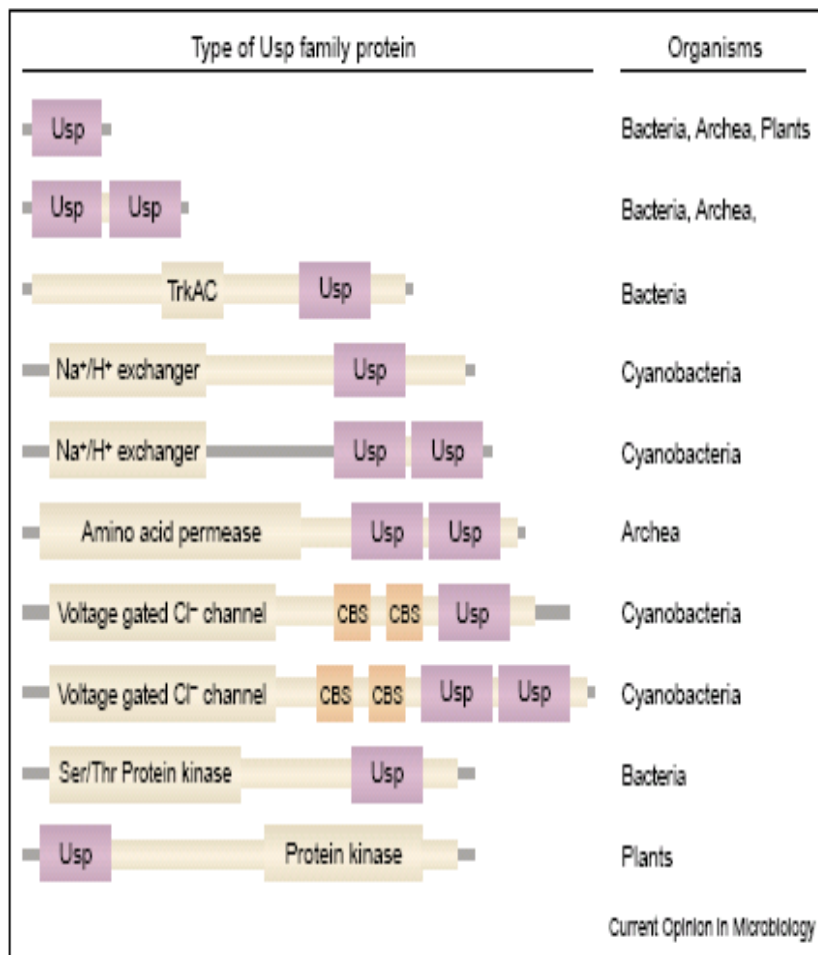


Figure 1.5. Schematic representation of different UspA family proteins and their presence in different organisms. Cystathionine-b synthase (CBS) domains are small intracellular modules of unknown function. The exact function of the TrkAC domain is unknown. (Kvint *et al.*, 2003).

1.11 USP Classification

Crystal structures of the UspA protein from *Haemophilus influenzae* (subfamily UspA) and MJ0577 protein from *Methanococcus jannaschii* (subfamily UspFG) show the proteins share similar structures, with both forming stable asymmetric homodimers (Sousa and McKay, 2001, Zarembinski *et al.*, 1998). Interestingly, when MJ0557 was crystallised it was shown to be bound to ATP. Combining structural and sequence alignment data analysis, Sousa and McKay (2001) identified the ATP-binding motif, G-X₂-G-X₉-G(S/T) from amino acid residues 127 to 141. Although most USP domains do contain the consensus sequence it should be noted that not all USP proteins have this motif, for example UspA from *Haemophilus influenzae*, which is unable to bind ATP (Sousa & McKay, 2001).

1.11.1.1 PP Loop Superfamily

The PP loop superfamily takes its name from the ATP ProPhosphatase domain originally found by analysis of conserved amino acid sequence motifs that were shown to be present in four groups of enzymes capable of hydrolysing ATP between the α - β phosphates. The four enzyme groups are GMP synthetases, asparagine synthetases, argininosuccinate synthetases and ATP sulfurylases (Bork & Koonin, 1994). The conservation of particular amino acids gave similar secondary structures within these proteins and it was hypothesised that the motif found was actually a modified version of the common P-loop nucleotide binding domain, which is involved in contact with the nucleotide.

The PFAM database describes the USP family (Pfam 00582) as belonging to the PP-loop clan (CL0039). With increased statistical analysis of amino acid sequences there are now 12 members of the clan including Argininosuccinate synthase, Asparagine synthase, ATB_bind_3, ATP_bind_4, DNA photolyase, Electron transfer flavoprotein, ExsB, NAD synthase, Phosphoadenosine phosphosulfate (PAP) reductase family, Thiamine biosynthesis protein (ThiI), tRNA methyl transferase and finally the USPs. The so called PP-motif does show alterations and modifications between the members of the clan, showing that evolution of new enzymatic functions for the PP-loop structure has lead to a diverse range of functional activities. An example is the homotetrameric argininosuccinate synthase, found in prokaryotes, archaeobacteria and eukaryotes, that is an enzyme involved in the urea cycle. The enzyme is involved in arginine biosynthesis and catalyses the ATP dependent ligation of citrulline and aspartate to form argininosuccinate, AMP and pyrophosphate.

The Universal Stress Proteins of Bacteria

Through studying the USP folds and ligand binding, Aravind *et al.* (2002) suggested that USPs, the electron transfer flavoprotein (ETFP) and the photolyase family may all be related and may form part of a much larger clade of USP-like domain proteins. The α/β folds within these proteins was termed the HUP domain (HIGH-signature proteins, USP-like domains and PP-loop ATPases) and it was proposed that these proteins have a common ancestor.

The Superfamily database describes the Adenine nucleotide alpha hydrolases-like superfamily as having a general functional role in metabolism with particular reference to nucleotide metabolism and transport. The SCOP (Structural Classification of Proteins) database places the USP protein family with this Adenine nucleotide alpha hydrolases-like superfamily. The database classifies the superfamily as having an adenine nucleotide alpha hydrolase-like fold consisting of a core with a three layer alpha/beta/alpha core and a parallel beta-sheet of five strands. This type of fold is then placed within a class called alpha and beta proteins (α/β) which consists of mainly parallel beta sheets (beta-alpha-beta units).

1.11.1.2 The G-X₂-G-X₉-G(S/T) Nucleotide Binding Motif

The conserved G-X₂-G-X₉-G(S/T) nucleotide binding motif perhaps gives some indication as to the role of the USPs. MJ0577 protein was first crystallised with ATP-bound nucleotide present (Figure 1.6A). Zarembinski *et al.* (1998) stated that MJ0577 itself did not hydrolyse ATP, although it can with the addition of *Methanococcus* cellular lysate; suggesting there maybe factors that are present to allow ATPase activity *in vivo*. The conserved residues were shown to assist in the coordination of the ATP nucleotide. Gly127 hydrogen bonds to ribose, Gly130 hydrogen bonds with the β -phosphate, Gly140 interacts with the α -phosphate and Ser141 hydrogen bonds to the γ -phosphate of ATP (Figure 1.6B). Other conserved residues which interact with ATP, such as aspartate at position 13 and valine at position 41, which is in contact with the base, were also described (Zarembinski *et al.*, 1998). However, the significance of the ATP-binding motif remains unknown (O'Toole & Williams, 2003).

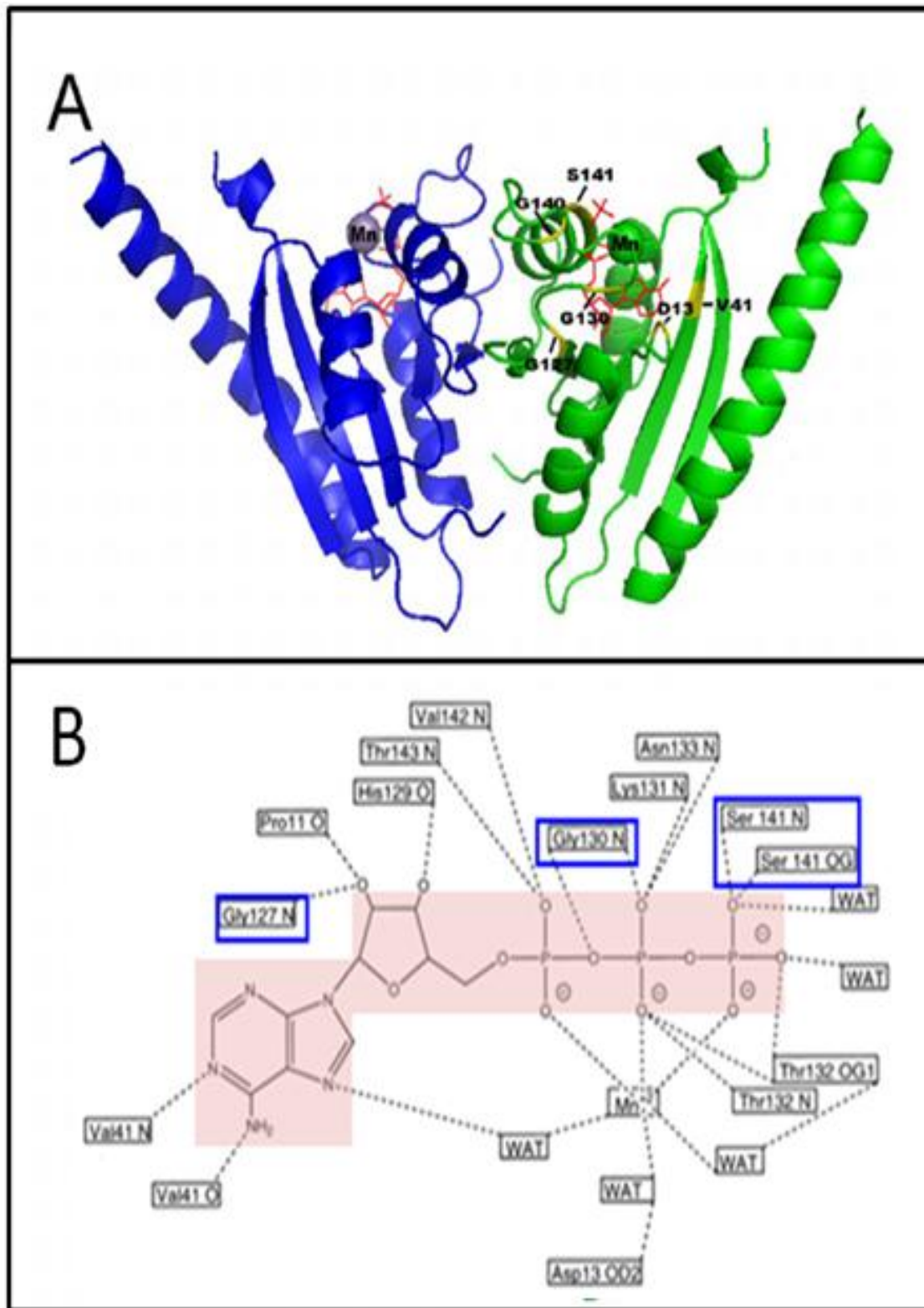


Figure 1.6. ATP binding of the *Methanococcus jannaschii* MJ0577 protein. (A) Crystal structure of MJ0577 protein with highlighted residues. (B) Conserved nucleotide binding residues are boxed. The ATP molecule is shaded in pink.

The Universal Stress Proteins of Bacteria

UspA from *H. influenzae* was shown not to have the ATP binding motif and, in the place of the glycine residues, were the bulky amino acids glutamine and methionine. These bulky residues are likely to impair the ATP-binding capabilities of the protein (O'Toole & Williams, 2003). The USP superfamily was suggested to be grouped in to two distinct classes; based on the presence/absence of the ATP-binding motif (Sousa & McKay, 2001; Kerk *et al.*, 2003). The two groups are likely to have evolved to perform different roles within the cell. The USPs which do not contain the putative nucleotide binding motif likely form a group of ancestors to UspA-like proteins (Aravind *et al.* 2002). Possible functions suggested for USPs include that they may serve as molecular switches, capable of relaying signals depending on the nucleotide bound state (Zarembinski *et al.*, 1998; Aravind *et al.*, 2002).

1.12 USP Function

1.12.1 Function of *E. coli* USPs

Studies of *E. coli* mutants which failed to induce the *uspA*, *uspC*, *uspD* and *uspE* genes revealed mutations in the genes *relA* and *spoT*. These two genes have important roles in the production of hyperphosphorylated guanosine (ppGpp) (Xiao *et al.*, 1991) and it has been shown ppGpp has a role in controlling the expression of *uspACD* and *E*. When *E. coli* is faced with nutrient starvation, the bacteria undergo a stringent response which involves ppGpp accumulation and repression of ribosomal and transfer RNA synthesis. Other work on ppGpp levels by Jones *et al.*, 1992, showed that a temperature downshift applied to *E. coli* resulted in lower levels of ppGpp production. The reduced levels of ppGpp may explain the lack of UspA induction under the cold shock stress condition.

The proteins UspACD and E form a class of proteins which have a role in protecting the cell against DNA-damaging agents (Kvint *et al.*, 2003). These USPs have been linked to survival in the presence of genotoxic agents, defence against oxidants and respiratory uncoupling agents (Nystrom & Neidhardt 1993, 1994, Farewell *et al.*, 1996, Diez *et al.*, 2000). The other USP proteins in *E. coli* are the UspFG type which are ATP-binding proteins and therefore potential ancestors of UspA-like proteins. *E. coli uspG* mutants are sensitive to respiratory uncouplers but not DNA-damaging agents (Bochkareva *et al.*, 2002) suggesting the roles of the two classes of USP may be distinct.

1.12.2 USP Function in Other Bacteria

There are 9001 USP sequences described in the Pfam database. Of these, 7655 USP domains are found in a total of 1306 species of bacteria. The number of USPs in different bacterial species varies widely, with the largest number of *usp* genes found in the bacterial genome of the *Halobacteriaceae*, where osmoregulation is a continued challenge. In addition, the organism *Natrialba magadii* ATCC 43099 has 54 USPs consisting of 66 domains. Evidence for the biological role of USPs thus far is circumstantial and the general role the family of proteins play remains largely unresolved. Studying the role of USPs has highlighted some examples where the USPs are involved in bacterial responses to stress.

Liu *et al.*, 2006, found that UspA of *Salmonella typhimurium* was induced by nutritional, oxidative and temperature stress. In another case, UspA mutants of the periodontal pathogen *Porphyromonas gingivalis* indicated that UspA was involved in biofilm formation and showed that UspA mutants were less tolerant to certain conditions including oxidative and antibiotic stresses (Chen *et al.*, 2006). A significant amount of work has been undertaken in the last few years by Max Schobert on *Pseudomonas* USPs. Work on the opportunistic pathogen *Pseudomonas aeruginosa* has shown the USPs PA3309 and PA4352 to be significantly involved in the survival of the bacteria under anaerobic conditions (Boes *et al.*, 2006; Schreiber *et al.*, 2006). One of these proteins, PA3309, is discussed in a later chapter (Chapter 5).

1.13 The USPs of *Mtb*

By studying *Mycobacterium* species in stationary phase, members of the UspA superfamily have been identified which are expressed under conditions of hypoxia or nutrient depletion (Boon *et al.*, 2001, O'Toole *et al.*, 2003). The number of USPs in other *Mycobacterial* species varies; however the USPs are highly conserved between *Mtb* and *M. bovis*, with both containing ten USPs. The difference between the numbers of USP is exemplified when comparing *M. smegmatis* and *M. leprae*. *M. smegmatis* is a fast growing *Mycobacterium* species, contains 17 USPs, whereas the obligate intracellular pathogen, *M. leprae*, contains a single USP, ML1390. *M. leprae* has lost a vast amount of its genome (~2000 genes) compared to *Mtb*, through extensive genome decay (Cole *et al.*, 2001) and the single domain USP, Rv1636, remaining may be vital to its survival.

Mtb H37Rv encodes ten proteins classified within the USP family in the Pfam database (Pfam accession number PF00582, www.sanger.ac.uk). The domain organisation of the USP proteins falls

The Universal Stress Proteins of Bacteria

into 5 classes, shown in Figure 1.7A. The first class encompasses Rv1636 (146 aa) which consists of a single USP domain. The second class comprising six proteins with tandem USP domains includes Rv1996 (317aa), Rv2319c (292aa), Rv2005c (295aa), Rv2026c (294aa), Rv2623 (297aa) and Rv2624c (272aa). Rv3134c (268aa) and Rv2028c (279aa) make up the third class which contain a single USP domain fused to another C-terminal domain of unknown function. It is possible the second domains are actually distantly related USP tandem domains. The homology of each of the USPs is shown in Figure 1.7B and shows two main clusters with Rv1636 and Rv2319c the most distantly related.

Finally the KdpD protein (860aa) at locus *Rv1028c* has a USP domain fused in between other catalytic domains. The KdpD protein has a role in regulating high-affinity K⁺ transport as part of a two-component sensor-effector regulation system (Walderhaug *et al.*, 1992). Figure 1.7B shows the sequence alignment relatedness and groups the *Mtb* USPs (excluding KdpD protein) into four clusters.

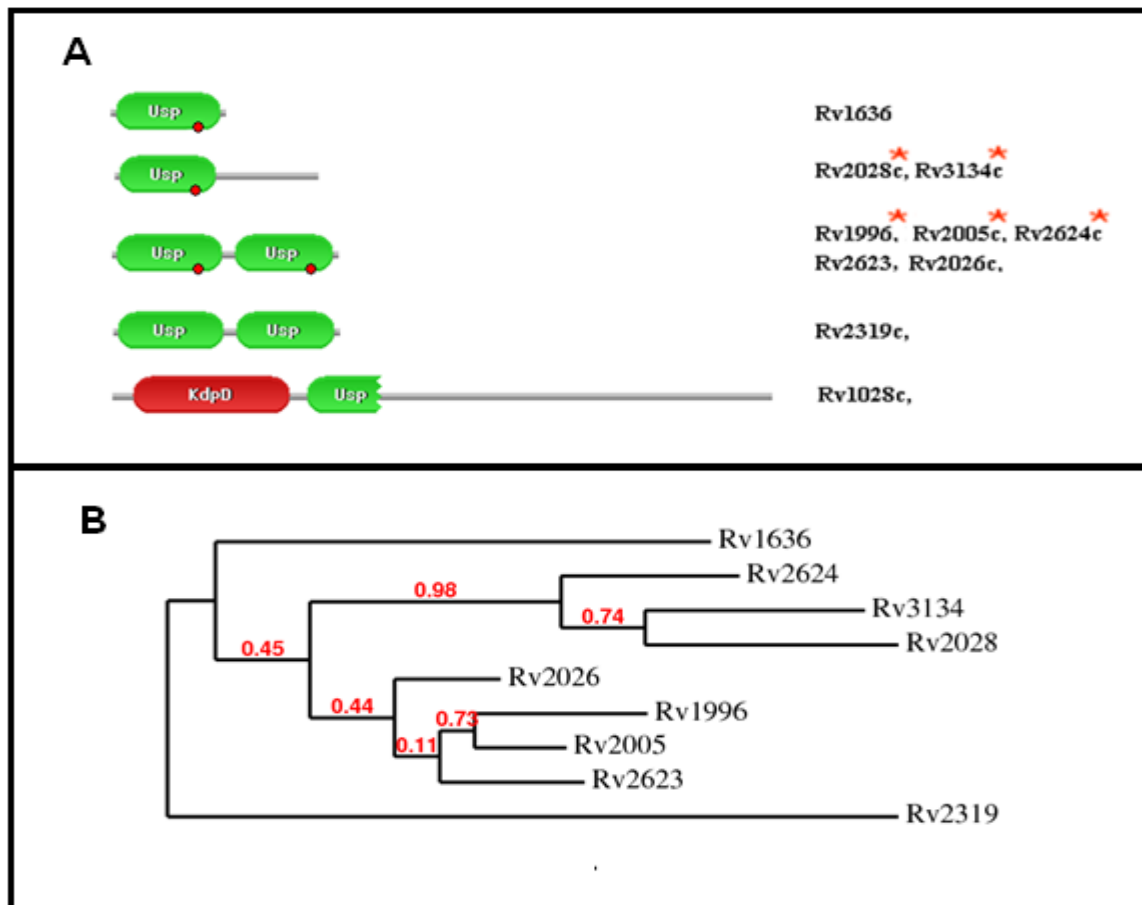


Figure 1.7. The USPs of *Mtb*. (A) Domain organisation of the USP's from *Mtb*. Red dot indicates the presence of the NTP-binding motif. Red star indicates induced under control of DosR regulon. (B) Phylogenetic tree of the USP proteins from *Mtb*. The red numbers are the branch support values (x100 to get percent) and the branch lengths indicate the degree of divergence.

The Universal Stress Proteins of Bacteria

Eight of the ten *Mtb* USP's contain at least one conserved motif G-X₂-G-X₉-G(S/T). The partial 103 bp USP domain in Rv1028c does not contain the conserved motif and Rv2319c contains the sequence G-X₂-P-X₉-GS twice within the domain where the middle conserved glycine is changed to a proline. The impact of this on function is not clear and remains to be determined. Of the tandem USP domains all contain two conserved motifs except the second USP domain in the Rv2624c protein, which contains the similar sequence G-X₁-G-X₉-GPS.

The number of USPs in other strains of *Mtb* can also be compared. Currently six phylogeographic lineages that make up the *Mtb* global population have been described (Gagneux *et al.*, 2006). The Haarlem genotype (taxID:395095), part of the Euro American phylogeographic group, which accounts for 25% of cases in Europe, central America and the Caribbean, has undergone at least four significant deletions (Cubillos-Ruiz *et al.*, 2010). The strain has 9 USPs and does not contain the tandem USP Rv2026c. The bacteria can therefore survive and cause disease without this USP, indicating there may be redundancy between these USPs.

The different genotypes of *Mtb* are known to exhibit different levels of pathogenicity and cause different immune responses in animal models (Lopez *et al.*, 2003) and in human disease. The differences between strains can give information about the importance of particular genes and through studying the strain variability of circulating lineages important considerations for drug development may arise.

1.14 Why Study the *Mtb* USPs?

Mtb is a highly successful pathogen, responsible for huge amount of morbidity around the world, and has developed many evasion and survival mechanisms to ensure survival within the human host. To discover the finer details of the survival of *Mtb* within macrophages remains a large challenge to the research community. The survival of the bacilli within the cells that are at the forefront of the immune defences is part of what allows *Mtb* to continue to be a global burden and understanding this is key to the development of new improved immunotherapeutic strategies.

Due to its global impact *Mtb* is a currently widely researched and relatively well funded with modern non-for-profit charities such as the Bill and Melinda Gates foundation recognising the global importance of TB research. Treatment of tuberculosis involving the current drug program is 40 years old and takes many months to administer, which is a huge issue regarding compliance and drug resistance. Policy makers and mandates are required to tackle the ongoing problems in the fight against tuberculosis, which remains one of the world's most deadly diseases.

The Universal Stress Proteins of Bacteria

The TB Alliance global report suggests in the next 10 years that TB alone will cost the world's poorest countries between 1 and 3 trillion dollars. TB is a devastating disease and is responsible for jeopardising healthcare systems and crippling economies and communities across the world. TB incidence is at an all time high (TB Alliance Annual report 2009) and to reverse the trend there are many scientific, regulatory and social hurdles to overcome.

With the challenges facing the control and containment of TB, research areas are required to be focused. One area of research key to the survival of the bacteria with the host over long periods of time is the ability of the bacteria to respond and endure the stresses that the bacteria encounter over the infection cycle. The *Mtb* proteins involved in surviving within the host macrophages and those which modify the cells' transcriptional activity all contribute to the persistence of the bacteria and to a successful infection/disease progression.

Finding out how the bacteria resist the toxicity of the macrophage can provide targets for drug development. The ability of *Mtb* to survive the numerous stresses previously described including ROIS, RNIs, pH and nutrient starvation, dictates that *Mtb* is an ideal candidate to study the USP superfamily, as these proteins have been linked to stress survival in a variety of organisms. Indeed, several USPs are upregulated in macrophages with Schnappinger *et al.*, 2003, showing USP Rv2623 is upregulated in human and mouse macrophage infections. The superfamily also are linked to the dormancy survival regulon (DosR) with six USPs having been shown to be part of the regulon. Dormancy of *Mtb* is currently a key research and the importance of the superfamily to *Mtb* survival is of great interest.

The key to understanding the role of USPs involves researching the biochemical characteristics, studying the structure and determining a role for the USPs within the cell. USPs are widely distributed; however a common function for the small proteins is yet to be elucidated. Any research on USPs can also be applied to other associated areas of biology and with the large amount of bioinformatic databases the potential for inter-linked research is high.

1.15 The USP Rv1636

Rv1636 is the smallest of the USPs in *Mtb* consisting of 146 aa, with a molecular mass of 15.312 kDa. Rv1636 is 100% identical to *M. bovis* protein Mb1662 and 89% identical to the sole USP, ML1390, of *M. leprae* TN (Figure 1.8). The retention of this protein in *M. leprae* suggests it may have an important function for intracellular survival in *M. leprae* and, by implication, *Mtb* also. Rv1636 contains the nucleotide binding sequence conserved motif G-X₂-G-X₉-G(S/T) and an alignment of

The Universal Stress Proteins of Bacteria

Rv1636 with the studied USPs MJ0577, UspA from *E. coli* and the UspA from *H. influenzae* is shown below (Figure 1.8). The protein is described as an iron regulated protein and has been shown, by Schnappinger *et al.* 2003, to be expressed in the presence of palmitic acid.

Rv1636 has been studied in Huw Williams' laboratory because it is the most similar to the well studied UspA of *E. coli* and because it has similarities to protein ML1390 of *M. leprae* as previously mentioned, and also as the protein has been shown to be soluble when over-expressed in *E. coli*, which is not the case with many of the other USPs (Williams, personal communication). It is also the only single domain USP protein in *Mtb* which may indicate it is a fundamentally valid place to start the analysis of the USPs of *Mycobacterium* species.

The Universal Stress Proteins of Bacteria

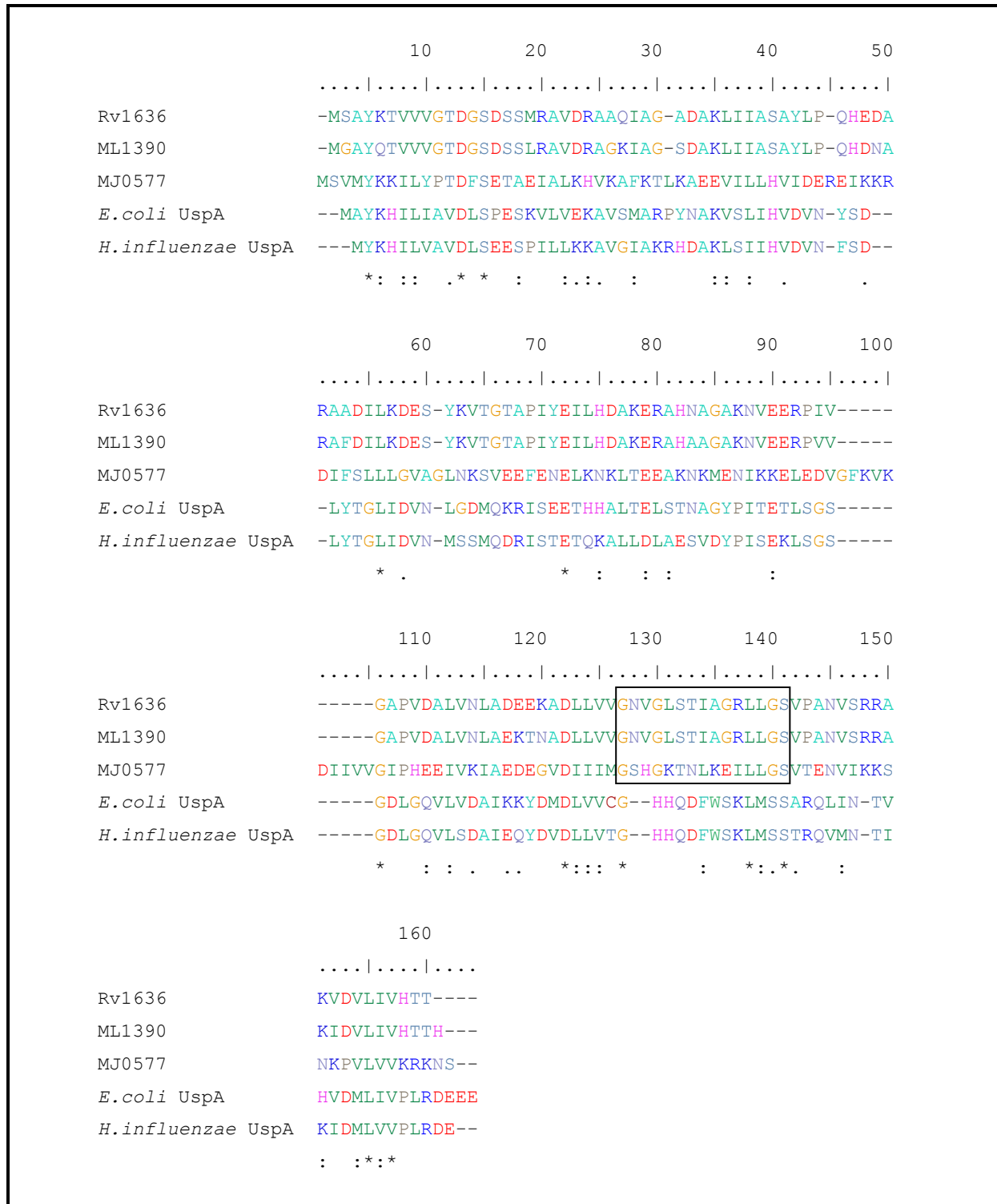


Figure 1.8. Protein sequence alignment of Rv1636 with ML1390 (*M. leprae*), MJ0577 (*Methanococcus jannaschii*), UspA (*E. coli*) and UspA (*H. influenzae*). The ATP binding motif first identified in MJ0577, is boxed. The alignment was constructed using BioEdit Sequence Alignment Editor.

1.16 Aims and Objectives

Universal Stress proteins (USPs) are widespread in Eukaryotes, Eubacteria and Archae, and yet the roles of the proteins are largely unknown. The research into USPs aims to determine the common role of the USP domains in biology, which have been described in over 9000 proteins. Studying the USPs of *Mtb* aims to determine the role USP play with particular reference to how the bacteria persist within granulomas in the lungs through enduring the stresses present. This project explores the role of USPs in mycobacterial metabolism, persistence, stress responses and pathogenesis, with particular emphasis on the USP Rv1636. Of the 48 genes in the DosR regulon, identified by Sherman *et al.*, 2001, most are conserved hypothetical proteins. Six of these are Universal Stress Proteins leading to the hypothesis that understanding the roles of USPs will provide an insight into the function of the DosR regulon and the mechanism of hypoxic adaptation.

Previous work in the Williams group has focused on the biochemistry of Rv1636 and initial experiments sought to confirm and extend previous observations. The protein had been shown to have a number of putative biochemical activities including autokinase activity, phosphatase activity and autonucleotidylation capabilities. Experiments are described with the aim of identifying key residues of the Rv1636 protein that are involved in these activities. Site-directed mutagenesis was employed to create a range of mutant proteins that were assayed for these activities in order to identify the function of the biochemical activities. A deposited crystal structure was analysed and further structural information was obtained through other experimental methods including multi-angle light scattering and analytical ultra centrifugation. In addition, the oligomeric state of the protein was proposed to provide clues to the role of the Rv1636.

The project also aimed to identify the biological role of Rv1636 and to investigate any proteins that specifically interact with the homologue of Rv1636 from *M. bovis*, Mb1662. Initially using a biochemical approach, work began with co-immunoprecipitation experiments to isolate any interacting proteins. Isolation of an *Rv1636* knockout would also aid characterisation and attempts to complete this line of work are described. Additional research has also focused on *Pseudomonas aeruginosa* USP PA3309, a USP for which a mutant and phenotype have been published. The *P. aeruginosa* work provided an opportunity to determine if the nucleotide binding motif was active *in vivo*.

The Universal Stress Proteins of Bacteria

The key questions to be addressed in order to characterise the functionality of USPs are:

- 1) Are the previously discovered putative autokinase, autonucleotidylation and RNA triphosphatase activities true biochemical activities of Rv1636? (Chapter 3)
- 2) Can the level of NTPase activity of Rv1636 preparations be increased? (Chapter 4)
- 3) What is the biological role of the USPs in *Mtb*? (Chapter 5)
- 4) Does the nucleotide binding motif of the *P. aeruginosa* PA3309 protein contribute to the anaerobic survival phenotype? (Chapter 5)

Chapter 2 - Materials and Methods

2.1 General

Radiochemical work was undertaken following general safety procedures and the specific local rules set out in the laboratory. Chemicals were purchased from Sigma unless otherwise stated.

2.1.1 Media

Table 2.1. Composition of media

Name	Composition per litre
Luria Bertani (LB) broth	10 g tryptone (Oxoid), 5 g yeast extract (Oxoid), 5 g NaCl.
LB agar	LB broth plus 15 g Agarose (Oxoid).
7H9 broth	4.7 g Middlebrook 7H9 broth powder (BD), 2 ml glycerol, 0.05% (v/v) Tween 80, 100 ml sterile Middlebrook OADC supplement (BD) added before use.
7H10 media	19.47 g Middlebrook 7H10 agar base (BD), 5 ml glycerol, 100 ml sterile Middlebrook OADC supplement (BD) added before use.
<i>Pseudomonas</i> Isolation Agar	45.03 g Fluka <i>Pseudomonas</i> Isolation Agar, 20 ml glycerol.
SOC media	20 g tryptone (Oxoid), 5 g yeast extract (Oxoid), 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl ₂ , 10 mM MgSO ₄ , 20 mM glucose.

LB media was autoclaved at 120 °C for 15 minutes. 7H9 and 7H10 were autoclaved in a pressure cooker for 7 minutes. Media was made with dH₂O. BD is Becton, Dickinson and Company.

The Universal Stress Proteins of Bacteria

2.1.2 Buffers and Solutions

Table 2.2. Buffers and solutions

Buffer	Composition per Litre
Phosphate-buffered saline (PBS)	200 mM NaCl, 1.8 mM KH_2PO_4 , 10 mM Na_2HPO_4 , 2.7 mM KCl (pH 7.3)
Tris-Acetate-EDTA (TAE)	40 mM Tris, 40 mM acetate, 1 mM EDTA (pH 8.0)
Tris-Glycine Buffer	25 mM Tris (pH 8.3), 250 mM glycine, 0.1% (w/v) SDS
Tris-Borate-EDTA (TBE)	89 mM Tris, 89 mM Boric acid, 2 mM EDTA (pH 8.0)
Transfer Buffer	1 M Tris-HCl (pH 8.3), 192 mM glycine, 20% methanol (v/v)
Thin layer Chromatography buffer	0.7 M Boric acid 0.4 M K_2HPO_4
Thin layer electrophoresis buffer	5% (v/v) acetic acid, 0.5% (v/v) pyridine, pH 3.5

Buffers were sterilised with 32mm Syringe Filter with 0.45 μm or 0.22 μm Supor® membranes (PALL life Sciences).

The Universal Stress Proteins of Bacteria

2.1.3 Antibiotics

Table 2.3. Antibiotics and solvents

Antibiotic (supplier)	Stock concentration (mg · ml ⁻¹) [Solvent]	Working concentration <i>E. coli</i> (µg · ml ⁻¹)	Working concentration <i>Mycobacterium</i> (µg · ml ⁻¹)	Working concentration <i>Pseudomonas</i> (µg · ml ⁻¹)
Ampicillin (Sigma)	50	50	N/A	N/A
Hygromycin (Roche)	50	150	50	N/A
Chloramphenicol (Sigma)	25	25	N/A	N/A
Kanamycin (Sigma)	40	40	40	N/A
Tetracycline (Sigma)	10	15	N/A	150
Carbenicillin (Sigma)	100	N/A	N/A	500

Antibiotics were dissolved in water before filter sterilisation using a 0.2µm filter. Chloramphenicol was dissolved in 100% ethanol.

The Universal Stress Proteins of Bacteria

2.1.4 Strains

Table 2.4. Strains

Strain	Description or genotype	Source
<i>M. tuberculosis</i> H37Rv	Wild type	-
<i>M. bovis</i> BCG Pasteur	Wild type	-
<i>M. smegmatis</i> mc ² 155	Wild type	-
<i>E. coli</i> XL-1 blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac</i> [F' <i>proAB lacI^f</i> Δ <i>M15</i> Tn10 (Tet ^r)] Subcloning competent cells	Stratagene™
<i>E. coli</i> Rosetta (DE3)	F <i>ompT hsdSB</i> (rB-mB-) <i>gal dcm</i> (DE3) pRARE (Cam ^r) BL21 derivative with additional copies of rare tRNA genes present on the pRARE plasmid. Chloramphenicol resistance.	Novagen™
<i>E. coli</i> S17.1	Donor strain derived from <i>E. coli</i> 294. <i>recA pro hsdR RP4-2-Tc::Mu-Km::Tn7</i>	Simon <i>et al.</i> , 1983
<i>E. coli</i> DH5α	<i>fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80</i> Δ(<i>lacZ</i>) <i>M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17</i>	Invitrogen™
<i>P. aeruginosa</i> PA01	Wild type	Schreiber <i>et al.</i> , 2006
<i>P. aeruginosa</i> KS17	PA01 ΔPA3309, Wild type with PA3309 knockout.	Schreiber <i>et al.</i> , 2006
<i>P. aeruginosa</i> KS28	KS17 attB::mini-CTX2	Schreiber <i>et al.</i> , 2006
<i>P. aeruginosa</i> KS29	KS17 attB::PA3309	Schreiber <i>et al.</i> , 2006
<i>P. aeruginosa</i> KS27	PA01 attB::mini-CTX2	Schreiber <i>et al.</i> , 2006

The Universal Stress Proteins of Bacteria

2.1.5 Plasmids

Table 2.5. Plasmids and origin

A. General

Plasmid	Description	Source
pGEM-Te	Direct PCR cloning vector, ampicillin resistance	Promega
pET15b	His-tag protein expression vector, ampicillin resistance	Novagen
pET41a	Protein Expression vector, kanamycin resistance	Novagen

The Universal Stress Proteins of Bacteria

B. Mycobacterial Vectors

Plasmid	Description	Source
pET15b- <i>Rv1636</i> .	pET-15b containing the <i>Rv1636</i> gene	Leiva, 2007
pGEM-Te- <i>Rv1636</i>	pGEM-Te containing the <i>Rv1636</i> gene	Leiva, 2007
GST- <i>Rv1636</i> vector	Vector containing <i>Rv1636</i> gene with GST tag	Leiva, 2007
pET41a- <i>Rv1636</i> _{MYKT}	pET41a containing <i>Rv1636</i> _{MYKT}	This study
pET41a- <i>Rv1636</i> _{MFKT}	pET41a containing <i>Rv1636</i> _{MFKY}	This study
pET41a- <i>Rv1636</i> _{MAKT}	pET41a containing <i>Rv1636</i> _{MAKT}	This study
pMV261	Mycobacterial non-integrating vector. Expression under control of HSP60 promoter.	Stover <i>et al.</i> , 1991
pMV361	Mycobacterial integrating vector, kanamycin resistance. Expression under control of HSP60 promoter	Stover <i>et al.</i> , 1991
pMV261:: <i>Rv1636</i>	pMV261 with <i>Rv1636</i> under the control of Hsp60 promoter.	This study
pMV361:: <i>Rv1636</i>	pMV361 with <i>Rv1636</i> under the control of Hsp60 promoter.	This study
pMV306:: <i>Rv1636</i>	pMV361 with <i>Rv1636</i> under the control of Hsp60 promoter.	Leiva, 2007
pSE100	Expression vector	Guo <i>et al.</i> , 2007
pSE100:: <i>Rv1636</i>	Expression vector containing <i>Rv1636</i> and ribosome binding site.	This study
pMC1s	Tetracycline repressor. tetR is expressed from a strong mycobacterial promoter	Guo <i>et al.</i> , 2007
pMC2m	Tetracycline repressor. tetR is expressed from a mycobacterial promoter of intermediate strength	Guo <i>et al.</i> , 2007
PSMT100:: <i>Rv1636</i>	Mtb suicide vector containing <i>Rv1636</i> gene region. PSMT100 vector from Dussurget <i>et al.</i> , 2001.	Leiva, 2007

The Universal Stress Proteins of Bacteria

C. *Pseudomonas* Vectors

Plasmid	Description	Source
Mini-CTX2	<i>Pseudomonas</i> integrating vector.	Schreiber <i>et al.</i> , 2006
Mini-CTX2:: <i>PA3309</i>	Mini-CTX2 vector containing the <i>PA3309</i> gene and upstream (255bp) and downstream (456bp) regions	Schreiber <i>et al.</i> , 2006
Mini-CTX2:: <i>PA3309</i> _{G118A}	Mini-CTX2:: <i>PA3309</i> with glycine at position 118 mutated to alanine.	This study
pFLP2	Flip recombinase (FLP) encoded on plasmid. Ampicillin resistant	Hoang <i>et al.</i> , 1998
pUCp18	<i>Pseudomonas</i> non-integrating vector. Expression under control of proteins natural promoter. Carbenicillin resistance.	Schweizer 1991
pUCp18:: <i>PA3309</i>	pUCp18 with <i>PA3309</i> gene with 456bp upstream and 255bp downstream.	This study

The Universal Stress Proteins of Bacteria

2.1.6 Oligonucleotides

Table 2.6. Primers used in site directed mutagenesis of Rv1636 (SDM) (Chapter 3). Base changes are highlighted in bold.

Primer	Primer Sequence 5'→3'
D12A sense	GTG GTA GGA ACC GCC GGT TCG GAC TCG
D12A anti-sense	CGA GTC CGA ACC GGC GCT TCC TAC CAC
G113A sense	CTG CTG GTC GTC GCC AAT GTC GGT CTG AG
G113A anti-sense	CT CAG ACC GAC ATT GGC GAC GAC CAG CAG
G116A sense	CGT CGG CAA TGT CGC TCT GAG CAC GAT CG
G116A anti-sense	CG ATC GTG CTC AGA GCG ACA TTG CCG ACG
G126A sense	GT CGG CTG CTC GCA TCG GTA CCG GC
G126A anti-sense	GC CGG TAC CGA TGC GAG CAG CCG AC
S127A sense	GT CGG CTG CTC GGA GCG GTA CCG GC
S127A anti-sense	GC CGG TAC CGC TCC GAG CAG CCG AC
S16A-S17A sense	GAC GGT TCG GAC GCG GCG ATG CGA GCG G
S16A-S17A antisense	C CGC TCG CAT CGC CGC GTC CGA ACC GTC
S118A sense	GC AAT GTC GGT CTG GCC ACG ATC GCG GGT C
S118A anti sense	G ACC CGC GAT CGT GGC CAG ACC GAC ATT GC
D139A sense	CGG GCC AAG GTC GCC GTG CTG ATC GTG
D139A antisense	CAC GAT CAG CAC GGC GAC CTT GGC CCG

The Universal Stress Proteins of Bacteria

Table 2.9. Primers used to clone Rv1636 into the over-expression vectors pMV261, pMV361 and pSE100 (Chapter 5). Restriction sites are underlined. RBS = ribosome binding site.

Primer	Primer Sequence 5'→3'
pMV261-Rv1636-For	ATAT <u>GGATCC</u> AATGAGCGCCTATAAGACCGTG BamHI Rv1366
pMV261-Rv1636-Rev	ATATA <u>AAGCTT</u> CTAGGTGGTGTGCACGATCAG HindIII Stop Rv1636
pMV361-Rv1636-For	ATATGAATTCATGAGCGCCTATAAGACCGTG EcoRI Rv1366
pMV361-Rv1636-Rev	ATATA <u>AAGCTT</u> CTAGGTGGTGTGCACGATCAG HindIII Stop Rv1636
pSE100-Rv1636-For	ATAT <u>GGATCC</u> GGGAGGGTAAATGAGCGCCTATAAGACC BamHI RBS Rv1636
pSE100-Rv1636-Rev	ATATA <u>AAGCTT</u> CTAGGTGGTGTGCACGATCAG HindIII Stop Rv1636
pMV261-Rv1636 _{MAKT} -For	ATATATGGCCAAGACCGTGGTGGTAGGAACCG MscI Rv1636 _{MAKT}
pMV261-Rv1636 _{MAKT} -Rev	ATATGGATCCCTAGTGATGGTGGTGGTGGTGGTGCACGATCAGCAGC BamHI Stop His tag Rv1636 _{MAKT}
pMV361-sequencing	AAGCGGAAGAGCGCCAATAC
pMV261-sequencing	GGCATAGGCGAGTGCTAAG

Table 2.10. Primers to amplify *M. smegmatis* attB site primers amplify 600 bp (Chapter 5)

Primer	Primer Sequence 5'→3'
Ms-attB-Forward	GAAGTGCACCAAGTGGCGAGAAC
Ms-attB-Reverse	GATCAGCTCTTCCACCGACTC

The Universal Stress Proteins of Bacteria

Table 2.11. Primers used for confirmation of miniCTX2::PA3309 integration into *P. aeruginosa* Chapter 5). attB is attachment site in *P. aeruginosa*. MCS is multiple cloning region of miniCTX2. Tet is tetracycline cassette.

Primer	Primer Sequence 5'->3'
Pa-attB Forward	CGAGTGGTTTAAGGCAACGGTCTTGA
Pa-attB Reverse	AGTTCGGCCTGGTGAACAACCTCG
Tet Forward	CGGGGTCTGACGCTCAGTGG
Tet Reverse	GGGCGCGGGGCATGACTATC
Pa3309 Forward	GGTAGCCGTGGACGGCAGC
Pa3309 Reverse	CAGCAGGACCGGACGCGAAG
MCS Forward	CGGGGCCGTCCTTGCTGAAT
MCS Reverse	AGAGCGCTTTTGAAGCTGATGTGC

Table 2.12. Primers used to clone PA3309 gene region (Chapter 5). Restriction sites are underlined.

Primer	Primer Sequence 5'->3'
Pa-3309-Up-Forward	ATATATCGGAATTCGCCATGGACGAGGAACTG
Pa-3309-Down-Reverse	ATATATCCAAGCTTATCCACGTGCCGATGGTC

2.1.7 Strain Storage and Culture

2.1.7.1 Glycerol Stocks

Individual colonies were picked from agar plates and grown in suitable medium. Glycerol was added to a concentration of 30%. Strains were stored in 2 ml Nunc cryovials at -80 °C. Plasmids were stored in H₂O at -20 °C or within *E. coli* at -20 °C.

2.1.7.2 Strain Culture

Stock cultures were streaked on agar plates containing appropriate selective antibiotics and incubated at 37 °C overnight. Plates were stored at 4 °C for up to 6 weeks. Single colonies were inoculated into liquid medium and 5 ml starter cultures, in 15 ml Falcon tubes, were grown with 220 r.p.m. shaking at 37 °C. These starter cultures were used to inoculate 50 ml cultures in 250 ml Erlenmeyer conical flasks. Flasks were incubated at 37 °C whilst shaking at 220 rpm. Optical density measurements of bacteria cultures were performed by measuring the OD at 600 nm in a spectrophotometer (WPA Lightwave II). OD readings were obtained below 0.6 and, to ensure the linearity of the reading, samples measured above this level were diluted.

2.1.7.2.1 *Escherichia coli* Culture

E. coli was cultured in LB broth and on LB agar plates. Plates were incubated overnight. Overnight broth cultures (5 ml) were grown for protein expression at 37°C with shaking at 220 rpm and used to inoculate one litre cultures.

2.1.7.2.2 *Mycobacterium* Culture

M. smegmatis was grown 3 days on LB plates and 3 days on LB broth. *Mtb* was grown 4 weeks on 7H10 plates, plus 0.2% glycerol and Middlebrook Oleic Albumin Dextrose Catalase (OADC) growth supplement, or two weeks in liquid culture 7H9, plus 0.5% glycerol and OADC at 37 °C with shaking at 220 rpm for liquid cultures. *M. bovis* BCG was grown in liquid culture for two weeks as with *Mtb*. The mild detergent Tween 80 was used in liquid cultures at 0.05% to ensure no clumping of the bacterial cells occurred. Manipulations of *Mtb* were undertaken in a category three

The Universal Stress Proteins of Bacteria

laboratory. Mycobacteria removed from the laboratory were heat killed by boiling for 10 minutes. CFU determination involved making serial dilutions in 96-well plates. Ten fold dilutions were prepared in PBS and 20 μ l spots were plated out in triplicate. Colonies were then counted by eye and the CFU determined. Hypoxic cultures were achieved using suba seals as bungs. A gradual depletion of oxygen tension within the 250 ml conical flasks created the desired hypoxic environment.

2.1.7.2.3 *Pseudomonas* Culture

P. aeruginosa was grown on *Pseudomonas* isolation agar (Fluka). Plates were left overnight at 37 °C while liquid cultures of LB were incubated overnight at the same temperature with shaking at 220 rpm. For anaerobic survival experiments, as described by Schreiber *et al*, 2006, *P. aeruginosa* was grown shaking at 37 °C to an OD of 0.3 in 50 ml potassium phosphate-buffered (100 mM) LB medium at pH 7.4. A sample of 15 ml was then put into test tube and then sealed with air-tight suba-seal bungs. Sodium pyruvate was added to a concentration of 40 mM and the cultures were incubated 37 °C without shaking. Before taking samples, the tubes were inverted a minimum of ten times to achieve a homogenous suspension of bacteria. Samples of 100 μ l were taken using a sterile needle through the rubber bung thus ensuring no oxygen was introduced to the anaerobic environment. Dilution series for CFU counts were done in 24 well plates containing 450 μ l of PBS, with 50 μ l transferred between wells and 40 μ l were plated out for CFU counting. Strains were grown in triplicate and CFU plating also done in triplicate.

2.2 Protein Methods

2.2.1 Protein Purification

The Rv1636_{WT} proteins were purified by nickel affinity chromatography using a Fast Protein Liquid Chromatography (FPLC) machine. Previous work in the laboratory involved cloning the *Rv1636* gene into a pET-15b vector and then introducing the vector into *E. Coli* Rosetta cells (Leiva, 2007). The *E. coli* cells were grown in LB broth to an OD₆₀₀ 0.6 and induced to express the 6x Histidine tagged fusion protein by addition of 0.1 mM isopropyl- β -D-thiogalactopyranoside (IPTG) for 2 hours. The over-expressing cells were pelleted by centrifugation and resuspended 5 ml per 1 g cells in a lysis solution containing Bugbuster Reagent (Novagen), Benzonase nuclease (Novagen) (100

The Universal Stress Proteins of Bacteria

U), Lysozyme (Sigma) (5 KU) and Complete Mini EDTA-free protease inhibitors (Roche). Benzonase nuclease is an endonuclease from *Serratia marcescens* that attacks and degrades all forms of DNA/RNA. After incubating the cell solution for 15 minutes at room temperature, on a rocking platform, the solution was centrifuged at 16000 *g* for 20 minutes at 4 °C, to pellet any cellular debris.

The supernatant containing the soluble fraction, typically 5 ml, was run through a 1 ml HisTrap[™] HP column using the automated liquid affinity chromatography system ÄKTA FPLC (Amersham Biosciences). To elute the His-tagged protein, an imidazole gradient was created by combining a wash buffer (20 mM sodium phosphate, 0.5 M NaCl and 10 mM imidazole, pH 7.4) and an elution buffer (20 mM sodium phosphate, 0.5 M NaCl and 0.5 M imidazole, pH 7.4). Samples were mixed with equal volume of 2X Laemmli buffer (Sigma) and heated at 95 °C for 5 minutes for analysis by SDS-PAGE. Electrophoresed proteins were stained with Coomassie stain or Imperial[™] Protein Stain (Pierce).

Fractions containing the target protein were pooled and subjected to buffer exchange, using Centriprep columns (3,000 mwco) (Millipore), into a protein storage buffer (20 mM Tris-HCl, pH 7.5, 100 mM NaCl and 10% glycerol), before being stored at -80 °C until use. Protein concentration was measured by a Bradford assay using the Better Bradford Assay[™] Reagent (Pierce) according to the manufacturer's instructions. Alternatively UV adsorption at 280 nm was performed.

2.2.2 Size Exclusion Chromatography (SEC)

Further purification of the proteins was performed using the HiLoad 26/60 Superdex 75 prep grade column (GE Healthcare/Amersham), which is ideal for separating proteins Mw 300-70000 Da. Size exclusion chromatography is also known as gel filtration chromatography. Samples were filtered through 0.22 µm filter before up to 5 ml was loaded on to the column. The buffer used for the chromatography was 50 mM Tris-HCl pH 7.4, 200 mM NaCl. A buffer of at least 150 mM NaCl avoids any unwanted non-specific interactions with the matrix. Samples were run overnight at flow rate of 0.3 ml/minute and eluted in 3 ml fractions. Prior to use with the FPLC, buffers were filtered through 0.45 µm filters and were degassed by being placed under a vacuum for 10 minutes.

2.2.3 SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

For SDS-PAGE gel electrophoresis, the solutions described in table 2.13 were mixed thoroughly and TEMED (N,N,N',N' – Tetramethylethylenediamine) was added last. Acrylamide monomer is a potent neurotoxin and to avoid potential danger a pre-mixed solution of acrylamide/bis acrylamide was used. Gel solutions were poured between assembled gel plates (8 cm x 10 cm x 1.5 mm) and allowed to set. The separating gel poured first and 50 µl n-butanol was used as an overlay before the stacking gel was poured on top. Protein samples were mixed with equal volume 2X Laemmli loading buffer (Sigma) before being heated 95 °C for 3-5 minutes.

The gels were run for 50 minutes at 200 V in a 1x Tris/Glycine/SDS buffer using a Biorad mini tank. Electrophoresed proteins were stained with Coomassie blue dye (0.5% Coomassie blue, 50% Methanol, 7.5% Acetic acid, 40% H₂O) and de-stained with a solution of 45% Methanol, 45% H₂O and 10% Acetic acid. The SeeBlue 2 Marker (Invitrogen) was used as a protein marker.

Table 2.13. Composition of SDS-PAGE gels.

Solutions	Separating Gel (12% gel)	Stacking Gel (4% gel)
30% Polyacrylamide	4 ml	0.67 ml
Distilled H ₂ O	3.4 ml	3.075 ml
1.5 M Tris-HCl, pH 8.8, 0.4% (w/v) SDS	2.5 ml	-
0.5 M Tris-HCl, pH 6.8, 0.4% (w/v) SDS	-	1.25 ml
Ammonium Persulfate 10% (w/v)	50 µl	50 µl
TEMED (N,N,N',N'-Tetramethylethylenediamine)	5 µl	5 µl

2.2.4 Western Blotting

Following the separation of proteins by SDS-PAGE, gels were then soaked in transfer buffer (1 M Tris-HCl, Glycine 192 mM, 20% methanol (v/v) adjusted to pH 8.3) for 10 minutes. The Hybond membrane (Amersham) was also soaked in transfer buffer as preparation for the transfer. The transfer of the proteins to the membrane was performed in a Biorad tank by electrophoresis for 2 hours at 30 V.

The Universal Stress Proteins of Bacteria

The membrane was then washed in PBS-T (PBS plus Tween 20 (0.1%)) before blocking and incubation with the primary antibody. Anti-His₍₆₎Rv1636 polyclonal antibodies, derived from rabbit antiserum, were previously produced by BioGenes GmbH (Freiburg, Germany). 0.5 µg primary antibody was added to 10 ml PSB-T milk solution (PBS-T plus 5% (w/v) skimmed milk powder) and the membrane was incubated overnight at 4 °C. The membrane was then washed in PBS-T and incubated with the secondary antibody to a dilution of 1:10000 (Anti-Rabbit IgG HRP Conjugate (Promega)) for 1 hour at RT. The Western blot was visualised using the Amersham ECL Western Blot Analysis System.

2.2.5 *In silico* Analysis of Proteins

Protein sequence translation was carried out using ExPaSy Translate Tool (<http://www.expasy.ch/tools/dna.html>). Alignment of protein sequences were performed using BioEdit Sequence Alignment Editor v7.0.9.0 with integrated ClustalW function (Hall, 1998). The Protein Homology/Analogy Recognition Engine (Phyre) server (<http://www.sbg.bio.ic.ac.uk/phyre>) was used to predict the tertiary structure of proteins (Kelley *et al.*, 2000) and visualised using Swiss-Pdb Viewer v3.7 (<http://www.expasy.org/spdbv>).

2.3 Nucleic Acid Methods

2.3.1 Transformation of Bacteria

2.3.1.1 *Escherichia coli*

2.3.1.1.1 Preparation of Competent Bacterial Cells

Starter cultures of *E. coli* cells (5 ml) were inoculated into 200 ml fresh LB (plus appropriate antibiotics) and incubated at 37°C with shaking at 220 rpm. When OD₆₀₀ 0.4–0.5 was reached, the culture was pelleted by centrifugation using Eppendorf® 5810R in pre-frozen 50 ml falcon tubes (7 minutes, 4 °C, 4000 g). The cells were then washed three times through re-suspension in 15 ml cold 50 mM calcium chloride with 15% (w/v) glycerol (Fluka). Finally, the pellet was re-suspended in 5 ml and the suspension transferred to Eppendorf tube in aliquots of 100 µl and stored at -80 °C till use.

The Universal Stress Proteins of Bacteria

2.3.1.1.2 Transformation

Plasmids were transformed into XL1 Blue *E. coli* cells or Rosetta *E. coli* cells. One hundred ng of plasmid DNA was mixed with 50 µl competent cells and left on ice for 20 minutes. The mix was then heat-shocked at 42 °C for 45 seconds followed by 2 minutes on ice, then 100 µl of LB was added to the transformed cells and the culture was incubated shaking at 37 °C for 1 hour. The cells were spread using a plate spreader on to the selective plates containing ampicillin (50 µg/ml) and were incubated overnight at 37 °C to allow colonies to develop.

2.3.1.2 Mycobacteria

2.3.1.2.1 Preparation of Competent Bacterial Cells

An 80 ml culture of *Mtb* or *M. smegmatis* was grown to late log phase in 7H9 OADC media plus 0.5% glycerol and 0.05% Tween 80. Cells were pelleted by centrifugation at 4000 *g* before being washed by resuspending in 10% glycerol. The wash step was repeated twice further before cells were resuspended in 2 ml of 10% glycerol and 200 µl aliquots were prepared ready for use. *M. smegmatis* competent cells were prepared on ice using cold 10% glycerol; *Mtb* cells were prepared at room temperature.

2.3.1.2.2 Electroporation

One µg DNA was mixed with a 200 µl aliquot of *M. smegmatis* or *Mtb* cells and incubated for 20 minutes (*Mtb* at room temperatures, *M. smegmatis* on ice). The suspensions were then electroporated, (in pre-cooled cuvettes for *M. smegmatis*), using the Biorad gene Pulser (Resistance 1000 Ω, Q= 25 µF, voltage 2.5 kV). Cells were resuspended in 1 ml of 7H9 media and allowed to recover at 37 °C for 3 hours shaking. Electroporated cells were then plated out (either 20 µl or 200 µl) on to selective plates and grown for 2 days (*M. smegmatis*) or 3-4 weeks (*Mtb*) at 37 °C before picking and re-streaking to isolate single colonies.

The Universal Stress Proteins of Bacteria

2.3.1.3 *Pseudomonas*

2.3.1.3.1 Heat Shock Transformation

Chemically-competent *P. aeruginosa* was produced according to Itoh *et al.* (1991) using CaCl_2 transformation buffer. Conical flasks with 50 ml LB broth were inoculated and grown at 37 °C on an orbital shaker at 220 rpm. When an OD_{600} reading of 0.4-0.5 was reached, the cells were pelleted by centrifugation at 4000 *g* for 10 minutes at 4 °C. The pellets were resuspended in 25 ml of ice cold transformation buffer and incubated on ice for 20 minutes. Cells were pelleted again and resuspended in 2.5 ml of ice cold transformation buffer and 0.2 ml aliquots were made in 1.5 ml Eppendorf tubes that were immediately used for transformation.

Chemically competent *P. aeruginosa* cells were mixed with DNA and incubated on ice for 20 minutes. The cells were heat-shocked in a 42 °C water bath for 45 seconds, then placed immediately back on ice for a further 2 minutes to allow recovery. Following the addition of 250 µl of fresh SOC media the samples were incubated for 1 hour at 37 °C on an orbital shaker at 220 rpm. Transformed cells were then plated onto selective agar plates and incubated overnight at 37 °C.

2.3.1.3.2 Biparental Mating

Biparental mating, using *E. coli* S17.1 as a donor strain, was used to introduce DNA into *P. aeruginosa* strains. The *E. coli* S17.1 was initially transformed with the plasmid DNA as described above. Both donor strain and *P. aeruginosa* were grown in 15 ml of LB in 250 ml conical flask until an OD_{600} reading of 0.4-0.5 was recorded. Antibiotics were used to select the plasmid in the *E. coli* S17.1. The cells were then pelleted in 50 ml Falcon tubes by centrifugation at 3000 *g* for 10 minutes at 4 °C in a Sigma 6K10 centrifuge. The pellets were then washed to remove trace amounts of the antibiotic by being resuspended in LB and re-pelleted.

The pelleted cells were resuspended in 400 µl of LB broth and mixed together thoroughly before being plated out in 100 µl spots onto LBA plates containing no antibiotics. As controls 100 µl of donor strain only and recipient strain only were plated out. Following incubation overnight at 37 °C the cultures on the mating plates and control plates were scraped from the plates using sterile disposable spreaders and re-suspended in 2 ml PBS. *P. aeruginosa* transformants were then selected by plating out 100 µl of the suspension onto PIA plates with appropriate antibiotics. After incubation overnight at 37 °C, potential transformants were picked off and further single colony purified on PIA selective plates before analysis.

2.3.2 DNA Plasmid Purification

Cells from a 5 ml overnight culture were harvested by centrifugation at 4000 *g* for 3 minutes. Plasmid DNA was purified using the QIAprep Spin Miniprep Kit (Qiagen) according to the manufacturers' specifications.

2.3.3 Agarose Gel Electrophoresis

Agarose gels (1.5% (w/v)) were prepared with 1 X TAE buffer. SYBR®Safe DNA gel stain (Invitrogen) (1:10,000 concentration) was added to the gel. Samples were mixed with 6x Gel Loading Solution (Promega) and loaded to the gel alongside a 1 kb DNA Ladder (Promega). Gels were run at 95 V in 1 X TAE buffer for 60 minutes and visualised on Safe Imager™ transilluminator (Invitrogen) using Gel-Doc™ 2000 (Bio-Rad) and Quantity One analysis software v4.5.2 (Bio-Rad) for image acquisition.

2.3.4 Sequencing DNA

Sequencing of plasmid DNA was carried out by Cogenics. Plasmid DNA at a concentration of 50 ng/μl was sequenced using universal primers T7, SP6, or T7 term or custom designed primers. BioEdit (www.mbio.ncsu.edu/bioedit) was used for alignments and viewing DNA sequences.

2.3.5 DNA Concentration

DNA concentration was measured using absorbance of UV light at 260 nm. DNA samples of 5 μl were diluted with 65 μl dH₂O. Concentration was determined by UV₂₆₀ measurement using an Eppendorf® Biophotometer with UV cuvettes (50 μg/ml plasmid per absorbance unit). The final concentration was determined by multiplying the absorbance by 50, followed by 14 to take into account the dilution factor.

2.3.6 Polymerase Chain Reaction

PCR reactions were carried out in a thermal cycler (Peltier DyadTM DNA Engine). The basic reaction involved using Taq polymerase and 1 µg of DNA as a template. Each reaction contained buffer (10X), dNTPs (2 mM), MgCl₂ (5 mM), forward primer (10 µM), reverse primer (10 µM), DNA template (1 µg), Taq DNA polymerase (1 U) and dH₂O to a final volume of 25 µl. The conditions for the reaction were 5 minutes at 95 °C, 30 cycles at 1 minute at 95 °C, 1 minute per kb at 55 °C and 1 minute at 72 °C. After 30 cycles the final step consisted of 7 minutes at 72 °C. Polymerases used included GoTaq (Promega), PCR Long Mix (Bioline), Pfu (Stratagene) and Taq (Promega). Where required, dimethyl sulfoxide (DMSO) was used at a concentration at 5%.

2.3.7 PCR Purification

PCR products were purified using a Qiagen MinElute PCR purification kit following the manufacturer's instructions. DNA was eluted in 30µl dH₂O.

2.3.8 Gel Extraction

Reactions were separated by agarose electrophoresis. PCR products and products of a restriction digest were gel extracted. The bands of interest were excised using a scalpel and the band processed according to the manufacturers' guidelines, using a Qiagen MiniElute Gel extraction kit. DNA was eluted in 30 µl dH₂O

2.3.9 DNA Restriction Digestion

Restriction digests were performed using Fast Digest (Fermentas) enzymes and a universal 10X buffer. Reactions were carried out for 30 minutes at 37 °C in 20 µl reaction volumes. Where Fast Digest enzymes were not available, Promega/NEB enzymes were used according to the manufacturer's instructions. Reactions were then visualised by agarose electrophoresis.

2.3.10 Genomic DNA extraction for PCR

2.3.10.1 *Pseudomonas*

Colony PCR was performed by picking a colony from *Pseudomonas* isolation agar and adding the cells to the PCR directly. Hot temperatures to denature the DNA result in the cells breaking open and DNA is then available as a template.

2.3.10.2 *Mycobacteria*

2.3.10.2.1 *Mtb*

For use as a template in PCR reactions, genomic DNA was extracted from *Mtb* using a CTAB (cetyltrimethylammonium bromide) method. Cultures of 50 ml of *Mtb* were pelleted by centrifugation and heat killed for 10 minutes at 100 °C. The pellet was then resuspended in TE buffer (10 mM Tris, pH 8.0, 1 mM EDTA) and 5 mg lysozyme and incubated at 37 °C for 1 hour, then 0.6 mg of proteinase K (Qiagen) and 700 µl of 10% SDS were added and the mixture was vortexed prior to incubation at 65 °C for 10 minutes. Following incubation 1 ml of NaCl 5 M and 800 µl CTAB (Sigma) was added and the incubation at 65 °C repeated. Finally 7 ml chloroform/ iso-amylalcohol mix (24:1) was then added and the mixture vortexed.

Following centrifugation at 14000 *g* for 5 minutes, the top aqueous phase was recovered. Isopropanol was mixed at an equal volume and the solution was incubated at -20 °C for 1 hour. The precipitated DNA was pelleted by centrifugation at 14000 *g* for 20 minutes at 4 °C and the pellet washed with cold 70% ethanol. DNA was air dried and finally resuspended in 500 µl of TE buffer.

2.3.10.2.2 *M. smegmatis*

For screening *M. smegmatis*, individual colonies were picked into 5 ml LB and 0.05% Tween 80 plus antibiotics for selection. Genomic DNA was extracted following a similar method to Talaat *et al.*, 1998. Cells were pelleted by centrifugation at 4000 *g* and resuspended in PBS before undergoing sonication. Cells were sonicated on ice using Sonic Vibra Cell™ (Model VCX 130) at 40% amplitude for 6 minutes with a pulse rate of 2 seconds on, 2 seconds off. Cells were then boiled for 10 minutes and following centrifugation at 15000 *g* for 20 seconds the supernatant was used for PCR amplification.

2.3.11 Site Directed Mutagenesis (SDM)

2.3.11.1 Primer Design

Generation of site directed mutants involved designing sense and anti-sense primers, harbouring the desired mutation, using the Stratagene primer design program. Producing optimal mutagenic primers involves taking into account any potential mismatches, any primer folding and secondary structures and the codon replacement. Primers were designed to be between 25 and 45 bases in length with a melting temperature (T_m) of greater or equal to 78 °C. T_m was estimated using the formula $T_m = 81.5 + 0.41(\%GC) - 675/N - \%mismatch$, where N is the primer length in bases.

The desired mutation was included in the middle of the primer with up to 15 bases of correct matched sequence flanking each side. Primers required an optimal GC content of 40% and were designed so they terminated in one or more C or G bases. The primers used in the site directed mutagenesis were produced and purified by Sigma Aldrich using polyacrylamide gel electrophoresis (PAGE) to ensure maximum purity to increase mutation efficiency.

2.3.11.2 Mutagenesis

SDM was carried out according to the specifications of the QuickChange™ II Site Directed Mutagenesis Kit (Stratagene). Reactions (50 µl) contained 10X reaction buffer, 5–50 ng of dsDNA template, 125 ng of each oligonucleotide primer, dNTP mix (Stratagene), 5% (w/v) DMSO (Sigma), dH₂O and 2.5 U of *PfuUltra* HF DNA polymerase. The control reactions used 10 ng of pWhitescript 4.5 kb control plasmid as the DNA template. *PfuUltra*® Hotstart High-Fidelity DNA polymerase, from the hyperthermophilic archaea *Pyrococcus furiosus*, was used as it has the required 3'-5' exonuclease activity and a high fidelity/proofreading capability. The modified thermal cycling steps used are detailed in Table 2.14.

The Universal Stress Proteins of Bacteria

Table 2.14. Cycling parameters for site-directed mutagenesis. Denaturation, annealing and extension reactions occur at 96 °C, 50 °C and 68 °C respectively.

Segment	Cycles	Temperature	Time
1	1	96 °C	30 seconds
2	22	96 °C	30 seconds
		50 °C	1 minute
		68 °C	6 minutes
3	-	4 °C	Infinite till end

The circular plasmid DNA used as a template was produced in a *dam*⁺ strain of *E. coli* usually XL1-Blue or DH5 α . Dam methylase methylates DNA and thus generates the sequence G^mATC, which the *DpnI* enzyme recognises and digests. Dam⁻ strains of *E. coli* are not suitable for the procedure as they propagate unmethylated plasmid DNA.

Following thermal cycling the reaction was then digested with the *DpnI* restriction enzyme, which specifically digests the methylated and hemi-methylated parental DNA, leaving the newly synthesised plasmid containing the desired mutation. Newly synthesised plasmids were then transformed into competent XL1-Blue cells (Stratagene). Plasmids were then isolated, using the Qiagen mini-prep kit, and sent for DNA sequencing to confirm the mutagenesis.

2.3.12 Polyacrylamide Gel Analysis of Oligonucleotides

Oligonucleotides 18-50 nucleotides in length were run on 10% polyacrylamide gels containing 10% urea (Table 2.15) using the Biorad mini-protean system. Urea is dissolved in the solution by gentle heating to 37 °C, before addition of ammonium persulfate and TEMED. After complete polymerisation of the gel (40 minutes) the gels were pre-run for 30 minutes at 200 V in 1X TBE buffer. Samples were mixed with 1.25X loading dye (Table 2.16) and heated for 2 minutes at 95 °C and chilled on ice. Samples containing 200 pmol of oligonucleotide were loaded into the wells and the gel run at 200 V for 1.5 hours.

The Universal Stress Proteins of Bacteria

Table 2.15. 10% Polyacrylamide denaturing urea gel.

Solution	Per Gel
40% Acrylamide/Bis-acrylamide (29:1)	3.75 ml
Urea	7.2 g
10X TBE	1.5 ml
Distilled H ₂ O	Up to 15 ml
Ammonium persulfate (10%)	75 µl
TEMED (N,N,N',N'-Tetramethylethylenediamine)	7.5 µl

Table 2.16. Loading Dye for denaturing urea PAGE.

Solution	Per ml
Formamide	900 µl
0.5 M EDTA pH 8.0	22 µl
Orange G (7.5%)	26 µl
Bromophenol blue (7.5%)	26 µl
Xylene cyanol FF (7.5%)	26 µl

2.4 Biochemical Assays

2.4.1 *In vitro* Kinase Assay

A typical assay was performed by incubating Rv1636_{WT} (5µM) in a 10 µl volume containing 10 mM Tris-HCl, 50 mM NaCl, and 5 µCi [γ -P³²] NTP (3,000 Ci/mmol) (Perkin Elmer). Reactions were incubated for 30 minutes at 37 °C before being terminated by addition 10 µl of 2X Laemmli loading buffer and heated at 95 °C for 5 minutes. Samples were run on SDS-PAGE gels, then the gels were then soaked in transfer buffer for 10 minutes. The Hybond membrane (Amersham) was also soaked in transfer buffer as preparation for the transfer. The transfer of the proteins to the membrane was performed in a Biorad tank by electrophoresis for 2 hours at 30 V. The membrane was then washed in water to remove nonspecifically bound nucleotides, before being subjected to phosphoimaging. Radioactive proteins were visualized by exposure of the membranes to imaging plates (BAS-IP MS2040, FujiFilm). The plates were scanned by a Fuji FLA-5000 phosphoimager and the signals quantified using the AIDA Image Analyser.

To test Rv1636 kinase activity, the autophosphorylation reaction was repeated with 5 µg of myelin basic protein (MBP_r) (NEB) added as a substrate. Other substrates tested include using 5 µg

The Universal Stress Proteins of Bacteria

of Bovine Serum Albumin (BSA) (Sigma) and Histone (H1 fraction) (Sigma). Where an alternative HEPES buffer was used the buffer contained 20 mM HEPES, 25 mM NaCl, 10 mM MgCl₂, 1 mM EGTA and 1 mM DTT as described by Walter *et al.*, 2000. As a positive control test for kinase activity, the Protein Kinase Catalytic Subunit A (Sigma) was used. Protein kinase A was reconstituted in water containing the reducing agent DTT at a concentration of 6 mg/ml.

2.4.2 Phosphoamino Acid Analysis

To determine whether serine, threonine, and/or tyrosine residues are phosphorylated in the autokinase activity, phosphoamino acid analysis was performed by semi-dry thin layer electrophoresis in the Multiphor II electrophoresis unit. *In vitro* phosphorylation was carried out as previously described. The proteins were separated by SDS-PAGE and the radiolabelled proteins were transferred to a polyvinylidene difluoride (PVDF) membrane (Immobilon P, Millipore). The area of the membrane containing the radiolabelled protein was excised and the protein hydrolysed by heating the sample at 100 °C in 200 µl 6 M hydrochloric acid for 2 hours. The membranes were then washed three times with 200 µl 30% methanol, 0.1 M HCl. The 6 M HCL and the three washes were pooled and dried using the speed-vac before being resuspended in 10 µl distilled water.

The sample was then subjected to thin layer electrophoresis and the migration of the radiolabelled phosphoamino acids compared to phosphoamino standards; O-Phospho-L-serine, O-Phospho-L-threonine and O-Phospho-L-tyrosine (Sigma). The Multiphor II electrophoresis unit was pre-cooled to 16 °C using the MultiTemp® III thermostatic circulator. Glass-backed TLC plates (Merck Cat No:1.05716.0000) were marked at the origin and migration points. Samples were spotted at the origin and dried using a hair dryer. 25 nmol of each phosphoamino acid standards were also spotted. Bromophenol blue (0.1%) was used as a tracking dye.

The plate was sprayed with thin layer electrophoresis buffer and placed in the Multiphor unit with the origin at the cathodic (-) end. Amersham Biosciences CleanGel™ electrode strips were pre-soaked with the electrophoresis buffer before being layered across the TLC plates. Three of the glass backs of the TLC plates were used to ensure adequate contact with the electrodes and proper humidity. The Multiphor II was run for 40 minutes at 1000 V, with current (30 mA) and power (30 W) limited. To detect the phosphoamino acids standards, the plates were sprayed with ninhydrin solution (0.2% (w/v) ninhydrin in ethanol, 10 ml) and heated using a hair dryer for 10 minutes. To detect the radiolabelled amino acids, the plate was subjected to phosphoimaging as described previously.

2.4.3 *In vitro* Nucleotidylylation Assay

Nucleotidylylation reactions were performed as described for the *in vitro* kinase reactions using [α - 32 P] GTP or [α - 32 P] ATP as the substrate. Reactions were resolved by SDS-PAGE and radiolabelled bands were excised, crushed and incubated in an elution buffer containing 50 mM NH_4HCO_3 , 0.1% (w/v) SDS and 1% (v/v) 2-mercaptoethanol at for 37 °C for 18 hours gently shaking. Eluates were cleared by centrifugation. Acetone precipitation was performed using 50 μg of BSA as a carrier protein. Samples were mixed with 1.5 volumes cold acetone (-20°C), vortexed and incubated 1 hour at -20°C.

The precipitate was pelleted by centrifugation at 14000 *g* for 10 minutes. Half the sample was then subjected to 0.5 M piperidine treatment at 37 °C for 2 hours. Piperidine was removed by speed-vac. Samples were resuspended in distilled water and spotted on to TLC cellulose polyethylenimine impregnated plates (Macherey-Nagel). The TLC was performed in 0.7 M boric acid and 0.4 M K_2HPO_4 . Standards of 20 nmole of NMP and NTP were run and viewed under short-wave UV light. The TLC plate was phosphoimaged as previously described.

2.4.4 *In vitro* RNA Triphosphatase Assay

For RNA work, all experiments were carried out DEPC treated water. RNA substrates consisting of 45 nucleotides were transcribed in the presence of 65 μCi [α - P^{32}] GTP (3,000 Ci/mmol) or [γ - P^{32}] GTP (3,000 Ci/mmol) using the AmpliScribe™ T7 RNA polymerase. One μg of pGEM-4Z (Promega) vector, previously linearised with *Sma*I and purified from an agarose gel, was used as the template. CTP, UTP and ATP were at a final concentration of 1 mM; whereas non-radio-labelled GTP was at a final concentration of 0.03 mM. Where required 40 units of RNasin (Promega) was used in the 50 μl reaction to prevent RNase activity.

Following RNA transcription for 2 hours at 37 °C, the template DNA then digested with DNase I for 15 minutes at 37 °C. Free nucleotides were then removed using Sephadex G-25 column (Amersham Biosciences). The quality and purity of the purified 45-mer were [α - P^{32}] or [γ - P^{32}] RNA was assessed on 10% polyacrylamide gels containing 7 M urea. RNA was located using the phosphoimager Fuji 5000 and visualised by comparison to the electrophoresis of the loading dyes (described in Chapter 2.3.12).

The 20 μl RNA triphosphatase reactions were performed as described by (Kim *et al.*, 2003). 0.4 μM RNA was mixed with up to 60 pmol Rv1636 and incubated at 37 °C for 1 hour. Reactions

The Universal Stress Proteins of Bacteria

were halted with loading dye containing formamide and EDTA. HK^(TM) Thermolabile Phosphatase was used as a control in the RNA triphosphatase experiments and required 5 mM CaCl₂ for activity. It is a phosphatase and can dephosphorylate RNA at the 5' end.

2.4.5 *In vitro* NTPase Assay

Purified Rv1636_{WT} (up to 10 µM) was incubated in a reaction buffer containing 10 mM Tris-HCl (pH 7.5), 50 mM NaCl, nucleotide triphosphate at 10 µM and 0.6 µCi/ml (3,000 Ci/mmol) [α -³²P] GTP/ATP. Where required cations were added at concentrations of 5 mM and EDTA was added at concentration of 1 mM. Reaction volumes of 20 µL were incubated for 30 minutes at 37 °C before being halted by addition of 100 µL 2 M formic acid. One µL was applied to polyethyleneimine cellulose (PEI) sheets (Macherey-Nagel) and mixtures were separated by TLC using 0.7 M boric acid and 0.4 M K₂HPO₄. Hydrolysis products were visualized using phosphoimaging as described previously. Unlabelled nucleotide standards (20 nmole) were viewed as under short-wavelength UV light.

2.4.6 Nucleotide Binding Experiments

Rv1636_{WT} (30 µM) was incubated in a reaction buffer (50 mM Tris-Cl, pH 7.5, 50 mM NaCl, 10% glycerol), and 5 µCi [α -³²P] GTP or 5 µCi [α -³²P] ATP. The total reaction volume was 20 µL. The reactions were then incubated on ice and subjected to either UV crosslinking or vacuum filtration to assay the nucleotide binding. Various concentrations of unlabelled nucleotides GTP, ATP, GDP and ADP, GMP and AMP were added where described. Bovine serum albumin, a stable non-enzymatically active protein, was used as a control protein. Radiolabelled nucleotide was mixed with unlabelled nucleotides prior to addition to the reaction mixture and proteins were added last to the reactions prior to incubation.

2.4.6.1 *UV Cross-linking Experiments*

UV light when directed at proteins bound to nucleotide causes the cross-linking of the protein to the nucleotide. Reactions were performed as described by Lin, 1999, and Preneta, 2004. Reactions were incubated at 4 °C for 40 minutes with a UV light 2 cm above. Samples were then

The Universal Stress Proteins of Bacteria

mixed with equal volume 2X Laemmli buffer and subjected to SDS-PAGE. Gels were washed three times with 5 minute washes in dH₂O. The gels were then blotted dry with Whatman filter paper, wrapped in cellophane and exposed to a BAS-IP MS2040 plate. Plates were read with a Fuji FLA-5000 phosphoimager. ³²P is a high energy emitter and can pass through cellophane.

2.4.6.2 Vacuum Filtration

Vacuum filtration as described by Ormo and Sjoberg, 1990, Rombel *et al.*, 1999 and Tameling *et al.*, 2002 was performed to measure the nucleotide binding capabilities of the Rv1636 proteins. A slot blot manifold (PR648, Amersham Biosciences) provided 48 wells, each with a capacity of 1 ml. 20 µl reactions, containing 30 µM Rv1636 protein, were incubated on ice for 40 minutes before being added to 180 µl reaction buffer. Increasing the reaction volume ensure consistency when filtering the samples through the slots membrane and gave equal distribution across the area.

Samples were loaded onto the slot blot and vacuum filtered through a nitrocellulose 0.45 µM membrane (Whatman). The membrane was previously prepared by soaking in H₂O, followed by reaction buffer for 1 minute. Hydrophobic interactions allowed the proteins to bind the membrane along with any bound nucleotide. Immobilized proteins were then washed twice with 500 µl water to remove any non-specific nucleotide. The membrane was removed, rinsed in 12 mM HCl to remove non-specific bound background radioactivity and air-dried before subsequent exposure to BAS-IP MS2040 plates and Fuji FLA-5000 phosphoimaging.

2.5 Biophysical Techniques

Analyses were undertaken at the Division of Molecular Structure at the National Institute for Medical Research (NIMR) with the guidance of Dr Ian Taylor and Dr Simon Pennell.

2.5.1 Multi Angle Laser Light Scattering (MALLS) Analysis

The absolute molar mass of particles in solution can be determined using multi-angle laser light scattering. The experiments described here were coupled to size exclusion chromatography (SEC) to allow the Mw determination of the protein peaks as they elute. The scattering of the laser

The Universal Stress Proteins of Bacteria

light was recorded at 18 different angles using a DAWN-HELEOS multi-angle laser light scattering detector (Wyatt Technology Corp., Santa Barbara, CA).

Calculating the apparent molecular weight (M_w) was achieved using the Debye plot method (Wyatt 1992, 1993) which involves creating plots of $R(\theta)/Kc$ versus $\sin^2(\theta/2)$, where θ is the scattering angle, $R(\theta)$ is the Rayleigh excess scattering ratio and K is the optical constant. The apparent M_w is then obtained from the intercept of this plot. The optical constant, K , is calculated using $4\pi^2 n^2 (dn/dc)^2 / (\lambda_0^4 N_A)$, where n is the solvent refractive index, dn/dc is the specific refractive index increment in mL/g, N_A is Avogadro's number and λ_0 is the wavelength of the scattered light in the vacuo in cm. The specific refractive index increment is calculated using the equation of Anthonsen *et al.* (1994).

One hundred μ l protein samples at 5 mg/ml in 50 mM Tris and 200 mM NaCl, previously flash frozen in liquid nitrogen, were injected into the chromatography system. Glycerol was not added as the protein concentration is measured using refractive index changes and these are affected by the presence of glycerol. As the protein sample eluted from the SEC column, the concentration was measured using the refractive index change ($dn/dc = 0.186$) at 1 second intervals. The light source for the refractive index was at wavelength 658 nm and was produced using an OPTILAB-rEX differential refractometer, coupled with a Peltier temperature-regulated flow cell to keep the temperature constant at 25 °C (Wyatt Technology Corp., Santa Barbara, CA).

Data from the light scattering was then analysed using the ASTRA software version 5.1 (Wyatt Technology Corp., Santa Barbara, CA). Debye plots ($R\theta/Kc$ vs $\sin^2(\theta/2)$) were constructed and the intercept of the individual plots were used to calculate the weight averaged molecular weight for each of the 1 second readings. The overall molecular weight and the polydispersity term for each species was then calculated by combination of the average the light scattering intensity with the protein concentration measurements.

2.5.2 Analytical Ultracentrifugation

Sedimentation coefficients were obtained using a Beckman Optima XLA analytical ultracentrifuge, using aluminium double sector quartz cells in an An-60 Ti rotor. The protein samples were made to a concentration up to 250 μ M and were in a buffer containing Tris 50mM and NaCl 200 mM. Samples were then centrifuged at a rotor speed of 55,000 rpm under a constant temperature of 293 kelvin. During centrifugation radial scans were collected every 300 seconds at $\lambda=280$ nm or $\lambda=230$ nm. The recorded data was analysed using the SEDFIT (Schuck, 2000) program, in terms of discrete species or to the continuous size distribution functions, $C(S)$ and $C(M)$.

Calculations of the solvent density and protein partial specific volumes were performed using tabulated values (Laue *et al.*, 1992)

2.6 Identification of Protein-Protein Interactions

2.6.1 Cell Fractionation

M. bovis BCG was cultured in Middlebrook 7H9 broth containing BBL™ Middlebrook OADC Enrichment media, 0.5% glycerol and 0.05% Tween 80. Cultures were grown in 250 ml conical flasks in 50 ml of media at 37 °C on an orbital shaker at 100 rpm. The optical density was recorded using the Eppendorf Biophotometer v.1.32. Cells were pelleted by centrifugation at 20000 *g*. The *M. bovis* BCG cells were fractionated in order to obtain the proteins present within the cells. The cells were resuspended in dH₂O and were washed using a hand-homogeniser. The cells were then pelleted by centrifugation for 10 minutes at 20000 *g* in J2-HS Beckman Centrifuge. The washing was repeated two further times.

The final pellet was resuspended in 5 ml PBS buffer with 0.4% Tween 80 (v/v) and 0.75 ml proteinase inhibitors (Complete mini tablet (Roche)). The cells were then broken open under high pressure (20 kps) using a cell disruptor (Constant Cell Disruption System; Constant Systems Ltd). Cells broken open using the cell disruptor were pressured three times with 1 minute on ice in between cycles. The broken cell mixture was then centrifuged at 20000 *g*. The supernatant containing the soluble cellular proteins was stored at -20 °C until use. Protein concentration was measured by a Bradford assay using the Better Bradford Assay™ Reagent (Pierce).

2.6.2 GST Pull-down

To perform the GST-pull-down, the Glutathione Sepharose™ 4B (GS4B) (Amersham) beads were prepared and the cytosolic fraction was pre-cleared to remove any nonspecific proteins that bound to GST or the GS4B beads. Preparation of the GS4B bead matrix 50% slurry was done by resuspending the beads by gentle shaking of the bottle. 1.33 ml GS4B was centrifuged in 20 ml falcon tube in the Beckman Coulter Allegra™ 6R Centrifuge at 500 *g*. The supernatant was removed and 10 ml cold PBS was added. The beads were resuspended and pelleted by repeating the centrifugation. Again, the supernatant was removed and 1 ml of PBS added. The matrix was stored at 4 °C. The cytosolic fraction was pre-cleared by mixing 1 ml cytosolic fraction with 50 µl GS4B 50%

The Universal Stress Proteins of Bacteria

slurry and 25 µg of GST added before incubation at 4 °C for 2 hours with gentle mixing. After a 2 minute centrifugation at 16000 g, the supernatant (precleared cytosolic fraction) was transferred to a new tube and stored at -20°C.

The pull-down assay was performed by mixing 1 ml precleared cytosolic fraction with 50 µl 50% GS4B slurry and 10 µg GST-Rv1636 protein before incubation for 2 hours at 4 °C. GST-Rv1636 protein was previously purified by IMAC purification by Leiva, 2007. The mixture was centrifuged at 16000 g in microcentrifuge for 2 minutes and the supernatant removed. The pellet was washed three times using 1 ml PBS buffer and 1 minute centrifugations. The washed pellet was mixed with 50 µl elution buffer (20 mM reduced glutathione, 50 mM Tris-HCl pH 8.0) and gently mixed. After 10 minutes at room temperature, the reactions were centrifuged at 16000 g for 2 minutes and the supernatant collected. The elution was repeated a further time and the eluates pooled. The eluted sample was then subjected to SDS-PAGE.

An alternative method of elution was boiling the beads with 2X Laemmli for 5 minutes and loading straight on to SDS-PAGE. Pull-down assays were performed under a range of conditions. Divalent ions were included by adding manganese chloride tetrahydrate ($\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$), magnesium chloride (MgCl_2) or calcium chloride (CaCl_2) at concentrations of 5 mM. Nucleotide triphosphates (ATP or GTP) were added at concentrations of 10 µM. These were added before the initial 2 hour incubation.

2.6.3 Co-immunoprecipitation

The cytosolic fraction was precleared by incubating 100 µl of Protein A:Agarose in every 1 ml of cellular protein mix for 30 minutes at 4 °C or alternatively using Merck Pansorbin® cells which are *Staphylococcus aureus* expressing Protein A. The beads were then removed by centrifugation at 3000 g for 2 minutes. Polyclonal rabbit anti-His₍₆₎Rv1636_{WT} antibodies, produced by BioGenes GmbH., Germany, were already available in the laboratory. Anti-His₍₆₎Rv1636_{WT} antibody was incubated in 750 µl *M. bovis* BCG cytosolic fraction overnight at 4 °C on a rocking platform (Grant Boekel, V5R23). 30 µl agarose beads (Protein A agarose suspension; Calbiochem®) were added and the solution was incubated for 1 hour at 4 °C. The beads were pelleted by centrifugation at 800 g for 1 minute. The beads were then washed with 1 ml breaking buffer four times with gentle inversions to resuspend.

An equal volume of 2X Laemmli was added and the mixture was heated for 5 minutes. The sample was then loaded on to SDS-PAGE gel. Anti-SecA antibody was used as a control antibody.

The Universal Stress Proteins of Bacteria

Attempts to scale up the procedure involved eluting the protein using 50 µl 0.1 M Glycine pH 2.5 vortex and incubating with agitation for 10 minutes at 4 °C. After centrifugation at 300 g, at 4 °C for 2 minutes of 5 µl 1M Tris pH 8.0 to each tube to neutralize the pH, before concentrating the samples using Centriprep Columns (3000 mwco) (Millipore). The immunoprecipitation sample was subsequently analysed by SDS-PAGE.

A modified Shevchenko stain (Shevchenko, 1996) was used that is compatible with mass spectrometry. In brief, the gel was fixed in 50% methanol/ 5% acetic acid for 20 minute before wash with 50% methanol followed by a wash overnight with water. The gel was then incubated in the sensitizer solution (0.02% sodium thiosulphate) for 1 minute followed by two 1 minute washes with water. The gel was then incubated at 4 °C in the silver reagent (0.1% silver nitrate) before being rinsed with water and left to develop for 5-8 minutes in the developer solution (0.04% formalin in 2% sodium carbonate). The development was halted by placing the gel in 5% acetic acid.

2.6.4 Macrophage Infection with *Mtb*

Murine derived J774 cells (ATCC number TIB-202) were grown in Dulbecco's Modified Eagle's Medium (DMEM) (Sigma) supplement with 10% heat inactivated calf serum. The experiments were carried out in a high containment category three laboratory in a class I safety cabinet. Screening the *Mtb* strains, including H37Rv and the putative mutants, for murine macrophage survival involved infecting J774 macrophages (plus interferon γ) with *Mtb* and taking samples at 0, 24, 48 and 72 hours.

Macrophage cells at 0.5×10^6 cells/ml in 2 ml DMEM were added to wells of the 24 well plates and left to grow confluent overnight. *Mtb* cells were pelleted from the cultures and resuspended in DMEM, without serum. *Mtb* was added to the macrophage monolayer at a multiplicity of infection (m.o.i.) of 10:1 (bacteria:cells). The infected cells were then transferred to 5% CO₂/37 °C incubator and incubated for at least 6hr. The media was then removed and the wells washed with PBS three times. Any extracellular bacteria were killed by addition of media containing 50 µg/ml gentamicin before washing.

Macrophages were then scraped from the wells with PBS/0.1 % Triton 100X to cause macrophage lysis. The samples were then plated out on 7H10/OADC/ 0.2% glycerol plates to allow determination of CFU/ml. Triplicate 20 µl aliquots of 1/10 dilutions in culture media were plated, then counted 14-21 days later after suitable growth was observed.

Chapter 3 - Biochemical Characterization of Rv1636

3.1 Introduction

USPs are widely distributed in the three domains of life. Although over 3000 proteins have been classed as USPs, the role of the proteins remains largely unknown and USPs are poorly characterised in terms of biological function and activity. In this chapter, the purification and biochemical characterization of Rv1636_{WT} and Rv1636 site directed mutants are described. The aim was to investigate the biochemical activities of Rv1636 more robustly. In the Williams group, previous work on the Rv1636_{WT} protein revealed some putative biochemical activities, including phosphorylation, nucleotidylylation and RNA triphosphatase activities. In this chapter, efforts to further characterise the USPs are described, with aim of linking the classification of USPs to their biochemical activities.

3.1.1 Biochemical Activities of USPs

The USPs of *E. coli* are the most widely studied USPs with respect to their biochemical activities, with previous experiments largely focusing on *E. coli* UspA and UspG. Freestone *et al.*, 1997, showed that UspA is a serine and threonine phosphoprotein. They detected three isoforms of UspA *in vivo* and determined, by alkaline phosphatase treatment, that two of these isoforms were the result of phosphorylation. The phosphorylation was shown to be dependent on the *typA* (*o591/Rv1165*) gene, which codes for an autophosphorylating tyrosine phosphoprotein. The protein has homology to known GTP-binding translation elongation factors, for example the TypA protein from *E. coli*. *In vitro* chase experiments, which followed the phosphate from donors including GTP and ATP, showed UspA to be an autophosphorylating protein.

The same authors showed that phosphorylation of UspA *in vitro*, with both [γ -³²P] ATP and [γ -³²P] GTP, was highest when divalent cations, either magnesium or manganese, were present in the reaction. No phosphorylation was observed when using [α -³²P] ATP, meaning the alpha-phosphate did not remain bound to the protein. The phosphorylation of UspA was stable to acid hydrolysis and unstable to base hydrolysis, which is consistent with serine or threonine phosphorylation. The protein TypA was shown to be involved with the autophosphorylation of UspA. A mutant lacking different isoforms of UspA was found to contain an insertional mutation in the *TypA* gene. They concluded that the gene *TypA* was required for the phosphorylation of UspA under carbon starved conditions.

The Universal Stress Proteins of Bacteria

Weber and Jung (2006) showed UspG from *E. coli* was also post-translationally modified and could be both phosphorylated and adenylylated when incubated with radiolabelled γ and α –ATP respectively (Figure 3.1). Acid stability of the isoforms indicated the modifications were at serine or threonine residues.

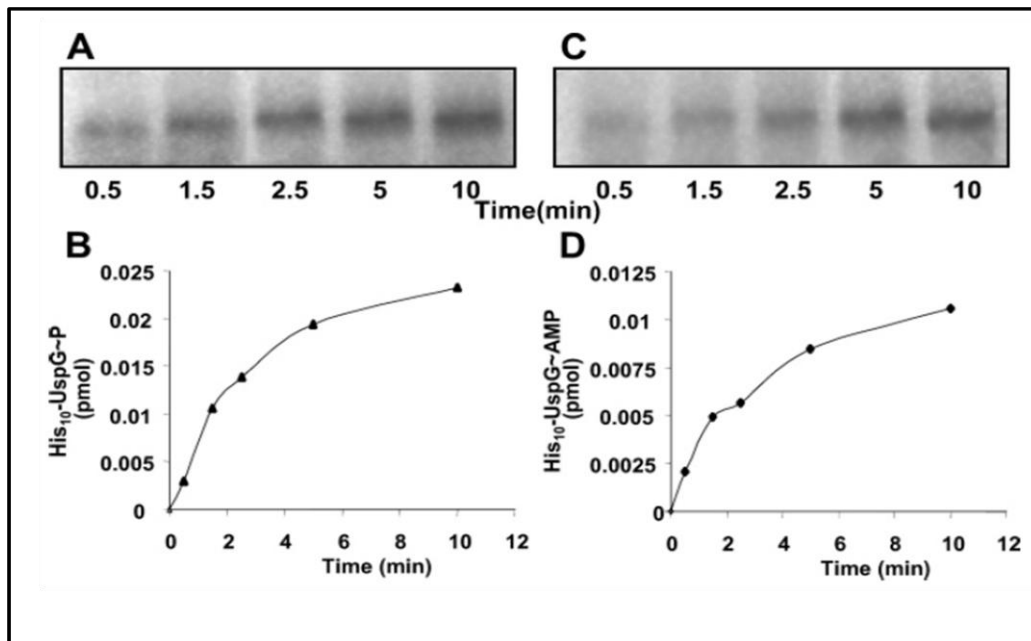


Figure 3.1. Time dependent modification of His₁₀-UspG *in vitro*. (A) Incubated with [γ -³²P] ATP to test for autophosphorylation (B) Quantified phosphorylation. (C) Incubated with [α -³²P] ATP to test for adenylylation. (D) Quantified adenylylation. Samples were run on 12.5% SDS and transferred to nitrocellulose membrane. Radiolabelling was detected by phosphoimaging. (Figure from Weber & Jung (2006)).

Post-translation modifications play key roles in regulatory mechanisms, often by altering the activity levels of enzymatic activities. Over-expression of UspA was shown to alter the isoelectric point of at least six other proteins (Nystrom & Neidhardt 1996), suggesting the UspA is involved in the post translation modification of other proteins.

3.1.2 Biochemical Characterisation of *Mtb* USP Rv1636

Rv1636_{WT} was previously shown to be modified by apparent nucleotidylation and preliminary evidence has shown that Rv1636_{WT} possesses an autokinase ability, allowing it to

The Universal Stress Proteins of Bacteria

undergo autophosphorylation. As previously mentioned, autoadenylation has been described for UspG (Weber & Jung, 2006), suggesting that this may be a general property of USPs. Similarly, *E. coli* UspA was shown to be an autophosphorylating serine/threonine phosphoprotein (Freestone *et al.*, 1997).

3.1.2.1 Autokinase Activity

It has been previously shown that incubation of Rv1636_{WT} with [γ -³²P] ATP/GTP leads to protein labelling with the γ -phosphate in a protein concentration dependent manner (Leiva, 2007). Figure 3.2 shows the radiolabelling of increasing concentrations of Rv1636_{WT} when incubated with [γ -³²P] ATP. In order to study the putative autokinase activity, further experiments were planned and included phosphoamino acid analysis and determination of whether Rv1636_{WT} could act as a kinase and transfer the phosphate group to other proteins.

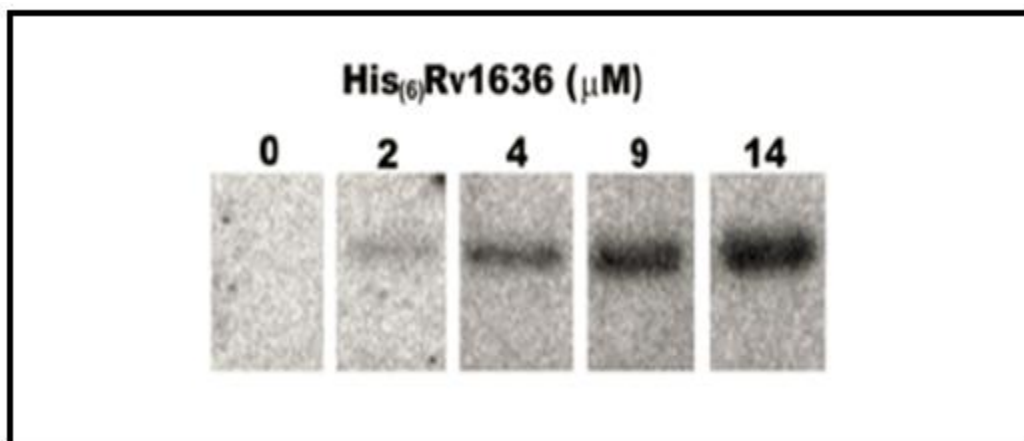


Figure 3.2. Apparent *in vitro* autophosphorylation of Rv1636_{WT}. Labelled band runs at same level as Rv1636_{WT} at 16 kDa. Rv1636_{WT} was incubated with [γ -³²P] labelled ATP for 30 minutes at 37 °C. Radioactivity was detected using phosphoimaging. Figure from Leiva (2007).

3.1.2.2 Nucleotidylation Activity

Previous work by Leiva, 2007, in the Williams group showed apparent nucleotidylation of Rv1636_{WT} when incubated with [α -³²P] ATP/GTP. The Rv1636 protein was covalently labelled with the α -phosphate. Figure 3.3 shows the radiolabelling of increasing concentrations of Rv1636_{WT} when

The Universal Stress Proteins of Bacteria

incubated with [α - 32 P] GTP. Furthermore, it was determined that the nucleotidylation was alkaline stable and acid labile. Further investigation into the putative auto-nucleotidylation activity would involve determining the bound state of the moiety and determining if the site directed mutants of the conserved nucleotide binding site would retain the activity.

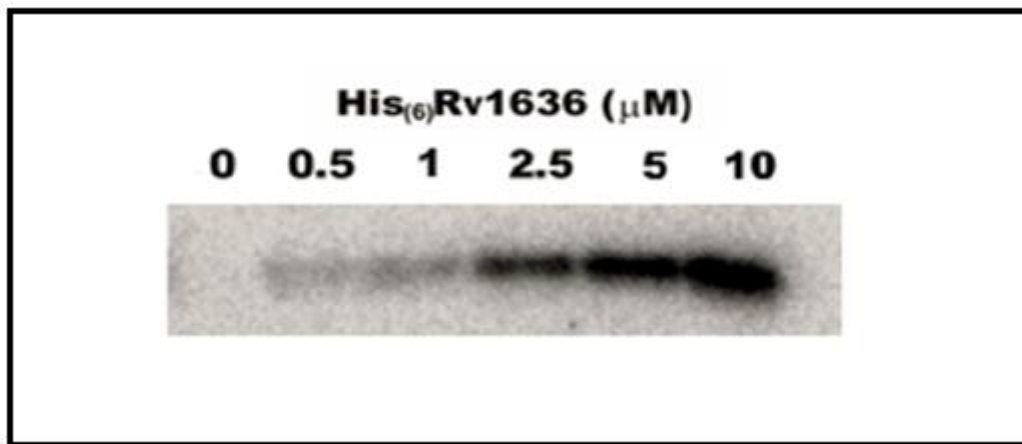


Figure 3.3. Apparent *in vitro* nucleotidylation of Rv1636_{WT}. Labelled band runs at same size as Rv1636_{WT}. Rv1636_{WT} was incubated with [α - 32 P] labelled GTP for 30 minutes at 37 °C. Radioactivity was detected using phosphoimaging. Figure from Leiva (2007).

3.1.2.3 RNA Triphosphatase Activity

Preliminary experiments by Leiva, 2007, showed that Rv1636_{WT} displayed RNA triphosphatase activity when incubated with [γ - 32 P] labelled RNA (as shown in Figure 3.4A). The protein appeared to cleave the terminal γ -phosphate from the RNA strand. RNA triphosphatase activity has been shown to be important in the baculovirus *Autographa californica* nuclear polyhedrosis virus (AcNPV) (Gross & Shuman, 1998a) and also the eukaryotic capping enzyme CEL-1 triphosphatase of *Caenorhabditis elegans* (Takagi *et al.*, 1997), during which the 5' triphosphate is cleaved to a diphosphate in the initial step of the capping reaction. In the experiments by Leiva, 2007, the radiolabelled RNA produced was in relatively low amounts, therefore it was necessary to incubate the reaction with a larger amount of unlabelled RNA in order to visualise that the RNA was not being degraded (Figure 3.4B). In the initial RNA triphosphatase activity experiments, it was not determined if the labelled RNA in the reactions was only cleaved at the 5' triphosphate end. Therefore, it was planned to achieve this by performing RNA triphosphatase experiments with [α - P^{32}] labelled RNA and end labelled ($[\gamma$ - P^{32}]-labelled) RNA.

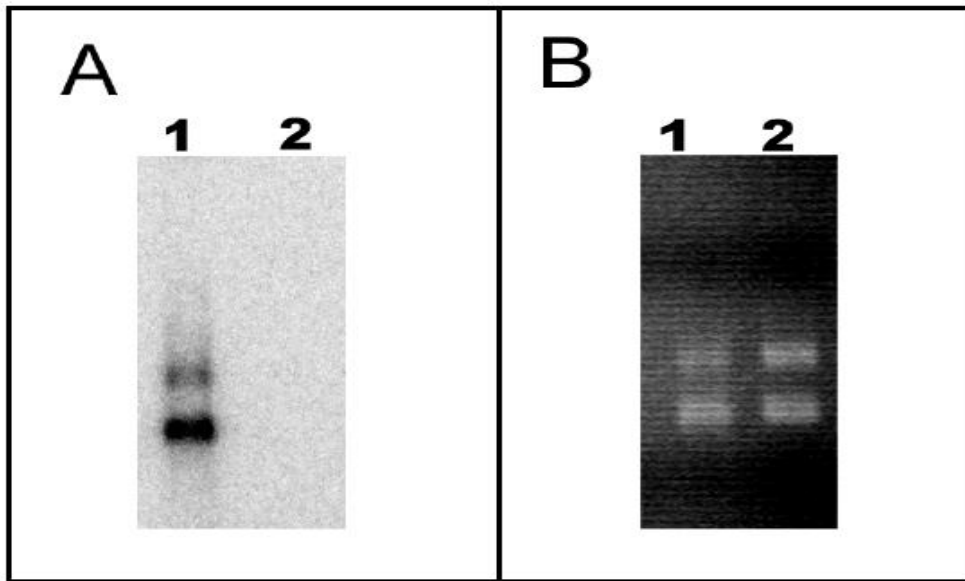


Figure 3.4. Putative RNA triphosphatase releases a radiolabelled phosphate from the γ -labelled RNA transcript. (A) 10 μ L reactions in the absence (lane 1) and the presence (lane 2) of 1 μ g Rv1636_{WT} containing 10 mM Tris-HCl (pH 7.5), 50 mM NaCl, [γ -P³²]-labelled RNA (~10 ng), were incubated for 30 minutes at 37 °C and separated by agarose gel electrophoresis. The gel was visualized by phosphorimaging. (B) Reactions repeated with ~100 ng unlabelled RNA and visualized under UV light. Figure from Levia (2007).

3.1.3 Aims

Rv1636 has previously been shown to have putative autonucleotidylation, autokinase activity and RNA triphosphatase activity. However, these results were from preliminary experiments from which controls were missing. In this chapter, attempts to further characterise these putative activities are described. Residues potentially important for the protein's activity were studied and, in order to further characterise the activities, specific site directed mutants were created. Phosphorylation of a protein occurs on specific residues; in eukaryotes this usually occurs on serine, threonine or tyrosine residues, whereas in prokaryotes phosphorylation commonly occurs on histidine residues. However, the increasing importance of the phosphorylation of serine, threonine and tyrosine residues in prokaryotes is becoming apparent. Nucleotidylation also occurs at specific residues and for further confirmation the release of a nucleotide monophosphate (NMP) moiety was assayed. The possibility that Rv1636 is an RNA triphosphatase is intriguing and in order to confirm this activity, RNA triphosphatase activity was tested using high concentrations of end labelled RNA and also RNA labelled throughout the molecule.

3.2 Results

3.2.1 Protein Purification

3.2.1.1 Rv1636_{WT} Purification

The *Rv1636* gene was originally cloned from *Mtb* H37Rv genomic DNA into pGEM-Teasy vector (Promega) between *NdeI* and *BamHI* restriction sites (Leiva, 2007). The gene was then cloned in the expression vector pET-15b (Novagen) and plasmids recovered and sequenced before being transformed into *E. coli* Rosetta (DE3) to allow protein expression. pET15b::*Rv1636* allows the transcription of *Rv1636* to be driven by a T7 promoter and resulted in the expressed protein being tagged with a N-terminal histidine tag to allow nickel affinity purification.

E. coli Rosetta cells containing the vector pET15b::*Rv1636* were grown to an OD₆₀₀ of 0.6 and induced with 0.1 mM IPTG at 37 °C. The cells were disrupted by Bugbuster Reagent (Novagen). The protein was purified by immobilized metal-ion affinity chromatography (IMAC) purification and the elution chromatograph is shown in Figure 3.5A. The soluble fraction obtained from the culture was run through a His-trap column (GE Healthcare) and eluted fractions analysed by SDS-PAGE as shown in Figure 3.5. The main peak corresponded to an over-expressed protein that migrated in a position similar to previously purified Rv1636_{WT} at 17 kDa.

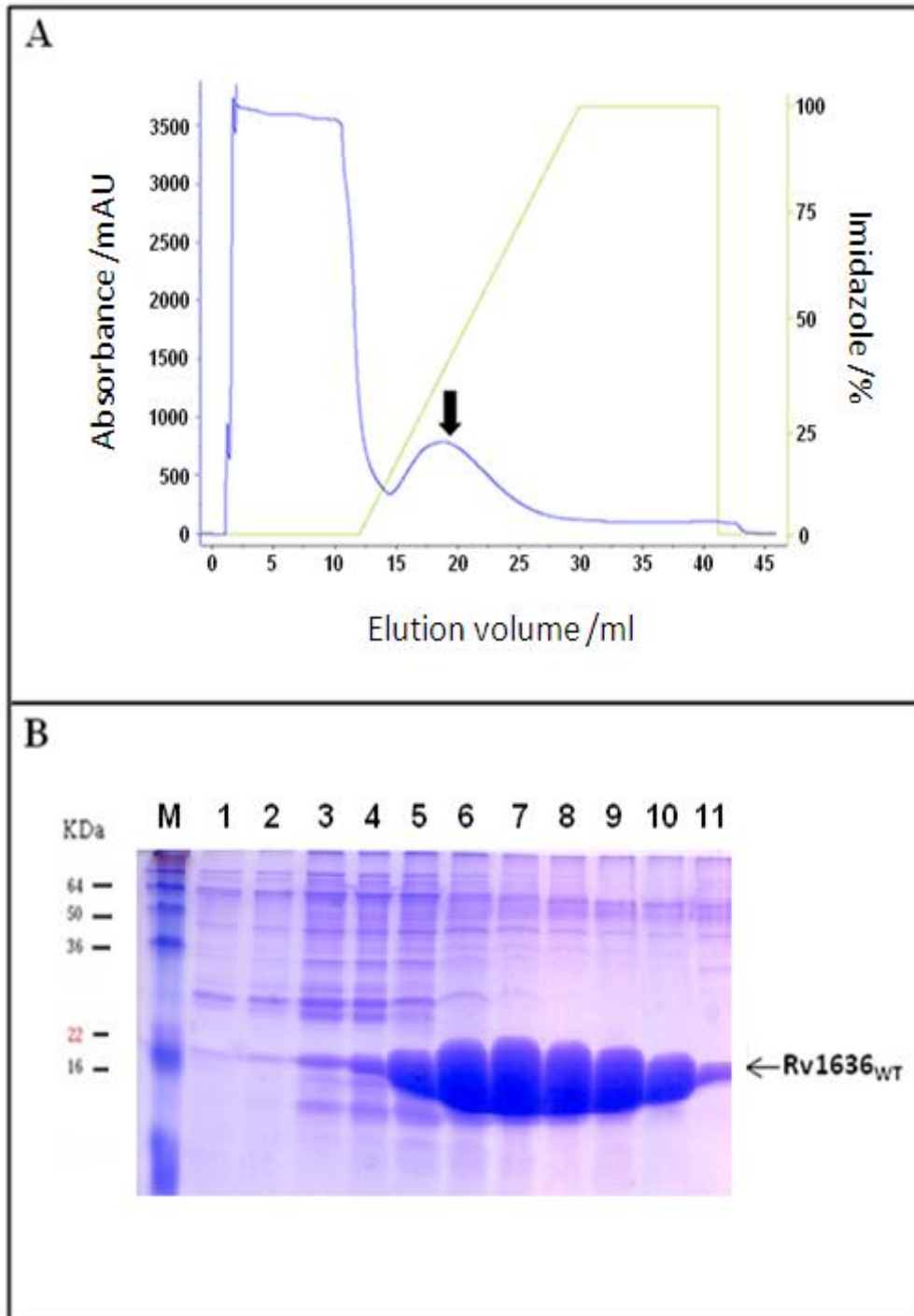


Figure 3.5. Purification of Rv1636_{WT} by immobilized metal-ion affinity chromatography. (A) FPLC elution chromatogram of Rv1636_{WT} purification showing UV_{280nm} reading (mAU, milli-absorbance units) and the imidazole gradient. Arrow shows the peak fractions that were analysed. (B) SDS-PAGE of the eluted fractions from peak. Lanes 1-11 representing elution volume 14-24 ml. Proteins were visualised with Imperial Stain (Pierce). Marker is Invitrogen SeeBlue®Plus2 (Lane M).

The Universal Stress Proteins of Bacteria

Fractions corresponding to a strong peak on the chromatogram were collected. These contained a protein of 17.3 kDa, which is equivalent to the expected weight of the His-tagged Rv1636_{WT} fusion protein. The fractions were desalted and buffer exchanged using an Amicon centrifugal filter device with a molecular weight cut off of 10,000 Da to remove any remaining imidazole. The final purified protein after desalting, buffer exchange and concentration is shown in Figure 3.6.

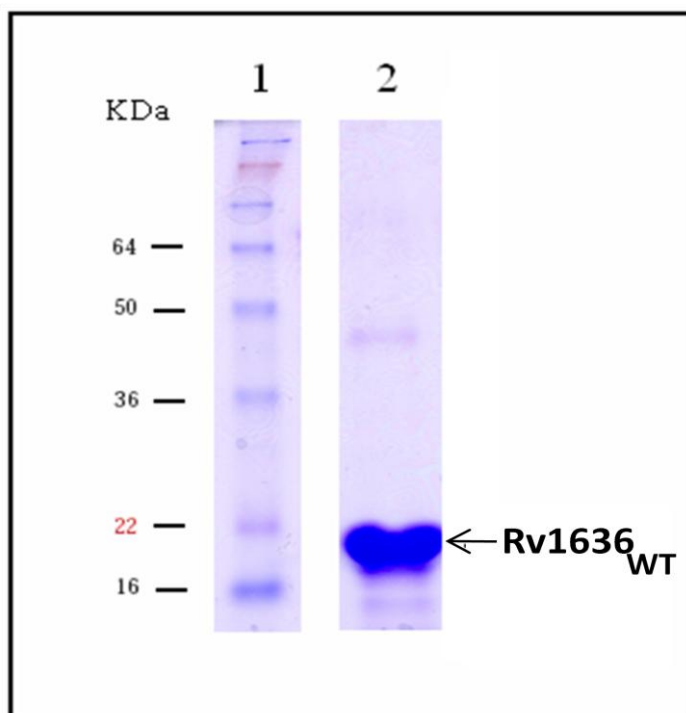


Figure 3.6. Purified His₍₆₎Rv1636_{WT} protein. Expressed in *E.coli* Rosetta cells and purified by IMAC. 12% SDS-PAGE stained with Imperial Stain. Lane 1: SeeBlue® Plus2 marker (Invitrogen), Lane 2: Purified His₍₆₎Rv1636_{WT}.

3.2.1.2 Generation of Site Directed Mutants in Rv1636

Site directed mutants in the putative nucleotide triphosphate binding site were created using the expression vector pET15b::Rv1636 as a template. The G-X₂-G-X₉-GS motif of the Rv1636 protein is highlighted in Figure 3.7. Any nucleotide binding pocket is likely to be disrupted with small changes to amino acid side chains. The conserved residues highlighted as having binding roles to nucleotide triphosphates were substituted to alanine using the Stratagene SDM Kit II according to the manufacturer's instructions.

The Universal Stress Proteins of Bacteria

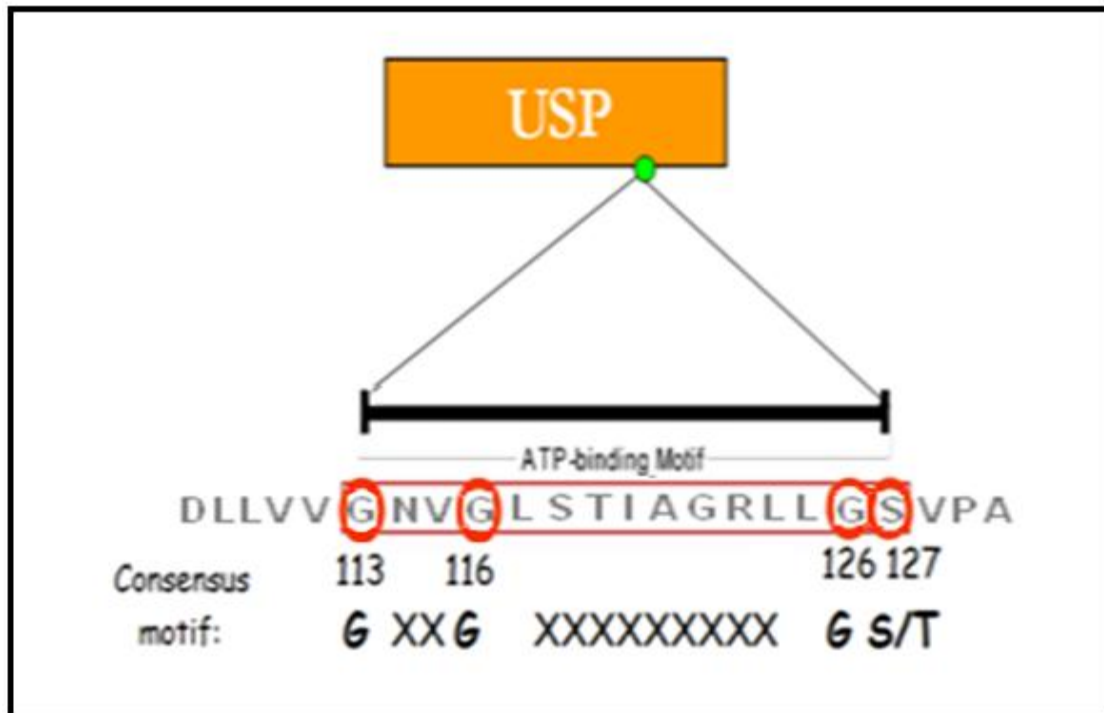


Figure 3.7. Putative nucleotide binding site of Rv1636 with residues selected for mutagenesis highlighted.

Substituting the side chain of the amino acids present in the putative binding site would provide important information about the binding of the nucleotide to the Rv1636 protein. Also substituting the D12 aspartic acid to glutamate was predicted to interfere with the putative nucleotide binding site conformation. The aspartic acid is highly conserved among USP proteins as can be seen in Figure 1.8 in Chapter 1, with the MJ0577 protein, UspA from *E. coli* and *H. influenzae* all containing this conserved residue. The site directed mutants made were therefore G113A, G116A, G126A and S127A in the putative nucleotide binding motif, and D12A. The replacement of the hydrogen side chain of the glycines, within the putative nucleotide binding site, with the larger methyl group of alanine would likely perturb both of the conserved loop regions which are in contact with the nucleotide.

Primers were designed using the Stratagene primer design program and SDM was achieved using Stratagene SDM Kit II according to the manufactures instructions. The SDM plasmids, once confirmed by sequencing to contain the desired mutation, were then transformed into competent *E. coli* Rosetta cells for expression. Expression was carried out as described previously (1.2.1.1 Rv1636_{WT} Purification) and the proteins were purified following solubility tests as shown in Figure 3.8. The concentration of the proteins was determined using absorbance at 280nm, which is dependent on tryptophan, tyrosine and cysteine residues. D12A did not overexpress soluble protein (Figure 3.8) indicating expression of the protein may have been toxic to the cell.

The Universal Stress Proteins of Bacteria

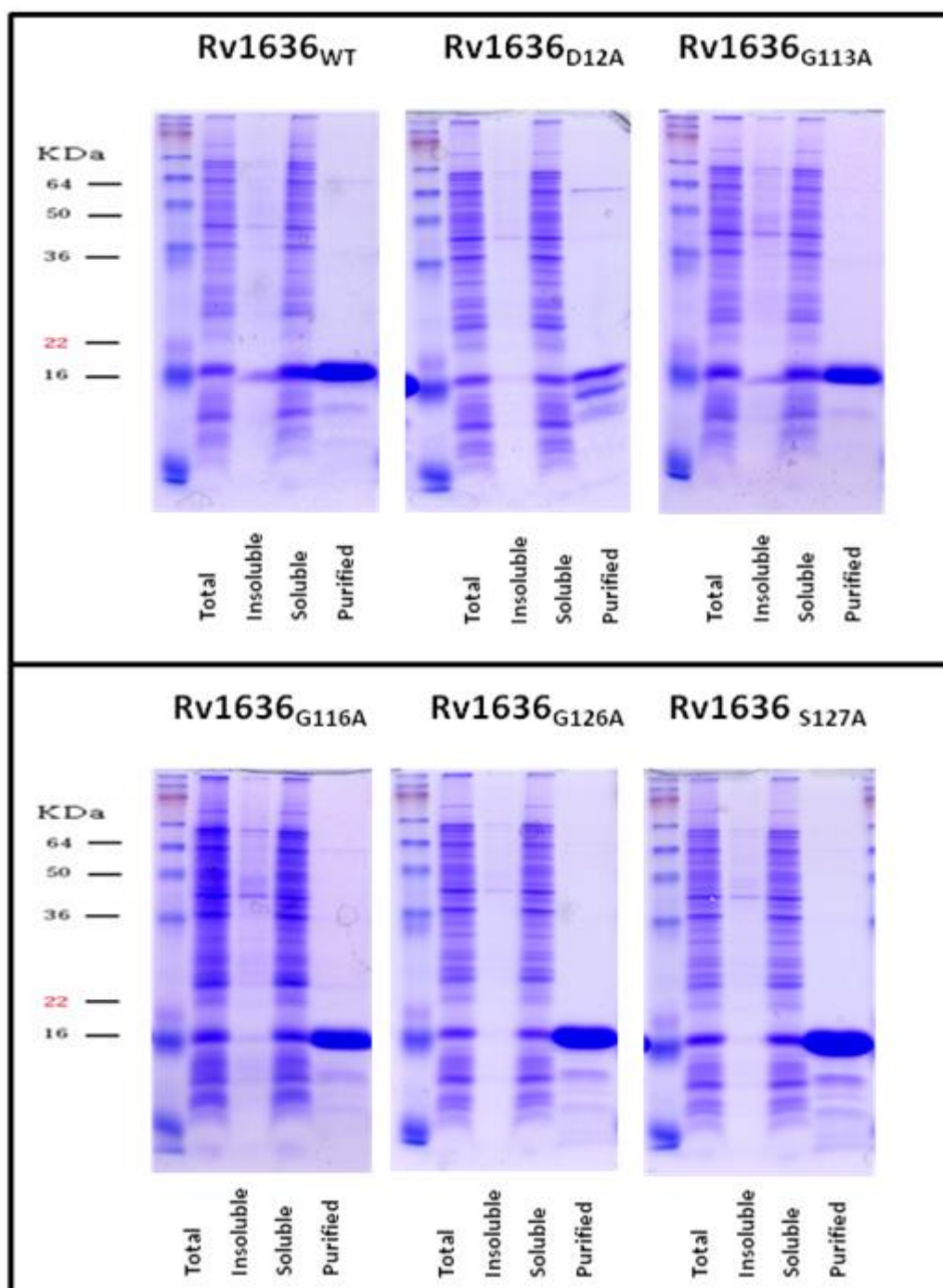


Figure 3.8. Expression and purification of his-tagged site direct mutant proteins. Samples were run on 12.5% SDS-PAGE and stained with Imperial Coomassie Stain (Pierce). Proteins were expressed in *E. coli* Rosetta cells at 37 °C and purified by IMAC. SeeBlue®Plus2 marker (Invitrogen) is shown.

3.2.2 Investigating the Biochemical Activities of Rv1636

3.2.2.1 Investigation of the Rv1636 Autokinase Activity

It was previously shown that Rv1636_{WT} has autokinase activity and is able to undergo autophosphorylation. Further evidence that suggested the Rv1636_{WT} protein can be covalently modified with the γ -phosphate was shown when the Rv1636_{WT} was incubated with [γ -P³²]-GTP and this produced two bands visualised by SDS-PAGE. The presence of a double band (Figure 3.9) suggests there is a modification that alters the migration of the protein. In the case of phosphorylation, the addition of a phosphate group can affect the binding of SDS-molecules to a protein.

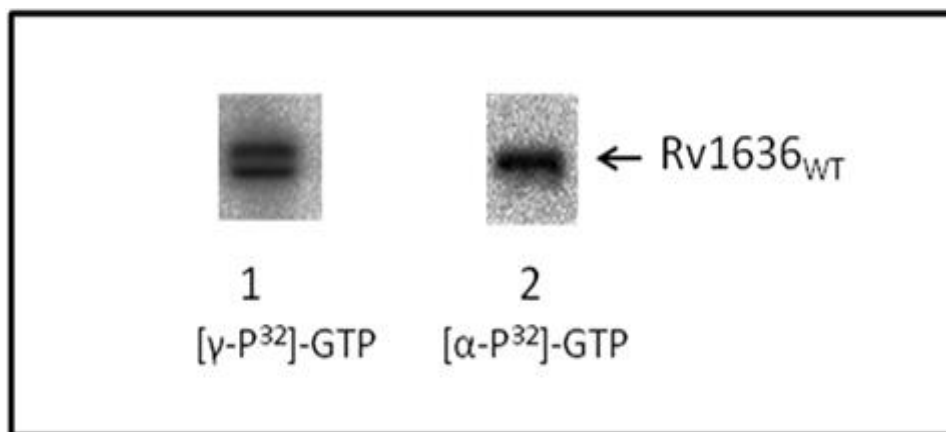


Figure 3.9. Autoradiograph of Rv1636_{WT} following incubation with radiolabelled GTP. The double band suggests different isoforms of Rv1636 as they pass through SDS-PAGE at a different rate. Reducing SDS-PAGE samples were boiled for 5 minutes with 2X Laemmeli.

3.2.2.1.1 Phosphoamino Acid Analysis of Labelled-Rv1636

In order to obtain further information on the kinase activity, phosphoamino acid analysis was performed to identify the nature of the apparent phosphorylated amino acid in Rv1636. [γ -P³²] ATP-incubated Rv1636_{WT} was run on an SDS-PAGE before being transferred to a PVDF membrane. The protein was then hydrolysed with 6 M HCL and the amino acids separated by semi-dry thin layer electrophoresis. When comparing the ninhydrin stain of the amino acid standards to the autoradiograph, shown in Figure 3.10, the labelled phosphate runs to the point equal to the phospho-serine, which suggests that the gamma-phosphate is attached to a serine. Free inorganic phosphate (Pi) was observed to run beyond the phosphoamino acids.

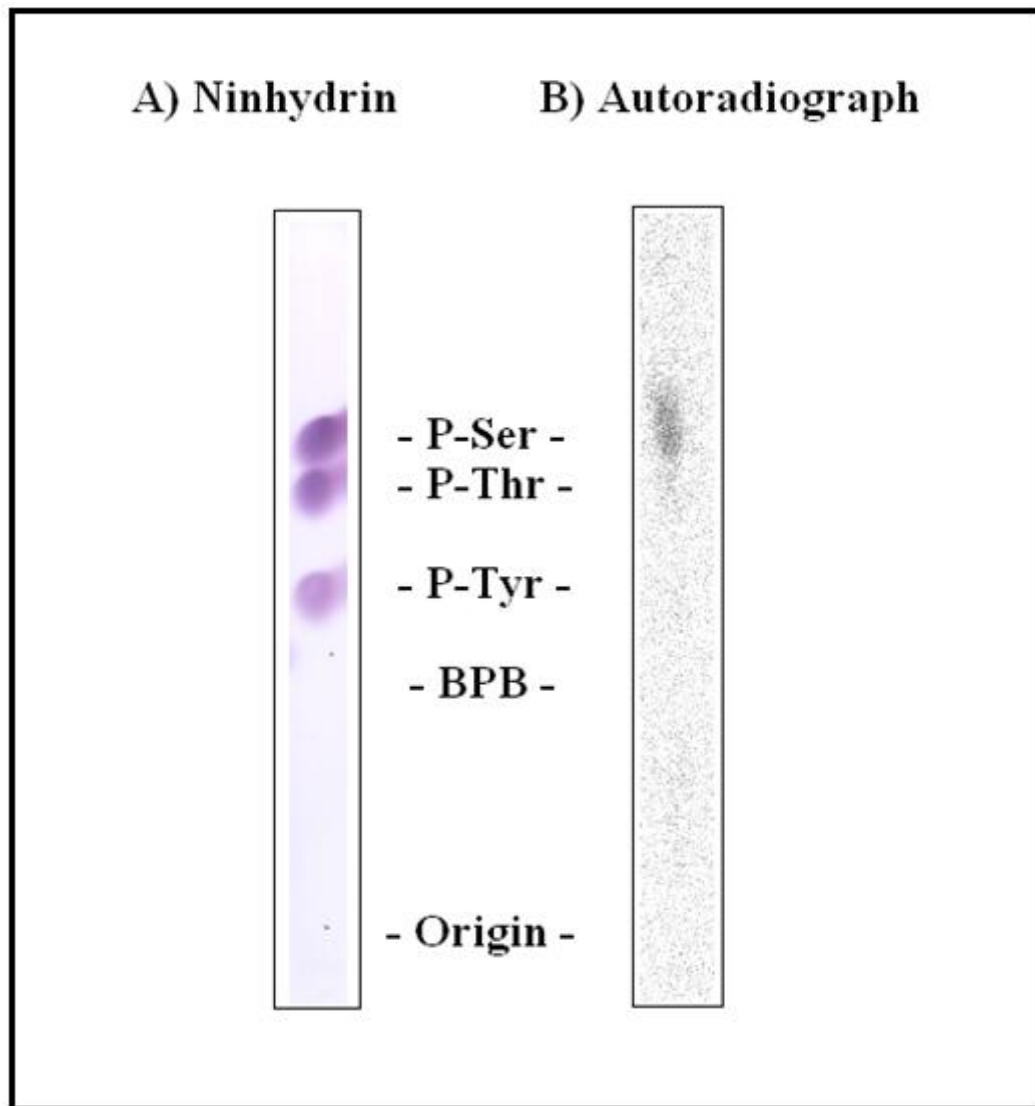


Figure 3.10. Phosphoamino acid analysis by semi-dry electrophoresis on a thin-layer chromatography electrophoretic plate. P^{32} -labelled Rv1636_{WT} was hydrolysed in 6M HCL, spiked with cold phosphoamino acids and electrophoresed. Detection of phosphoamino acids by ninhydrin colour reaction (A) and Fuji BAS 2000 phosphoimaging analysis (B).

As the autophosphorylation of Rv1636 was shown to involve a serine residue, SDM of serine residues was undertaken in order to determine which serine(s) were involved. Rv1636 encodes nine serines and to determine which of these residues may have a role in the phosphorylation activity. Rv1636 was analysed using a phosphorylation site predicting software. Serine residues that could potentially be phosphorylated were highlighted using the program NetPhos (Figure 3.11). NetPhos predicted that the serines at the four sites S16, S17, S57 and S118 were potential sites of phosphorylation. The program looks for serines that are near the surface and are accessible for phosphorylation, although it should be noted it was designed for the analysis of eukaryotic proteins.

The Universal Stress Proteins of Bacteria

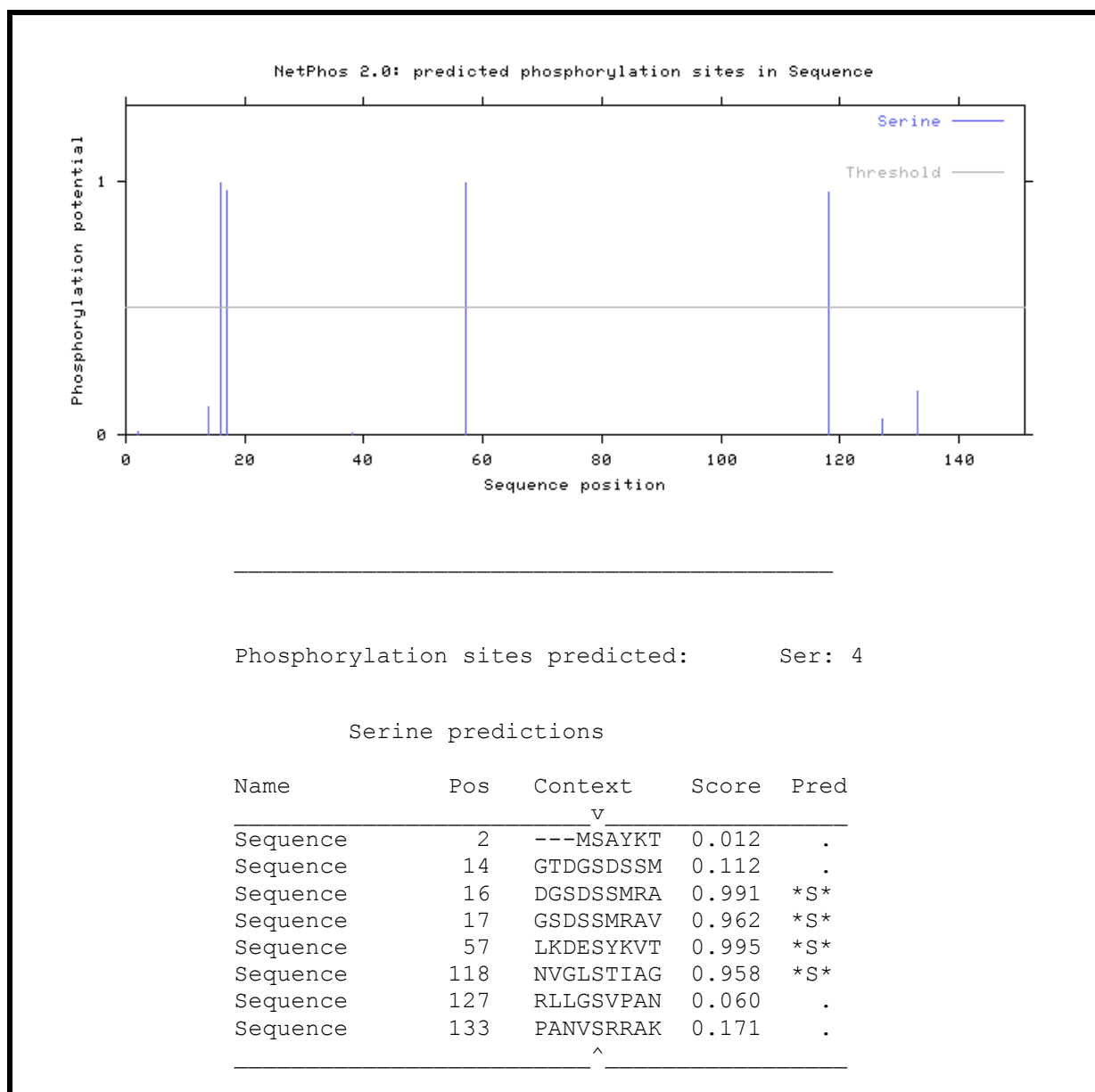


Figure 3.11. NetPhos Prediction of the serine phosphorylation sites in Rv1636. The threshold is set by default. www.cbs.dtu.dk/services/NetPhos/.

Through a combination of the Netphos program and sequence alignment of Rv1636 with homologues from other *Mycobacterium* species (Figure 3.12), four serines were identified. The serines at positions S16A, S17A and S118A were selected for mutagenesis to alanine residues; whereas the construction of a SDM of the serine at position 127 was already described previously (Chapter 1.2.1.2) as it is part of the conserved nucleotide binding motif. NetPhos highlighted S57 as a possible site of phosphorylation, however, sequence alignment showed the residue was not conserved in all the other mycobacterial homologues and therefore, likely not of functional significance (Figure 3.12).

The Universal Stress Proteins of Bacteria

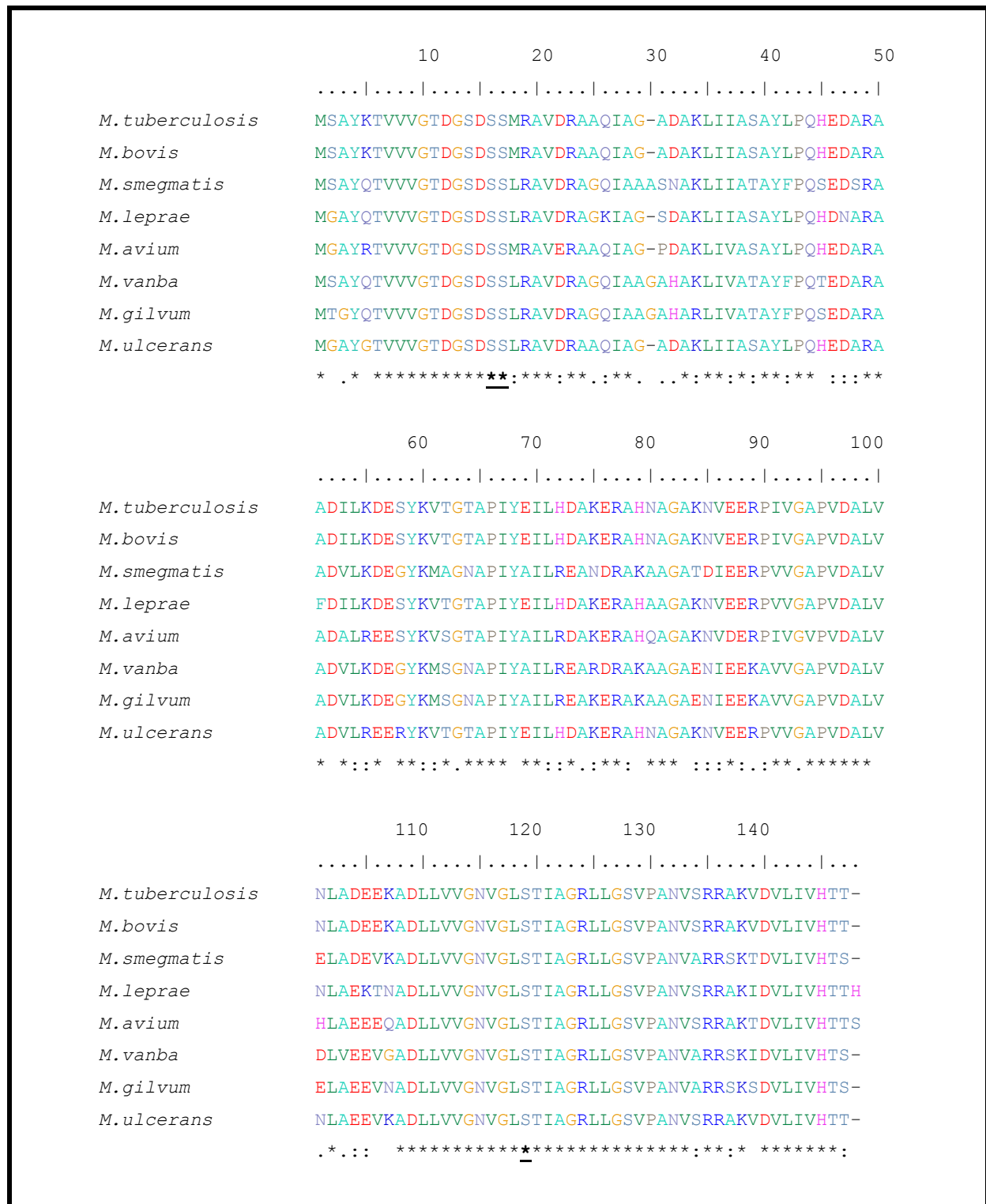


Figure 3.12. Multiple alignment of Rv1636 from *Mtb* with homologues from *M. bovis* (Mb1662), *M. smegmatis* (MSMEG3811), *M. leprae* (ML1390), *M. avium* (MAP1339), *M. vanbaalenii* (Mvan3342), *M. gilvum* (Mflv3551) and *M. ulcerans* (MUL1619). Serines chosen for SDM are highlighted by underlining of the consensus sequence.

The Universal Stress Proteins of Bacteria

PAGE purified primers were used to introduce point mutations into *Rv1636*, using the QuickChange II site-directed mutagenesis kit (Stratagene) according to the supplier's instructions (Chapter 2.3.11). pGEM-Te::*Rv1636* was used as the template. Sequencing confirmed the successful mutagenesis. The mutated *Rv1636* genes were digested from pGem-Teasy::*Rv1636* and were ligated into the expression vector pET15b between the restriction sites *Bam*HI and *Nde*I. The resulting expression vectors pET15b::*Rv1636*_{S16A-S17A} and pET15b::*Rv1636*_{S118A} were then transformed into competent *E. coli* Rosetta cells for expression and IMAC purification (Figure 3.13).

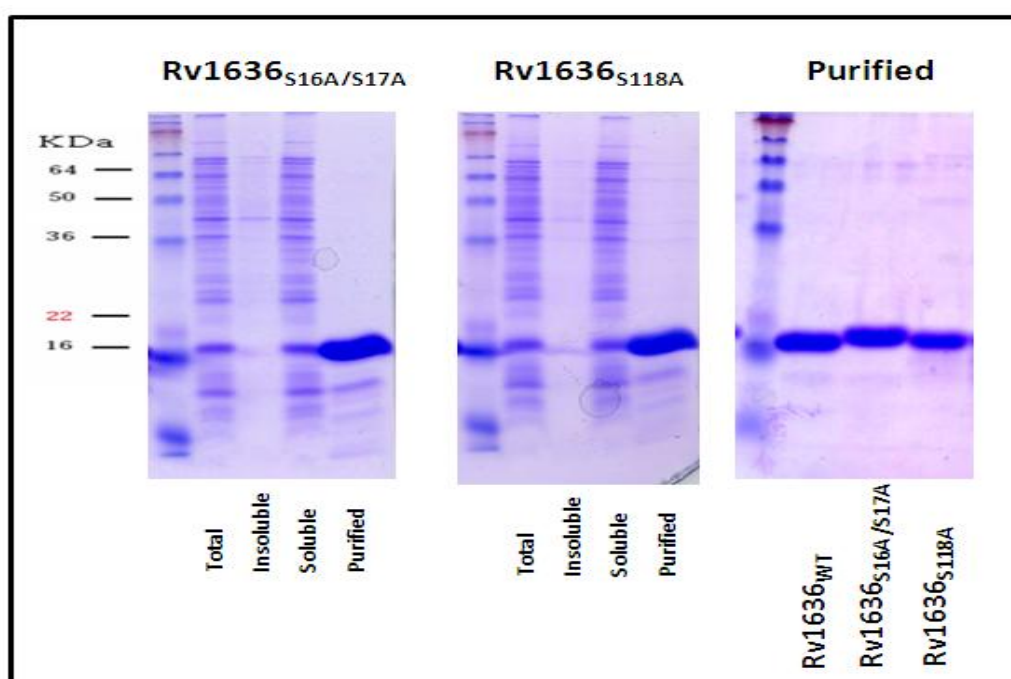


Figure 3.13. Expression and purification of *Rv1636*_{S16A-S17A} and *Rv1636*_{S118A} proteins resolved on 12.5% SDS-PAGE and stained with Imperial Coomassie Stain (Pierce). SeeBlue®Plus2 marker (Invitrogen) is shown.

3.2.2.1.2 Radio-labelling of *Rv1636* was shown to be Non-specific

Subsequent to the creation of the site directed mutant proteins, the *Rv1636*_{WT} was assayed to determine if the protein could phosphorylate other proteins. The myelin basic protein (MBP), which is an 18.5 kDa protein substrate of a wide array of kinases, is similar in size to *Rv1636*_{WT}, therefore a recombinant fusion protein (NEB) with a size of 29 kDa was used. The fusion protein consisted of the MBP attached with a chitin binding domain on the N-terminus.

The Universal Stress Proteins of Bacteria

Rv1636_{WT} was incubated with 5 µg of myelin basic protein (MBP) and phosphoimaging unexpectedly revealed that the MBP became radiolabelled in the absence of Rv1636_{WT} (Figure 3.14A). The potentially non-specific radiolabelling of the MBP was in comparable levels to the radiolabelling of the Rv1636_{WT} indicating the apparent autophosphorylation of the Rv1636_{WT} observed may be biologically insignificant.

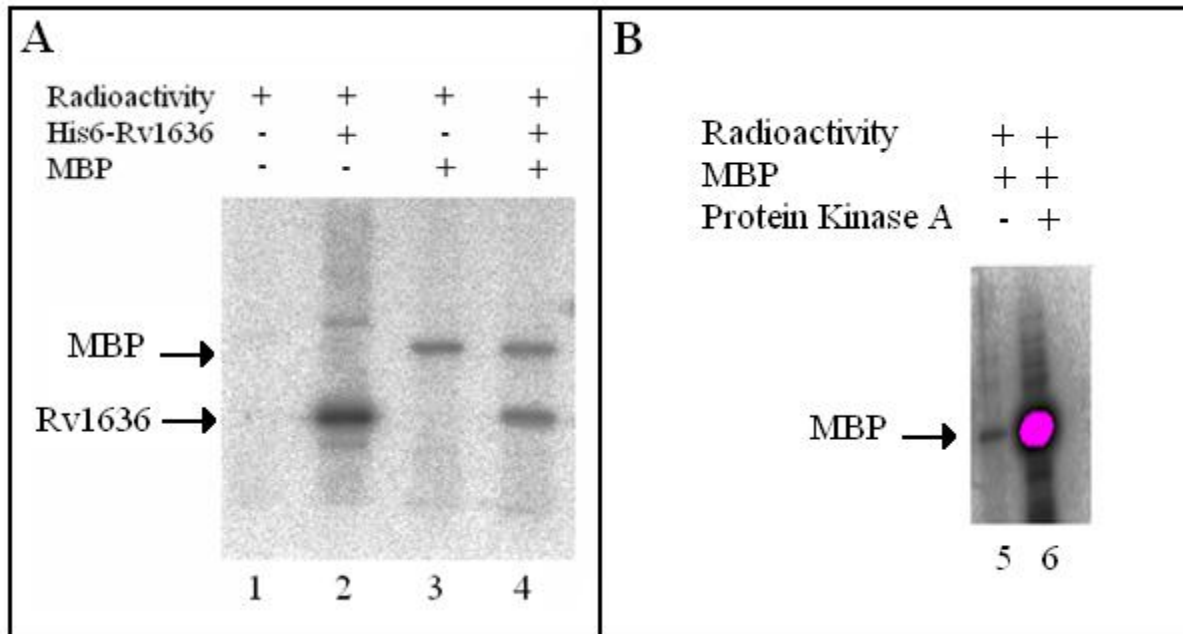


Figure 3.14. Phosphorylation assays with Rv1636 and MBP incubated with [γ - 32 P] ATP. (A) Rv1636_{WT} and MBP were incubated with [γ - 32 P] ATP. Lane 1: No protein, Lane 2: Rv1636_{WT}, Lane 3: MBP, Lane 4: Rv1636_{WT} and MBP. (B) MBP incubated with protein kinase A and [γ - 32 P] ATP. Lane 5: MBP, Lane 6: Protein kinase A and MBP. Reactions were run for 30 minutes at 37°C. Radioactivity was detected using phosphoimaging with Fuji Film 5000 phosphoimager.

Further investigation into the apparent non-specific labelling included performing a variety of trial experiments. This included changing the *in vitro* reaction buffer from a Tris-salt buffer to HEPES buffer, as described by Walter *et al.*, 2000, which showed that although HEPES buffer reduced the background labelling, the MBP remained radiolabelled. MBP protein was also replaced with the histone H1 fraction, another known kinase substrate, and BSA, a stable protein that lacks any effect in many biochemical reactions. The experiments showed the two control proteins also became radiolabelled in the absence of Rv1636_{WT}.

The experiment was then repeated with protein kinase A, which is known to phosphorylate MBP. Upon incubation with protein kinase A, MBP became significantly radiolabelled (Figure 3.14B).

The Universal Stress Proteins of Bacteria

From these experiments it could be concluded that the observed labelling of Rv1636 was non-specific and phosphoamino acid analysis was not performed on the serine mutants created. *In vitro* phosphorylation is not directly an indication of *in vivo* activity. The presence of phospho-serine may suggest that the Rv1636 can be phosphorylated on these residues *in vivo*. Given these results further experiments were not pursued with the serine site directed mutants.

3.2.2.2 Investigation of Autonucleotidylation Activity

Nucleotidylation is defined as the addition of an NMP nucleotide to an amino acid by the creation of a α -phosphodiester bond between the hydroxyl group of the amino residue and the phosphate group of the nucleotide. In order to disrupt the phosphodiester bond and allow analysis of the bound moiety, the organic compound piperidine ($(\text{CH}_2)_5\text{NH}$) was used. Piperidine is a heterocyclic amine that is made up of a six-membered ring containing five methylene units and a single nitrogen atom and which is commonly used in chemical degradation reactions as it is able to cleave particular modified nucleotides. Piperidine is a strong base that can efficiently cleave phosphodiester bonds at both the 3' and 5' sides of bases (Maxam and Gilbert, 1977) through alkaline hydrolysis and which has also previously been used to cleave off an AMP moiety, as described by Weber and Jung (2006) and Usuki *et al.* (1991). The use of piperidine should result in the release of any nucleotide monophosphate bound from the Rv1636 protein.

3.2.2.2.1 α -NTP Labelled Rv1636 Releases a NMP Moiety

To further the analysis, an experiment was performed to determine if the radiolabelled phosphate was attached in the form of a GMP moiety. Piperidine treatment of the $[\alpha\text{-P}^{32}]$ GTP incubated Rv1636_{WT}, followed by thin layered chromatography (TLC), determined the release of a $[\alpha\text{-P}^{32}]$ GMP (Figure 3.15), showing that the reaction of Rv1636_{WT} with $[\alpha\text{-P}^{32}]$ GTP leads to a covalent adduct between the protein and GMP. The experiment was repeated with $[\alpha\text{-}^{32}\text{P}]$ ATP-incubated Rv1636_{WT} which determined that Rv1636_{WT} could also release an AMP moiety.

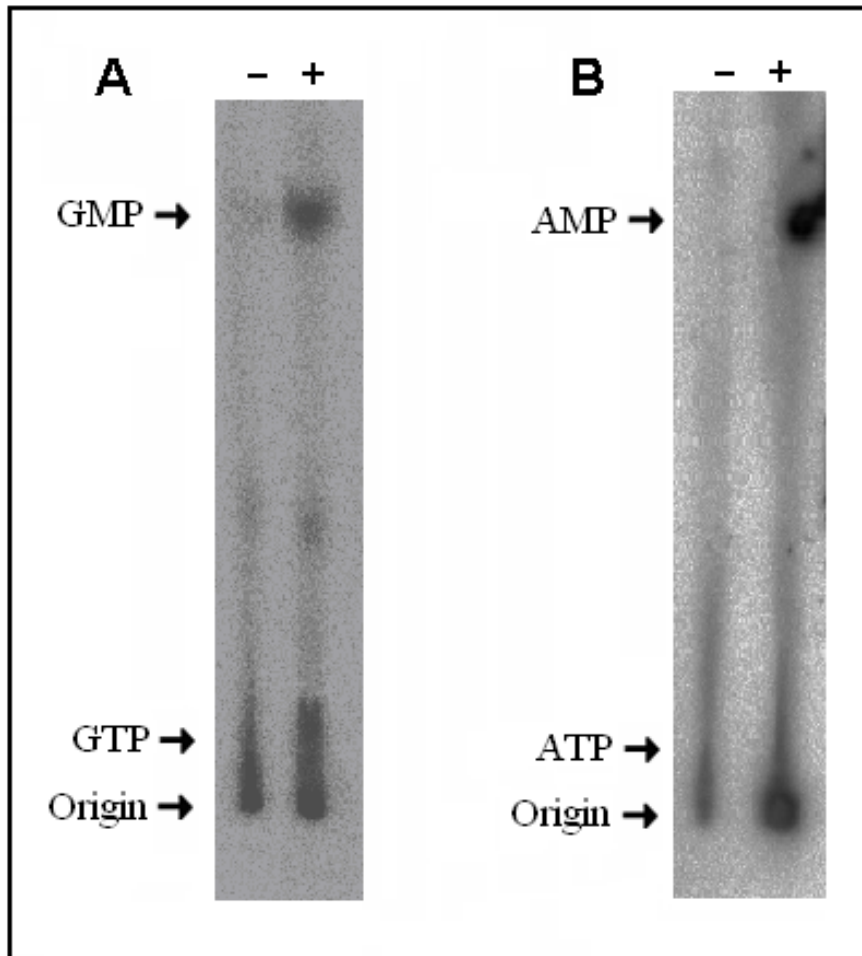


Figure 3.15. Analysis of the nucleotidylylation of Rv1636 using piperidine release. (A) [^{32}P]-GMP release from [$\alpha\text{-}^{32}\text{P}$] GTP-incubated Rv1636_{WT}, following piperidine treatment. Rv1636_{WT} (5 μg) was incubated with 10 μM [$\alpha\text{-}^{32}\text{P}$] GTP (0.5 $\mu\text{Ci}/\mu\text{l}$) for 1 hour at 37 $^{\circ}\text{C}$ before gel purification and acetone precipitation. Piperidine treated (+) and non-treated (-) samples were resolved by TLC. Standards of 20 nmol GMP and GTP were run and were visualised under short-wave UV light to determine the migration patterns. (B) [^{32}P] AMP release when Rv1636_{WT} incubated with [$\alpha\text{-}^{32}\text{P}$] ATP. 20 nmol of AMP and ATP were run as standards.

3.2.2.2.2 Radiolabelled Nucleotide Attaches to Denatured Rv1636

To determine if the Rv1636 mutants were able to be nucleotidylylated, the proteins were incubated with [$\alpha\text{-}^{32}\text{P}$] ATP and subjected to SDS-PAGE analysis. As a further control, Rv1636_{WT} was denatured by boiling in SDS prior to incubation with the ATP. Figure 3.16 shows all the proteins became labelled including the SDS-treated Rv1636_{WT}. As a denatured protein would not be expected to autoadenylylate, it appeared the labelling in the experiment was unspecific and autonucleotidylylation was not occurring to significant levels. In a separate experiment, lysozyme

The Universal Stress Proteins of Bacteria

(14.7 kDa) also becomes labelled when incubated with [α - 32 P] ATP (Figure 3.15). Lysozyme is known not to be nucleotidylated.

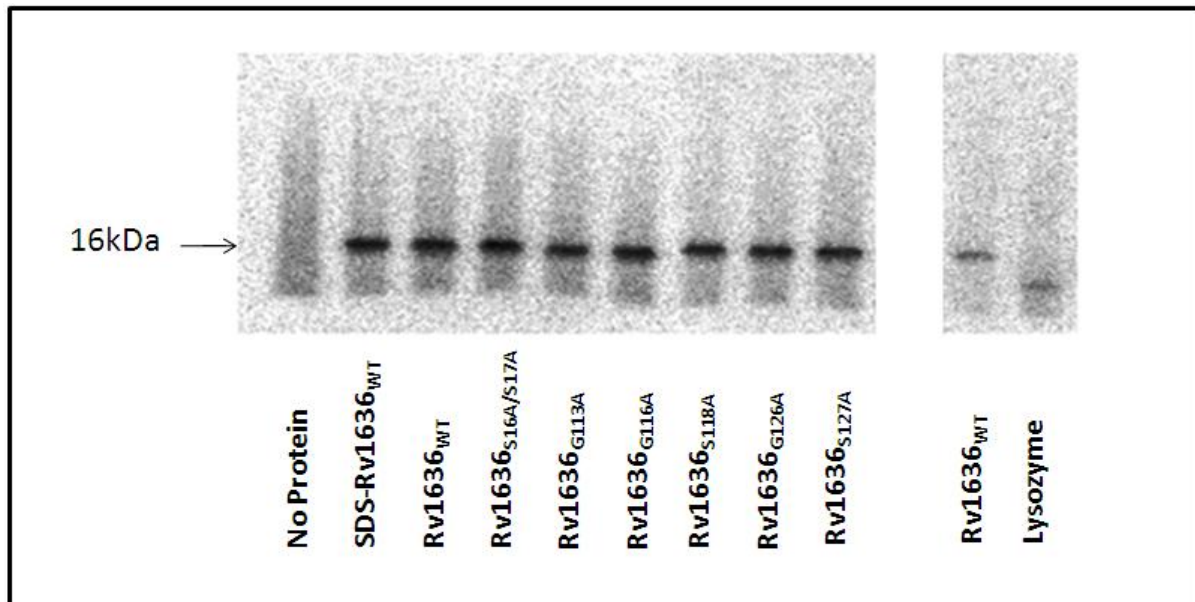


Figure 3.16. *In vitro* nucleotidylation of Rv1636 mutants. Labelled bands runs at the same level as Rv1636_{WT} on SDS-PAGE stained with Coomassie. Rv1636 proteins were incubated with [α - 32 P] ATP and ran on 12.5% SDS Page. Radioactivity was detected using Fuji-film 5000 phosphoimager.

3.2.2.3 Is Rv1636 an RNA Triphosphatase?

Previous data suggested that Rv1636 had RNA triphosphatase activity as the protein was shown to hydrolyse the terminal phosphate from the 5' end of RNA molecules (Leiva, 2007). The property suggests that *in vivo* it may have a role interacting with RNA and that Rv1636 may represent a novel class of prokaryotic RNA triphosphatases, as the activity would be a unique biochemical property if described for a prokaryotic protein.

The noted preliminary evidence which suggested that Rv1636 was a RNA triphosphatase was based on the observed removal of the 5' terminal phosphate from RNA. The original RNA triphosphatase experiments were performed on a low concentration of labelled RNA. However, RNA is very unstable and it was attempted to repeat the experiment using higher concentrations of labelled RNA in order to visualise whether the RNA was still intact. Initial attempts to produce labelled RNA transcript with T4 polynucleotide kinase (PNK) after HK Thermolabile Phosphatase dephosphorylation were unsuccessful and the level of RNA obtained was still too low to visualise.

The Universal Stress Proteins of Bacteria

To further these experiments, 45-nucleotide RNA substrates were transcribed in the presence of [α -P³²] GTP or [γ -P³²] GTP using the AmpliScribe™ T7 RNA polymerase. The incorporation of the alpha label in the generated strand meant there were labelled phosphates throughout the strand, whereas the gamma labelled phosphate would only be found at the 5' terminal end of the RNA. The template used was the pGEM-4Z (Promega) vector that had been linearised with *Sma*I and purified from an agarose gel. Following RNA transcription, the template DNA was digested with DNaseI and free nucleotides were then removed using a Sephadex G-25 column (Amersham Biosciences). The purified 45-mer was [α -P³²] or [γ -P³²] labelled and the quality of the RNA produced was assessed on 10% polyacrylamide gels containing 7M urea.

RNA triphosphatase experiments were then performed on the radiolabelled RNA. As a further control, HK^(TM) Thermolabile Phosphatase was used as it is known to dephosphorylate RNA at the 5' end. Reactions were halted with formaldehyde and EDTA loading buffer as the formic acid previously used was found to inhibit running of the polyacrylamide gels. It was determined that when previous samples of Rv1636_{WT} and Rv1636_{G113A} were mixed with the alpha labelled RNA, the RNA was totally degraded (Figure 3.17). This suggests that Rv1636 is not directly a RNA triphosphatase, and was unable to specifically cleave the 5' terminal phosphate from RNA strands.

It remains to be determined if the breakdown of the RNA is a true catalytic activity of Rv1636. RNase inhibitors were added to the reactions to minimise RNA breakdown by a range of RNases. However, the ribonuclease inhibitor RNasin does not always block all RNase contamination. Upon boiling of Rv1636 in SDS the RNA was no longer degraded, indicating the breakdown was due to either catalytic action of Rv1636 or low level contamination of the purified protein.

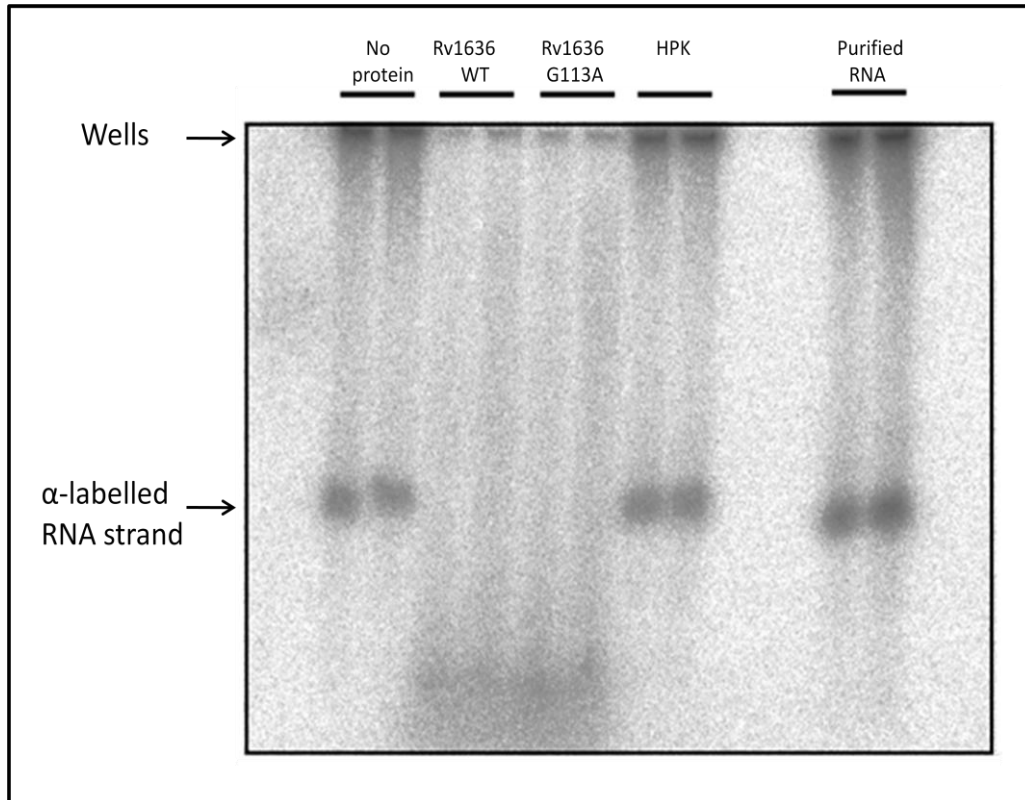


Figure 3.17. RNA triphosphatase control assay. Rv1636_{WT} and Rv1636_{G113A} protein were incubated with [α -P³²] labelled RNA, which was shown to be totally degraded post incubation. 10% polyacrylamide gel with 7 M urea and exposed with the Fuji-film 5000 phosphoimager. HPK is HK^(TM) Thermolabile Phosphatase, a phosphatase that can cleave the 5' phosphate of RNA.

3.2.2.1 Rv1636 Binds Adenosine Triphosphate

To continue the biochemical analysis of Rv1636, the nucleotide binding properties of the protein was assayed. The conserved G-X₂-G-X₉-GS NTP-binding motif is found within the Rv1636 protein and in many other USPs. To determine whether Rv1636_{WT} could bind nucleotide triphosphate the protein was incubated with [α -P³²] ATP in a 20 μ l reaction containing a simple Tris (100 mM) and NaCl (20 mM) buffer. Reactions were kept at 4 °C to ensure further enzymatic reactions did not occur. Two methods were trialled to test for nucleotide triphosphate binding. Firstly, crosslinking of the protein to any bound nucleotide using UV light and secondly, a filter binding method which removes any free nucleotide in the reaction by passing the reaction through a nitrocellulose membrane under vacuum.

The Universal Stress Proteins of Bacteria

3.2.2.1.1 UV Crosslinking Approach

UV crosslinking is an ideal mechanism to study the binding of nucleotides as it introduces covalent binding between the protein and the nucleotide. When a protein initially binds a nucleotide, the interactions are non-covalent interactions, such as hydrophobic interactions, Van der Waals forces and ionic forces. This means that nucleotide binding would not be seen after reducing SDS-PAGE. Irradiation with UV light of nucleotides produces purine and pyrimidine free radicals. Subsequently any protein molecule in close proximity to the free radical can then form a covalent bond, thus crosslinking the nucleotide to the protein.

UV crosslinking of any bound nucleotide to Rv1636 involves exposing the reaction to UV light (Preneta, 2004). When the protein is then subjected to SDS-PAGE, the nucleotide remains attached and can then be assayed by phosphoimaging. In this case, the reactions were carried out in the absence and the presence of cations to determine their role. UV crosslinking determined that Rv1636_{WT} could bind nucleotides (Figure 3.18). The gels showed radioactive bands at the appropriate molecular weight and corresponded to the level of Rv1636_{WT} when electrophoresed gels were stained with Coomassie. When no UV light was applied to the sample there was no evidence of nucleotide binding suggesting no post-translational modifications involving the [α -P³²] phosphate were taking place at 4 °C.

The BSA protein was added as a control and was shown not to remain bound to any nucleotide when subjected to UV crosslinking. This result suggests that the Rv1636_{WT} can bind nucleotide triphosphate and that the addition of cations magnesium and manganese did not result in an increase in binding. Boiling the Rv1636_{WT} in 2% SDS prior to incubation with the [α -P³²] ATP also led to the negation of binding properties as expected.

The Universal Stress Proteins of Bacteria

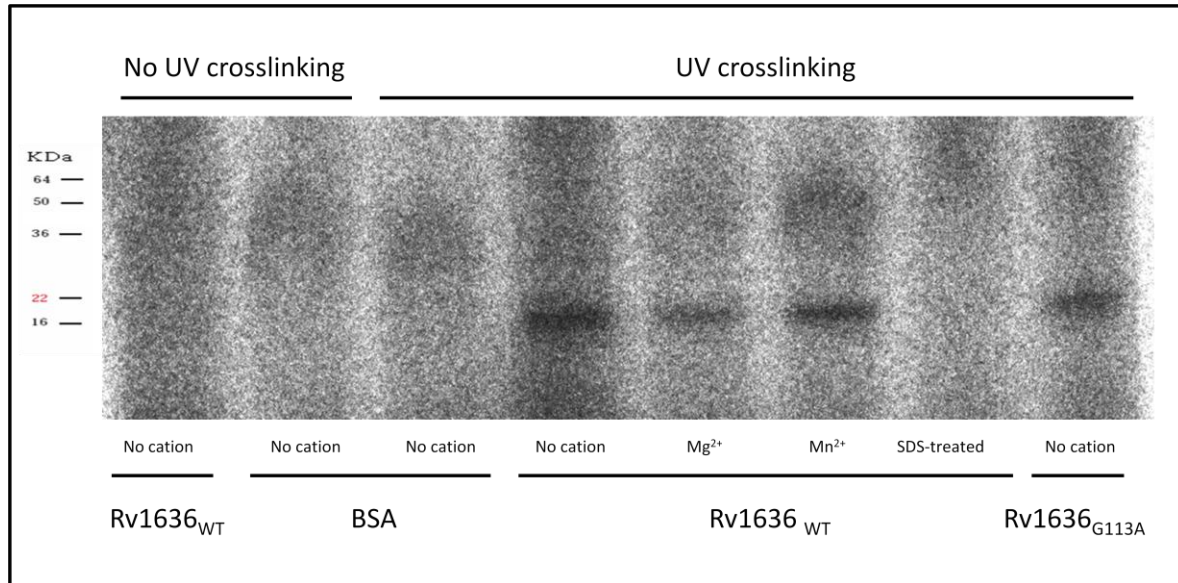


Figure 3.18. UV crossing linking of Rv1636 proteins. Rv1636_{WT} (30 μ M) was incubated with [α -P³²] ATP for 40 minutes at 4 °C under UV light. BSA is Bovine Serum Albumin. Reactions were then subjected to SDS-PAGE and phosphoimaging.

To further analyse the nucleotide binding capabilities of Rv1636, multiple reactions were needed in order to statistically quantify the levels of binding. Attempts to quantify the radiolabel remained difficult and were hindered by a lack of reproducibility of the reactions, due to the variability observed when the similar reactions were cross-linked. To get reproducible data, experiments were performed with a base level of cold nucleotide (30 μ M); however, six replication reactions for the Rv1636_{WT} produced a similar amount of variation as seen in competition experiments with GTP, GDP, GMP, ATP, ADP and AMP, and therefore, a more accurate method to study nucleotide binding was required. In order to continue studying the nucleotide binding capabilities of Rv1636, a separate method of nucleotide binding was developed based on the vacuum filter binding assay.

3.2.2.1.2 Filter Binding Vacuum Approach

Filter binding allows the non-covalent interactions between nucleotides and proteins to be quantified and is a well established method for measuring the binding of nucleotide (Rombel *et al* 1999, Tameling *et al.*, 2002). The proteins with any bound nucleotide remain attached whilst any unbound nucleotide is eluted through a nitrocellulose membrane.

Twenty μ l reaction volumes were passed through a nitrocellulose membrane allowing the proteins to bind due to their hydrophobic nature while free nucleotides pass through the

The Universal Stress Proteins of Bacteria

membrane. The amount of bound radiolabelled nucleotide is then represented by the strength of the radioactive signal remaining. The results of the vacuum filter binding first showed that the blank reaction, containing no protein, when washed with the wash buffer described in Visweswariah *et al.* (2010) of 150 mM NaCl and 10 mM Tris-Cl, gave a positive signal. After trying a variety of wash buffers, it was discovered that replacing the wash buffer with dH₂O alone cleared the background.

Figure 3.19 shows the levels of bound nucleotide associated with the proteins that are bound to the nitrocellulose membrane. Figure 3.19A shows a number of control reactions that indicate the binding of the nucleotide to Rv1636_{WT} is specific and above the background level. The blank (no protein), the BSA and the Rv1636_{WT} boiled in SDS detergent are negative controls and would not be expected to bind any nucleotide. Boiling the Rv1636_{WT} in SDS detergent denatures the protein by disrupting the tertiary and quaternary structure, thus rendering the protein unable to bind nucleotide. BSA was used as a control protein as it is known to be a stable protein with no biochemical activities described. The radiolabelled signal remained with the Rv1636_{WT} protein indicating the protein could bind the radiolabelled nucleotide. Figure 3.19B shows the nucleotide binding capabilities of Rv1636_{WT} and some of the site directed mutants, which suggests that some of the mutated bases may be involved in binding.

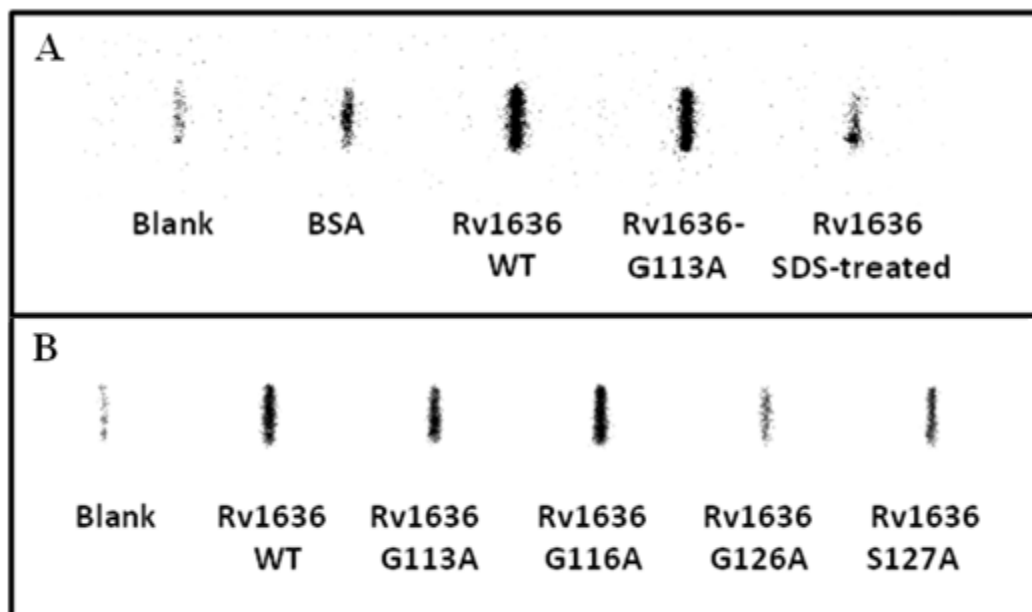


Figure 3.19. Nucleotide binding assay using a filter vacuum. Rv1636_{WT} (30 μ M) was incubated with [α -P³²] ATP for 40 minutes at 4 °C under UV light. Samples were passed through a 0.45 μ M membrane, washed, dried and phosphorimaged to determine levels of radioactive nucleotide bound to the proteins. (A) Control Reactions including BSA and SDS-treated Rv1636_{WT}. (B) Rv1636 and site directed mutants.

The Universal Stress Proteins of Bacteria

The preliminary results of the vacuum binding suggest that the mutant proteins bind ATP (Figure 3.19B) and Rv1636_{G126A} and Rv1636_{S127A} do exhibit lower amounts of bound radioactive nucleotide. Across the slot blot there was again high variation between reactions containing the same protein. An example is shown between the Rv1636_{WT} protein in Figure A and B where the reaction was performed at the same time in the same slot blot and exhibits a high level variation. Further efforts, including changing the wash buffer and doing further repeats to obtain high reproducibility did not improve reproducibility. This prevented further characterisation of the binding properties. Competition experiments would have aimed to show if the protein had a preferred substrate such as NMP, NDP or NTP for the nucleosides guanine, adenine, cytosine and thymine. Also, the effects of cyclic nucleotides and cations would have been of interest; however comparisons were unable to be made due to the high levels of variation encountered.

3.3 Discussion

3.3.1 Purification of Rv1636 Proteins

Rv1636_{WT} and the Rv1636 site directed mutants were successfully expressed as fusions with a small six-histidine tag. The expression produced soluble proteins which were then purified from the *E. coli* cellular extract by one-step ion-metal affinity chromatography (IMAC). The proteins were purified to a high level, with washing with imidazole removing the majority of contaminants that bind the nickel in the column due to unspecific interactions. Once purified, the proteins were stored at -20 °C in 30% glycerol until use in the *in vitro* biochemical assays described. The site directed mutants including S16A/S17A, G113A, G116A, S118A, G126A and S127A were successfully produced.

An important part of SDM is to ensure that the mutations produced have local structural changes and do not affect the native fold of Rv1636. As the site directed mutants of Rv1636_{WT} were found to be soluble, the proteins are likely to be folded when expressed in *E. coli*. A further observation noted in the 89% identical homologue protein ML1390, from *M. leprae*, is that it has an asparagine (N) instead of an alanine in position 84 compared to Rv1636. The change is likely to be a conservative mutation and when comparing mutants could provide a further reference position.

The aspartic acid at position D12 within Rv1636_{WT} is highly conserved in the USPs and the protein D12A was found to be difficult to over-express in *E. coli*. Other approaches that could be used in the future to obtain Rv1636_{D12A} could include purifying the protein within inclusion bodies or optimising the protein yield through *in vitro* refolding methods, the use of other tag fusions systems, such as GST, Maltose binding protein (MBP) or expressing the protein along with chaperones which aid in the proper folding of the proteins (Villaverde & Carrio, 2003). Difficulty in expressing USPs in *E. coli* has been found before in the Williams group.

In the Williams group, the expression of other several other *Mtb* USPs including Rv1996, Rv2005c, Rv2026c, Rv2028c, Rv2623 and Rv2624c in *E. coli* resulted in a large amount of insoluble protein accumulation and no soluble protein, thus making further analysis difficult. The size of the native Rv1636 in comparison to these USPs may have contributed to the solubility of this protein. The six USPs noted are tandem proteins and may have associated folding issues that do not apply to Rv1636. The cysteine content of the tandem USPs proteins may also play a role with the formation of unwanted cysteine intermolecular disulphide bridges as each tandem USP has between two and four cysteines, compared to no cysteines found in Rv1636. Zanier *et al.* (2007) described how, although cysteine bridges have a role in stability of tertiary/quaternary structure, any unspecific

The Universal Stress Proteins of Bacteria

cross-linking of the residues could result in the mis-folding and aggregation of recombinant proteins in *E. coli* during the purification process.

Any aggregated proteins are often compartmentalised by cells, within insoluble inclusion bodies, making further work with the proteins difficult. As an alternative the proteins could be purified from a native environment such as from *Mtb*. A category level 2 *Mtb* strain has recently been obtained by the laboratory and could be used to further study the USPs and Rv1636.

3.3.2 Biochemical Activities of Rv1636

3.3.2.1 Post-translational Modifications of Rv1636

3.3.2.1.1 Investigation of Autokinase Properties of Rv1636

The transfer of phosphate groups from nucleotide triphosphates to amino acids have important functions in biological pathways and can be used to modulate protein activity. Phosphorylation can be used as a molecular switch, as is common in two component systems, and can be involved in both signalling and metabolism networks. Phosphorylation is defined as the addition of a phosphate group (PO_4^-). When added to the R group of an amino acid it can make the hydrophobic nature change to a polar hydrophilic nature. These changes are often important for protein-protein interactions, often due to the conformational changes that can often result from phosphorylation. Phosphorylation is widely used in signal transduction networks that allow the bacteria to sense environmental signals and implement adaptive changes.

Determination of a double band of Rv1636_{WT} by SDS-PAGE after incubation with [γ - ^{32}P] GTP suggested there are two forms of Rv1636 bound to the gamma phosphate. A phosphate group (PO_4^{3-}) has a molar mass of 94.973 Da and once bound can prevent SDS molecules binding, thus giving the phosphate bound form a difference in charge, resulting in slower migration on SDS-PAGE. An example of a protein that migrates differently on SDS-PAGE after a phosphorylation reaction is the C-Jun protein. When phosphorylated at serine 63/73 the migration changes and there is a 39 to 43KDa jump. However, when the C-Jun protein is phosphorylated at 3' terminal sites no shift is observed (Pulverer *et al.*, 1991). Steinberg *et al.*, 2003, also showed the catalytic (Co) subunit of Cyclic AMP-Dependent Protein Kinase could be phosphorylated in two forms and that the two phosphorylated forms migrate differently when separated by SDS-PAGE.

To further the analysis of the putative kinase activity, phosphoamino acid analysis was undertaken. This method was used to determine whether autophosphorylation occurs at serine,

The Universal Stress Proteins of Bacteria

threonine or tyrosine residues. These amino acids contain a hydroxyl group in their side chains which is necessary for phosphate binding. At physiological pH these hydroxyl groups are deprotonated, meaning an oxyanion can act as a nucleophile and attack the phosphate from nucleotide triphosphates. The hydroxyl groups on serine, threonine, and tyrosine are neutral, and when phosphorylated the neutral charge is replaced by negatively charged phosphates.

The result of the phosphoamino acid analysis indicated the radiolabelled gamma-phosphate was attached to the Rv1636_{WT} at a serine residue. Similarly the *E. coli* UspA protein has been shown to exhibit autophosphorylation activity on serine/threonine residues (Freestone *et al.*, 1997). Rv1636 contains nine serine residues and through a combination of the NetPhos program and sequence analysis of other *Mycobacterium* homologues, four serine residues were highlighted as possible candidates for the phosphorylation site(s). One limitation of the NetPhos is that it is specifically designed for eukaryotic proteins, where serine/threonine and tyrosine residues are the usual sites of phosphorylation. Although many prokaryotic proteins are phosphorylated on histidine residues, Boitel *et al.*, 2003 described that phosphorylation of serine/threonine and tyrosine residues has an increasingly evident role in prokaryotic biochemistry.

The successful creation and purification of the serine SDM mutants was carried out with the aim of further elucidating apparent autophosphorylation. Site directed mutagenesis was successful in mutating the four serine residues identified as possible sites of phosphorylation, including S16, S17, S118 and S127. The double mutant at positions S16 and S17 was produced because individually the primers for the mutagenesis had strong secondary structures. However, before phosphoamino acid analysis on the Rv1636 serine mutants was undertaken, the autophosphorylation of Rv1636 was unexpectedly shown to occur at a low level.

The analyses using the MBP fusion protein as a substrate protein showed the Rv1636 protein did not possess any significant kinase activity. MBP is routinely used as a substrate for phosphorylation assays because the protein has no autokinase activity. The MBP alone became labelled in the autophosphorylation assay and therefore put an uncertainty to whether the radioactivity observed on Rv1636_{WT} is biologically relevant. A recent screen mass-spectrometry based by Prsic *et al.* (2010) did not highlight Rv1636 as being a phosphorylated protein when the *Mtb* was screened for phosphoproteins. However, the list is not exhaustive and gene regulation and stage of growth are important factors to consider.

The previous phosphoamino acid analysis, which showed phosphoserine was present and not phosphothreonine or phosphotyrosine, suggested the protein did become labelled with some degree of specificity. Interestingly, Freestone *et al.* (1997) showed that although most UspA could become phosphorylated *in vivo*, the phosphorylation was likely to be dependent on some unknown

The Universal Stress Proteins of Bacteria

factor as only 1:10000 molecules of UspA were phosphorylated in their *in vitro* assay. Indeed, the *in vivo* phosphorylation of *E. coli* UspA, was shown to be dependent on protein TypA, a ribosomal tyrosine phosphoprotein (Freestone *et al.*, 1998). Significant levels of Rv1636 phosphorylation may therefore require the presence of another, as yet unknown, factor. Chapter 5 describes experiments to search for interacting partners of Rv1636.

If the Rv1636 protein was shown to be a serine kinase the protein would be added to the number of serine/threonine kinase already identified in *Mtb*. Within the *Mtb* genome, 11 serine/threonine protein kinases (STPKs) are encoded. The number of these kinases varies between bacterial species and Pérez *et al.*, 2008, suggested the presence of the Ser/Thr or Tyr kinases may correspond with the complexity of the bacterial lifecycle, and that these kinases are required to allow regulation of essential processes allowing the bacteria to adapt to a variety of stresses and promote development and virulence. Future work could involve direct kinase assays, such as those by Mukhopadhyay *et al.* (1999), on immunoprecipitated Rv1636 from *Mtb*, to determine if any proteins that co-immunoprecipitates with Rv1636 can allow the phosphorylation of the protein.

3.3.2.1.2 Putative Nucleotidylylation of Rv1636

Nucleotidylylation is important in signalling recognition and regulation, which are two processes likely to be related to the survival of the bacteria under stress. Nucleotidylylation has been shown to function as a regulatory mechanism for glutamine synthase in *E. coli* and often post translation modifications allow multifunctional regulation. Other known proteins that have both autophosphorylation and autonucleotidylylation properties include the MAFP protein [major acidic fibroblast growth factor (aFGF) stimulated phosphoprotein], which is a glycoprotein from bovine liver (Oda *et al.*, 1993). Interestingly, in the MAFP protein the modifications both occur at the same threonine residue.

The observation that Rv1636_{WT} was also labelled when incubated with [α -P³²] ATP/GTP following denaturation in SDS suggested that all of the labelling was not due to enzymatic activity. The SDS-treated Rv1636_{WT} protein was labelled to a similar degree as all of the site directed mutants tested. Weber and Jung (2006) write that the denatured UspG protein in 2% SDS exhibited no enzymatic activities with no significant labelling detected. The authors incubated UspG with 10 μ M [α -P³²] or [γ -P³²] nucleotide triphosphate at 37 °C for 30 minutes and then analysed the bound GMP stability. They found the bound GMP was acid stable (1 M HCl) and alkaline labile (3 M NaOH); which is in contrast to the stability of the nucleotidylylation described by Leiva (2007).

The Universal Stress Proteins of Bacteria

However, it remains a possibility that the proteins can be modified *in vivo*. The proportion of adenylylated His-UspG after purification in *E. coli* was found to be approximately 5% (Weber & Jung, 2006). As this protein was described to have autoadenylylation properties, it is clear that not all of the units underwent autoadenylylation and other factors must be involved. Weber and Jung (2006) explained how the partial adenylylation of the UspG that they observed *in vitro* could be due to the half-life time of the modification.

The release of an [P^{32}] NMP when the [α - P^{32}] NTP incubated Rv1636_{WT} was subjected to piperidine treatment implied an NMP moiety was covalently bound. The finding that NMP moieties are released could suggest that a low level of nucleotidylylation was occurring. Post-translational modifications such as phosphorylation and nucleotidylylation are known to alter the isoelectric point (pI) of the proteins and therefore future work should involve separating the different isoforms of Rv1636 by 2D-PAGE analysis. The anti-Rv1636 antibody could then be used to detect different isoforms of Rv1636 which may be present.

3.3.2.2 RNA Triphosphatase Activity

The Rv1636_{WT} and Rv1636_{G113A} mutant proteins were both shown to degrade the total RNA in the RNA triphosphatase assay. The cleavage of the 5' phosphate was not specific and the backbone of the RNA molecule was also degraded in the reaction. The creation of α - P^{32} labelled RNA and γ - P^{32} labelled allowed the phosphatase activity to be visualised. The use of the positive control HK Thermolabile Phosphatase indicated a 5' RNA triphosphatase would not degrade the total RNA. The degradation of the RNA was shown to be catalytic and could be a result of Rv1636 or a contaminating protein that co-purified with Rv1636 from the *E. coli*. The use of RNasin did not prevent the degradation of the RNA species and it remains a possibility that catalytic action of Rv1636 resulted in the degradation of the RNA. One potential assay that could be performed in the future is the immunoprecipitation of Rv1636 and assays on the native protein. Following immunoprecipitation a phenol chloroform extraction could determine if any nucleic acids, such as RNA/DNA, are bound to the Rv1636 protein. Investigation into the total RNA pool in a knockout or overexpressing strain of *Mtb* could further link Rv1636 to RNA biochemistry.

3.3.2.3 Nucleotide Binding of Rv1636

Rv1636 was shown to bind nucleotide in two different experiments. The first involved crosslinking the protein to any bound nucleotide using UV light. The second was a filter binding assay where the protein was incubated with radioactive nucleotide and any binding could be assayed after the protein had been filtered on to a nitrocellulose membrane. The SDS-treated denatured protein did not show binding implying the binding is a true biological activity of the Rv1636. When Preneta *et al.* (2004) measured ATP binding it is unclear why an incubation step of 10 minutes at 37 °C was used prior to UV crosslinking at 4 °C. The step was omitted in similarity to Van der Wolk (1997) and the more recent, Ostapchuk & Hearing (2008) who showed the Adenovirus Iva2 protein could bind ATP.

The stable protein BSA was also used as a control and was shown not to bind the nucleotide when subjected to UV crosslinking. Both methods have been used in published articles (Van der Wolk, 1997; Ostapchuk & Hearing, 2008; Rombel *et al.*, 1999; Tameling *et al.*, 2002) to determine levels of nucleotide binding; however, it was found that the reproducibility of identical experiments resulted in large variation. Future studies could involve studying the retention time of boiled supernatant of the proteins when passed through a high-performance liquid chromatography (HPLC) system. Different compounds can be compared to standards and it can be determined if the proteins when expressed by *E. coli* are purified with nucleotide bound.

3.3.3 Biochemical Studies: *M. smegmatis* USPs

A recent study published by Nambi *et al.* (2010) has given further insight into the possible biochemical function of USPs. The study of protein acetyltransferases in *M. smegmatis* disclosed further information about the USPs present in *M. smegmatis* and linked the USPs to cyclic adenosine monophosphate (cAMP). cAMP is known to play a role in *Mtb* pathogenesis, with the adenylyl cyclase Rv0386, which produces cAMP from ATP, shown to be required for the successful establishment of an infection in mammalian hosts (Agarwal, *et al.*, 2009). Intracellular cAMP is found at a high level in both pathogenic and non-pathogenic mycobacterial species, and cAMP has a regulatory role in other biological processes in addition to a role in *Mtb* pathogenesis. cAMP is often used in signalling and regulatory pathways in bacteria and higher eukaryotes, where the small molecule can elicit diverse cellular responses. In *M. smegmatis*, where intracellular cAMP can reach

The Universal Stress Proteins of Bacteria

between 100 μ M-1 mM, the levels of cAMP are affected by various stress conditions such as with SDS and acid stress (Dass *et al.*, 2008).

Nambi *et al.* (2010) investigated the proteins MSMEG_5458 and Rv0998, from *M. smegmatis* and *Mtb* respectively, which consist of a cAMP binding domain fused to a domain with similarity to the GNAT family of acetyltransferases. Interestingly, in biochemically characterising these proteins, it was discovered that protein lysine acetylation occurred and a USP was the substrate of the MSMEG_5458 *in vitro* and *in vivo*. The USP was first identified by performing GST pull-down using GST-5458 and further to this the acetylation of the epsilon (ϵ) amino group of a lysine residue in the USP MSMEG_4207 was found to be regulated by cAMP. *In vitro* acetyltransferase assays involved incubation of MSMEG_5458 with acetyl-CoA and MSMEG_4207 USP as a substrate. Western blotting for acetyl-lysine antibodies showed that the acetylation of the USP increased in the presence of cAMP.

Further analysis with tandem mass spectrometry showed the acetylated residue was lysine 104 and the position was confirmed with MSMEG_4207 USP mutant K104R showing no acetylation upon Western blot analysis. Attempts to recreate the study with the *Mtb* protein Rv0998 were undertaken, but an interacting partner to Rv0998 was not found. The Rv0998 did show strict cAMP-dependent acetylation of the *M. smegmatis* USP; however MSMEG_4207 does not have an orthologue in *Mtb*. MSMEG_4207 has 130 amino acids and has a modified binding motif of G-X₂-H-X₉-GS, which encompasses the K104 residue, but there is not a lysine within the conserved nucleotide binding motif region of Rv1636. Rv1636 has 8 lysines positioned at K5, K33, K54, K59, K74, K83, K106 and K137, which is more than the four possessed by MSMEG_4207 at positions K24, K104, K120 and K126. The finding that USP MSMEG_4207 can be acetylated adds a further biochemical characteristic to the known diverse number of modifications attributed to the Mycobacterial USPs.

A recent meeting with Sandhya Visweswariah, responsible for the work on the USP MSMEG_4207, provided some further lines of research possibilities to be undertaken in the future. They highlighted a potential role for Rv1636 and cyclic AMP as the USP homologue in *M. smegmatis*, MSMEG_3811 (see Chapter 5.2.1 - Homologues of Rv1636), was shown to bind to a cAMP binding column when passed through the agarose matrix (Personal communication, Visweswariah, 2010). The potential biochemical activities of Rv1636 in *Mtb* and the 100% identical homologue Mb1662 in *M. bovis* have yet to be identified. With a deeper understanding of the biochemical activities of Rv1636 and USPs in general, it is more likely a specific biochemical function of the USPs will be able to be determined.

3.3.3.1 cAMP in *Mtb*

cAMP is a nucleotide in which the single phosphate group is bonded to two of the ribose hydroxyl groups, thus forming a cyclic or ring structure. cAMP is synthesised from ATP by adenylyl cyclases that are commonly found in bacteria. The genome of *Mtb* H37Rv is predicted to encode 16 adenylyl cyclases genes (Cole *et al.*, 1998) including the gene *Rv0386* which was recently shown to be required for virulence (Agarwal *et al.*, 2009).

cAMP is likely to be an important signalling molecule in *Mtb* as shown by the abundance of adenylyl cyclases genes, the high concentrations cAMP reaches within the bacteria cell and that the cyclic AMP receptor protein Rv3676 is a global regulator also required for virulence (Rickman *et al.* 2005; Stapleton *et al.*, 2010). Rv3676 has adapted to function in high cAMP concentrations and sensing of cAMP is linked to regulation of the key virulence component ESAT-6. Two examples of genes regulated by the Rv3676 are *rpfA* and *whiB1*. *rpfA* encodes a protein that has a suggested role in reviving dormant bacteria (Mukamolova *et al.*, 2002), whereas the gene *whiB1*, is an essential nitric oxide responsive regulator of gene expression (den Hengst & Buttner, 2008; Smith *et al.* 2010).

The importance of cAMP mediated gene regulation was indicated by Gazdik & McDonough (2005) who added cAMP to cultures of *Mtb* and recorded alterations in gene expression. Mycobacterial infection studies with macrophages by Lowrie *et al.* (1975) first showed macrophage intracellular cAMP levels increase when infected with *M. microti* and *M. bovis* BCG. Bai *et al.* (2009) studied the burst of production and highlighted that an increase in production of cAMP resulted in alterations in macrophage signalling and bacterial gene regulation. cAMP is likely to be an important secondary messenger in the *Mtb* infection cycle and the possibility that Rv1636 is involved with the network remains of great interest.

3.3.4 Summary

Rv1636 had previously been described as undergoing autophosphorylation, nucleotidylation and to exhibit a putative role as a RNA triphosphatase. The phosphorylation and nucleotidylation have been determined to be not significantly above the level of an inactivated protein and indeed, post-translational modification of Rv1636 may well be dependent on other factors within the *Mtb* cell. RNA was shown to be degraded when incubated with Rv1636, suggesting the putative RNA triphosphatase activity was not specifically cleaving the 5' phosphate off the RNA.

The Universal Stress Proteins of Bacteria

The successful creation of site directed mutants in key amino acid residues allowed further characterisation of the putative biochemical activities. UV crosslinking and filter binding assays showed evidence of nucleotide binding by Rv1636_{WT} and the potential cAMP binding action of Rv1636 may indicate a key role linking Rv1636 and cAMP in signalling within the bacterial cell. Signalling within a cell is likely to be of significant importance during infection and stresses are known to form cascades of signalling that allow the bacteria to modify their gene regulation and promote continued survival of the bacilli in challenging conditions such as those encountered by the tubercle bacillus *in vivo*.

Chapter 4 - Characterisation of the Putative NTPase Activity of Rv1636

4.1 Introduction

The first USP to be crystallised was the MJ0577 protein from *M. jannaschii* (Zarembinski *et al.*, 1998) and an additional 11 USP structures have since been deposited in the Protein Data Bank. The second USP to be crystallised was *Haemophilus influenzae* UspA protein (1jmv) (Sousa & McKay, 2001), which revealed an asymmetric dimer structure similar to that of *M. jannaschii* MJ0577. The first two structures of USPs suggested the formation of higher oligomeric structures may be a common property of proteins in the UspA superfamily. The oligomeric structure of a protein can often provide an insight into biochemical function. The Structural Classification of Proteins (SCOP) describes USPs as belonging to the alpha and beta protein (α/β) class of adenine nucleotide alpha hydrolase-like proteins. The core of these proteins consists of three layers, $\alpha/\beta/\alpha$, and a parallel beta-sheet of five strands. Figure 4.1 shows a representation of the two early crystallised proteins MJ0577 and UspA. Crystallised MJ0577 was found to contain a bound ATP molecule; however experimental data indicated that the protein could not hydrolyse the ATP (Zarembinski *et al.*, 1998).

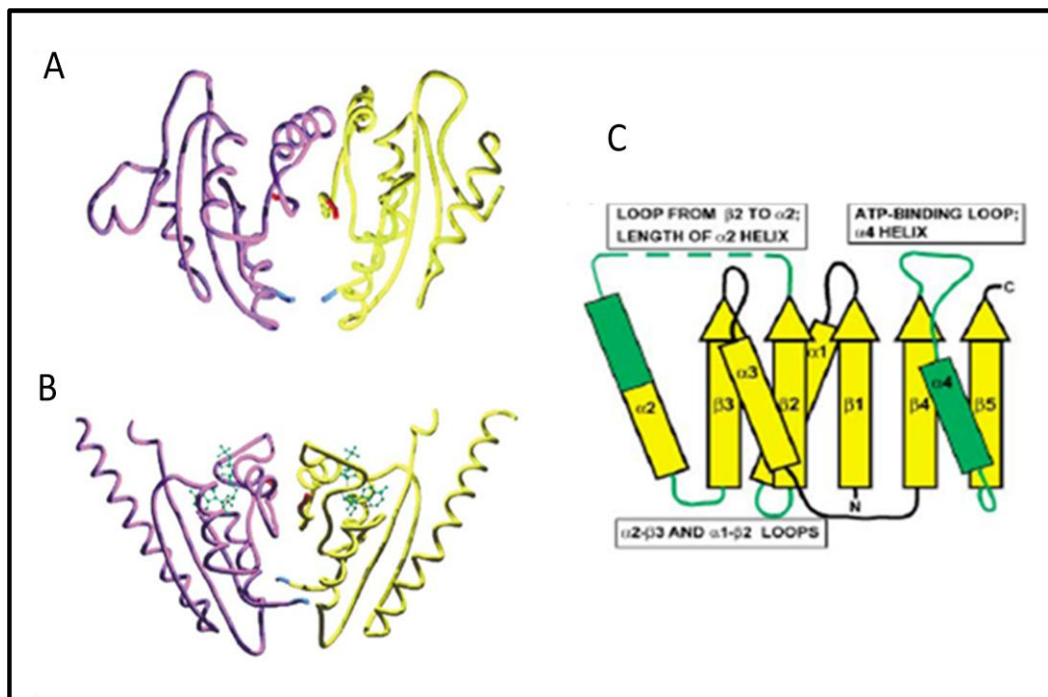


Figure 4.1. The crystal structures of *H. influenzae* UspA (A) and *M. jannaschii* MJ0577 (B) proteins. Monomeric units are shown in yellow and purple. Bound ATPs are shown in MJ0577 and N/C terminals are coloured blue and red respectively. (C) Schematic topology of the Usp proteins showing the α/β folds. Green regions show where UspA and MJ0577 differ. (Figures from Kvint *et al.*, 2003; Sousa & McKay, 2001).

In order to provide further details about the mode of action of Rv1636 the structure of the protein has been investigated. The crystal structure of Rv1636 available in the protein data bank (1TQ8) by Rajashankar *et al.*, 2004, showed a strand swap dimer conformation that appears to block the putative nucleotide binding motif (see Figure 4.7). Due to the nature of the strand swap, it is not known if the deposited structure is the biochemically active state of Rv1636. Of interest is if the Rv1636 can bind nucleotide and subsequently Leiva, 2007 described an NTPase activity for Rv1636 in the absence of cations.

4.1.1 Putative NTPase Activity of Rv1636

Biochemical characterisation of N-terminal His-tagged Rv1636_{WT} has shown that Rv1636 expresses *in vitro* GTPase and ATPase activity (Leiva, 2007). Interestingly, the protein was found to have the very unusual property of exhibiting the NTPase activity only in the absence of divalent cations. The major hydrolysis product of this cation-independent NTPase activity is the nucleotide diphosphate and a phosphate group. The reaction was also shown to be sequential, with nucleotide monophosphates produced.

Most nucleotide triphosphatases belong to the well-characterised class of P-loop NTPases, containing the Walker A [GXXXXGK(S/T)] and Walker B [DXXG] sequence motifs; regions involved in NTP binding and NTP hydrolysis, respectively (Dever *et al.*, 1987; Saraste *et al.*, 1990). The presence of Mg²⁺ is required to complete the hydrolysis of the β - γ phosphate ester bond of the nucleotide (Saraste *et al.*, 1990). Rv1636 does not contain these classical motifs (Walker *et al.*, 1982; Bourne *et al.*, 1991), thus the protein may potentially represent a novel NTP hydrolysis mechanism. Indeed divalent cations were shown to inhibit the NTPase activity of Rv1636.

Rv1636_{WT} was previously shown to hydrolyse nucleotide triphosphates including ATP and GTP. When incubated with [α -³²P] nucleotide triphosphate at 37 °C, the hydrolysis could be visualised when the reaction mixture was separated by thin layer chromatography (TLC) (Figure 4.2). The hydrolysis occurred in the absence of cations and was inhibited when cations were added at a concentration of 5 mM. NTPase kinetic data fits a classic parabolic curve when 10 μ M Rv1636_{WT} was incubated with concentrations of ATP ranging from 1-25 μ M. However, the K_m for the NTPase activity was low at 13.02 $\pm$ 1.4 μ M. It was hypothesised that the proportion of active Rv1636 in the sample was low. Traditionally, kinetic studies to see a saturation curve towards V_{max} are performed with at least 10-fold excess of substrate to enzyme; however, in the case of the Rv1636 experiments, the molar concentrations used were equal.

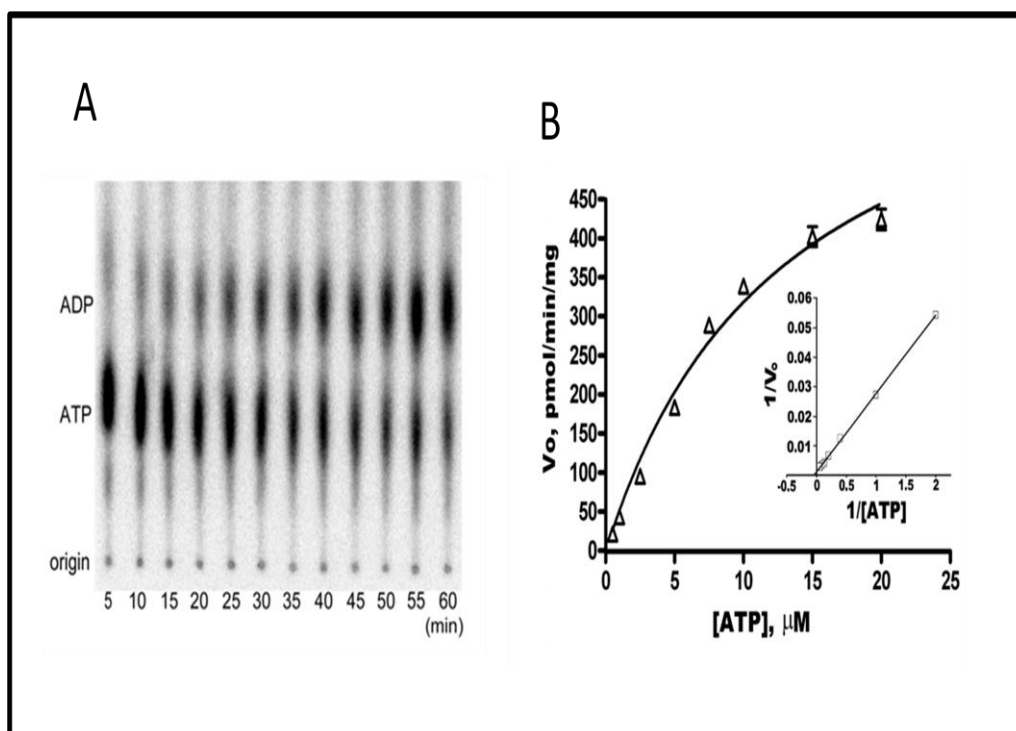


Figure 4.2. Michaelis-Menten kinetics of Rv1636_{WT} ATPase activity over 60 minute reaction at 37 °C. (A) 10 μM Rv1636_{WT} was incubated with 10 μM ATP in 10 mM Tris-HCl (pH 7.5) and 50 mM NaCl. Reactions were separated by TLC. **(B)** Lineweaver-Burk plot of Rv1636_{WT} ATPase activity. The initial velocities, V_o , were calculated by measuring time course of ADP released for fixed concentrations of ATP from 1 to 20 μM . The plot gave an r^2 of 0.98. The K_m and V_{max} constants were calculated to be $13.02 \pm 1.7 \mu M$ and $653 \pm 44.9 \text{ pmol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$, respectively. Figure from Levia (2007).

4.1.2 Structural Methods Used in this Study

In this chapter, the use of Multi-angle Laser Light Scattering (MALLS) and Analytical Ultracentrifugation (AUC) to study the structure Rv1636 is described. A brief introduction to these methods is given below.

4.1.2.1 Multi-Angle Laser Light Scattering

Classical light scattering experiments can be used to provide a direct measure of molecular mass. SEC-MALLS is a method that allows the calculation of molecular mass irrespective of the shape of the particle (Ye, 2006). This allows the native state of a protein to be determined and the

The Universal Stress Proteins of Bacteria

presence of higher oligomeric forms or aggregates detected. The method involves directing light at the protein solution and detecting the amount of scattered light at a range of angles relative to the incident laser beam. Molecular masses determined by light scattering using a single detector are prone to significant error due to the experimental limitations. Multi-angle light scattering detectors measure light scattering from a range of angles and combine the data to give estimates of molecular mass with a vastly increased accuracy.

Multi-angle light scattering is often performed in conjunction with size exclusion chromatography (SEC). The sample proteins pass through the column and any reading of the molecular masses are independent of the elution volume. Samples are injected at concentrations that are inversely related to the molecular mass. With an increased amount of sample loaded the possibility of having large background signal:noise ratio is reduced.

The samples pass through the SEC and have the concentration recorded using UV absorbance. The light-scattering detector receives a signal that is directly proportional to the protein concentration multiplied by the molecular mass of the protein. Through a combination of the light detector and the concentration detector the molecular mass of each peak coming off the column can be accurately calculated. The molecular mass calculated by the SEC-MALLS technique is generally accurate to 3% or better. The calculated molar mass in the SEC-MALLS involves a number of parameters analysed by the software. The more horizontal the line indicating the calculated molar mass, the higher the confidence of the program is to the calculated molar mass.

4.1.2.2 Analytical Ultracentrifugation

Molecular data can be gathered about a protein sample depending on the sedimentation velocity when placed under a centripetal force. Analytical centrifugation is an important technique allowing the study of the oligomerisation state of macromolecules. The procedure can allow the molecular mass to be calculated and, as species of different oligomeric forms sediment differently, the method can be used to analyse the homogeneity of molecules in a solution. Information about the strength of the interactions between molecules can also be calculated. Figure 4.3 shows the system setup.

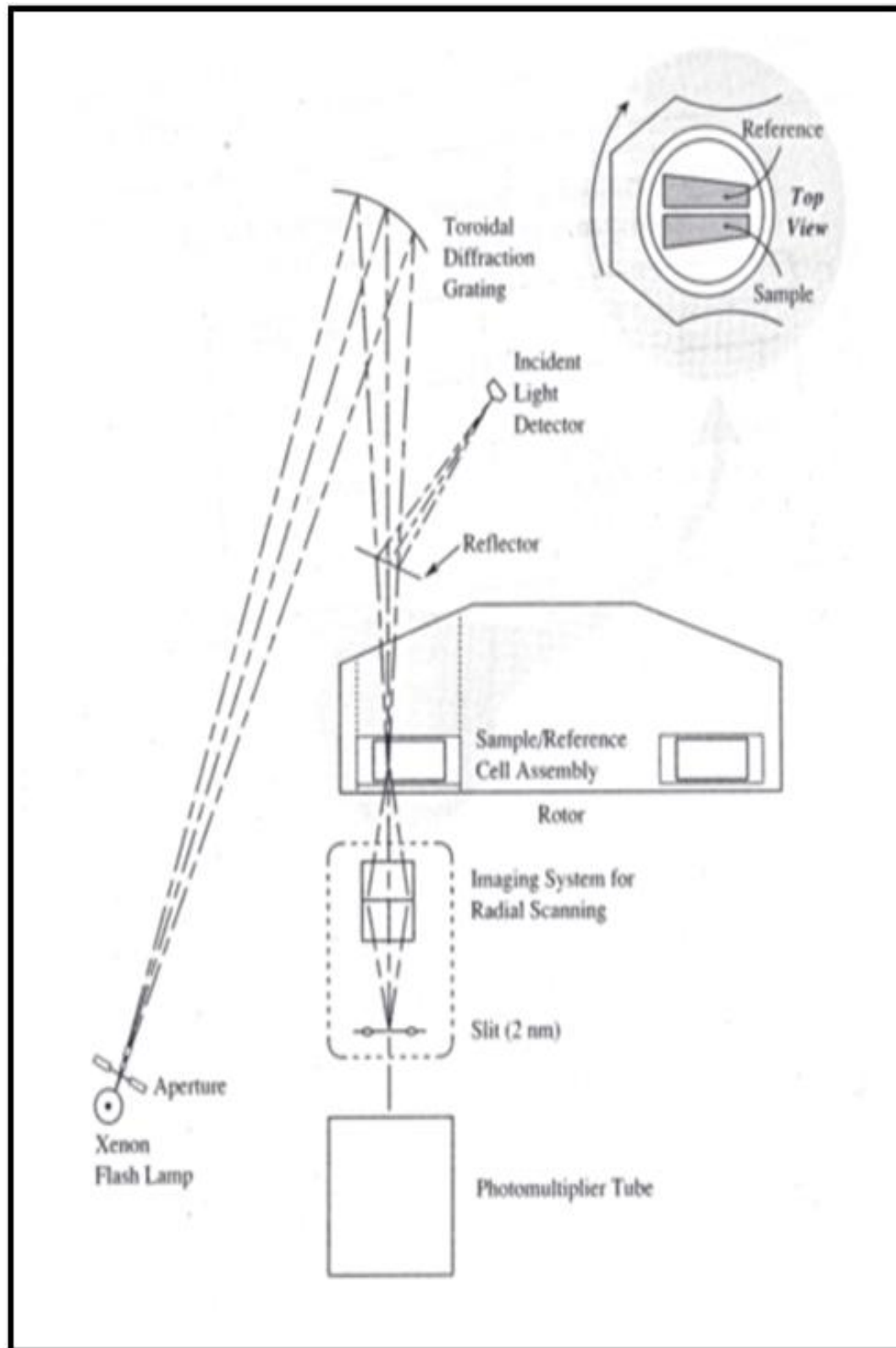


Figure 4.3. Schematic diagram of the Beckman Optima XL-A absorbance system. Diagram from Ralston (1993).

4.1.3 Aims

The NTPase activity of Rv1636 has previously been characterised to determine the effects of temperature, cations and nucleotide competitors (Leiva, 2007). However, an explanation for the low NTPase activity remained to be determined. The kinetic data suggested that a large proportion of the sample was in an inactive form. It was hypothesised that the dimeric form of Rv1636 observed in the deposited crystal structure did not represent the biologically active form. In the crystallised form, the strand swapped structure would lead to the putative nucleotide binding motif being blocked and this could potentially explain the low activity observed. Therefore, it was proposed that disrupting the strand swap would result in an increase in NTPase activity. A number of mutated forms of Rv1636 were designed with the aim of disrupting the strand swap and these were characterised biophysically using a range of techniques to elucidate further details of the oligomeric state of the proteins. The chapter describes the investigations into the NTPase activity and nucleotide binding capabilities of these newly constructed Rv1636 mutants.

4.2 Results

4.2.1 USP Ligand Binding

Investigating folding patterns and tertiary structures of proteins can provide insights into possible cellular functions. The mode of ligand binding, such as adenosine in ATP, is highly linked to the structure of the protein. Table 4.1 lists the 12 crystallised USPs and details the ligand binding and probable quaternary structures according to the PDB database.

The Universal Stress Proteins of Bacteria

Table 4.1. Ligand binding and probable quaternary structure of the 12 crystallised USP structures from the Pfam Data base. *Mtb* Rv2623 is the first tamden domain USP to be crystallised and therefore the homodimeric structure contains 4 individual USP domains.

Organism	Gene name	PDB ID	No. of Chains	Ligands/Cofactors	Probable Quaternary Structure (PDBe)
<i>Klebsiella pneumoniae</i>	<i>kpn_01444</i>	3fdx	2	Mg ion ATP Formic acid (CH ₂ O ₂)	Homodimeric
		3fh0	2	ADP 1,2-Ethanediol	Homodimeric
<i>Mtb</i>	<i>Rv1636</i>	1tq8	6	-	Homotetrameric
<i>Mtb</i>	<i>Rv2623</i>	3cis	8	Mg ion ATP	Homodimeric
		2jax	1	Mg ion Unknown atom or ion	Homodimeric
<i>Archaeoglobus fulgidus</i>	<i>AF_0826</i>	3dlo	4	Cl ion Acetate Ion (C ₂ H ₃ O ₂) Unknown (C ₄ H ₉ NO ₂)	Homotetrameric
<i>Pyrococcus horikoshii</i>	<i>PH0823</i>	2dum	4	-	Homodimeric
<i>Thermus thermophilus</i>	<i>TTHA0895</i>	2z09	1	Mg ion	Homotetrameric
		2z3v	1	ACP S-1,2-Propanediol	Homodimeric
		2z08	1	Mg ion ATP	Homotetrameric
		1wig	1	-	Homodimeric
<i>Nitrosomonas europae</i>	<i>NE1028</i>	2pfs	1	Cl ion	Homodimeric
<i>Lactobacillus plantaru</i>	<i>lp_3663</i>	3fg9	6	Mg ion Formic acid (CH ₂ O ₂)	Homodimeric
<i>Haemophilus influenza</i>	<i>UspA (HI_0815)</i>	1jmv	4	Sulphate (SO ₄)	Homodimeric Homotetrameric
<i>Arabidopsis thaliana</i>	<i>AT3G01520</i>	2gm3	6	AMP	Homodimeric
<i>Aquifex aeolicu</i>	<i>USP Aq_178</i>	1q77	2	Sulphate (SO ₄)	Homodimeric Homotetrameric
<i>Methanococcus jannaschii</i>	<i>MJ0577</i>	1mjh	2	ATP Mn Ion	Homodimeric

ACP: Phosphomethylphosphonic acid adenylate ester (C₁₁H₁₈N₅O₁₂P₃)

Mg ion: Magnesium ion.

Cl ion: Chloride ion.

Mn ion: Manganese ion.

4.2.2 Analysis of the Crystal Structure of Rv1636

Six chains of Rv1636 make up the crystal structure of Rv1636 deposited in the Protein Data Bank. The schematic in Figure 4.4 shows that the order is a homotetrameric structure with a dimer attached. The structure consists of a trio of dimers with A-B, C-D and E-F having 27, 27 and 26 respective H-bonds between the two units. The dimers are then bound to each other with A-C and B-D having 2 H-bonds between them. These bonds/interactions described are the likely biologically significant ones rather than the weak interactions between units F and C, which maybe an artefact of crystallisation. Table 4.2 shows the number of bonds/contacts and the number of residues interface between the different monomeric units.

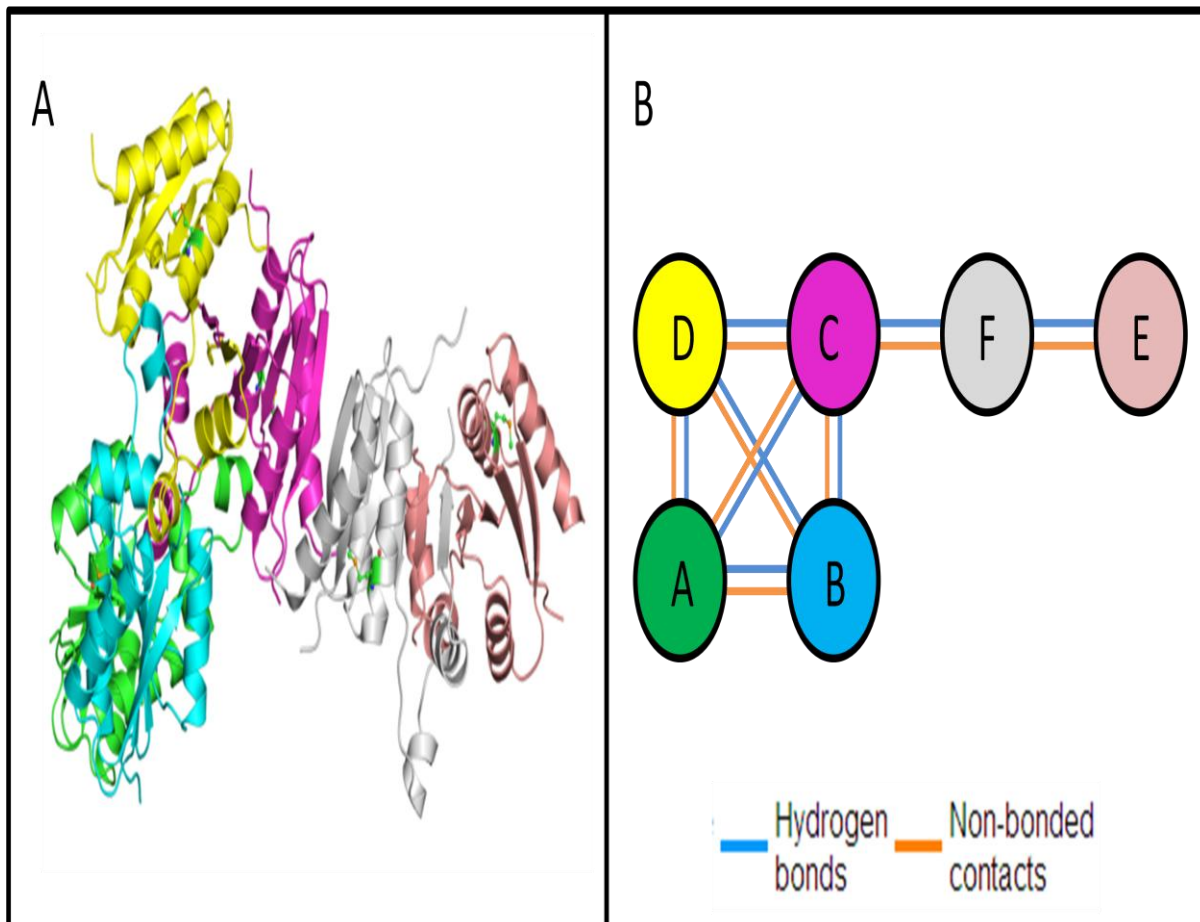


Figure 4.4. Deposited crystal structure of Rv1636. Shows Rv1636 forms a homotetrameric structure consisting of 4 subunits of 1 distinct polymer. (A) Crystal structure. (B) Schematic structure. Structure was analysed using PDBsum.

The Universal Stress Proteins of Bacteria

Table 4.2. Interface statistics of Rv1636 deposited crystal structure

Chains	No. of interface residues	Interface area	No. of H-bonds	No. of non-bonded contacts
A=B	34:37	2094:2081	27	248
C=D	36:36	2126:2110	27	236
E=F	36:36	2082:2090	26	233
A=C	19:21	1143:1133	2	52
B=D	20:21	1142:1158	3	59
A=D	9:9	411:402	2	44
B=C	9:9	407:407	2	47
C=F	15:17	862:856	5	101

It appears the monomeric units come together to form a tightly associated dimer. The dimers are held tightly together by numerous H-bonds and these dimeric units form tetramers with non-bonded contacts. Figure 4.5A shows the secondary structure of the Rv1636 protein and Figure 4.5B shows possible tetramer conformation as predicated by the PDB sum database. Rv1636 was not shown to be crystallised with any ligands.

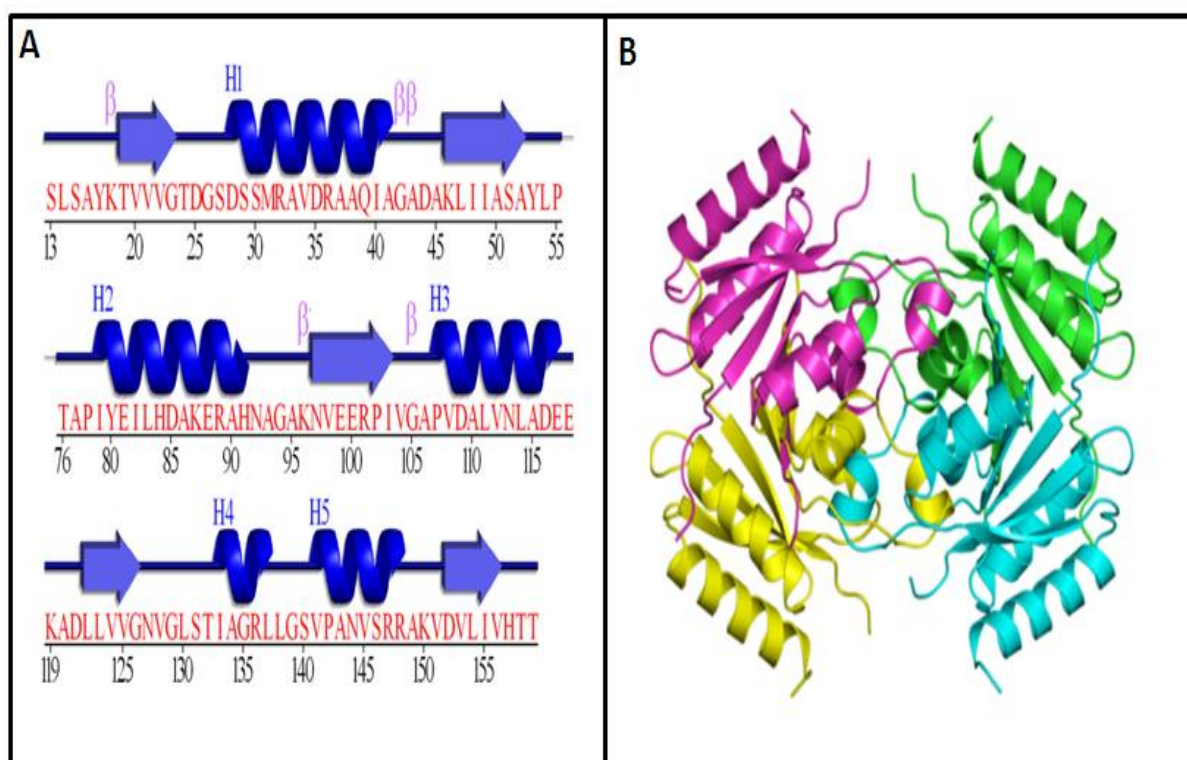


Figure 4.5. The structure of Rv1636. (A) Sequence and secondary structure alignment. H1-5 are helices. Arrows represent beta sheets, β = beta turn. (B) Ribbon diagram of the PDB sum probable quaternary state.

4.2.3 Classification of Rv1636

Previously, Zarembinski *et al.*, 1998, suggested the USPs can be subdivided into those which bind ATP and those that do not. These subdivisions were based on the first two structures of the USP group, MJ0577 and UspA from *H. influenza*. Studying the alignments of the USPs, it was observed that there are five regions of homology, shown in Figure 4.6. It is these regions that are used to classify USPs and are responsible for the USPs having a shared fold. Rv1636 is similar to the MJ0577 protein as it possesses the binding motif as highlighted in conserved region 4.

The Universal Stress Proteins of Bacteria

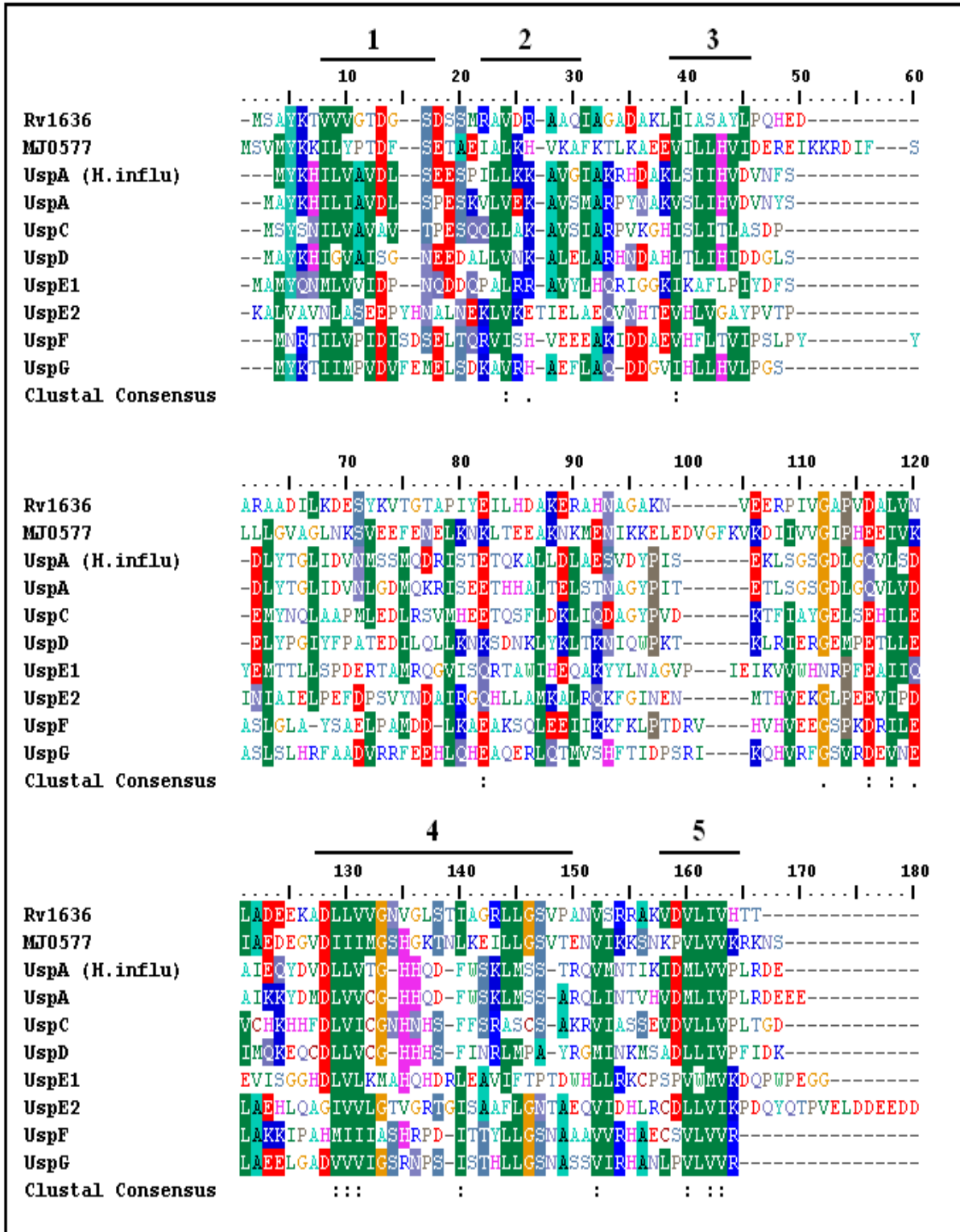


Figure 4.6. Alignment of Rv1636 with the MJ0577, UspA from *H. influenza* and the USPs of *E. coli*. The consensus sequence is shown and conserved regions are highlighted 1-5. Alignment made with Bioedit.

The Universal Stress Proteins of Bacteria

The USPs of *E. coli* were initially classed into two subgroups: Class 1 (UspACD) and Class II (UspFG) as they contain some identifiable differences in the structure. Of the *E. coli* USPs, it is difficult to assign the closest homologue to Rv1636 as most have only around 25% shared identity. When the Rv1636 sequence is compared against *E. coli* non-redundant protein sequences, the top hits are UspG with 28% shared homology and UspF with 27% homology. Of the *E. coli* USPs, only UspFG contain the conserved GS in the nucleotide binding motif, G-X₂-G-X₉-GS. When compared with UspA, which has no G-X₂-G-X₉-GS motif, the protein shares 25% similarity.

Nachin *et al.*, 2008 subdivided the *E. coli* USPs based on their structure in relation to their close homologues in a select few other species. A conserved aspartic acid residue at position D139 was suggested to be important in forming salt bridges in the UspACD subgroup. USP Rv1636 also contains this conserved aspartic acid residue at position D139. Comparing the USPs of *E. coli* to MJ0577, UspA from *H. influenza* and Rv1636, it appears that Rv1636 is sufficiently different for it to be unclear if Rv1636 fits into one of these distinct groups.

4.2.4 Rv1636 is a Strand Swapped Dimer

Inspection of the Rv1636 crystal structure indicated there was a strand-swap dimer conformation between individual dimeric units (A=B, C=D or E=F of Figure 4.4), with strands from two monomers interlocking between the two units. The strand swap can be seen to be in close proximity to the highlighted residues of the nucleotide binding pocket meaning that, in this state, the binding of nucleotides would be reduced or completely blocked as shown in Figure 4.7 of one dimeric unit. It was decided to test the hypothesis that the strand swap affects NTPase activity. In order to disrupt the strand swapped dimer, the sequence and structure of Rv1636 was analysed using molecular graphics program PDB viewer. Disrupting the H-bonding was proposed as a way to destabilise the strand swapped dimer and therefore free access to the putative nucleotide binding motif. Studying the structure with PDB viewer it can be seen that Tyr4 makes a hydrogen bond with Asp139 to stabilise the strand swap.

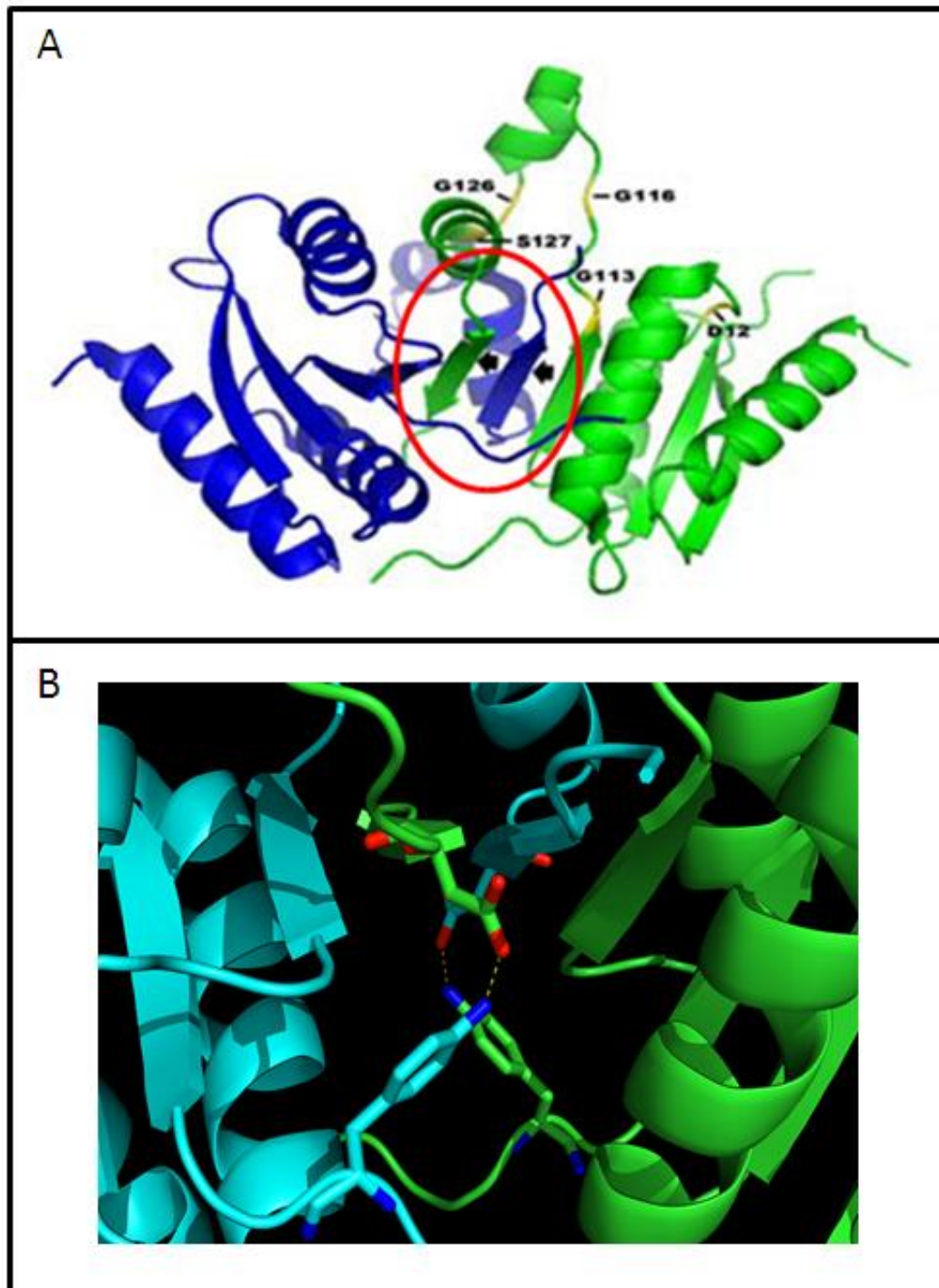


Figure 4.7. Strand swap of the Rv1636 protein. (A) Crystal Structure of Rv1636 from Protein Database (PDB:ITQ8). Strand swap is highlighted in red circle. Dimeric strand swapped form is shown with one unit in green and one in blue. (B) Magnified strand swap of the Rv1636 dimer. Dashed yellow lines signify the H-bonds between the residues of different monomers.

4.2.4.1 Disrupting the Strand Swapped Rv1636

Constructs were designed to disrupt the strand swap. The amino acid sequence of the Rv1636 construct used to produce the crystals for diffraction analysis was analysed. It was observed that the crystallised Rv1636 had extra amino acids added prior to the N-terminal histidine tag and also the additional amino acids EGGS added to the C-terminus, due to the choice of expression vector.

The truncated constructs were designed to include the central 4-144 amino acid residues with a C-terminal His-tag. The histidine tag replaced the additional EGGS residues at the C-terminal, and the methionine, serine and alanine before the tyrosine at position 4 were removed from the N-terminal end. The new construct was constructed to start MYKT as shown in Figure 4.8, including a methionine at the start to allow expression in *E. coli*. The new histidine tag at the C-terminus involved replacing the terminal two threonine (TT) residues with six histidines, forming a small seven-histidine tag to allow purification. The TT was shown not to be involved in any interactions when the structure was analysed, justifying their removal.

The Universal Stress Proteins of Bacteria

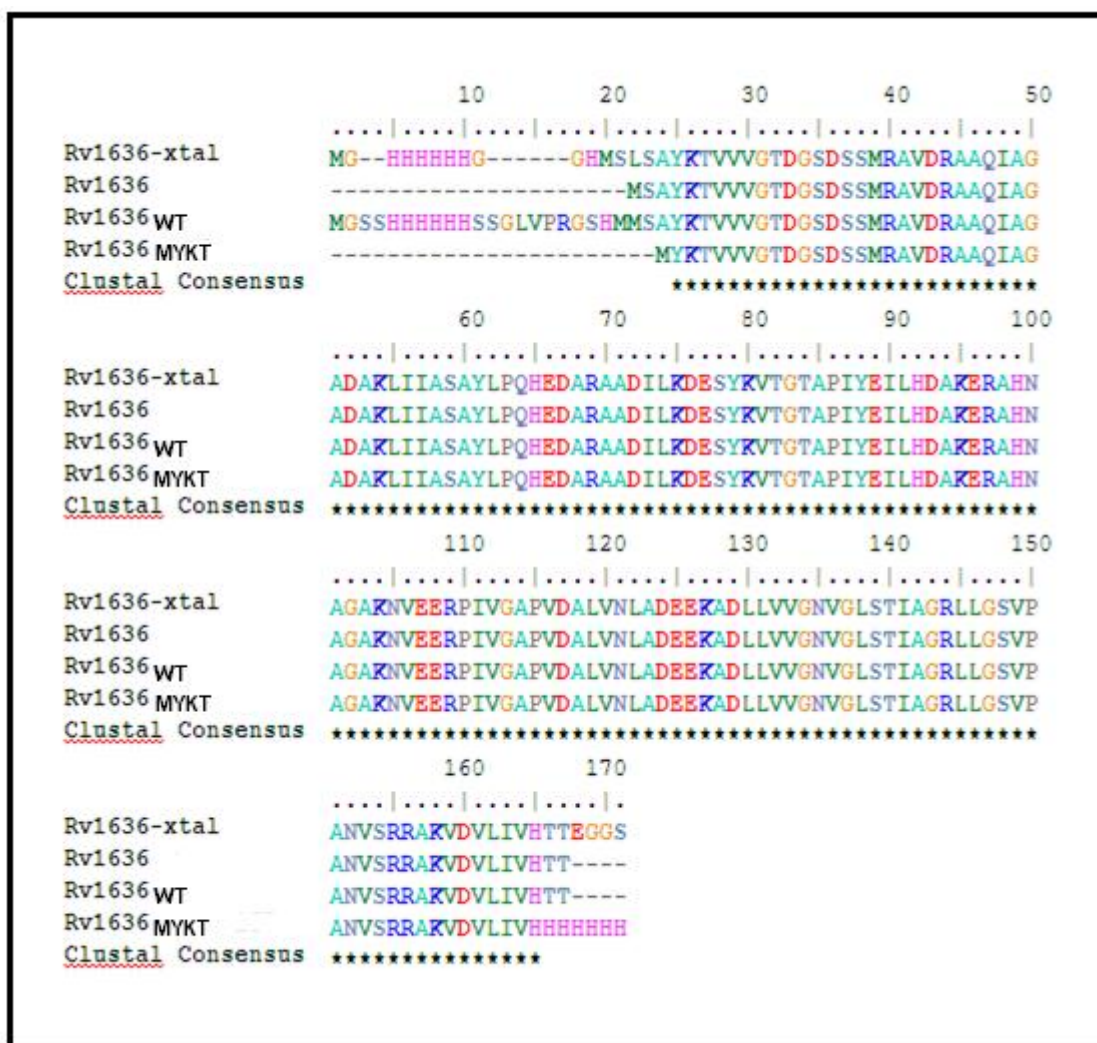


Figure 4.8. Sequence alignment of Rv1636 and the truncated version of Rv1636. The sequence of the deposited crystal structure sequence (Rv1636-Xtal), Rv1636, His-tagged wild type (Rv1636_{WT}), and the truncated version Rv1636_{MYKT} (Rv1636 17-157) are all shown. The consensus sequence is shown below. Alignment constructed with BioEdit.

The truncated Rv1636_{MYKT} constructs were designed to remove any potential terminal residues that could interfere with the native structure. The Rv1636_{MYKT} construct was based on the sequence of the wild type and still contained the tyrosine responsible for the H-bond with Asp139. Once the truncated constructs were confirmed, replacing the tyrosine residue at position 4 was the next priority. Changing tyrosine at position 4 to alanine removed the hydroxyl group (OH) and therefore the H-bond with aspartic acid at position 139. In order to replace the tyrosine, constructs were designed in which the Tyr4 was changed to a phenylalanine with C-terminal his-tag (Rv1636_{MYKT}) and in which Tyr4 was changed to an alanine with C-terminal his-tag (Rv1636_{MAKT}). Schematic diagrams of the constructs are shown in Figure 4.9.

The Universal Stress Proteins of Bacteria

Alanine has a non-polar aliphatic side chain (CH_3) that lacks the OH group and has a neutral charge. Changing the polar tyrosine to the non-polar phenylalanine is a more conservative mutation because both contain aromatic R groups and are of a similar molecular mass. An additional protein was constructed with a point mutation in Asp139, replacing it with an alanine (Rv1636_{MYKT D139A}). It should be noted in the Rv1636_{MAKT} mutant, the initiator methionine should be cleaved off automatically *in vivo* in *E. coli* by endogenous methionine aminopeptidase action.

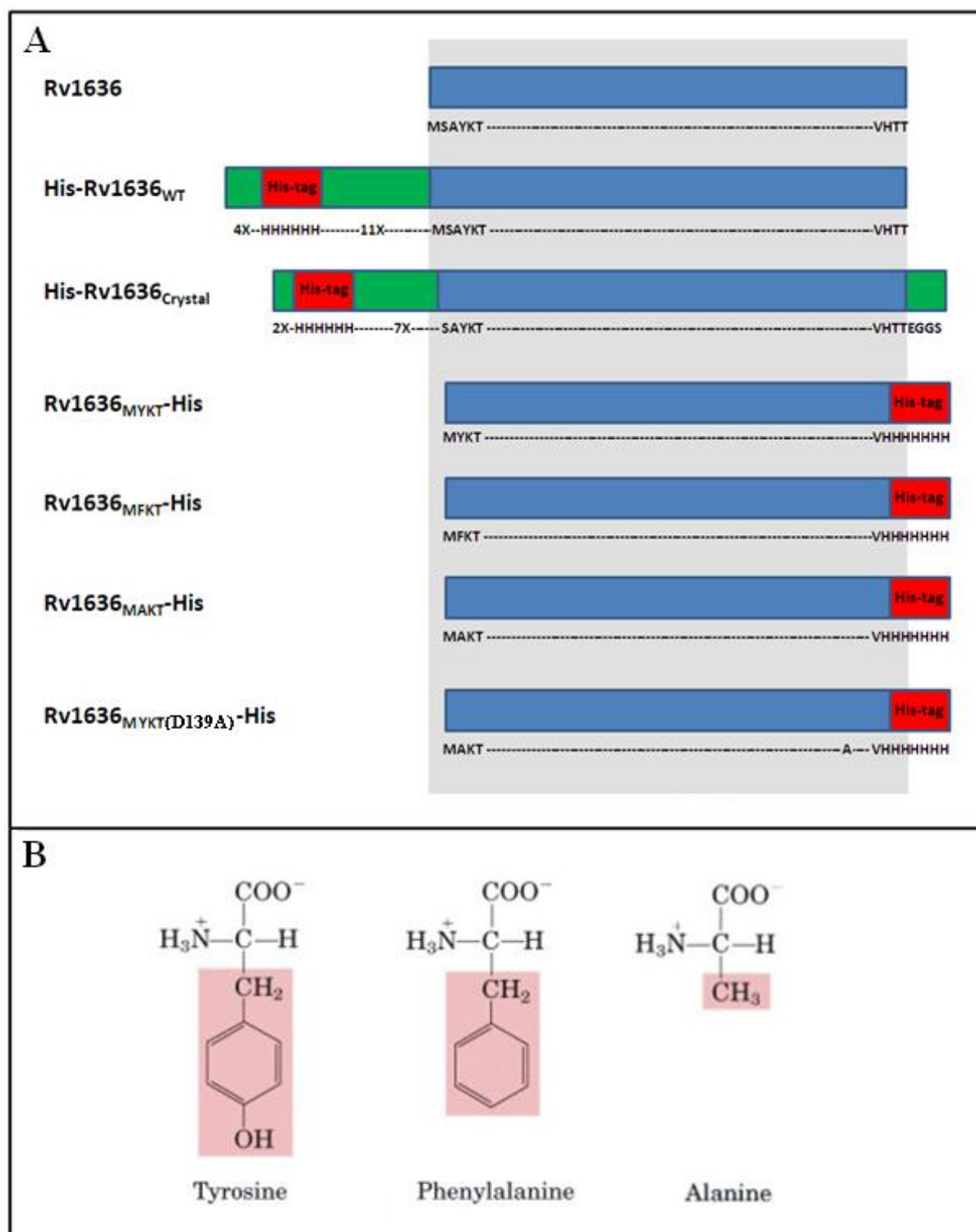


Figure 4.9. Schematic diagram of the R1636 constructs designed and the amino acid side chains. (A) Construct design. Blue indicates the Rv1636 wild type sequence, green are additional linker amino acids and red is the histidine tags. (B) Amino acid side chains of tyrosine, phenylalanine and alanine are highlighted in pink.

The Universal Stress Proteins of Bacteria

The His-tagged *Rv1636* gene in the pGem-Teasy vector was used as a template for the construct design. Primers were designed (see Chapter 2.1.6) to amplify the *Rv1636* gene flanked by *NdeI* and *AvrII* sites. PCR was performed on pGem-Te-*Rv1636* and the products ligated into pGem-Teasy. *Rv1636* was then cloned into pET41a using *NdeI* and *AvrII*. The cloning of the gene in this way removed the pET41a features including the His-tag, GST-tag, thrombin cleavage site and the extra amino acid linker at the beginning, resulting in the gene being 8 base pairs downstream of the ribosome binding site AAGGAG with no additional amino acid additions to the sequence. Translation could therefore occur directly from the methionine codon within the *NdeI* restriction site. pET41a vectors containing the truncated constructs were sequenced to confirm the final sequence. Site directed mutagenesis, as described in chapter 3, was then performed to introduce the D139A change.

4.2.4.2 Purification of Truncated *Rv1636* Constructs

The vectors containing the truncated constructs were transformed into Rosetta cells, the proteins over expressed and purified by IMAC in a similar manner to that described earlier. Solubility assays to determine if the newly designed constructs produced soluble protein were performed as shown in Figure 4.10. *Rv1636*_{MYKT D139A} was found to be completely insoluble. Purification of *Rv1636*_{MFKT} was also complicated by the relatively low levels of protein retrieved following lysis of *E. coli*.

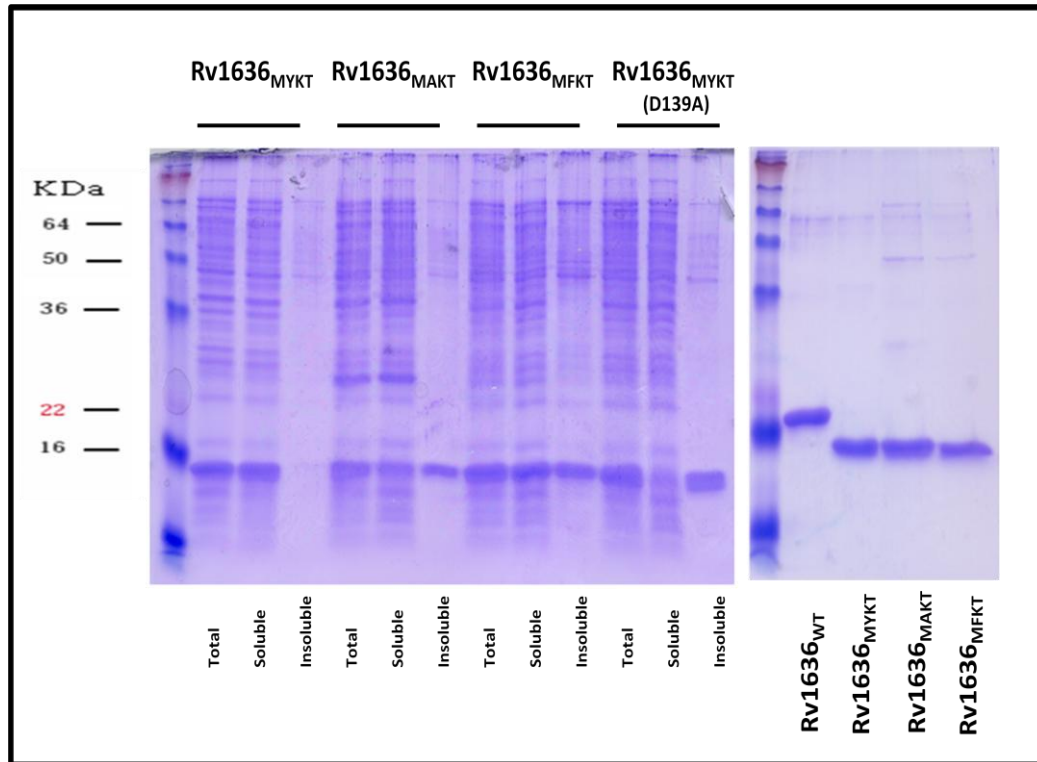


Figure 4.10. Solubility of truncated proteins expressed in *E. coli* Rosetta cells. Expression induced using IPTG and occurred at 37 °C for three hours. Total, soluble and insoluble protein is showed (left) along with the purified soluble protein (right). MYKT_{D139A} was shown to be totally insoluble when expressed in *E. coli*.

4.2.5 Size Exclusion Chromatography - Multi-Angle Laser Light Scattering

SEC-MALLS is a method for absolute molar mass determination and facilitates the direct determination of the characteristic properties of the macromolecular species. The determined molecular mass of the proteins within a solute depends on the oligomeric states formed. The mutagenesis of the Rv1636 in producing the truncated constructs aimed to disrupt the strand swap organisation of Rv1636_{WT}. SEC-MALLS was undertaken to determine if the alterations to Rv1636_{WT} resulted in further oligomeric species being found.

Rv1636_{WT} has been previously characterised by SEC-MALLS and it was shown that the predominant species had a mass calculated by MALLS of ~32 kDa. This mass correlated well with the predicted molecular mass of the dimeric protein. A second peak showed a molecular mass of 66kDa corresponding directly to tetrameric form of the protein. The data suggest the Rv1636_{WT} exists in a

The Universal Stress Proteins of Bacteria

dimer/tetramer equilibrium. The absence of a monomeric unit suggests tight binding of the units in the dimeric conformation.

4.2.5.1 SEC-MALLS shows MAKT Exists in a Monomeric Form

The SEC-MALLS graphs show the UV reading of the elution profile. The calculated molar masses are then overlaid. The void (run off below 14ml) would include any protein too large to interact with the matrix. The SEC-MALLS experiments described were performed on protein that had previously been subjected to SEC and re-concentration. Therefore, any material too large to enter the matrix is derived from newly formed higher order aggregates upon concentration.

The results of the SEC-MALLS on the truncated constructs show that the Rv1636_{MYKT} and Rv1636_{MFKT} remain mostly in a dimeric conformation. Figure 4.11 shows the SEC-MALLS result with Rv1636_{MYKT} sample containing a predominant species of 29.20 kDa. The mass corresponds well with the predicted molecular mass of the dimer being 31.56 kDa. A small peak of protein with a mass 65.02 kDa corresponded to the 63.12 kDa calculated mass for the tetramer. The Rv1636_{MFKT} protein shows evidence of a dimer with the main lower order species being 27.2 kDa. Both samples have higher order oligomeric states corresponding to mass of over 300 kDa. It is likely these are artefacts of concentrating the protein samples prior to loading.

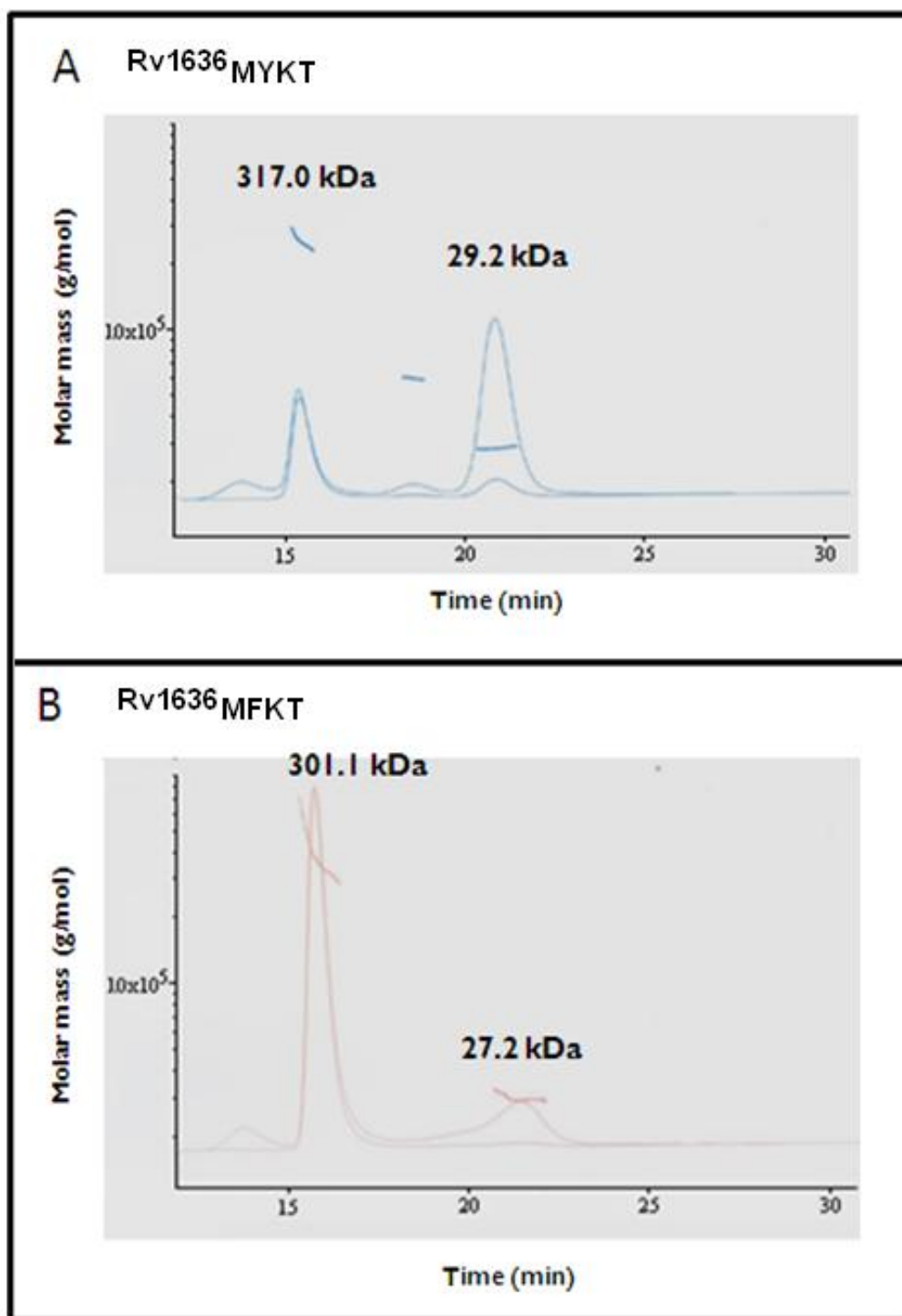


Figure 4.11. SEC-MALLS analysis of the molecular mass of the Rv1636_{MYKT} protein (A) and Rv1636_{MFKT} protein (B) in solution. His-Rv1636 was applied to the SEC column (Superdex 200 10/300 GL column) and subsequently through the light scattering unit. Calculated molar mass of the peaks was done using the Astra 5.1 software.

The Universal Stress Proteins of Bacteria

The SEC-MALLS result for the Rv1636_{MAKT} protein sample provided evidence of a possible monomeric unit (Figure 4.12), suggesting the strand swapped conformation had been disrupted. The lower order species had a calculated mass of 20.59 kDa which is slightly higher than the predicted mass of 15.68 kDa. The protein still appears to have a concentration dependent propensity to form multimers although the equilibrium is shifted downwards to between the dimer and monomer. The molecular mass is therefore is higher than the mass of the monomer and under that of the dimeric form. It appears the dimer/monomer peaks are unresolved suggesting that there is still a tight association between the proteins which exchange with each other rapidly. The peak at 229.4 kDa is also likely to be an artefact of concentrating the proteins for the SEC-MALLS analysis.

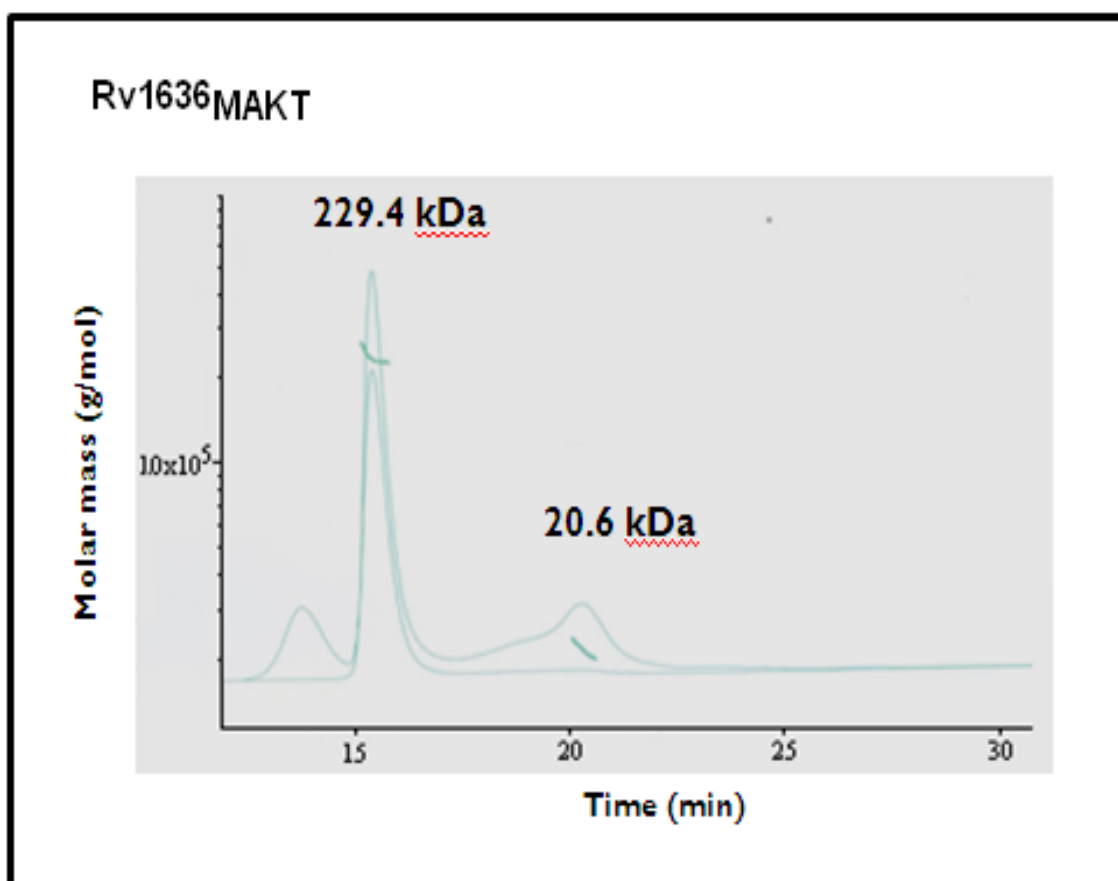


Figure 4.12. SEC-MALLS analysis of the molecular mass of the Rv1636_{MAKT} proteins in solution. Rv1636_{MAKT} was applied to the SEC column (Superdex 200 10/300 GL column) and subsequently through the light scattering unit. Calculation of the molar mass of the peaks was performed using the Astra 5.1 software.

4.2.6 Analytical Ultracentrifugation

Analytical ultracentrifugation provides information on the gross shape of macromolecules, the conformational changes in macromolecules, and size distributions of macromolecular mixtures. Sedimentation velocity experiments can be used to study reversible chemical equilibria between macromolecular species. Prior to performing analytical ultracentrifugation, the protein solutions were separated by size exclusion chromatography.

4.2.6.1 Size Exclusion Chromatography

The proteins were concentrated and then subjected to size exclusion chromatography to prepare pure samples suitable for analytical ultracentrifugation. The AUC technique allows the analysis of the state of the proteins following purification without the need for further concentration of the sample. In size exclusion chromatography, the larger proteins interact less with the matrix pores and are eluted first. The smaller proteins take longer to migrate through the matrix and are eluted at the end. Both Rv1636_{WT} and Rv1636_{MAKT} produced readouts with two main peaks as shown in Figure 4.13. The Rv1636_{WT} main peaks were eluted later compared to the Rv1636_{MAKT} elution which showed the presence of separate distinct peaks.

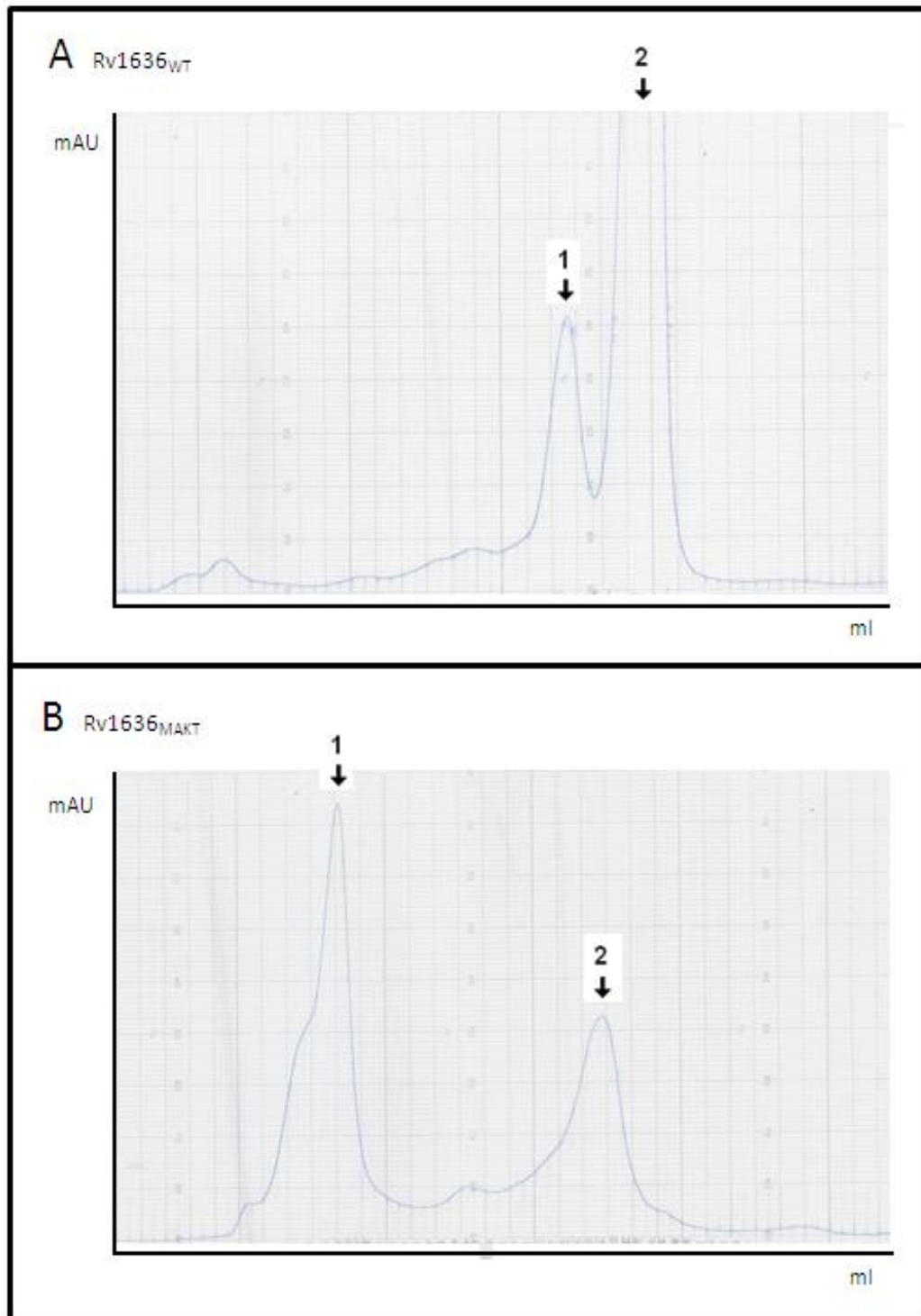


Figure 4.13. Size exclusion chromatography of Rv1636_{WT} (A) and Rv1636_{MAKT} (B) using HiLoad 26/60 Superdex 75 prep grade column (GE Healthcare/Amersham). Samples were run at flow rate of 0.3 ml/minute and eluted in 3 ml fractions. Samples were taken from the peak fractions highlighted and subjected to analytical centrifugation.

4.2.6.2 AUC of Rv1636_{WT}

The analytical ultracentrifugation data for the Rv1636_{WT} protein shows the presence of dimeric and tetrameric states as shown in Figure 4.14. The dimer/tetramer equilibrium observed is in agreement with the SEC-MALLS data. The best-fit C(S) function for the Rv1636_{WT} sedimentation velocity experiments contains two species that sediment at 0.5S and 1S respectively, and correlate to the dimeric and tetrameric forms. The presence of two individual resolved peaks suggests the interaction between the oligomeric states is a tight association and the exchange between the two forms is relatively low. Figure 4.14A shows peak 1 contains particles of size 57.0 kDa which represents the tetramer and accounts for the largest proportion of the protein present with a C(S) reading of 1S. The tetramer can then split into the dimer over time and this is represented by the presence of smaller amounts of dimeric protein with a molecular mass of 32 kDa and C(S) reading of 0.5. Peak 2 AUC data in Figure 4.14B shows evidence of a pure dimer with a mass of 29.8 kDa and a C(S) value of 2S.

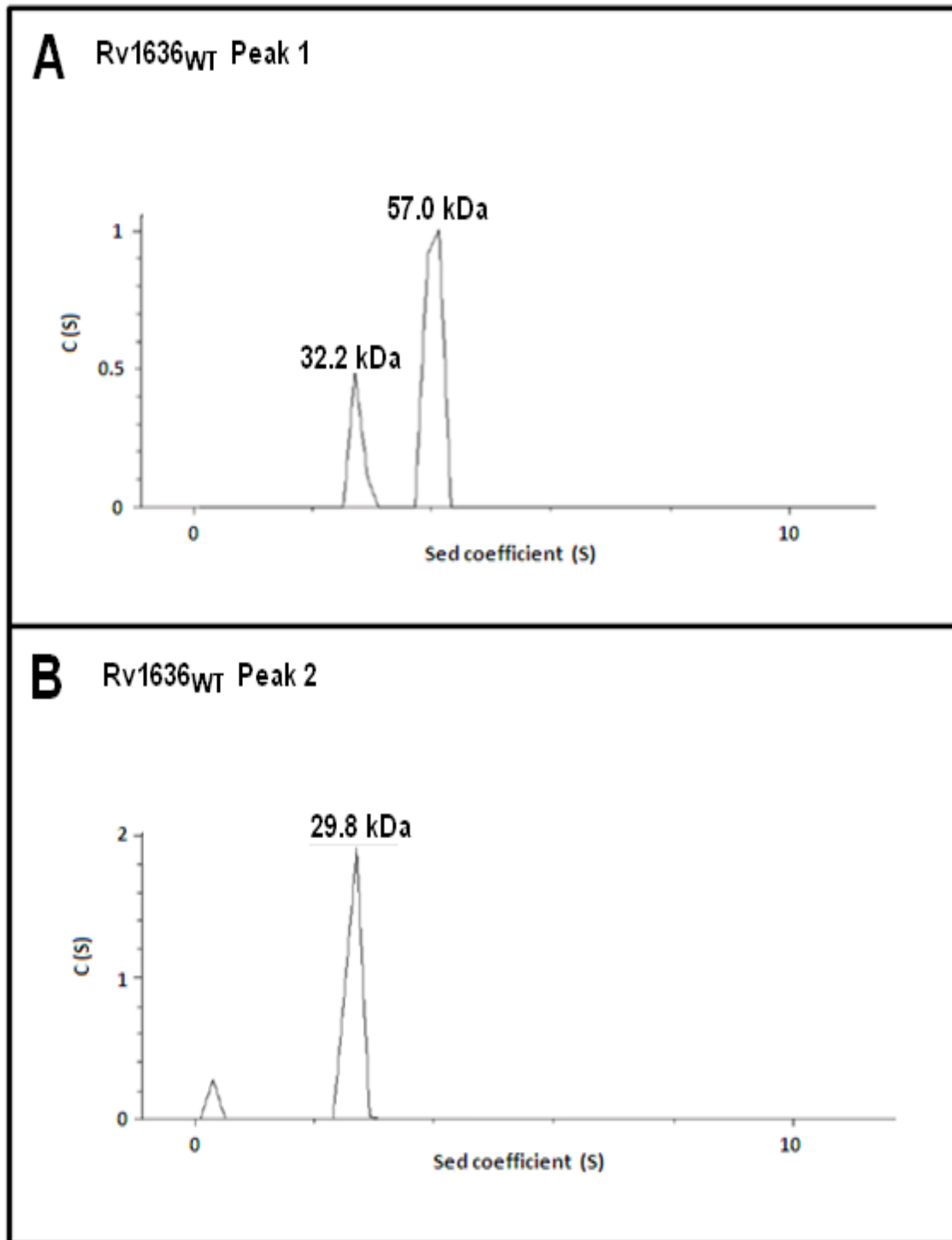


Figure 4.14. Analytical centrifugation data of Rv1636_{WT}. Samples from peak 1 (A) and peak 2 (B) from the size exclusion chromatography (Figure 4.13A) were analysed by AUC in the Beckman Ultracentrifuge. The C(S) function that best fits the sedimentation velocity data for Rv1636_{WT} is shown.

4.2.6.3 AUC of Rv1636_{MAKT} Construct

The MAKT construct was also passed through a size exclusion chromatography column and produced two distinct peaks (Figure 4.13B). When subjected to SEC-MALLS, the Rv1636_{MAKT} peak 1 gave sizes of 30.3 kDa, 56.4 kDa, 111 kDa and 221 kDa. The first of these is likely to be a dimer, with a predicted molecular mass of 31.36 kDa, followed by the tetramer, octomer and 16mer (Figure 4.15A). The 111 kDa peak, presumably representing an octomeric formation is the predominant species. Rv1636_{MAKT} Peak 2 shows the presence of monomeric species at 15.4 kDa, a tetramer peak 63.8 kDa and peak representing the octomer at 108 kDa (Figure 4.15B). There is a small evidence of dimer formation however this is hidden by the wide tetramer peak. In conclusion, the biophysical data suggests that there is a monomeric form of Rv1636_{MAKT}; however the protein is also found in higher oligomeric states.

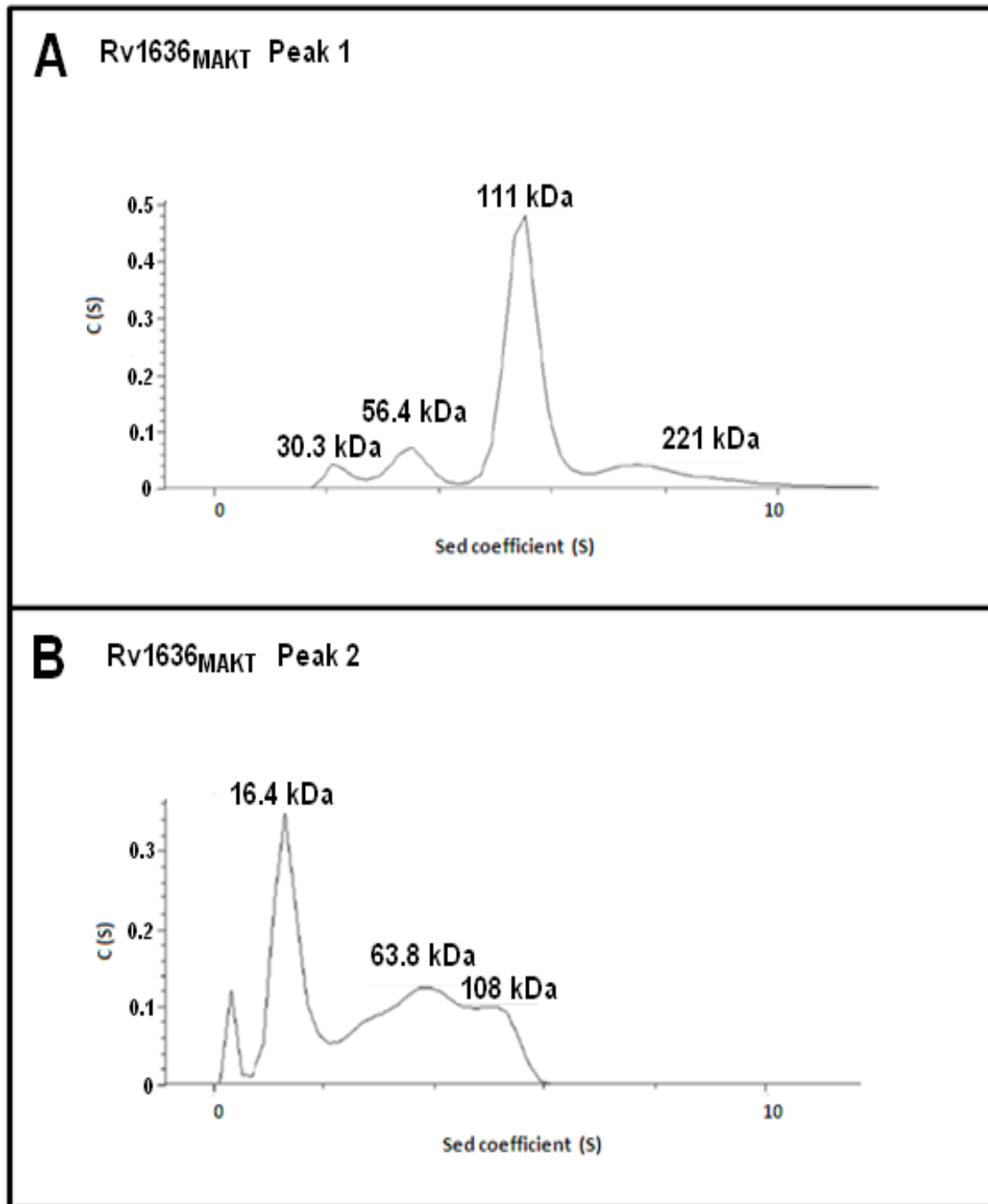


Figure 4.15. Analytical centrifugation of Rv1636_{MAKT}. Samples from peak 1 (A) and peak 2 (B) from the size exclusion chromatography (Figure 4.13B) were analysed by AUC in the Beckman Ultracentrifuge. The C(S) function that best fits the sedimentation velocity data for Rv1636_{MAKT} is shown.

4.2.6.4 Does Rv1636 Hydrolyse NTPs?

Previous work by S. Leiva (2007) showed that Rv1636_{WT} can hydrolyse nucleotides in the absence of cations. However, the activity was shown to be low and the kinetics were measured using equimolar concentration of nucleotide to protein. The kinetic data suggested only a small fraction of the protein sample was active. The site directed mutants created in Chapter 3 were tested for their ability to hydrolyse nucleotide triphosphate (Figure 4.16). The mutants all had below 15% activity of the Rv1636_{WT}. Reactions were performed with equimolar concentrations of nucleotide to the protein concentration. Prior to further analysis, the truncated constructs were assayed, to determine the levels of NTPase activity. Site-directed mutants in the truncated constructs may be of more interest if the kinetic reactions could be performed with a higher concentration of nucleotide.

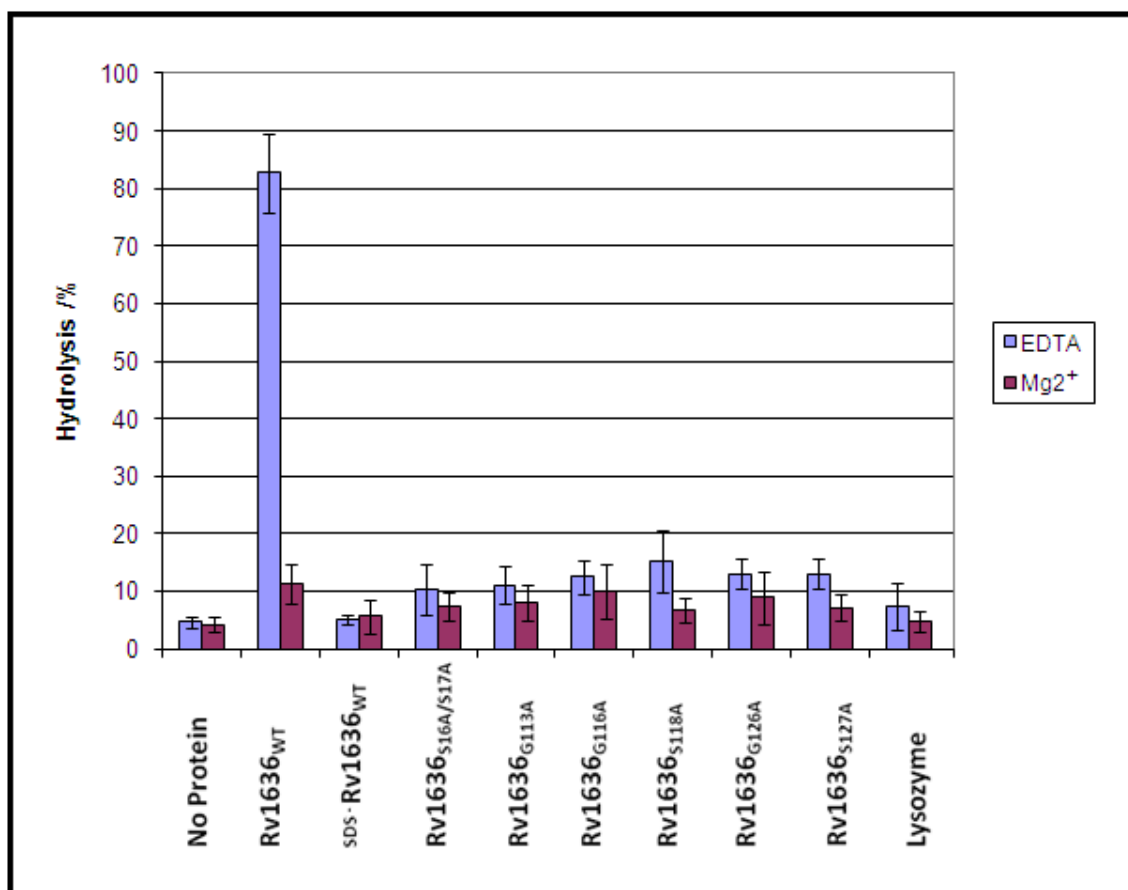


Figure 4.16. Nucleotide hydrolysis by Rv1636_{WT} and the site direct mutants. Reactions were carried out for 30 minutes at 37 °C in a reaction buffer containing [α -P³²] GTP, 10 mM Tris (pH 7.5), 50mM NaCl and 10 μ M protein. Reactions contained either 1 mM EDTA or 5 mM MgCl₂. Hydrolysis was visualised by separation on TLC plates and quantification was achieved using Aida Analyser Software.

The Universal Stress Proteins of Bacteria

To compare the level of NTP hydrolysis between the Rv1636_{WT} and the newly designed constructs, the hydrolysis of [α -P³²] GTP was measured with increasing concentrations of non-radiolabelled competitor GTP from 0 mM to 1 mM (Figure 4.17).

The results of the hydrolysis suggested that the proteins Rv1636_{MYKT} and Rv1636_{MFKT} had comparable level of hydrolysis to the Rv1636_{WT} protein. However Rv1636_{MAKT} appeared to show an increased level of hydrolysis with all the GTP being hydrolysed with cold competitor GTP at 0.1 mM. The hydrolysis was also shown to be sequential with the hydrolysis of the β phosphate also occurring producing GMP.

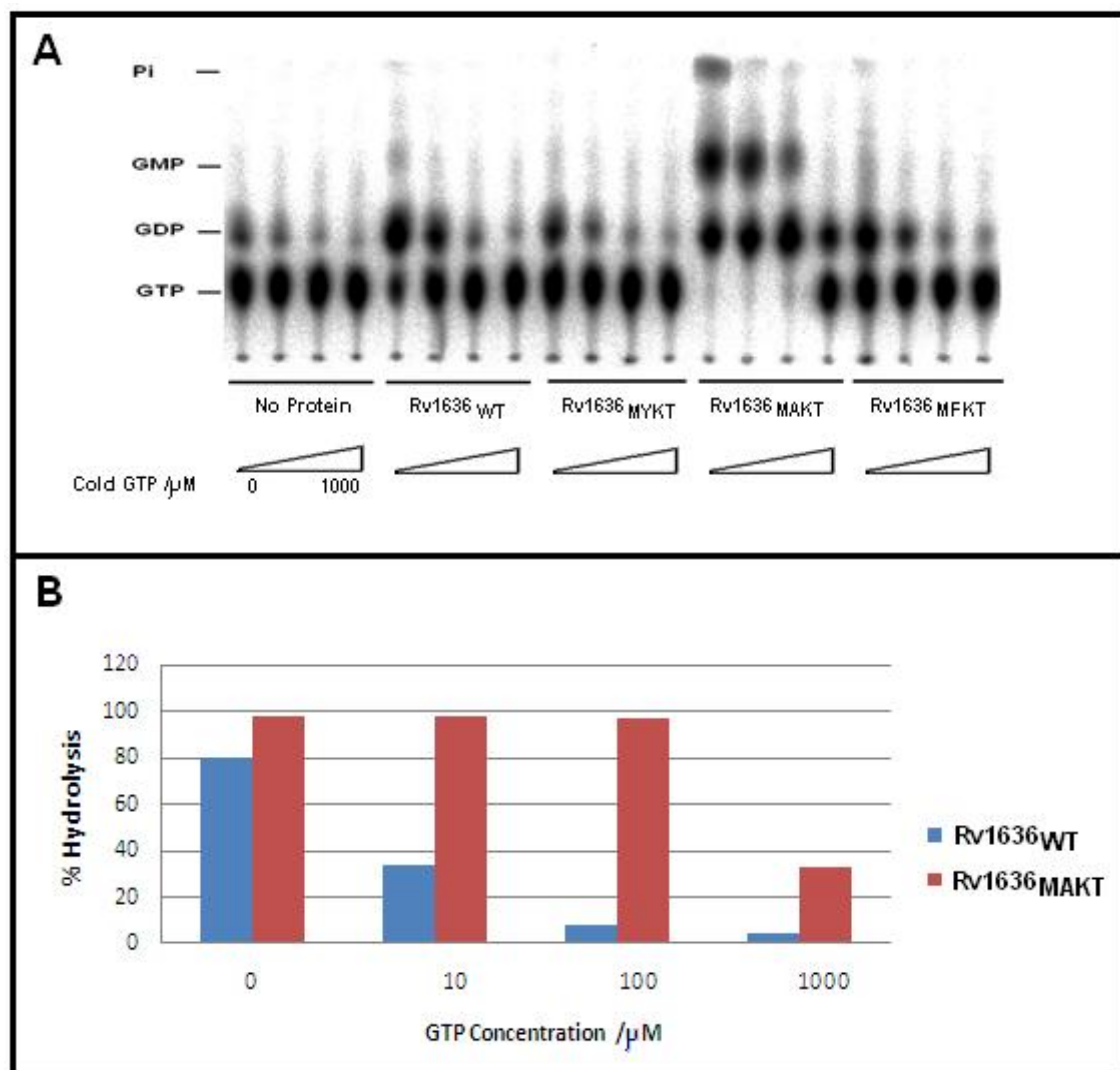


Figure 4.17. Nucleotide hydrolysis by Rv1636_{WT} and the Truncated constructs. (A) Hydrolysis visualised on TLC plate to separate the nucleotides. (B) Comparison of hydrolysis between Rv1636_{WT} and Rv1636_{MAKT}.

The Universal Stress Proteins of Bacteria

To investigate the kinetics of the hydrolysis of NTP by Rv1636_{WT} and Rv1636_{MAKT}, the protein samples were to be assayed immediately following size exclusion chromatography. These samples would be expected to provide the most accurate measurement of the reaction kinetics. Size exclusion chromatography separates proteins based on size and would also remove any aggregated protein which is likely to have a lower hydrolysing activity. Following the size exclusion chromatography, NTPase assays were performed on the eluted proteins in the major peaks for the proteins Rv1636_{WT}, Rv1636_{WT(G113A)}, Rv1636_{MAKT} and Rv1636_{MAKT(G113A)}. Surprisingly, no NTPase activity was found following size exclusion chromatography.

To investigate the NTPase activity further, the experiments were repeated. Again it was found that the SEC elution profile peaks corresponding to Rv1636_{WT} (Peak 1 and 2) did not have any hydrolysing activity (Figure 4.18). The NTPase activity was located in one of the small peaks that had been pooled together and termed sample 3.

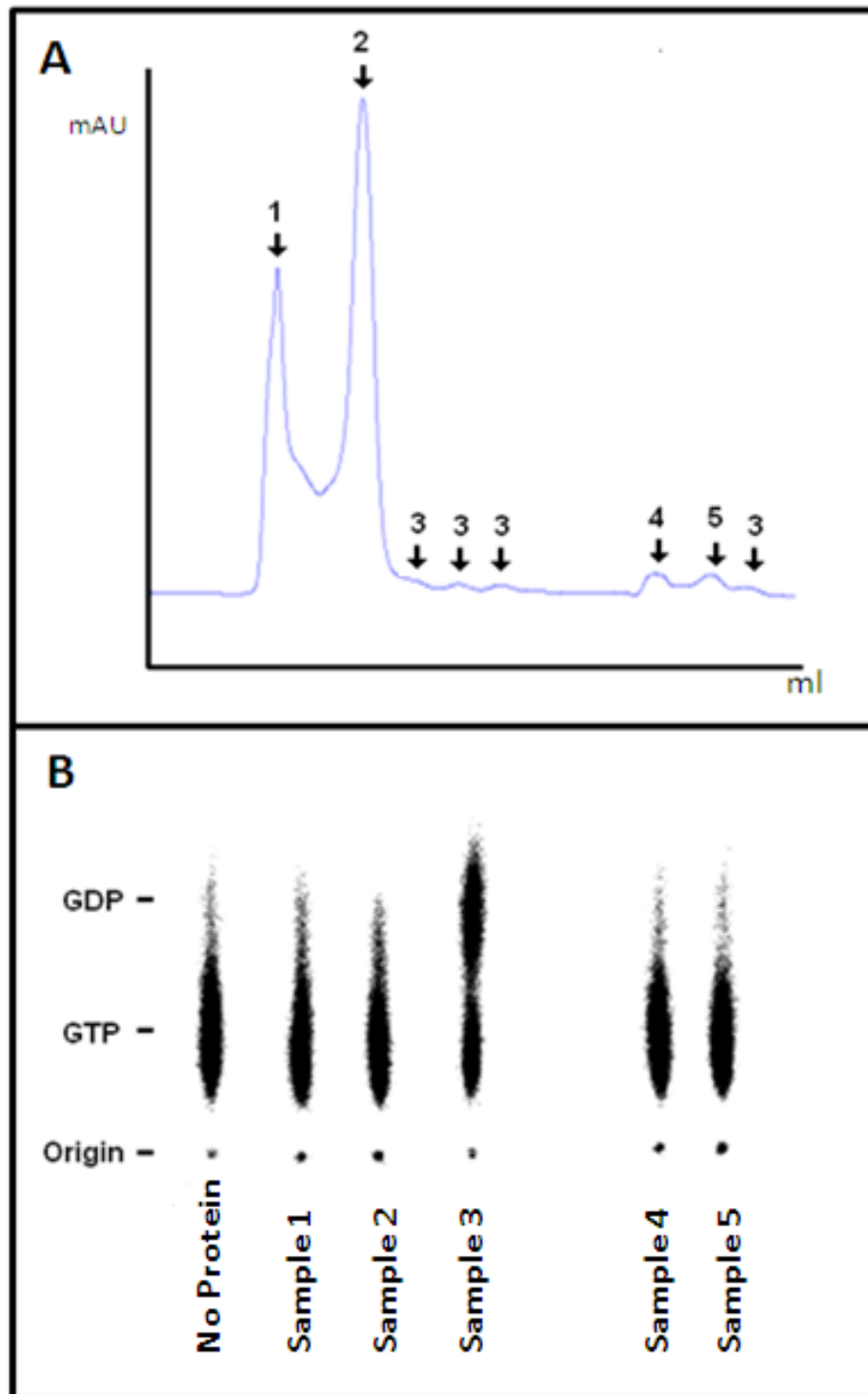


Figure 4.18. Size exclusion chromatography of Rv1636_{WT} followed by an NTPase assay. (A) FPLC elution chromatogram of Rv1636_{WT} SEC showing UV_{280nm} reading (mAU, milli absorbance units). Pooled samples are indicated by arrows. Sample 3 was collected from four small peaks. (B) NTPase assay directly on the peak samples with [α -P³²]-GTP. Reactions were separated by TLC and visualised by phosphoimaging using Fuji Film 5000 phosphoimager.

The Universal Stress Proteins of Bacteria

To further investigate the NTP hydrolysis, the fractions gathered in sample 3 were concentrated and separated by SDS-PAGE. However, no protein species was abundant enough to be visualised on the gel. To identify a protein by mass spectrometry, a Coomassie stained band is required. Further assays were performed following IMAC purification of Rv1636_{WT} which involved pooling fractions of Rv1636_{WT} and directly assaying the samples for activity. Figure 4.19 shows that the level of Rv1636_{WT} present did not correlate with the hydrolysing activity observed. The experiments indicated the activity was not linked to the proportion of Rv1636_{WT} in the sample and could, instead, be associated with a modified form of Rv1636, or a contaminating *E. coli* protein, or the hydrolysis could be dependent on unknown factors. It was, therefore, reasoned that the activities found were not directly associated with the common species of Rv1636_{WT} obtained by purification. To continue the investigation of Rv1636 biochemical activities, the nucleotide binding capabilities of the Rv1636 were analysed.

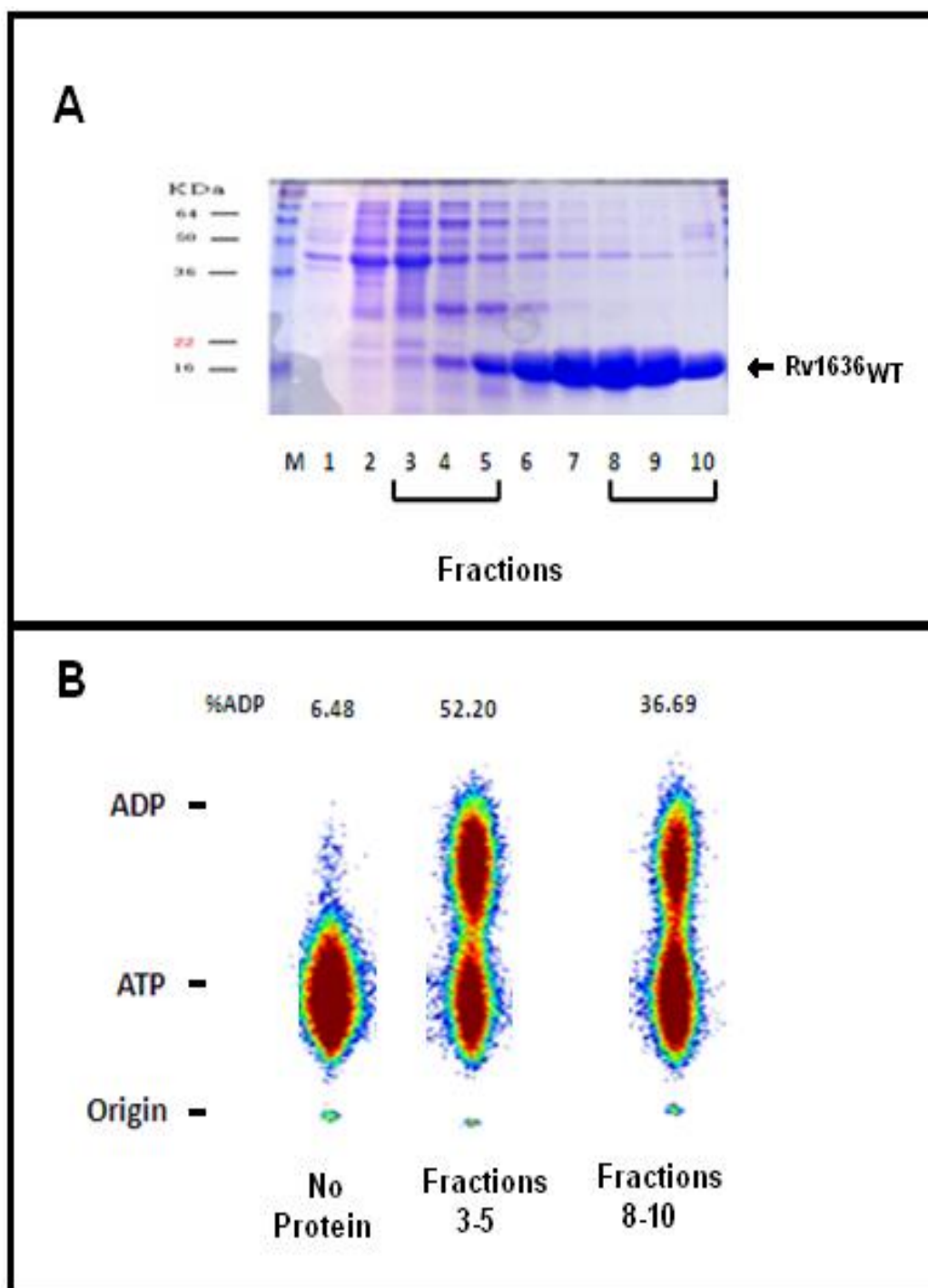


Figure 4.19. NTPase assay on fractions after IMAC purification. (A) Purification of Rv1636_{WT} showing the fractions (1-10) electrophoresed by SDS-PAGE and stained with Coomassie. (B) NTPase assay with [α - P^{32}]-ATP. Reactions were separated by TLC and visualised by phosphoimaging using Fuji Film 5000 phosphoimager.

4.2.6.5 Do the Truncated Constructs Bind Nucleotide Triphosphate?

A filter binding assay was used to investigate the nucleotide binding properties of the truncated constructs. Radio-labelled ATP and proteins were incubated together at 4 °C and passed through a nitrocellulose membrane. The proteins, with any bound nucleotide, remain attached to the membrane whilst any unbound nucleotide is washed away. The results of the vacuum filter binding for the Rv1636_{WT}, Rv1636_{MYKT} and Rv1636_{MAKT} are shown in Figure 4.20. Initial binding tests with the proteins Rv1636_{MAKT} and Rv1636_{MYKT} showed the proteins could bind nucleotides. Rv1636_{MAKT} appeared to bind a higher amount of nucleotide as quantified by the phosphoimaging. The Rv1636_{WT} and Rv1636_{MYKT} have been shown to exist in an equilibrium between dimer and tetramer (Figure 4.11). The lower binding activity could be a result of the active sites being blocked as shown in the crystal structure. The control proteins including BSA and SDS-denatured Rv1636 showed similar levels to background radiation observed.

The phosphoimages were quantified and are able to give the relative binding in terms of photo-stimulated luminescence (PSL). PSL gives the intensity of the radioactivity bound to the membrane and therefore is indicative of the amount of nucleotide bound to the protein present. Reactions were performed twice and average values were calculated. The Rv1636_{MFKT} protein was found to be unstable and, when stored in glycerol at -20 °C, the protein precipitated and was therefore not assayed.

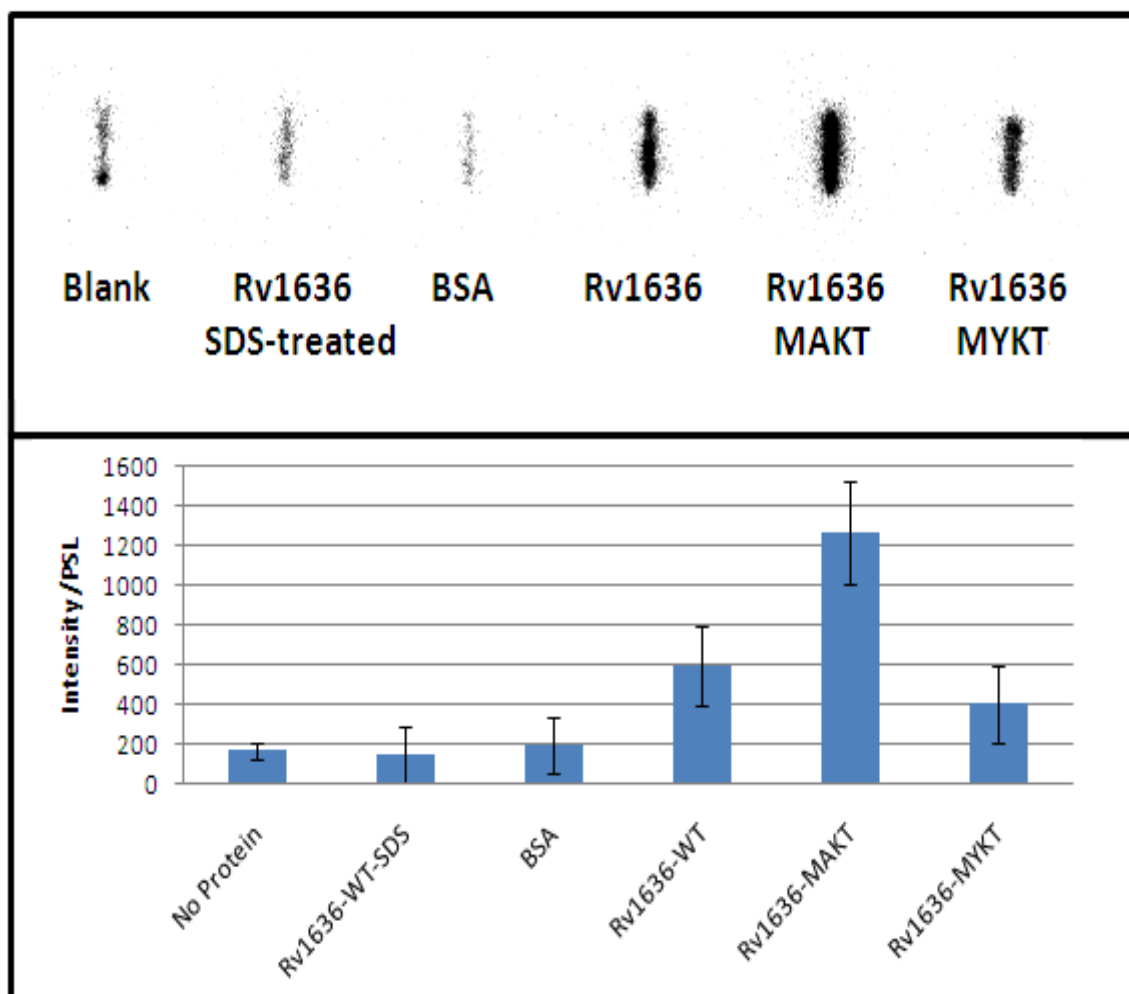


Figure 4.20. Nucleotide binding of Rv1636_{WT} and truncated constructs using vacuum filter binding. 30 μ M proteins were incubated with [α -P³²]-ATP at 4 °C under UV light for 40 minutes. Reactions were passed through a nitrocellulose membrane before washing. Visualised using a Fuji film 5000 phosphorimager.

4.3 Discussion

4.3.1 Rv1636 Crystal Structure

The crystal structure of the Rv1636 protein (PDB 1TQ8) showed the protein formed a tetrameric unit, composed of two strand exchanged dimers. A third dimeric unit was then lightly bonded to the tetrameric structure. The swapping of the beta sheets between adjacent monomers appears to disrupt the putative nucleotide binding motif. The subsequent formation of a tetramer seems to block the site further. It is possible the tetramer configuration observed may represent a crystallisation artefact and may not be the oligomeric state of the protein *in vivo*. Crystallisation uses high concentration of proteins in solution and this can result in forcing the proteins present into unnatural forms. Conditions of crystallisation may have favoured the formation of the tetramer; although the presence of nucleotide and/or addition of a cation may have changed the state observed. In order to achieve crystallisation, many additives are often added to buffer solutions and crystallisation may not form structures that are biologically relevant.

In the crystal structure of Rv1636, it was observed that the tag used was modified slightly at both the N- and the C-terminus due to the choice of expression vector, including the addition of EGGS at the N-terminal end. The affinity tag chosen was a basic his-tag, which should be suitable for crystallisation. Bucher *et al.*, 2002 discussed the use of tags in crystallisation and concluded that his-tags have not been shown to be detrimental to crystallization of proteins and the quality of crystals is equal to those where removal of the tag has taken place. The crystal structure deposited shows the number of hydrogen bonds between adjacent units. The dimer is clearly the most tightly associated unit with up to 28 hydrogen bonds connecting the two. The tetramer bond is likely to be a more transient association, however the large buried surface area suggests the tetramer form is quite tight.

4.3.2 Design and Purification of Truncated Rv1636

Truncated proteins were successfully produced and did not contain surplus amino acids that could influence the structure, except the addition of a small 6-his tag to allow purification. Rv1636_{MYKT}, Rv1636_{MAKT}, Rv1636_{MFKT}, and Rv1636_{MYKT D139A} were successfully expressed in *E. coli*. Rv1636_{MYKT D139A} was found to be totally insoluble and work did not continue with this construct. Rv1636_{MYKT}, Rv1636_{MAKT} and Rv1636_{MFKT} were expressed and purified to comparable levels to

The Universal Stress Proteins of Bacteria

Rv1636_{WT}. The Rv1636_{MFKT} was found not to be stable for long term storage and precipitated when stored in 30% glycerol at -20°C. The expression of Rv1636_{MFKT} also inhibited the growth of the *E. coli* Rosetta cells. It is likely the increased production of an unstable protein was detrimental to the growth of the bacteria as the protein may become unstable and form insoluble inclusion bodies within the cell. The unstable nature of the proteins highlights the importance of the internal interactions and bonds to the structural integrity of the protein. Thermal denaturation profiling to obtain T_M values could be undertaken to determine if the mutation destabilises the native structure.

4.3.3 Biophysical Analysis of Rv1636

The oligomeric states of Rv1636 and the truncated proteins were analysed using a range of biophysical techniques including SEC-MALLS, AUC and size exclusion chromatography. The results for Rv1636_{WT} were in agreement with the crystal structure of Rv1636. The results of the sedimentation velocity experiments and the SEC-MALLS supported that the protein exists in a dimer-tetramer equilibrium. Overall the data studying the Rv1636_{WT} protein showed no evidence of a monomeric unit present.

To further the analysis, truncated constructs were designed and purified, before being subjected to the biophysical techniques. Data for Rv1636_{MAKT} provided evidence of the presence of a monomeric unit, with a molecular mass of ~16 kDa, and was the first time the singular species was observed in the biophysical studies. The strength of a dimer-dimer interaction can influence what state the protein is in at cellular concentrations and small amino acid modifications can change the strength of the intermolecular interactions. The replacement of tyrosine with adenine resulted in a dramatic change in the state of the proteins in solution. The Rv1636_{MFKT} protein containing a phenylalanine did not produce any significant changes from the Rv1636_{MYKT} version. Rv1636_{MFKT} was however found to precipitate when stored in glycerol at -20 °C and was not subjected to SEC-MALLS analysis.

The SEC-MALLS of the Rv1636_{MAKT} indicated that there may be a monomeric unit present however the mass was calculated to be 20 kDa, above that of the individual unit. The AUC confirmed the presence of a monomer unit in the solution with a species at 15.4 kDa observed. In the AUC the Rv1636_{MAKT} protein did show some evidence of forming higher order structures; however it is hard to envisage how multiple tetramers would bind together if the higher structure were indeed biologically relevant.

4.3.4 Biochemical Properties of Rv1636

The hydrolysis previously described in the laboratory on kinetics of the NTPase activity including the Michaelis-Menten constant (K_m) dynamics suggests the activity of the Rv1636_{WT} purified protein stock was relatively low. With further kinetic analysis using a size exclusion chromatography purified product, this activity has not been clearly reproduced. The cation-independent activity could therefore be due to an endogenous *E. coli* protein that co-purifies as a contaminant. It is also possible that Rv1636_{WT} is hydrolysing the NTPs, although the active form of the protein may not be known and may pass through the matrix at a different rate to the inactive abundant common species of Rv1636_{WT}. Attempts to identify the protein responsible for the NTPase hydrolysis were unsuccessful and the activity could not be directly associated with one protein that could be analysed further by mass spectrometry.

The combination of biophysical techniques often provides greater insights into the structural forms of proteins. For example, Alliance Protein Laboratories (www.ap-lab.com) showed when studying antibodies that the elution position through the SEC could indicate the presence of a trimer or a tetramer; however on analysis with MALLS it was seen the two distinct peaks were actually different conformations of a dimer. Through combination of the data of sedimentation velocities and SEC-MALLS it is possible to detect different conformations of specific oligomeric states. It is possible that the active conformation of Rv1636 travels through the SEC at a different rate and it is not represented by the major peaks.

Nucleotide triphosphatase activity involves the hydrolysis of the β - γ phosphate bond of the nucleotide. Thermodynamically and kinetically unfavourable conformational changes of biological substrates are often required to allow directed and ordered biological processes to occur. These conformational changes are often initiated by so called 'molecular machines' that couple nucleotide hydrolysis, such as the energy available in ATP, to direct interactions with the substrate molecules. Zarembinski *et al.*, 1998, first described the ATP binding capabilities of USPs and suggested that MJ0577 may function as an ATP-binding molecular switch.

It was thought Rv1636 may have capabilities to hydrolyse ATP in the absence of cations. Hydrolysis in the absence of cations has been reported in only three NTPases to date. Cel-1 of *Caenorhabditis elegans* (Takagi *et al.*, 1997), Mcel of the mouse (Ho *et al.*, 1998) and BVP of AcNPV baculovirus (Gross & Shuman, 1998a) are all triphosphatases that are involved in forming part of the capping apparatus of these organisms. These RNA triphosphatases, with the described cation-independent hydrolysis activity, have been described as eukaryotic mRNA capping enzymes (Takagi *et al.*, 1997; 2003). Sequence analysis indicates that the N-terminal section of the enzymes has

The Universal Stress Proteins of Bacteria

similarities to protein tyrosine phosphatases. Importantly, there is a strict conservation motif: (I/V)HCXAGXGR(S/T)G. The cysteine residue is thought to be the active site for phosphoryl transfer (Takagi *et al.*, 2003; Martins & Shuman, 2000) and is therefore essential for the RNA triphosphatase activity. There is no sequence similarity between these proteins and Rv1636, which, in addition, contains no cysteine residues.

Whilst the cysteine residue appears important for the described cation-independent NTPase activity, proteins that can hydrolyse nucleotides without the classical Walker motifs described in P-loop have been described. The P42 protein from *Mycoplasma mobilis* exhibits ATPase activity; however it does not contain the Walker A sequence (Ohtani & Miyata, 2007). Rv1636 possesses the ATP binding consensus motif G-X₂-G-X₉-G(S/T) which is also found in many other USPs; however, to date, hydrolysis has not been shown. The apparent cation-independence of the hydrolysis observed is rare and discovering the enzyme responsible would be of interest. Patton and Thompson, 2004, describe precautions about using chelators to buffer metals as chelators cannot bring metal ion concentration to zero.

4.3.5 Ligands of USPs

The range of ligands shown to be bound to the 12 crystallised USPs indicates that USPs might possess diverse functions and suggests the binding properties of the USPs to different ligands may have evolved in order for the proteins to fulfil distinct roles. Ligands found bound include all three states of phosphorylated adenosine. Monophosphate was found in *Arabidopsis thaliana* AT3G01520 gene product, diphosphate found in *Klebsiella pneumonia* gene kpn_01444 product and triphosphate found in *M. jannaschii* MJ0577, *Thermus thermophilus* gene product TTHA0895 and Rv2623 of *Mtb*. Magnesium, manganese and chloride ions were also found implying the ions may have a role in the proteins' function. Magnesium was discovered in four structures, chloride in two structures and manganese only in the MJ0577 protein. When visualising the Rv1636 protein structure, it was seen that the nucleotide binding pocket has a complex fold that forms a hydrophobic section. This section might be involved in expelling H₂O and creating an environment where binding to ligands can occur.

T. thermophilus (2z08), *M. jannaschii* (1mjh), *Klebsiella pneumonia* (3fdx) and *Mtb* (3cis) show tetrameric assemblies with crystal structures showing USP-ATP complexes. Most USPs without ATP bound crystallise as dimers. The crystal structure of *K. pneumonia* (3fdx) which lacks the ATP-binding motif still crystallised with ATP and similarly the structure of 3fh0 crystallised with ADP. The

The Universal Stress Proteins of Bacteria

presence of the motif is not detrimental to ATP binding and the lack of the motif cannot be used as a criterion to decide if the ATP binding is possible.

The crystal structure of Rv1636 in the PDB did not show any ligands bound. However, the UV crosslinking discussed in the previous chapter suggests the protein can bind nucleotides. Initial nucleotide binding assays using vacuum filtration through a nitrocellulose membrane suggested that Rv1636_{MAKT} protein can bind radiolabelled ATP to a higher degree than the wild type protein. An expression vector containing a Rv1636_{MAKT G113A} mutation has already been created, however nucleotide binding assays have not been performed. Rv1636_{MAKT} has also been cloned into an over-expression system in vector pMV261 along with the Rv1636_{WT} for expression studies in mycobacteria (see Chapter 5). Total analysis of the nucleotide binding could again be performed using a HPLC based method to measure the level of binding.

4.3.6 Structural Insights into USPs

Proteins can also undergo conformational changes which can directly be involved in the biological role of the protein. The function of the proteins can change with these reversible structural changes. The loops of Rv1636 are likely to play a role in the manipulation of any bound ligand. Although the crystallisation studies have shown only homodimer and homotetramer formation, heterodimer formation within the USP classes was recently studied by Nachin *et al.*, 2008, showing that USP domains of different USPs could interact with each other. Through analysis of the USPs by multiple sequence alignments, the residues at the dimer-dimer interfaces were studied. The physiochemical properties of these interface residues showed that, as well as forming the predicted homodimers, the USPs can also form heterodimers. The Nachin group's studies of the *E. coli* USPs confirmed the existence of these heterodimeric interactions in both *in vivo* and *in vitro*.

Studying UspA and UspC, both class I USPs, Nachin *et al.* (2008) showed the proteins can form these homo- and heterodimers. It was hypothesised that the dynamic nature of these interactions could help the USPs fulfil a variety of roles within the cell and possibly provide details as to the extensive range of functions that have been assigned to USP previously described; including stress survival, metabolism of carbon, virulence, motility and cellular adhesion. The ability of USPs to undergo heterodimeric interactions and potentially perform diverse functional roles within the cell, could explain why many organisms have multiple copies of USPs within their genomes. *Mtb* for example contains 9 or 10 USPs and the domains are present in numerous organisational modes.

The Universal Stress Proteins of Bacteria

Chromosomal deletions did indeed suggest the USPs of *E. coli* have both distinct and redundant functions and Nachin *et al.* (2005) suggested USP roles may overlap as they showed mutants in UspA and UspD have cumulative affects when studying oxidative stress. It is possible to envisage the USPs working as accessory factors as well as within pathways in which they may have more individual, distinct roles. The USP partners have the potential to form protein complexes that can then play a role in further pathways. The results of the physiochemical studies by Nachin *et al.* (2008) investigating the interface residues suggested heterodimerisations between classes was unlikely and heterodimers are only likely between members of the same class. The conservation of key interface residues could have contributed to the functional expansion of the roles of USPs into many biological pathways, through the ability to form hetero- and homodimers.

The laboratory grouped the USPs of *E. coli* into four classes. Class I included UspACD, Class II UspFG and the two domains of the tandem domain UspE, called E1 and E2, were described as falling into Class III and Class IV respectively as they show distinct features to the USP domains of Class I and II. Rv1636 does not seem to fit into these groups as the groups were formed between close homologues of the *E. coli* proteins. Studying the USPs of *Mtb* and seeing which have the same structure could provide clues as to whether Rv1636 is involved in formation of hetero-oligomeric structures within *Mtb*.

4.3.7 USP Structure Related to Function

Heermann *et al.* (2009a) recently published a possible scaffolding role for the UspC protein in *E. coli*. The KdpD/KdpE signalling cascade is a two component system consisting of a sensor kinase (KdpD) and a response regulator (KdpE). The system regulates the *KdpFABC* operon which encodes the high-affinity potassium transport system. The system is highly regulated and is upregulated in response to potassium limitation and salt stress. The protein KdpD has kinase activity that is inhibited under the high K⁺ conditions that occur in response to salt stress (Jung *et al.*, 2000). In order to overcome this inhibition it was showed that UspC provides a scaffolding role and was shown to interact with the USP domain found at the N-terminus stimulus-perceiving end of the KdpD sensor. The interaction and scaffolding then permitted phosphorylation of the KdpE response regulator, leading to the *KdpFABC* proteins being expressed. UspC was not shown to directly influence the enzymatic activities of KdpD, but was shown to stabilise the KdpD/KdpE~P/DNA complex that resulted in a shift in the equilibrium towards the phosphorylated form of KdpE.

The Universal Stress Proteins of Bacteria

Biochemical evidence of the interaction between UspC and KdpD was shown using surface plasmon resonance analysis. KdpD is membrane integrated, however the N-terminus containing the USP domain protrudes into the cytoplasm. The KdpD USP domain, consisting of amino acids 1253-1373, shared the same predicted tertiary structure as USPs; five beta sheets surrounded by four alpha helices, and shared 15-25% homology with UspA and UspC. The interaction was shown not to occur in KdpD1-240. Additionally studying a *uspC* knockout mutant, *kdpFABC* expression in response to salt stress was shown to be significantly reduced. It should be noted KdpD autophosphorylation can also occur in response to other conditions including an increase in ionic strength and also in relation to the level of ATP within the cell. This multi-level control of KdpD activation allows the *kdpFABC* response to be activated in conditions distinct to those that induce expression of UspC.

Further hydropathy plot analysis of the KdpD allowed the creation of a detailed secondary structure model. Combining the analysis with an array of domain-swaps, Heermann *et al.* (2009b) showed that the Usp-Usp interactions are dependent on the net charge of the domains, with the electrostatic nature of the domain playing an important role in the signalling of KdpD. The USP domain in KdpD was swapped with USP domains from a variety of other bacterial USP proteins including the six *E. coli* USPs. Chimeras with UspFG, that could still exhibit kinase and phosphatase activity *in vitro*, could not express the *kdpFABC* operon under limited K⁺ conditions and under salt stress. The net positive charge of the Usp domain in the KdpD protein could not be complemented by the negative charged surface of the UspFG. The usual positive charge interaction of the domains was hypothesised to be important to create the electrostatic repulsion required to shift the KdpD histidine kinase into 'ON' state.

This leads to the suggestion that the sole single domain USP of Rv1636 could provide a scaffolding role to the USP domain in the KdpD protein found in *Mtb*. Rv1028c is a KdpD protein that is known to contain a partial USP domain of 103 base pairs. The KdpD protein is a sensor and is associated with the cell membrane for its role as an osmosensitive K⁺ channel. Protein Rv1028c from *Mtb* is 860 amino acids and catalyses the following reaction: ATP + protein L-histidine = ADP + protein N-phospho-L-histidine. It contains a KdpD sensor kinase domain and is a member of the two component regulatory system KdpD/KdpE which controls the Kdp operon of genes *kdpFABC*. It is likely to function as a membrane associated protein kinase that is able to phosphorylate KdpE (Rv1027c) in response to environmental stimulus. Amino acids 256 to 377 is annotated as a USP domain region named 'USP_OKCHK'. The domain lies between the N-terminal sensor domain and the C-terminal catalytic domain. Important residues for the domain to be highlighted include 260-262, 290, 344-345, 347-348 and 358-361. The C-terminal domain contains the Histidine Kinase A (dimerisation/phosphoacceptor) region between amino acids 630 and 696. Also in the region 753-

The Universal Stress Proteins of Bacteria

853 is the histidine kinase-like ATPase region that contains an ATP binding site, Mg^{2+} binding site (amino acid 763) and a G-X-G motif. The protein is similar to others including the KdpD sensor protein from *E. coli* strain K12 (894aa). The two share 32.3% homology in 888 amino acid overlap.

This is interesting as Rv1636 was also shown to be present in the cell wall protein fraction when Rosenkrands *et al.* (2000) mapped *Mtb* proteins by 2D gel electrophoresis. In the co-immunoprecipitation experiments, the cell wall fraction would be pelleted as cell debris leaving the soluble cytosolic proteins present. Western blots on the cytosolic fraction did confirm the presence of the Rv1636 within the cytosol. Other studies by Mawuenyega (2005) showed Rv1636 to be located in the cell wall and cytosolic fraction of *Mtb* cellular proteins, as determined by protein profiling using 2DLC/MS. As the protein is in the cytosol this suggests it may be soluble, or alternatively may be released from the cell wall or membrane during disruption process (Mawuenyega *et al.*, 2003). If the protein was shown to have a scaffolding role in the case of Rv1636, it is unclear if the nucleotide binding motif plays a role in the scaffolding. If it is not involved, it is very possible that the putative nucleotide binding motif allows the proteins to have other functional roles within the bacteria.

4.3.8 Summary

Rv1636 appears to form higher oligomeric structures as a dimer and tetramer. The number of hydrogen bonds in the deposited crystal structure suggest the dimeric form of Rv1636 is the most stable. Production of the truncated constructs was successful and, in the case of the Rv1636_{MAKT} protein, disrupted total dimerisation of the protein. However, the Rv1636_{MAKT} protein also showed evidence of forming higher order structures. It remains unclear if nucleotide binding is dependent on the formation of higher oligomeric structures. The formation of high oligomeric states of USPs is likely to play an important role in the USP function. The UspC and KdpD role is the only evidence of two USP structures coming together and being linked to a direct function. The scaffolding is required to ensure phosphotransfer to the KdpE protein in response to the environmental signals. Any oligomeric states observed are likely to be important for the enzymes' biochemical activities and therefore relevant to the role that the USPs have in mycobacteria. The abundance of the common species of Rv1636 does not correlate with the NTPase activity suggesting, if the NTPase activity is a result of Rv1636 protein, the active form of Rv1636 may not be highly represented in the sample. Further experimental evidence is required to confirm the putative NTPase activity of Rv1636.

Chapter 5 - Intracellular Studies of Universal Stress Proteins

5.1 Introduction

Studying the biochemical properties of Rv1636 and the structure of the protein has aimed to provide an insight into the mechanisms of action of the USP family. To find a biological role for the USP *in vivo* would provide greater insight into the biological significance of the USPs and indeed the importance of any putative biochemical activities. The identification of interacting partners within the mycobacterial cell and the creation of a knockout Rv1636 *Mtb* strain are seen as important tools to study the *in vivo* importance of Rv1636.

During the course of this study, a phenotype was published by Max Schobert's group (Scheriber *et al.*, 2006) which defined the role of the USPs of *Pseudomonas aeruginosa*, and with particular reference to PA3309. This chapter details the work undertaken to link the biochemical activities described towards detailing a biological role of the USPs. The *in vivo* actions of the mycobacterial USPs are discussed, followed by an overview of the opportunistic pathogen *P. aeruginosa* and the USPs it encodes.

5.1.1 Mycobacterial USPs *in vivo*

The Williams group has studied the USPs of *Mtb* and the potential role these proteins undertake in the survival of the bacterium in response to stress. The study of the mycobacterial USPs has been challenging with extensive phenotype screens of single USP knockouts not highlighting any phenotypes in *Mtb*. Genetic knockouts created in the Williams' group include Rv2005c, Rv2026c, Rv2028c, and Rv2623. Phenotype screens using UV, mitomycin C, heat shock, cold shock, hypoxia and nitric oxide did not reveal any increased susceptibilities for any of these mutants. In addition, expressing these tandem-domain USPs also proved difficult and soluble protein was severely limited, preventing the biochemical characterisation of the tandem-domain proteins.

The Sasseti *et al.* (2003) TraSH screen for essential genes in *Mtb* did not suggest that any of USPs are essential. It is possible there remains a degree of redundancy between the proteins and to associate a particular phenotype with the USP family may require the study of double or even multiple knockout mutants. Figure 5.1 shows the USPs of *Mtb* and categorises them according to published gene expression results. The grouping of the USPs may provide ideas and possible routes that lead towards investigating the possible redundancy observed.

The Universal Stress Proteins of Bacteria

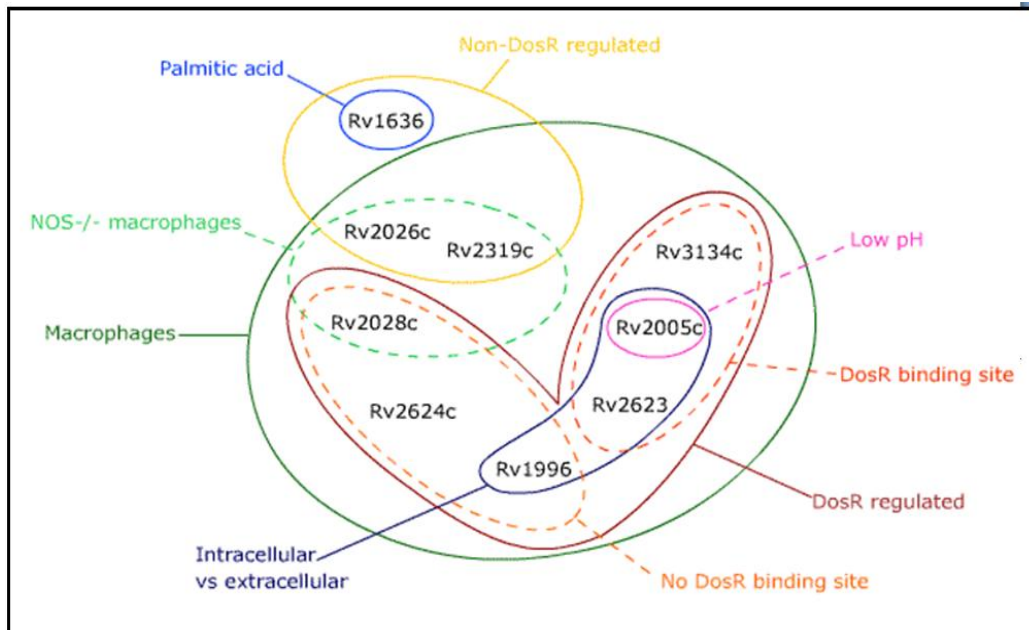


Figure 5.1. The regulation of the USPs of *Mtb*. The USPs have been grouped by similarities in upregulation. Data was gathered from published microarray studies. Diagram by K. Loughheed.

5.1.2 *Pseudomonas aeruginosa* USPs

5.1.2.1 *P. aeruginosa* Overview

The gram negative bacterium *P. aeruginosa* is an important opportunistic pathogen in humans and is ubiquitously found in soil, water and in most man-made environments. As an opportunistic pathogen, the bacterium causes a variety of nosocomial infections, wound infections and burn-related sepsis. Furthermore, *P. aeruginosa* infections are a significant cause of morbidity and mortality in cystic fibrosis (CF) patients (Lyczak *et al.*, 2000; Murray *et al.*, 2007). *P. aeruginosa* is an aerobic rod, although the bacteria can perform anaerobic growth using a variety of substrates including using nitrate as a terminal electron acceptor in the respiration chain. The bacilli can therefore survive in low oxygen conditions, such as within the lungs of CF sufferers.

The CF disease is caused due to an inherited mutation in the cystic fibrosis transmembrane conductance regulator (CFTR) gene which encodes a chloride ion (Cl⁻) channel of epithelial cells. Without the proper functioning ion channels, an ionic imbalance leads to the re-absorption of water from the lining of the lung, creating a viscous mucus layer. This thickened dehydrated mucosal layer prevents the mucocilliary escalation and compromises the lung's first line of defence against bacterial infection.

The Universal Stress Proteins of Bacteria

Within these low-oxygen mucus layers, biofilms of bacteria are formed and have been shown to establish a microaerobic environment to maximize growth (Sabra *et al.*, 2002; Worlitzsch *et al.*, 2002). The biofilms harbour a variety of species of bacteria and provide a protective environment for the *P. aeruginosa* against host immune response and antibiotics (Singh *et al.*, 2000). The CF lung is, therefore, characterised by a high level of bacterial infection. *P. aeruginosa* causes chronic infection in around 80% of CF sufferers and ultimately leads to respiratory failure and death.

Within the created niche in the CF lung mucus, *P. aeruginosa* is able to use nitrate or nitrite as alternative electron acceptors to permit anaerobic growth (Zumft & Korner, 1997). Van der Wauven *et al.* (1984) also showed that, when nitrate/nitrite were limited, arginine could also be used in this manner through substrate-level phosphorylation.

More recently, a separate pyruvate fermentation pathway has been identified as depicted in Figure 5.2 (Eschbach *et al.*, 2004). The fermentation does not permit anaerobic growth, but allows the bacteria to persist and was shown to be essential for long-term anaerobic stationary phase survival. Studying the regulation of the anaerobic survival by Max Schobert's group (Boes *et al.*, 2006; Schreiber *et al.*, 2006) showed that two USPs of *P. aeruginosa* were heavily upregulated. The two USP were genes *PA3309*, upregulated 72 fold, and *PA4352*, upregulated 22 fold.

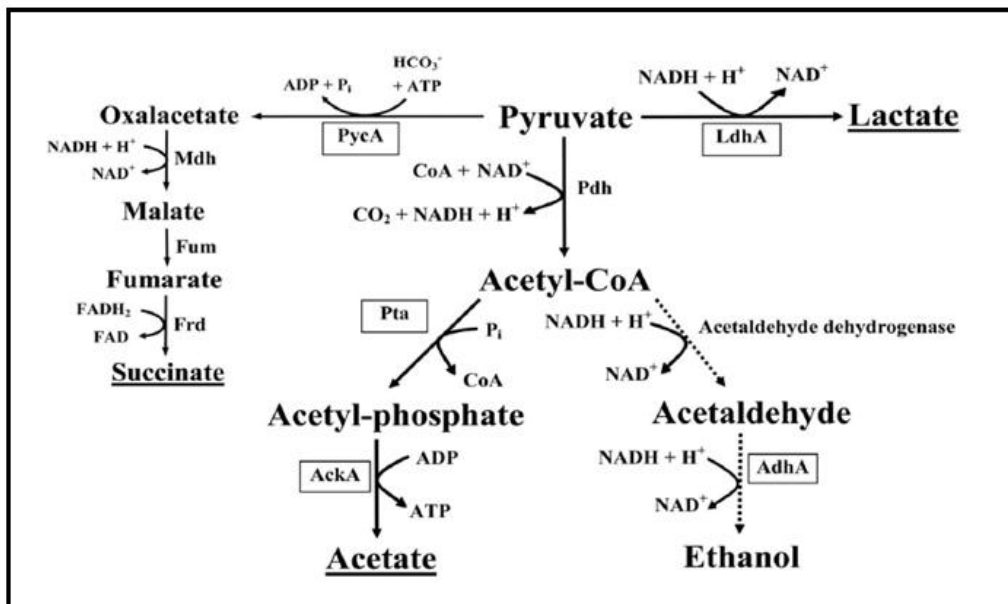


Figure 5.2. Pyruvate fermentation pathway model by *P. aeruginosa*. Full lines indicate identified fermentation routes; whilst dashed lines represent potential fermentation pathways. PycA, pyruvate carboxylase; LdhA, fermentative lactate dehydrogenase; Pdh, pyruvate dehydrogenase; Pta, phosphotransacetylase; AckA, acetate kinase; AdhA, alcohol dehydrogenase; Mdh, malate dehydrogenase; Fum, Fumarase; Frd, fumarate reductase. Figure by Eschbach *et al.* (2004).

5.1.2.3 The Universal Stress Protein PA3309

Through studying the anaerobic metabolic activities of the *P. aeruginosa*, two USPs were identified by Schrebier *et al.* (2006) to be potential virulence factors that allowed the harsh environment of the cystic fibrosis lung to be colonised. When Schrebier *et al.* (2006) published that USPs, PA4352 and PA3309, were involved in long-term anaerobic survival it was of interest to the work on *Mtb* USPs. The Williams research group already had experience of working with *P. aeruginosa* and had been investigating the chronic infections seen in CF patients with particular reference to the production of cyanide. Schreiber *et al.*, demonstrated that USP PA3309 is required for survival via pyruvate fermentation, however, the exact role remains unknown. Interestingly, the protein has also been shown to be present in anaerobic biofilms (Schreiber *et al.*, 2006).

An initial infecting *P. aeruginosa* isolate can infect a patient's lung for 20 years or more. The long term infection by *P. aeruginosa* mirrored the survival of *Mtb* within the lung in a way that allowed latent infections to be maintained over extreme long periods of time. It could, therefore, be postulated that USP PA3309 was important for the survival of *P. aeruginosa* within the cystic fibrosis lung. PA3309 had not been studied biochemically and the only information on the protein was the survival role described. Sequence analysis showed the PA3309 protein contained the nucleotide binding motif previously highlighted in the *Methanococcus* USP MJ0577. The protein was therefore suitable to study the biochemical role of the nucleotide binding domain and to see if the nucleotide binding domain was linked to the ability of *P. aeruginosa* to survive through pyruvate fermentation.

The biochemical function of USPs, especially those containing the G-X₂-G-X₉-G(S/T) nucleotide binding motif, is likely to be related even if the regulation and overall function may differ between USPs within and between different organisms. Working with *P. aeruginosa* aimed to confirm any findings with Rv1636 in a separate USP protein from a distinct organism. With biofilm models already established in the laboratory the potential models for study are readily available, and these may provide excellent opportunities to further the biochemical studies of the PA3309 and USPs in general.

5.1.2.4 PA3309 Function

PA3309 plays an important role in anaerobic growth, shown by the 100-fold decrease of the growth PA3309 knockout strain compared to the wild-type, when cultured in the absence of nitrate,

The Universal Stress Proteins of Bacteria

nitrite or pyruvate (Schreiber *et al.*, 2006). The protein Anr, an oxygen-regulating protein, was shown to be responsible for the induction of the PA3309 protein (Schreiber *et al.*, 2006), which was independent of pyruvate. PA3309 was also shown to have increased expression in deep layers of biofilms when *P. aeruginosa* was grown on nitrate. The promoter induction in this case was shown to be independent of Anr in the same study.

Under long term aerobic and anaerobic stationary phase conditions, the induction of the PA3309 protein was shown to be independent of Anr, and did not involve a range of other common regulators of promoter activity including RelA, RpoS, GacA and quorum sensing regulators RhIR and LasR (Schreiber *et al.*, 2006). The regulation of PA3309, therefore, may provide clues to the role of the USP's and the biochemical function these proteins possess. The fact that RelA is an important regulator of *E. coli* UspA (Kvint *et al.*, 2000) and that GacA shares similarities to DevR of *M. smegmatis* indicates the regulation of the USPs across the different organisms may vary considerably (Schreiber *et al.*, 2006).

5.1.3 Aims

This chapter describes studies to link the biochemistry and structure, all with the aim of finding an *in vivo* biological role for USPs. USPs *in vivo* may be redundant as there are often many encoded within bacterial genomes and single mutations in several USP's in *Mtb* made by the Williams group revealed no observable phenotypes. The retention of many USPs within the genomes of a range of bacteria suggests the presence of USPs may have been beneficial in their evolutionary history. Rv1636 is the sole single domain USP in *Mtb* and, therefore, may have a distinct role separate from the other USPs found in *Mtb*; it is also worth noting that Rv1636 is the only USP retained in the reduced genome of *M. leprae*. Knocking out the gene will be attempted by continuation of a suicide vector protocol started by Leiva (2007).

In order to study the action of Rv1636 *in vivo*, the genomic context of Rv1636 and the homologues in mycobacteria were also considered. Protein interactions were then studied and over-expression of the proteins within mycobacteria was attempted. Later in the chapter, the *P. aeruginosa* gene PA3309 is discussed including work to determine if the predicted nucleotide binding motif present is involved in the anaerobic pyruvate survival described by Schreiber *et al.* (2006). This was to be investigated through complementation of the PA3309 knockout with a site directed mutant in the putative nucleotide binding motif of PA3309.

5.2 Results

5.2.1 Rv1636 Homologues

Phylogenomic analysis of Rv1636 using Uniprot/TrEMBL shows the family members of Rv1636 are exclusively from members of the *Mycobacterium* genus. *Mtb* Rv1636 has clear homologues within other mycobacteria ranging from 100% identities in *M. bovis* to 78% for the homologue in *M. smegmatis* (Table 5.1). A sequence alignment is shown in Figure 3.2 in Chapter 3. Figure 5.4 shows the evolution of the protein cluster CLSK791295 as demonstrated by a fastest minimal evolution phylogenetic tree made using NCBI phylogenetic tree viewer. The Rv1636 is clearly the most related USP between the members of the *Mtb* complex and is highly conserved between different strains of *Mtb*. Interestingly, *M. smegmatis* also contains the protein MSMEG_3308 that is 160 amino acids in length and has 52% shared homology.

Table 5.1. Homologues of Rv1636 across *Mycobacterium* species

Organism	Homologue	Identity
<i>M. tuberculosis</i>	Rv1636	100%
<i>M. bovis</i>	Mb1662	100%
<i>M. kansasii</i>	MkanA1_01785	94%
<i>M. marinum</i>	Mmar_2440	93%
<i>M. leprae</i>	MIbr_01390	89%
<i>M. avium</i>	MaviaA2_13103	88%
<i>M. intracellulare</i>	MintA_10486	85%
<i>M. smegmatis</i>	Msmeg_3811	78%

The Universal Stress Proteins of Bacteria

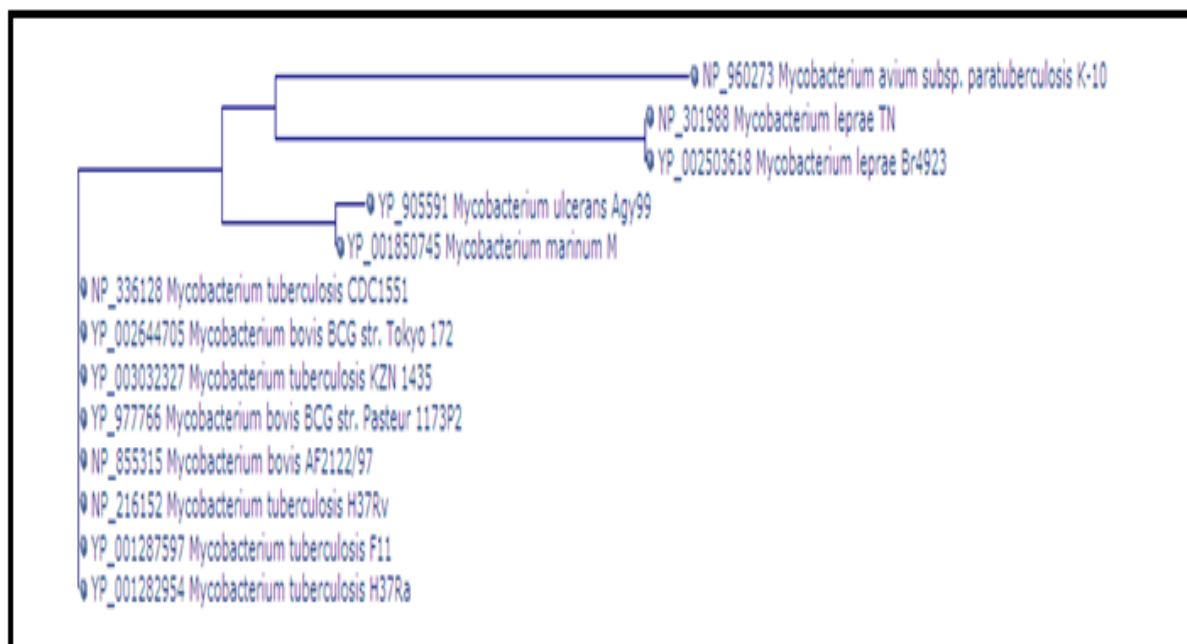


Figure 5.4. Phylogenetic tree view of the fastest minimal evolution of Rv1636 and the homologues which form protein cluster CLSK791295. Distance method used mPAM-20. Constructed using NCBI phylogenetic tree viewer. Accession numbers are indicated.

Within the order *Actinomycetales*, which contains the sub orders *Corynebacterineae*, *Streptomyceinae*, *Pseudonocardineae*, *Streptosporangineae* and *Frankineae*, there are related proteins with approximately 50% identities to the Rv1636 protein of *Mtb*, as shown by BLAST analysis (data not shown). Within the suborder *Corynebacterineae*, which contains the Family *Mycobacteriaceae*, consisting of the sole genus *Mycobacterium*, there are a range of proteins which have significant similarities to Rv1636 with up to 65% homology (Table 5.2). The protein Rv1636 is highly conserved and was clearly present earlier in the evolutionary past within the order.

Table 5.2. Related proteins to Rv1636 within the *Corynebacterineae* suborder.

Suborder	Family	Genus	Species	Protein	Homology (Positives)
<i>Corynebacterineae</i>	<i>Mycobacteriaceae</i>	<i>Mycobacterium</i>	<i>Mycobacterium tuberculosis</i>	Rv1636	100% (100%)
<i>Corynebacterineae</i>	<i>Nocardiaceae</i>	<i>Rhodococcus</i>	<i>Rhodococcus erythropolis PR4</i>	RER_33060	64% (81%)
<i>Corynebacterineae</i>	<i>Gordoniaceae</i>	<i>Gordonia</i>	<i>Gordonia bronchialis</i>	Gbro_2864	65% (79%)
<i>Corynebacterineae</i>	<i>Tsukamurellaceae</i>	<i>Tsukamurella</i>	<i>Tsukamurella paurometabola</i>	Tpau_2426	57% (77%)

The Universal Stress Proteins of Bacteria

The genomic context of the Rv1636 homologues does not provide further clues to the role of the protein *in vivo*, except a possible role in UV radiation protection. In *Mtb*, the protein is flanked in proximity to the *UvrA* and *UvrB* genes, which are involved in protection against UV radiation. The *E. coli* USPs subtype have been shown to be induced by UV light; however, Rv1636 was not highlighted as having strong upregulation following exposure to exposure to UV light (Boshoff *et al.*, 2003). It is possible the *UvrA* and *UvrB* genes may have protected Rv1636 and its homologues as mutations or deletions to Rv1636 could have influence the neighbouring genes. *UvrA* and *UvrB* are upregulated in response to UV induced DNA damage and are part of the nucleotide excision pathway.

Comparisons of the genomic context of the homologue Rv1636 within the mycobacteria are shown in Figure 5.5. The genomic location is conserved in *Mtb*, *M. bovis* and *M. leprae* however difference in transcription direction of neighbouring genes are seen in *M. marinum* and *M. smegmatis*. In *Mtb*, the USPs are scattered seemingly randomly across the chromosome with neighbouring genes providing no clues to the function of the proteins. Adjacent genes either have varying functions or are hypothetical proteins with unknown function.

The Universal Stress Proteins of Bacteria

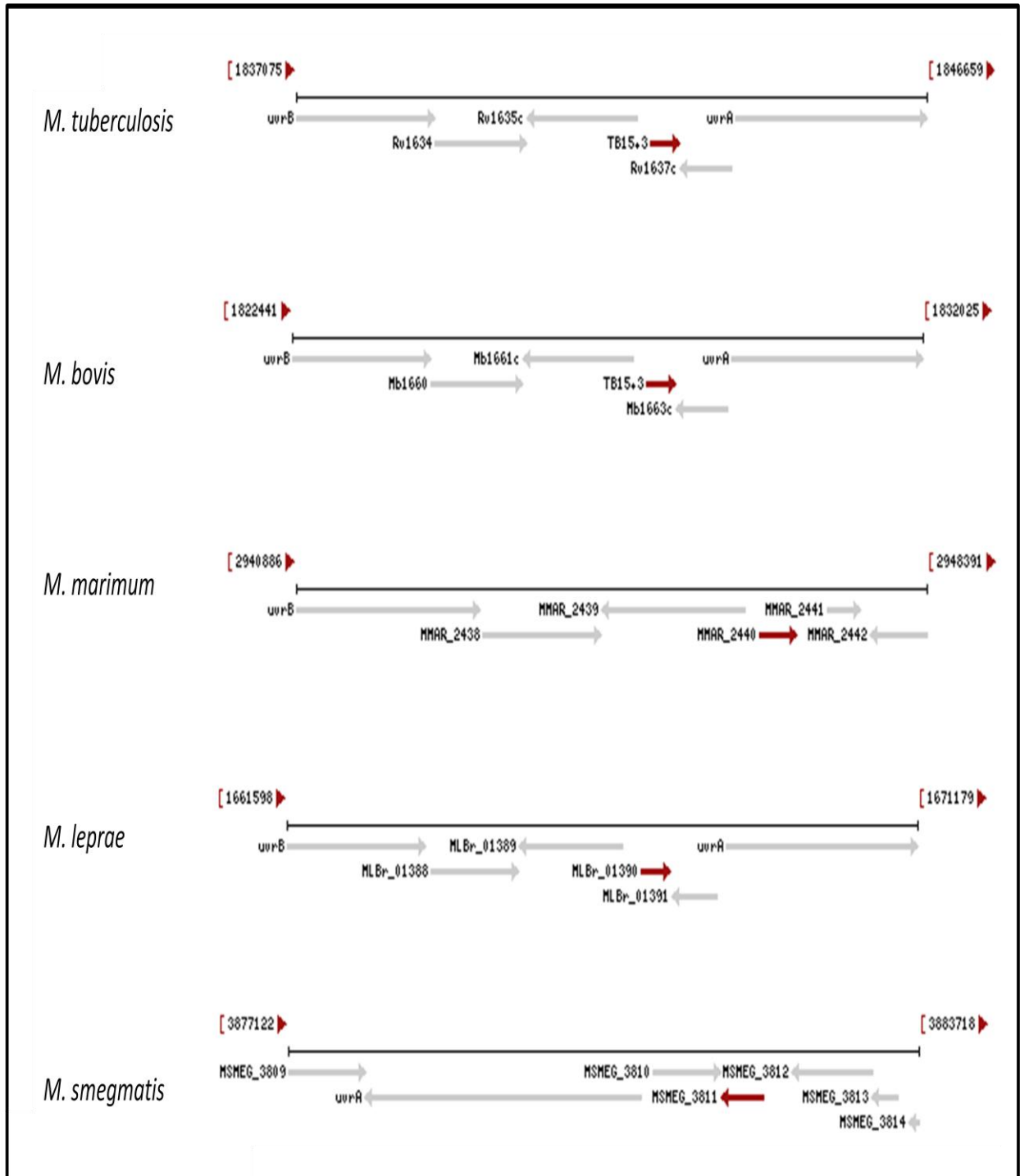


Figure 5.5. Genomic context of *Rv1636* homologues in selected *Mycobacterium* species. Homologues are shown with red arrows. Arrows indicate transcriptional direction. *Mtb* (*Rv1636*), *M. bovis* (*Mb1662*), *M. marimum* (*MMAR_2440*), *M. leprae* (*ML1390*) and *M. smegmatis* (*MSMEG_3811*) genomic contexts are all shown.

5.2.2 Creation of a *Mtb* Rv1636 Mutant

Previous work by Leiva (2007) involved using a suicide vector system to attempt to create a *Rv1636* knockout mutant in *Mtb* H37Rv. The pSMT100 vector system (Dussurget *et al.*, 2001) relies on the presence of the *hygromycin resistance* gene with *sacB* gene as a counterselective marker (Figure 5.6). Using this system, single crossovers were selected by hygromycin resistance and confirmed by PCR analysis. Work continued from this point in order to isolate double cross-over mutants.

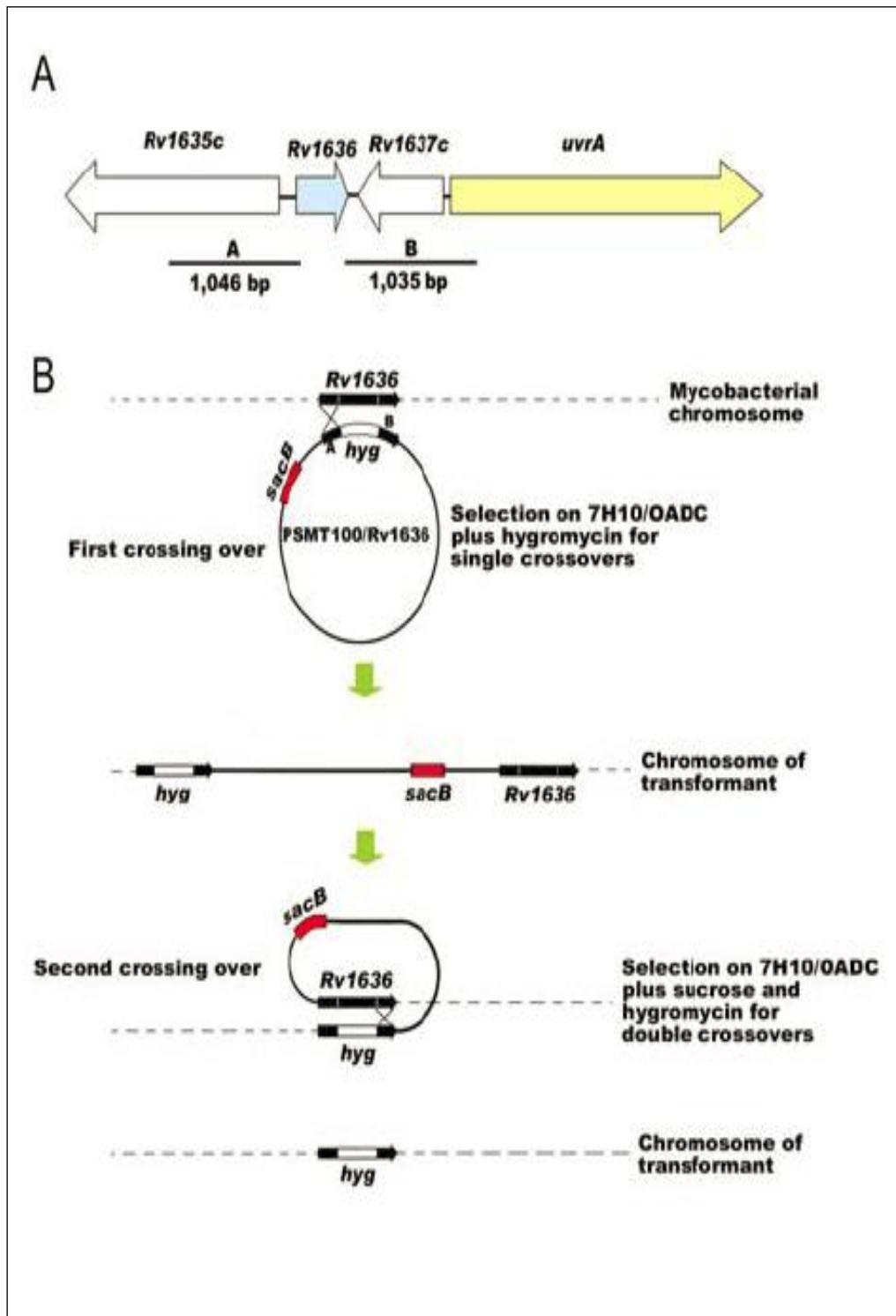


Figure 5.6. General strategy for mutagenesis of *Rv1636* gene of *M. tuberculosis*. (A) 1,046 and 1,035 bp fragments, upstream and downstream of *Rv1636*, respectively, were cloned around a hygromycin resistance gene in the pSMT100 vector. (B) *Rv1636* gene replacement is achieved through a two-step method using *SacB* as counter selective marker allowing growth on sucrose (Levia, 2007).

The Universal Stress Proteins of Bacteria

These mutants were selected on sucrose for double crossovers, selecting for cells which have lost, via homologous recombination, the *sacB* gene-containing plasmid backbone, leaving the hygromycin cassette in the place of the *Rv1636* gene. From the confirmed single crossovers, over 50 colonies were sub-cultured and selection for double crossovers performed with hygromycin and sucrose. Genomic DNA was isolated from the resulting colonies and tested both for the presence of the hygromycin cassette and the presence of the *Rv1636* gene. The primers used to screen the mutant (see Chapter 2.1.6) include Hyg primers which amplify an 800 bp fragment coding for hygromycin resistance, and *Rv1636* primers, which amplify the *Rv1636* gene to produce a fragment of approximately 450 bp.

It was confirmed that none of the 50 colonies screened carried the *Rv1636* complete knockout. The colonies on the plate were small, suggesting that they were still single crossovers containing the *sacB* gene. More suitable primers would have spanned the *Rv1636* gene and produced a PCR product of different sizes depending upon whether the double crossover had occurred to avoid relying on a negative result. Work is currently underway to knockout the wild type copy of *Rv1636* in a merodiploid strain containing a second copy of the gene to take into account the possibility that *Rv1636* is essential.

5.2.2.1 Macrophage Infection with Putative *Mtb* H37Rv *Rv1636* Mutants Reveals a Possible Survival Defect Phenotype

Before the screening of potential double crossovers detailed above had been performed, putative *Mtb* mutants, were tested for their ability to survive in murine J774 macrophages. Following infection, samples were taken at 0, 24, 48 and 72 hours. The intracellular *Mtb* bacilli were then plated out after removal of extracellular bacilli and macrophage lysis. The results of the experiment (Figure 5.7) show the putative *Rv1636* mutants have a potential survival defect compared to the wild type and also a *Mtb* H37Rv mutant in the USP *Rv2623* gene. Unfortunately the mutants were unable to be confirmed as double crossovers, indicating that the infection had been performed with single crossovers. The experiment was not able to be repeated with a confirmed knockout of the *Rv1636* gene. These studies are ongoing, but this preliminary data suggests that $\Delta Rv1636$ may exhibit a survival defect and as such this USP may be of importance in intracellular survival within the host macrophage.

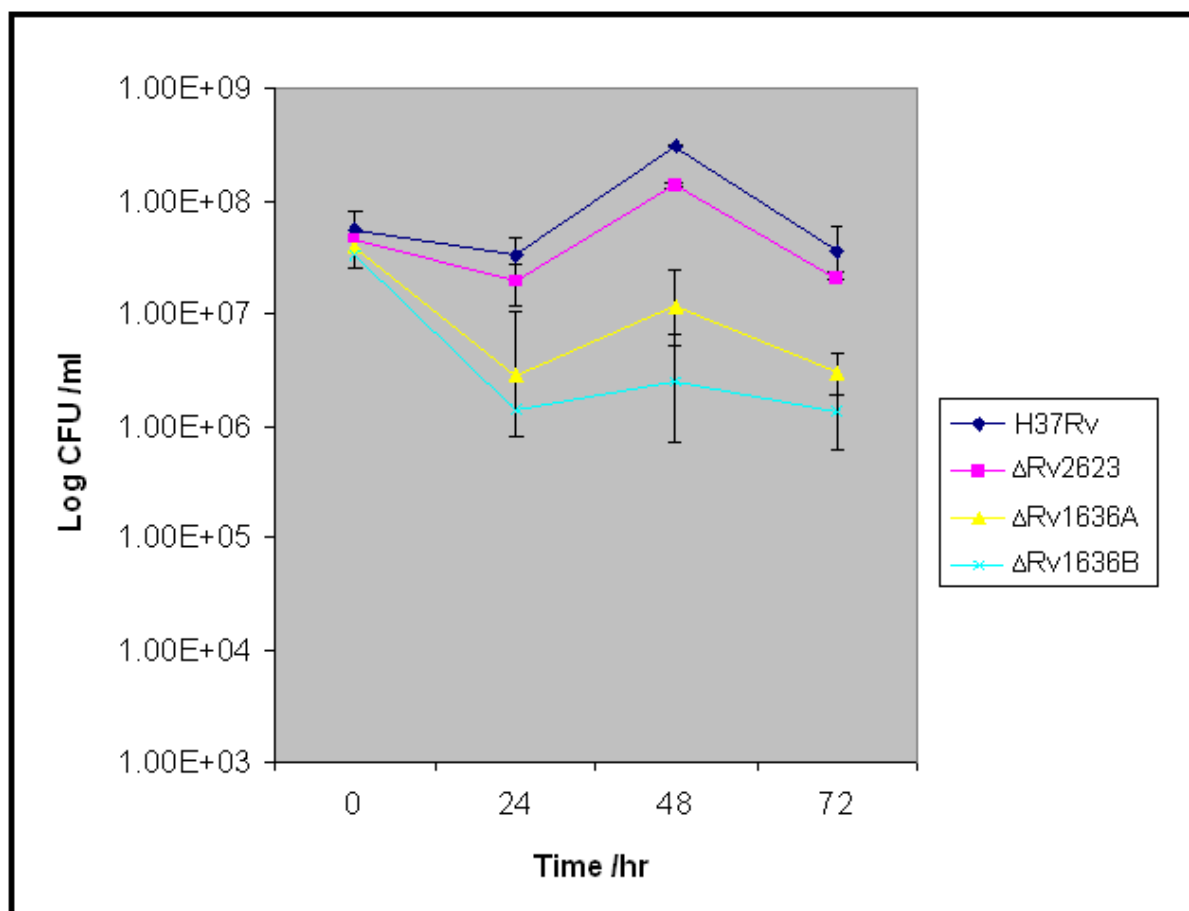


Figure 5.7. Survival curve of intracellular bacilli from a range of *Mtb* mutants as indicated. Strains included WT (H37Rv), mutant *Rv2636*, putative mutants *Rv1636 A* and *B*. Murine J774 macrophages (plus interferon γ) were infected with the bacilli and the survival of intracellular bacilli assayed by plating out the bacteria and calculating the cfu/ml.

5.2.3 Identification of Proteins Interacting with Rv1636

USPs have been shown to have interactions with other proteins such as the *E. coli* UspG with GroEL (Bochkareva *et al.*, 2002), UspC with KdpD (Heermann *et al.*, 2009a) and USP1 of *Lactobacillus plantarum* that can interact with transcriptional regulators involved in acid stress response (Gury *et al.*, 2009). Protein interactions within cells can provide large insights into the role of a protein. To begin the work, a database search was performed using the String Database. In order to further investigate the Rv1636 interaction, GST-pull-downs and co-immunoprecipitation reactions were performed. Interaction studies were performed with *M. bovis* BCG as it contains the 100% identical

The Universal Stress Proteins of Bacteria

homologue to Rv1636, Mb1662, and can be used under Category II conditions thus making it a suitable alternative to *Mtb* to investigate the cellular interactions.

5.2.3.1 Database Search

The String Data Base shows all known and predicted protein-protein interactions for individual proteins. Searching the database for Rv1636 (Figure 5.8) and the homologues, Mb1662 from *M. bovis* and MSMEG_3811 from *M. smegmatis*, showed that little is known about the interactions of these proteins, with any links shown being tentative links. Rv1636 was shown to have concurrence with the USP *Rv2026c* (*MT2085* gene) as homologous genes from *Roseiflexus castenholzii* are genomic neighbours.

Similarly *Methanothermobacter thermoautotrophicus* has homologues of the genes *Rv1636* and *Rv2642*, a DNA binding transcriptional regulator protein, that are found adjacent to each other in the *M. thermoautotrophicus* genome. *M. thermoautotrophicus* is an anaerobic lithoautotrophic thermophilic archaeon that is often found in biodegrading facilities and only requires carbon dioxide, hydrogen and salts to grow. Growth occurs at temperatures ranging from 40 to 70 °C and is optimal at 65 °C. The neighbouring genes to *Rv1636*, which include acetylglutamate kinase (*argB*), the probable arsenic transport protein (*argC*), *UvrA* and *UvrB*, have not been linked to the role of *Rv1636* as shown in Figure 5.8.

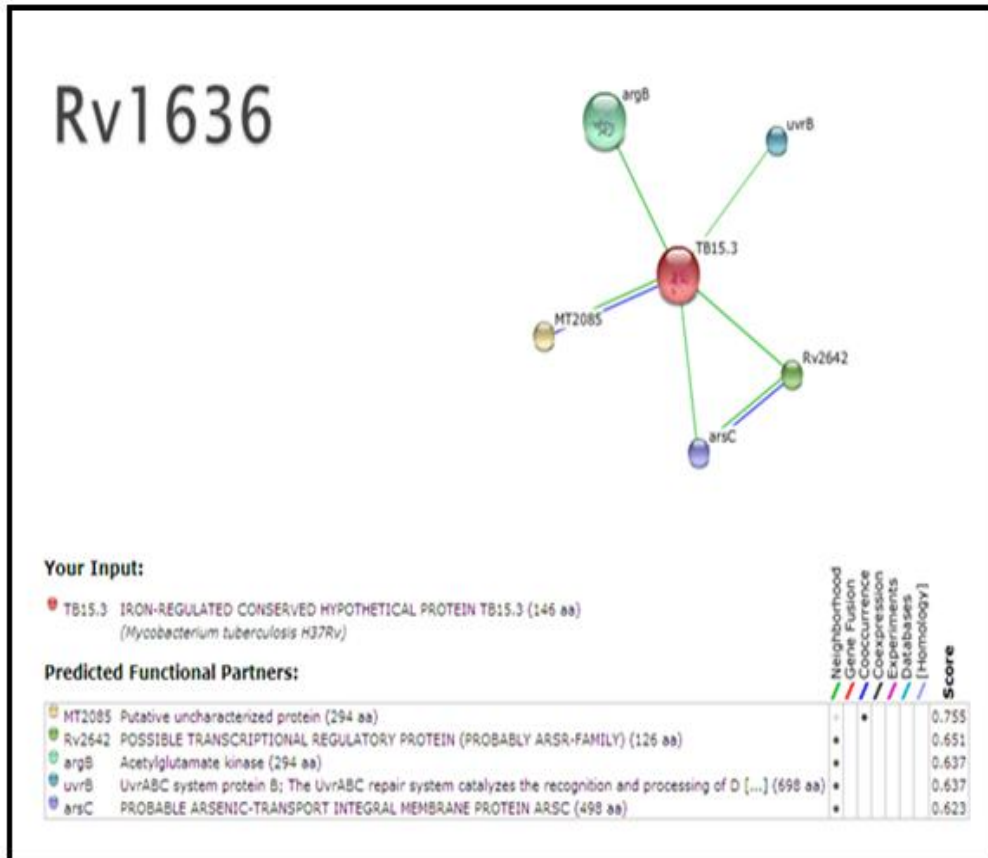


Figure 5.8. String database search for predicted protein-protein interactions of Rv1636. Predicted functional partners highlighted include MT2085, Rv2642, argB, UvrB and argC. No genes were highlighted as being co-expressed.

5.2.3.2 Rv1636 Expression

Interaction studies were performed in *M. bovis* BCG. To ensure that this was a suitable model for *Mtb*, the expression of Mb1662 in *M. bovis* BCG was confirmed by performing Western blots using anti-His6-Rv1636_{WT} antibodies. USPs are highly conserved between *M. tuberculosis* and *M. bovis* BCG, with both containing the ten USPs. *M. bovis* BCG is often used as a model organism instead of the pathogen *M. tuberculosis* because *M. bovis* BCG is safer, meaning the project can be undertaken in a containment level two laboratory. *M. bovis* BCG cells were cultured for up to 12 days in order to obtain cytosolic fractions. Growth curves were plotted by recording the optical density of the cultures in a time course (Figure 5.9).

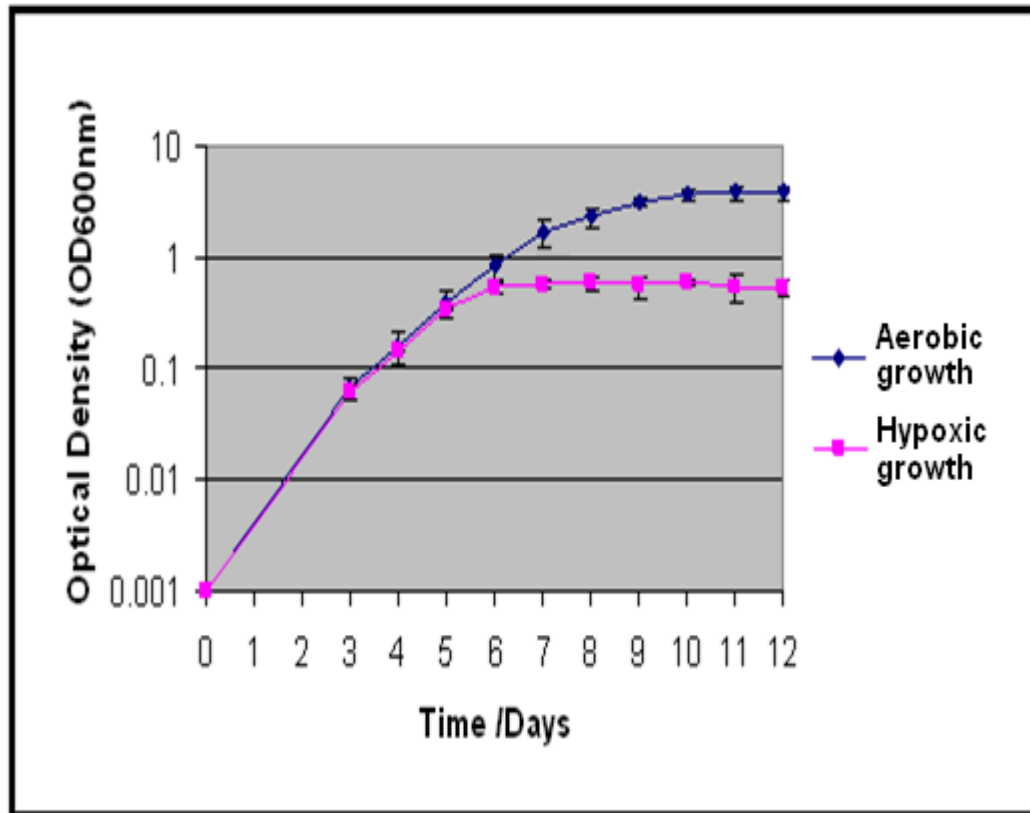


Figure 5.9. Growth curves for *M. bovis* BCG grown at 37°C in Middlebrook 7H9 media in aerobic and hypoxic conditions. Error bars show standard deviation. Hypoxic cultures were grown in conical flask with suba seal bungs resulting in a gradual decrease in available oxygen. Optical density at 600 nm was recorded.

Western blots on cell protein extracts were performed using an anti-His₆-Rv1636_{WT} antibody. Western blot analysis was performed with 15 µl of the cytosolic fraction of *M. bovis* BCG cells in stationary phase (day 12/OD₆₀₀ 3.5), mid-exponential phase (OD₆₀₀ 0.5) and oxygen starved stationary phase. Figure 5.10 shows the results of the Western blot on the cytosolic lysates. The Western blot indicates that Mb1662 is up-regulated when the *M. bovis* BCG cells enter stationary phase from mid-exponential phase (lanes 2 and 3). Stationary phase cytosolic lysates were used to investigate the protein-protein interactions of Rv1636. It should be noted that the Mb1662 protein (15.3kDa) is smaller than the His tagged-Rv1636 version.

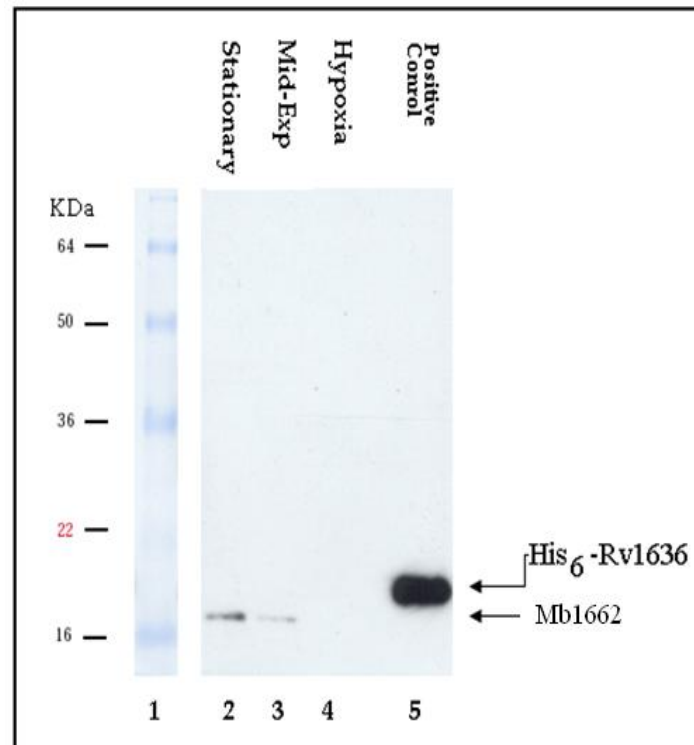


Figure 5.10. Western Blot using anti-His₆-Rv1636 antibody against samples of cytosolic fractions produced from *M. bovis* BCG cells in different growth conditions. Lane 1: SeeBlue Pre-stained 2 Marker Lane 2: Stationary phase cytosolic fraction. Lane 3: Mid-exponential phase cytosolic fraction. Lane 4: Oxygen starved stationary phase cytosolic fraction. Lane 5: His₆-Rv1636 (1 µg) as a positive control. In order to compare the level of expression cytosolic fractions were diluted to 1.5 mg/ml.

5.2.3.3 GST Pull-down of Rv1636

Pull-down assays are based on the methods used for affinity purification of proteins. The method is similar to immunoprecipitation except that, in pull-down assays, a bait protein is used instead of an antibody. The use of a tagged bait protein allows the protein to be captured by immobilized affinity, thus generating a secondary affinity support that allows the purification of additional proteins that interact with the bait protein. The GST-Rv1636 fusion protein was previously generated in the laboratory (Leiva, 2007) and had been purified by affinity chromatography using glutathione agarose. The GST-tag is 26kDa.

The GST-fusion protein was added to cytosolic fractions produced from *M. bovis* BCG and after incubation the GST protein was pulled down by the GS4B beads. Any protein interacting with the GST-Rv1636 would also be pulled down by the beads. Different methods to release the pulled down proteins from the GS4B beads were tested. Boiling the beads in 2X Laemmli and loading the

The Universal Stress Proteins of Bacteria

sample produced a lot of background and made identifying a pulled down partner difficult. An elution buffer containing glutathione allowed the pulled down protein to be released gently and resulted in a sample with little background (Figure 5.11).

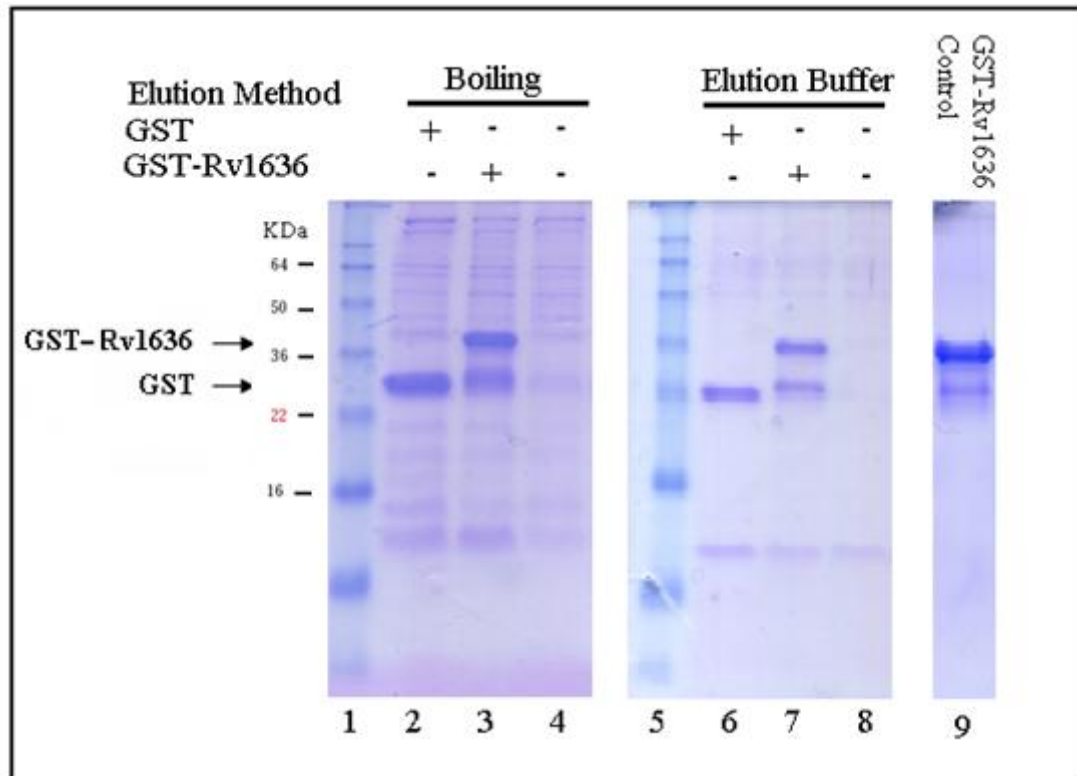


Figure 5.11. GST-pull-down reaction with GST-Rv1636. Cytosolic fraction of *M. bovis* BCG grown in stationary phase was used. Pull-down visualised on 15% SDS-PAGE by staining with Imperial Coomassie stain. Lane 1 and 5: Pre-stained SeeBlue Marker, Lane 2 and 6: GST pull-down, Lane 3 and 7: GST-Rv1636 pull-down, Lane 4 and 8: Negative Control (no GST added), Lane 9: Purified GST-Rv1636 (20 µg).

Numerous GST-pull-downs were performed on the different cytosolic fractions and under different conditions. Following experimental evidence that Rv1636 may be a GTPase, pull-downs were repeated in the presence of nucleotides such as GTP, ATP, and GDP and in the presence of different cations including Mg^{2+} , Mn^{2+} and Ca^{2+} . Pull-downs were performed using a range of concentrations of cations and nucleotides; however, no partner was found in any of the tested conditions. Some proteins were found that were able to bind non-specifically to the beads only in the presence of GTP. In oxygen starved condition proteins bound non-specifically to the beads that were not present in the stationary phase cytosolic fraction.

Staining of the gels was important because the more sensitive the staining the more likely that any partner pulled down would be visualised and identified. The Imperial Coomassie provided

The Universal Stress Proteins of Bacteria

extra sensitivity than standard Comassie stain (<3 ng of protein could be visualised). However, when silver staining was used, it was apparent that the method provided greater resolution and it was decided that this was the method of choice. Figure 5.12 shows a silver stained gel with pull-down reactions using cytosolic fractions with added Mn^{2+} and GTP. Unfortunately, no interacting partner of Mb1662 was found.

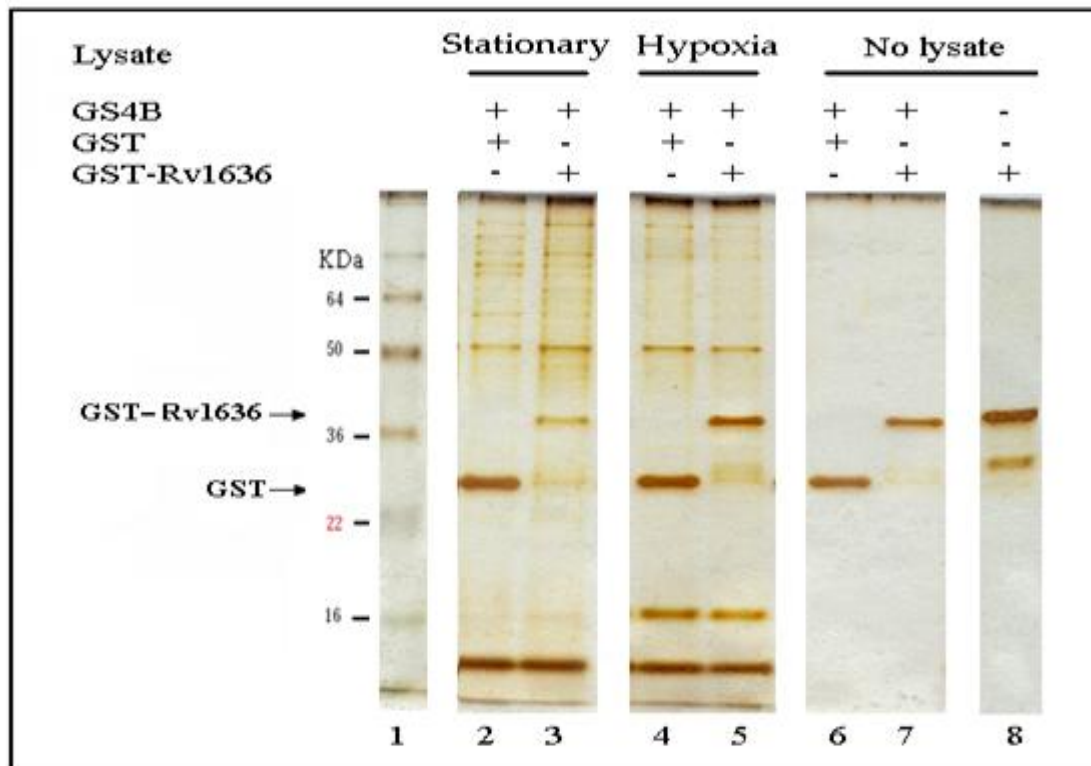


Figure 5.12. GST-pull-down reaction of GST-Rv1636 from cytosolic fractions of *M. bovis* BCG grown in stationary phase (lane 2 and 3), hypoxia (lanes 4 and 5) and buffer control (lanes 6 and 7). Lane 1: Pre-stained SeeBlue Marker, Lane 2, 4 and 6: GST controls, Lane 3, 5 and 7: GST-Rv1636 pull-down, Lane 8 Purified GST-Rv1636 protein. No interacting partner was found. 15% SDS-PAGE gel stained by silver staining.

5.2.3.4 Co- immunoprecipitation of Rv1636

GST-pull-downs did not highlight any pull-down partners. To continue studying the protein-protein interactions, co-immunoprecipitation assays were performed. Antibodies against His₆-Rv1636 were added to the cytosolic fractions and, after an incubation to allow their specific binding, beads were added that were able to bind the antibodies, thus allowing the precipitation of Mb1662

The Universal Stress Proteins of Bacteria

and any interacting protein if present. Initial experiments precipitated high levels of background and silver staining did not identify the immunoprecipitated Mb1662. It was very unlikely that any partner co-immunoprecipitated would be visible if Mb1662 was not visible.

In order to confirm that Mb1662 was being precipitated, a different ratio beads/antibody and two different beads, Protein A plus G and Protein A, were trialled. Protein A plus G can bind a wider range of antibodies, whereas Protein A binds specifically to the Fc region of immunoglobulin molecules. Protein A is especially strong binding to IgG of rabbit antibodies and is therefore more specific to the antibodies being used to precipitate Mb1662. Western blots were performed to determine if Mb1662 was being precipitated. The result of the Western blot analysis showed that numerous bands were present in lanes. The specificity of the Western blot was low and in order to increase the specificity for future Western blots only 0.5 µg of primary antibody was used in the PBS-T/milk solution. A band at 55 kDa was a result of the secondary antibody binding to the heavy chain of the antibody and producing a positive signal.

Instead of using Western blots to determine if Mb1662 was being precipitated, control reactions were performed using beads, anti-His₆-Rv1636 antibody and purified His₆-Rv1636 in 750 µl of breaking buffer. The reactions were performed to test the ability of the antibodies to precipitate the Mb1662 protein from a solution. The combination of variables that gave better results was using 20 µl Protein A beads with an antibody dilution of 1 in a 100 for the total reaction. Further experiments showed that up to 5 µg of the anti-His₆-Rv1636 antibody could pull-down a suitable amount of Mb1662 protein (Figure 5.13). The silver stained SDS-PAGE showed the presence of the heavy and light chain of the antibody are 55 kDa and at 25 kDa respectively. The beads alone when boiled and subjected to SDS-PAGE produced a number of bands (lane 2).

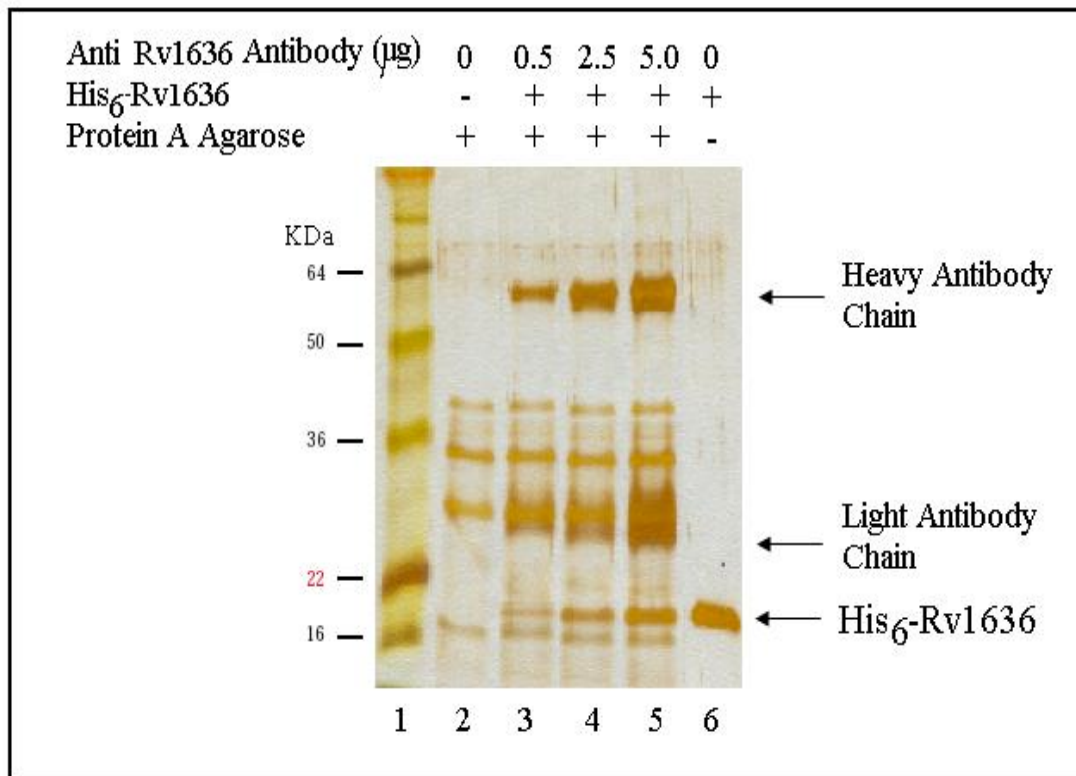


Figure 5.13. Immunoprecipitation of His₆-Rv1636_{WT} protein using protein A agarose beads and different amount of anti-His₆-Rv1636_{WT} antibody. Lane 1: Marker Pre-stained Seeblue 2. Lane 2: Protein A agarose beads boiled with 2X Laemmli and the supernatant loaded, Lanes 3-5: Immunoprecipitation of His₆-Rv1636_{WT} by different levels of antibody, Lane 6: Purified His₆-Rv1636_{WT} (5 μ g) as a marker. 12% SDS-PAGE gel stained by silver staining.

Co-immunoprecipitation reactions were performed on cytosolic fractions from stationary phase, mid-exponential phase and hypoxia stationary phase (Figure 5.14A). The results showed that Mb1662 was precipitated from the stationary and mid-exponential phase but not hypoxia conditions. No co-immunoprecipitating partner could be found with visual examination of the silver stained gel. Figure 5.14B shows a Western blot that was performed on the same samples to confirm the precipitation of Mb1662. The result showed Mb1662 was precipitated as expected from the stationary and mid-exponential cytosolic fractions. Other background includes heavy antibody chains, proteins from the beads and also secondary antibodies reacting non-specifically with BSA (Bovine serum albumin) which is present in the agarose solution.

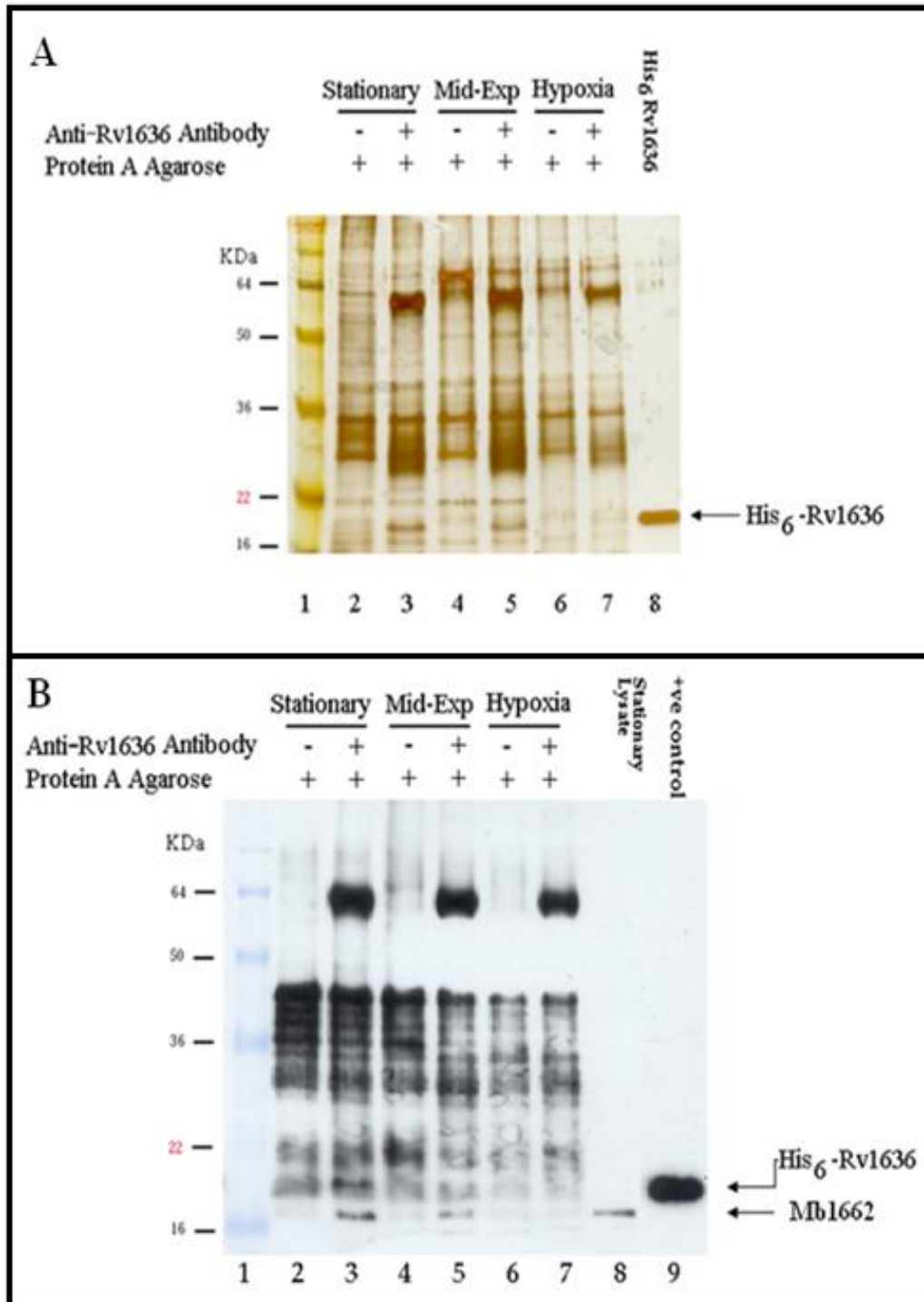


Figure 5.14. Co-immunoprecipitation of the Mb1662 protein from *M. bovis* BCG cytosolic fractions produced from cells grown under stationary phase (lanes 2 and 3), mid exponential phase (lanes 4 and 5) and hypoxia stationary phase (lanes 6 and 7). Protein A agarose beads and 5 µg anti-His₆-Rv1636_{WT} antibody were used. Lane 1: Marker Pre-stained Seeblue. (A) 12% SDS-PAGE gel stained by silver staining. Lane 8: Purified His₆-Rv1636_{WT} (5 µg) as a marker. (B) Western blotting with anti-His₆-Rv1636_{WT} antibody. Lane 8: Stationary phase cytosolic fraction, Lane 9: Purified His₆-Rv1636_{WT} (1 µg) as a positive control.

5.2.4 Over-expression of Rv1636

It was hypothesised that over-expression of Rv1636 would provide an insight into the role of the protein and also allow any knockouts, if generated, to be complemented. Over-expressing the protein can also give further scope to study the protein-protein interactions as over-expressing the protein would result in a higher level of 'native' protein. A range of over-expression systems of mycobacteria have previously been published in the literature including mycobacterial promoters upregulated by acetamide (ACE) (Dziadek *et al.*, 2003), tetracycline inducible (pMIND) (Blokpoel *et al.*, 2005) and heat shock promoters (HSP) (Stover *et al.*, 1991).

5.2.4.1 Cloning of Rv1636 into Over-expression Vectors

Rv1636 gene was cloned into the expression vectors pMV216, pMV361 and pSE100. Cloning the *Rv1636* gene into a number of systems aimed to provide a variety of options to allow the over-expression of Rv1636_{WT}. pMV261 and pMV361 were originally described by Stover *et al.* (1991). pMV261 is a non-integrating plasmid with an origin of replication in both *E. coli* and mycobacteria and pMV261 allows expression of the selected protein under the control of the constitutively expressed promoter from the mycobacterial heat shock protein, Hsp60. pMV361 contains the attP site which allows integration into the mycobacterial attB site and also encodes the integrase gene (*int*) to allow insertion. Protein expression is driven by the GroEL promoter in this vector. Both pMV261 and pMV361 encode the kanamycin resistance gene, which for the integrating pMV361 can be excised from the mycobacterial genome using the flp recombinase. The expression vector pMV261 is a self-replicating episomal plasmid at 3-5 copies per cell and offers high levels of over-expression.

The pSE100 vector, originally described by Schnappinger *et al.* (2007), allows controlled expression of target protein due to the strong (pMC1s) and medium (pMC2m) tet repressors, which integrate into the *Mtb* genome at the attB site, allow finer control of expression depending on the level of tetracycline used. The episomal pSE100 expression is driven by pmc1tetO which is a tet operator-containing promoter for mycobacteria.

5.2.4.1.1 Cloning into pMV261

Rv1636 was cloned from pGem-Teasy-*Rv1636* using primers pMV261-*Rv1636* For/Rev (Chapter 2.1.6). The gene was cloned between the *Bam*HI and *Hind*III sites and then inserted

The Universal Stress Proteins of Bacteria

directionally in frame and downstream of the Hsp60 promoter (P_{Hsp60}) of the multi-copy plasmid pMV261 to produce pMV261::Rv1636. pMV261 adds the amino acid residues MAKTIADP to the start of the Rv1636 protein. To ensure the correct cloning had taken place pMV261 vectors were sequenced across the Rv1636 gene with the pMV261 sequencing primer 5'-GGCATAGGCGAGTGCTAAG-3' that is located approximately 150 bases upstream of the *Bam*HI restriction site.

Rv1636_{MAKT} was also cloned, from pET41a-Rv1636_{MAKT}, into the pMV261 vector to produce pMV261::Rv1636_{MAKT}, using the primers pMV-261-Rv1636_{MAKT} For/Rev (see Chapter 2.1.6). The cloning of Rv1636_{MAKT} allowed the protein to be cloned directly as it is produced with no additional amino acids added at either end, due to the additional amino acids pMV261 adds on being complementary to the initial amino acids of Rv1636_{MAKT}. The Rv1636_{MAKT} also contained a histidine tag at the C-terminus and could therefore be a useful protein to express in the future. In the design of the pMV261::Rv1636_{MAKT} the forward primer has one codon mismatch with pET41a-Rv1636_{MAKT}, with the codon GCC present instead of GCT, to create the restriction site *Msc*I. GCC and GCT both code for the amino acid alanine. *Msc*I cuts a blunt end restriction site.

5.2.4.1.2 Cloning into pMV361

Rv1636 was cloned, from pGem-Teasy-Rv1636, into frame between the *Eco*RI and *Hind*III sites of the pMV361 to produce pMV361::Rv1636. The primers used for the amplification were pMV-361-Rv1636-For/Rev (see Chapter 2.1.6). The *Eco*RI-Rv1636-*Hind*III PCR fragment was sub-cloned in frame with P_{Hsp60} of the integrating plasmid pMV361. To ensure the correct frame reading of Rv1636, the gene was cloned in frame with the additional amino acids, MDVLYS, that pMV361 adds to the start. To ensure the correct cloning had taken place, pMV361 vectors were sequenced with the pMV361 sequencing primer 5'-AAGCGGAAGAGCGCCCAATAC-3' that is approximately 150 bp upstream of the *Bam*HI restriction site.

pMV361::Rv1636 and the pMC1s/pMC2m plasmids were transformed into *M. smegmatis*. PCRs across the integration site did not show the presence of any insertions. The primers used to amplify the attB site were those successfully used in Murry *et al.* (2005). The primer pair (Msmeg-attB-Forward/Reverse (see Chapter 2.1.6) are designed to amplify a region of 250 bp across the attB integration site. No integration was found at the attB site (Figure 5.15).

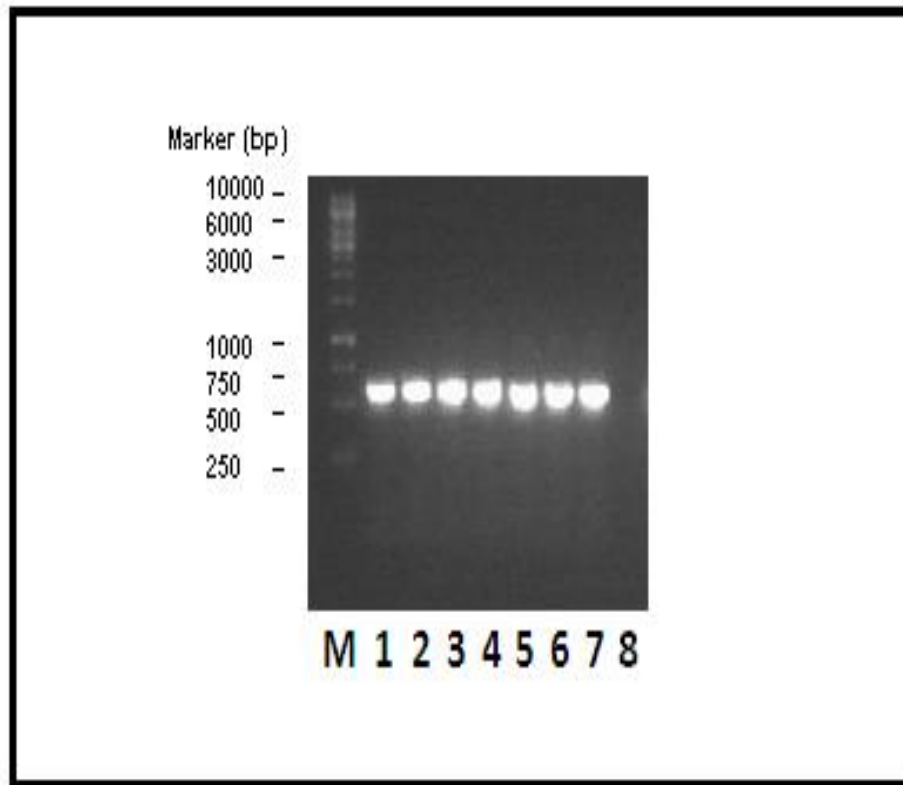


Figure 5.15. PCR analysis of across *M. smegmatis* attB site. Lane 1. *M. smegmatis* WT. Lane 2-7. *M. smegmatis* transformed with pMV361::Rv1636_{WT}. Lane 8. Negative control. M is marker Promega 1 kb hyperladder. No site-specific insertion was observed in the *M. smegmatis* attB site.

5.2.4.1.3 Cloning into pSE100

The cloning into pSE100 to create pSE100::Rv1636 involved amplifying the Rv1636 with part of the promoter region to ensure the ribosome binding site was present to allow translation. Genomic DNA from *Mtb* was used for the cloning that was extracted from wild type H37Rv *Mtb* using CTAB method. The primers pSE100-Rv1636-For/Rev are detailed in Chapter 2.1.6 and show the presence of the ribosome binding site. Over-expression using pSE100 allows the over-expression of the native protein with no additional amino acids added at the termini. The construct was not used for over-expression studies as confirmation of pMC1s and pMC2m insertion is first required to ensure tetracycline controlled regulation in *Mycobacterium* species.

5.2.4.2 Over-expression of Rv1636 with pMV261::Rv1636

An initial screen was performed to over-express the Rv1636 protein in *M. smegmatis* using the pMV261::Rv1636 vector. The plasmid is non-integrating and is selected for under the presence of kanamycin due to the resistance gene encoded. Minipreps on *M. smegmatis* confirmed the presence of the plasmid and Western blot analysis confirmed the over-expression of Rv1636 (Figure 5.16).

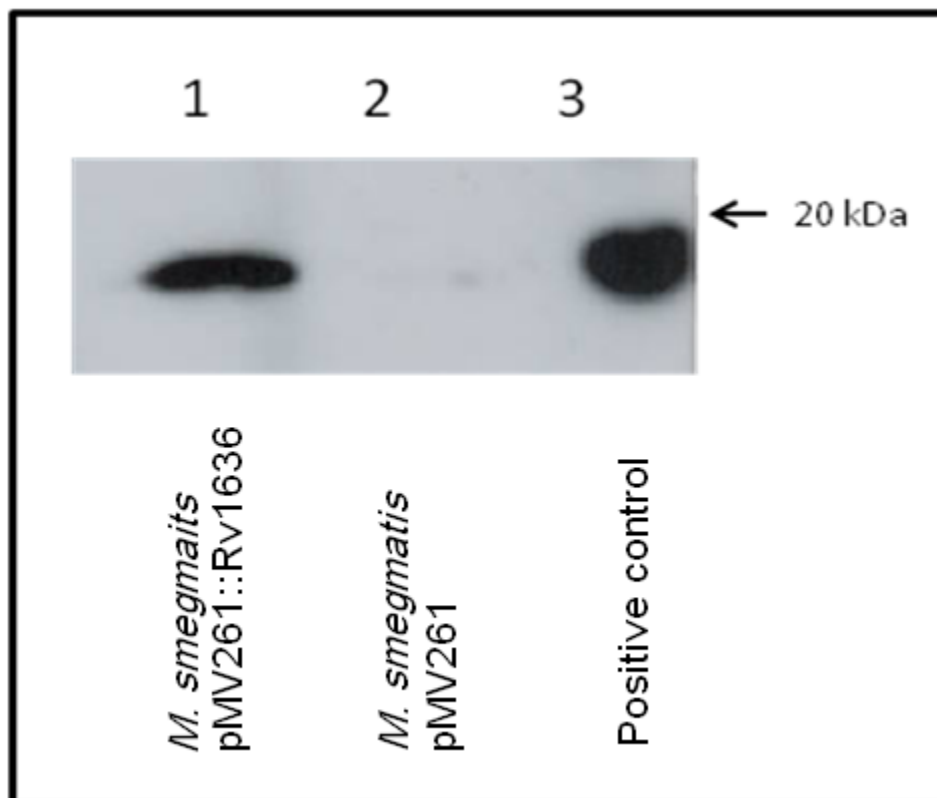


Figure 5.16. Western blot confirming over-expression of Rv1636 in *M. smegmatis* from the pMV261::Rv1636 vector. *M. smegmatis* in late log phase was boiled and the supernatant subjected to SDS-PAGE prior to Western blotting with anti-Rv1636 antibodies. The positive control was previously purified Rv1636 protein.

The results of the over-expression showed there was no significant effect on growth of over-expressing the USP in *M. smegmatis* (Figure 5.17). *M. smegmatis* containing the pMV261::empty had been previously shown not to cause any survival defect (Stover *et al.*, 1991). *M. smegmatis* WT and *M. smegmatis* pMV261::Rv1636 were grown to stationary phase and did not show any significant differences in growth.

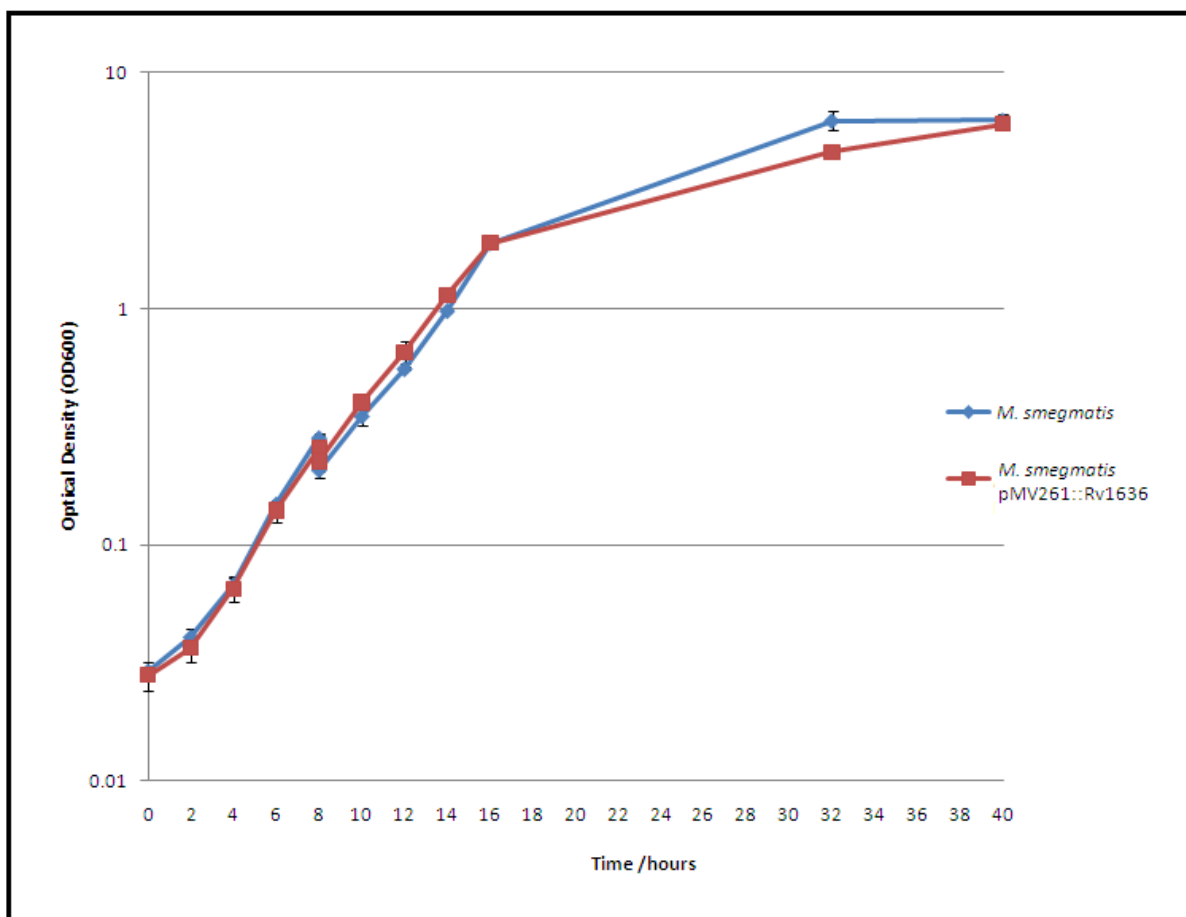


Figure 5.17. Over-expression of Rv1636 in *M. smegmatis* using the expression vector pMV261::Rv1636. *M. smegmatis* was cultured in 50 ml LB with 0.05% Tween 80 and kanamycin 40 µg/ml. OD₆₀₀ reading were recorded every two hours.

5.2.5 The Universal Stress Protein PA3309 from *P. aeruginosa*

5.2.5.1 PA3309 Structure

PA3309 is a single domain USP from *P. aeruginosa* that consists of 151 amino acids and has a calculated molecular mass of 16496 Da. It has 37% sequence identity to UspA of *E. coli*. Sequence alignment of PA3309 with Rv1636 is shown in Figure 5.18. The proteins have 27% homology and both contain the G-X₂-G-X₉-GS predicted nucleotide binding motif.

The Universal Stress Proteins of Bacteria

Rv1636	4	YKTVVVGTDGSDSSMRAVDRAAQIAG-ADAKLIIAS-----AYLPQHEDARAADILKDE	56
		Y ++V DGS S RA++ A ++AG + KL + + +LP + + L++	
PA3309	2	YDRILVAVDGSPVSDRALEEAIKLAGLSGGKLRVITIVDSPLRHLDPYAVYYNPEPLREA	61
Rv1636	57	SYKVTGTAPIYEILHDAKERAHNAGAKNVEERPIVGAPVDALVN-LADEE-----	105
		+ K ++L A+E+ A NVE GA D + A EE	
PA3309	62	ALKAAD-----DVLAKAREKVE---ASNVE-----GATFDRVCQERASEEIAERIELEAQ	108
Rv1636	106	--KADLLVVGNVGLSTIAGRLLGSVPANVSRRAKVDVLIVH	144
		AD++V+G G + +LGSV + R A VL+V	
PA3309	109	ASDADVVMGTHGRGMRRLMLGSVAEGLIRVASRFPVLLVR	149

Figure 5.18. Sequence comparison of *Mtb* USP Rv1636 with *P. aeruginosa* USP PA3309.

Using the fold recognition-based structure prediction server Phyre, the predicted tertiary structure of PA3309 was generated (Figure 5.19). The top-scoring model was based on the 1mjh crystal structure of MJ0577 (e-value=2.628228e-18). PA3309 has a 23% sequence identity with MJ0577 protein. The G-X₂-G-X₉-G(S/T) NTP-binding motifs is highlighted.

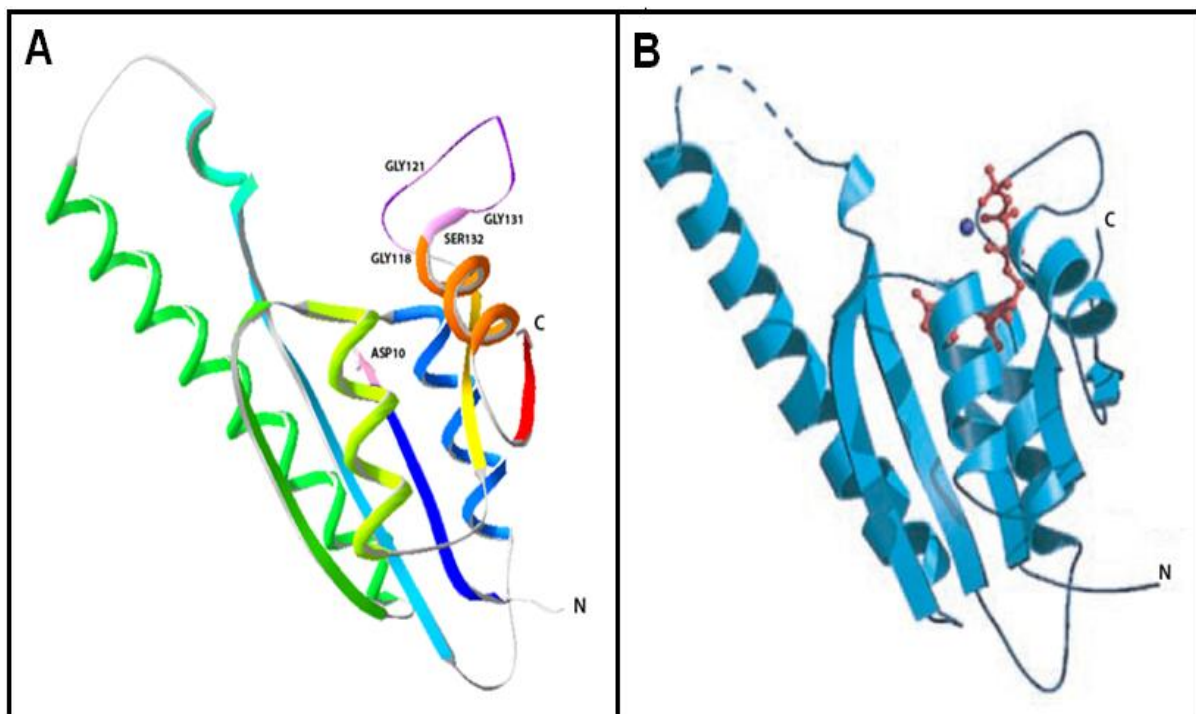


Figure 5.19. Predicted tertiary structure of *P. aeruginosa* USP PA3309. (A) Ribbon diagram of the PA3309 protein. Structure generated by PHYRE program based on the top-scoring model MJ0577 (1mjh) crystal structure. Visualised using Swiss-PDB viewer. N and C termini of each monomer are labelled. G-X₂-G-X₉-GS nucleotide binding motif is highlighted. (B) Ribbon diagram of the crystal structure of MJ0577 with bound ATP (Zarembinski *et al.*, 1998).

5.2.6 Does the Predicted PA3309 Nucleotide Binding Motif have a Direct Biological Role?

5.2.6.1 Expressing PA3309 in *P. aeruginosa* using pUCp18 vector

To determine if the predicted nucleotide binding motif had a role in the anaerobic survival, the viability of *P. aeruginosa* with pyruvate was investigated. By complementing the PA3309 knockout with a site directed mutant of PA3309 and assaying anaerobic survival, the role played by the predicted nucleotide binding in anaerobic growth could be determined. In addition, in order to test if the nucleotide binding motif had a direct relevant role in the pyruvate survival, SDM of the predicted nucleotide binding motif was performed. To express the proteins in *P. aeruginosa*, the non-integrating vector pUC18, as described by Schweizer *et al.* (1991) was initially trialled.

The PA3309 gene region (3709770-3710969), that contained the PA3309 gene flanked between the 455 bp upstream region and 289 bp downstream region of PA3309, was successfully cloned from *P. aeruginosa* genomic DNA and ligated into pUCp18, between restriction sites *EcoRI* and *HindIII*, to create pUCp18::PA3309. The primers used for the amplification were Pa-3309-Up-forward and Pa-3309-Down-Reverse (Chapter 2.1.6). The pUCp18 vector is 4500 bp and contains the ampicillin resistance gene. The whole PA3309 gene region, including native promoter and ribosome binding site, was cloned into the vector.

The experiment performed by Schreiber *et al.* (2006) was initially repeated with the pUC vectors being used to ensure that the phenotype could be replicated in our experimental conditions. The pUC vectors do not need integration and are present in the *P. aeruginosa* in an episomal form. Experimentation with the pUC vectors allowed the method and experimental details to be trialled, to ensure the results of the anaerobic pyruvate survival could be replicated.

Anaerobic survival, similar to the anaerobic culture conditions described in Schreiber *et al.* (2006) was measured by CFU counts across a time period of 11 days. *P. aeruginosa* strains were grown in potassium-phosphate buffered LB medium to an OD₆₀₀ of 0.3 before being transferred to rubber stoppered test tubes when 40 mM pyruvate was added. The survival was assayed by counting the colonies to determine viability and the results are shown in Figure 5.20. The PA3309 knockout and the knockout containing the empty pUCp18 vector both had a survival defect, as expected. The complemented strain Δ PA3309 with pUCp18::PA3309 had the survival restored to wild-type level.

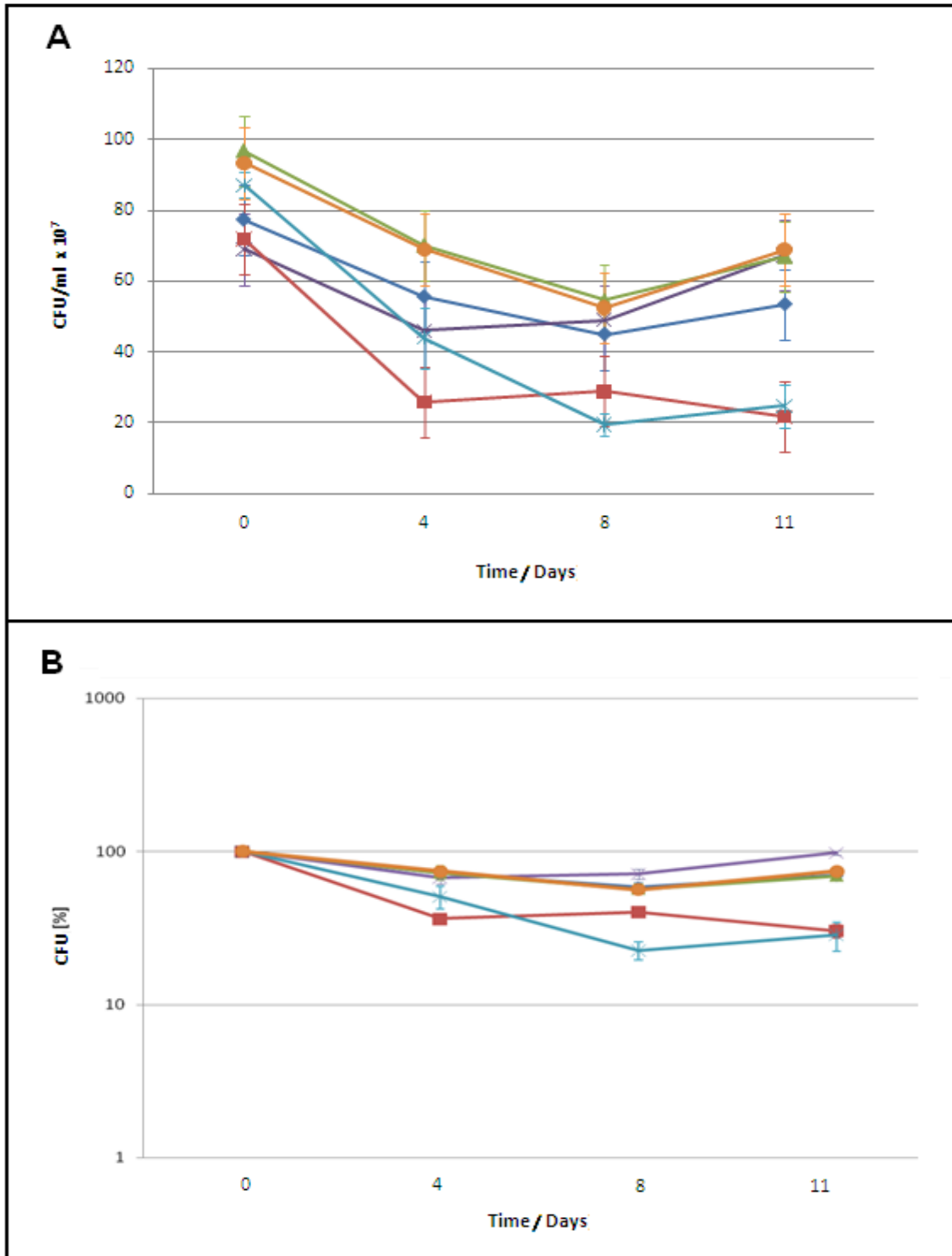


Figure 5.20. Anaerobic survival of *P. aeruginosa* on 40 mM pyruvate-supplemented media. Data is expressed as CFU/ml, with counts calculated using 10-7 dilution (A) and as percentages of the original CFU count (B). Strains based on a PAO1 strain include PAO1 (●), PAO1 + empty pUCp18 vector (×), PAO1 + pUCp18::PA3309 (▲). Strains based on a PAO1 Δ PA3309 include PAO1 Δ PA3309 (■), PAO1 Δ PA3309 + empty pUCp18 vector (✱) and PAO1 Δ PA3309 + pUCp18-PA3309 (◆).

5.2.6.2 Expression of PA3309 Using the Integrative Mini-CTX2 Vector

Following the initial experiments detailed above, the experiment was to be repeated using the mini-CTX2 vector, an integrating vector, made available by Max Schobert. The mini-CTX2 vector is able to integrate, due to the presence of the site-specific recombinase ϕ C31, into the *P. aeruginosa* genome at the site attB and, therefore, produce stable transformants. A flip recombinase then allows the antibiotic resistance cassette to be removed and therefore no antibiotic is required for selection. The removal of the unwanted plasmid backbone sequences also removes the integrase gene from the genome of integrants.

Max Schobert sent strains of *P. aeruginosa* including the PA01 wild type, PA01 containing the empty integrated mini-CTX2 (PA01 attB::mini-CTX2), PA3309 knockout, PA3309 knockout containing the empty integrated mini-CTX2 (PA01 Δ 3309 attB::mini-CTX2) and the PA3309 knockout complemented with miniCTX2::PA3309 (PA01 Δ 3309 attB::PA3309). The aim was to mutate the G residue in the G-X₂-G-X₉-GS motif of the PA3309 and integrate it into the *P. aeruginosa* PA3309 knockout to create PA01 Δ PA3309 attB::PA3309_{G118A}.

5.2.6.3 Construction of Mini-CTX2::PA3309_{G118A}

The mini-CTX2::PA3309 contained 455 bp of sequence upstream of PA3309 and 289 bp downstream of PA3309. Attempts to mutate the mini-CTX2 plasmid directly were unsuccessful. Instead, the whole gene region was removed using restriction enzymes *EcoRI* and *HindIII*. Following restriction digest and agarose electrophoresis the gene region was extracted and ligated into pET15b vector and SDM performed on this vector. Following a successful SDM reaction, the PA3309 gene and surrounding regions were again digested and ligated back into mini-CTX2 vector to produce mini-CTX2::PA3309_{G118A}.

The plasmids were sequenced and were confirmed to contain the PA3309 with G118A mutation in the correct reading frame. A further check showed was that *BamHI* and *EcoRV* both cut once in Mini-CTX2::PA3309_{G118A}. Double digestion confirmed the presence of the mini-CTX2:PA3309 and gave as expected two bands at 3.6kbp and 4.2 kbp (Figure 5.21).

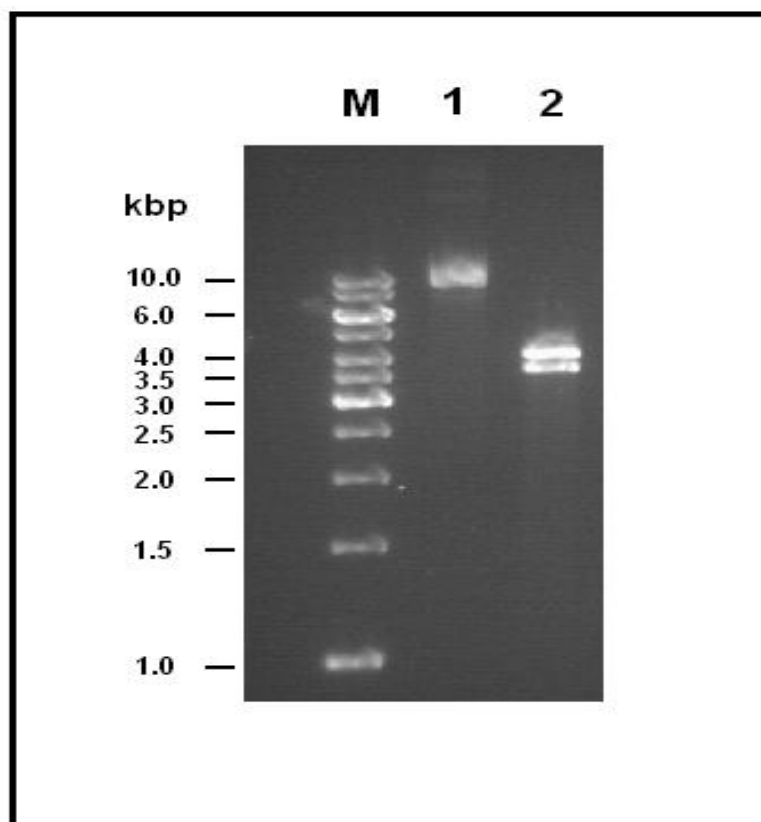


Figure 5.21. Double digestion of the mini-CTX2::PA3309_{G118A} plasmid. Lane 1 undigested mini-CTX2::PA3309_{G118A}. Lane 2. Double digestion with restriction enzymes *Bam*HI and *Eco*RV. Lane M is marker GeneRuler™ 1 kb DNA Ladder (Fermentas). Digestions were carried out at 37°C in 20 µl reactions using Fermentas fast digest restriction enzymes.

5.2.6.4 Confirmation of Insertion Site in *P. aeruginosa*

The mini-CTX2 vector was transformed into *P. aeruginosa* wild type and the ΔPA3309 knockout strain in order to obtain stable integrants. The *P. aeruginosa* attB site is at 2.94 Mb on the 6.4 Mb PA01 chromosome. Primer pairs used in the characterisation, included amplification of the attB region, the internal PA3309 region, amplification of the inner tetracycline cassette and amplification from the multiple cloning site and surrounding regions, are shown in Chapter 2.1.6. Table 5.3 shows the primer pairs used and the expected band sizes produced after PCR on different *P. aeruginosa* strains. It was decided integration should be at the same attB site as the Mini-CTX2:PA3309_{WT} strain that was sent from Max Schobert to allow direct comparison in the studies.

The Universal Stress Proteins of Bacteria

Table 5.3. PCR primer pairs used for analysis of the *P. aeruginosa* transformed strains. (F forward and R reverse). Strains sent by Max Schobert have tetracycline cassette removed by flip recombinase action.

Strain	Size of fragment amplified between primer pairs (bp)			
	attB F/R	Tet F/R	PA3309 F/R	MCS F/R
PA01 WT	279	-	424	-
PA01 Δ 3309	279	-	-	-
PA01 WT attB::mini-CTX2::empty	1160	-	-	485
PA01 Δ 3309 attB::mini-CTX2::empty	1160	-	-	485
PA01 WT attB::mini-CTX2::PA3309	2639	-	424	1685
PA01 Δ 3309 attB::mini-CTX2::PA3309	2639	-	424	1685
PA01 WT attB::mini-CTX2::PA3309 _{G118A}	~7000	632	424	1685
PA01 Δ 3309 attB::mini-CTX2::PA3309 _{G118A}	~7000	632	424	1685
S17.1 -miniCTX2::PA3309 _{G118A}	-	632	424	1685

Control amplifications on the strains sent by Max Schobert showed the PA01 wild-type and the Δ PA3309 knockout both produced a band across the attB site of the expected 279 bp (Figure 5.22A, lane 1 and 2). The strains transformed with an empty mini-CTX2 vector produced a band at the expected 1160 bp (lane 3 and 4). The 1160 bp is made up of 881 bp of the vector back bone between the two flip recombinase sites and the 279 bp of the attB site. The complemented strain did not produce a band in the PCR of the expected 2639 bp (lane 5). However, the strain did not produce a band at 279 bp suggesting there was an insertion present.

When screening the *P. aeruginosa* transformants for the desired insertion there would not be a band present at 279 bp if insertion at the attB site had occurred. All 250 mutants screened, from a number of mating and transformation attempts produced a band at 279 bp suggesting no insertion at the attB site was found. The presence of the tetracycline gene indicated insertion may have taken place at non-specific sites (Figure 5.22C).

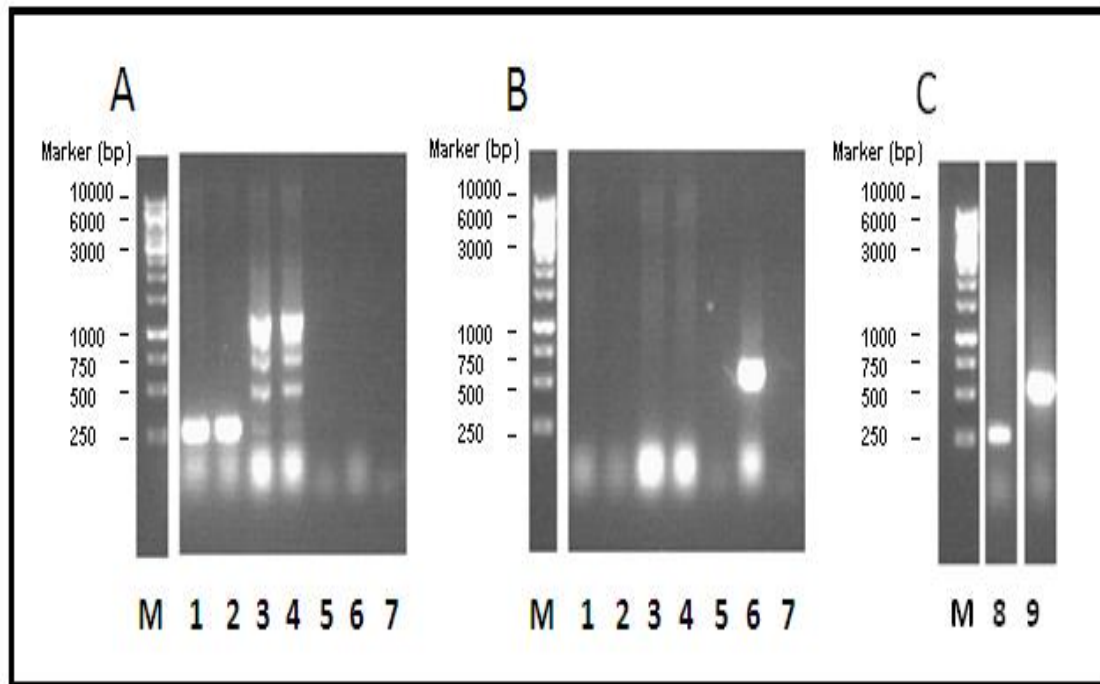


Figure 5.22. PCR analysis of *P. aeruginosa* strains covering *attB* site and the tetracycline resistance cassette. (A) PCR of *attB* site. (B) PCR of tet cassette. (C) PCR on putative PA01Δ3309 attB::mini-CTX2 3309_{G118A} mutant across attB (Lane 8) and tet cassette (lane 9). Lanes 1. PA01 WT, 2. PA01 Δ3309, 3. PA01 WT attB::mini-CTX2 4. PA01 Δ3309 attB::mini-CTX2, 5. PA01 Δ3309 attB::PA3309, 6. *E. coli* S17.1 containing episomal mini-CTX2::PA3309_{G118A} 7. Negative control, 8 and 9. Putative PA01 Δ3309 attB::mini-CTX2::PA3309_{G118A} mutant.

Attempts with heat shock transformation, chemical transformation and mating all were unsuccessful at obtaining any site specific integration. The mating controls were all successful with no colonies on the tetracycline plates when the colonies were not mated. *P. aeruginosa* did grow on the high tetracycline up to 200 μg/ml; however PCRs showed no attB integration. Instead of amplifying the whole region primer pairs using the T7 primer within the mini-CTX2 vector and either Pa-attB-For/Rev which would produce PCR products of around 2.5kb did not show any integration. Attempts to get mini-CTX2::PA3309_{G118A} into *P. aeruginosa* chromosome at attB site were unsuccessful and no survival complementation experiment with the mutated PA3309 gene was able to be performed.

5.3 Discussion

In this chapter, work is described which was carried out with the aim of identifying the *in vivo* role of USPs and the potential role that the nucleotide binding motif plays. Attempts to find an *in vivo* role for USPs in the biology of *Mtb* and *P. aeruginosa* are discussed.

5.3.1 Genomic context and lineage

The genomic context of Rv1636 and the homologues in other mycobacteria does not provide any obvious clues to the *in vivo* role of Rv1636. The *UvrA* and *UvrB* genes are both in the same genomic location of *Mtb*, *M. bovis* and *M. leprae* and *M. smegmatis*, suggesting a possible link to survival under UV radiation; however UV radiation has not been linked to alterations in *Rv1636* transcription and the transcriptional direction of the genes does not suggest co-regulation. The number of homologues found in other mycobacterial species point towards the potential importance of the protein. Within the suborder Corynebacterineae, which contain the genus *Mycobacterium*, proteins with over 60% homology are found such as protein DSM_43247, from *Gordonia bronchialis*, which shares 65% homology. *Gordonia bronchialis* can cause a pulmonary disease in humans.

Interestingly a recent paper highlighted the potential importance of USP promoters. The cotton plant, *Gossypium arboreum*, contains a USP gene called *GUSP* gene that plays a role in the response to water stress. Interestingly, the promoter region of 949bp upstream of the *GUSP* gene contained numerous cis-acting elements found to be involved in stress survival in plants including E-box, ABRE, DPBF-box and MYB-core elements. This is the first such study where a USP promoter region has been linked to responses to stress. Stresses included response to dark, heavy metals, dehydration, salt stress and acid stress. The promoter elements of *GUSP* are likely to allow regulation of this USP to be tightly controlled. Rv1636 is preceded by Rv1635c which is transcribed in the opposite orientation to Rv1636. There are 209 bp in between the two start codons and therefore the promoter elements are likely to overlap.

5.3.2 Creation of a Rv1636 knockout

The creation of a mutant is an important tool for studying gene function. PCR screening on the putative mutant confirmed none of the 50 potential mutants were positive for the hygromycin cassette coupled with being negative for the *Rv1636* gene. The potential double crossovers were all shown to not be true *Rv1636* deletions. The colonies did appear very small compared to the wild type and could therefore be spontaneous mutations in the *sacB* gene, which is present to allow for negative selection on sucrose.

Although the mutants screened for survival within macrophages were not double crossovers, it remains possible that the confirmed single crossover could have partially disrupted the expression of the *Rv1636* gene, meaning the survival defect could be a true representation of a *Rv1636* mutant. The construction of stable double crossovers by homologous recombination can be difficult in the slow growing mycobacterial species. Low levels of homologous recombination and high levels of non-homologous recombination, combined with variation at different loci are some of the potential reasons why the stable double crossovers were difficult to obtain (McFadden 1996).

To isolate *Rv1636* mutants, the work can be revisited at the single cross-over stage and attempts made to re-select for double crossovers. An essentiality test will also be performed by using the partial diploid that is available in the laboratory, alongside the work of creating an *Rv1636* mutant. Further work will involve returning to the confirmed single crossover strains and with further selection on sucrose isolate more colonies which will again be screened for the presence of hygromycin and the absence of the *Rv1636* gene. An alternative approach of creating the mutants using specialised transduction could also be used as described in Bardarov *et al.*, 1997, or by using the Hatfull lab's recombineering method (van Kessel and Hatfull, 2007).

Further work will also aim to determine if the *Rv1636* gene is essential. By making partial diploids with the *Rv1636* gene and creating a knockout in one of these genes this will allow determination of essentiality essential, when compared to the wild type.

When a knock-out mutant can be isolated, the phenotypic characterisation of the mutant can be studied in stress conditions including macrophage survival, nutrient and oxygen depletion and sensitivity to nucleic acid damaging agents such as mytomycin C, UV light and hydrogen peroxide. Complementing any *Rv1636* knockout mutant with site directed mutants in *Rv1636* would also allow the biological significance of the nucleotide binding motif to be investigated.

5.3.3 Identification of Protein Interactions

Protein interactions have specific roles in the biochemistry of protein molecules, signal transduction and networks. They allow communication between the cells exterior and the interior of the cell that is mediated by signalling molecules. To find an interacting protein would provide a large amount of potential for future research. The identity of interacting proteins could provide an insight into the role of the protein by providing information on the network the protein is involved in.

USPs have been shown to interact with each other when studying heterodimerisation within the classes of *E. coli* (Nachin *et al.*, 2008; Nachin *et al.*, 2009). UspC of *E. coli* has also been shown to interact with the KdpD protein in *E. coli* through its own USP domain. The significance of the USP-USP interactions were discussed in the Chapter 4 and the protein interactions discussed here are those of USPs with other mycobacterial proteins. A recent study of global protein-protein interactions by Wang *et al.* (2010) did not highlight any significant interacting partners for the USPs of *Mtb*.

5.3.3.1 Rv1636 Expression

S. Leiva (2007) previously showed that Rv1636 was expressed in stationary phase using q-rtPCR and Western blotting. Due to this result, stationary phase lysates were then used for interaction studies. The Mb1662 protein was found to be present in cytosolic fractions obtained from *M. bovis* BCG grown aerobically in mid-exponential growth and in stationary phase. The immunoblot shows a stronger signal present in the stationary phase cytosolic fraction indicating the protein is more abundant compared to the mid-exponential phase.

The blots on the cytosolic fractions also showed that the antibody did not cross-react with any other protein from the cytosol. No Mb1662 was found in the oxygen starved stationary phase cytosolic fraction. The expression results were backed up by the immunoprecipitation of the Mb1662 protein when performing Co-IP.

From studying *M. bovis* BCG expression of the Mb1662 protein, it is thought that the Rv1636 protein in *Mtb* is expressed in the same way and by the same control mechanisms. The protein is 100% identical and is most likely to have the same role within the cell. It is feasible that other interacting proteins are present in *Mtb* that have been lost in *M. bovis* BCG and to confirm any interactions, the experiments should be repeated with *Mtb*.

The Universal Stress Proteins of Bacteria

Bahk *et al.* (2004) compared protein levels between cultures of *Mtb* K-strain from Korea and the laboratory strain H37Rv. Rv1636 was one of the eight proteins found to be more abundant in the K-strain. The finding that there are differences in expression between individual strains of *Mtb*, leads to the suggestion that there could also be differences in expression between *M. bovis* BCG and *Mtb*.

The Wayne-Sohaskey (2001) model suggests granuloma conditions where the bacteria persist within infected macrophages and also extracellularly in a mix of lysed cell debris, resulting in decreasing oxygen levels until the bacilli enter the non-replicating persistent state. If protein is not expressed under oxygen depleted conditions it does not imply the protein is not expressed within the granulomas. It's possible that other USPs permit survival of the cells under the hypoxic conditions and Rv1636 contributes to cell persistence through allowing survival against other stresses. The ten USPs of *Mtb* are not all under the same control systems suggesting their roles may be distinct. Another possibility is the protein may not be essential when *Mtb* remains within granulomas in a non-replicating persistent state, but plays a crucial role in other circumstances, such as survival in the macrophage upon initial infection or survival when transmitted.

5.3.3.2 GST Pull-down of Rv1636

The GS4B beads were shown to retrieve the GST-Rv1636 protein from the cytosolic fractions as expected. No interacting partner(s) were found under the range of tested conditions. GST-pull-down benefit's include that the approach can purify low abundant protein complexes and is applicable to very weak interactions. Possible reasons why no interacting partner was found include the fact that Rv1636 attached to the GST is subject to competition for any potential partners because the identical protein Mb1662 is present in the cytosolic fraction and may already be present in an *In vivo* pre-assembled complex, thus preventing the GST-Rv1636 protein from binding. The need for *in vitro* complex formation is one of the major draw-backs of the GST-pull-down method.

A further disadvantage of GST-pull-down is the tag can influence the function of protein being tested for interactions. The GST-tag at 26 kDa is large compared to the Rv1636 protein (15 kDa) and the tag may alter the binding capabilities of the protein. Obviously the size could alter any oligomeric structures being formed thus preventing complex formation and subsequent pulling down of the partner protein(s).

Early in the project, SDS-PAGE gels were being stained with the Imperial Coomassie stain which claimed sensitivity down to 3 ng of protein. The stain was used to highlight the GST and GST-

The Universal Stress Proteins of Bacteria

fusion protein in the pull-downs; however the possibility remained that any partner might not be in sufficient quantity to be able to produce a significant band on the gel and therefore may be overlooked. Silver staining was tested and found to be significantly more sensitive than Imperial stain, highlighting bands that were not visible on the Coomassie stained gels. The disadvantages of silver staining are that the procedure is expensive, sensitive to background and time-consuming. The benefit of increased resolution outweighed the disadvantages and the stain was therefore chosen for future experimentation.

5.3.3.3 Co-immunoprecipitation of Rv1636

Immunoprecipitation of the Mb1662 from the cytosolic fraction was successful with western blots confirming the presence of the protein within the precipitated sample. Even with modifications to the procedure including increasing the antibody concentration, an interacting protein was not discovered. Although no nucleotides or cations were added to the cytosolic fractions in the Co-IPs that found the partner, the fractions are likely to have low level of these present due to the destructive nature of the cell fractionation. It is possible post-translational modification has an effect on the interaction Rv1636 undergoes within the bacterial cell *in vivo*. Merck Pansorbin® cells (*Staphylococcus aureus* cells bearing fixed Protein A) were used to pre-clear the reaction lysates and remove any unspecific interactions. Earlier in the project a band of interest at 10 kDa was shown to bind the Protein A in a non-specific manner and was not an interacting partner of Rv1636.

Co-IP benefits include that it is independent of cloning and there is no competition with *in vivo* pre-assembled complex. Using the antibodies to retrieve protein complexes from endogenous components of the *M. bovis* BCG cells does not require the use of an ectopic bait protein or an over-expressed protein and therefore more closely resembles the native physiological conditions that usually exist within the cytosol (Bauer and Kuster, 2003). The disadvantages of Co-IP include that it requires an antibody against the protein of interest. The laboratory has polyclonal antibodies against the Rv1636 which are specific against different epitopes because they are produced from several lines of B-lymphocytes within the rabbit.

Monoclonal antibodies on the other hand are derived from a single cell line and are specific against a single epitope; however they are very expensive because of the long time taken to produce them. When using antibodies, there can be problems with cross-reactivity; however, the Western blot on the cytosolic fraction using the anti-His₆-Rv1636 antibody only highlighted one band at the

The Universal Stress Proteins of Bacteria

expected molecular mass; therefore the antibody was confirmed to be specific with no significant cross-reactivity.

Future interaction studies would involve crosslinking any complexes within the cellular lysates. If the interaction is too weak for the washing steps any weak association may be lost. Crosslinking could be performed before protein purification and even straight before co-immunoprecipitation. If Rv1636 could be over-expressed within the mycobacteria, crosslinking could provide more material with the aim of isolating the interacting partners within the cells. Performing assays in the presence of cAMP may also be of interest as the nucleotide could play a role in the interactions of the Rv1636 protein.

Crosslinking reactions are currently being performed within the laboratory using formaldehyde. Other examples of *in vivo* crosslinking described by Nachin *et al.*, 2008, involves using the cross-linking agents bismaleimido-hexane or disuccinimidyl glutarate along with DMSO as a control, which dissolves these cross-linking agents in solution. Control proteins to be used include DosR, which have known interacting partners DosS/T (Roberts, *et al.*, 2004). Other methods also allow protein-protein interactions to be studied. Using a bacterial-two-hybrid (B2H) system any interaction found could be tested *in vivo*. More recently Singh *et al.*, 2006, described a method called mycobacterial protein fragment complementation, which can be used to confirm protein-protein association in mycobacteria.

5.3.3.4 Known USP Interactions

The interaction between *M. smegmatis* MSMEG_5458 and the USP MSMEG-4207, found by Nambi *et al.*, 2010, is an example of an USP interaction identified by GST-pull-down. The interaction was shown to be a tight interaction with the *M. smegmatis* version of the protein shown to have dissociation equilibrium constant of 200 μ M. However, the GST pull-down performed in the same manner as Nambi *et al* here was unsuccessful in identifying a precipitated partner.

Attempts by Nambi *et al*, 2010, to identify an interacting partner for the homologue in *Mtb*, Rv0998, were not successful. The result might suggest that the interactions of the homologous protein might have different interaction strength across different species of mycobacteria. Any proteins Rv1636 interacts with might be of a more transient nature. It was found that the Rv0998 protein can complement the knockout in *M. smegmatis* even though there is 57% homology over the 333 amino acids of the two proteins (Personal communication, Visweswariah, 2010). This suggests that the *Mtb* protein can interact with proteins and perform a similar function to the homologue in *M. smegmatis*.

The Universal Stress Proteins of Bacteria

A well described protein which interacts with a USP is the *E. coli* protein GroEL, which was shown to release UspG from complex on addition of ATP (Bochkareva *et al.*, 2002). The molecular chaperone GroEL is a member of the highly conserved Hsp60 family of heat-shock proteins. Molecular chaperones are able to bind to non-native structural forms of polypeptides and promote them towards forming their native conformation (Bukau & Horwich, 1998). Under heat shock and other stress conditions they prevent mis-folding and aggregation of mis-folded proteins (Gragerov, 1992). It is thought that UspG may require GroEL to ensure proper protein folding conformation (Bochkareva *et al.*, 2002). The post translational modifications of proteins can have roles in marking proteins by tagging, scaffolding, folding, alter signalling events, and activation of other proteins.

UspG (also called UP12) is induced by heat shock and was found to be in a complex with GroEL and was able to be isolated by performing co-immunoprecipitation experiments. Mutants in UspG recover slowly from stationary phase and are more sensitive to the toxic agent carbonyl cyanide m-chlorophenyl hydrazone (CCCP) (Bochkareva *et al.*, 2002). As a molecular chaperone it is possible that the GroEL interaction with UspG ensures the proper folding state of UspG is achieved when the cell is under stress. Alternately, UspG could be aiding the GroEL in its role as a stress response protein.

5.3.4 Over-expression of USPs

With no mutant of Rv1636 being created and protein-protein interactions failing to identify an interacting partner, the role of Rv1636 was to be investigated by determining if over-expression of Rv1636 produced any phenotype. The Rv1636 gene was successfully cloned into a range of expression vectors for mycobacteria including under the constitutive heat shock promoter P_{hsp60} of pMV261 and pMV361. The cloning into pSE100 vector would allow finer level of control of the expression. Rv1636_{MAKT} was also cloned into the pMV261 expression vector with the aim of determining if the modifications to the Rv1636 had any phenotype on the cell and would allow comparison to the Rv1636_{WT}.

Over-expressing Rv1636 did not show any growth phenotype in *M. smegmatis*. *M. smegmatis* encodes the homologous protein MSMEG_3811 (78% identity) and also the similar MSMEG_3308 (52% identity). It is possible that the *Mtb* protein Rv1636 does not interact with the same partners of the *M. smegmatis* proteins. Small modifications in the protein structure may prevent Rv1636 having an effect in the *M. smegmatis*. The *M. smegmatis* was able to grow on the selective medium and overexpression of Rv1636 was confirmed by Western blot analysis, showing

The Universal Stress Proteins of Bacteria

the over-expressing gene was also not toxic to the cell. It would also be interesting to test for other phenotypes such as resistance to heat shock, hypoxia, nutrient starvation, UV radiation, osmotic stress, ROIs, RNIs and DNA damaging agents.

An example of expressing a gene from *Mtb* in *M. smegmatis* was shown with Rv0998, where the gene was able to complement the MSMEG_5458 knockout in *M. smegmatis* (Personal communication, Visweswariah, 2010). The protein MSMEG-5458 is the protein responsible for acetylation of the USP MSMEG_4207. The result shows that an *Mtb* protein was able to have an effect in *M. smegmatis*. Similar to the experiment described, a knockout of the homologue to Rv1636, MSMEG_3811, might provide a system that allows the Rv1636 to have a greater effect when over-expressed in *M. smegmatis*.

Drumm *et al.* (2009) only saw an effect in mycobacteria by over-expressing the *Mtb* USP Rv2623 protein in the fast growing *M. smegmatis* strain mc²155. The introduction of the multi-copy plasmid pMV261::Rv2623, resulted in growth deficiency of the *M. smegmatis* strain mc²155 on solid 7H10 Middlebrook medium and in the liquid medium-based BD BACTEC 9000MB system. Expressing a protein to high levels during growth would likely be detrimental to cells. However, over-expression of Rv2623 site directed mutants G117A and D15E did not result in a marked retardation of growth, suggesting the nucleotide binding plays a role in the growth attenuation in *M. smegmatis* strain mc²155.

5.3.5 *P. aeruginosa* Aerobic Survival and the PA3309 Protein

There were certain problems with the experimental design for the anaerobic survival experiment described in Schreiber *et al.* (2006). Firstly, it was not clear why the anaerobic samples were left static. A static culture has a larger amount of variation across the vessel compared to a continuously shaking culture. Even taking the sample from a test tube that had been inverted would likely not provide the most accurate results, as homogeneity would not be achieved. Individual eppendorfs for each measurement may be a way to keep the culture static and obtain reliable readings from samples without disturbing the continuing experiment. The experiment should be repeated with a shaking culture and samples taken in a similar manner.

The *P. aeruginosa* cultures were initially grown aerobically due to the fact growth is not supported on pyruvate. Schreiber *et al.* (2006) did do a control experiment in which it was shown that pyruvate added during aerobic growth was not used and remained in the medium. PA3309

The Universal Stress Proteins of Bacteria

mutants were 400-fold reduced in survival rates whereas complemented mutants were reduced 100-fold in the PA3309 study.

The integration-proficient miniCTX2 plasmid was designed to allow the efficient and stable integration of exogenous sequences at the defined attP location in the *P. aeruginosa* genome. Transcription of the integrase gene in the mini-CTX2 vector is driven by the strong *trc* promoter. The *trc* promoter is regulated by the Lac repressor that is encoded by *lacIq* also contained on the plasmid. Hoang *et al.*, 2000 showed that the addition of IPTG to the mating mixture had no significant effect on integration efficiency. They showed that after conjugal transfer the mini-CTX2 integrated at frequencies of 1.46×10^{-7} . Unfortunately for our study, no successful integration was observed. The controls on the mating plates gave the expected results, with strain not mated not producing any *P. aeruginosa* that could survive on the tetracycline 200 µg/µl plates.

The mini-CTX2 vector contains the CTX integrase that promotes integration of the vector into the *attB* site of the *P. aeruginosa* chromosome. The modified ϕ CTX integrase (*int*) gene is able to catalyze the recombination between the phage attachment site (attP) on the plasmid and the attachment site in the *P. aeruginosa* genome (attB) to allow integration of Mini-CTX2::PA3309 into the *P. aeruginosa* chromosome. To achieve stable transformation with episomal plasmids, continuous antibiotic selection is required.

It was desirable to be able to remove the resistance marker to make the *P. aeruginosa* strains conducive to future infection models. Some anaerobic survival biofilm models commonly used with *P. aeruginosa* would not be suitable for continuous antibiotic selection. In these infection models, the bacteria may lose plasmids as infection progresses. Thus to obtain stable chromosomal recombinants it was decided to use the mini-CTX vectors that are able to achieve site-specific integration. High copy number episomal plasmid and genomic DNA also have different extent of supercoiling, which has been shown to have an effect in gene expression regulation (Dorman, 1991).

The advantages of the site specific recombination includes that the gene sequence reproducibly integrates at a single known site. The insertion at the known site does not disrupt further gene regions and the effects of the integration are equal to all the strains that contain the directed insertion. Being chromosomal, the gene is therefore found in single copy and any detrimental effects of a multicopy plasmid being present within a strain are reduced. Multicopy plasmids can produce artefacts as they are present in large quantities. Once inserted, the integrants are then often a significant level of stability compared to episomes. The mini-CTX vector was able to provide these advantages and was used to further the study.

Further work could involve cloning the G118A mutant into the pUCp18 vector and assaying the survival. Re-attempting to complement the knockout PA3309 strain with the PA3309_{G118A} using

The Universal Stress Proteins of Bacteria

the mini-CTX2 vector would provide the most robust evidence that the nucleotide binding motif is required to allow aerobic survival. The experiment with the pUCp18 vectors confirmed that there was a survival defect in the PA3309 mutant and showed the system was in place to test the survival of a knockout mutant complemented with the PA3309_{G118A}.

5.3.6 Summary

Investigation into the intracellular role of Rv1636 and PA3309 attempted to determine the direct biological role of the USPs and to supplement the biochemical and structural analyses. Attempts to make a genetic knockout mutant in Rv1636 were unsuccessful, suggesting that the protein might be essential; however further investigation is required to confirm this. No-interacting partner has been found through co-immunoprecipitation or GST-pull-downs; however Western blotting confirmed protein expression of Rv1636 was at highest levels in stationary phase. Over-expression of Rv1636 in *M. smegmatis* did not result in any changes in growth in LB broth and further work could be trialled in *Mtb* or *M. bovis* BCG to express the protein in the native environment and under other conditions, including certain stresses. *P. aeruginosa* continues to provide an excellent opportunity to study the USPs predicted nucleotide binding motif's role in anaerobic survival. The PA3309_{G118A} protein when integrated in *P. aeruginosa* should provide key information as to whether the predicated nucleotide binding site is biologically relevant to the survival of *P. aeruginosa* during pyruvate fermentation.

Chapter 6 - General Discussion

6.1 Introduction

The original aim behind this thesis was to characterise the USP Rv1636 of *Mtb*. Later, our area of interest broadened to include the USP PA3309 of *P.aeruginosa*. The USPs have been linked to a variety of stress survival phenotypes in a range of bacteria and plants, with the suggestion that they may have a role in signalling pathways. To elucidate the role of USPs in the pathogen *Mtb*, biochemical and biophysical analysis was carried out on Rv1636 protein, the sole single-domain USP of *Mtb*. This general discussion provides a summary of the main findings of the project and the work is discussed with reference to recent studies on other USPs, with the aim of better understanding the potential roles of USPs in *Mtb*.

6.2 Biochemical Activities

6.2.1 Activities of Rv1636

As a result of this work, it has been demonstrated that Rv1636 is able to bind nucleotides. To further develop this finding, it will be important to determine the specificity of the nucleotides bound, including cyclic nucleotide monophosphates. The recent publication and personal communication with Visweswariah (2010) has provided evidence that it may be worth investigating the potential cAMP binding capabilities of Rv1636.

The putative biochemical activities of Rv1636 previously described by Leiva (2007), including autokinase, autonucleotidylation, RNA triphosphatase and NTP triphosphatase activities, were not confirmed by the more in-depth investigations described in this thesis. Although evidence was provided for the presence of specific phosphorylation of Rv1636 on serine residue(s), the levels of labelling in the autokinase reaction was found to be similar to that of biochemically inactive proteins. Similarly, the labelling as a result of the nucleotidylation reaction was found to occur on SDS-denatured Rv1636, indicating the addition of a NMP moiety was not catalytic in nature. The potential phosphatase activity of Rv1636 acting on RNA or nucleotide triphosphates substrates also remains unresolved. It remains a possibility that Rv1636 can degrade RNA, which could be intriguing were this a true, catalytic activity of the protein with implications for the involvement of Rv1636 in RNA turnover and adaptation to changing conditions. It has also been demonstrated that a purified Rv1636 preparation has cation-independent NTPase activity; however further work is required to fully understand this activity. Particularly intriguing was the observation that the NTPase activity

The Universal Stress Proteins of Bacteria

was shown not to be associated with the predominant species of Rv1636 after size exclusion chromatography. It is possible that this was a result of a contaminating protein, however the possibility that a different form of Rv1636 present in small amounts was responsible for the activity, has not been ruled out. Expressing the proteins in a mycobacterial species could allow the purification of Rv1636 in a state more closely resembling that seen *in vivo*.

It should be noted that the post-translational modifications of Rv1636 *in vivo* have not been studied and it is likely that these would provide a more focused approach to determining the true biochemical activities of the protein. One possibility is that the biochemical activities of Rv1636 may require alternative factors that are provided *in vivo*. This hypothesis is somewhat supported by the observation that the *in vitro* phosphorylation of *E. coli* UspA requires the TypA protein (Freestone, 2007).

Successful SDM has provided a variety of mutant Rv1636 proteins that will aid the future study of this USP. Mutations in potentially key serine residues and the putative nucleotide binding site can be screened rapidly for specific activities, thus providing information about essential residues.

6.2.2 PP-loop Members

The structure of the USP proteins can provide information about the potential biochemical activities of these proteins. The majority of NTP hydrolysing enzymes cleave the β - γ bond, of which the best characterised are the p-loop proteins that contain the conserved Walker NTP-binding motifs A and B. Other mechanisms of β - γ bond cleavage are present in other proteins including some protein kinases (Hanks and Quin, 1991) and other selected proteins such as the hsp70 heat shock protein (Bork *et al.*, 1992). The detailed study of the Rv1636 structure and biochemical activities can be used to investigate the function of these proteins. Rv1636 is a member of the PP-loop clan (Bork and Koonin, 1994) and investigation of these members and the structures may provide clues to shared roles.

Members of the PP-loop clan originally described by Bork and Koonin (1994) were found to have distinct catalytic mechanisms to hydrolyse the α - β phosphodiester bond. The PP-loop clan were likely not to have a single phylogenetic origin as shown by two distinct aminoacyl-tRNA synthases that showed different topologies (Nagel, & Doolittle, 1991). Members used cofactors such as acyl-CoA synthase that use CoA, whereas others, such as DNA and RNA ligases, rely on covalently linking AMP moieties to other enzymes. The members of the PP-loop do contain a glycine rich loop

that follows an interior β -strand. Although it is drastically reduced in the PP-loop proteins it is likely a modified version of the P-loop found in β - γ NTPases (Bork and Koonin, 1994). Comparing the two sulphate binding proteins PAPS reductase and ATP sulfurylase, Bork and Koonin (1994) showed they had significant overall similarity even though PAPS reductase did not exhibit ATPase activity. The two are an example where selection pressures can allow the protein function of a conserved motif to shift from one role to another.

The linking of the PP-loop protein to a catalytic mechanism of α - β phosphodiester bond hydrolysis may be relevant to the RNA being degraded when incubated with a Rv1636 preparation in a RNA triphosphatase assay. The Rv1636 may therefore be able to hydrolyse the α - β phosphodiester bond nucleotide monophosphates. This is more interesting following the suggestion that Rv1636 may have cAMP binding properties as a result of the results described by Nambi (2010); providing one potential link between nucleotide monophosphates and Rv1636.

6.3 Biophysical Studies

6.3.1 Structure of Rv1636

Successful disruption of the Rv1636 strand swapped dimer did appear to lead to an increase in nucleotide binding. Rv1636 was shown to form high order structures with evidence of an octomer observed in the SEC-MALL analysis. A monomeric form of the truncated protein Rv1636_{MAKT} was identified by analytical ultracentrifugation, although the species gave a molecular mass slightly above that seen in the SEC-MALLs analysis suggesting associations between units were still present. It remains to be determined if the oligomeric state of Rv1636 is affected by nucleotide binding or another factor. In other species, studies of USP structures have allowed predicted interactions between USP to be modelled.

With Rv1636 being the only single domain USP in *Mtb*, it is clear that the tandem domains are the dominant conformation in this species. The potential role that a single domain USP may have in the monomeric form may be a distinct role compared to the tandem domain USPs and even distinct to the role of the dimer/tetramer that Rv1636 can form through association of multiple Rv1636 species.

Of further interest would be attempting to purify Rv1636 with the addition of different nucleotides, such as ATP, to the purification buffers. Nucleotides may bind to the monomer, thus

preventing the formation of the strand swapped dimer. The distribution of the oligomeric species may shift and be observed after SEC. It would also be worth purifying Rv1636 from *Mtb* or a related mycobacterial strain to understand whether the oligomeric species are the true *in vivo* species or an artefact of overexpression in *E. coli*.

6.3.2 Functional Oligomerisation

A recent paper by Schweikhard *et al.* (2010) describes the structural and functional organisation of the TeaD USP from the halophilic bacteria, *Halomonas elongate*. The regulation of the ectoine transporter by the USP TeaD, allows survival under the harsh conditions experienced by this bacillus. The bacilli are able to both synthesise and take up the natural compound ectoine, which can serve as a protective substance aiding the bacteria in its resistance to temperature and salt stress.

The USP *TeaD* is the fourth open reading frame in the *TeaADCD* operon and is a negative regulator of the *TeaABC* operon which encodes an osmoregulated tripartite ATP-independent periplasmic (TRAP) transporter. The negative regulation does not occur via DNA binding and Schweikhard *et al.* (2010) in fact showed that ATP-dependent oligomerisation might play a functional role in the regulatory mechanism of TeaD. By analysing HPLC retention times, the authors showed that TeaD was able to bind ATP and the ATP was shown to stabilise the dimer-dimer complex of TeaD.

The interactions between the monomers that make up the dimers are also of interest in TeaD. Schweikhard *et al.* (2010) showed that the Arg135 residue provided by a neighbouring dimer forms a hydrogen bond with the gamma phosphate of the ATP molecule. The crystallised USPs from *K. pneumonia* and *T. thermophilus* also both have a positive charged residue (either lysine or arginine) present on an adjacent dimer involved in mediating the binding to the gamma phosphate of the ATP molecule. The property can be likened to the R-finger as described for the AAA+ classical Walker motif ATPases. Phosphoryl transfer can generate negative charge and the R-finger is involved in offsetting the charge build up, therefore allowing hydrolysis. No ATPase activity was shown for TeaD alone *in vitro*. The bound ligand Mg^{2+} also had interactions with the phosphates and the interaction with the α and β phosphate was shown to be stronger than the γ phosphate. Water molecules were found to complete the octahedral coordination of the magnesium ion.

TeaD has similarities to eukaryotic USP Arabidopsis 2dum, the archaeal USPs, MJ0577 and *Aquifex aeolicus* (1q77), and the *Pyrococcus horikoshii* protein PH0823, which has a short α 2-helix.

The Universal Stress Proteins of Bacteria

Mtb USPs do not have this helical secondary structure element (3cis), as with the *Archaeoglobus fulgidus* (3dlo). Structural studies also showed monomer-monomer interactions between $\alpha 5$ and $\beta 4$ are conserved in the USPs. Sequence alignment of the proteins that were found to be crystallised with an ATP cofactor allowed Schweikhard *et al.* (2010) to identify highly conserved residues involved in ATP binding (Figure 6.1). Three regions were highlighted as being responsible for adenosine binding and one region, containing the G-X₂-G-X₉-G(S/T) motif, was found shown to be responsible for phosphate binding.

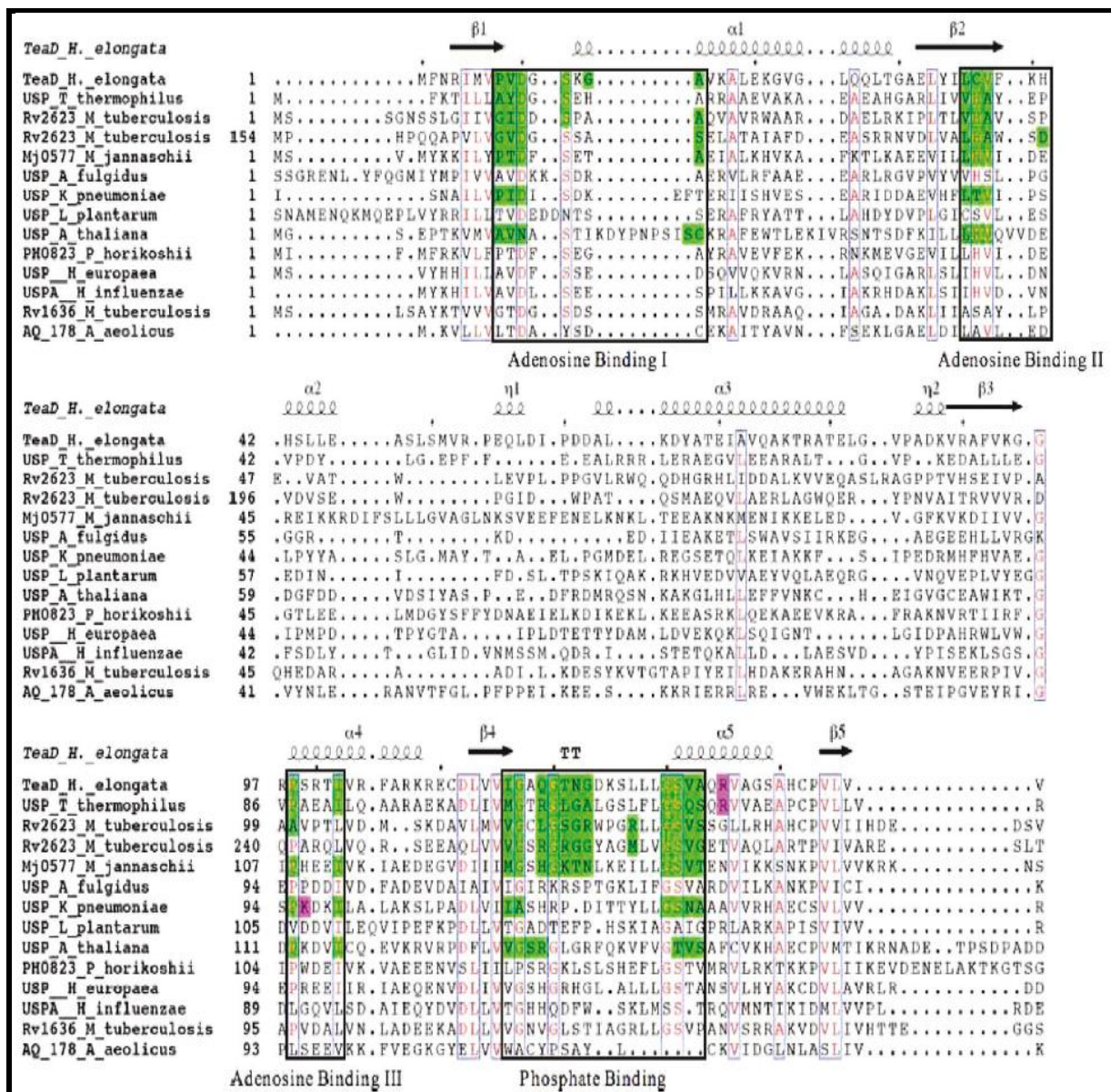


Figure 6.1. Sequence alignment of crystallised USPs. Alignment by Schweikhard *et al.*, 2010.

The Universal Stress Proteins of Bacteria

The alignments attempts to identify key regions of TeaD protein that are in contact with the bound nucleotide. The corresponding regions of Rv1636 are highlighted to play a role in binding include the three regions responsible for adenosine binding, regions I, II and III. The regions incorporated residues G10 to S17, A37 to P42 and P94 to L98, respectively. Phosphate binding site V112 to P129 encompasses the nucleotide binding motif. The region II of Rv1636 has the highest level of difference compared to the other crystallised bound structures. An alanine, serine, alanine and tyrosine do not match the LXV that is suggested to be highly conserved on the $\beta 2$ loop at positions 36-38. Region I, between $\beta 1$ and $\alpha 1$, correlates reasonably with the Rv2623 protein that was bound to ATP. Adenosine binding III contains the P94 which was found to be an important residue in the corresponding P98 in the TeaD protein.

The USP structures are twisted, five stranded, parallel B sheet flanked by four α helices, ($\beta\alpha\beta$ motifs) which have four loops between these beta sheets and a helices. The loops are likely to be responsible for the ATP recognition. It is likely that Rv1636 has two main loops responsible for ATP binding; the loop around the aspartic acid at position D12A and around the G-X₂-G-X₉-GS motif, and the G-X₂-G-X₉-GS loop that contains the phosphoryl/ribosyl-binding residues. A hydrophobic cluster of residues is also present which along with the loops forge the likely adenine/ribose-binding scaffold. The dimerisation of Rv1636 is predicted to bring together two monomers to face each other and the tetramer would bring together the two dimers, along a 2-fold axis that juxtaposes ATP binding pockets from the opposing dimers. The crystalised strand swap observed may be biologically relevant and have a role in the nucleotide binding capabilities of the Rv1636 protein.

Comparison of the Rv1636 structure with other published structures supports the nucleotide binding observed for Rv1636 and suggests which residues are involved. In addition, it is likely that the oligomerisation has a role in Rv1636's function. Comparison with the TeaD USP suggests that the interaction between dimers is required for the nucleotide binding of the protein.

6.4 Intracellular Role of USPS

6.4.1 Role of Rv1636

The biological role of Rv1636 remains unknown. Overexpression of the USP in *M. smegmatis* did not reveal any immediate phenotype under the growth conditions tested. In addition, co-immunoprecipitation of Rv1636 and GST-pull-downs did not reveal any interacting partners.

The Universal Stress Proteins of Bacteria

Rv1636 is described as an iron-regulated protein by Wong *et al.* (1999), who, using two-dimensional (2-D) gel electrophoresis, studied the effects of iron on the protein expression profile of *Mtb*. The use of iron has been linked to expression of full virulence in some gram negative organisms (Litwin & Calderwood 1993). *Mtb* has multiple proteins described as iron-regulated and the acquisition of the iron is likely to be required for growth and long term survival of the mycobacterial (Wong *et al.*, 1999). *Mtb* encodes two siderophores that are small iron chelating molecules able to sequester iron from other iron-containing molecules. Indeed mutants disrupted in the *mbtB* gene, involved in synthesising the cell surface associated siderophore mycobactin, display growth defects in iron-limited conditions and in the human monocyte cell line THP-1 (Voss *et al.*, 2000). Iron is extremely important in the mycobacterial cell and is highly linked to the oxidative status of a cell. Low iron levels can cause oxidative stress as many enzymes involved in protection against oxidative damage contain iron as a ligand. High iron levels can result in the production of free oxygen radicals as a result of the Fenton reaction.

Iron storage proteins and KatG are up-regulated to prevent the toxicity associated with high iron. The *ideR* gene is a major regulator of iron metabolism within mycobacteria and was also shown by Hobson *et al.* (2002), to be up-regulated in high oxidative stress conditions that are encountered by *Mtb* bacilli inside macrophages. The role of IdeR as a positive regulator in response to oxidative stress was supported by Rodriguez *et al.* (2002) who discovered *ideR* mutants of *Mtb* are more susceptible to hydrogen peroxide and superoxides. Also linking iron with oxidative stress is the Dps iron binding protein which can additionally bind DNA to protect the DNA from the ROIs by preventing access to the susceptible nucleic acids (Gupta *et al.*, 2002).

The DosS protein, one part of the DosR/S sensor kinase two component system, can also bind ferrous iron and thus can sense a shift to ferric iron allowing the protein to give a net indicator to the net redox state of the cell (Rustad *et al.*, 2009). DosT binds to a haem bound to O₂ and therefore directly senses the oxygen tension (Cho *et al.*, 2009, Sardiwal *et al.*, 2005). It is interesting to note that the DosR regulon contains six of the tandem USPs. With the Dos regulon responding to iron signals and iron being involved in the regulation of Rv1636, a role for the USPs in iron metabolism or toxicity is an intriguing possibility.

6.4.2 USP Rv2623 has a role in Establishing Chronic Infection in Mice

During the progress of our own studies, evidence was published that suggested an *Mtb* USP, Rv2623, was required for the establishment of a chronic infection in mouse and guinea pig infection models (Drumm *et al.*, 2009). Rv2623 is one of the most highly induced genes when *Mtb* is subjected to hypoxia and nitrosative stress (Voskuil *et al.*, 2003; Park *et al.*, 2003; Ohno *et al.*, 2003). The protein is also upregulated when the bacillus undergoes growth arrest such as when grown in macrophages and under latent conditions during chronic lung infection of mice. Drumm *et al.* (2009) described how Rv2623 over-expression was shown to attenuate the growth of *M. smegmatis* *in vitro*, which is in contrast to the over-expression of Rv1636, and a 2.9Å resolution crystal structure shows a nucleotide binding pocket within the structure. Through site directed mutagenesis, based on the structure of the nucleotide binding pocket, the USP Rv2623 was shown to regulate bacillary growth dependent on its ATP binding.

The USP was required to cause chronic lung infection in guinea pigs. In fact, the group reported the Rv2623 deficient mutant had a hypervirulence phenotype leading to a higher bacterial load and an increase in mortality in both guinea pigs and mice with weakened immune systems (C3H/HeJ mice). The group hypothesised that Rv2623 could act as an ATP-dependent signalling intermediate in a pathway that promotes the persistent infection by *Mtb*. Rv2623 is one of the six USP genes previously mentioned to be induced via the DosR dormancy regulon. Drumm *et al.* (2009) created a knockout mutant in *Mtb* Δ Rv2623 by specific allelic exchange. The Rv2623 gene was replaced with a hygromycin resistance cassette delivered via recombinant mycobacteriophage phAE159. When the *in vitro* growth over 14 days was compared to the wild type *Mtb* H37Rv in 7H9 Middlebrook medium plus OADC supplement, glycerol and Tween 80, and in minimal Sauton's media, it was found that the growth kinetics were comparable. This indicates there may be a redundancy among the USPs of *Mtb* as previously mentioned. Indeed, partial functional overlap of *E. coli* USPs has been shown (Gustavsson *et al.*, 2002; Nachin *et al.*, 2005) and previous work in the Williams group has shown that single *Mtb* USP mutants have no phenotype under a range of conditions.

The regulation of Rv2623 has been characterised in more depth than for the regulation of Rv1636. Rv2623 is upregulated in murine lungs, correlated with the induction of Th1-mediated immunity (Shi *et al.*, 2003). Rv2623 is upregulated in wild type IFN- γ activated macrophages (but not in naïve and NOS2^{-/-} macrophages) in response to nitric oxide (Schnappinger *et al.*, 2003). Rv2623 is

The Universal Stress Proteins of Bacteria

upregulated approximately 20-fold from days 4-30 in non-replicating persistence (NRP) compared to approximately 3-fold in days 6-24 in stationary phase under aerated growth (Voskuil *et al.*, 2004).

Much of the information on the conditions resulting in an upregulation of Mtb USPs comes from the mining of a range of published microarray studies under a large number of conditions. To date, Rv2623 has appeared more often than Rv1636, leading to more being known about the regulation of Rv2623 compared to Rv1636. The regulation information available supports the potential role that the USPs may play in aiding bacterial survival in challenging conditions.

6.4.3 Role of PA3309

The *P. aeruginosa* USP PA3309 was mutated and inserted into the integrating vector, mini-CTX2. Mini-CTX2 was designed to integrate into the *P. aeruginosa* genome at the attachment B (attB) site. Integration of the PA3309 gene containing a site directed mutation was performed with the aim of determining if the putative nucleotide binding motif played a role in the survival via pyruvate fermentation as described by Scheriber *et al.* (2006). A contribution to the role could imply that binding of a nucleotide substrate has a role in the bacterium's survival. Alternatively, it may be that the physiological characteristics of the residue are important; something which could be tested by creating alternative residue substitutions.

Unfortunately, the role of the PA3309 predicted nucleotide binding motif in the anaerobic survival by pyruvate fermentation was not able to be tested. The use of *P. aeruginosa* to study the USPs is potentially a very important resource. The publication of phenotypes for the USP mutants in PA4352 and PA3309 make the organism an important future tool in the research of the biochemical activities. The *P. aeruginosa* knockout library became available to the Williams' group in 2010, consisting of transposon mutants in non-essential genes (Liberati *et al.*, 2006). The knockout library contains 4000 strains, each with a transposon mutation within non-essential genes. The USPs of *P. aeruginosa* were included in the library and the use of the strains may provide a valuable research tool.

6.5 Future work

The challenge to determine the function role of the USPs of *Mtb*, and especially Rv1636, remains. For the biochemical work, the proteins' states *in vivo* could be analysed by 2D-PAGE. The

The Universal Stress Proteins of Bacteria

ability of Rv1636 to interact with nucleotides and the biochemical activities that may operate *in vivo* is likely to influence the interactions of the protein within the cell. Eastern blotting and HPLC analysis may provide further details of the modification of Rv1636 and the *in vivo* state of the protein in different growth phases of *Mtb*.

This work was unsuccessful in identifying any interacting proteins of Rv1636. Identification of interacting partners can provide a large amount of information on the *in vivo* function of a protein. An alternative method could be to use *in vivo* crosslinking. This would involve expressing tagged versions of the protein in a mycobacterial species, crosslinking interacting proteins with formaldehyde or an alternative crosslinker, followed by purification of the USP and identification of co-purified proteins. The over-expression vectors containing RV1636 provide a means to express a higher level of Rv1636 in *M. bovis* BCG or *Mtb* and provide a higher level of protein material for interaction studies.

A key area of future research is to obtain an Rv1636 knockout. Complementation can then be achieved with the over-expression vectors designed and the site directed mutants used to test the *in vivo* role of potentially important residues. The category 2 strain of *Mtb* may provide the opportunity to study the Rv1636 in its native environment. The strain does have certain limitations; however the opportunity to study the bacteria at a lower containment level should be exploited.

Work with *P. aeruginosa* continues to offer the opportunity to determine the role of the predicted nucleotide binding site. With the successful integration of the site directed mutant PA3309 proteins, the project will provide details on the pyruvate fermentation survival.

6.6 General Role of USPs

The physiological state of the persistent bacteria within the latently infected human host is not well understood and the entry into the state of persistence is also largely uncharacterized. Any host and pathogen interaction is likely to be a finely balanced and intricate affair. A thorough understanding of the dynamics of the interaction between the host/pathogen over the time course of the infection, which can be decades for *Mtb*, will allow the identification of targets which may allow the design of novel therapies and treatments. The establishment of the long term latent infection is an ideal area to study. If an advantage to the immune system could be gained, the result could be complete clearance of the bacteria from the human host and loss of the human reservoir of infection. This in turn would limit the huge number of cases of active tuberculosis disease due to re-activation of a previously contained bacterial infection.

The Universal Stress Proteins of Bacteria

Mtb is a well adapted intracellular pathogen. The conditions within the intracellular environment need to be sensed by the bacteria to ensure survival. In order to survive it is clear a tightly controlled system of signalling is required to ensure environmental cues are acted upon in a finely controlled manner. Transducing complex environmental signals would require a highly tuned signalling network to be possessed by *Mtb* to enable the continued survival in the demanding conditions within the cell and to enable the evasion of the array of defence mechanisms that are deployed by the host immune system. The universal stress proteins have been linked to enhanced survival of numerous bacteria. The diverse nature of roles and abundance of copies of this superfamily of proteins within individual organisms may suggest a wide ranging role and it is a challenging proposition to fully comprehend the function of these small ubiquitous proteins.

If we were to speculate on the role of Rv1636, and *Mtb* USPs in general, we could propose that they are acting as signalling proteins involved in the adaptation of the bacillus to its changing environmental conditions. This appears to be the case in *Halomonas elongate* where the TeaD USP is involved in gene regulation, and *Mtb* Rv2623, which regulates growth in an ATP dependent manner. One possibility is that the USPs act to modify other proteins, such as was suggested by Nystrom *et al.* (1996). In addition, the USP's own functions may be regulated by the modification of these proteins, such as the acetylation of a USP described by Nambi *et al.* (2010). There is evidence from *E. coli* that USP domains can be important for signalling – for example, *E. coli* KdpD requires its USP domain to respond to K⁺ limitation and salt stress.

In addition, USP domains have been proposed to bind a range of ligands and cofactors, perhaps allowing the USPs to perform distinct roles under a range of conditions. A signalling role would fit with the phenotypes of USP mutants in a number of species, where it has been seen that the cells lacking certain USPs have difficulty in adapting to various stress conditions. Although this study was unable to confirm some of the proposed activities previously associated with Rv1636, it is still possible that the protein does have some interesting biochemical activities that could allow it to act in some kind of signalling role. It is appealing to consider how the higher order oligomers made by Rv1636 might be involved in its function. However, there is still a large amount of research required before we can fully understand this protein.

Chapter 7 - Bibliography

The Universal Stress Proteins of Bacteria

- Abdool-Karim, S., Churchyard, G., Karim, Q. & Lawn, S. (2009) HIV infection and tuberculosis in South Africa: an urgent need to escalate the public health response. *Lancet*, **6736**: 41–43.
- Abu-Raddad, L. J., Sabatelli, L., Achterberg, J. T., Sugimoto, J. D., Longini, I. M., Jr. Dye, C. & Halloran, M. E. (2009) Epidemiological benefits of more-effective tuberculosis vaccines, drugs, and diagnostics. *Proc. Natl. Acad. Sci. U.S.A.*, **106**: 13980–13985.
- Agarwal, N., Lamichhane, G., Gupta, R., Nolan, S., & Bishai, W. R. (2009) Cyclic AMP intoxication of macrophages by a *Mycobacterium tuberculosis* adenylate cyclase. *Nature*, **460**: 98–102.
- Ahearn, J. M. & Fearon, D. T. (1989) Structure and function of the complement receptors, CR1 (CD35) and CR2 (CD21). *Adv. Immunol.*, **46**: 183–219.
- Ahmed, A. B., Abubakar, I., Delpech, V., Lipman, M., Boccia, D., Forde, J., Antoine, D. & Watson, J. M. (2007) The growing impact of HIV infection on the epidemiology of tuberculosis in England and Wales: 1999–2003. *Thorax*, **62**: 672–676.
- Albalak, R., O'Brien, R. J., Kammerer, J. S., O'Brien, S. M., Marks, S. M., Castro, K. G. & Moore, M. (2007) Trends in tuberculosis / human immunodeficiency virus comorbidity, United States 1993–2004. *Arch. Inter. Med.*, **167**: 2443–2452.
- Andries, K., Verhasselt, P., Guillemont, J., Gohlmann, H., Neefs, J., Winkler, H., Van Gestel, J., Timmerman, P., Zhu, M., Lee, E., Williams, P., De Chaffoy, D., Huitric, E., Hoffner, S., Cambau, E., Truffot-Pernot, C., Lounis, N. & Jarlier, V. (2005) A diarylquinoline drug active on the ATP synthase of *Mycobacterium tuberculosis*. *Science*, **307**: 223–227.
- Anuchin, A. M., Mulyukin, A. L., Suzina, N. S., Duda, V. I., El-Registan, G. I. & Kaprelyants, A. S. (2009) Dormant forms of *Mycobacterium smegmatis* with distinct morphology. *Microbiology*, **155**: 1071–1079.
- Bai, G., Schaak, D. D., & McDonough, K. A. (2009) cAMP levels within *Mycobacterium tuberculosis* and *Mycobacterium bovis* BCG increase upon infection of macrophages. *FEMS Immunol. Med., Microbiol.*, **55**: 68–73.
- Bannon, M. (1999) BCG and tuberculosis. *Arch. Dis. Child.*, **80**: 80–83
- Barry, C.E. 3rd, Boshoff, H. J., Dartois, V., Dick, T., Ehrt, S., Flynn, J., Schnappinger, D., Wilkinson, R. J. & Young, D. (2009) The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. *Nat. Rev. Microbiol.* **7**:845–855.
- Bauer, A. & Kuster, B. (2003) Affinity purification-mass spectrometry. Powerful tools for the characterization of protein complexes. *Eur. J. Biochem.*, **270**: 570–578.
- Ben Amor, Y., Nemser, B., Singh, A., Sankin, A. & Schluger, N. (2008) Underreported threat of multidrug-resistant tuberculosis in Africa. *Emerging Infectious Diseases*, **14**: 1345–1352.

The Universal Stress Proteins of Bacteria

- Bernstein, J. Lott, W. A. Steinberg, B. A. & Yale, H. L. (1952) Chemotherapy of experimental tuberculosis. V. Isonicotinic acid hydrazide (nydrazid) and related compounds. *Am. Rev. Tuberc.*, **65**: 357-364.
- Betts, J. C., Lukey, P. T., Robb, L. C., McAdam, R. A. & Duncan, K. (2002) Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling. *Mol. Microbiol.*, **43**: 717-31.
- Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-ides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, A., Goeden, M. A., Rose, D. J. & Mau, B. (1997) The complete genome sequence of *Escherichia coli* K-12. *Science*, **277**: 1453–1474.
- Blokpoel, M., Murphy, H., O'Toole, R., Wiles, S., Runn, E., Stewart, G., Young, D. & Robertson, B. (2005) Tetracycline-inducible gene regulation in mycobacterial. *Nuc. Acid. Res.*, **33**(2): e22
- Bloom, B. & Murray, C. (1992) Tuberculosis: commentary on a reemergent killer. *Science*, **257**: 1055-64.
- Bochkareva, E.S., Girshovich, A.S. & Bibi, E. (2002) Identification and characterization of the *Escherichia coli* stress protein UP12, a putative *in vivo* substrate of GroEL. *Eur. J. Biochem.*, **269**: 3032-3040.
- Boes, N., Schreiber, K., Hartig, E., Jaensch, L. & Schobert, M. (2006) The *Pseudomonas aeruginosa* universal stress protein PA4352 is essential for surviving anaerobic energy stress. *J. Bacteriol.*, **188**: 6529-6538.
- Boitel, B., Ortiz-Lombardía, M., Durán, R., Pompeo, F., Cole, S. T., Cerveñansky, C. & Alzari, P. M. (2003) PknB kinase activity is regulated by phosphorylation in two Thr residues and dephosphorylation by PstP, the cognate phospho-Ser/Thr phosphatase, in *Mycobacterium tuberculosis*. *Molecular Microbiology*, **49**: 1493-1508.
- Boon, C., Li, R., Qi, R. & Dick, T. (2001) Proteins of *Mycobacterium bovis* BCG induced in the Wayne dormancy model. *J. Bacteriol.*, **183**: 2672-2676.
- Bork, P., & Koonin, E. V. (1994) A P-loop-like motif in a widespread ATP pyrophosphatase domain: implications for the evolution of sequence motifs and enzyme activity. *Proteins*, **20**: 347-355.
- Bork, P., Sander, C., & Valencia, V. (1992) An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and hsc70 heat shock proteins. *Proc. Natl. Acad. Sci. U.S.A.*, **89**: 7290-7294.
- Boshoff, H., Reed, M., Barry, C., & Mizrahi, V. (2003) DnaE2 polymerase contributes to *in vivo* survival and the emergence of drug resistance in *Mycobacterium tuberculosis*. *Cell*, **113**: 183-193.
- Brennan, M. J. (2005) The tuberculosis vaccine challenge. *Tuberculosis*, **85**: 7–12.

The Universal Stress Proteins of Bacteria

Bukau, B. & Horwich, A.L. (1998) The Hsp70 and Hsp60 chaperone machines. *Cell*, **92**: 351–366.

Cappelli, G., Volpe, E., Grassi, M., Liseo, B., Colizzi, V. & Mariani, F. (2006) Profiling of *Mycobacterium tuberculosis* gene expression during human macrophage infection: upregulation of the alternative sigma factor G, a group of transcriptional regulators, and proteins with unknown function. *Res. Microbiol.*, **157**: 445-455.

Chan, J., Fan, X., Hunter, S. W., Brennan, P. J. & Bloom, B. R. (1991) Lipoarabinomannan, a possible virulence factor involved in persistence of *Mycobacterium tuberculosis* within macrophages. *Infect. Immun.*, **59**: 1755-1761.

Chatterjee, D. (1992) Lipoarabinomannan. Multiglycosylated form of the mycobacterial mannosylphosphatidylinositols. *J. Biol. Chem.*, **267**: 6228–6233.

Chen, W., Honma, K., Sharma, A. & Kuramitsu, H. K. (2006) A universal stress protein of *Porphyromonas gingivalis* is involved in stress responses and biofilm formation. *FEMS Microbiol. Lett.*, **264**: 15-21.

Cho, H. Y., Cho, H. J., Kim, Y. M., Oh, J. I. & Kang, B. S. (2009) Structural insight into the heme-based redox sensing by DosS from *Mycobacterium tuberculosis*. *J. Biol. Chem.*, **284**: 13057-13067.

Clemens, D. L., Lee, B. & Horwitz, M. A. (2000) Deviant expression of Rab5 on phagosomes containing the intracellular pathogens *Mycobacterium tuberculosis* and *Legionella pneumophila* is associated with altered phagosomal fate. *Infect. Immun.*, **68**: 2671–2684.

Cohen-Gonsaud, M., Barthe, P., Bagneris, C., Henderson, B., Ward, J., Roumestand, C. & Keep, N. H. (2005) The structure of a resuscitation-promoting factor domain from *Mycobacterium tuberculosis* shows homology to lysozymes. *Nat. Struct. Mol. Biol.*, **12**: 270–273.

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

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

Daffe, M. & Draper, P. (1998) The envelope layers of mycobacteria with reference to their pathogenicity. *Adv. Microb. Physiol.*, **39**: 131–203.

The Universal Stress Proteins of Bacteria

- Dao, D. N., Sweeney, K., Hsu, T., Gurucha, S. S., Nascimento, I. P., Roshevsky, D., Besra, G. S., Chan, J., Porcelli, S. A. & Jacobs, W. R. (2008) Mycolic acid modification by the *mmaA4* gene of *M. tuberculosis* modulates IL-12 production. *PLoS Pathog.*, **4**(6): e1000081.
- Dass, B. K., Sharma, R., Shenoy, A. R., Mattoo, R., & Visweswariah, S. S. (2008) Cyclic AMP in mycobacteria: characterization and functional role of the Rv1647 ortholog in *Mycobacterium smegmatis*. *J. Bacteriol.*, **190**: 3824–3834.
- Delaude, A., Jorda, M. F. & Moreau, G. (1962) Action of ethambutol on experimental tuberculosis in the guinea pig. *Rev. Tuberc. Pneumol.*, **26**: 519-524.
- DeMaio, J., Zhang, Y., Ko, C., Young, D. B., Bishai, W. R. (1996) A stationary-phase stress-response sigma factor from *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U.S.A.*, **93**: 2790-2794.
- den Hengst, C. & Buttner, M. (2008) Redox control in actinobacteria. *Biochem. Biophys. Acta.*, **1780**: 1201-1216.
- Deretic, V. & Fratti, R. A. (1999) *Mycobacterium tuberculosis* phagosome. *Mol. Microbiol.*, **31**: 1603-1609.
- Deretic, V., Delgado, M., Vergne, I., Master, S., De Haro, S., Ponpuak, M. & Singh, S. (2009) Autophagy in immunity against mycobacterium tuberculosis: a model system to dissect immunological roles of autophagy. *Curr. Top. Microbiol. Immunol.*, **335**:169–188.
- Diez, A., Gustavsson, N. & Nystrom, T. (2000) The universal stress protein A of *Escherichia coli* is required for resistance to DNA damaging agents and is regulated by a RecA/FtsKdependent regulatory pathway. *Mol. Microbiol.*, **36**: 1494-1503.
- Doherty, M. Wallis, R. & Zumla, A. (2009) WHO-Tropical Disease Research/European Commission joint expert consultation group. *Curr. Opin. Pulm. Med.*, **15**: 181.
- Doukhan, L., Predich, M., Nair, G., Dussurget, O., Mandic-Mulec, I., Cole, S. T., Smith, D. R. & Smith, I. (1995). Genomic organization of the mycobacterial sigma gene cluster. *Gene*, **165**: 67-70.
- Drumm, J. E., Mi, K., Bilder, P., Sun, M., Lim, J., Bielefeldt-Ohmann, H., Basaraba, R., So, M., Zhu, G., Tufariello, J. M., Izzo, A. A., Orme, I. M., Almo, S.C., Leyh, T. S. & Chan, J. (2009) *Mycobacterium tuberculosis* universal stress protein Rv2623 regulates bacillary growth by ATP-Binding: requirement for establishing chronic persistent infection. *PLoS Pathog.* Epub **5**(5): e1000460.
- Dussurget, O., Stewart, G., Neyrolles, O., Pescher, P., Young, D. & Marchal, G. (2001) Role of *Mycobacterium tuberculosis* copper-zinc superoxide dismutase. *Infect. Immun.*, **69**: 529-533.
- Dye, C., Scheele, S., Dolin, P., Pathania, V., & Raviglione, M. C. (1999) *J. Am. Med. Assoc.*, **282**: 677-686.

The Universal Stress Proteins of Bacteria

- Dziadek, J., Rutherford, S., Madiraju, M., Atkinson, M. & Rajagopalan, M. (2003) Conditional expression of *Mycobacterium smegmatis* ftsZ, an essential cell division gene. *Microbiology*, **149**: 1593-1603.
- Eschbach, M., Schreiber, K., Trunk, K., Buer, J., Jahn, D. & Schobert, M. (2004) Long-term anaerobic survival of the opportunistic pathogen *Pseudomonas aeruginosa* via pyruvate fermentation. *J. Bacteriol.*, **186**: 4596-4604.
- Farewell, A., Diez, A. A., DiRusso, C. C. & Nystrom, T. (1996) Role of the *Escherichia coli* FadR regulator in stasis survival and growth phase-dependent expression of the uspA, fad, and fab genes. *J. Bacteriol.*, **178**: 6443–6450.
- Fisher, M. A., Plikaytis, B. B., & Shinnick, T. M. (2002) Microarray analysis of the *Mycobacterium tuberculosis* transcriptional response to the acidic conditions found in phagosomes. *J. Bacteriol.*, **184**: 4025-4032.
- Flynn, J. & Chan, J. (2003). Immune evasion by *Mycobacterium tuberculosis*: living with the enemy. *Curr. Opin. Immunol.*, **15**: 450-455.
- Flynn, J. L. (2004) Immunology of tuberculosis and implications in vaccine development. *Tuberculosis (Edinb)*, **84**: 93-101.
- Flynn, J. (2006) Lessons from experimental *Mycobacterium tuberculosis* infections. *Microbes Infect.*, **8**: 1179–1188.
- Flynn, J. & Chan, J. (2005) What's good for the host is good for the bug. *Trends Microbiol.*, **13**: 98–102.
- Fox, H. (1952) The chemical approach to the control of tuberculosis. *Science*, **116**: 129-134.
- Freestone, P., Nyström, T., Trinei, M. & Norris, V. (1997) The Universal Stress protein UspA, of *Escherichia coli* is phosphorylated in response to stasis. *J. Mol. Biol.*, **274**: 318-324.
- Freestone, P., Trinei, M., Clarke, S., Nyström, T. & Norris, V. (1998) Tyrosine phosphorylation in *Escherichia coli*. *J. Mol. Biol.*, **279**: 1045-1051.
- French, M. A., Lenzo, N., John, M., Mallal, S. A., McKinnon, E. J., James, I. R., Price, P., Flexman, J. P. & Tay-Kearney, M. L. (2000) Immune restoration disease after the treatment of immunodeficient HIV-infected patients with highly active antiretroviral therapy. *HIV. Med.*, **1**: 107–115.
- Furesz, S., Arioli, V., & Pallanza, R. (1965) Antimicrobial properties of new derivatives of rifamycin SV. *Antimicrob. Agents. Chemother.*, **5**: 770-777.
- Gagneux, S., DeRiemer, K., Van, T., Kato-Maeda, M., de Jong, B. C. Narayanan, S. Nicol, M., Niemann, S. Kremer, K., Gutierrez, M. C., Hilty, M., Hopewell, P. C. & Small, P. M. (2006) Variable host-pathogen compatibility in *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U.S.A.*, **103**: 2869-2873.

The Universal Stress Proteins of Bacteria

- Ganguli, S., Wang, H., Matsumura, P., & Volz, K. (1995) Uncoupled phosphorylation and activation in bacterial chemotaxis. The 2.1-A structure of a threonine to isoleucine mutant at position 87 of CheY. *J. Biol. Chem.*, **270**: 17386-17393.
- Ganguly, N., Giang, P. H., Gupta, C., Basu, S. K., Siddiqui, I., Salunke, D. M. & Sharma, P. (2008) *Mycobacterium tuberculosis* secretory proteins CFP-10, ESAT-6 and the CFP10:ESAT6 complex inhibit lipopolysaccharide-induced NF-kappaB transactivation by downregulation of reactive oxidative species (ROS) production. *Immunol. Cell Biol.*, **86**: 98-106.
- Garton, N. J., Waddel, S. J., Sherratt, A. L., Lee, S., Smith, R. J. Senner, C., Hinds, J., Rajakumar, K., Adegbola, R. A., Besra, G. S., Butcher, P. D. & Barer, M. R. (2008) Cytological and transcript analyses reveal fat and lazy persister-like bacilli in tuberculous sputum. *PLoS Med.*, **5**: 634–645.
- Gazdik, M. & McDonough, K. (2005) Identification of cyclic AMP-regulated genes in *Mycobacterium tuberculosis* complex bacteria under low-oxygen conditions *J. Bacteriol.*, **187**: 2681–2692.
- Ghosh, J., Larsson, P., Singh, B., Pettersson, B. M., Islam, N. M., Sarkar, S. N., Dasgupta, S. & Kirsebom, L. A. (2009) Sporulation in mycobacteria. *Proc. Natl. Acad. Sci. U.S.A.*, **106**: 10781–10786.
- Gill, W., Harik, N., Whiddon, M., Liao, R., Mittler, J., & Sherman, D. R. (2009) A replication clock for *Mycobacterium tuberculosis*. *Nat. Med.*, **15**: 211–214.
- Gordon, A. H. & Hart P. D. (1994) Stimulation or inhibition of the respiratory burst in cultured macrophages in a mycobacterium model: initial stimulation is followed by inhibition after phagocytosis. *Infect Immun.*, **62**: 4650-4651.
- Goren, M. B. (1977) Phagocyte lysosomes: interactions with infectious agents, phagosomes, and experimental perturbations in function. *Annu. Rev. Microbiol.*, **31**: 507–533.
- Gragerov, A., Nudler, E., Komissarova, N., Gaitanaris, G.A., Gottesman, M. E. & Nikiforov, V. (1992) Cooperation of GroEL/GroES and DnaK/DnaJ heat shock proteins in preventing protein misfolding in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.*, **89**: 10341–10344.
- Guo, X., Monteleone, M., Klotzsche, M., Kamionka, A., Hillen, W., Braunstein, M., Ehrt, S. & Schnappinger, D. (2007) Silencing *Mycobacterium smegmatis* by using tetracycline repressors. *J. Bacteriol.*, **189**: 4614-4623.
- Gupta, S., Pandit, S. B., Srinivasan, N., & Chatterji, D. (2002) Proteomics analysis of carbon-starved *Mycobacterium smegmatis*: induction of Dps-like protein. *Protein Eng.*, **15**: 503-512.
- Gury, J., Seraut, H., Tran, N. P., Barthelmebs, L., Weidmann, S., Gervais, P. & Cavin, J. F. (2009) Inactivation of PadR, the repressor of the phenolic acid stress response, by molecular interaction with Usp1, a universal stress protein from *Lactobacillus plantarum*, in *Escherichia coli*. *App. Env. Microbio.*, **75**: 5273-5283.

The Universal Stress Proteins of Bacteria

- Gustavsson, N., Diez, A. & Nystrom, T. (2002) The universal stress protein paralogues of *Escherichia coli* are co-ordinately regulated and co-operate in the defence against DNA damage. *Mol. Microbiol.*, **43**: 107-117.
- Hackam, D. J., Roptstein, O. D., Zhang, W. J., Cruenheid, S. P. & Grinstein, S. (1998) Host resistance to intracellular infection: mutation of natural resistance-associated macrophages protein 1 (Nrap 1) impairs phagosomal acidification. *J. Exp. Med.* **188**: 351–364.
- Hanekom, W. A. Lawn, S. D. Dheda, K. & Whitelaw, A. (2010) Tuberculosis research update. *Trop. Med. Intern. Health*, **15**: 981-989.
- Hanks, S. K. & Quinn, A. M. (1991) Protein kinase catalytic domain database: Identification of conserved features of primary structure and classification of family members. *Methods Enzymol.*, **200**: 38-62
- Hanscheld, T. (2008) The future looks bright: low-cost fluorescent microscopes for detection of *Mycobacterium tuberculosis* and Coccidia. *T. Roy. Soc. Trop. Med. H.*, **102**: 520–521.
- Harding, C. V. & Boom, W. H. (2010) Regulation of antigen presentation by *Mycobacterium tuberculosis*: a role for Toll-like receptors. *Nat. Rev. Microbiol.*, **8**: 296-307.
- Hart, D., Young, M. R., Jordan, M. W., Perkins, W. J. & Geisow, M. J. (1983) Chemical inhibitors of phagosome–lysosome fusion in cultured macrophages also inhibit salutary lysosomal movements. A combined microscopic and computer study. *J. Exp. Med.*, **158**: 477–492.
- Heermann, R., Weber, A., Mayer, B., Ott, M., Hauser, E., Gabriel, G., Pirch, T. & Jung, K. (2009a) The universal stress protein UspC scaffolds the KdpD/KdpE signaling cascade of *Escherichia coli* under salt stress. *J. Mol. Biol.* **386**: 134-48.
- Heermann, R., Lippert, M. L. & Jung, K. (2009b) Domain swapping reveals that the N-terminal domain of the sensor kinase KdpD in *Escherichia coli* is important for signaling. *BMC Microbiol.*, **9**: 133-145.
- Hegymegi-Barakonyi, B., Szekely, R., Varga, Z., Kiss, R., Borbely, G., Nemeth, G., Banhegyi, P., Pato, J., Greff, Z., Horvath, Z., Meszaros, G., Marosfalvi, J., Eros, D., Szantai-Kis, C., Breza, N., Garavaglia, S., Perozzi, S., Rizzi, M., Hafenbradl, D. K. M., Av-Gay, Y., Klebl, B. M., Orfi, L. & Keri, G. (2008) Signalling inhibitors against *Mycobacterium tuberculosis* - early days of a new therapeutic concept in tuberculosis. *Curr. Med. Chem.* **15**: 2760-2770.
- Hoang, T. T. Karkhoff-Schweizer, R. R. Kutchma, A. J. & Schweizer, H. P. (1998) A broad-host-range Flp-FRT recombination system for site-specific excision of chromosomally-located DNA sequences: application for isolation of unmarked *Pseudomonas aeruginosa* mutants. *Gene*, **212**: 77-86.
- Hoang, T. T., Kutchma, A. J., Becher, A. & Schweizer, H. P. (2000) Integration-proficient plasmids for *Pseudomonas aeruginosa*: site-specific integration and use for engineering of reporter and expression strains. *Plasmid*, **43**: 59-72.

The Universal Stress Proteins of Bacteria

- Hobson, R. J., McBride, A. J. Kempself, K. E. & Dale, J. W. (2002) Use of an arrayed promoter-probe library for the identification of macrophage-regulated genes in *Mycobacterium tuberculosis*. *Microbiology*, **148**: 1571-1579.
- Hoft, D. (2008) Tuberculosis vaccine development: goals, immunological design, and evaluation. *Lancet*, **372**: 164–175.
- Horwitz, M. A., Harth, G., Dillon, B. J. & Maslesa-Galic, S. (2006) A novel live recombinant mycobacterial vaccine against bovine tuberculosis more potent than BCG. *Vaccine*, **24**: 1593-1600.
- Hu, Y. & Coates A. R. (1999) Transcription of the stationary-phase-associated hspX gene of *Mycobacterium tuberculosis* is inversely related to synthesis of the 16-kilodalton protein. *J. Bacteriol.*, **181**: 1380-1387.
- Hu, Y., Mangan, J. A., Dhillon, J., Sole, K. M., Mitchison, D. A., Butcher, P. D., & Coates, A. R. (2000) Detection of mRNA transcripts and active transcription in persistent *Mycobacterium tuberculosis* induced by exposure to rifampin or pyrazinamide. *J. Bacteriol.*, **182**: 6358-6365.
- Jaffe, A. B. & Hall, A. (2005) Rho GTPases: biochemistry and biology. *Annu. Rev. Cell and Devel. Bio.*, **21**: 247-269.
- Jenkins, D. E., Scshultz, J. E. & Matin, A. (1988) Starvation-induced cross-protection against heat or H₂O₂ challenge in *Escherichia coli*. *J. Bacteriol.*, **170**: 3910-3914
- Jones, P. G., Cashel, M., Glaser, G. & Neidhardt, F. C. (1992) Function of a relaxed-like state following temperature downshifts in *Escherichia coli*. *J. Bacteriol.*, **174**: 3903-3914.
- Jung, K., Veen, M. & Altendorf, K. (2000) K⁺ and ionic strength directly influence the autophosphorylation activity of the putative turgor sensor KdpD of *Escherichia coli*. *J. Biol. Chem.*, **275**: 40142-40147.
- Kana, B. D. & Mizrahi, V. (2010) Resuscitation-promoting factors as lytic enzymes for bacterial growth and signaling. *FEMS Immunol. Med. Microbiol.*, **8**: 39-50.
- Karakousis, P. C., Yoshimatsu, T., Lamichhane, G., Woolwine, S. C., Nuermberger, E. L., Grosset, J., & Bishai, W. R. (2004) Dormancy phenotype displayed by extracellular *Mycobacterium tuberculosis* within artificial granulomas in mice. *J. Exp. Med.*, **200**: 647–657.
- Karimova, G. (2001) Protein-protein interaction between *Bacillus stearothermophilus* tyrosyl-tRNA synthase subdomains revealed by a bacterial two hybrid system. *J. Mol. Microbiol. Biotech.*, **3**: 73-82.
- Kaufmann, S. H. (2001) How can immunology contribute to the control of tuberculosis? *Nat. Rev. Immun.*, **1**: 20–30.
- Kim, J. Parker, J. S. Murray, K. E. & Nibert, M. L. (2003) Nucleoside and RNA triphosphatase activities of orthoreovirus transcriptase cofactor mu2. *J. Biol. Chem.*, **279**: 4394-403.

The Universal Stress Proteins of Bacteria

- Kolter, R., Siegele, D. A. & Tormo, A. (1993) The stationary phase of the bacterial life cycle. *Annu. Rev. Microbiol.*, **47**: 855–874.
- Kumar, A., Deshane, J. S., Crossman, D. K., Bolisetty, S., Yan, B. S., Kramnik, I., Kramnik, I., Agarwal, A. & Steyn, A. J. (2008) Heme oxygenase-1-derived carbon monoxide induces the *Mycobacterium tuberculosis* dormancy regulon. *J. Biol. Chem.*, **283**: 18032–18039.
- Kvint, K., Hosbond, C., Farewell, A., Nybroe, O. & Nyström, T. (2000) Emergency derepression: stringency allows RNA polymerase to override negative control by an active repressor. *Mol. Microbiol.*, **35**: 435–443.
- Kvint, K., Nachin, L., Diez, A. & Nyström, T. (2003) The bacterial universal stress protein: function and regulation. *Cur. Opin. Microbiol.*, **6**: 140–145.
- Lange, R. & Hengge-Aronis, R. (1991) Identification of a central regulator of stationary phase gene expression in *Escherichia coli*. *Mol. Microbiol.*, **5**: 49–59.
- Lazarus, J., Olsen, M., Ditiu, L. & Matic, S. (2008) Tuberculosis-HIV co-infection: policy and epidemiology in 25 countries in the WHO European region. *HIV Medicine*, **9**, 406–414.
- Leistikow, R. L., Morton, R. A., Bartek, I. L., Frimpong, I., Wagner, K., & Voskuil, M. I. (2010) The *Mycobacterium tuberculosis* DosR regulon assists in metabolic homeostasis and enables rapid recovery from nonrespiring dormancy. *J. Bacteriol.*, **6**: 1662–1670.
- Leiva, J. (2007) The universal stress protein Rv1636 of *Mycobacterium tuberculosis*. PhD Thesis, Imperial College.
- Liberati, N., Urbach, J., Miyata, S., Lee, D., Drenkard, E., Wu, G., Villanueva, J., Wei, T. & Ausubel, F. (2006) An ordered, nonredundant library of *Pseudomonas aeruginosa* strain PA14 transposon insertion mutants. *Proc. Natl. Acad. Sci. U.S.A.*, **103**: 2833–2838.
- Lillebaek, T., Dirksen, A., Baess, I., Strunge, B. & Thomsen, V. (2002) Molecular evidence of endogenous reactivation of *Mycobacterium tuberculosis* after 33 years of latent infection. *J. Infect. Dis.*, **185**: 401–404.
- Lin, P. L., Rodgers, M., Smith, L., Bigbee, M., Myers, A., Bigbee, C., Chiosea, I., Capuano, S. V., Fuhrman, C., Klein, E. & Flynn, J. L. (2009) Quantitative comparison of active and latent tuberculosis in the cynomolgus macaque model. *Infect. Immun.*, **77**: 4631–42.
- Litwin, C. M. & Calderwood, S. B. (1993) Role of iron in regulation of virulence genes. *Clin. Microbiol. Rev.*, **6**: 137–149.
- Liu, W. T., Karavolos, M. H., Bulmer, D. M., Allaoui, A., Hormaeche, R. D., Lee, J. J. & Anjam Khan, C. M. (2006) Role of the universal stress protein UspA of *Salmonella* in growth arrest, stress and virulence. *Microb. Pathog.*, **42**: 2–10.

The Universal Stress Proteins of Bacteria

Lodish, H., Berk, H., Zipursky, L. S., Matsudaira, P. & Baltimore, D. (2000) Molecular Cell Biology, 4th Ed. Chapter. 24, Cancer.

Lopez, B., Aguilar, D., Orozco, H., Burger, M., Espitia, C., Ritacco, V., Barrera, L., Kremer, K., Hernandez-Pando, R., Huygen, K. & van Soolingen, D. (2003) A marked difference in pathogenesis and immune response induced by different *Mycobacterium tuberculosis* genotypes. *Clin. Exp. Immunol.*, **133**: 30-37.

Lowrie, D. B., Jackett, P. S., & Ratcliffe, N. A. (1975) *Mycobacterium microti* may protect itself from intracellular destruction by releasing cyclic AMP into phagosomes. *Nature*, **254**: 600–602.

Lyczak, J. B., Cannon, C. L. & Pier, G. B. (2000) Establishment of *Pseudomonas aeruginosa* infection: lessons from a versatile opportunist. *Microbes and Infection*, **2**: 1051-1060.

Manabe, Y. C. & Bishai, W. R. (2000) Latent *Mycobacterium tuberculosis*-persistence, patience, and winning by waiting. *Nat. Med.*, **6**: 1327-1329.

Manganelli, R., Dubnau, E., Tyagi, S., Kramer, F. R. & Smith, I. (1999) Differential expression of 10 sigma factor genes in *Mycobacterium tuberculosis*. *Mol. Microbiol.*, **31**: 715-724.

Manganelli, R., Provvedi, R., Rodrigue, S., Beaucher, J., Gaudreau, L. & Smith, I. (2004) Sigma factors and global gene regulation in *Mycobacterium tuberculosis*. *J. Bacteriol.*, **186**: 895-902.

Manganelli, R., Voskuil, M. I., Schoolnik, G. K., Dubnau, E., Gomez, M. & Smith, I. (2002) Role of the extracytoplasmic-function sigma factor sigma(H) in *Mycobacterium tuberculosis* global gene expression. *Mol. Microbiol.*, **45**: 365-374.

Markham, K., Bai, Y. & Schmitt-Ulms, G. (2007) Co-immunoprecipitations revisited: an update on experimental concepts and their implementation for sensitive interactome investigations of endogenous proteins. *Anal. Bioanal. Chem.*, **398**: 461-473.

Mawuenyega, K., Forst, C. V., Dobos, K. M., Belisle, J. T., chen, J., Bradbury, E. M., Bradbury, A. R. M. & Chen, X. (2005) *Mycobacterium tuberculosis* functional network analysis by global subcellular protein profiling. *Mol. Biol. Cell*, **16**: 396-404.

Maxam, A. M. & Gilbert, W. (1977) A new method for sequencing DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **74**: 560-563.

McDonough, K. A., Kress, Y. & Bloom, B. R. (1993) Pathogenesis of tuberculosis Interaction of *Mycobacterium tuberculosis* with macrophages. *Infect. Immun.*, **61**: 2763–2773.

McFadden, J. (1996) Recombination in mycobacterial. *Mol. Microbiol.*, **21**: 205–211.

McKinney, J. D., Bentrup, K. H. Z., Munoz-Elias, E. J., Miczak, A., Chen, B., Chan, W., Swenson, D., Sacchettinl, J. C., Jacobs, W. R. & Russel, D. G. (2000) Persistence of *Mycobacterium tuberculosis* in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase. *Nature*, **406**: 735–738.

The Universal Stress Proteins of Bacteria

- McShane, H., Pathan, A. A., Sander, C. R., Keating, S. M., Gilbert, S. C., Huygen, K., Fletcher, H. A. & Hill, A. V (2004) Recombinant modified vaccinia virus Ankara expressing antigen 85A boosts BCG-primed and naturally acquired antimycobacterial immunity in humans. *Nat. Med.*, **10**: 1240–1244.
- Metchnikoff, E. (1905) Immunity to Infective Disease. Cambridge University Press; Cambridge, London, New York.
- Mitchison, D. A. (2000) Role of individual drugs in the chemotherapy of tuberculosis. *Int. J. Tuberc. Lung. Dis.*, **4**: 796-806.
- Mohan, V. P., Scanga, C. A., Yu, K., Scott, H. M., Tanaka, K. E., Tsang, E., Tsai, M. C., Flynn, J. L. & Chan, J. (2001) Tumor necrosis factor- α is required to prevent reactivation and limit pathology in latent murine tuberculosis. *Infect. Immun.*, **69**: 1847-1855.
- Mukamolova, G., Turapov, O., Malkin, J., Woltmann, G., & Barer, M.R. (2010) Resuscitation promoting factors reveal an occult population of tubercle bacilli in sputum. *Am. J. Respir. Crit. Care Med.* **181**: 174–180.
- Mukamolova, G. V., Turapov, O. A., Young, D. I., Kaprelyants, A. S., Kell, D. S. & Young, M. (2002) A family of autocrine growth factors in *Mycobacterium tuberculosis*. *Mol. Microbiol.*, **46**: 623–635.
- Mukhopadhyay, S., Kapatral, V., Xu, W. & Chakrabarty, A. M. (1999) Characterization of a Hank's type serine/threonine kinase and serine/threonine phosphoprotein phosphatase in *Pseudomonas aeruginosa*. *J. Bacteriol.*, **181**: 6615-6622.
- Murray, T., M. Egan & B.I. Kazmierczak (2007) *Pseudomonas aeruginosa* chronic colonization in cystic fibrosis patients. *Curr. Opin. Pediatr.*, **19**: 83-88.
- Murry, J., Sassetti, C. M., Moreira, J., Lane, J., & Rubin, E. J. (2005) A new site-specific integration system for mycobacterial. Tuberculosis, **85**: 317-323.
- Nachin, L., Nannmark, U. & Nystrom, T. (2005) Differential roles of the universal stress proteins of *Escherichia coli* in oxidative stress resistance, adhesion, and motility. *J. Bacteriol.*, **187**: 6265–6272.
- Nagel, G. M., & Doolittle, R. F. (1991) Evolution and relatedness in two aminoacyl-tRNA synthetase families. *Proc. Natl. Acad. Sci. U.S.A.*, **88**: 8121-8125.
- Nambi, S., Basu, N. & Visweswariah, S. (2010) cAMP-regulated protein lysine acetylases in mycobacteria. *J. Biol. Chem.*, **285**: 24313–24323.
- Nathan, C., & Shiloh, M. U. (2000) Reactive oxygen and nitrogen intermediates in the relationship between mammalian hosts and microbial pathogens. *Proc. Natl. Acad. Sci. USA*, **97**: 8841–8848.
- Nesbitt, N. M., Yang, X., Fontan, P., Kolesnikova, I., Smith, I., Sampson, N. S. & Dubnau, E. (2010) A thiolase of *Mycobacterium tuberculosis* is required for virulence and production of androstenedione and androstadienedione from cholesterol. *Infect. Immun.*, **78**: 275–282.

The Universal Stress Proteins of Bacteria

- Nunn, P., Zignol, M., Jaramillo, E., Wright, A. & Getahun, H. (2007) Tb drug resistance: is it really a threat to Africa? *Ethiopian Medical Journal*, **45**: 399–404.
- Nyka, W. (1974) Studies on the effect of starvation on mycobacteria. *Infect. Immun.*, **9**: 843–850.
- Nystrom, T. & Neidhardt, F. C. (1992) Cloning, mapping and nucleotide sequencing of a gene encoding a universal stress protein in *Escherichia coli*. *Mol. Microbiol.*, **6**: 3187–3198.
- Nystrom, T. & Neidhardt, F. C. (1996) Effects of overproducing the universal stress protein, UspA, in *Escherichia coli* K-12. *J. Bacteriol.*, **178**: 927–930.
- Nystrom, T. & Neidhardt, F.C. (1994) Expression and role of the universal stress protein, UspA, of *Escherichia coli* during growth arrest. *Mol. Microbiol.*, **11**: 537–544.
- Nyström, T., & F. C. Neidhardt. (1993) Isolation and properties of a mutant of *Escherichia coli* with an insertional inactivation of the *uspA* gene, which encodes a universal stress protein. *J. Bacteriol.*, **175**: 3949–3956.
- O'Toole, R. & Williams, H. D. (2003) Universal stress proteins and *Mycobacterium tuberculosis*. *Res. Microbiol.*, **154**: 387–392.
- O'Toole, R., Smeulders, M. J., Blokpoel, M. C., Kay, E. J., Loughheed, K. & Williams, H. D. (2003) A two-component regulator of universal stress protein expression and adaptation to oxygen starvation in *Mycobacterium smegmatis*. *J. Bacteriol.*, **185**: 1543–1554.
- Oda, Y., Kuo, M. D., Huang, S. S. & Huang, J. S. (1993) The major acidic fibroblast growth factor (aFGF)-stimulated phosphoprotein from bovine liver plasma membranes has aFGF-stimulated kinase, autoadenylation, and alkaline nucleotide phosphodiesterase activities. *J. Biol. Chem.*, **268**, 27318–27326.
- Offe H., & Domagk, G. (1952) The tuberculostatic activity of hydrazine derivatives from pyridine carboxylic acids and carbonyl compounds. *Naturforsch.* **7b**: 462.
- Ohno, H., Zhu, G., Mohan, V. P., Chu, D., Kohno, S., Jacobs, W. R. & Chan, J. (2003) The effects of reactive nitrogen intermediates on gene expression in *Mycobacterium tuberculosis*. *Cell. Microbiol.*, **5**: 637–648.
- Ostapchuk, P. & Hearing, P. (2008) Adenovirus IVa2 protein binds ATP. *J. Virology*, **82**: 10290–10294.
- Pandey, A. K., & Sassetti, C. M. (2008) Mycobacterial persistence requires the utilization of host cholesterol. *Proc. Natl. Acad. Sci. USA*, **105**: 4376–4380.
- Pang, X., Vu, P., Byrd, T.F., Ghanny, S., Soteropoulos, P., Mukamolova, G.V., Wu, S., Samten, B. & Howard, S. T. (2007) Evidence for complex interactions of stress-associated regulons in an *mprAB* deletion mutant of *Mycobacterium tuberculosis*. *Microbiology*, **153**: 1229–1242.

The Universal Stress Proteins of Bacteria

- Park, H. D., Guinn, K. M., Harrell, M. I., Liao, R., Voskuil, M. I., Tompa, M., Schoolnik, G. K. & Sherman, D. R. (2003) Rv3133c/dosR is a transcription factor that mediates the hypoxic response of *Mycobacterium tuberculosis*. *Mol. Microbiol.*, **48**: 833–843.
- Parrish, N. M., Dick, J.D. & Bishai, W.R. (1998) Mechanisms of latency in *Mycobacterium tuberculosis*. *Trends in Microbiology*, **6**: 107-112.
- Patton, C. Thompson, S. & Epel, D. (2004) Some precautions in using chelators to buffer metals in biological solutions. *Cell Calcium*, **35**: 427-31.
- Payne, D. J., Gwynn, M. N., Holmes, D. J. & Pompliano, D. L. (2007) Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat. Rev. Drug. Discov.*, **6**: 29-40.
- Peirs, P., De Wit, L., Braibant, M., Huygen, K. & Content, J. (1997) A serine / threonine protein kinase from *Mycobacterium tuberculosis*. *Eur. J. Biochem.*, **244**: 604–612.
- Pérez, J., Castañeda-García, A., Jenke-Kodama, H., Müller, R. & Muñoz-Dorado, J. (2008) Eukaryotic-like protein kinases in the prokaryotes and the myxobacterial kinome. *Proc. Natl. Acad. Sci. U.S.A.*, **105**: 15950–15955.
- Pethe, K., Sequeira, P. C., Agarwalla, S., Rhee, K., Kuhen, K., Phong, W. Y. Patel, V., Beer, D., Walker, J. R., Duraiswamy, J., Jiricek, J. Keller, T. H., Chatterjee, A., Tan, M. P., Ujjini, M., Rao, S. P., Camacho, L., Bifani, P., Mak, P. A., Ma, I., Barnes, S. W., Chen, Z., Plouffe, D., Thayalan, P., Ng, S. H., Au, M., Lee, B. H., Tan, B. H., Ravindran, S., Nanjundappa, M., Lin, X., Goh, A., Lakshminarayana, S. B., Shoen, ., Cynamon, M., Kreiswirth, B., Dartois, V., Peters, E. C., Glynn, R., Brenner, S., & Dick, T. (2010) A chemical genetic screen in *Mycobacterium tuberculosis* identifies carbon-source-dependent growth inhibitors devoid of in vivo efficacy. *Nat. Commun.*, **1**: 1-8.
- Prabhakaran, K., Harris, E. B. & Randhawa, B. (2000) Regulation by protein kinase of phagocytosis of *Mycobacterium leprae* by macrophages. *J. Med. Microbiol.*, **49**: 339–342.
- Preneta, R. Papavinasasundaram, K. G. Cozzone, A. J. & Duclos, B (2004) Autophosphorylation of the 16 kDa and 70 kDa antigens (Hsp 16.3 and Hsp 70) of *Mycobacterium tuberculosis*. *Microbiology*, **150**: 2135-2141.
- Prisic, S., Dankwa, S., Schwartz, D., Chou, M. F., Locasale, J. W., Kang, C. M., Bemis, G., Church, G. M., Steen, H. & Husson, R. N. (2010) Extensive phosphorylation with overlapping specificity by *Mycobacterium tuberculosis* serine/threonine protein kinases. *Proc. Natl. Acad. Sci. U.S.A.*, **107**: 7521-7526
- Pulverer, B. J., Kyriakis, J. M., Aruch, J., Nikolakaki, E. & Woodgett, J. R. (1991) Phosphorylation of c-jun mediated by MAP kinases. *Nature*, **353**: 670-674.
- Rajajashankar, R., Kniewel, R., Solorzano, V., Lima, C. D. & Burley, S. K. (2004) Crystal Structure of Rv1636. Protein data bank (1TQ8). New York.

The Universal Stress Proteins of Bacteria

Ralston, G. (1993) Introduction to Analytical Ultracentrifugation. Beckman.

Raman, S. Hazra, R. Dascher, C. C. & Husson, R. N. (2004) Transcription regulation by the *Mycobacterium tuberculosis* alternative sigma factor SigD and its role in virulence. *J. Bacteriol.*, **186**: 6605-6616.

Ratledge, C. (1982) Part 1 C. Ratledge, Nutrition, growth and metabolism. In: C. Ratledge and J. Stanford, Editors, *The Biology of the Mycobacteria* Vol. 1, Academic Press, La Habana (1982).

Rickman, L., Scott, C., Hunt, D. M., Hutchinson, T., Menéndez, M. C., Whalan, R., Hinds, J., Colston, M. J., Green, J. & Buxton, R. S. (2005) 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.* **56**: 1274-1286.

Roberts, D., Liao, R., Wisedchaisri, G., Hol, W. & Sherman, D. (2004) Two sensor kinases contribute to the hypoxic response of *Mycobacterium tuberculosis*. *J. Biol. Chem.*, **279**: 23082–23087

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

Rohde, K., Yates, R. M., Purdy, G. E., & Russell, D. G. (2007) *Mycobacterium tuberculosis* and the environment within the phagosome. *Immunol. Rev.*, **219**: 37–54.

Rombel, I. North, A. Hwang, I. Wyman, C. Kustu, S. (1999) The bacterial enhancer-binding protein NtrC as a molecular machine. *Cold Spring Harb. Symp. Quant. Biol.*, **63**: 157-166.

Rosenkrands, I., Weldingh, K., Jacobsen, S., Hansen, C. V., Florio, W., Gianetri, I. & Andersen, P. (2000) Mapping and identification of *Mycobacterium tuberculosis* proteins by two-dimensional gel electrophoresis, microsequencing and immunodetection. *Electrophoresis*, **21**: 935-948.

Russell, D. G., Cardona, P. J., Kim, M. J., Allain, S., & Altare, F. (2009) Foamy macrophages and the progression of the human tuberculosis granuloma. *Nat. Immunol.*, **10**: 943–948.

Rustad, T. R. Sherrid, A. M. Minch, K. J. & Sherman, D. R. (2009) Hypoxia: a window into *Mycobacterium tuberculosis* latency. *Cell. Microbiol.*, **11**: 1151-1159.

Rustad, T. R., Harrell, M. I., Liao, R. & Sherman, D. R. (2008) The enduring hypoxic response of *Mycobacterium tuberculosis*. *PLoS One*, **3**: e1502.

Sabra, W., Kim, E. J. & Zeng, A. P. (2002) Physiological responses of *Pseudomonas aeruginosa* PAO1 to oxidative stress in controlled microaerobic and aerobic cultures. *Microbiology*, **148**: 3195-3202.

Sardiwal, S., Kendall, S. L., Movahedzadeh, F., Rison, S. C., Stoker, N. G. & Djordjevic, S. (2005) A GAF domain in the hypoxia/NO-inducible *Mycobacterium tuberculosis* DosS protein binds haem. *J. Mol. Biol.*, **353**: 929-936.

The Universal Stress Proteins of Bacteria

- Saunders, B. & Britton, W. (2007) Life and death in the granuloma: immunopathology of tuberculosis. *Immunol. Cell Biol.*, **85**: 103–111.
- Sauter, M., Rzewuski, G., Marwedel, T. & Lorbiecke, R. (2002) The novel ethyleneregulated gene OsUsp1 from rice encodes a member of a plant protein family related to prokaryotic universal stress proteins. *J. Exp. Bot.*, **379**:2325–31.
- Schatz, A., Bugie, E. & Waksman, S. A. (2005) Streptomycin, a substance exhibiting antibiotic activity against gram-positive and gram-negative bacteria. 1944. *Clin. Orthop. Relat. Res.*, **437**: 3-6.
- Schlesinger, L.S., Kaufmann, T.M., Iyer, S., Hull, S.R. & Marchiando, L.K. (1996) Differences in mannose receptor mediated uptake of lipoarabinomannan from virulent and attenuated strains of *Mycobacterium tuberculosis* by human macrophages. *J. Immunol.*, **157**: 4558–4575
- Schnappinger, D., Ehrt, S., Voskuil, M. I., Liu, Y., Mangan, J. A., Monahan, I. M., Dolganov, G., Efron, B., Butcher, P. D., Nathan, C. & Schoolnik, G. K. (2003) *Transcriptional adaptation of Mtb within macrophages*. *J. Exp. Med.*, **198**: 693-704.
- Schreiber, K., Boes, N., Eschbach, M., Jaensch, L., Wehland, J., Bjarnsholt, T., Givskov, M., Hentzer, M. & Schobert, M. (2006) Anaerobic survival of *Pseudomonas aeruginosa* by pyruvate fermentation requires an Usp-type stress protein. *J. Bacteriol.*, **188**: 659-668.
- Schweikhard, E., Kuhlmann, S., Kunte, H., Grammann, K. & Ziegler, C. (2010) Structure and function of the universal stress protein TeaD and its role in regulating the ectoine transporter TeaABC of *Halomonas elongata* DSM 2581(T). *Biochemistry*. **49**:2194-2204.
- Schweizer, H. P. (1991) *Escherichia-Pseudomonas* shuttle vectors derived from pUC18/19. *Gene*. **97**: 109-112.
- Siegele, D. A. (2005) Universal Stress Proteins in *Escherichia coli*. *J. Bacteriol.*, **187**: 6253-6254.
- Seiler, P., Ulrichs, T., Banderhmann, S., Pradl, L., Jorg, S., Krenn, V., Morawietz, L., Kaufmann, S. H. & Aichele, P. (2003) Cell-wall alterations as an attribute of *Mycobacterium tuberculosis* in latent infection. *J. Infect. Dis.* **188**:1326–1331.
- Sherman, D. R., Voskuil, M., Schnappinger, D., Liao, R., Harrell, M. I. & Schoolnik, G. K. (2001) Regulation of the *Mycobacterium tuberculosis* hypoxic response gene encoding a-crystallin. *Proc. Natl. Acad. Sci. U.S.A.*, **98**: 7534-7539.
- Shevchenko, A., Wilm, M., Vorm, O., & Mann, M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* **68**: 850-858.
- Shi, L., Jung, Y. J., Tyagi, S., Gennaro, M. L. & North, R. J. (2003) Expression of Th1-mediated immunity in mouse lungs induces a *Mycobacterium tuberculosis* transcription pattern characteristic of nonreplicating persistence. *Proc. Natl. Acad. Sci. U.S.A.*, **100**: 241-246.

The Universal Stress Proteins of Bacteria

- Shleeve, M. O., Bagramyan, K., Telkov, M. V., Mukamolova, G. V., Young, M., Kell, D. B. & Kaprelyants, A. S. (2002) Formation and resuscitation of “nonculturable” cells of *Rhodococcus rhodochrous* and *Mycobacterium tuberculosis* in prolonged stationary phase. *Microbiology*, **148**: 1581–1591.
- Simon, R., Priefer, U. & Puhler, A. (1983) A broad host range mobilization system for in vivo genetic engineering – Transposon mutagenesis in Gram-negative bacteria. *Bio. Technology*, **1**: 784–791.
- Singh, A., Mai, D., Kumar, A., & Steyn, A. J. (2006) Dissecting virulence pathways of *Mycobacterium tuberculosis* through protein–protein association. *Proc. Natl. Acad. Sci. U.S.A.*, **103**: 11346–11351.
- Singh, P. K., Schaefer, A. L. Parsek, M. R., Moninger, T. O., Welsh, M. J. & Greenberg, E. P. (2000) Quorum-sensing signals indicate that cystic fibrosis lungs are infected with bacterial biofilms. *Nature*, **407**: 762–764.
- Small, P., Blankenhorn, D., Welty, D., Zinser, E. & Slonczewski, J.L. (1994) Acid and base resistance in *Escherichia coli* and *Shigella flexneri*: role of rpoS and growth pH. *J. Bacteriol.* **176**: 1729–1737.
- Smeulders, M. J., Keer, J., Speight, R. A. & Williams, H. D. (1999) Adaptation of *Mycobacterium smegmatis* to stationary phase. *J. Bacteriol.* **181**: 270–283.
- Smith, L. J., Stapleton, M. R., Fullstone, G. J., Crack, J. C., Thomson, A. J., Le Brun, N. E., Hunt, D. M., Harvey, E., Adinolfi, S., Buxton, R. S. & Green, J. (2010) *Mycobacterium tuberculosis* WhiB1 is an essential DNA-binding protein with a nitric oxide-sensitive iron-sulfur cluster. *Biochem J.* **432**: 417–27.
- Sousa, M.C. & McKay, D. B. (2001) Structure of the universal stress protein of *Haemophilus influenzae*. *Structure (Cambridge)*, **9**: 1135–1141.
- Stapleton, M., Haq, I., Hunt, D. M., Arnvig, K. B., Artymiuk, P. J., Buxton, R. S. & Green, J. (2010) *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. Biolo. Chem.*, **285**: 7016–27.
- Steinberg, R. A., Cauthron, R. D., Symcox, M. M. & Shuntoh, H. (1993) Autoactivation of catalytic (C alpha) subunit of cyclic AMP-dependent protein kinase by phosphorylation of threonine 197. *Mol. Cell Biol* **13**: 2332–2341.
- Steiner-Lange, S., Fischer, A., Boettcher, A., Rouhara, I., Liedgens, H., Schmelzer, E. & Knogge, W. (2003) Differential defense reactions in leaf tissues of barley in response to infection by *Rhynchosporium secalis* and to treatment with a fungal avirulence gene product. *Mol. Plant Micro. Interact.*, **16**: 893–902.
- Stuehr, D. J. & Nathan, C. F. (1989) Nitric oxide. A macrophage product responsible for cytostasis and respiratory inhibition in tumor target cells. *J. Exp. Med.*, **169**: 1543–1555.

The Universal Stress Proteins of Bacteria

- Sturgill-Koszycki, S., Schlesinger, P. H., Chakraborty, P., Haddix, P. L., Collins, H. L., Fok, A. K., Allen, R. D., Gluck, S. L., Heuser, J. & Russell, D. G. (1994) Lack of acidification in *Mycobacterium* phagosomes produced by exclusion of the vesicular proton-ATPase. *Science*. **263**: 678-681.
- Talaat, A. M., Reimschuessel, R., Wasserman, S. S. & Trucksis, M. (1998) Goldfish, *Carassius auratus*, a novel animal model for the study of *Mycobacterium marinum* pathogenesis. *Infect. Immun.*, **66**: 2938-42.
- Tameling, W. I., Elzinga, S. D., Darmin, P. S., Vossen, J. H., Takken, F. L., Haring, M. A. & Cornelissen, B. J. (2002) The tomato R gene products I-2 and MI-1 are functional ATP binding proteins with ATPase activity. *Plant Cell*, **14**: 2929-2939.
- Traag, B. A., Driks, A., Stragier, P., Bitter, W., Broussard, G., Hatfull, G., Chu, F., Adams, K. N., Ramakrishnan, L. & Losick, R. (2010) Do mycobacteria produce endospores? *Proc. Natl. Acad. Sci. U.S.A.*, **107**: 878–881.
- Tsai, M.C., Chakraavarty, S., Zhu, G., Xu, J., Tanaka, K., Koch, C., Tufariello, J., Flynn, J. & Chan, J. (2006) Characterization of the tuberculous granuloma in murine and human lungs: cellular composition and relative tissue oxygen tension. *Cell Microbiol.* **8**: 218–232.
- Ulrichs, T., Kosmiadi, G.A., Trusov, V., Jorg, S., Pradl, L., Titukhina, M., Mishenko, V., Gushina, N. & Kaufmann, S. H. (2004) Human tuberculous granulomas induce peripheral lymphoid follicle-like structures to orchestrate local host defence in the lung. *J. Pathol.* **204**: 217–228.
- Usuki, K., Miyazono, K. & Heldin, C. H. (1991) Covalent Linkage between Nucleotides and Platelet-derived Endothelial Cell Growth Factor. *J. Biol. Chem.*, **266**: 20525-20531.
- van der Wauven C., Piérard, A., Kley-Raymann M. & Haas, D. (1984) *Pseudomonas aeruginosa* mutants affected in anaerobic growth on arginine: evidence for a four-gene cluster encoding the arginine deiminase pathway. *J. Bacteriol.*, **160**: 928-934.
- van der Wel, N., Hava, D., Houben, D., Fluitsma, D., van Zon, M., Pierson, J., Brenner, M. & Peters, P. J. (2007) *M. tuberculosis* and *M. leprae* translocate from the phagolysosome to the cytosol in myeloid cells. *Cell*, **129**:1287–1298.
- van der Wolk, J. P., Boorsma, A., Knoche, M., Schafer, H. J., & Driessen, A. J. (1997) The low-affinity ATP binding site of the *Escherichia coli* SecA dimer is localized at the subunit interface. *Biochemistry*, **36**: 14924-14929.
- van Kessel, J. C. & Hatfull, G. F. (2007) Recombineering in *Mycobacterium tuberculosis*. *Nat. Methods.*, **4**: 147-152.
- Vandal, O., Pierini, M., Schnappinger, D., Nathan, C. & Ehrt, S. (2008) A membrane protein preserves intrabacterial pH in intraphagosomal *Mycobacterium tuberculosis*. *Nat. Med.*, **14**: 849–854.

The Universal Stress Proteins of Bacteria

- Velayati, A., Farnia, P., Masjedi, M., Ibrahim, T., Tabarsi, P., Haroun, R., Kuan, H., Ghanavi, J., & Varahram, M. (2009) Totally drug-resistant tuberculosis strains: evidence of adaptation at the cellular level. *European Respiratory Journal*, **34**: 1202-1203.
- Via, L. E., Lin, P. L., Ray, S. M., Carrillo, J., Allen, S. S., Eum, S. Y., Taylor, K., Klein, E., Manjunatha, U., Gonzales, J., Lee, E. G., Park, S. K., Raleigh, J. A., Cho, S. N., McMurray, D. N., Flynn, J. L., & Barry, C. E., 3rd. (2008) Tuberculous granulomas are hypoxic in guinea pigs, rabbits, and nonhuman primates. *Infect. Immun.*, **76**: 2333–2340.
- Voskuil, M. I., Schnappinger, D., Harrell, M. I., Visconti, K. C., Dolganov, G., Sherman, D. R. & Schoolnik, G.K. (2003) Inhibition of respiration by nitric oxide induces a *Mycobacterium tuberculosis* dormancy program. *J. Exp. Med.*, **198**: 705–713.
- Voskuil, M. I., Visconti, K. C. & Schoolnik, G. K. (2004) *Mycobacterium tuberculosis* gene expression during adaptation to stationary phase and low oxygen dormancy. *Tuberculosis*, **84**: 218-227.
- Voss, D., Rutter, J. K., Schroeder, B. G., Su, H., Zhu, Y. & Barry, C. E. (2000) The salicylate derived mycobactin siderophores of *Mycobacterium tuberculosis* are essential for growth in macrophages. *Proc. Natl. Acad. Sci. U.S.A.*, **97**: 1252–1257.
- Walburger, A., Koul, A., Ferrari, G., Nguyen, L., Prescianotto-Baschong, C., Huygen, K., Klebl, B., Thompson, C., Bacher, G. & Pieters, J. (2004) Protein Kinase G from pathogenic mycobacteria promotes survival within macrophages. *Science*, **304**: 1800-1804.
- Walderhaug, M. O., Polarek, J. W., Voelkner, P., Daniel, J. M., Hesse, J. E., Altendorf, K. & Epstein, W. (1992) KdpD and KdpE, proteins that control expression of the *kdpABC* operon, are members of the two-component sensor-effector class of regulators. *J. Bacter.*, **174**: 2152–2159.
- Walter, A. O. Seghezzi, W. Korver, W. Sheung, J. & Lees, E. (2000) The mitotic serine/threonine kinase Aurora2/AIK is regulated by phosphorylation and degradation *Oncogene*, **19**: 4906-4916.
- Wang, Y., Cui, T., Zhang, C., Yang, M., Huang, Y., Li, W., Zhang, L., Gao, C., He, Y., Li, Y., Huang, F., Zeng, J., Huang, C., Yang, Q., Tian, Y., Zhao, C., Chen, H., Zhang, H. & He, Z. G. (2010) Global protein-protein interaction network in the human pathogen *Mycobacterium tuberculosis* H37Rv. *J. Proteome Res.* **9**: 6665-6677.
- Wayne L. (1954) Growth of *Mycobacterium tuberculosis* from resected specimens under various atmospheric conditions. *An. Rev. Tuberc.*, **70**: 910–911.
- Wayne L. G., Kubica G. P. (1986) Genus *Mycobacterium*. Lehmann and Neumann 1896, 363AL. In: Sneath P. H. A., Mair N. S., Sharpe M. E., Holt J. G. (eds.): Bergey's Manual of Systematic Bacteriology, 2. The Williams & Wilkins Co., Baltimore. 1436-1457.

The Universal Stress Proteins of Bacteria

- Wayne, L. G. & Hayes, L. G. (1996) An *in vitro* model for sequential study of shutdown of *Mycobacterium tuberculosis* through two stages of non-replicating persistence. *Infect. Immun.*, **64**: 2062–2069.
- Wayne, L. G. (1994) Dormancy of *Mycobacterium tuberculosis* and latency of disease. *Eur. J. Clin. Microbiol. Infect. Dis.*, **13**: 908–914.
- Wayne, L. G. & Sohaskey, C. D. (2001) Non-replicating persistence of *Mycobacterium tuberculosis*. *Ann. Rev. Microbiol.*, **55**: 139-163.
- Weber, A. & Jung, K. (2006) Biochemical Properties of UspG, a Universal Stress Protein of *Escherichia coli*. *Biochemistry*, **45**: 1620-1628.
- Weber, I. Fritz, C. Ruttkowski, S. Kreft, A. & Bange, F. C. (2000) Anaerobic nitrate reductase (narGHJI) activity of *Mycobacterium bovis* BCG *in vitro* and its contribution to virulence in immunodeficient mice. *Mol Microbiol.*, **35**: 1017-25.
- WHO (2007) Revised BCG vaccination guidelines for infants at risk for HIV infection. *Wkly. Epidemiol. Rec.*, **83**: 19-196.
- WHO (2009) Report – Global Tuberculosis Control: Epidemiology, Strategy, Financing. www.who.int/tb
- WHO (2010) Report – Global Tuberculosis Control: A short Update to the 2009 Report. www.who.int/tb
- Wong, D. K., Lee, B. Y., Horwitz, M. A. & Gibson, B. W. (1999) Identification of fur, aconitase, and other proteins expressed by *Mycobacterium tuberculosis* under conditions of low and high concentrations of iron by combined two-dimensional gel electrophoresis and mass spectrometry. *Infect. Immun.*, **67**: 327-336.
- Worlitzsch, D., Tarran, R., Ulrich, M., Schwab, U., Cekici A., Meyer, K. C., Birrer, P., Bellon, G., Berger, J., Weiss, T., Botzenhart, K., Yankaskas, J. R., Randell, S., Boucher, R. C. & Doring, G. (2002) Effects of reduced mucus oxygen concentration in *Pseudomonas* infections of cystic fibrosis patients. *J. Clin. Invest.*, **109**: 317-325.
- Xiao, H., Kalman, M., Ikehara, K., Zemel, S., Glaser, G. & Cashel, M. (1991) Residual guanosine 30,50-bispyrophosphate synthetic activity of relA null mutants can be eliminated by spoT null mutations. *J. Biol. Chem.*, **266**: 5980-5990.
- Yang, S. & Frey, P. (1979) Nucleophile in the Active Site of *Escherichia coli* Galactose- 1 -phosphate Uridyltransferase: Degradation of the Uridyl-enzyme Intermediate to N3-Phosphohistidine. *Biochemistry*, **18**: 2980-2984.
- Yeager, R. L., Munroe, W. G. & Dassaifu, F.I. (1952) Pyrazinamide (aldinamide*) in the treatment of pulmonary tuberculosis. *Trans. Annu. Meet. Natl. Tuberc. Assoc.*, **65**: 523-46.

The Universal Stress Proteins of Bacteria

Yuan, Y., Lee, R. E., Besra, G. S., Belisle, J. T. & Barry, C. E. (1995) Identification of a gene involved in the biosynthesis of cyclopropanated mycolic acids in *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U.S.A.*, **92**: 6630–6634.

Zarembinski, T. I., Hung, L. W., Mueller-Dieckmann, H. J., Kim, K. K., Yokota, H., Kim, R. & Kim, S. H. (1998) Structure-based assignment of the biochemical function of a hypothetical protein: A test case of structural genomics. *Proc. Natl. Acad. Sci. U.S.A.*, **95**: 15189-15193.

Zegzouti, H., Jones, B., Marty, C., Lelièvre, J.-M., Latché, A., Pech, J.-C. & Bouzayen, M. (1997). ER5, a tomato cDNA encoding an ethylene-responsive LEA-like protein: characterization and expression in response to drought, ABA and wounding. *Plant. Mol. Biol.*, **35**: 847-54.

Zhou, B., He, Y., Zhang, X., Xu, J., Luo, Y., Wang, Y., Franzblau, S., Yang, Z., Chan, R., Liu, Y., Zheng, J. & Zhang, Z. (2010) Targeting mycobacterium protein tyrosine phosphatase B for antituberculosis agents. *Proc. Natl. Acad. Sci. U.S.A.*, **107**: 4573-4578.

Zink, A. R., Sola, C., Reischl, U., Grabner, W., Rastogi, N., Wolf, H., & Nerlich, A. G. (2003) Characterization of *Mycobacterium tuberculosis* complex DNAs from Egyptian mummies by spoligotyping. *J. Clin. Microbiol.*, **41**: 359-367.

Zumft, W. G. & Korner, H. (1997) Enzyme diversity and mosaic gene organization in denitrification. *A. Van. Leeuw.*, **71**: 43-58.