



The
University
Of
Sheffield.

Role of ECF sigma factors in stress responses of *Burkholderia cenocepacia*

By:
Sayali Haldipurkar

A thesis submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy.

The University of Sheffield
Faculty of Medicine, Dentistry and Health
Department of Infection, Immunity and Cardiovascular Disease
April 2017

Acknowledgements

Firstly, I would like to thank my supervisor Dr. Mark S Thomas for his guidance during this project and for always finding the time to be available in spite of his busy schedule. I am grateful to the university for granting me 'The University of Sheffield Prize Scholarship' without which I could not have carried out this research. I would also like to thank Dr. Paul Heath and Matthew D Wyles from SITraN, University of Sheffield, for their technical support with sequencing the next-generation libraries. I am thankful to everyone I've worked with in the LU103 lab, particularly, Aaron, Helena, Jamie, Saeedah, Ruyue, Sabella and Rebecca for always being so kind and helpful.

I have made some wonderful friends in last three and half years that have genuinely been there for me. Especially, I would like to thank Apoorva for her advice and support and for standing by me through thick and thin; Furaha for practically being my shrink and especially being there for me emotionally during the very challenging stipend-less writing up period; Lucy for always being there to listen and also for being a reassuring voice when the going got tough (plus for the innumerable movie nights!). I would like to thank Emily, Chloe, Jess, Billy and Renata amongst so many other research students I met through the doctoral development activities who have been an important part of the support system I found in Sheffield. Not forgetting the gym/zumba buddies, flatmates and lovely people I met here who have made my time in the UK special - with camps, picnics and London trips, giving me the much needed break now and then.

Lastly and most importantly, I would like to thank my family – Mum, Dad and sister, Ketaki for constantly supporting and loving me beyond measure. I owe most of where I stand in life today to them. Unexpectedly and unfortunately, I lost my Dad, Mr. Shantaram Haldipurkar, during the course of this PhD. So, in my own humble way I would like to dedicate this thesis to him and would specially thank him for always believing in me. He continues to live with us in our memories through his teachings and principles... *Look how far I've come, daddy, - I did it!*

Abstract

Transcription initiation in bacteria requires the dissociable sub-unit σ to bind to core RNA polymerase (RNAP) for recognition of key promoter elements. Bacteria harbour an assortment of different types and numbers of σ factors that they can selectively utilize depending on its adaptive requirement. One such group of alternative σ factors that are used for transcribing genes in response to stressors from outside the cytoplasm are called extracytoplasmic function (ECF) σ factors that are often regulated by anti- σ factors. However, alternative anti- σ devoid models also exist. The opportunistic pathogen, *Burkholderia cenocepacia*, has 13 ECF σ factors, of which 11 are currently uncharacterized.

In this study, the role of four previously uncharacterized ECF σ factors, PrtI, Ecf41_{Bc1}, Ecf41_{Bc2} and Ecf42_{Bc} of *B. cenocepacia* were investigated. Using ChIP-seq and reporter-fusion assays PrtI was found to regulate the *prtI* promoter itself and the promoter of a putative catalase. The -10 and -35 promoter elements for P_{prtI} were experimentally shown to be 'GTC' and 'GGAATA', respectively, separated by a 16 bp spacer. A predicted anti- σ factor encoded by the gene downstream of *prtI* was shown to be inhibitory to PrtI activity and to interact with the C-terminal domain of PrtI through its N-terminal domain.

Ecf41_{Bc1} and Ecf41_{Bc2} were shown to regulate promoters of putative alkyl-hydroperoxide D and LysR-type transcriptional regulator, respectively. Ecf41_{Bc1} and Ecf41_{Bc2} lack cognate anti- σ factors, but instead, have a C-terminal extension which was hypothesized to be regulatory. However, based on experiments from C-terminally truncated derivatives of Ecf41_{Bc1} and Ecf41_{Bc2}, the CTE exerting such an effect cannot be concluded. No phenotypes were obtained for deletion mutants of any of the above σ factors.

In case of the iron-starvation ECF σ factor OrbS, in the OrbS-dependent promoter, 'TAAA' and 'CGTC' were demonstrated to form the -35 and -10 promoter elements, respectively, separated by 17 bp. The length of the spacer was found to be critical. The bases located near the transcription start site were important for efficient promoter utilization when substituted in a block of three. OrbS is regulated transcriptionally by an iron regulator Fur and does not have an anti- σ factor. Presence of an alternative regulatory mechanism was investigated using a *fur* mutant. However, it was not possible to draw conclusions from these studies as the *fur* mutant used, unexpectedly, retained Fur activity.

Table of Contents

Acknowledgements	2
Abstract	3
List of Tables	11
List of Figures.....	12
List of abbreviations	16
Chapter 1. Introduction.....	19
1.1 General introduction	20
1.2 σ factor: structure.....	20
1.3 σ factor: function in transcription initiation.....	22
1.3.1 Structural interaction of σ factor with RNAP in transcription initiation.....	22
1.3.2 Kinetics of the σ factor-RNAP interaction in transcription initiation	25
1.3.3 σ release	26
1.4 Classification of σ factors.....	27
1.5 Extra-cytoplasmic function σ factors.....	28
1.5.1 Structure	28
1.5.2 Mechanism	29
1.5.3 Classification	32
1.5.4 Regulatory mechanisms	33
1.5.5 Role in stress response	37
1.6 The <i>Burkholderia cepacia</i> complex (<i>Bcc</i>)	48
1.6.1 Interaction with the environment/host	49
1.6.2 Virulence determinants	51
1.6.3 ECF σ factors of <i>B. cenocepacia</i>	54
1.7 Hypotheses and aims.....	65
Chapter 2: Materials and methods.....	67
2.1 Bacteriological techniques	68
2.1.1 Bacterial strains	68

2.1.2 Bacteriological media and supplements	69
2.1.3 Bacterial storage glycerol stocks.....	74
2.1.4 Determination of bacterial growth rate	74
2.1.5 Efficiency of plating (EOP).....	75
2.2 Recombinant DNA techniques	76
2.2.1 Plasmids	76
2.2.2 Primers and oligonucleotides:	81
2.2.3 Plasmid extraction using a nucleic acid isolation kit.....	87
2.2.4 DNA electrophoresis	87
2.2.5 Restriction digestion	88
2.2.6 Ligation.....	88
2.2.7 Phosphatase treatment	88
2.2.8 Annealing of complementary oligonucleotides	89
2.2.9 DNA purification.....	89
2.2.10 Amplification of DNA by the polymerase chain reaction.....	90
2.2.11 Techniques for plasmid transfer	94
2.3 Recombinant protein overproduction and purification techniques:.....	96
2.3.1 Protein induction protocol (small scale)	96
2.3.2 Determination of overproduced protein solubility.....	96
2.3.3 Purification of His-tagged proteins by nickel affinity chromatography column using FPLC	97
2.3.4 Protein storage.....	98
2.4 Protein separation and quantification techniques:	99
2.4.1 SDS PAGE.....	99
2.4.2 Western blotting	100
2.4.3 Protein quantification assay	100
2.5 Enzyme assays.....	101
2.5.1 β -galactosidase assay.....	101
2.5.2 Catalase assay	103
2.6 Interaction studies:	103
2.6.1 BACTH assay (protein-protein interaction).....	103
2.6.2 Chromatin immunoprecipitation and deep sequencing (DNA-protein interaction)	107

2.7 Phenotypic studies:	108
2.7.1 CAS assay	108
2.7.2 Motility and swarming assays	108
2.7.3 Protease assay	108
2.7.4 Biofilm formation.....	109
2.8 Peroxide sensitivity stress tests.....	109
2.8.1 Peroxide sensitivity test in liquid media (survival following 20 minute exposure)	109
2.8.2 Peroxide sensitivity test in liquid media (monitoring survival time).....	110
2.8.3 Peroxide sensitivity test on solid culture medium (CFU assay).....	110
2.8.4 Zone inhibition assay to test peroxide sensitivity	110
 Chapter 3. Investigation of the role of the ECF σ factor I35_RS16290 (PrtI).....	 111
3.1 Introduction.....	112
3.2 Investigation of the role of PrtI by identification of PrtI dependent promoters using ChIP-seq and experimental verification of identified promoters.....	116
3.2.1 Construction of plasmid to express FLAG-tagged PrtI.....	116
3.2.2 Validation of antibody specificity	118
3.2.3 Validation of ChIP-seq experimental design by reporter fusion assay: Analysis of putative PrtI regulated promoter activity under induced and non-induced conditions	120
3.2.4 Construction of pKAGd4-P _{prtI}	120
3.2.5 Analysis of the activity of PrtI-FLAG at the P _{prtI} promoter.....	120
3.2.6 Optimisation of conditions for chromatin immunoprecipitation.....	123
3.2.7 PCR validation of ChIP efficiency	128
3.2.8 Quality control of DNA samples library preparation	132
3.2.9 Data analysis of ChIP-seq output: Quality assessment of sequencing output	135
3.2.10 Mapping DNA sequence reads using Bowtie to the <i>Burkholderia cenocepacia</i> genome to reveal PrtI binding sites.....	146
3.2.11 Peak calling and motif identification	151
3.2.12 <i>In silico</i> identification of putative PrtI regulated promoters at the <i>prtI</i> locus.....	151
3.2.13 Construction of reporter vector pKAGd4 harbouring putative PrtI regulated promoters	153
3.2.14 Analysis of putative PrtI regulated promoter activity in <i>E. coli</i> and <i>B. cenocepacia</i>	153
3.2.15 Construction of plasmids expressing native PrtI	157

3.2.17 Analysis of the activity of P _{prtI} ^{full} , P _{prtI} , and P _{cat} in the presence and absence of native PrtI in <i>E.coli</i>	157
3.3 Systematic study of the DNA sequence requirements for promoter recognition by PrtI	159
3.3.1 Construction of plasmids harbouring single base substituted P _{prtI} promoter derivatives	159
3.3.2 Analysis of DNA sequence requirements for promoter utilisation by PrtI	159
3.4 FUZZNUC based promoter search in other bacteria.....	162
3.5 Genetic analysis of the role of PrtI.....	162
3.5.1 Construction of a <i>prtI</i> in-frame deletion mutant	163
3.5.2 Construction of a <i>prtR</i> in-frame deletion mutant.....	164
3.5.3 Construction of a plasmid for complementing the Δ <i>prtR</i> mutant.....	168
3.5.4 Effect of deletion of <i>prtI</i> or <i>prtR</i> on bacterial growth under normal conditions	168
3.5.5 Effect of deletion of <i>prtI</i> or <i>prtR</i> on bacterial growth under oxidative stress conditions	170
3.5.6 Effect of deletion of <i>prtI</i> or <i>prtR</i> on protease activity.....	175
3.5.7 Effect of deletion of <i>prtI</i> or <i>prtR</i> on bacterial motility and swarming	177
3.5.8 Effect of deletion of <i>prtI</i> or <i>prtR</i> on biofilm formation	177
3.5.9 Effect of deletion of <i>prtI</i> or <i>prtR</i> on EPS production.....	179
3.6 Investigation of the role of PrtI by studying the role of its target gene – I35_RS16285	182
3.6.1 Testing the enzyme activity of purified I35_RS16285.....	182
3.6.2 Identifying the function of I35_RS16285 by testing for complementation of an <i>E. coli</i> catalase deficient mutant	189
3.6.3 Bioinformatic analysis of I35_RS16285.....	191
3.7 Regulation of PrtI by the putative anti- σ factor PrtR.....	194
3.7.1 Construction of plasmids expressing PrtI and PrtR hybrid proteins for two-hybrid analysis.....	194
3.7.2 Analysis of PrtI-PrtR interaction and self-interaction using the bacterial two-hybrid assay	195
3.7.3 Effect of PrtR on the activity of PrtI in <i>E. coli</i>	198
3.7.4 PrtR on the activity of PrtI in <i>B. cenocepacia</i>	200
3.8 Defining the extent of the I35_RS16285 and <i>prtI</i> operons.....	201
3.9 Discussion.....	204
3.9.1 PrtI recognises at least two promoters.....	204

3.9.2 Deletion of <i>prtI</i> or <i>prtR</i> did not affect stress responses in <i>B. cenocepacia</i>	207
3.9.3 Does PrtI regulate the production of a catalase?	208
3.9.4 PrtR, an inhibitory σ factor regulatory protein.....	210
3.9.5 The PrtI operon.....	211
Chapter 4. Investigation of the roles of ECF σ factors I35_RS16330 (Ecf41 _{Bc1}), I35_RS29600 (Ecf41 _{Bc2}) and I35_RS20320 (Ecf42 _{Bc})	215
4.1 Introduction.....	216
4.2 Genomic context of <i>ecf41</i> _{Bc1} and <i>ecf41</i> _{Bc2}	217
4.3 Investigation of the role of Ecf41 _{Bc1} and Ecf41 _{Bc2} by of identification target promoters.....	222
4.3.1 Identification of Ecf41 _{Bc1} and Ecf41 _{Bc2} -regulated promoters <i>in silico</i>	222
4.3.2 Construction of plasmids harbouring either of the σ factors Ecf41 _{Bc1} /Ecf41 _{Bc2} or putative promoter derivatives P _{ahpD} /P _{lysR}	228
4.3.3 Analysis of putative Ecf41 _{Bc1} and Ecf41 _{Bc2} -dependent promoter activity in the presence of the respective ECF σ factors in <i>E. coli</i>	229
4.3.4 Attempted construction of plasmids expressing FLAG-tagged Ecf41 _{Bc} derivatives for CHIP-seq.....	231
4.3.5 FUZZNUC based identification of additional ECF41-dependent promoters	231
4.4 Genetic analysis of the role of ECF41 group σ factors in <i>B. cenocepacia</i>	233
4.4.1 Construction of an <i>ecf41</i> _{Bc1}	233
4.4.2 Construction of an <i>ecf41</i> _{Bc2} mutant.....	234
4.4.3 Construction of an <i>ecf41</i> _{Bc1} and <i>ecf41</i> _{Bc2} in-frame deletion mutant	235
4.4.4 Effect of deletion of <i>ecf41</i> genes on growth of <i>B. cenocepacia</i> under normal conditions	238
4.4.5 Effect of deletion of <i>ecf41</i> genes on bacterial growth under oxidative stress conditions	240
4.5 Regulatory mechanism of Ecf41 _{Bc1} and Ecf41 _{Bc2}	242
4.5.1 Construction of plasmids encoding truncated derivatives of Ecf41 _{Bc1} and Ecf41 _{Bc2}	242
4.5.2 Analysis of putative Ecf41 _{Bc1} and Ecf41 _{Bc2} regulated promoter activity in the presence of the respective C-terminally truncated ECF σ factors in <i>E. coli</i>	243
4.5.3 Analysis of putative Ecf41 _{Bc1} and Ecf41 _{Bc2} regulated promoter activity in the presence of the respective C-terminally truncated ECF σ factors in <i>B. cenocepacia</i>	245
4.6 Genomic context of ECF42 _{Bc}	247
4.7 Investigation of the role of ECF42 by identification of target promoters	251

4.7.1 <i>In silico</i> identification of ECF42 regulated promoters.....	251
4.7.2 Construction of plasmids harbouring ECF42 _{Bc} and putative ECF42 _{Bc} -dependent promoter	253
4.7.3 Analysis of putative ECF42-dependent promoter activity in the presence of ECF42 in <i>E. coli</i> and <i>B. cenocepacia</i>	254
4.7.4 FUZZNUC based identification of additional putative ECF42-dependent promoters....	254
4.7.4 <i>In silico</i> identification of putative ECF42-dependent promoters in the genome of other bacterial species.....	258
4.8 Genetic analysis of the role of Ecf42 _{Bc}	258
4.8.1 Construction of an <i>ecf42_{Bc}</i> in-frame deletion mutant	259
4.8.2 Effect of <i>ecf42_{Bc}</i> deletion on growth of <i>B. cenocepacia</i> in standard lab conditions.....	260
4.8.3 Effect of <i>ecf42_{Bc}</i> deletion on growth of <i>B. cenocepacia</i> under oxidative stress	262
4.9 Investigation of the role of the ECF42 _{Bc} C-terminal domain in the regulation of Ecf42 _{Bc}	263
4.9.1 Construction of a plasmid expressing C-terminally truncated Ecf42 _{Bc}	263
4.9.2 Analysis of putative Ecf42 _{Bc} -dependent promoter activity in the presence of the C-terminally truncated Ecf42 _{Bc} in <i>E. coli</i> and <i>B. cenocepacia</i>	263
4.10 Discussion.....	264
4.10.1 Ecf41 _{Bc1} and Ecf41 _{Bc2} -dependent promoters	264
4.10.2 Role of Ecf41 _{Bc1} and Ecf41 _{Bc2}	265
4.10.3 Regulation of Ecf41Bc activity.....	266
4.10.4 Role and regulation of Ecf42 _{Bc}	267
Chapter 5. Investigation of the iron starvation ECF σ factor, OrbS	269
5.1 Introduction	271
5.2 Investigation of DNA sequence requirements for OrbS activity.....	273
5.2.1 Activity of OrbS dependent promoters under inducing conditions in <i>B. cenocepacia</i> ..	274
5.2.2 Analysis of the P _{orbH} promoter by studying the effect of single base substitutions on OrbS-dependent activity in <i>B. cenocepacia</i>	276
5.2.3 Analysis of GC rich spacer region in <i>B. cenocepacia</i>	280
5.2.4 Analysis of the A+G rich transcription start site region in <i>B. cenocepacia</i>	282
5.2.5 Re-construction of one P _{orbH} variant promoter and its analysis in <i>E. coli</i>	284
5.2.6 Testing the functionality of <i>in silico</i> determined putative OrbS-dependent promoters P _{fpr} and P _{ureA}	286

5.3 Studying the regulatory mechanisms for the activity of OrbS using a <i>B. cenocepacia fur</i> mutant	293
5.3.1 Phenotypic characterisation of a <i>B. cenocepacia fur</i> mutant	295
5.3.2 Complementation of the <i>fur</i> mutant in the investigation of Fur regulated promoters	304
5.3.3 Iron dependent regulation of OrbS activity in the absence of Fur.....	306
5.3.4 Analysis of the activity of iron-independent promoters in response to iron in <i>B. cenocepacia</i>	308
5.4 Investigation of iron-dependent regulation of OrbS activity using an <i>E. coli fur</i> mutant	310
5.4.1 Activity of Fur regulated promoter in an <i>E. coli fur</i> mutant	310
5.4.2 Activity of an OrbS-regulated promoter in an <i>E. coli fur</i> mutant	312
5.4.3 Construction of an iron-independent promoter and its activity in <i>E. coli</i>	314
5.5 Identification of Fur binding regions on the <i>B. cenocepacia</i> genome.....	316
5.6 Discussion	318
5.6.1 DNA sequence requirements for effective promoter utilisation by OrbS.....	318
5.6.2 OrbS may have any additional promoter targets	321
5.6.3 Regulation of OrbS activity	321
Chapter 6. Final discussion	326
6.1 General discussion.....	327
6.2 Future perspectives	329
References	333
Appendices	353
Appendix I:.....	354
Appendix II:.....	355

List of Tables

Chapter 2:

Table 2.1. Bacterial strains.....	68
Table 2.2. Concentration of antibiotics used in selective media.....	73
Table 2.3. Concentration of supplements used in selective media.....	74
Table 2.4. Plasmids.....	76
Table 2.5. Primers and oligonucleotides used in this study.....	80
Table 2.6. Components added in PCR amplification.....	90
Table 2.7. Components added to colony screening PCR.....	93
Table 2.8. SDS-Polyacrylamide gel ingredients.....	99
Table 2.9. Combinations of plasmids used in a BACTH study of the interaction between two proteins A and B.....	105

Chapter 3:

Table 3.1. Optimisation of medium for EPS production by <i>B. cenocepacia</i>	180
---	-----

Chapter 4:

Table 4.1. Putative Ecf42-dependent promoters identified in the <i>B. cenocepacia</i> H111 genome.....	255
--	-----

Chapter 5:

Table 5.1. Colony size comparison of 715j (WT) and 715j- <i>fur</i> ::Tp.....	300
---	-----

List of Figures

Chapter 1:

Figure 1.1. Structure of σ^{70} from <i>E. coli</i> when interacting with RNAP.....	22
Figure 1.2 Structure of RNAP holoenzyme with the corresponding promoter recognition regions.....	25
Figure 1.3. Transcription initiation process.....	27
Figure 1.4 Domain organisation of ECF σ factors compared to σ^{70} (not drawn to scale).....	29
Figure 1.5 Promoter melting at the -10 promoter element by σ^{70} and an ECF σ	32
Figure 1.6. Promoters that drive transcription of genes encoding ECF σ factors.....	32
Figure 1.7 Schematic representation of some mechanisms for regulation of ECF σ factors.....	36
Figure 1.8. Summary of the σ^E system in <i>E. coli</i>	40
Figure 1.9. Summary of the Pvd and Fec iron acquisition systems.....	47
Figure 1.10 Taxonomic classification of the <i>Bcc</i>	49
Figure 1.11. Predicted PrtI-dependent promoter P_{aprX}	58
Figure 1.12. The Orb system in <i>B. cenocepacia</i>	64

Chapter 2:

Figure 2.1 Principle of SOE PCR deletion mutagenesis.....	92
Figure 2.2 Principle of BACTH assay.....	106

Chapter 3:

Figure 3.1. Genes present at the <i>prtI</i> locus in a variety of Gram-negative bacteria.....	114
Figure 3.2. Amino acid sequence alignment of PrtI from <i>B. cenocepacia</i> with the corresponding protein in other Gram-negative species.....	115
Figure 3.3. Construction of pSRK-Km-PrtI-FLAG.....	117
Figure 3.4. Production of PrtI-FLAG in cells containing pSRK-Km-PrtI-FLAG.....	119
Figure 3.5 Construction of pKAGd4- P_{prtI} full.....	121
Figure 3.6. Activity of the <i>prtI</i> promoter in the presence of PrtI-FLAG.....	122
Figure 3.7. ChIP-seq work flow for identification of PrtI binding sites.....	126
Figure 3.8. Optimisation of cell lysis and shearing genomic DNA by sonication.....	127
Figure 3.9. Validation of IP and CONTROL samples using PCR.....	130
Figure 3.10. Validation of IP and CONTROL samples using PCR with additional positive control.....	131
Figure 3.11. Quality assessment of samples before library preparation.....	133
Figure 3.12. Quality assessment of samples after library preparation.....	134
Figure 3.13. Quality metrics of reads from sequencer using SAV.....	136
Figure 3.14. Quality metrics of reads mapped to index using SAV.....	137
Figure 3.15. Quality metrics of sequenced data using GALAXY (1).....	140
Figure 3.16. Quality metrics of sequenced data using GALAXY (2).....	141

Figure 3.17. Quality metrics of sequenced data using GALAXY (3).....	142
Figure 3.18. Quality metrics of sequenced data using GALAXY (4).....	143
Figure 3.19. Quality metrics of sequenced data using GALAXY (5).....	144
Figure 3.20. Quality metrics of sequenced data using GALAXY (6).....	145
Figure 3.21. Mapping DNA sequence reads to the <i>B. cenocepacia</i> genome using IGV(1).	148
Figure 3.22. Mapping DNA sequence reads to the <i>B. cenocepacia</i> genome using IGV(2).	149
Figure 3.23. Mapping DNA sequence reads to the <i>B. cenocepacia</i> genome using IGV(3).	150
Figure 3.24. Clustal Omega alignment for identification of putative PrtI recognition sequences.....	152
Figure 3.25. Activity of P _{prtI} and P _{cat} in the presence and absence of PrtI-FLAG in <i>E. coli</i> ...	155
Figure 3.26 Activity of P _{prtI2} and P _{cat} in the presence and absence of PrtI-FLAG in <i>B. cenocepacia</i>	156
Figure 3.27. Activity of P _{prtIfull} , P _{prtI} and P _{cat} in the presence and absence of native PrtI in <i>E. coli</i>	158
Figure 3.28. Identification of the -35 and -10 promoter elements of a PrtI-dependent promoter.....	161
Figure 3.29. Schematic representation of mutant construction using pEX18TpTer-Scel- <i>pheS</i>	166
Figure 3.30. PCR screening of candidate H111Δ <i>prtI</i> and H111Δ <i>prtR</i> mutants.....	167
Figure 3.31. Effect of deletion of <i>prtI</i> and <i>prtR</i> on growth under normal conditions.....	169
Figure 3.32. Effect of peroxide stress on the <i>B. cenocepacia</i> Δ <i>prtI</i> mutant.....	172
Figure 3.33. Effect of oxidative stress due to deletion of <i>prtI</i>	173
Figure 3.34. Effect of oxidative stress due to deletion of <i>prtR</i>	174
Figure 3.35. Effect of deletion of <i>prtI</i> and <i>prtR</i> on extracellular protease production by <i>B. cenocepacia</i>	176
Figure 3.36. Effect of deletion of <i>prtI</i> or <i>prtR</i> on motility, swarming and biofilm formation.....	178
Figure 3.37. Effect of inactivation of <i>prtI</i> and <i>prtR</i> on EPS production by <i>B. cenocepacia</i> on EPS medium (pH 6.0) at 30°C.....	181
Figure 3.38. Purification of His ₆ . I35_RS16285.....	185
Figure 3.39. Testing the catalase activity of I35_RS16285.....	188
Figure 3.40 Complementation assay on an <i>E. coli</i> catalase and NADH-dependent peroxidase.....	190
Figure 3.41. Alignment of KatE, I35_RS16285 and SrpA.....	193
Figure 3.42. Two hybrid analysis of PrtI-PrtR interaction.....	197
Figure 3.43. Activity of a PrtI-dependent promoter in the absence and presence of PrtR in <i>E. coli</i>	199
Figure 3.44. Activity of a PrtI-dependent promoter in the absence and presence of PrtR in <i>B. cenocepacia</i>	200
Figure 3.45. Identification of the I35_RS16285 (<i>catalase</i>) and <i>prtI</i> operons by reverse transcriptase PCR.....	203
Figure 3.46. Current understanding of the PrtI-PrtR system in <i>B. cenocepacia</i>	213

Chapter 4:

Figure 4.1 Genomic context of ECF41 σ factors.....	218
Figure 4.2. Amino acid sequence alignment of Ecf41 _{Bc1} and Ecf41 _{Bc2} from <i>B. cenocepacia</i> H111 with related proteins in other species.....	220
Figure 4.3. DNA sequence alignment for identification of putative Ecf4 _{Bc1} recognition sequences.....	223
Figure 4.4. Clustal Omega alignment for identification of putative Ecf41 _{Bc2} recognition sequences.....	225
Figure 4.5. Clustal Omega alignment for identification of putative <i>ecf41_{Bc2}</i> promoter.....	227
Figure 4.6. Activity of putative Ecf41-dependent promoters in the presence and absence of the ECF41 σ factors in <i>E. coli</i>	230
Figure 4.7. PCR screening of candidate H111 Δ <i>ecf41_{Bc1}</i> , H111 Δ <i>ecf41_{Bc2}</i> and H111 Δ <i>ecf41_{Bc1}</i> - Δ <i>ecf41_{Bc2}</i> mutants.....	237
Figure 4.8. Effect of deletion of <i>ecf41_{Bc1}</i> and/or <i>ecf41_{Bc2}</i> on growth under normal conditions.....	239
Figure 4.9. Representative images showing effect of oxidative stress due to deletion of <i>ecf41_{Bc1}</i> or <i>ecf41_{Bc2}</i> or both <i>ecf41_{Bc1}</i> and <i>ecf41_{Bc2}</i>	241
Figure 4.10. ECF41-dependent promoter activity in the presence and absence of full-length and truncated ECF41 σ factor derivatives in <i>E. coli</i>	244
Figure 4.11. ECF41-dependent promoter activity in the presence and absence of full-length and truncated derivatives of ECF41 σ factors in <i>B. cenocepacia</i>	246
Figure 4.12. Genomic context of <i>ecf42</i> σ factor genes.....	248
Figure 4.13. Amino acid sequence alignment of Ecf42 _{Bc} from <i>B. cenocepacia</i> H111 with related proteins in other species.....	249
Figure 4.14. Clustal Omega alignment for identification of putative Ecf42 recognition sequences.....	252
Figure 4.15. Clustal Omega alignment for identification of putative P _{phnB} promoter.....	257
Figure 4.16. PCR screening of a candidate H111 Δ <i>ecf42_{Bc}</i> mutant.....	260
Figure 4.17. Effect of deletion of <i>ecf42_{Bc}</i> on growth of <i>B. cenocepacia</i> under normal conditions.....	261
Figure 4.18. Effect of deletion of <i>ecf42_{Bc}</i> on the oxidative stress response of <i>B. cenocepacia</i>	262
Figure 4.19 Some proposed models of ECF41 regulation by the C-terminal extension (CTE).....	267

Chapter 5:

Figure 5.1. Promoter activities of OrbS dependent promoters in <i>B. cenocepacia</i>	275
Figure 5.2. P _{OrbHds6} substitution derivatives.....	278
Figure 5.3. Effect of single base substitutions on P _{OrbH} activity.....	279
Figure 5.4. Effect of alteration of the GC-rich spacer region of P _{OrbH}	281
Figure 5.5. Effect of alterations to the A+G-rich region at the TSS on P _{OrbH} activity.....	283

Figure 5.6. Promoter activity of selected P _{orbH} variant promoters in <i>E. coli</i>	285
Figure 5.7. Promoter activity of putative promoter P _{fpr} in <i>E. coli</i>	288
Figure 5.8. Promoter activity of putative promoters P _{fpr} and P _{ureA} in <i>B. cenocepacia</i>	290
Figure 5.9. Cloning FLAG tagged <i>orbS</i> into pSRK-Km.....	292
Figure 5.10. Experimental design to test for the presence of an alternate mechanism for regulating OrbS activity other than transcriptional regulation by Fur.....	294
Figure 5.11 CAS plate analysis of siderophore production by <i>B. cenocepacia</i> WT and <i>fur</i> ::Tp mutant.....	297
Figure 5.12. Growth rate analysis of <i>B. cenocepacia</i> WT Vs <i>fur</i> ::Tp mutant.....	299
Figure 5.13. Efficiency of plating of <i>B. cenocepacia</i> WT and <i>fur</i> ::Tp mutant.....	300
Figure 5.14. Activity of iron dependent Fur regulated promoters P _{orbS} , P _{orbS-69} , P _{flrS} and P _{bfd} in <i>B. cenocepacia</i> 715j and a <i>fur</i> mutant 715j- <i>fur</i> ::Tp	303
Figure 5.15. Complementation of the iron-regulation defect of a <i>B. cenocepacia fur</i>	305
Figure 5.16. Iron-dependent activity of the OrbS-regulated promoter P _{orbH} in <i>B. cenocepacia</i> WT and <i>fur</i> mutant strains	307
Figure 5.17. Activity of iron independent promoters P _{tac} and P _{cysI} in response to iron in <i>B. cenocepacia</i> in the presence and absence of <i>fur</i>	309
Figure 5.18. Activity of the Fur regulated P _{orbS} promoter in response to iron in <i>E. coli</i> WT and <i>fur</i> mutant strains	311
Figure 5.19. Activity of the OrbS-dependent promoter P _{orbHds6} in <i>E. coli</i> WT and <i>fur</i> mutant strains containing OrbS in response to iron.....	313
Figure 5.20. Activity of the variant <i>tac</i> promoter, P _{tac2} , in response to iron in <i>E. coli</i> WT and <i>fur</i> mutant strains.....	315
Figure 5.21. Cloning an SOE-PCR product encoding FLAG-tagged Fur into pEX18-TpTer- <i>pheS</i>	317

List of abbreviations

1CS: 1 Component system
2CS: 2 Component system
AHL – Acyl homoserine lactone
ASD – Anti-sigma domain
ATP – Adenosine tri phosphate
BACTH - bacterial adenylate cyclase two-hybrid system
Bcc – *Burkholderia cepacia* complex
BcCV – *Burkholderia cepacia* containing vacuole
BHI – Brain heart infusion
bp – base pairs
BSA – Bovine serum albumin
CAA - Casamino acids
cAMP – cyclic adenosine monophosphate
CAP - Catabolite activator protein
CF – Cystic fibrosis
CFU – Colony forming units
CLP – Cyclic Lipopeptide
COE - Carboxymuconolactone decarboxylases, oxidoreductases & epimerases
cPhe – p-chlorophenylalanine
CRE - Core recognition element
CRP – Catabolite repressor protein
CSS – Cell surface signalling
CTD – C-terminal domain
CTE – C-terminal Extension
D-BHI – Dialysed brain heart infusion
Da – Daltons
ddH₂O – double distilled H₂O (HPLC grade sterile water)
DMSO – Dimethyl sulfoxide
DNA – Deoxyribonucleic acid
DNase – Deoxyribonuclease
dNTP - Deoxynucleotide triphosphate
ECF – Extra-cytoplasmic function
EDTA – Ethylenediaminetetraacetic acid
HTH - Helix turn helix
IMAC – Immobilized metal ion affinity chromatography
IPTG – Isopropyl β-D-1-thiogalactopyranoside
IST – Iso-sensitest
kbp – Kilobase pairs
kDa - Kilodaltons
L – Litres

LB – Lysogeny broth
LPS – Lipopolysaccharide
LTTR – LysR-type transcriptional regulator
M – Molar
Mbp – Megabase pairs
ml – Millilitres
mM – Millimolar
MWCO - Molecular weight cut-off
NCR - Non-conserved region
NTD – N-terminal domain
°C - Degrees centigrade
OD – Optical density
OHL – Octanoyl homoserine lactone
OMP – Outer membrane protein
ORF – Open reading frame
PAGE – Polyacrylamide gel electrophoresis
PBS – Phosphate buffered saline
PCR – Polymerase chain reaction
PMF – Proton motive force
PVDF – Polyvinylidene fluoride
QS – Quorum sensing
Rmsd - Root mean square deviation
RNA – Ribonucleic acid
RNAP – RNA polymerase
ROS – Reactive oxygen species
RP_o - RNAP-Promoter Open complex
RP_c - RNAP-Promoter Closed complex
RP_i - RNAP-Promoter intermediate
SDS – Sodium dodecyl sulphate
SOE-PCR – Slice overlap extension PCR
TBDT – TonB-dependent transducer
TBDR – TonB-dependent receptor
TEC - Transcription elongation complex
TPR – Tetratricopeptide repeat
UP - Upstream promoter
V – Volts
WT – Wild-type
X-gal – 5-bromo-4-chloro-3-indolyl-beta-D-galacto-pyranoside
ZAS - Zinc anti- σ factor
 Δ – Deletion
 μ g – Microgram
 μ l – Microlitre

Chapter 1. Introduction

1.1 General introduction

Bacteria require efficient regulatory mechanisms to respond to changing environments and adapt themselves for survival, especially within hosts. The use of extra-cytoplasmic function sigma (ECF σ) factors is one such mechanism employed by bacteria to respond to external stress signals where the ECF σ factor associates with core RNA polymerase (RNAP) to form a RNAP holoenzyme and recognise promoters to initiate transcription of stress responsive genes. Particularly, opportunistic pathogens such as *Burkholderia cenocepacia*, use a number of different strategies for infection, pathogenesis, survival and inter-species competition but these mechanisms are still not fully elucidated at the molecular level. Although *B. cenocepacia* harbours 13 ECF σ factors, only two are currently characterised: OrbS and RpoE (Agnoli et al., 2006, Flannagan and Valvano, 2008). Hence, there is a need to study the roles of the other ECF σ factors to develop a better understanding of the bacterial stress response strategies and whether they may play a role in infection.

1.2 σ factor: structure

σ factors are bacterial proteins that function as dissociable sub-units of the RNAP. They bind to core RNAP to form a RNAP holoenzyme which is capable of transcription initiation (Murakami, 2015). A variety of σ factors are present which are functionally and structurally very diverse not only across different genera but within the same micro-organism. There are two families of σ factors based on amino acid sequence conservation that are termed as σ^{70} (or σ^A) and σ^{54} families. σ^{70} is the major σ factor in all bacteria and is encoded by *rpoD* in *E. coli* hence the *E. coli* protein is also referred to in some earlier literature as σ^D . It is more usually called σ^A or the 'primary σ factor' and is responsible for transcription initiation of most housekeeping genes, making it essential for cell growth and survival (Paget and Helmann, 2003). σ^{54} is also present in both Gram-positive and Gram-negative bacteria but is structurally and functionally unrelated to σ^{70} and is also dispensable for survival under laboratory conditions (Wosten, 1998).

The primary σ factor, shown in Figure 1.1, consists of 4 domains connected by linkers which give it flexibility (Campbell et al., 2002). These domains are further divided into sub-regions based on amino acid sequence conservation and function. Domain 1 (σ_1 or sometimes referred to as $\sigma_{1.1}$) consists of conserved region 1.1 and is auto-inhibitory in function (Schwartz et al., 2008). Domain 2 (σ_2) is made up of 5 helices and a non-conserved region (NCR) consisting of multiple helices. This domain interacts strongly with the core RNAP and recognizes a conserved hexameric nucleotide consensus sequence (TATAAT) that is called the '-10 promoter element' located approximately ten basepairs upstream of the transcription start site (+1). Conserved regions 2.3-2.4 are involved in this -10 element recognition (Campbell et al., 2002, Paget, 2015). One of the long helices contains region 2.4 while region 2.3 contains aromatic residues involved in DNA melting and basic residues for specific DNA binding. Domain 3 (σ_3) is composed of 3 helices and contains region 3.0, previously known as region 2.5, which is involved in recognition of the extended -10 region that is present in some promoters (Murakami et al., 2002). Domain 4 (σ_4) is present at the C-terminal end of σ^{70} and is known to interact strongly with RNAP. The linker between σ_3 and σ_4 is relatively long. σ_4 is composed of 4 helices and contains conserved regions 4.1 and 4.2. Region 4.2 recognizes a conserved hexameric nucleotide consensus sequence (TTGACA) located approximately 35 basepairs upstream of +1 called the '-35 promoter element' through its helix-turn-helix motif (Campbell et al., 2003).

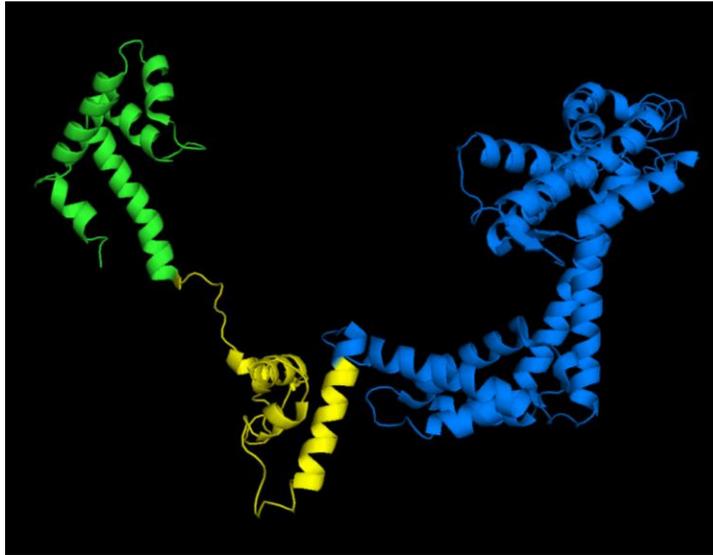


Figure 1.1. Structure of σ^{70} from *E. coli* when interacting with RNAP

σ_2 , σ_3 and σ_4 are shown in blue, yellow and green, respectively (σ_1 is absent in this crystal structure). This image was generated from the '4YG2' file deposited on PDB and the software PyMOL® Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC.

1.3 σ factor: function in transcription initiation

Bacterial core RNAP is the fundamental component required for transcription and it consists of five subunits: 2α , β , β' and ω (Zhang et al., 2012, Murakami, 2015). The defining features of a σ factor are its ability to bind to the core RNAP, specifically recognise promoter elements and initial DNA unwinding/strand separation in the -10 region forming the transcription bubble to initiate transcription (Feklistov et al., 2014).

1.3.1 Structural interaction of σ factor with RNAP in transcription initiation

β and β' form the two pincers of the crab-claw like structure of RNAP. When the σ subunit associates with the core RNAP, σ_2 interacts with a coiled coil within β' . This is one of the largest interactions based on surface area. σ_3 , in contact with the wide internal channel of RNAP above the active catalytic site with Mg^{2+} , simultaneously recognizes and

interacts with β (Young et al., 2001). The long inter-domain linker between σ_3 and σ_4 passes under the β flap and the tip of the β flap fits into the concave pocket-like structure of σ_4 . This causes a separation between σ_4 and σ_2 of approximately 11 Å so as to allow them to interact simultaneously with the -35 and -10 promoter elements (Kuznedelov et al., 2002). This σ -RNAP interaction consists of a series of events which takes place throughout the transcription initiation process causing conformational changes as extensive as the movement of region 1 (residue 59 in *E. coli* σ^{70}) by about 20 Å relative to region 2.4 (Callaci et al., 1999). The cumulative result of these interactions is the formation of a RNAP holoenzyme with the σ factor bound across one side of the polymerase which juxtaposes the specific regions of the σ factor involved in DNA interaction with the two conserved promoter elements (Burgess and Anthony, 2001).

The promoter search by RNAP is a very multifaceted process during transcription initiation. Extensive evidence suggests that RNAP holoenzyme slides along DNA by a one-dimensional movement to scan for the promoters (Park et al., 1982, Ricchetti et al., 1988, Sakata-Sogawa and Shimamoto, 2004, Feklistov, 2013). However, another study argues that three dimensional collisions by a 3D diffusion mechanism might be the only process by which RNAP searches for the promoter (Wang et al., 2013). Additionally, various activators and repressors influence the specificity of transcription initiation by RNAP (Browning and Busby, 2004).

A multi-step DNA transcription initiation model has been proposed based on structural, biochemical and biophysical studies that includes promoter recognition and melting of the -10 promoter element, transcription bubble formation and scrunching of the DNA as mRNA synthesis initiates. In the first step, RNAP makes contact with DNA through specific and non-specific interactions to first form a RNAP-promoter closed complex (RP_c) (Chen et al., 2010). Factors like temperature, salts and solutes affect K_d (dissociation rate constant) during RP_c formation (Saecker et al., 2011). The double stranded DNA is unwound with the help of the σ factor to allow it to pass through the narrow cleft of RNAP containing the catalytic site. This process of promoter recognition is suggested to be coupled with the

double strand opening step (Feklistov and Darst, 2011). Recognition of the -10 element (5' T₋₁₂A₋₁₁T₋₁₀A₋₉A₋₈T₋₇ 3') in the non-template strand by σ is crucial because it triggers the promoter melting step. σ_2 , σ_3 and β form a positively charged trough. A₋₁₁ is unstacked, in which amino acids T429 and W433 of σ^A of *T. aquaticus* were shown to be involved, followed by flipping the base into a shallow pocket in σ_2 where it interacts with tyrosine residues such as Y430 (Karpen and deHaseh, 2015). Similarly T₋₇ is flipped and fits into a complementary positively charged pocket of σ_2 (Feklistov and Darst, 2011, Liu et al., 2011, Feklistov, 2013). Residues like Q437 and T440 interact with T₋₁₂ (Zhang et al., 2012). Recognition of the -35 promoter element occurs through interaction with residues in the helix turn helix motif of region 4.2 (Pabo and Sauer, 1992). Positively charged arginine residues interact with the bases at the -31 and -33 positions on the template strand and extensive interaction between region 4.2 and the phosphate backbone of the non-template strand takes place (Campbell et al., 2002).

Although the -10 and -35 promoter elements are the most conserved and common promoter recognition motifs, other regions of the promoter are involved in contacts with the RNAP. In the case of *E. coli* σ^{70} , region 1.2 makes direct contacts with the promoter discriminator element 'GGGA' located at positions -6 to -3 at some promoters; especially M102 makes contacts at the -5 position (Haugen et al., 2006, Barinova et al., 2008, Haugen et al., 2008b). G₋₆ is unstacked and flipped by region 1.2 of σ (Miropolskaya et al., 2012, Zhang et al., 2012). Another important interaction is carried out by region 3.0 of σ^3 that interacts with the extended -10 element at some promoters (Mitchell et al., 2003). RNAP also recognizes the Z-element, located within the spacer between the -10 and -35 promoter elements, and the core recognition element (CRE), located at the -4 to +2 position (Yuzenkova et al., 2011, Zhang et al., 2012). The latter is a sequence-specific recognition in which RNAP unstacks and flips G at position +2 and inserts it into a pocket where it makes contacts with positively charged residues of the β subunit. The C-terminal domain of the α subunit recognizes an A+T-rich upstream promoter (UP) element of up to 20 bps located upstream of the -35 element (Gourse et al., 2000).

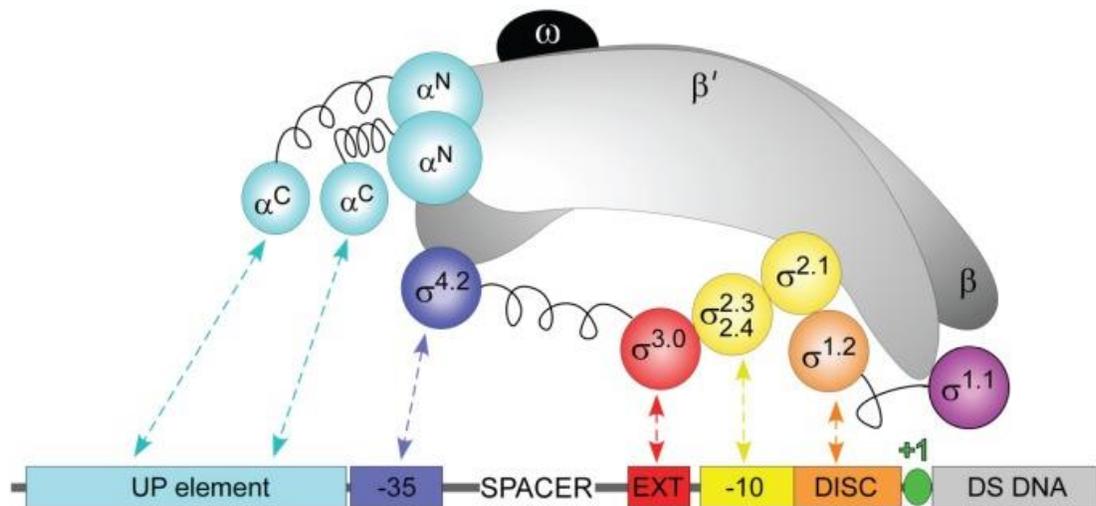


Figure 1.2 Structure of RNAP holoenzyme with the corresponding promoter recognition regions
 The sub-units of core RNAP, β and β' are shown in grey, α in light blue and ω in black. The regions of σ are shown in purple, orange, yellow, and dark blue colour co-ordinated with the promoter elements they interact with directed by the dashed arrows. Other interactions of the RNAP (not illustrated here) include promoter regions – core recognition element and Z-element (see text for further details). Figure credits: RUFF, E. F., RECORD, M. T. & ARTSIMOVITCH, I. 2015. Initial Events in Bacterial Transcription Initiation. *Biomolecules*, 5, 1035-1062. Copyright covered by creative Commons By 4.0.

1.3.2 Kinetics of the σ factor-RNAP interaction in transcription initiation

Recognition of these promoter elements differs in different promoters though the prerequisite in all instances is to fine-tune the promoter recognition process. These interactions cause isomerization of residues forming multiple short-lived intermediates (RP_i), to eventually form a transcription bubble by unwinding of about 12 nucleotides to form RNAP-promoter open complex (RP_o). An important step is the formation of the second intermediate RP_{i2} from the first intermediate RP_{i1} ($RP_{i1} \rightarrow RP_{i2}$) and the reverse ($RP_{i2} \rightarrow RP_{i1}$) which is a kinetically significant rate determining step (Davis et al., 2007, Ruff et al., 2015). Recognition of upstream promoter elements by σ_4 and α CTD RNAP influences the kinetics of this bottleneck isomerization step. The formation of an open promoter complex by RNAP

occurs through use of free energy by electrostatic force between the positively charged RNAP and negatively charged DNA to unwind the dsDNA from -11 to +2 position, bend it and allow the active site cleft to grip the template DNA strand. A hypothesis based on structural evidence suggests that DNA strand opening occurs above the active site cleft after which the template strand is placed at the active site (Murakami et al., 2002, Vassylyev et al., 2002). However, more recent kinetic and footprinting data suggest that the dsDNA is actively opened by RNAP after being loaded as dsDNA into the cleft (Kontur et al., 2010, Gries et al., 2010). Once the bases of the template strand can be read at the active site of the polymerase, binding of NTPs and formation of phosphodiester bonds commences.

1.3.3 σ release

Elongation can be either productive or abortive. DNA melting downstream of the +1 site does not take place actively and it has been proposed that the initial transcription steps cause 'DNA scrunching stress' which leads to either abortive initiation (mRNA release) or promoter escape and productive initiation (Kapanidis et al., 2006). Region 1.2 along with β (186-433) and perhaps region 1.1 are required for this process and play key roles in promoter escape to ensure that abortive initiation does not occur (Revyakin et al., 2006, Hook-Barnard and Hinton, 2009, Pupov et al., 2010, Bochkareva and Zenkin, 2013). Typically, the σ factor is stochastically released from the complex after 12-15 nucleotides are incorporated into mRNA and the elongation phase of transcription begins. Conserved region 3.2 of the σ factor plays an important role in this promoter escape and has been proposed to be required for σ -dependent promoter proximal pausing (Pupov et al., 2014). Nevertheless, studies also suggest that σ factor is held along with RNAP during the elongation process (Bar-Nahum and Nudler, 2001, Mukhopadhyay et al., 2001, Raffaele et al., 2005, Saecker et al., 2011). A summary of the isomerization steps in this process is depicted in Figure 1.3.

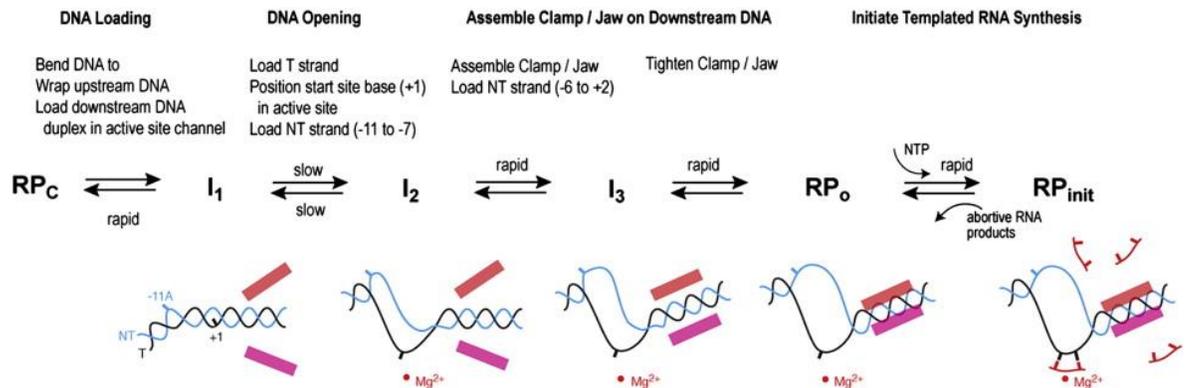


Figure 1.3. Transcription initiation process

RNAP forms a closed complex with the promoter region RP_c (DNA unmelted) followed by subsequent intermediate steps before forming an open promoter complex (RP_o) to then form RP_{init} , the initiation complex where bases on the DNA strand can be read for mRNA synthesis (and may involve abortive initiation). Red and pink rectangles represent the clamp/jaw of the RNAP. Reprinted by permission from Elsevier: Journal of Molecular Biology, (Saecker et al., 2011), copyright 2017.

1.4 Classification of σ factors

The σ^{70} family consists of σ factors that possess regions of conserved amino acid sequence with the major σ factor σ^{70} (σ^A). This family is further divided into sub-families based on phylogenetic relationships (Helmann, 2002). Group 1 consists of primary σ factors that are orthologues of σ^{70} of *E. coli* and are involved in transcription initiation of housekeeping genes. Group 2 consists of σ factors closely related to σ^{70} but dispensable for growth and present in only some bacterial species. They are involved in some stress responses and include RpoS, the general stress response or stationary phase σ factor. Group 3 σ factors are further divergent in sequence conservation from σ^{70} and are divided into sub-groups depending on their function which include roles in flagella formation (e.g. SigD of *B. subtilis*), response to heat shock (e.g. RpoH of *Caulobacter crescentus*) and sporulation (e.g. SigG of *B. subtilis*). The most distantly related sub-family is Group 4, an extensive group of σ factors that are also referred to as ECF σ factors (described in the next section).

Examples include RpoE and FecI of *E. coli*, SigW of *B. subtilis*, PvdS of *Pseudomonas aeruginosa* and OrbS of *B. cenocepacia* (Butcher et al., 2008).

1.5 Extra-cytoplasmic function σ factors

Since Group 4 σ factors were recognized as being involved in transcription initiation of genes in response to stresses sensed outside the cytoplasm they were originally termed extracytoplasmic function (ECF) σ factors (Lonetto, et al 1994). They consist of a very diverse and complex group of proteins which have multiple different modes of functioning across different bacteria (Helmann, 2002, Gruber and Gross, 2003, Osterberg et al., 2011, Feklistov et al., 2014, Paget, 2015).

1.5.1 Structure

Although an alignment of Group 1 and Group 4 σ factors shows very little sequence similarity, the structural conservation is surprisingly very high (Wosten, 1998, Gruber and Gross, 2003). ECF σ factors do not have $\sigma_{1.1}$ and lack most of σ_3 , but σ_2 and σ_4 are similar to those of σ^{70} (Figure 1.4) and recognize the -10 and -35 promoter elements in a similar way to the primary σ factor which explains this maintenance of the structural conservation (Gruber and Gross, 2003, Busby, 2009). ECF σ factors have a linker that connects σ_2 and σ_4 that follows the same path as region 3.2 which serves as the σ_3 - σ_4 interdomain linker in σ^{70} (Campbell et al., 2002).

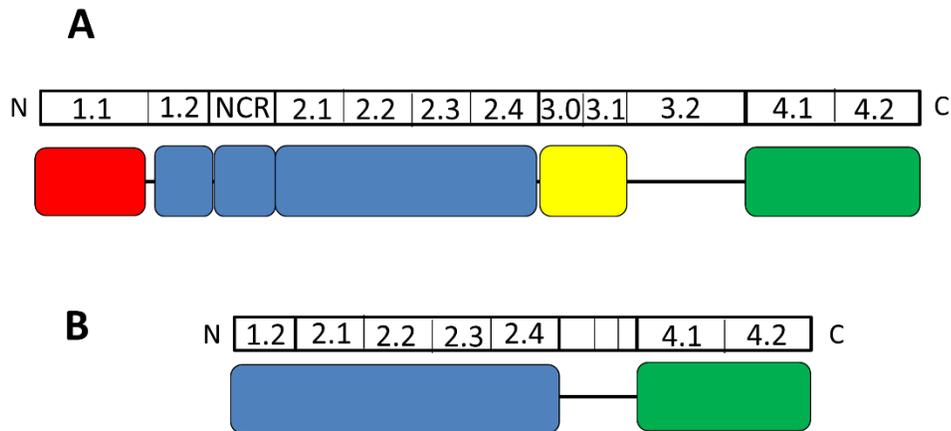


Figure 1.4 Domain organisation of ECF σ factors compared to σ^{70} (not drawn to scale)

(A) Structural domain organization of σ^{70} and (B) a typical ECF σ factor where σ^1 , σ^2 , σ^3 and σ^4 are shown in red, blue, yellow and green, respectively. The sub-regions comprising the σ factors are given above each domain organization.

1.5.2 Mechanism

The key functional distinction between σ^{70} and a typical ECF σ is that the former are required to recognize many different housekeeping genes whereas the latter are required for the transcription initiation of only a relatively small set of genes in response to very specific conditions. As a result of this difference, ECF σ factors have developed a more stringent recognition mechanism but has reduced efficiency in promoter melting compared to σ^{70} (Darst et al., 2014). As in σ^{70} , σ_2 and σ_4 of ECF σ factors associate with RNAP and recognize conserved promoter elements for transcription initiation. However, as most of σ_3 is absent in ECF σ factors this indicates that recognition of extended -10 promoter element does not take place.

A mutational analysis performed on the *P. aeruginosa* ECF σ factor PvdS revealed that region 2.1 and 2.2 interacts with core RNAP as in Group 1 σ factors. Highly conserved residues within these regions were found to massively affect binding when substituted by other amino acids. Unlike σ^{70} which has many aromatic residues in region 2.3 whose role is associated with DNA melting, ECF σ factors have fewer aromatic residues but one particular aromatic residue seemed to be specifically important. Amino acid substitutions in region 2.4 resulted in a significant loss of the activity of the σ factor, particularly, three amino acids (N73, D77, R80) seemed to be important for the -10 promoter element recognition. The HTH motif of region 4.2 which is involved in DNA interaction during the -35 promoter element recognition, was shown to contain several residues that were critical for this interaction (Wilson and Lamont, 2006). Promoter recognition by ECF σ factors require relatively fewer activators compared to σ^{70} -dependent promoters and also in contrast to σ^{70} the recognition of the -35 promoter element cannot be compensated for by recognition of the UP element by the α subunit C-terminal domain (Rhodius et al., 2012).

Recent structural studies in σ^E show that ECF σ factors select only one base of the non-template strand equivalent to A₋₁₁ in the case of σ^{70} that is flipped within a pocket in σ_2 and there is no equivalent to the other base, T₋₇ that is flipped by σ^{70} (Figure 1.5). Moreover, the amino acids forming the positively charged pocket where the base is flipped are not conserved among different ECF σ factors, allowing for diversity in the -10 promoter element recognition (Campagne et al., 2014, Campagne et al., 2015). Experimental evidence derived from analysis of the ECF σ factor Tt-RpoE1 of *Thermoanaerobacter tencongensis* show that a 'DXXR' motif present in region 2.4 of σ^2 interacts with CGTC (-13 to -10) bases. The conserved aspartic acid residue specifically was reported to be required for the recognition of G₋₁₂ (Liu et al., 2012). The *Mycobacterium tuberculosis* ECF σ factor σ^C , when superposed on the primary σ factor σ^A of *T. aquaticus* showed σ_4 to be highly structurally conserved with a root mean square deviation (rmsd) of 1.6 Å, however, superposition of σ_2 of the two σ factors had an rmsd of 2.97 Å. An rmsd of 1.4 Å of the corresponding domains was found when σ^C was superposed with ECF σ factor σ^E of *E. coli* (Thakur et al., 2007). Another study

proposed that σ_4 of Group 1 and Group 4 interact with the -35 promoter element differently due to the differences in the residues in their HTH. σ^E unlike primary σ factors recognizes a 7 bp motif, GGAACTT, at the -35 position. The 'AA' present at -33 and -32 positions are indirectly critical since their geometry allows specific binding but no direct protein-DNA contact occurs at these two positions. The HTH motif of σ^E interacts more extensively at the major groove as compared to that of the primary σ factor and further it is not involved in bending of DNA (Lane and Darst, 2006). On the other hand, in the ECF σ factor SigR of *Streptomyces coelicolor*, a methionine residue in region 4.2 was found to specifically recognize A₋₃₁ of the -35 promoter element (GGGAAT) (Kim et al., 2016). It has been suggested that in the -10, -35 and UP elements in promoter recognition, although some bases are preferentially recognized in the single-stranded form and some in the double-stranded form, this recognition can be mixed and matched for effective transcription initiation by housekeeping as well as ECF σ s (Hook-Barnard and Hinton, 2007, Guzina and Djordjevic, 2017). In *Corynebacterium glutamicum* it has been reported that certain promoters are capable of being recognised by both the housekeeping and the ECF σ factors which is an unusual phenomenon (Silar et al., 2016).

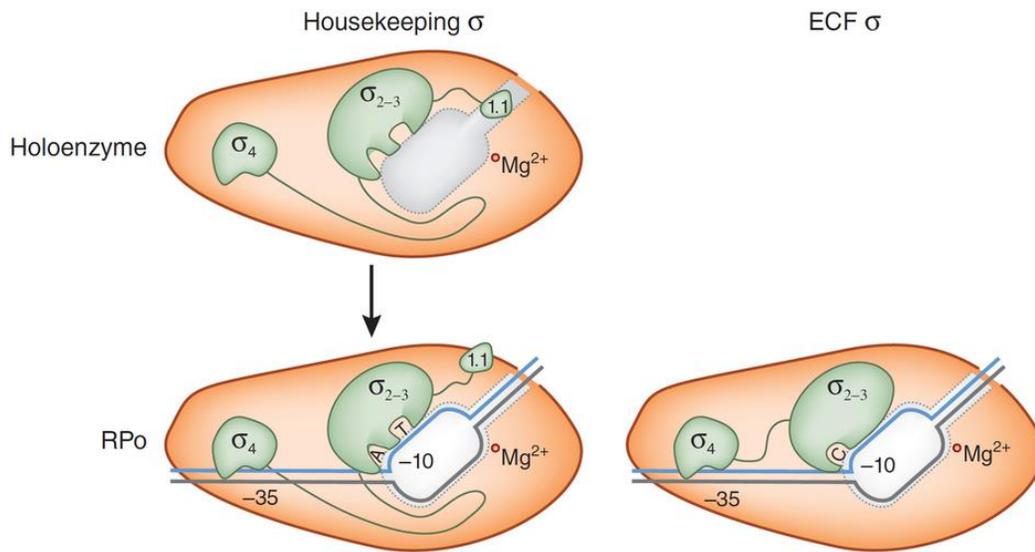


Figure 1.5 Promoter melting at the -10 promoter element by σ^{70} and an ECF σ

σ_2 of σ^{70} flips two bases at the -10 promoter element (A₋₁₁) and (T₋₇) while σ_2 of the ECF σ RpoE uses only one protein pocket to flip one base (in this case cytosine residue) when initiating the DNA unwinding step. Reprinted by permission from Nature publishing group: Nature Structural and Molecular Biology, (Darst et al., 2014), copyright 2017.

Often, under inducing conditions, ECF σ factors are auto-regulated as they initiate transcription by recognising their own promoter (Paget and Helmann, 2003). However, a low-level constitutive promoter also exists as shown in Figure 1.6 to maintain a low level of initial ECF σ for initial response. When required under 'stressful' conditions, the auto-regulatory promoter is utilized, to appropriately provide an amplified amount of ECF σ .

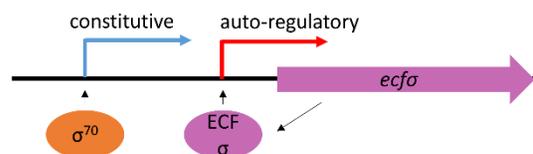


Figure 1.6 Promoters that drive transcription of genes encoding ECF σ factors

Often, ECF σ factors have a weak constitutive promoter represented by the blue arrow, which can be used for initial transcription using the ECF σ dependent promoter represented by the red arrow for an amplified response during inducing conditions.

1.5.3 Classification

The nomenclature and classification of this diverse family of proteins is complex. They are named as either a four letter word, with the first three mostly being 'Sig'/'Rpo' or σ with a superscript that may be a letter or a number signifying the molecular weight in kDa (Gruber and Gross, 2003, Staron et al., 2009). All ECF σ factors are included in Group 4 but are functionally and mechanistically very diverse (Paget and Helmann, 2003). In 2009, a comprehensive study of all known ECF σ factors was carried out using comparative genomics and phylogenetic analysis. This approach led to identification and classification of ECF σ factors into 43 different subgroups based on genomic context, structural similarities and roles (Staron et al., 2009). This systematic classification simplifies the process of elucidating regulatory mechanisms of this previously ambiguous and constantly expanding group of proteins. However, it is still an inconsistent classification because in some cases σ factors belonging to the same group based on their structure/genomic context may have divergent roles and vice versa. This classification was further extended in another bioinformatics study that identified a further 8 subgroups of the ECF σ factor group (Jogler et al., 2012).

1.5.4 Regulatory mechanisms

ECF σ factors need to compete with primary σ factors and other alternative σ factors for binding to core RNAP as well as compete with regulators such as ppGpp that can bind the RNAP and modulate its activity (Haugen et al., 2008a). In the presence of inducing conditions, ECF σ factor activity is regulated by sophisticated mechanisms to increase their concentration and affinity towards RNAP (Osterberg et al., 2011).

The conventional model of ECF σ factor activity regulation suggests that in the absence of stimuli from the external environment, the ECF σ factor remains inactive while bound in a stoichiometric complex to an inhibitory protein called the anti- σ factor that is anchored in the cytoplasmic membrane by a centrally located transmembrane domain

(Helmann, 1999). An external signal causes this anti- σ factor to release its cognate bound σ factor. However, in the past few years, experimental evidence suggests that the anti- σ factor may also be involved in σ activation under inducing conditions or alternative mechanisms exist by which ECF σ factors are regulated that do not involve anti- σ factors. Some mechanisms for regulating ECF σ factors are described below.

In the presence of inducing conditions, some anti- σ factors can undergo **regulated intra-membrane proteolysis (RIP)** to release its cognate σ factor. Well studied ECF σ factors that are regulated by this mechanism include σ^E of *E. coli*, σ^V and σ^W of *B. subtilis*, RpoE or σ^E (AlgU) of *P. aeruginosa* and CarQ of *Myxococcus xanthus* (Heinrich and Wiegert, 2009, Ho and Ellermeier, 2012, Hastie et al., 2013, Helmann, 2016). RIP in σ^E of *E. coli* is further explained in section 1.5.5.1. Some iron-starvation ECF σ factor regulators are also auto-proteolysed (Bastiaansen et al., 2015a).

Some ECF σ factors are regulated by a **partner-switching mechanism**. Here the σ factor and an 'anti-anti σ factor' compete with each other in order to bind to the anti- σ factor. Examples include σ^B in *B. subtilis* and σ^T in *Caulobacter crescentus*. NepR the anti- σ factor of σ^T involved in general stress-response, in the α -proteobacterium *Caulobacter crescentus* has an antagonist PhyR (anti-anti- σ). In addition, PhyK, a histidine kinase plays an important role in by phosphorylating PhyR (Francez-Charlot et al., 2009, Lourenco et al., 2011). In another α -proteobacterium *Bartonella henselae*, the ECF σ factor RpoE also positively regulates this system in addition to PhyR (Tu et al., 2016).

Certain ECF σ factors are reported to be bound to **soluble anti- σ factors** where the σ factor activity is regulated by conformational changes of the associated anti- σ factor. For example, in *Streptomyces coelicolor*, SigR and RsrA (σ and anti- σ) are involved in detecting and responding to thiol oxidative stress (Paget et al., 1998, Paget et al., 2001). RsrA, a zinc binding anti- σ factor (ZAS) is bound to a zinc ion and in addition SigR is bound to the hydrophobic core of RsrA. Thiol-induced oxidative stress conditions are detected by the conserved cysteine thiols of RsrA, causing the release of the bound zinc ion (Zn^{2+}). The conformational change due to this formation of a disulphide bond between two cysteine

residues of RsrA allows the release of SigR (Li et al., 2002, Zdanowski et al., 2006, Heo et al., 2013, Rajasekar et al., 2016). In case of the ECF σ factor CnrH of *Cupriavidus metallidurans* that responds to metal stress, the heterodimeric anti- σ factor CnrXY is anchored to the membrane. When it detects cobalt or nickel in the external environment, through its metal-detecting domain CnrX, CnrY undergoes conformational changes and releases the bound σ to initiate transcription of genes involved in metal resistance (Trepreau et al., 2011a, Trepreau et al., 2011b).

A **two component system** (2CS), CseBC, regulates SigE activity in *Streptomyces coelicolor* at the transcriptional level thus replacing the requirement for an anti- σ factor. Under stress induction, CseC auto-phosphorylates and the phosphate is then transferred to CseB through a two component system pathway. Phosphorylated CseB induces CseB expression of the gene encoding the ECF σ factor SigE, involved in responding to cell envelope stress (Hong et al., 2002). Another related example includes SigP of *Porphyromonas gingivalis* that regulates secretion of pathogenically important proteins via the type IX secretion system where the activity of the σ factor is regulated by a two-component system consisting of PorX (response regulator) and PorY (histidine kinase) (Kadowaki et al., 2016).

ECF σ factors such as those belonging to the ECF41 group lack a canonical anti- σ factor, but instead contain a large C-terminal extension of about 100 amino acids. A study by Wecke, *et al.*, (2012) identified that the C-terminal extension acts as an **auto-regulatory** domain to increase binding affinity to RNAP and σ factor-dependent promoter activation thus replacing the need for an anti- σ factor. While the C-terminal extension plays both a positive and negative role since the full-length extension is inhibitory and a partially truncated extension acts as an activator, the exact mechanism of auto-regulation remains to be established (Wecke et al., 2012). Other examples of ECF σ factors C-terminal regulatory domain include CorE and CorE2 involved in regulation of copper and cadmium uptake in *Myxococcus xanthus*. They have a C-terminal extension containing a cysteine rich domain where the cysteine residues are required for both activation and inactivation of the σ factor by identifying the redox state of copper/cadmium. It was shown that Cu¹⁺ caused

the inactivation and Cu^{2+} caused the activation of CorE, respectively. The C-terminal extension thus behaves as an on board fused regulator similar to the conventional anti- σ factor (Gomez-Santos et al., 2011, Marcos-Torres et al., 2016). Examples of ECF σ factors that do not have an extension include SigX and SigM from *B. subtilis* but are reported to be induced to auto-regulate by CshA (an RNA helicase) that associates with the RNA polymerase core enzyme in response to increased glucose concentrations. (Ogura and Asai, 2016).

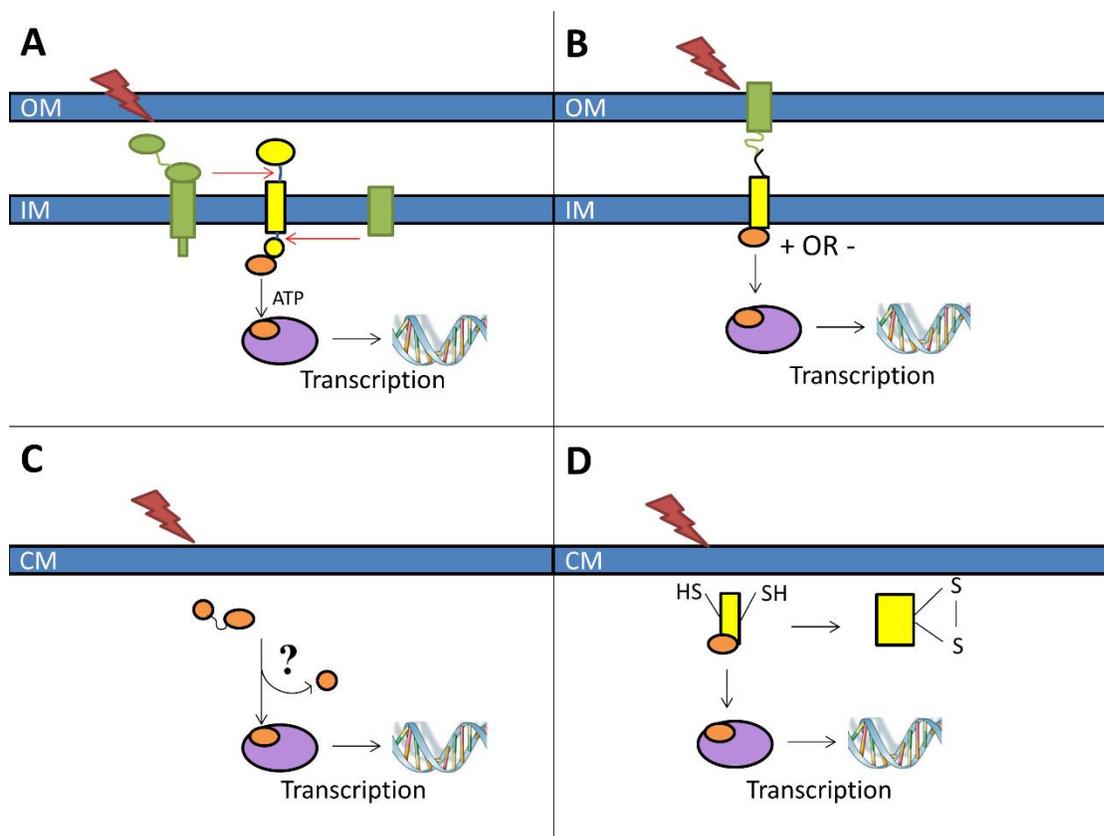


Figure 1.7 Schematic representation of some mechanisms for regulation of ECF σ factors

The outer (OM) and inner/cytoplasmic membranes (IM) are labelled accordingly. The external stress is shown in maroon, ECF σ factor in orange, RNAP core enzyme in purple, anti- σ factor in yellow and other proteins in green. (A) Regulated intra-membrane proteolysis where red arrows indicate proteolysis of the anti- σ factor to release σ factor. (B) Iron-starvation sub-class (CSS-cell surface signalling) where a receptor protein passes the signal for σ release where the anti- σ factor may or may not play a role of an activator. (C) Autoregulation where a domain of the ECF σ factor itself

regulates the activity. (D) Conformational changes in the cytoplasmic anti- σ factor under inducing conditions release the σ factor.

1.5.5 Role in stress response

Numerous studies so far show that ECF σ factors play roles in response to cell envelope stress, oxidative stress, iron starvation, biosynthesis of exopolysaccharides / antibiotics / carotenoids / fatty acids, resistance to heavy metals, induction of virulence factors, etc. (Bashyam and Hasnain, 2004, Kazmierczak et al., 2005, Woods and McBride, 2017). For example, the ECF σ factor, AlgU (σ^E), of the opportunistic pathogen *Pseudomonas* has roles in regulating alginate (exopolysaccharide) production and hence mucoidy, induction of efflux pumps and β -lactamase production, anti-microbial resistance, oxidative stress response and countering the host immune response by regulating the transcription of more than 30 genes including those involved in virulence (Martin et al., 1994, Firoved et al., 2002, Firoved and Deretic, 2003, Balasubramanian et al., 2011, Markel et al., 2016). SigR from *Streptomyces coelicolor* has roles in antibiotic resistance in addition to responding to thiol-oxidative stress (Yoo et al., 2016). RpoE (σ^E) in *Salmonella enterica* serovar Typhi has been reported to play a role in antibiotic resistance where an increase in resistance was observed in *rpoE* mutants whereas in serovar Typhimurium a decrease was observed (Humphreys et al., 1999, Crouch et al., 2005, Xie et al., 2016). Thus, the use of ECF σ factors allows the cell to survive in a competitive niche in the environment. Since it is critical for pathogens to recognise, adapt and survive in the unfavourable conditions present inside their hosts, many ECF σ factors prove to be important virulence determinants (Raivio, 2005, Boor, 2006, Rowley et al., 2006).

More recently, many unconventional ECF σ factor functions have been described. One such ECF σ factor is σ^{ShbA} in *Streptomyces griseus*. Here σ^{ShbA} controls transcription of the gene encoding the primary σ factor σ^{HrdB} and is the first of its kind to be reported. σ^{ShbA} -deficient mutants showed defects in morphology, streptomycin production, down-regulation of many housekeeping genes and extensive loss of mycelia in the stationary phase. *S. griseus* might use its alternative σ factor to fine tune regulation of housekeeping

genes such as those involved in mycelial growth in response to the external environment (Otani et al., 2013). Another unusual example of σ factor-mediated transcription initiation was elucidated in a study carried out on *B. subtilis*. Here two independent σ subunits (SigO and RsoA) function as separately encoded σ_2 and σ_4 domains instead of being fused as in a conventional ECF σ factor. Under acid stress conditions, the anti- σ factor RsiO which is usually bound to SigO, releases it, allowing SigO to associate with RsoA and subsequently to bind core RNAP (MacLellan et al., 2009, Davis et al., 2016). Presently, hundreds of ECF σ factors and their roles are being researched which will eventually help in understanding the mechanism of cross-talking and non-cross-talking (orthogonal) σ factors within the same bacterium (Feklistov et al., 2014).

The role and detailed molecular mechanisms of three ECF σ factors that serve as paradigms in the study of these proteins, σ^E of *E. coli* and Fecl and PvdS of the iron-starvation class, are described below.

1.5.5.1 The cell envelope stress σ factor, σ^E , of *E. coli*

σ^E (RpoE) of *E. coli* is involved in responding to disruption of the cell membrane due a variety of factors. The anti- σ factor, RseA inhibits its activity through binding between σ^2 and σ^4 of σ^E (Campbell et al., 2003). Stresses that result in the presence of misfolded outer membrane proteins (OMPs) are recognized by the periplasmic protease DegS in order to initiate transcription of σ^E -dependent genes (Heinrich and Wiegert, 2009). DegS, present on the inner face of the outer membrane, recognizes the 3 C-terminal amino acids of denatured/misfolded OMPs (Walsh et al., 2003, Brooks and Buchanan, 2008). As a result, regulated intra-membrane proteolysis of RseA is triggered by introducing a cut by DegS at position 148 and 149 of RseA. This specificity is regulated by the PDZ domain of DegS that allosterically activates the proteolytic affinity of DegS towards RseA (Wilken et al., 2004). This cleavage releases a protein RseB that also responds to misfolded OMPs and protects RseA. A RIP protease called RseP located at the inner membrane can now cleave RseA at the cytosolic side because RseB, the inhibitor of RseP once cleaved can no longer be

effective (Grigorova et al., 2004). The cytoplasmic remnant of RseA bound to σ^E is then cleaved by ATP driven cytosolic proteases such as ClpXP and which have the capability of unfolding proteins (Flynn et al., 2004). RseA is subsequently degraded by energized proteases and σ^E is free to bind core RNAP and initiate transcription (Heinrich and Wiegert, 2009). Deletion of the gene encoding this σ factor caused death of laboratory strains; however, some clinical isolates could persist in the absence but with a deficiency in efficiently resisting polymyxin B, an antibiotic that destabilises the bacterial outer membrane (Hayden and Ades, 2008, Kulesus et al., 2008). This mechanism is schematically depicted in Figure 1.8.

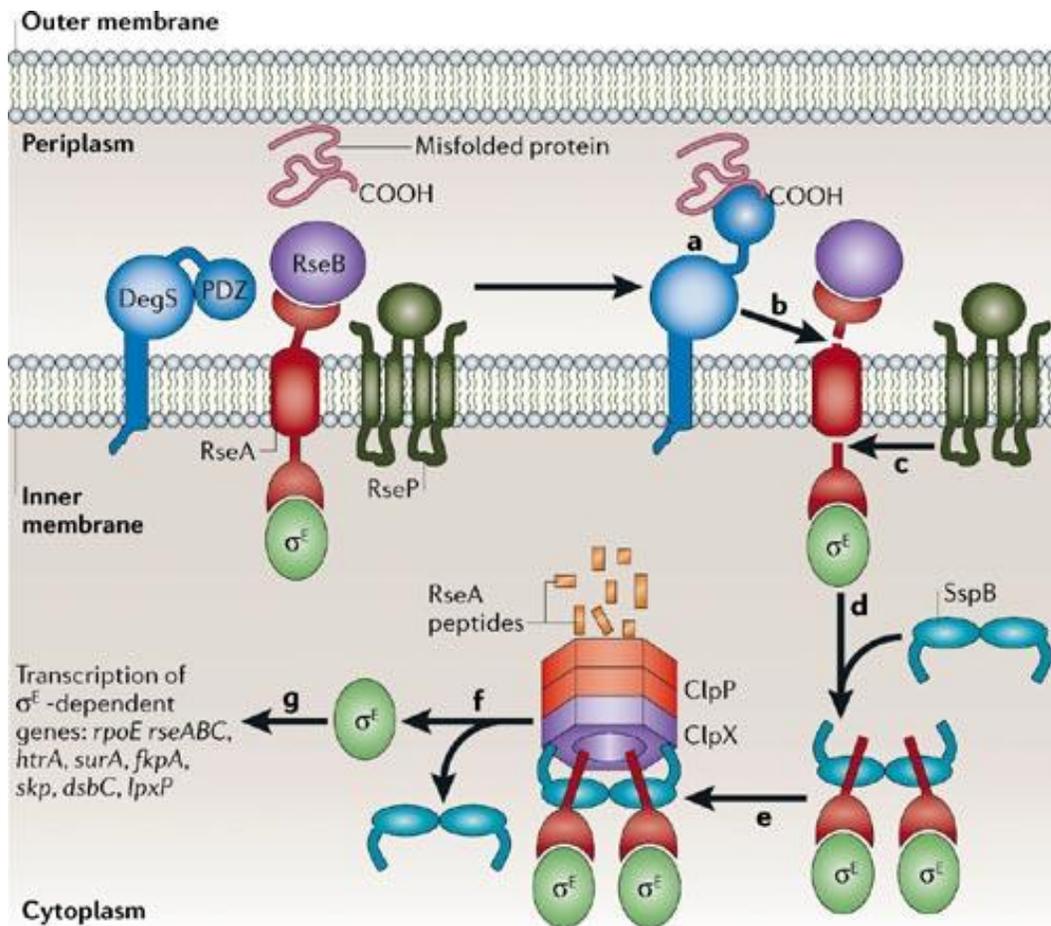


Figure 1.8. Summary of the σ^E system in *E. coli*

(a) The carboxyl end of a misfolded OMP interacts with the PDZ domain of the DegS protease (b) DegS cleaves RseA on recognizing mis-folded protein. (c) RseP cleaves RseA as RseB is cleaved. Enzymes like (d) SspB and (e) ClpXP further cleave RseA. (f) σ^E can associate with core RNA polymerase and form a RNAP holoenzyme. (g) Transcription of ECF σ factor-dependent genes is initiated. See text for further details. Reprinted by permission from Nature publishing group: Nature Structural and Molecular Biology, (Rowley et al., 2006), copyright 2017.

1.5.5.2 Role in iron regulation (Fecl of *E. coli* and PvdS of *P. aeruginosa*)

Iron is an essential micronutrient utilised by all forms of life. To carry out vital functional and metabolic processes such as enzymatic reactions, energy generation, DNA synthesis, etc. bacteria require iron; but it is not usually available in a biologically usable form in the environment and in plant/animal hosts (Ratledge and Dover, 2000). Although present in ample amounts in the environment, iron is usually in the insoluble ferric oxyhydroxide form due to oxidation under normal conditions (Neilands, 1995). In host tissues and fluids iron is bound by the transferrin and lactoferrin family proteins. Therefore, bacteria have developed so-called direct or indirect mechanisms for iron uptake, serving as an evolutionary advantage for survival in iron starved niches in natural environment as well as within host tissues (Schaible and Kaufmann, 2004). Lactoferrin, haem, ferritin, etc. can be directly sourced for iron intake or, indirectly, bacteria can secrete siderophores and haemophores as iron scavengers (Biville et al., 2004, Wandersman and Delepelaire, 2004, Choby and Skaar, 2016). Additionally, excess iron is harmful for cells due to formation of hydroxyl radicals via the Fenton reaction. As surplus iron cannot be excreted and as there is a limited capacity for iron storage this leads leading to the requirement for effective regulatory mechanisms for controlling iron uptake in order to maintain appropriate levels of intracellular iron. Transcriptional regulation of siderophore-mediated iron uptake systems is often facilitated through the action of ECF σ factors, two-component systems or AraC-type regulators (Saha et al., 2013).

ECF σ factors of the iron starvation class regulate genes involved in the acquisition of iron, and in case of internally synthesized siderophores, regulate the genes involved in its assembly and export. **Siderophores** are low molecular weight (approximately 400-1000 kDa on average), generally water soluble iron chelators that have very high affinity for ferric iron (Neilands, 1995). Bacteria are capable of utilising internally synthesized as well as externally available (xenosiderophore) forms of siderophores (Wandersman and Delepelaire, 2004). Siderophores largely possess bidentate catecholate, hydroxamate or hydroxycarboxylate iron binding groups, subject to how the oxygen ligand is coordinated

during iron capture (Saha et al., 2013). A smaller number of siderophores contain a phenolate group located adjacent to a heterocyclic ring that also serves as a bidentate ligand. Siderophores may contain one, two or three identical bidentate groups or they may contain combinations of these groups (mixed class). Siderophore structure and diffusibility are reported to have evolved based on the habitat of the bacterium. Less diffusible siderophores seem to be a characteristic of bacteria occupying unstructured environmental niches like sea water, whereas more diffusible siderophores might be an evolutionary characteristic of bacteria occupying structured niches such as soil and hosts, thus allowing nutrient sharing (Kummerli et al., 2014). Biosynthesis and assembly of the siderophore compound can be either carried out via non-ribosomal peptide synthetases (NRPSs) or without them (NRPS-independent synthesis, NIS) (Barry and Challis, 2009). NRPSs are multi-domain enzyme structures primarily involved in synthesis of siderophores composed of amino acid constituents (Crosa and Walsh, 2002).

Siderophores are reported to be involved in virulence since they can compete with lactoferrin in humans for iron capture thereby allowing colonisation of the host and are reported to be directly involved in biofilm formation where iron acts as the signal for formation of biofilms (Xiao and Kisaalita, 1997, Banin et al., 2005, Condon et al., 2014). Siderophores also have wide bioremediation, biotechnological and therapeutic applications (Miethke and Marahiel, 2007, Ahmed and Holmstrom, 2014).

Iron-siderophore complexes are transported from outside the cells to the periplasm of Gram-negative bacteria via **TonB-dependent receptors (TBDRs)**. These outer-membrane spanning proteins derive energy for internalisation of the ferric-siderophore complex from the TonB-ExbB-ExbD complex present at the inner membrane. TBDRs consist of a 22 strand beta-barrel structure with a globular plug/hatch domain in their centre. At the N-terminal end of this domain, which faces the periplasm, is located a short amino acid sequence known as the TonB box, which is required for recognition of the TBDR by the C-terminal domain of TonB (Noinaj et al., 2010). When a siderophore bound to iron is recognised by a TBDR, massive conformational changes influence exposure of the TonB box. At least one

TonB subunit with an ExbB pentamer and ExbD dimer is present in the TonB complex anchored at the inner membrane. Energy derived from the proton motive force is derived through interactions between ExbD and ExbB causing conformational changes to TonB (Celia et al., 2016). The TonB box forms an extended β -sheet with the TonB CTD.

Next, to transport the internalised Fe-siderophore complex from the periplasm into the cytoplasm, cytoplasmic membrane transport systems are required. ATP-binding cassette transporters (**ABC transporters**) are one such mechanisms. The energy-driven transport by ATP hydrolysis is carried out through the cytosolic components of this multi-component system. Gram-negative bacteria have one or more of the four types of ABC transporters classified based on the number and types of components for this transporter (Miethke and Marahiel, 2007). In some systems iron is released from the Fe-siderophore complex within the cytoplasm, often by a chemical modification of the siderophore. The TBDR and ABC transporter encoding genes are together regulated by the ECF σ factor which is in turn regulated at the transcriptional level by the global regulator of iron called **Ferric uptake regulator** (Fur).

Fur is a conserved small low molecular weight (17 kDa in *E. coli*) protein which has affinity for Fe^{2+} (Escolar et al., 1998). In iron rich conditions its dimeric form negatively regulates transcription of target genes by recognising a consensus 19 bp inverted repeat, thus blocking promoter regions. In iron-starved conditions, iron dissociates from Fur, which then causes it to dissociate from the promoter regions thereby allowing their transcription. In addition to Fe^{2+} , Fur is also capable of binding and being activated by Fe^{2+} and Fe^{3+} as well as Zn^{2+} , Co^{2+} , Mn^{2+} (Mills and Marletta, 2005, Fillat, 2014). Fur regulates multiple genes in addition to those involved in iron regulation. In *E. coli* Fur regulates more than 90 genes involved in iron homeostasis, oxidative stress response, metabolism and virulence factors while in *Pseudomonas syringae* Fur has more than 300 genomic targets, including iron responsive genes and two component systems (Hantke, 2001, Butcher et al., 2011). In some Gram-positive bacteria the protein DtxR carries out the same function as Fur (Hantke, 2001).

Although they have similar structures, Fur has an additional N-terminal helix involved in specific DNA-binding (Pohl et al., 2003).

FecI from *E. coli* and PvdS from *P. aeruginosa* are profoundly researched ECF σ factors involved in iron acquisition, thus serving as conventional models to study components of this type of signal transduction system.

ECF σ factor mediated iron uptake is carried out in *E. coli* via the '**Fec**' system that uses diferric-dicitrate as an iron source. The Fec system illustrated in Figure 1.9B consists of regulatory components that correspond to the σ and anti- σ factor pair FecI/FecR and a ferric citrate transport system that is composed of 5 proteins encoded by genes *fecABCDE* (Braun et al., 2005). FecA, an outer membrane protein that recognizes diferric dicitrate, is composed of a β -barrel with a plug spanning the outer membrane. FecBCDE are components of the ABC transporter required for ferric citrate transport across the cytoplasmic membrane (Braun, 2003). FecB functions as the periplasmic binding protein, FecC and -D as membrane spanning permeases and FecE as the inner membrane ATPase. *fecIR* are located upstream of the genes encoding *fecABCDE*. Furthermore, Fur regulates the *fecI* and *fecA* promoters depending on the availability of iron (Braun et al., 2003). Under iron-replete conditions, Fur associated with Fe^{2+} inhibits the transcription of *fecIR* and *fecABCDE*. Conversely, under iron starvation conditions, it dissociates from these promoters, allowing the transcription of these genes.

Crystal structure studies reveal that up to 10 residues within the extra-cellular loops of FecA bind to external diferric dicitrate causing conformational changes in the structure of FecA. Two loops in its β -barrel structure shift and thus lock the ferric citrate (Ferguson et al., 2002). As with all siderophore receptors in Gram-negative bacteria, FecA contains a 'TonB box' on its periplasmically oriented N-terminal domain which interacts with the C-terminal region of TonB (Braun and Mahren, 2005). This further causes major conformational changes allowing the ferric citrate to pass through the β -barrel due to movement of the 'plug/hatch' of FecA, straight into the periplasm. FecA not only transports

ferric citrate but is also required for initiation of transcription of *fecABCDE* (Ferguson et al., 2002). Some TBDRs such as FecA in *E. coli* have an additional domain of the N-terminus which is used in signalling (Ferguson et al., 2002). Such receptors are termed TonB dependent transducers (TBDTs). This additional domain located at the N-terminus of FecA interacts with the C-terminal periplasmic domain of the cytoplasmic membrane anchored protein FecR, which serves as the anti- σ factor for FecI, and passes a signal to FecR when bound to ferric citrate. Conformational changes as well as energy derived from the protonmotive force transduced by the TonB-ExbB-ExbD complex is required for this signal transduction. Six leucine residues (leucine zipper) in the C-terminal domain (CTD) of FecR are highly conserved in FecR-like regulatory proteins and are critical for this interaction. The cytoplasmically located N-terminal domain (NTD) of FecR (residues 1-85) interacts with the CTD (σ_4) of FecI. This region contains three tryptophan residues which are conserved across different anti- σ factor homologues in the iron-starvation class and are critical for the FecI-FecR interaction (Enz et al., 2000, Stiefel et al., 2001, Mahren et al., 2002).

FecR is proposed to activate FecI under inducing conditions. The N-terminal domain of FecR is reported to be bound to FecI while the latter interacts with RNAP by formation of a tertiary FecR1-85-FecI- β' 1-317 complex allowing FecI to interact with RNAP with higher affinity (Braun, 2003). Also, FecI is completely inactive in the absence of FecR and it is suggested that FecR might act as a chaperone to protect it from degradation. Due to this requirement to activate FecI, FecR is now termed ' σ factor regulator' rather than an ' σ ' factor (Braun, 2003, Braun et al., 2003, Brooks and Buchanan, 2008).

There are multiple ECF σ factors of *P. aeruginosa* that are involved in iron transport (Llamas et al., 2014). These are sometimes termed as cell surface signalling systems (CSS) (Llamas et al., 2009). One of these systems, consists of FoxA/I/R (similar to FecA/I/R of *E. coli*) utilises the xenosiderophore ferrioxamine. It has been shown that the anti- σ factor FoxR undergoes auto-proteolysis by a N-O acyl rearrangement for the activation of FoxI under inducing conditions (Bastiaansen et al., 2015b).

One of the most extensively studied iron-uptake ECF σ factor of *P. aeruginosa* is **PvdS**, which mediates iron transport through the endogenously produced siderophore pyoverdine. In this system, the siderophore pyoverdine bound to iron is recognized by the TBDT FpvA, in a TonB-dependent manner, which then passes a signal to FpvR, the anti- σ factor (Potvin et al., 2008). FpvR plays strictly an inhibitory role and in addition, controls the activity of PvdS and another ECF σ factor, FpvI. The anti- σ factor encoding gene *fpvR* and σ factor encoding genes *fpvI* and *pvdS* are not co-transcribed. Although *fpvR* and *fpvI* are located next to each other, they are in divergent orientations. PvdS is responsible for transcription of at least 26 genes involved in pyoverdine biosynthesis and utilisation, including *ptxR*, as well as virulence associated proteins such as exotoxin A and PrpL (Leoni et al., 2000). When PvdS forms a stable complex with RNA polymerase it recognises the consensus sequence TAAAT-N16/17-CGT at the promoter regions to initiate transcription (Ochsner et al., 2002). The TAAAT signature sequence present 33 bases upstream relative to the transcription start site forms the iron starvation (IS) box critical for promoter utilisation. PtxR is a protein that regulates transcription of the *pvc* genes which are involved in biosynthesis of the chromophore component of pyoverdine that gives it its characteristic property to fluoresce (Ravel and Cornelis, 2003, Visca et al., 2007, Cornelis et al., 2009). FpvI, is responsible for initiation of transcription of *fpvA*. Finally, the master regulator Fur negatively regulates all the above genes except *fpvA* under non-inducing conditions (Ochsner and Vasil, 1996). As a result, the combination of switching on pyoverdine biosynthesis and transport genes allows an elegant mechanism for *P. aeruginosa* to capture iron by employing two of its ECF σ factors as outlined in Figure 1.9A (Cornelis et al., 2009, Llamas et al., 2014). Other than iron regulation, *pvdS* expression is influenced by multiple factors such as those involved in sulphur regulation, oxidative stress response, biofilm factors, secreted factors, etc. (reviewed in (Llamas et al., 2014). Since these are beyond the scope of this subject they are not discussed here.

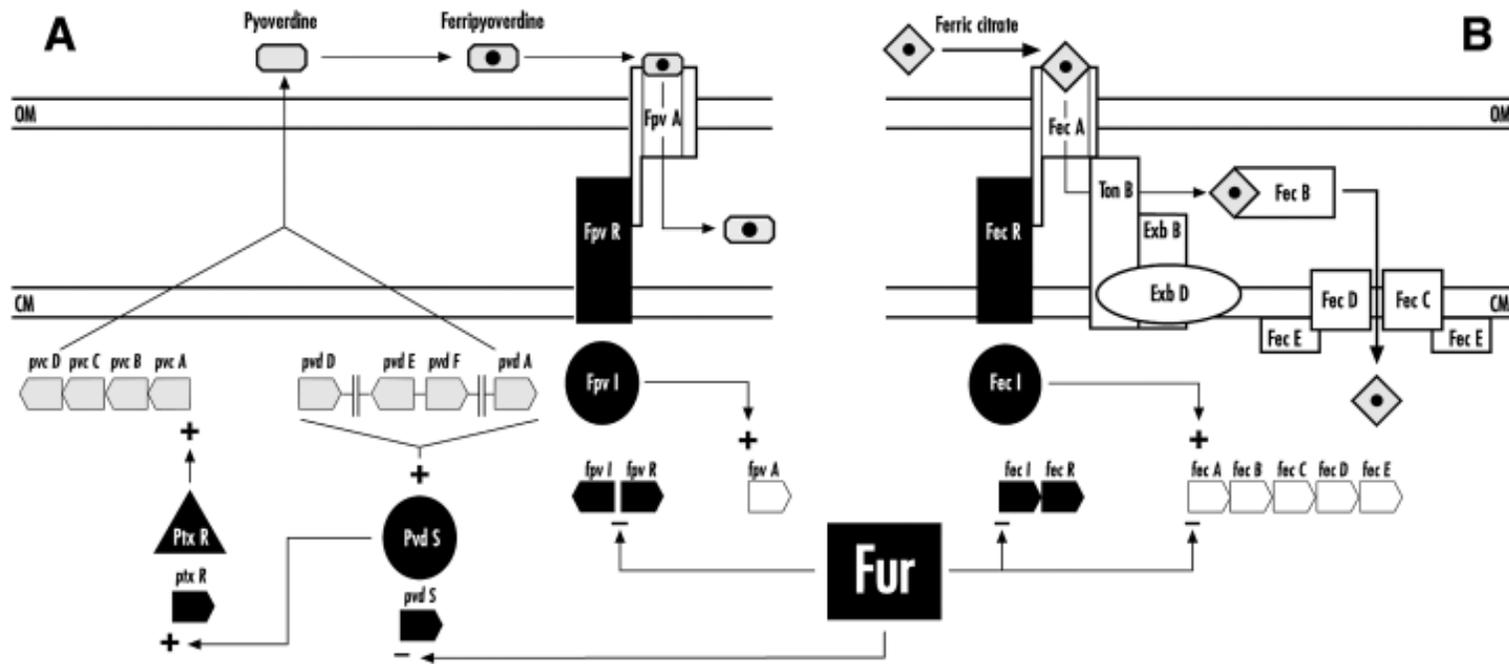


Figure 1.9. Summary of the Pvd and Fec iron acquisition systems

(A) In response to a signal transduced by FpvA the TonB-dependent transducer FpvR regulates two σ factors, PvdS and FpvI, involved in switching on genes involved in pyoverdine biosynthesis and uptake. Fur regulates the transcription of *fpvI*, *fpvR* and *pvdS*. (B) The signalling cascade to switch on the *fecABCDE* genes occurs through recognition of ferric citrate by FecA passing the signal in a TonB-dependent mechanism to FecR thereby causing FecI to associate with the RNA polymerase core enzyme. FecR is said to be an activator for FecI (not shown in figure, see text for details). Fur negatively regulates *fecI* and *fecR*. The proteins and genes shown in the figure are not drawn to scale. Reprinted by permission from John Wiley and Sons: Molecular Microbiology, (Visca et al., 2002) copyright 2016.

1.6 The *Burkholderia cepacia* complex (*Bcc*)

The *Burkholderia cepacia* complex (*Bcc*) are a group of approximately 20 genetically related but phenotypically distinct species belonging to the *Burkholderia* genus (Sousa et al., 2017). As such they are Gram-negative, rod-shaped non-spore forming bacteria of the β -proteobacterium class. The 'type' species *Burkholderia cepacia*, initially classified with the *Pseudomonads* as *P. cepacia*, was identified as the cause of 'sour skin' in onions (onion rot disease) and later classified into the new genus termed *Burkholderia* based on its phenotypic characteristics, 16S RNA sequences, fatty acid composition, etc. (Burkholder, 1950, Yabuuchi et al., 1992). The criteria for taxonomy includes *recA* sequencing, DNA-DNA hybridization, biochemical profiling, etc. for this genus currently consisting of approximately 100 species distributed within three clades (Depoorter et al., 2016). More recently, a new genus called *Paraburkholderia* has been assigned for species not associated with human infections (Dobritsa and Samadpour, 2016). Taxonomy of the *Bcc* is shown in Figure 1.10. Members of the *Bcc* have relatively large, GC-rich (>66%) genomes of approximately 7 - 9 Mb consisting of three chromosomes and one plasmid and have relatively high genome plasticity (Holden et al., 2009). They are mainly known to be opportunistic pathogens in humans, animals and plants, and have a high innate resistance to antibiotics and disinfectants, making them difficult to kill (Loutet and Valvano, 2010, Sass et al., 2011). Moreover, they are nutritionally versatile and can utilize a variety of compounds including penicillin G as carbon sources (Beckman and Lessie, 1979). However, they are also beneficial to the environment since they can degrade toxic compounds and can biosynthesize antimicrobials thus having wide biotechnological and bio-remedial applications [reviewed in (Depoorter et al., 2016)]. The niche occupied by the *Bcc* and the virulence factors they harbour are described below, especially focusing on *B. cenocepacia*, the organism that is the host for the ECF σ factors investigated in this study.

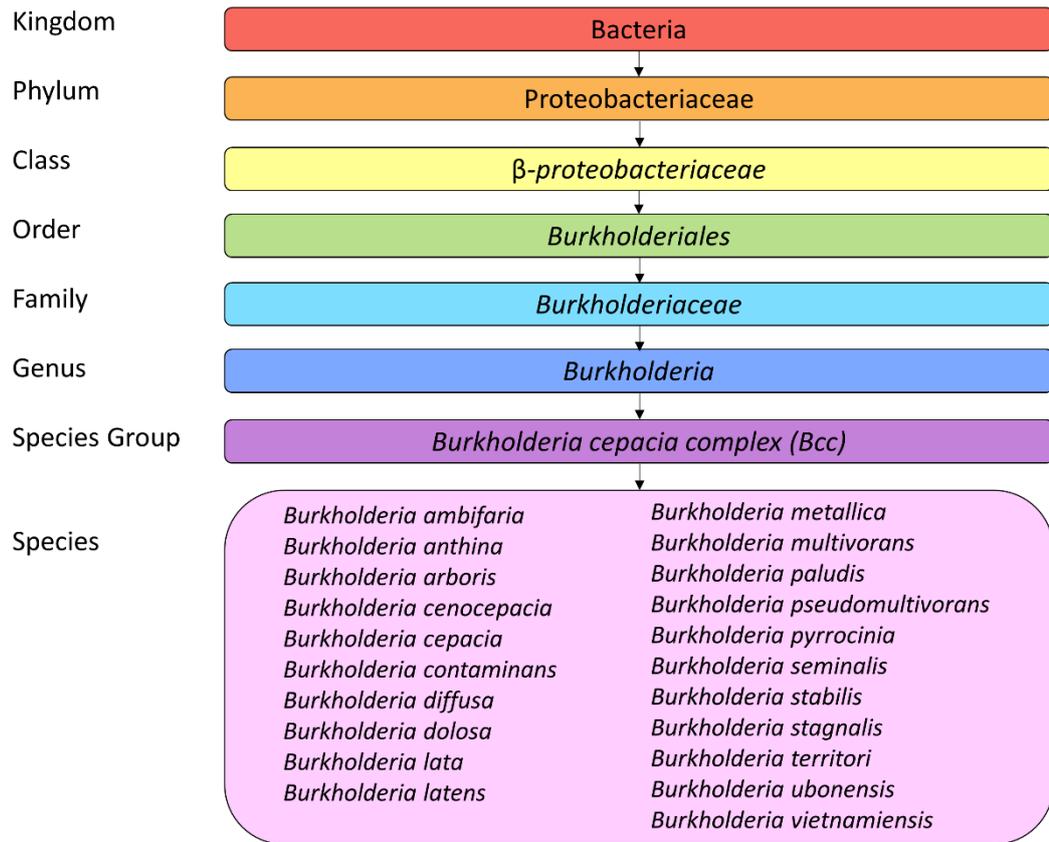


Figure 1.10. Taxonomic classification of the Bcc

The taxonomic hierarchical organisation of the currently known 21 species forming the *Burkholderia cepacia* complex is shown.

1.6.1 Interaction with the environment/host

Members of the *Bcc* are naturally found in soil, water and in association with plants, particularly in the rhizosphere, forming a symbiotic relationship, as some of them are capable of fixing nitrogen (Coenye and Vandamme, 2003). In humans, they do not cause infections in healthy individuals although maybe carried by them. These bacteria are associated with lung infections in immunocompromised patients. Cystic fibrosis (CF) patients have defective alleles of the cystic fibrosis transmembrane conductance regulator (CFTR) encoding gene. CFTR which is involved in secretion of fluids, sweat and mucous, when defective, causes these secretions to thicken and increases susceptibility to infections (Isles et al., 1984, Goldmann and Klinger, 1986). *Bcc* members are also

present in infections in individuals with the genetic disorder chronic granulomatous disease where the immune system of these individuals have a reduced capacity of reactive oxygen species production to kill pathogens (Johnston, 2001). *B. cenocepacia* and *B. multivorans* are the predominant *Burkholderia* species (70% of total *Bcc* infections) in cystic fibrosis infections (LiPuma, 2010). The *B. cenocepacia* epidemic strain ET-12 was reported to be more dominant in the UK and Canada (Johnson et al., 1994) while the PHDC strain was more dominant in Mid-west USA (Pitt et al., 1996). *Bcc* infections in CF patients are mainly caused by hospital based transmissions including patient-patient transmission as well as directly from the environment (Horsley et al., 2016). In some cases *B. cenocepacia* significantly contributes to morbidity and mortality resulting from a rapid decline of the immune system known as 'cepacia' syndrome which is characterized by necrotizing pneumonia and septicaemia (Isles et al., 1984). An abundant amount of mucus formation in lungs is a hallmark feature of cystic fibrosis in which opportunistic pathogens from different genera can thrive. Bacteria belonging to the *Pseudomonas* species can form mixed biofilms with the members of the *Bcc* (Tomlin et al., 2001). However, another study suggests that *Bcc* bacteria inhibit the formation of biofilms by *Pseudomonas aeruginosa in vivo* and are prevalent in late stage CF more often as single cells in macrophages rather than as biofilms (Schwab et al., 2014). During the course of lung infections, these bacteria undergo multiple mutations commonly in genes required for iron uptake, cell-membrane constituents and antibiotic resistance, although these mutations do not tend to fix in the population within the same patient (Lieberman et al., 2014). Patient segregation, strict use of disinfectants for equipment and treatments consisting of a combination of antibiotics are employed to keep transmission under control. But as the *Bcc* bacteria can mutate to develop antibiotic resistance in addition to their intrinsic antibiotic resistance and recurrent infections by different strains, there is no straightforward treatment in the current clinical practice (Horsley et al., 2016). Therefore, it is important to study their molecular mechanisms of infection and survival within the host which yet remain to be fully understood. Although several virulence determinants have been identified, most have been experimentally

verified in only one or two host models and the significance of each of these factors is found to be strain dependent (Loutet and Valvano, 2010).

1.6.2 Virulence determinants

Despite the host immune response imposing oxidative stress and the presence of iron starvation conditions, pathogens use multiple virulence factors to adapt to these conditions for colonisation and infection. One of the virulence determinants includes formation of **biofilms** that are multi-cell communities on solid surfaces. Biofilm formation has been reported to be required for resistance to antibiotics and for responding to host immune cells (Caraher et al., 2007, Fazli et al., 2013, Van Acker et al., 2013, Messiaen et al., 2014, Murphy and Caraher, 2015). Transcriptomic based studies revealed that multiple genes involved in biofilm formation were upregulated during the infection process (Peeters et al., 2010, Van Acker et al., 2014, Sass et al., 2015). However, it has also been shown that in some cases biofilm formation by bacteria does not necessarily improve the ability to respond to antibiotics as compared to non-biofilm forming bacteria (Peeters et al., 2009). Production of **exopolysaccharide (EPS)** that gives the appearance of mucoidy to bacterial colonies is involved in biofilm formation (Cunha et al., 2004, Herasimenka et al., 2007). The EPS cepacian has been reported to be involved in countering the ROS induced stress generated by neutrophils, metal ion stress and desiccation, as well as increased persistence in infections (Conway et al., 2004, Bylund et al., 2006, Sousa et al., 2007, Ferreira et al., 2010). However, EPS producing strains (mucoidy) are not necessarily associated with virulence as some clinical isolates of *B. cenocepacia* are non-mucoid (Zlosnik et al., 2008).

A robust **oxidative stress response** is likely to be a key virulence determinant as production of reactive oxygen species (ROS) is prevalent in the airway epithelia of cystic fibrosis patients (Galli et al., 2012). *B. cenocepacia* harbours at least one monofunctional catalase, two bifunctional catalase-peroxidases and a superoxide dismutase (Holden et al., 2009). A *B. cenocepacia* superoxide dismutase (*sodC*) mutant was reported to be more susceptible to macrophage mediated killing and this was shown to be complementable *in trans* (Keith and Valvano, 2007). The two bifunctional catalase-

peroxidase enzymes KatA and KatB have been experimentally shown to exhibit catalase activity (Lefebvre et al., 2005, Charalabous et al., 2007). In addition to these, a pyomelanin pigment in *B. cenocepacia* C2454 was shown to be protective against oxidative stress (Keith et al., 2007).

High affinity **iron uptake systems** allow *Bcc* pathogen to grow in low iron conditions present inside the host's lungs thus indirectly contributing to virulence (Mahenthalingam et al., 2005). Members of the *Bcc* produce combinations of the siderophores ornibactin, pyochelin, cepabactin and cepaciachelin (Darling et al., 1998, Thomas, 2007). Based on infection studies ornibactin and pyochelin are reported to be significant players in the pathogenicity of *B. cenocepacia*, where ornibactin was found to be more important (Visser et al., 2004, Uehlinger et al., 2009). *B. cenocepacia* can also utilise xenosiderophores ferrichrome and ferrioxamine B as well as alternate iron sources ferritin and haem (Whitby et al., 2006, Thomas, 2007). *B. cenocepacia* can competitively capture iron by using a combination of siderophore based uptake and alternate iron uptake (Tyrrell et al., 2015). Moreover, *B. cenocepacia* is capable of surviving under iron limiting conditions through a siderophore-independent mechanism of iron acquisition, involving the FtrABCD system that has also been reported in *Bordetella* and *Brucella* species (Brickman and Armstrong, 2012, Ahmed and Holmstrom, 2014, Mathew et al., 2014).

Quorum sensing, as originally proposed, is a mechanism used for intercellular communication where the external concentration of an autoinducer secreted by the cell itself, when beyond a specific threshold, causes the regulation of certain genes by a receptor. An alternative hypothesis posits that the purpose of QS is for diffusion sensing (Redfield, 2002). Members of the *Bcc* possess the CepRI system that uses N-octanoylhomoserine lactone or N-hexanoylhomoserine lactone as the inducer produced by CepI, the acyl homoserine lactose (AHL) synthase, and CepR acting as the transcriptional regulator (Lewenza et al., 1999, Lutter et al., 2001). CepR is reported to control the expression of a variety of genes involved in biofilms, motility, siderophores, toxins, proteases, lipases, etc. and was shown to be essential for virulence based on experiments in multiple plant and animal models (Sokol et al., 2003, Venturi et al., 2004,

Eberl, 2006, Uehlinger et al., 2009, Subramoni and Sokol, 2012, Suppiger et al., 2013). In addition, the *Burkholderia* genomic island 11 harbours genes encoding CciR and CciI which function similar to CepR and CepI, respectively (Malott et al., 2005). CepR2 is an 'orphan' regulator as it does not have an associated CepI2 and is antagonised by octanoylhomoserine lactone (OHL) (Malott et al., 2009, Ryan et al., 2013). It is also proposed that cis-2-dodecenoic acid/*Burkholderia* diffusible signal factor (BDSF) acts as a QS mechanism (Deng et al., 2009, Deng et al., 2010).

Detecting stress and regulating specific responsive genes allows the bacterium to adapt for survival within host tissues. **Alternative σ factors** often control the expression of virulence determinants thus being important for pathogenicity themselves. Inactivation of *rpoE* (encoding σ^E) in *B. cenocepacia* K56-2 caused defects in the outer-membrane protein profile, reduced the ability to delay phagolysosomal fusion within macrophages in addition to reducing the capacity of the bacterial cells to survive under elevated temperatures and high osmotic environments (Flannagan and Valvano, 2008). These effects were complementable *in trans*. Another alternative σ factor, RpoN (σ^{54}), also showed a role in delaying phagolysosomal fusion in macrophages in addition to biofilm formation and motility (Saldias et al., 2008). In *B. cenocepacia* H1111, σ^{54} was shown to regulate genes involved in nitrogen starvation as well as upregulate EPS production and increase the accumulation of the polymer polyhydroxybutyrate (Lardi et al., 2015). A transcriptomics study on biofilm-grown *B. cenocepacia* J2315 exposed to oxidative stress by hydrogen peroxide and sodium hypochlorite showed that the alternative σ factors BCAL3478 (ECF σ PrtI-like) and BCAL0787 (σ^{32}) and two iron-starvation class of ECF σ factors BCAL1688 (OrbS) and BCAL1369 (FecI-like) were upregulated (Peeters et al., 2010).

B. cenocepacia is proposed to cause a delay in macrophages to clear pathogens by surviving within vacuoles inside macrophages thus forming *B. cenocepacia* containing vacuoles (BcCV) and delaying phagosomal NADPH oxidase complex recruitment (Keith et al., 2009).

Recently, an effector of the type-six secretion system has been identified, TecA, that has deamidation activity that disrupts host actin cytoskeleton and through

modification of Rho-GTPases activates Pyrin inflammasome leading to host cell death (Aubert et al., 2016).

Other potential virulence factors of the *Bcc* include flagella (motility) and pili, lipopolysaccharide (cell surface presenting antigens), protein secretion systems, secretion of hydrolytic enzymes (zinc metalloproteases and lipases), toxin-anti-toxin systems and haemolysin production amongst others. Studies on the experimental evidence on the molecular regulation of the known *Bcc* virulence factors and their effect on the host-pathogen relationship during infections are described and reviewed extensively in (Mahenthiralingam et al., 2005, Loutet and Valvano, 2010, Drevinek and Mahenthiralingam, 2010, Sousa et al., 2017).

1.6.3 ECF σ factors of *B. cenocepacia*

Burkholderia cenocepacia has a genome of 8.06 Mb consisting of three circular chromosomes and a plasmid. Chromosome 2, an accessory chromosome, encodes one of the two copies of the primary σ factor which is an unusual phenomenon (Holden et al., 2009). *B. cenocepacia* is predicted to encode 13 ECF σ factors, only three of which have been investigated (Menard et al., 2007, M. Thomas, unpublished). One of them is homologous to RpoE (σ^E) of *E. coli*, the cell-envelope stress response σ factor, but it might also play a role in contributing to the pathogenicity of *B. cenocepacia* (Flannagan and Valvano, 2008). The other two ECF σ factors are OrbS (BCAL1688 in *B. cenocepacia* J2315) and FlrS (BCAL1369 in *B. cenocepacia* J2315), both belonging to the iron starvation type of σ factors. FlrS is predicted to be regulated by the anti- σ factor, FlrR, and regulates transcription of a gene encoding a putative TBDT, FlrA. However, unlike the Fec system in *E. coli*, ABC transporter genes are absent downstream of *flrA*. Importantly, the absence of genes that encode synthesis of a siderophore that might be transported by this system suggests utilization of an external siderophore (xenosiderophore) by this system (M. Thomas, unpublished). The other iron starvation σ factor, OrbS, was subject to a previous investigation (Agnoli et al., 2006). Four currently uncharacterised *B. cenocepacia* ECF σ factors together with OrbS are the subject of study in this project and are described below.

1.6.3.1 I35_RS16290 (PrtI)

I35_RS16290 in *B. cenocepacia* H111 is a putative ECF σ factor based on its amino acid sequence that includes regions homologous to σ_2 and σ_4 of other ECF σ factors. I35_RS16290 belongs to the ECF26 group based on the current classification of ECF σ factors (Staron et al., 2009). ECF26 σ factors are mainly found in Proteobacteria and some in Acidobacteria. ECF26 σ factors encoding genes tend to appear with a cognate anti- σ factor encoding gene termed *prtR* in the current literature (Burger et al., 2000). Defined by their genomic context, the *ecf26* and the *anti-ecf26* pair are usually present in the vicinity of sets of genes encoding the following products – (i) cytochrome c oxidase, metallophosphoesterase (ii) cytochrome, catalase (iii) lipoprotein (Staron et al., 2009, Stockwell et al., 2012). The role of one the σ factors from this group, SigE of *Starkeya novella*, is implicated in thiosulphate oxidation (Kappler et al., 2001). The role of another ECF26 σ factor, PrtI, has been broadly explored in *Pseudomonas* species, particularly in *P. fluorescens*, a psychotrophic strain involved in food spoilage where extracellular enzymes play a key role (Liao and McCallus, 1998). As I35_RS16290 has >60% amino acid similarity with *P. fluorescens* PrtI, I35_RS16290 is referred to as PrtI. The first published study on *P. fluorescens* LS107d2 PrtI demonstrated that a mutation in the gene encoding this σ factor resulted in loss of production of an extracellular metalloprotease AprX at the elevated temperature of 29°C compared to the WT (Burger et al., 2000). Since this gene seemed to be involved in regulating protease production it was termed ‘Prt’ (Prt Regulator). Surprisingly, inactivation of *prtR* also resulted in loss of protease production at 29°C as compared to WT. Further, complementing each of the *prtI* and *prtR* mutants with WT *prtI* and *prtR in trans* restored the WT phenotype. Moreover, ablation of these genes did not have an effect on protease production at 23°C, the ambient growth temperature (Burger et al., 2000).

Contrastingly, Okrent and colleagues observed in *P. fluorescens* WH6 that while the *prtR* mutant showed a loss of protease function, the *prtI* mutant did not (Okrent et al., 2014). Further, complementing the *prtR* mutant with WT *prtR in trans* restored protease production but when WT *prtI* was introduced *in trans* into the *prtI* mutant a loss in protease production was obtained. This suggested that PrtI may negatively

regulate protease production while PrtR, by virtue of its inhibitory effect on PrtI regulates it positively. However, the exact regulatory mechanism by which PrtIR exert their effect on AprX production has not yet been established. Negative regulation by the σ factor suggests an indirect effect on *aprX* expression. PrtI was also suggested to negatively regulate production of the secondary metabolite germination-arrest factor (GAF) in *P. fluorescens*. However, the exact mechanism of regulation of the GAF biosynthetic genes by PrtIR is not known. Real time qPCR analysis with a *prrR* mutant showed that PrtR controls transcription of two genes encoding an aminotransferase and a formyltransferase which are involved in the biosynthesis of GAF (Halgren et al., 2013). Inactivation of *prrI*, however, did not have any effect on the transcription of these two genes and did not affect GAF production. The loss of GAF production in the *prrR* mutant was complemented by WT *prrR*. When the mutant *prrI* allele was introduced into the *prrR* mutant to give a double (*prrIR*) mutant, it restored GAF production. Nonetheless, when WT *prrI* was introduced into the double mutant *in trans*, the mutant failed to produce GAF (i.e. it behaved as a *prrR* mutant) and conversely when *prrR* was introduced into the double mutant GAF production was restored (i.e. it behaved as a *prrI* mutant). Also, in the double (*prrIR*) mutant levels of transcription of these two genes were the same as the WT levels but transcript levels were suppressed when WT *prrI* was introduced (Kimbrel et al., 2010, Okrent et al., 2014). These results suggest that PrtI is inhibitory to GAF production and is not a direct regulation of genes required for GAF biosynthesis.

In *Pseudomonas* sp. PCL1171, PrtR positively regulated extracellular polysaccharide (EPS) production thereby affecting colony morphology, while PrtI seemed to negatively regulate EPS production (van den Broek, 2005). In 2006, a study on *Pseudomonas entomophila* revealed that PrtI-PrtR regulate the *aprA-inh-aprDEF* operon. *aprA* is an orthologue of *aprX* from *P. fluorescens*. In this study, mutation in the *prrR* gene reduced pathogenicity of *P. entomophila* towards *Drosophila* as the protease AprA is proposed to degrade anti-microbial peptides produced by the host (Vodovar et al., 2006, Liehl et al., 2006).

In *P. fluorescens* HC1-07, PrtR was found to have a role in regulation of protease production as seen previously in *P. fluorescens* LS107d2. However, it was found to be temperature independent. In addition, it was shown to have a role in regulation of CLP (cyclic lipopeptide) production and surface motility (Yang et al., 2014). The *prtR* mutant produced significantly less CLP compared to WT, and was therefore defective in forming biofilms and motility as CLP is required for both (Raaijmakers et al., 2006, D'Aes et al., 2010). These effects were complementable when *prtR* was introduced *in trans* (Yang et al., 2014). However, this study does not elucidate the mechanism of regulation by PrtR.

PrtR was also reported to be involved in regulation of cyclic lipopeptide (massetolide) biosynthesis and extracellular protease production in *P. fluorescens* SS01. RT-qPCR studies showed a decrease in transcription of 3 massetolide biosynthesis genes (*massABC*) and 2 *luxR*-type regulatory genes (*massAR*, *massBCR*) in a *prtR* mutant (Song et al., 2014).

Burger and colleagues proposed that (i) PrtR is a transmembrane activator rather than an inhibitory anti- σ factor, (ii) elevated temperature might be a stress signal recognised by PrtR and (iii) PrtR may regulate multiple σ factors, although none of these proposals have been experimentally evidenced (Burger et al., 2000).

Overall, studies so far described above have shown that mutating this σ /anti- σ factor pair affected one or more of the following phenotypes: temperature-dependent or -independent protease (*aprX/aprA*) production, EPS production, GAF production, cyclic lipopeptide production, biofilm formation, surface motility and lipase production. Moreover, PrtI appears to exert a negative, and therefore possibly indirect effect on their expression. It was suggested that enzymes like proteases and lipases play a role in phytopathogenesis (Burger et al., 2000). GAF from *P. fluorescens* has anti-microbial and herbicidal properties (Halgren et al., 2013). EPS production (mucoidy) contributes to biofilm formation while cyclic lipopeptide (CLP)/massetolide has anti-microbial properties and has functions in surface motility and biofilm formation (Raaijmakers et al., 2006, de Bruijn et al., 2007, D'Aes et al., 2010). Biofilm formation and surface motility both form important aspects of bacterial infection (Yang et al., 2014, Olsen, 2015). The metalloprotease AprA is a virulence factor, providing protection against the host

immune response (Liehl et al., 2006, Vallet-Gely et al., 2010). Since PrtI regulates genes that affect production of such proteins involved in pathogenesis it can be speculated that PrtI might indirectly be a virulence determinant.

PrtR functions along with the two component system 'GacA/GacS' in its roles in regulation of *aprA* in *P. entomophila* to escape anti-microbial peptide activity (Liehl et al., 2006) and regulation of lipopeptide biosynthesis in *P. fluorescens* SS01 (Song et al., 2014). In addition to GacA/GacS, PrtI-PrtR also co-functions with the ECF σ factor RpoS in the regulation of 'phase changes', i.e. colony morphology, in *Pseudomonas* PCL1171 (van den Broek, 2005).

Despite these extensive observations, none of the above studies outlines the exact mechanism as to how PrtI and PrtR function together to cause the phenotypic changes that were observed. Even though some qPCR studies show that transcription of certain genes changed based on mutating *prtI/prtR*, additional experiments are required to address the questions of what promoters are actually targeted by PrtI and how it functions at the molecular level.

Okrent and colleagues proposed that the *prtR* promoter lies 108 bps upstream of *prtR*, within the *prtI* gene while the *prtI* promoter lies 163 bps upstream of its ATG start codon. It was also suggested that *prtI* and *prtR* can be transcribed independently and also as a single transcript in *P. fluorescens* WH6, suggesting that there might be more than one way of regulating *prtI/R* transcription initiation (Okrent et al., 2014).

It was suggested that the *aprX* promoter might be regulated by PrtI although no experimental evidence was provided to support this. The sequence shown in Figure 1.11 was suggested to be the *aprX* promoter (Burger et al., 2000). However, this was not experimentally tested. In addition, even though this might well be the *aprX* promoter with its defined promoter elements there is no published evidence that PrtI recognises it and directs transcription from this sequence.

P_{aprX}: A **CTGTAGAC** GAAAGATTTCA **AGTTGTGTCAGTTT** TCTTA **A** CGCCTTCAAAA

Figure 1.11. Predicted PrtI-dependent promoter P_{aprX}

The predicted -35 and -10 promoter elements are highlighted in yellow and blue, respectively. The predicted transcription start site (+1) 'A' is highlighted in green (Burger et al., 2000).

While the direct or indirect role of PrtI on expression of certain genes has been ascertained by studying the phenotypic effects and transcription of such genes upon inactivation of *prtI* or *prtR*, so far, no studies have been reported where the location or sequence of PrtI-dependent promoters have been experimentally determined. Hence, the exact genes targeted by PrtI remain unresolved.

Perhaps, PrtI regulates one or more regulatory genes that are involved in different processes in different species and strains of *Pseudomonas*, explaining the many different observed phenotypes. In contrast, PrtI may have different roles in different species/strains, although this is difficult to explain from a bioinformatics and evolutionary perspective. Further, the regulation of PrtI by PrtR needs to be specifically investigated to identify the exact mode of regulation. Additionally, *aprX/aprA* –like genes are not present in *Burkholderia*. Therefore, PrtI may play a different role in this genus. Since there is no protease in the proximity of *prtI/R* in *B. cenocepacia*, perhaps they are involved in the regulation of protein families other than proteases (Martinez-Bueno et al., 2002, Menard et al., 2007).

1.6.3.2 I35_RS16330 (Ecf41_{Bc1}) and I35_RS29600 (Ecf41_{Bc2})

Based on their amino acid sequences, I35_RS16330 and I35_RS29600 from *B. cenocepacia* H111 appear to belong to the ECF41 group of σ factors and therefore are addressed in this thesis as 'Ecf41_{Bc1}' and 'Ecf41_{Bc2}', respectively. ECF41 σ factors do not seem to be regulated by a conventional anti- σ factor. Also, they contain a C-terminal extension of about 100 amino acids that is not present in other ECF σ factors. It has been reported that ECF41 σ factors are distributed across 10 phyla and are predominantly present in Actinobacteria and Proteobacteria. It has also been observed that genes encoding ECF41 σ factors are often present in the vicinity of genes that encode one or more proteins belonging to the carboxymuconolactone decarboxylase, oxidoreductase or epimerase/NmrA family of proteins (COE) (Staron et al., 2009). Carboxymuconolactone decarboxylases are reported to be involved in protection against oxidative damage (Bryk et al., 2002, Koshkin et al., 2003). Oxidoreductases are a large group of enzymes that are NAD⁺-dependent and include members with anti-

oxidant activity. Epimerases are isomerising enzymes while NmrA is similar to reductases and dehydrogenases and is involved in nitrogen metabolism (Wecke et al., 2012). Out of the 373 ECF σ factor genes analysed, approximately 30% did not appear to be associated with genes encoding the above proteins. Instead they were located close to genes encoding a hypothetical protein or cupin (dioxygenase) (Wecke et al., 2012).

One ECF41 member, SigJ from *Mycobacterium tuberculosis*, has been proposed to be involved in oxidative stress response where the *sigJ* mutant was found to be more sensitive to hydrogen peroxide compared to WT and this phenotype was complementable *in trans* (Hu et al., 2004). However, this mutant was as competent as the WT in causing infections. The precise role of SigJ in oxidative stress tolerance is currently not known as the SigJ-dependent promoters have not been identified. The ECF41 σ factors from *Bacillus licheniformis* (*Ecf41_{Bli}*) and *Rhodobacter sphaeroides* (*Ecf41_{Rsp}*) did not seem to have a role in oxidative stress response based on mutational analysis of their respective genes (Wecke et al., 2012). A phenotypic screen that tested over 900 conditions for these σ factor deletion mutants did not reveal any phenotypic differences compared to WT and the role of these σ factors still remains unknown. However, it was found they drive the transcription of the respective carboxymuconolactone decarboxylase encoding genes present next to them. DNA fragments (approximately 100 bp) containing the *Ecf41_{Bli}*-dependent promoter or the *Ecf41_{Rsp}*-dependent promoter were experimentally shown to be regulated by the respective σ factor. A manual bioinformatic analysis using WEBLOGO was carried out to predict that a TGTCACA motif formed the -35 promoter element followed by a spacer of 16 bp and a CGTC motif formed the -10 promoter element. A larger bioinformatic scan using a virtual footprint algorithm in other species involving a total of 285 sequences showed the conservation of these promoter motifs with a 16 ± 1 bp spacer present upstream of COE encoding genes and in some cases upstream of the ECF σ factor itself (Wecke et al., 2012).

Studies on the regulation of these two σ factors show that the C-terminal extension behaves as a fused on-board regulatory domain, where a DGGG motif within

the domain is critical for activity and the extension plays a dual role of an inhibitor and an activator. The exact molecular mechanism of this regulation is yet to be ascertained (Wecke et al., 2012).

1.6.3.3 I35_RS20320 (Ecf42_{Bc})

I35_RS29320 from *B. cenocepacia* H111 appears to belong to the ECF42 group of σ factors and is therefore addressed here as 'Ecf42_{Bc}'. ECF42 σ factors do not seem to be regulated by a conventional anti- σ factor and contain a C-terminal extension of about 200 amino acids that is not present in other ECF σ factors. Additionally, a tetra tricopeptide repeat (TPR) domain is present within the C-terminal extension (Staron et al., 2009). The TPR motif is present usually present as 3 to 16 tandem-repeats of 34 amino acids and is involved in playing different roles within cells (D'Andrea and Regan, 2003). TPR domains are sometimes associated with virulence in pathogens and might play a regulatory role through protein-protein interactions (Blatch and Lassel, 1999, Cervený et al., 2013). ECF42 σ factors are distributed across Actinobacteria, Cyanobacteria, Planctomycetes, Acidobacteria and Proteobacteria phyla (Staron et al., 2009). ECF42 σ factor encoding genes are often present near genes that belong to a family named 'YCII-related' (named after YCII domain) and their protein products are termed as 'PhnB' or 'DGPF' (Staron et al., 2009). These proteins have a highly conserved 'DGPF' motif which is a part of a fairly conserved 'DGPFAETKE' motif (Staron et al., 2009), M Thomas unpublished). The role of these genes is not yet known although it is speculated that they have an enzymatic function (Yeats et al., 2003). In case of *ecf42_{Bc}*, a gene encoding putative rubrerythrin is present downstream. Rubrerythrins belong to the oxidoreductase family of proteins and have been shown to have a protective role in oxidative stress response (Sztukowska et al., 2002).

A study on a ECF 42 σ factor termed 'ECF10' in *Pseudomonas putida* KT2440 reported it to be involved in biofilm formation and drug resistance (Tettmann et al., 2014). Deletion of the gene encoding ECF10 increased the ability of *P. putida* to form biofilms and upregulated the expression of *ttgA*, encoding a component of the TtgABC efflux pump involved in extrusion of antibiotics, thus increasing antibiotic resistance.

This implies that activating ECF10 would make *P. putida* more susceptible to antibiotics, but the specific role of the σ factor in this process was not explicitly clarified. The molecular mechanism for the activation of ECF42, nature of its target promoters as well as its exact function are as yet unknown. Moreover, the role of the TPR domain in the regulation of ECF42 activity and the function of PhnB-type protein in this system has not been explored.

1.6.3.4 OrbS

A well characterised ECF σ factor of *B. cenocepacia* is OrbS, belonging to the iron starvation class of alternative σ factors. It facilitates iron acquisition via the siderophore ornibactin. Ornibactin is similar to malleobactin from *Burkholderia pseudomallei* except that in the *orb* operon two genes encoding acyltransferases are present which are absent in the malleobactin biosynthesis gene cluster (Franke et al., 2013). In addition to a role in iron transport, ornibactin is reported to have a role in protecting cells against toxic metals such as aluminium, copper, zinc and lead, thereby serving as a defensive mechanism (Mathew et al., 2016). The *orb* operon consists of 14 genes shown in Figure 1.12A. OrbS regulates 3 promoters in response to iron starvation: P_{orbH} , P_{orbE} and P_{orbI} (Agnoli et al., 2006). RT-PCR analysis showed that genes *orbH* to *orbB* are transcribed from P_{orbH} , *orbI* to *orbL* are transcribed via P_{orbI} , and P_{orbE} is required for transcription of *orbE* only. P_{orbS} is σ^{70} -dependent and drives the same genes as P_{orbH} along with *orbS*. A systematic base substitution study of the OrbS-dependent promoter P_{orbH} revealed that TAAA (positions -30 to -33 with respect to TSS) and CGTC (positions -12 to -9 with respect to TSS) formed the -10 and -35 promoter elements separated by a 17 bp spacer that included a 9 bp GC-rich tract (Agnoli, 2007). However, these experiments have been carried out in *E. coli* and some unexpected contribution by bases at the A+G rich transcription start site were found which would not be expected to be engaged by OrbS. As the effect of mutating these bases is likely to reflect alterations in the core RNAP-promoter DNA interaction, the assays need to be repeated in a *B. cenocepacia* host strain where the effect of these substitutions can be examined in the presence of *B. cenocepacia* core RNAP.

The ferric-ornibactin complex is transported from the outer membrane to the periplasmic space via OrbA, a TBDR (Sokol et al., 2000). Since it lacks the N-terminal signalling domain it does not function as a TBDT that engages with an anti- σ factor. The ferric-ornibactin is proposed to be pumped into the cytoplasm through the ABC transporter composed of OrbB, OrbC and OrbD, following which iron is then released from the siderophore by reduction by the predicted reductase, OrbF. Biosynthesis of ornibactin is carried out by PvdA, PvdF, OrbG, OrbK and OrbL along with the nonribosomal peptide synthetases OrbI and OrbJ that are involved in the assembly of the ornibactin tetrapeptide backbone (Thomas, 2007). As there is no anti- σ factor for OrbS and OrbA is not a TBDT, the regulation of OrbS activity was proposed to be carried out exclusively by the conventional iron regulator Fur at the transcriptional level in response to iron concentrations (Agnoli et al., 2006). Currently, the exact mode of post-translational regulation of OrbS, if any, remains to be fully elucidated.

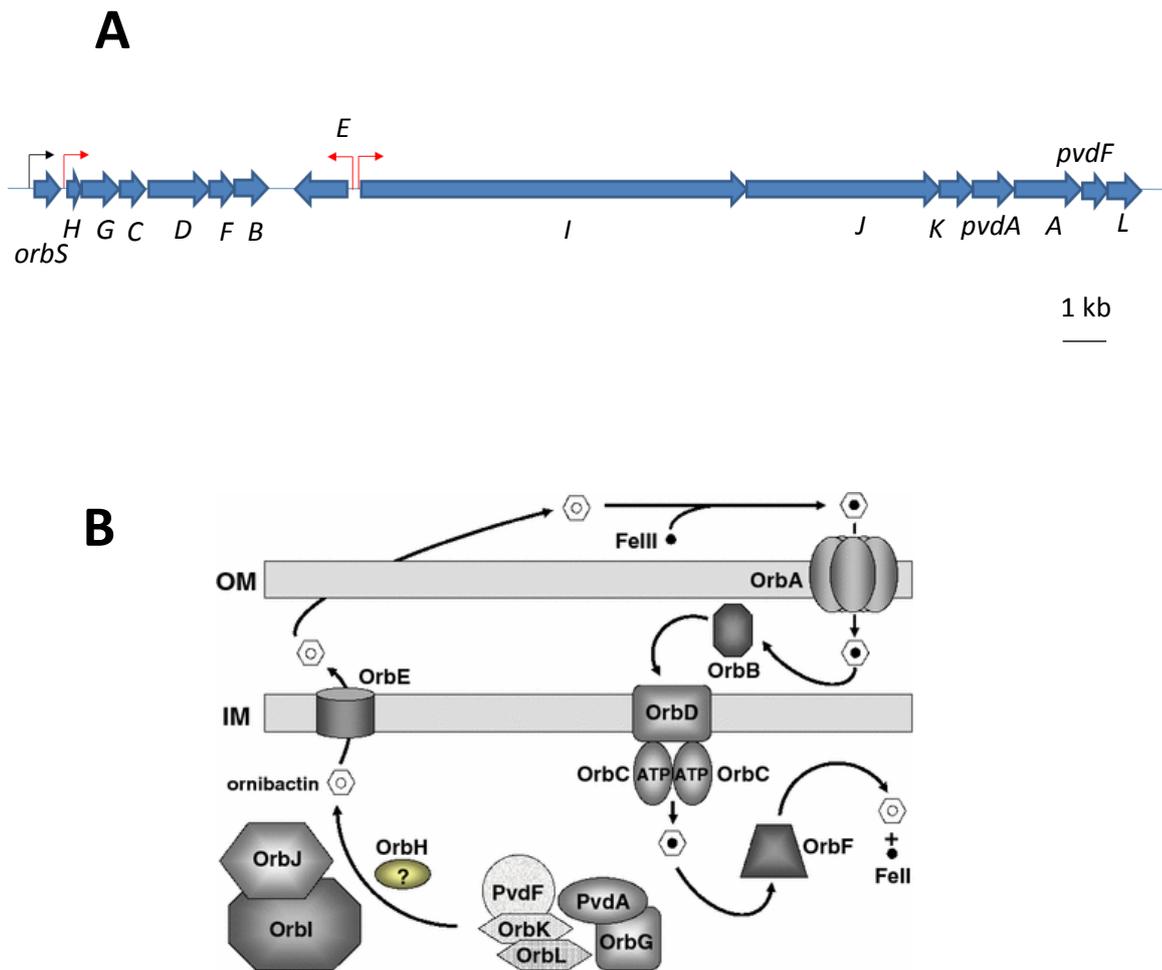


Figure 1.12. The Orb system in *B. cenocepacia*

(A) Gene map of the *orb* operon. Genes are shown as blue block arrows with the name underneath each gene. Transcription initiation from the σ^{70} -dependent promoter is shown as black bent arrow while that from the OrbS-dependent promoters are shown as red bent arrows where the direction of the arrow represents the direction of transcription. (B) Model for OrbS dependent iron acquisition through ornibactin. OrbJ, OrbI, PvdF, OrbK, OrbL, PvdA and OrbG are involved in the synthesis and assembly of ornibactin exported through OrbE. The ferric-ornibactin complex is transported through the TBDR OrbA into the periplasm. Through the ABC transporter composed of OrbB, OrbC, OrbD and the reductase OrbF, ferric siderophore is transported inside the cytoplasm and iron is dissociated in a form that can be biologically utilised. The function of OrbH is not known. Reprinted by permission from Springer: Biometals, (Thomas, 2007), copyright 2016.

1.7 Hypotheses and aims

Based on the current understanding of the above 5 ECF σ factors of *B. cenocepacia*, the following hypotheses and aims were developed. They have been sectioned below according to the chapters that describe the experiments carried out to address them.

Chapter 3:

Hypothesis - I35_RS16290 (PrtI) is an ECF σ factor of *B. cenocepacia* H111, which is released from its anti- σ factor I35_RS16295 (PrtR) in response to an appropriate extra-cytoplasmic stress signal, whereupon PrtI initiates transcription of genes which are required to cope with the extra-cytoplasmic stress, by recognising specific promoter sequences located upstream of such genes.

Aims -

- To study the role of PrtI in *B. cenocepacia* by (1) identifying the PrtI-dependent promoters to reveal its target genes and (2) studying the effects of deleting *prtI* or *prtR* on stress responses of *B. cenocepacia*.
- To study the specific DNA sequence requirements of PrtI for efficient promoter utilisation.
- To investigate the regulation of PrtI activity by putative anti- σ factor PrtR.

Chapter 4:

Hypothesis -

- Ecf41_{Bc1} and Ecf41_{Bc2} belong to the ECF41 group of σ factors that are activated in response to an extra cytoplasmic stress signal whereupon these σ factors initiate the transcription of genes that are required to cope with the extracytoplasmic stress, by recognising key sequences within the promoters present upstream of such genes. Ecf41_{Bc1} and Ecf41_{Bc2}, like ECF41 group σ factors possess an on board regulatory domain in the form of a C-terminal extension.
- Ecf42_{Bc} belongs to the ECF42 group of σ factors that is activated in response to an extra cytoplasmic stress signal whereupon the sigma factor initiates the

transcription of genes that are required to cope with the extracytoplasmic stress, by recognising key sequences within the promoters present upstream of these genes. Ecf42_{Bc} possesses an on board regulatory domain in the form of a C-terminal extension.

Aims -

- To study the role of Ecf41_{Bc1}, Ecf41_{Bc2} and Ecf42_{Bc} in *B. cenocepacia* by (1) identifying the respective σ -dependent promoters to reveal their target genes (2) studying the effects of deleting each σ factor encoding gene on stress responses of *B. cenocepacia*.
- To investigate the regulation of Ecf41_{Bc1}, Ecf41_{Bc2} and Ecf42_{Bc} activity by their respective C-terminal extensions.

Chapter 5:

Hypotheses -

- At OrbS-dependent promoters, 'TAAA' and 'CGTC' form the -35 and -10 promoter regions.
- The length and GC richness of the spacer and the A+G richness near the transcription start site is important for promoter utilisation by OrbS.
- In addition to the genes within the OrbS operon, OrbS may regulate other genes on the *B. cenocepacia* genome.
- In addition to transcriptional regulation of *orbS* by Fur, an alternative post-translational mechanism for regulation of OrbS may operate.

Aims -

- To test the effect of base-specific alterations in the OrbS-dependent promoter PorbH on promoter utilisation by Orbs in *B. cenocepacia*.
- To verify novel OrbS-dependent promoters previously identified *in silico*
- To investigate the post-translational regulation of OrbS in a *B. cenocepacia fur* mutant

Chapter 2: Materials and methods

Some methods and materials have been taken and adapted/modified with minor amendments from K. Agnoli PhD thesis, 2007, S. Haldipurkar MSc thesis, 2012 and H. Spiewak PhD thesis, 2016.

2.1 Bacteriological techniques

2.1.1 Bacterial strains

All bacterial strains used in this investigation are listed in table 2.1

Table 2.1. Bacterial strains

Bacterial strains	Genotype or description ^a	Source or reference
<i>Burkholderia cenocepacia</i>		
715j	CF isolate, prototroph (Orb ⁺ Pch ⁺)	Darling <i>et al.</i> , 1998
715j- <i>fur</i> ::Tp	715j with Tp ^R cassette inserted in <i>fur</i>	Han and Thomas, unpublished
H111	Clinical isolate, prototroph	Römling <i>et al.</i> , 1994
H111Δ <i>prtI</i>	H111 containing an in-frame deletion of 486 bp within <i>prtI</i> (<i>I35_RS16290</i>)	This study
H111Δ <i>prtR</i>	H111 containing an in-frame deletion of 717 bp within <i>prtR</i> (<i>I35_RS16295</i>)	This study
H111Δ <i>ecf41</i> _{Bc1}	H111 containing an in-frame deletion of 855 bp within <i>I35_RS16330</i>	This study
H111Δ <i>ecf41</i> _{Bc2}	H111 containing an in-frame deletion of 870 bp within <i>I35_RS16335</i>	This study
H111Δ <i>ecf41</i> _{Bc1} -Δ <i>ecf41</i> _{Bc2}	H111 containing an in-frame deletion of 855 bp within <i>I35_RS16330</i> and an in-frame deletion of 870 bp within <i>I35_RS16335</i>	This study
H111Δ <i>ecf42</i> _{Bc}	H111 containing an in-frame deletion of 1230 bp within <i>I35_RS20320</i>	This study
<i>E. coli</i>		
JM83	F- <i>ara</i> Δ(<i>lac-proAB</i>) <i>rpsL</i> φ80d <i>lacZ</i> Δ <i>M15</i> (Sm ^R)	Yanisch-Perron <i>et al.</i> , 1985

MC1061	<i>hsdR araD139 Δ(ara-leu)7697 ΔlacX74 galU galk rpsL (Sm^R)</i>	Casadaban and Cohen, 1979
SM10(λ pir)	<i>thi-1 thr leu tonA lacY supE recA RP4-2-Tc::Mu (Km^R) (λpir)</i>	Simon <i>et al.</i> , 1983
S17-1(λ pir)	<i>thi proA hsdR recA RP4-2-tet::Mu-1 kan::Tn7 integrant (Tp^R Sm^R)</i>	Simon <i>et al.</i> , 1983
BTH101	<i>F⁻ cya-99 araD139 galE15 galK16 hsdR2 mcrA1 mcrB1 rpsL1 (Sm^R)</i>	Karimova <i>et al.</i> , 1998
BL21 (λD3)	<i>F⁻ ompT gal dcm lon hsdS_B (rB⁻ mB⁻) (λDE3 [lacI lacUV5-T7 gene 1 ind1 Sam7 nin5])</i>	Studier & Moffatt, 1986
MG1655	<i>F⁻ wild-type</i>	Seaver & Imlay, 2004
LC106	<i>MG1655 ΔahpC-ahpF kan::ahpF, Δ(katG17::Tn10)1 (Tet^S), Δ(katE12::Tn10)1</i>	Seaver & Imlay, 2004

a: Orb⁺, ornibactin-positive phenotype; Pch⁺, pyochelin-positive phenotype; Tp^R, trimethoprim resistance; Sm^R, streptomycin resistance; Km^R, kanamycin resistance; Tet^S, tetracycline sensitive.

2.1.2 Bacteriological media and supplements

2.1.2.1 Bacteriological media

To make agar plates, agar (VWR) was added to liquid media at a final concentration of 1.5% and autoclaved at 120°C for 20 minutes at 15 psi (standard procedure). After cooling the medium to approximately 50°C, required antibiotics and other supplements were added and approximately 25 ml agar was poured into each sterile Petri dish and was kept at room temperature to solidify before use.

2.1.2.1.1 LB medium

To prepare LB broth tryptone, yeast extract and NaCl were added to ddH₂O to final concentration of 1%, 0.5% and 1% respectively and autoclaved as standard procedure

and stored at room temperature. To prepare LB agar, agar was added to a final concentration of 1.5%.

2.1.2.1.2 M9 glucose medium

Glucose was added to ddH₂O to a final concentration of 0.5% and autoclaved as standard procedure. After cooling the medium to approximately 50°C, x10 M9 salts was added to a final concentration of 10%. Then 1 ml each of 1 mM MgSO₄ and 0.1M CaCl₂ per litre of medium were added. To make M9 agar, agar was added to a final concentration of 1.5%. X10 M9 salts: 60 g Na₂HPO₄, 30 g KH₂PO₄, 5 g NaCl and 10 g NH₄Cl were added to 1000 ml deionized H₂O and autoclaved at 120°C for 20 minutes at 15 psi and stored at room temperature.

1M MgSO₄: 24.6 g MgSO₄·7H₂O was dissolved in 100 ml ddH₂O, filter sterilized and stored at room temperature.

0.1M CaCl₂: 15 g CaCl₂·2H₂O was dissolved in 100 ml ddH₂O, filter sterilized and stored at room temperature.

2.1.2.1.3 M9-CAA medium

Cas-amino acids (BD Difco) to final concentration of 0.1-1% as required was added to M9 minimal glucose medium

2.1.2.1.4 M9 glycerol CAA-X-gal-IPTG medium

M9 CAA (1%) agar was made as described above without glucose. The following components with final concentrations of 0.5% glycerol, 0.0005% thiamine freshly prepared in ddH₂O, 100 μM CaCl₂, 1 mM MgSO₄, IPTG 100 μM and X-gal 40 μg/mL were then added. Any required antibiotics were also added before pouring into sterile Petri dishes.

2.1.2.1.5 0.1% chlorophenylalanine medium

4-Chloro-DL-phenylalanine (Sigma-Aldrich) to final concentration of 0.1 was added to M9 minimal glucose medium.

2.1.2.1.6 EPS (γ-glutamate-glycerol) medium

To make Exopolysaccharide (EPS) producing media the following components to the final concentrations of 0.169 % glutamic acid (sodium salt), 0.3 % Tris base, 0.1 ml MgSO₄·7H₂O (10 % w/v), 0.1 ml CaCl₂·6H₂O (22 % w/v), 0.1 ml K₂HPO₄·3H₂O (22 % w/v) were added to ddH₂O. After adjusting the pH to 6.8 (or 6.0), the final volume was adjusted to 99 ml. A final concentration of 1.5 % agar (Oxoid) was added prior to autoclaving. 50% sterile glycerol solution was added to a final concentration of 1% to the medium and poured into sterile Petri dishes.

2.1.2.1.7 IST medium

Iso-Sensitest broth (Oxoid) was added to ddH₂O at a final concentration of 23.4% and autoclaved as standard procedure and stored at room temperature. To prepare IST agar, a final concentration of 1.5% agar was added.

2.1.2.1.8 Lennox medium

To prepare Lennox broth tryptone, yeast extract and NaCl were added to ddH₂O to final concentration of 1%, 0.5% and 0.5% respectively and autoclaved as standard procedure and stored at room temperature. To prepare Lennox agar, a final concentration of 1.5% agar was added.

2.1.2.1.9 BHI medium

Brain heart infusion (BD Difco) powder to a final concentration of 3.7% was added to ddH₂O and autoclaved as standard procedure. To make BHI agar a final concentration of 1.5 % agar was added.

2.1.2.1.10 Protease agar

Protease agar was made by combining dialysed BHI (D-BHI) agar with skimmed milk. D-BHI (100 ml) was prepared by mixing 1.85 g in 5 ml ddH₂O, transferred to a 12000-14000 MWCO dialysis tube and dialysed against 100 ml ddH₂O overnight at 4°C thus

giving half strength BHI. Agar was added to a final concentration of 3% and autoclaved as standard procedure.

An equal quantity (100 ml) of 3% (w/v) skimmed milk powder (Sigma) in ddH₂O was separately autoclaved for 5 minutes, at 120°C for 20 minutes at 15 psi.

The dialysed BHI agar and skimmed milk solution were then combined and poured in sterile Petri dishes, hence giving final concentrations of quarter strength BHI and 1.5% agar.

2.1.2.1.11 Maltose MacConkey agar

MacConkey agar (BD Difco) was dissolved in ddH₂O to give a final concentration of 4% and autoclaved as standard procedure. After cooling the medium to approximately 50°C, 1% glucose-free maltose and appropriate antibiotics were added and poured in sterile Petri dishes.

2.1.2.1.12 CAS (chrome azurol S) agar

(Taken from K. Agnoli PhD thesis, 2007, with minor amendments)

CAS agar was prepared by adding CAS mix to Y minimal agar.

CAS mix was prepared by combining 10 ml of 1 mM FeCl₃.6H₂O (dissolved in 10mM HCl) and 60.5 mg CAS powder in 50 ml ddH₂O. 72.9 mg hexadecyltrimethylammonium bromide (HDTMA) was dissolved in 40 ml ddH₂O, which was then added to the CAS/FeCl₃ mixture with constant stirring. CAS mix was stored at room temperature in the dark.

Y minimal agar was prepared by adding the following components to the final concentrations of 0.169 % glutamic acid (sodium salt), 0.3 % Tris base, 0.1 ml MgSO₄.7H₂O (10 % w/v), 0.1 ml CaCl₂.6H₂O (22 % w/v), 0.1 ml K₂HPO₄.3H₂O (22 % w/v) were added to ddH₂O. After adjusting the pH to 6.8, the final volume was adjusted to 90 ml. A final concentration of 1.5 % agar (Oxoid) was added prior to autoclaving.

CAS mix and Y minimal agar were prepared and autoclaved separately. Following cooling to approximately 55°C, 10 ml CAS mix was slowly poured into 90 ml Y minimal agar, and mixed thoroughly but slowly until homogeneous. On combining the CAS mix

with Y minimal agar, the resultant CAS agar contained 10 μM FeCl_3 . For iron-rich CAS agar, CAS mix containing 50 μM additional $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$ was added to the medium prior to pouring, to give a final concentration of 60 μM Fe^{3+} .

2.1.2.1.13 Motility and swarming agar

Motility agar was made by adding tryptone, NaCl and agar at final concentrations of 1%, 0.5% and 0.25% to ddH₂O, autoclaved and poured in sterile Petri dishes as standard procedure. Swarming agar was made by adding nutrient broth (BD Difco), NaCl and Eiken agar at final concentrations of 0.8%, 0.5% and 0.6% to tap water and autoclaved and poured in sterile Petri dishes as standard procedure.

2.1.2.2 Bacteriological media supplements

2.1.2.2.1 Antibiotics

Stock solutions of antibiotics used for making selective media are listed below. They were stored in polypropylene universal bottles at -20°C. When ddH₂O was used as a solvent the stocks were filter sterilized using a 0.22 μm syringe filter.

Table 2.2: Concentration of antibiotics used in selective media

Antibiotic	Stock concentration (mg/ml)	Final concentration ($\mu\text{g/ml}$)	
		<i>E. coli</i>	<i>B. cenocepacia</i>
Ampicillin	100 in ddH ₂ O	100	n/a
Chloramphenicol	25 in 100% ethanol	25	50
Kanamycin	50 in ddH ₂ O	25-50	50
Tetracycline	10 in 50% ethanol	10	200
Trimethoprim	25 in DMSO	25	25
Gentamycin	25 in ddH ₂ O	25	n/a

2.1.2.2.2 Other media supplements

Stock solutions of other supplements added to bacteriological media are listed below.

Table 2.3: Concentration of supplements used in selective media

Supplement	Stock	Concentration	
		<i>E. coli</i>	<i>B. cenocepacia</i>
2,2'-dipyridyl	0.1 M in 100% ethanol	175 μ M	100 μ M (M9), 200 μ M (LB)
FeCl ₃	0.1 M in 10 mM HCl	n/a	50 μ M
IPTG	0.1M in ddH ₂ O	0.2mM	n/a
X-gal	25 mg/ml in DMSO	40 μ g/ml	40 μ g/ml
Thiamine	10 mg/ml in ddH ₂ O	0.005 mg/ml	0.005 mg/ml
DL-4-Chlorophenylalanine (cPhe)	n/a	1 mg/ml	1 mg/ml
Glycerol solution	50% (v/v) in ddH ₂ O	0.5 % (v/v)	0.5 % (v/v)

2.1.3 Bacterial storage glycerol stocks

For long term storage of bacterial cultures, 0.7 ml of a broth culture grown overnight was mixed with 0.3 ml of 50% glycerol, vortexed and stored at -80°C in sterile cryogenic vials.

2.1.4 Determination of bacterial growth rate

(Taken from K. Agnoli PhD thesis, 2007, with minor amendments).

In order to determine the growth rate of bacterial strains in liquid medium, the strains to be analysed were grown overnight at 37°C with shaking. The overnight cultures were then diluted in the same medium to an OD₆₀₀ of 0.5. Growth was monitored by taking 1 ml samples immediately following inoculation and at 30-60 minutes' intervals thereafter, while the cultures were incubated at 37°C with shaking. The OD₆₀₀ of each sample was determined at the time of sampling. Experiments were carried out in biological triplicates and experimental replicates.

Growth curves were plotted using Microsoft Excel from the data obtained, using a logarithmic scale for the optical density measurements, and a line of best fit calculated using the 'exponential trendline' facility for the exponential growth phase (characterised by a linear progression). The equation of each line of best fit was used in conjunction with the following equations to determine growth rate.

The equation of each line was in the following format:

$$y = me^{cx}$$

The elements of this equation were inserted into the following:

$$T_n = (\ln n - \ln m) / c$$

$$T_{2n} = (\ln 2n - \ln m) / c$$

Where T_n = time (hours) at which the culture reached an OD₆₀₀ of n

And T_{2n} = time (hours) taken for the OD₆₀₀ to double at T_n

Growth rate (doublings per hour) was calculated as:

$$1/(T_{2n} - T_n)$$

2.1.5 Efficiency of plating (EOP)

The required strain of bacterium was grown overnight in nutrient rich broth medium (LB/BHI) for approximately 16 hours at 37°C with shaking. Then, the OD₆₀₀ of each sample (1 ml) was adjusted to 0.500 by diluting in sterile saline (0.85% NaCl in ddH₂O). This was further subjected to ten-fold serial dilutions serially diluted in sterile saline and 0.1 ml amounts from the dilutions were spread out on required agar plates and incubated at 37°C overnight. Alternatively, 5 µl aliquots of the dilutions were spotted on required agar plates and incubated at 37°C overnight. The number of colonies were counted the next day to determine number of colony forming units per ml of bacterial culture (CFU/ml). Experiments were carried out in biological triplicates and experimental replicates.

2.2 Recombinant DNA techniques

2.2.1 Plasmids

The plasmids used in this study are listed in table 2.4

Table 2.4: Plasmids

Plasmids	Description ^a	Source or reference
pBBR1MCS (pBBR)	Broad host range cloning vector, (Cm ^R)	Kovach et al., 1994
pBBR-fur2	pBBR carrying <i>fur</i> in opposite orientation to <i>lacZ</i> promoter cloned between sites <i>SmaI</i> and <i>BamHI</i> (Cm ^R)	Haldipurkar and Thomas, unpublished
pBBR-I35_RS16285-FLAG	pBBR containing <i>I35_RS16285</i> modified to encode a C-terminal FLAG epitope cloned between sites <i>HindIII</i> and <i>XbaI</i> (Cm ^R)	This study
pBBR1MCS2 (pBBR2)	Mobilisable broad host-range cloning vector (Km ^R)	Kovach et al. 1995
pBBR2-fur3	pBBR2 containing <i>fur</i> cloned between sites <i>HindIII</i> and <i>BamHI</i> (Km ^R)	Agnoli et al., 2006
pBBR2-orbS	pBBR2 containing <i>orbS</i> cloned between sites <i>HindIII</i> and <i>BamHI</i> (Km ^R)	Agnoli et al., 2006
pBBR2-Ecf41 _{Bc1}	pBBR2 containing full length <i>I35_RS16330</i> cloned between sites <i>HindIII</i> and <i>BamHI</i> (Km ^R)	This study
pBBR2-Ecf41 _{Bc1-213}	pBBR2 containing first 213 amino acids of <i>I35_RS16330</i> cloned between sites <i>HindIII</i> and <i>BamHI</i> (Km ^R)	This study
pBBR2-Ecf41 _{Bc1-201}	pBBR2 containing first 201 amino acids of <i>I35_RS16330</i> cloned between sites <i>HindIII</i> and <i>BamHI</i> (Km ^R)	This study
pBBR2-Ecf41 _{Bc1-176}	pBBR2 containing first 176 amino acids of <i>I35_RS16330</i> cloned between sites <i>HindIII</i> and <i>BamHI</i> (Km ^R)	This study
pBBR2-Ecf41 _{Bc2}	pBBR2 containing full length <i>I35_RS29600</i> cloned between sites <i>HindIII</i> and <i>BamHI</i> (Km ^R)	This study
pBBR2-Ecf41 _{Bc2-215}	pBBR2 containing first 215 amino acids of <i>I35_RS29600</i> cloned between sites <i>HindIII</i> and <i>BamHI</i> (Km ^R)	This study

pBBR2-Ecf41 _{Bc2-198}	pBBR2 containing first 198 amino acids of <i>I35_RS29600</i> cloned between sites <i>HindIII</i> and <i>BamHI</i> (Km ^R)	This study
pBBR2-Ecf41 _{Bc2-175}	pBBR2 containing first 175 amino acids of <i>I35_RS29600</i> cloned between sites <i>HindIII</i> and <i>BamHI</i> (Km ^R)	This study
pBBR2-PrtI	pBBR2 containing <i>prtI</i> cloned between sites <i>HindIII</i> and <i>BamHI</i> (Km ^R)	This study
pBBR2-PrtI(b)	pBBR2 containing <i>prtI</i> with an improved Shine Dalgarno sequence before the promoter cloned between sites <i>HindIII</i> and <i>BamHI</i> (Km ^R)	This study
pBBR2-PrtIR	pBBR2 containing full length <i>prtIR</i> cloned between sites <i>HindIII</i> and <i>XbaI</i> (Km ^R)	This study
pBBR2-Ecf42 _{Bc}	pBBR2 containing full length <i>I35_RS20320</i> cloned between sites <i>HindIII</i> and <i>SacI</i> (Km ^R)	This study
pBBR2-Ecf42 _{Bc-175}	pBBR2 containing first 525 amino acids of <i>I35_RS20320</i> cloned between sites <i>HindIII</i> and <i>BamHI</i> (Km ^R)	This study
pBBR1MCS5	Mobilisable broad host-range cloning vector. IncP- and ColE1-compatible (Gm ^R)	Kovach et al., 1995
pBBR-5-orbS	pBBR1MCS-5 containing <i>orbS</i> cloned between <i>HindIII</i> and <i>BamHI</i> (Gm ^R)	K. Agnoli, 2007
pKAGd4	Broad host-range <i>lacZ</i> transcriptional fusion vector (Cm ^R , Ap ^R)	Agnoli et al., 2006
pKAGd4-P _{orbH}	pKAGd4 containing P _{orbH} - <i>lacZ</i> fusion cloned between sites <i>HindIII</i> and <i>BamHI</i> (Cm ^R , Ap ^R)	Agnoli et al., 2006
pKAGd4-P _{orbHds (1-51)}	pKAGd4 containing P _{orbH} promoter- <i>lacZ</i> fusions, containing corresponding ds oligonucleotide cloned between sites <i>HindIII</i> and <i>BamHI</i> (see Appendix for sequences) (Cm ^R , Ap ^R)	K. Agnoli, 2007
pKAGd4-P _{tac}	pKAGd4 bearing P _{tac} - <i>lacZ</i> fusion cloned between sites <i>HindIII</i> and <i>BamHI</i> (Cm ^R , Ap ^R)	K. Agnoli, 2007
pKAGd4-P _{tac2}	pKAGd4 bearing minimal version of P _{tac} - <i>lacZ</i> fusion cloned between sites <i>HindIII</i> and <i>BamHI</i> (Cm ^R , Ap ^R)	This study
pKAGd4-P _{cysl}	pKAGd4 bearing <i>B. cenocepacia</i> P _{cysl} - <i>lacZ</i> fusion cloned between sites <i>HindIII</i> and <i>BamHI</i> (Cm ^R , Ap ^R)	K. Agnoli, 2007
pKAGd4-P _{bfd}	pKAGd4 bearing <i>B. cenocepacia</i> P _{bfd} - <i>lacZ</i> fusion cloned between sites <i>HindIII</i> and <i>BamHI</i> (Cm ^R , Ap ^R)	A. Asghar, 2002

pKAGd4-P _{firs}	pKAGd4 bearing <i>B. cenocepacia</i> P _{firs} - <i>lacZ</i> fusion cloned between sites <i>Hind</i> III and <i>Bam</i> HI (Cm ^R , Ap ^R)	J. Mohanlal and M. Thomas, unpublished
pKAGd4-P _{orbS}	pKAGd4 containing P _{orbS} full length- <i>lacZ</i> fusion cloned between sites <i>Hind</i> III and <i>Bam</i> HI (Cm ^R , Ap ^R)	Agnoli <i>et al.</i> , 2006
pKAGd4-P _{orbEds1}	pKAGd4 containing P _{orbE} - <i>lacZ</i> fusion cloned between sites <i>Hind</i> III and <i>Bam</i> HI (Cm ^R , Ap ^R)	Agnoli <i>et al.</i> , 2006
pKAGd4-P _{orblds1}	pKAGd4 containing P _{orbl} - <i>lacZ</i> fusion cloned between sites <i>Hind</i> III and <i>Bam</i> HI (Cm ^R , Ap ^R)	Agnoli <i>et al.</i> , 2006
pKAGd4-P _{orbS-69}	pKAGd4 containing P _{orbS} with upstream endpoint at -69 relative to <i>orbS</i> start codon cloned between sites <i>Hind</i> III and <i>Bam</i> HI (Cm ^R , Ap ^R)	Agnoli <i>et al.</i> , 2006
pKAGd4-P _{orbS-40}	pKAGd4 containing P _{orbS} with upstream endpoint at -40 relative to <i>orbS</i> start codon cloned between sites <i>Hind</i> III and <i>Bam</i> HI (Cm ^R , Ap ^R)	Agnoli <i>et al.</i> , 2006
pKAGd4-P _{orbS+5}	pKAGd4 containing P _{orbS} with upstream endpoint at +5 relative to <i>orbS</i> start codon cloned between <i>Hind</i> III and <i>Bam</i> HI (Cm ^R , Ap ^R)	Agnoli <i>et al.</i> , 2006
pKAGd4-P _{fpr}	pKAGd4 containing P _{fpr} cloned between sites <i>Hind</i> III and <i>Bam</i> HI (Cm ^R , Ap ^R)	This study
pKAGd4-P _{ureA}	pKAGd4 containing P _{ureA} cloned between sites <i>Hind</i> III and <i>Bam</i> HI (Cm ^R , Ap ^R)	
pKAGd4- P _{prtl} (1-15)	pKAGd4 containing P _{prtl} promoter- <i>lacZ</i> fusions, containing corresponding ds oligonucleotide cloned between sites <i>Hind</i> III and <i>Bam</i> HI (see table 2.5) (Cm ^R , Ap ^R)	This study
pKAGd4-P _{aphD}	pKAGd4 bearing P _{aphD} - <i>lacZ</i> fusion cloned between sites <i>Hind</i> III and <i>Bam</i> HI (Cm ^R , Ap ^R)	This study
pKAGd4-P _{lysR}	pKAGd4 bearing P _{lysR} - <i>lacZ</i> fusion cloned between sites <i>Hind</i> III and <i>Bam</i> HI (Cm ^R , Ap ^R)	This study
pKAGd4-P _{phnB2}	pKAGd4 bearing P _{phnB2} - <i>lacZ</i> fusion cloned between sites <i>Hind</i> III and <i>Bam</i> HI (Cm ^R , Ap ^R)	This study
pKAGd4-P _{prtI}	pKAGd4 containing P _{prtI} promoter- <i>lacZ</i> fusion, consisting of the full length <i>prtI</i> promoter cloned between sites <i>Hind</i> III and <i>Bam</i> HI (Cm ^R , Ap ^R)	This study
pKAGd4-P _{prtI-43}	pKAGd4 containing P _{prtI} promoter- <i>lacZ</i> fusion – 43 bases relative to putative <i>prtI</i>	This study

	start codon, cloned between sites <i>HindIII</i> and <i>BamHI</i> (Cm ^R , Ap ^R)	
pKAGd4-P _{cat}	pKAGd4 containing P _{I35_RS16285} promoter- <i>lacZ</i> fusion cloned between sites <i>HindIII</i> and <i>BamHI</i> (Cm ^R , Ap ^R)	This study
pSRK-Km	Broad host-range expression vector (Km ^R)	
pSRK-Km-PrtI-FLAG	pSRK-Km containing full length <i>prtI</i> modified to encode a C-terminal FLAG epitope cloned between sites <i>NdeI</i> and <i>XbaI</i> (Km ^R)	This study
pETDuet-1	<i>E. coli</i> specific vector for high-level expression of cloned genes. Contains two T7 promoters and multiple cloning sites (Ap ^R)	Novagen
pETDuet-1-I35_RS16285 _{His6}	pETDuet-1 containing <i>RS16285</i> between <i>BamHI</i> and <i>HindIII</i> sites of MCS1, incorporating an N-terminal His tag. (Ap ^R)	This study
pHLS16	pETDuet-1 containing BCAS0667.CTD (C-terminal 258 codons) cloned into sites <i>BamHI</i> and <i>Sall</i> sites, incorporating an N-terminal His tag.	H. Spiewak, 2015
pEX18TpTer- <i>pheS</i>	Derivative of allelic replacement vector pEX18Tp- <i>pheS</i> with <i>dfrB2</i> gene replaced by a version fused to the <i>rrnBT1T2</i> terminators. ColE1- derived replicon, <i>oriT</i> ⁺ , <i>mob</i> ⁺ and mutated α -subunit of phenylalanyl tRNA synthase, <i>pheS</i> , for cPhe counterselection. (Tp ^R)	H. Spiewak, 2015
pEX18TpTer- <i>pheS</i> - Δ <i>prtI</i>	pEX18Tp- <i>pheS</i> containing a Δ <i>prtI</i> allele cloned between sites <i>BamHI</i> and <i>HindIII</i> . (Tp ^R)	This study
pEX18TpTer- <i>pheS</i> - Δ <i>Ecf41</i> _{Bc1}	pEX18Tp- <i>pheS</i> containing a Δ <i>I35_RS16330</i> allele cloned between sites <i>BamHI</i> and <i>HindIII</i> . (Tp ^R)	This study
pEX18TpTer- <i>pheS</i> - <i>I</i> Scel	pEX18TpTer- <i>pheS</i> with an <i>I</i> - <i>Scel</i> site (Tp ^R)	H. Spiewak and M. Thomas, unpublished
pEX18TpTer- <i>pheS</i> - <i>I</i> Scel- Δ <i>PrtI</i>	pEX18Tp- <i>pheS</i> - <i>I</i> Scel containing a Δ <i>prtI</i> allele cloned between sites <i>BamHI</i> and <i>HindIII</i> . (Tp ^R)	This study
pEX18TpTer- <i>pheS</i> - <i>I</i> Scel- Δ <i>Ecf41</i> _{Bc2}	pEX18Tp- <i>pheS</i> - <i>I</i> Scel containing a Δ <i>I35_RS29600</i> allele cloned between sites <i>BamHI</i> and <i>HindIII</i> . (Tp ^R)	This study

pEX18TpTer- <i>pheS</i> -ISce1- Δ Ecf42 _{Bc}	pEX18Tp- <i>pheS</i> -ISceI containing a Δ I35_RS20320 allele cloned between sites <i>Xba</i> I and <i>Hind</i> III. (Tp ^R)	This study
pEX18TpTer- <i>pheS</i> -ISceI-Cm	pEX18TpTer- <i>pheS</i> -ISceI with a chloramphenicol resistance gene (Tp ^R , Cm ^R)	H. Spiewak and M. Thomas, unpublished
pEX18TpTer- <i>pheS</i> -ISceI-Cm- Δ Ecf41 _{Bc1}	pEX18Tp- <i>pheS</i> -ISce1-Cm containing a Δ I35_RS16330 allele cloned between sites <i>Bam</i> HI and <i>Hind</i> III. (Tp ^R , Cm ^R)	This study
pEX18TpTer- <i>pheS</i> -ISceI-Cm- Δ prtR	pEX18Tp- <i>pheS</i> -ISce1-Cm containing a Δ prtR allele cloned between sites <i>Bam</i> HI and <i>Hind</i> III. (Tp ^R)	This study
pUT18	High copy number BACTH vector, allows in-frame fusion of gene to N-terminal coding sequence of T18 (Ap ^R)	EUROMEDEX
pUT18-PrtI _{CTD}	pUT18 containing PrtI _{CTD} coding sequence cloned between sites <i>Xba</i> I and <i>Acc</i> 65I (Ap ^R)	This study
pUT18-PrtR _{NTD}	pUT18 containing PrtR _{NTD} coding sequence cloned between sites <i>Xba</i> I and <i>Acc</i> 65I (Ap ^R)	This study
pUT18C	High copy number BACTH vector, allows in-frame fusion of gene to C-terminal coding sequence of T18 adenylate cyclase domain (Ap ^R)	EUROMEDEX
pUT18C-PrtI _{CTD}	pUT18C containing PrtI _{CTD} coding sequence cloned between sites <i>Xba</i> I and <i>Acc</i> 65I (Ap ^R)	This study
pUT18C-PrtR _{NTD}	pUT18C containing PrtR _{NTD} coding sequence cloned between sites <i>Xba</i> I and <i>Acc</i> 65I (Ap ^R)	This study
pKT25	Low copy number BACTH vector, allows in-frame fusion of gene to C-terminal coding sequence of T25 adenylate cyclase domain (Km ^R)	EUROMEDEX
pKT25-PrtI _{CTD}	pKT25 containing PrtI _{CTD} coding sequence cloned between sites <i>Xba</i> I and <i>Acc</i> 65I (Km ^R)	This study
pKT25-PrtR _{NTD}	pKT25 containing PrtR _{NTD} coding sequence cloned between sites <i>Xba</i> I and <i>Acc</i> 65I (Km ^R)	This study
pKNT25	Low copy number BACTH vector, allows in-frame fusion of gene to N-terminal coding sequence of T25 adenylate cyclase domain (Km ^R)	EUROMEDEX
pKNT25-PrtI _{CTD}	pKNT25 containing PrtI _{CTD} coding sequence cloned between sites <i>Xba</i> I and <i>Acc</i> 65I (Km ^R)	This study
pKNT25-PrtR _{NTD}	pKNT25 containing PrtR _{NTD} coding sequence cloned between sites <i>Xba</i> I and <i>Acc</i> 65I (Km ^R)	This study

a: Tp^R, trimethoprim resistance; Sm^R, streptomycin resistance; Km^R, kanamycin resistance; Cm^R, chloramphenicol resistance; Ap^R, ampicillin resistance; Gm^R, gentamycin resistance

2.2.2 Primers and oligonucleotides:

Table 2.5: Primers and oligonucleotides used in this study

Primer/oligonucleotide name	Sequence (5' to 3')	Restriction site (underlined)
Primers for Screening/Sequencing		
M13 Forward	TGTAACACGACGGCCAGT	
M13 Reverse	CAGGAAACAGCTATGACC	
M13revBACTH	GTGTGGAATTGTGAGCGGAT	
pETDuet-T7-1for	ATGCGTCCGGCGTAGA	
pACYCDuet-T7-1rev	GATTATGCGGCCGTGTACAA	
pEX18Tpfors	GAAGCCAGTTACCTTCGGA	
pEX18Tprev	TTGTCGGTGAACGCTCTCCT-	
AP10	GAACGCTCTCCTGAGTA	
AP11	CCTCTATAGTGAGTCGG	
Primers for cloning required to carry out ChIP-seq/ PCR validation of ChIP-seq samples		
FurFLAGa	<u>GCGCGGATCCTCGATGTCGGCCAGCAGGTC</u>	<i>Bam</i> HI
FurFLAGb	TCATTTATCATCGTCGTCCTTTATAATCGTGCTTGCG GTGCGGGCAGT	
FurFLAGc	GATTATAAAGACGACGATGATAAATGAAGCGCGC CACCCGCGCA	
FurFLAGd	<u>GCGCTCTAGACGCGCGGCATTTCCCGATCA</u>	<i>Xba</i> I
RpoEFLAGFor	<u>GCGCCATATGAGTGAAAAAGAAATCGATCA</u>	<i>Nde</i> I
RpoEFLAGRev	<u>GCGCTCTAGATTATTTATCATCGTCGTCCTTTATAAT</u> CCCAGCGCTTGCCTTCGG	<i>Xba</i> I
BCAL3486FLAGFor	<u>GCGCCATATGCCGAGCCTCGACCCGGCCA</u>	<i>Nde</i> I
BCAL3486FLAGRev	<u>GCGCTCTAGATTATTTATCATCGTCGTCCTTTATAAT</u> CGCTTGCCGCCCGCTCGGAAT	<i>Xba</i> I
BCAM2748FLAGFor	<u>GCGCCATATGACCGACGACGCACAAACCTT</u>	<i>Nde</i> I
BCAM2748FLAGRev	<u>GCGCTCTAGATTATTTATCATCGTCGTCCTTTATAAT</u> CGAAGCTCTGGACGTGGCCGC	<i>Xba</i> I
OrbSFLAGFor	<u>GCGCCATATGGCGGAAGTGCTCGACCGACC</u>	<i>Nde</i> I
OrbSFLAGRev(A)	<u>GCGCTCTAGATTATTTATCATCGTCGTCCTTTATAAT</u> CGGCGTCTGCGTGCTCTTCTCCAGGAACACCGG GCAGGCG	<i>Xba</i> I

OrbSFLAGRev(B)	GCGCTCTAGATTATTTATCATCGTCGTCCTTTATAAT CCCGCCGCCGCGCGCGGC	<i>XbaI</i>
PrtIFLAGFor	GCGCCATATGCCCTCCAACGCCCTCGA	<i>NdeI</i>
PrtIFLAGRev	GCGCTCTAGATTATTTATCATCGTCGTCCTTTATAAT CTTTCTTCATCAGCCGCAAGG	<i>XbaI</i>
Bcam0849flagfor	GCGCCATATGACGCATGAGGCCACGCATCG	<i>NdeI</i>
Bcam0849flagrev	GCGCTCTAGATTATTTATCATCGTCGTCCTTTATAAT CCGCATCCATCGCGCGCTTGA	<i>XbaI</i>
pIOTVfor	GCGCAAGCTTGTAGCGAATCATGGACGATC	<i>HindIII</i>
pIOTVrev	GCGCGGATCCACCGGCGCGTGAGTTTTAAA	<i>BamHI</i>
pFlrSfor2	CGCAAGCTTATGTGATAACCGAACGGGGC	<i>HindIII</i>
pFlrSrev	CGCGGATCCGACAGCTTGTCAGCGGACAT	<i>BamHI</i>
Primers for construction of in-frame deletion mutants		
PrtIdelA	GCGCGGATCCTGTCCGGCGCGTACGGATTT	<i>BamHI</i>
PrtIdelB	GTTCAATTTCTTCATCAGCCGAGGGCGTTGGAGGG CATGG	
PrtIdelC	CCATGCCCTCCAACGCCCTCCGGCTGATGAAGAAA TGAAC	
PrtIdelD	GCGCAAGCTTCCCGCTGTCCGTCGTGAACAA	<i>HindIII</i>
Prtlout-for	GCGAGCGTCTGCTCGTAGAA	
Prtlout-rev	CCCGGTCATCCTGATAGAG	
Prtlinfor	CCCGATCGTCAGATCATTGT	
Prtlinrev	CATCTGATAAGCCTTGACCG	
Prtloutfor2	TTCAATGCGAGGCTTCGGAT	
Prtloutrev2	TAGCAGCCGATGTTCCGATT	
prtRdelA	GCGCGGATCCAAGCATTGCGCGAACTCATC	<i>BamHI</i>
prtRdelB	CTCCGCTTCAAGTCTGCTGAAGGTCGTGTTTCGTT GGA	
prtRdelC	TCCGAACGAACACGACCTTCAGCAGACTTGAAGCG GAG	
prtRdelD	GCGCAAGCTTTGAAACCCTGTGCATCCGAT	<i>HindIII</i>
prtRoutfor	GCTGGAGGGAATAATCGATG	
prtRoutrev	AATGCAAATCGGCTCGGGAA	
3486delA	GCGCGGATCCGGGGCTCGCTCGTCATATCA	<i>BamHI</i>
3486delB	GCGCGCCGCTTCAAGTGCAGCGGGTCGAGGCTCG GCATCA	
3486delC	TGATGCCGAGCCTCGACCCGCTGCACTGAAGGCG GCGCGC	
3486delD	GCGCAAGCTTTACGTGCGCAAGAACGCCGC	<i>HindIII</i>
3486out-for	CCCGCTTCGGCGGTCCAAA	
3486out-rev	GCTGCATCGCTGCGAGCAA	
2748delA	GCGCGGATCCTCGATCGTGCTGCGCACGAA	<i>BamHI</i>
2748delB	GCGATCCGGGCGAGCTTCTCTGCGTCGTCGGTCAT GACGG	

2748delC	CCGTCATGACCGACGACGCAGAGAAGCTCGCCCG GATCGC	
2748delD	GCGCAAGCTTTTTCCGAGAGCTGGCCGCG	<i>HindIII</i>
2748out-for	TGATCTCGGACGGCGACGAC	
2748out-rev	ACCGAAGGCCTTGTCGGACG	
BCAM2748outfor2	ATGGAAGCTGATGCGGGAAA	
BCAM2748outrev2	GAACGTCAGGTCCGACAATT	
BCAM2748infor	GCTGTTTCCGCAGAAAACCA	
BCAM2748inrev	TCATCTTAGGAGTGTGCACC	
TPRdelA	GCGCTCTAGACGCCGAACCTCGACTGCGAGA	<i>XbaI</i>
TPRdelB	CACGCATCCATCGCGCGCTTCGTGGCCTCATGCGT CACGA	
TPRdelC	TCGTGACGCATGAGGCCACGAAGCGCGCGATGGA TGCGTG	
TPRdelD	GCGCAAGCTTGGTCTTCGCCATGCCCGGAT	<i>HindIII</i>
TPRout-for	CGCCGGCTATACGCTGATCC	
TPRout-rev	GCCTTCGTGTAGCGGTTTCGC	
BCAM0849outfor2	GCCGAAACGAAGGAACTGAT	
BCAM0849outrev2	CGATTTTCGTCAAACCTTCG	
BCAM0849infor	TTTTCCAGACCTTCACCGAC	
BCAM0849inrev	CGATCAGATATGTGGGGAAC	
Primers for Orbs		
fpr_in_fwd	GTCAGCCCCGAACCTACGAGGAGC	
fpr_in_rev	ATGTCGCCAGGTAATCGTGCC	
Primers for I35_RS16330 (BCAL3486 in J2135) and I35_RS29600 (BCAM2748 in J2315)		
BCAL3486for	GCGCAAGCTTAGTCGACCCGGCCGATGCGAA	<i>HindIII</i>
BCAL3486rev	GCGCGGATCCGCATGCGCACGCCGCTTCA	<i>BamHI</i>
Bcal3486.bam.213.rev	GCGCGGATCCTTAGCGCTTGCCGCCACCGTCGGA	<i>BamHI</i>
Bcal3486.bam.201.rev	GCGCGGATCCTTACTCGGCGAGCATCGCGCCGA	<i>BamHI</i>
Bcal3486.bam.176.rev	GCGCGGATCCTTACCCGCGCTGCTTCTCGACAT	<i>BamHI</i>
BCAM2748for	GCGCAAGCTTACGCGGAGTGTGGCCGTCATGA	<i>HindIII</i>
BCAM2748.bam.rev	GCGCGGATCCTGCACCAATCCGGGCTCGGC	<i>BamHI</i>
bcam2748.bam.215.rev	GCGCGGATCCTTAGAAGCTCTGGACGTGGCCGC	<i>BamHI</i>
bcam2748.bam.198.rev	GCGCGGATCCTTATTCGGCGAGCAACGCCTGGA	<i>BamHI</i>
bcam2748.bam.175.rev	GCGCGGATCCTTACCGCGATGCGTTTCATGCG	<i>BamHI</i>
Primers for I35_RS20320 (BCAM0849 in J2315)		
BCAM0849For	GCGCAAGCTTGATGGCCATGCCGCTGGATG	<i>HindIII</i>
BCAM0849Rev(full)	GCGCGAGCTCTCACGCATCCATCGCGCGCT	<i>SacI</i>
BCAM0849w/oCTDRev	GCGCGGATCCCAGCGAGCGTGCGTTTCGCGC	<i>BamHI</i>
Primers for PrtI		
prtIproForPCR	GCGCAAGCTTTCAATGCACTCCCGATCGTC	<i>HindIII</i>
prtIproRevPCR	GCGCGGATCCTGGCCGGGCGGTTGCCGAAA	<i>BamHI</i>
PrtIforA	GCGCAAGCTTATGATCCACGACCTTTCGCGAA	<i>HindIII</i>

PrtIforB	GCGCAAGCTTATGAAAAGGAGATATACCATGCCCT CCAACGCCCTCGA	<i>HindIII</i>
PrtIrev	GCGCGGATCCTCATTCTTCATCAGCCGCA	<i>BamHI</i>
Catduetfor	GCGCTCTAGACCCGGCGCTGACCCC	<i>XbaI</i>
Catduetrev	GCGCAAGCTTTCATCGTCCCTCCCCACTC	<i>HindIII</i>
PrtRrev	GCGCTCTAGATCAAGTCTGCTGATCGAGCG	<i>XbaI</i>
Cat-compFor	GCGCAAGCTTCTGACGCGGTTTCGACAATGATC	<i>HindIII</i>
Cat-compRevflag	GCGCTCTAGATCATTATCATCGTCTTTATAAT CTCGTCCCTCCCCACTCAGT	<i>XbaI</i>
PrtRfor	GCGCAAGCTTATGAACTTCTGCTCCTTCCTTG	<i>HindIII</i>
Primers for RT-PCR		
Cat-cytFor2	GCCGCGACGTGAATTACGA	
Cat-cytFor	ACGCGTTCTATTTCTGTCGAC	
Cat-cytRev	AAACAGCATCGCGACAACCA	
PrtIRfor	AGGGCTTCACCTATCAGGAA	
PrtIRRev	TGCCATTGCTTCACCTGTTC	
Primers for two-hybrid assay		
PrtICTDfor	GCGCTCTAGAGGCCGAACGCGAATTCGCG	<i>XbaI</i>
PrtICTDrev	GCGCGGTACCTCATTCTTCATCAGCCGCAA	<i>Acc65I</i>
PrtICTDrev2	GCGCGGTACCCGTTTCTTCATCAGCCGCAAGG	<i>Acc65I</i>
PrtRNTDfor	GCGCTCTAGAGATGAACACGCCTCCGAACGA	<i>XbaI</i>
PrtRNTDrev	GCGCGGTACCTCACCGCATCCGCGAGCGTTC	<i>Acc65I</i>
PrtRNTDrev2	GCGCGGTACCCGCCGCATCCGCGAGCGTTC	<i>Acc65I</i>
Self-complimentary oligonucleotides for cloning		
PBCAM0847For	AGCTTACGTCGCGACTCGGTATTTACCCTGATTTGT CGTGCGCATGTCGATCGGGCCATCTTCC	<i>HindIII</i>
PBCAM0847For2	GCCCGTCGTCGGATAAGGCGCCGCGAACGGGC GCCGCCTTCCACCGACGACAAGGAGTAAACGG	<i>BamHI</i>
PBCAM0847Rev	TTATCCGACGACGGGCGGAAGATGGCCCGATCG ACATGCGCCACGACAAATCAGGGTAAATACCGAG TCGCGACGTA	<i>HindIII</i>
PBCAM0847Rev2	GATCCGTTTACTCCTTGTGTCGCTGGAAGGCGG CGCCGTTCCGCGGCGCC	<i>BamHI</i>
PBCAM2751for	AGCTTCCGGCTCGCACGCCGCGCCCTTTTTTTG AAAAAATGCGGCGGGCGG	<i>HindIII</i>
pBCAM2751for2	TGTCACAGGCGGCGAGCGTGAACCGTCTAGTCGC ATACGGACACCACGATTCCAGG	<i>BamHI</i>
pBCAM2751rev2	GATCCCTGGAATCGTGGTGTCCGTATGCGACTAGA CGGTT	<i>BamHI</i>
PBCAM2751rev	CAGCGTGCCGCTGTGACACCGCCCGCCGATTT TTTCAAAAAAGGGCGCGGCGGCGTGCGAGCCGG A	<i>HindIII</i>
pBCAL3487for	AGCTTCCGCTTCGGCGGTCCAAAAATTTTTTGCATC GA	<i>HindIII</i>

PBCAL3487for2	TGTCACACACGCGGGGCGCTCGCTCGTCATATCAG CGGAACCCACACGAAAGG	<i>Bam</i> HI
PBCAL3487rev	AGCGCCCCGCGTGTGTGACATCGATGCAAAAATT TTTGGACCGCCGAAGCGGA	<i>Hind</i> III
PBCAL3487rev2	<u>GATCCCTTT</u> CGTGTGGGTTCCGCTGATATGACGAG CG	<i>Bam</i> HI
Pfprds1For	<u>AGCTTAGGGT</u> AAAATAACGAGTTCCTCAAACGTCT GTCATCGAGAAGG	<i>Hind</i> III
Pfprds1Rev	<u>GATCCCTT</u> CTCGATGACAGACGTTTGAGGAACTCG TTATTTTACCCTA	<i>Bam</i> HI
prtIproFor2	<u>AGCTTGGAGGGA</u> AATAATCGATGCCGGCCGGCGTC CTACGCGGCATCCAG	<i>Hind</i> III
prtIproRev2	<u>GATCCTGGAT</u> GCCGCGTAGGACGCCGGCCGGCAT CGATTATTCCCTCCA	<i>Bam</i> HI
prtIproFor	<u>AGCTTGTTTTT</u> TCGTATGCTGGAGGGAATAATCGA TGCCGGCCGGCGTCTACGCGGCATCCACGACCG	<i>Hind</i> III
prtIproRev	<u>GATCCGGT</u> CGTGGATGCCGCGTAGGACGCCGGCC GGCATCGATTATTCCCTCCAGCATAACGAAAAACA	<i>Bam</i> HI
PcatFor	<u>AGCTTCGAAAAA</u> ACTTTTCGTGGAATAAAGCGGCG CGGGCGGCGTCTACGCGGTTGACAATGATG	<i>Hind</i> III
PcatRev	<u>GATCCATCATT</u> GTCGAACGCGCGTAGGACGCCGCC CGCGCCGCTTATTCCACGAAAAGTTTTTTTCGA	<i>Bam</i> HI
Ptacfor	<u>GTACCCCGTT</u> CTGGATAATGTTTTTTCGCGCGACAT CATAACGGTTCTGGCAAATATTCTGAAA	<i>Hind</i> III
Ptacrev	ATTAATTGTCAACAGCTCATTTTCAGAATATTTGCCA GAACCGTTATGATGTCGGCGCAAAAACATTATCC AGAACGGG	<i>Hind</i> III
Ptacfor2	TGAGCTGTTGACAATTAATCATCGGCTCGTATAAT GTGTGGAATTGTGAGCGGATAACAATTTACACAC	<i>Bam</i> HI
Ptacrev2	TCGAGTGTGTGAAATTGTTATCCGCTCACAATTCCA CACATTATACGAGCCGATG	<i>Bam</i> HI
Ptacshortfor	GTACCGAAATGAGCTGTTGACAATTAATCATCGGC TCGTATAATGTGTGGAATTGC	<i>Hind</i> III
Ptacshortrev	TCGAGCAATTCCACACATTATACGAGCCGATGATT AATTGTCAACAGCTCATTTTCG	<i>Bam</i> HI
PprtiDs1for	<u>AGCTTGGAT</u> GGAATAATCGATGCCGGCCGGCGTC CTACGCGGCATCCAG	<i>Hind</i> III
PprtiDs1rev	<u>GATCCTGGAT</u> GCCGCGTAGGACGCCGGCCGGCAT CGATTATTCCATCCA	<i>Bam</i> HI
PprtiDs2for	<u>AGCTTGGAGT</u> GAATAATCGATGCCGGCCGGCGTC CTACGCGGCATCCAG	<i>Hind</i> III
PprtiDs2rev	<u>GATCCTGGAT</u> GCCGCGTAGGACGCCGGCCGGCAT CGATTATTCACTCCA	<i>Bam</i> HI
PprtiDs3for	<u>AGCTTGGAGGTA</u> AATAATCGATGCCGGCCGGCGTC CTACGCGGCATCCAG	<i>Hind</i> III

PprtiDs3rev	<u>GATCCTGGATGCCGCGTAGGACGCCGGCCGGGCAT</u> CGATTATTACCTCCA	<i>Bam</i> HI
PprtiDs4for	<u>AGCTTGGAGGGCATAATCGATGCCGGCCGGCGTC</u> CTACGCGGCATCCAG	<i>Hind</i> III
PprtiDs4rev	<u>GATCCTGGATGCCGCGTAGGACGCCGGCCGGGCAT</u> CGATTATGCCCTCCA	<i>Bam</i> HI
PprtiDs5for	<u>AGCTTGGAGGGACTAATCGATGCCGGCCGGCGTC</u> CTACGCGGCATCCAG	<i>Hind</i> III
PprtiDs5rev	<u>GATCCTGGATGCCGCGTAGGACGCCGGCCGGGCAT</u> CGATTAGTCCCTCCA	<i>Bam</i> HI
PprtiDs6for	<u>AGCTTGGAGGGAAGAATCGATGCCGGCCGGCGTC</u> CTACGCGGCATCCAG	<i>Hind</i> III
PprtiDs6rev	<u>GATCCTGGATGCCGCGTAGGACGCCGGCCGGGCAT</u> CGATTCTTCCCTCCA	<i>Bam</i> HI
PprtiDs7for	<u>AGCTTGGAGGGAATCATCGATGCCGGCCGGCGTC</u> CTACGCGGCATCCAG	<i>Hind</i> III
PprtiDs7rev	<u>GATCCTGGATGCCGCGTAGGACGCCGGCCGGGCAT</u> CGATGATTCCCTCCA	<i>Bam</i> HI
PprtiDs8for	<u>AGCTTGGAGGGAATACTCGATGCCGGCCGGCGTC</u> CTACGCGGCATCCAG	<i>Hind</i> III
PprtiDs8rev	<u>GATCCTGGATGCCGCGTAGGACGCCGGCCGGGCAT</u> CGAGTATTCCCTCCA	<i>Bam</i> HI
PprtiDs9for	<u>AGCTTGGAGGGAATAAGCGATGCCGGCCGGCGTC</u> CTACGCGGCATCCAG	<i>Hind</i> III
PprtiDs9rev	<u>GATCCTGGATGCCGCGTAGGACGCCGGCCGGGCAT</u> CGCTTATTCCCTCCA	<i>Bam</i> HI
PprtiDs9.1for	<u>AGCTTGGAGGGAATAATAGATGCCGGCCGGCGTC</u> CTACGCGGCATCCAG	<i>Hind</i> III
PprtiDs9.1rev	<u>GATCCTGGATGCCGCGTAGGACGCCGGCCGGGCAT</u> CTATTATTCCCTCCA	<i>Bam</i> HI
PprtiDs10for	<u>AGCTTGGAGGGAATAATCGATGCCGGCCGTCGTC</u> CTACGCGGCATCCAG	<i>Hind</i> III
PprtiDs10rev	<u>GATCCTGGATGCCGCGTAGGACGACGGCCGGGCAT</u> CGATTATTCCCTCCA	<i>Bam</i> HI
PprtiDs11for	<u>AGCTTGGAGGGAATAATCGATGCCGGCCGGAGTC</u> CTACGCGGCATCCAG	<i>Hind</i> III
PprtiDs11rev	<u>GATCCTGGATGCCGCGTAGGACTCCGGCCGGGCAT</u> CGATTATTCCCTCCA	<i>Bam</i> HI
PprtiDs12for	<u>AGCTTGGAGGGAATAATCGATGCCGGCCGGCTTC</u> CTACGCGGCATCCAG	<i>Hind</i> III
PprtiDs12rev	<u>GATCCTGGATGCCGCGTAGGAAGCCGGCCGGGCAT</u> CGATTATTCCCTCCA	<i>Bam</i> HI
PprtiDs13for	<u>AGCTTGGAGGGAATAATCGATGCCGGCCGGCGGC</u> CTACGCGGCATCCAG	<i>Hind</i> III

PprtiDs13rev	<u>GATCCTGGATGCCGCGTAGGCCGCCGGCCGGGCAT</u> CGATTATTCCTCCA	<i>Bam</i> HI
PprtiDs14for	<u>AGCTTGGAGGGAATAATCGATGCCGGCCGGCGTA</u> CTACGCGGCATCCAG	<i>Hind</i> III
PprtiDs14rev	<u>GATCCTGGATGCCGCGTAGTACGCCGGCCGGGCAT</u> CGATTATTCCTCCA	<i>Bam</i> HI
PprtiDs15for	<u>AGCTTGGAGGGAATAATCGATGCCGGCCGGCGTC</u> ATACGCGGCATCCAG	<i>Hind</i> III
PprtiDs15rev	<u>GATCCTGGATGCCGCGTATGACGCCGGCCGGGCAT</u> CGATTATTCCTCCA	<i>Bam</i> HI

2.2.3 Plasmid extraction using a nucleic acid isolation kit

GeneJet plasmid mini prep kit (Thermo Scientific) was used according to the manufacturer's instructions. When the DNA was required to be sequenced, a sample of 10 µl (100 ng/µl) plasmid per primer was sent to the Core Genomics Facility, University of Sheffield.

2.2.4 DNA electrophoresis

Depending on the size of DNA samples to be electrophoresed, an 0.8-1.5% agarose gel was prepared in X1 TAE buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA, pH 8.0). The mixture was boiled in a microwave oven until the agarose was completely melted. Midori Green dye (Nippon Genetics Europe GmbH) was added to the gel at a final concentration of 0.005% and dissolved which was then poured into a cast to set. The DNA samples were combined with an equal volume of 2X DNA loading dye (4.2 mg/ml bromophenol blue, 5.5% (v/v) glycerol) and loaded on the gel. For plasmid DNA, 1 µl Supercoiled DNA ladder (Biolabs) and for linear DNA pre-stained DNA ladder mix (Thermo Scientific) was used. As an alternative to using Midori Green, the gel was stained in an ethidium bromide solution (5 µg/ml in ddH₂O) for 30 minutes with gentle shaking following electrophoresis. The gel was then washed in H₂O to remove excess

ethidium bromide. DNA was visualized on a UV transilluminator and the gel image was captured using U:Geneius3 (Syngene).

2.2.5 Restriction digestion

Promega/New England Biolabs restriction enzyme and buffers were used according to manufacturers' instructions. Plasmid DNA/PCR product (> 200 ng) along with the appropriate buffer, one or more restriction enzymes, BSA (0.1 mg/ml) and ddH₂O were mixed in a final reaction volume of 50 µl. Restriction digestion was carried out at the recommended temperature for ≤ 2 hours after which the sample was purified if required for a subsequent digestion. For two restriction enzymes working optimally in different buffers, digestion was carried out using one of the enzymes in its optimal buffer, followed by purification and then a digestion using the second enzyme in its optimal buffer.

2.2.6 Ligation

Promega ligation buffer and enzyme was used according to manufacturers' instructions where the DNA to be inserted and plasmid DNA were mixed in a 5:1 molar ratio along with T4 ligase buffer, ligase enzyme and ddH₂O in a final reaction mix of 30 µl. As a vector control to assess whether the plasmid has been completely cut, DNA to be inserted and ligase enzyme were not added. As a ligation control to assess self-ligation of the plasmid and whether the ligation conditions were appropriate, DNA to be inserted was not added. Ligation and control samples were incubated at room temperature for approximately 16 hours. Competent *E. coli* cells were transformed with ligation samples as described in Section 2.2.12.

2.2.7 Phosphatase treatment

In some cases, to prevent re-circulation or re-ligation of plasmid cut with restriction enzymes, calf intestinal alkaline phosphatase (Promega) treatment to remove phosphate groups from both 5' ends of the vector was carried out. At least 10 pmol

plasmid DNA cut with restriction enzyme(s) was combined with alkaline phosphatase buffer and 0.01 units of alkaline phosphatase as per the manufacturer's instructions. This was incubated at 37°C for 30 minutes after which another aliquot of alkaline phosphatase was added and the sample was incubated for a further for 30 minutes at 37°C. It was then purified and used for further cloning procedures.

2.2.8 Annealing of complementary oligonucleotides

45µl of each of two oligonucleotides (100 µM each) or 22.5µl of each of four oligonucleotides (100 µM each) and 10 µl of x10 annealing buffer [10 mM Tris-HCl (pH 7.5), 50 mM NaCl, 1 mM EDTA] were combined in a PCR tube and incubated at 90°C for 10 minutes. After cooling the mixture at room temperature for one hour to anneal the oligonucleotides, the tube was stored at -20°C until use.

2.2.9 DNA purification

2.2.9.1 Spin column

Qiagen/Thermo Scientific DNA purification kits were used according to manufacturer's instructions to inactivate the enzyme and remove of unwanted reagents/buffer components from DNA after restriction digestion or PCR amplification for downstream steps.

2.2.9.2 Gel extraction

To purify DNA fragments from an agarose gel, the required DNA fragment was visualized by placing the agarose gel on a UV transilluminator on low power and a gel slice containing the DNA fragment was cut using a clean scalpel quickly enough to minimize damage of the DNA due to UV radiation. The required DNA fragment was purified from the gel slice using Qiagen/Macherey-Nagel gel extraction kit according to the manufacturer's instructions.

2.2.10 Amplification of DNA by the polymerase chain reaction

2.2.10.1 Standard PCR for cloning

For PCR amplification where the resulting amplified product was to be cloned into a vector, KOD polymerase or Q5 polymerase was used since they have proofreading activity. To prepare template DNA, one CFU of the required strain was re-suspended in 200 μ l TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA). This was boiled for 10 minutes followed by centrifugation at 14,000 rpm for 5 minutes. The supernatant was transferred to another tube and used as template DNA. To amplify the required region of DNA, the components added to the PCR are given in Table 2.6

Table 2.6: Components added in PCR amplification

Component	Volume (μ l) (KOD)	Volume (μ l) (Q5)
Template DNA (colony lysate)	3	3
x10 reaction buffer	5	5-10
DMSO	2.5	0
25 mM MgSO ₄	4	0
GC enhancer	0	10
10 μ M Forward primer	3	2.5
10 μ M Reverse primer	3	2.5
2mM each dNTPs (dNTP mix)	5	5
KOD/Q5 polymerase	0.5 (2 units/ml)	0.5 (2 units/ml)
Sterile ddH ₂ O	To make up final volume to 49.5	To make up final volume to 49.5
Total:	50	50

The DNA was subjected to PCR amplification using the thermocycler 'Expand High Fidelity PCR System' by Roche or the 'T100 thermal cycler' by Bio-rad. The programme was initiated using a 2 minute at 95°C step to heat the lid. This was followed by 30 cycles

of reaction where each cycle consisted of denaturation/separation of primers (95°C for 30 seconds), annealing of DNA strands (for a period of 30 seconds and at an appropriate temperature) and elongation at 70°C for a duration of 30 seconds per 500 bases of DNA for KOD polymerase or at 72°C for 30 seconds per 1000 bases of DNA for Q5 polymerase.

Calculation of annealing temperature:

Annealing temperature of primer = $[4(G+C) + 2(A+T)] - 5^{\circ}\text{C}$

where,

G+C = Total number of guanine and cytosine bases in the complementary region of the primer

A+T = Total number of adenine and thymidine bases in the complementary region of the primer

After completion of the PCR, 5 μ l of the PCR product was electrophoresed in an agarose gel.

2.2.10.2 PCR mutagenesis - Splice overlap extension

In order to construct markerless in-frame deletion mutants allelic replacement was carried out. Regions upstream (AB) and downstream (CD) of the region to be deleted were separately amplified using a high-fidelity DNA polymerase in the first PCR step (section 2.2.10.1). The primers B and C were designed such that they had complementary 5' tails so that in the second PCR step when the two amplified products were used as template in equimolar concentration, using primers A and D, the final product consisted of the deleted gene as shown in Figure 2.1. The product was electrophoresed in an agarose gel and passed through a clean-up column prior to ligation into a suitable plasmid.

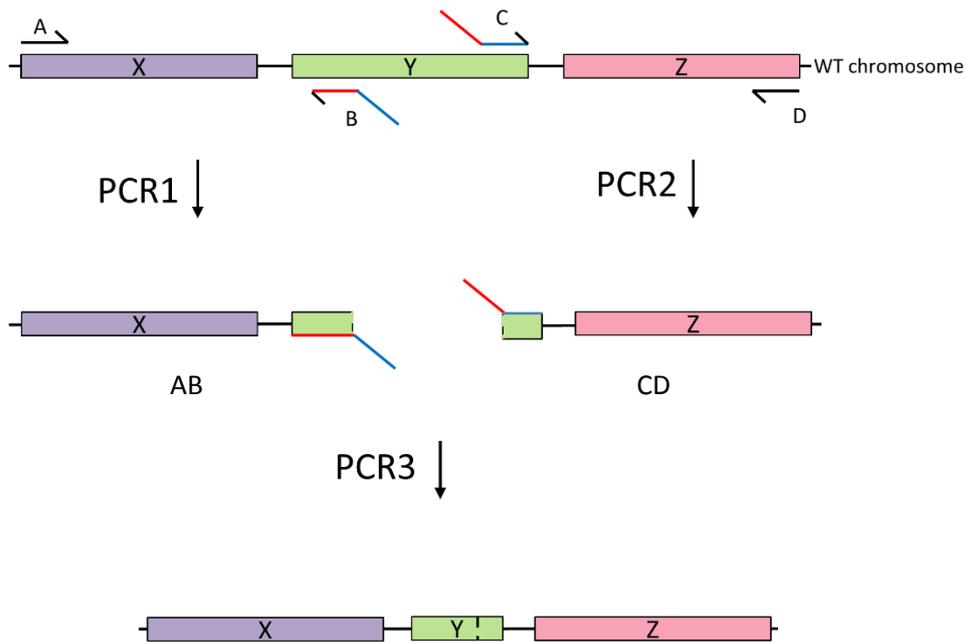


Figure 2.1 Principle of SOE PCR deletion mutagenesis

Fragment AB was generated using primers A and B while fragment CD was generated using primers C and D. Products AB and CD were used as template in a subsequent PCR, PCR3, using primers A and D resulting in the gene lacking the central portion because primers B and C were designed such that they had overlapping complementary bases indicated by red and blue colours.

2.2.10.3 Screening for recombinant plasmids or genomic mutants using colony PCR

When screening for the presence of the required recombinant plasmid among transformants or replacement of chromosomal genes by mutant copies, the PCR product was not to be used further. Hence there was no requirement for a thermostable DNA polymerase with proofreading activity. A sterile toothpick was touched to the colony to be screened and smeared on the inside of a PCR tube. The components added in order to set up a PCR reaction are given in Table 2.7

Table 2.7: Components added to colony screening PCR

Component	Volume (μl)
x10 Taq buffer	2.5
DMSO	1.25
50 mM MgSO ₄	1
10 μ M Forward primer	1.5
10 μ M Reverse primer	1.5
2.5 mM each dNTPs (dNTPs mix)	0.5
ddH ₂ O	16.5
Taq DNA polymerase	0.25
Total:	25

The DNA was subjected to PCR amplification by using the thermocycler 'Expand High Fidelity PCR System' by Roche or the 'T100 thermal cycler' by Bio-rad. The programme was initiated using an "Initial denaturation" step for 5 minutes at 95°C to heat the lid. This was followed by 30 cycles of amplification. Each cycle consisted of denaturation/separation of DNA strands at 95°C for 30 seconds, annealing of DNA strands for a period of 30 seconds at a temperature appropriate for the polymerisation of primers and elongation at 72°C for duration of 30 seconds per 500 bases of DNA. After completion of the PCR, 10 μ l of the PCR product was electrophoresed on an agarose gel.

2.2.10.4 Reverse transcriptase (RT)-PCR

RT-PCR was used to determine whether two adjacent genes were co-transcribed. The bacterial strains of interest were grown overnight at 37°C in LB broth containing appropriate antibiotics. The following day the OD₆₀₀ was measured and an amount of culture containing approximately 10⁹ bacterial CFU assuming OD₆₀₀ of 1.0 = 8 x 10⁸ cells/ml. was used to extract mRNA using an mRNA extraction kit (Machery-Nagel) according to the manufacturer's instructions. cDNA was prepared by combining 5 μ l mRNA with 0.5 μ l 10 μ M reverse primer. The tubes were incubated at 70°C for 5 minutes and quickly cooled on ice for 5 minutes. Then, using the Promega MLV RT enzyme kit, buffer, dNTPS, enzyme and DMSO were added according to the manufacturer's

instructions and incubated at room temperature for 10 minutes followed by 55°C for one hour. The tubes were further incubated at 70°C for 15 minutes to inactivate the reaction. The resultant cDNA (5 µl) was used as template to carry out a standard PCR as described in section 2.2.10.1. PCR using mRNA template as a control was carried out to check for DNA contamination.

2.2.11 Techniques for plasmid transfer

2.2.11.1 Competent cell preparation (Hanahan's method)

The required *E. coli* strain was inoculated from a fresh colony in 5ml of LB medium supplemented with the appropriate concentration of antibiotic. This was incubated overnight at 37°C with shaking so that the cells were grown to stationary phase. The following day, 0.5 ml of the overnight culture was inoculated to 50 ml LB broth and incubated in a 37°C shaker incubator until the cells reached exponential phase ($OD_{600} = 0.3-0.5$). The culture was then chilled on ice for 15 minutes. The cells were harvested by centrifugation at 4,000 rpm for 10 minutes at 4°C in a refrigerated bench top centrifuge whereupon the supernatant was discarded and the cell pellets re-suspended by moderate vortexing in 16 ml ice cold RF1 solution. The cells were then incubated on ice for 30 minutes followed by centrifugation for 10 minutes at 4,000 rpm at 4°C. The supernatant was discarded and the cell pellets were carefully re-suspended in 4 ml RF2 solution and chilled on ice for 15 minutes. The competent cells were then aliquoted in 225 µl or 425 µl volumes into Eppendorf tubes and stored frozen at - 80 °C until use.

RF1: 7.46 g KCl, 9.90 g $MnCl_2 \cdot 4H_2O$, 2.94 g potassium acetate, 1.50 g $CaCl_2 \cdot 2H_2O$ and 150 ml glycerol were dissolved in ~750 ml ddH₂O, and the resultant solution was adjusted to pH 5.8 using 0.2 M acetic acid (glacial acetic acid is 17.4 M). The final volume was then made up to 1000 ml by adding ddH₂O. The solution was filter sterilized through a 0.22 µm membrane, and stored at 4°C.

RF2: Solution A: 0.5 M MOPS, pH 6.8

Solution B: 10 mM KCl, 75 mM $CaCl_2 \cdot 2H_2O$ and 15% (w/v) glycerol.

Both solutions A and B were made up in deionised H₂O, and stored at 4°C. RF2 was prepared by combining RF2A and RF2B in a 2:98 ratio.

2.2.11.2 Transformation

Plasmid DNA or ligation reactions to be transformed were mixed with 100 µl ice-cold competent cells and chilled on ice for 30 minutes, flicking the micro-centrifuge tubes every few minutes to avoid settling of cells to the bottom of the tube. A tube containing only competent cells was used as a control. The cells were given a heat shock at 42°C for a period of 150 seconds by placing the tubes in a water bath. The tubes were subsequently transferred to ice for 5-6 minutes and then 1 ml LB broth was added to each tube. The transformation mixtures were incubated at 37°C for one hour to allow expression of the antibiotic resistance gene contained in the plasmid. 100 µl of the transformed cells was spread on selective media. The plates were inverted and incubated overnight at 37°C for colonies to form.

2.2.11.3 Conjugation

Day 1: A single bacterial colony of the donor strain or the recipient strain were inoculated into separate universals containing 5 ml LB supplemented with appropriate antibiotics and the cultures were grown overnight in a shaker incubator at 37°C.

Day 2: 1 ml of each overnight culture was centrifuged at 13,500 rpm for 2 minutes. The supernatant was discarded and the cell pellets were each re-suspended in 120 µl 0.85 % (w/v) sterile saline. Using sterile forceps, 0.22 µm nitrocellulose membrane filters were placed on LB agar (one per plate). As controls, 25 µl of the donor and recipient cultures were separately mixed with 25 µl sterile saline each and for the conjugation, 25 µl of donor strain and 25 µl of recipient strain were mixed together. Each of the cell suspensions were pipetted directly onto the filters and spread over the membrane using a pipette tip. When the filters had dried, the plates were inverted and incubated at 37°C for 8-16 hours.

Day 3: Using sterile forceps, the filters with bacterial growth were transferred to sterile universal bottles containing 3 ml sterile saline, and the bottles were vortexed to re-

suspend the bacteria. 10^{-1} , 10^{-2} and 10^{-3} dilutions of the resultant cell suspension were spread onto selective media and incubated at 37°C for 48 to 72 hours.

Day 5/6: When individual colonies had attained a large enough size, a number of them were either patched onto selective media or streaked to give single colonies.

2.3 Recombinant protein overproduction and purification techniques:

(Taken from Helena Spiewak, PhD thesis, 2016, with minor amendments)

2.3.1 Protein induction protocol (small scale)

A single bacterial colony per strain was used to prepare an overnight culture. The following day, 8 – 10 ml of LB broth was inoculated with 80-100 µl of the bacterial culture including appropriate antibiotics and grown at 37°C until the OD₆₀₀ reached 0.3-0.5. A 1 ml aliquot was removed as an un-induced control. To the remaining bacterial culture 1µl/ml of 1 M IPTG was added and the culture was allowed to grow at 37°C for 3 hours. The cultures were then kept on ice and the OD₆₀₀ of the induced culture was measured. An appropriate volume of induced culture equivalent to the OD₆₀₀ value of the un-induced sample was centrifuged at 13,000 rpm for 3 mins to harvest the bacteria. The supernatant was discarded and the cell pellet re-suspended in 50 µl of 2x Laemmli buffer (125 mM Tris-HCl (pH 6.8), 4% SDS, 20% glycerol 0.004% bromophenol blue, 10% β-mercaptoethanol). The samples were boiled at 100°C in a heat block for 10 mins and then centrifuged at 13,000 rpm for 10 mins to remove cell debris. Samples thus prepared were visualised following electrophoresis in a SDS PA-gel.

2.3.2 Determination of overproduced protein solubility

A single bacterial colony per strain was used to prepare an overnight culture. The following day, two sterile flasks per strain were used to inoculate 25 ml of LB broth with 250 µl of the bacterial culture including appropriate antibiotics and grown at 37°C until OD₆₀₀ reached 0.3-0.5. IPTG was added to the final concentration of 1 mM to one of the flasks (induced culture) and the other flask was un-induced. The cultures were allowed to grow for a further 3 hours at 30°C/37°C. The cells were harvested at 12,500 rpm for

20 minutes at 4°C and then washed with 20 ml/g bacterial cells of wash buffer (25 mM Tris-HCl (pH 7.5), 150 mM NaCl, 2 mM EDTA). The cells were re-suspended in 5 ml/g bacterial cells of cell resuspension buffer (50 mM Tris-HCl (pH 8), 200 mM NaCl, 2 mM EDTA, 5% glycerol). 50 µl of culture was set aside as 'whole cell sample'. Lysozyme was then added to a final concentration of 200 µg/ml and the suspension was gently mixed at 4°C for 30 minutes on a roller. PMSF and sodium deoxycholate were then added at final concentrations of 25 µg/ml and 500 µg/ml respectively and the suspension was gently mixed at 4°C for 30 minutes on a roller. The remainder was sonicated for 30 seconds (20% amplitude) with 2 minute intervals until the cell suspension was clear using a Sonics Vibracell VCX750 Ultrasonic Cell Disrupter with a microtip probe. 50 µl was set aside as 'crude lysate' and remainder of the lysate was centrifuged at 12,500 rpm for 30 minutes at 4°C to remove insoluble material. The 'supernatant', containing soluble proteins, was retained and the pellet was discarded. An equal volume of 2X Laemlli buffer was to the 'whole cell sample', 'crude lysate' and 'supernatant' from both induced and un-induced cultures and proteins were visualised following electrophoresis in a SDS PA gel. If the protein of interest was soluble, it was observed in the induced 'whole cell sample', 'crude lysate' and 'supernatant'. If the protein of interest was insoluble, it was present only in the induced 'whole cell sample' and 'crude lysate'.

2.3.3 Purification of His-tagged proteins by nickel affinity chromatography column using FPLC

A single bacterial colony per strain was used to prepare an overnight culture. The following day, 200 ml of LB broth was inoculated with 2 ml of the bacterial culture including appropriate antibiotics as needed and grown at 37°C until an OD₆₀₀ of 0.3-0.5. IPTG was added at a final concentration of 1 mM to induce the cultures and were grown for 3 hours at 30°C/37°C. The cells were harvested at 12,500 rpm for 20 minutes at 4°C and then washed with 20ml/g bacterial cells of wash buffer (25 mM Tris-HCl (pH 7.5), 150 mM NaCl, 2 mM EDTA). The cells were re-suspended in 5 ml/g bacterial cells of low-imidazole buffer (50 mM Tris-HCl (pH 8), 200 mM NaCl, 10 mM imidazole, 5% glycerol). Lysozyme was then added to a final concentration of 200 µg/ml and the suspension was

gently mixed at 4°C for 30 minutes on a roller. PMSF and sodium deoxycholate were then added at final concentrations of 25 µg/ml and 500 µg/ml respectively and the suspension was gently mixed at 4°C for 30 minutes on a roller. The suspension was sonicated for 30 seconds (20% amplitude) with 2 minute intervals until clear using a Sonics Vibracell VCX750 Ultrasonic Cell Disrupter with a microtip probe. The cell suspension was centrifuged at 12,500 rpm for 30 minutes at 4°C to remove insoluble material. The 'supernatant' was retained and the pellet was discarded. A 1 ml HisTrap HP pre-packed nickel sepharose column (GE healthcare) was washed with 10 ml of ddH₂O and then equilibrated with 10 ml of low imidazole buffer at the rate of 1ml per minute using an ÄKTA purifier 10 FPLC purification system (GE Healthcare) controlled by UNICORN control software (GE Healthcare). The 'supernatant' containing His-tagged protein was manually loaded onto the nickel column using a syringe driven 0.22 µM filter and 50 µl of the 'flow through' was kept aside. The column was then washed with 10-13 ml of low imidazole buffer at the rate of 1 ml per minute to remove non-specifically bound and unbound proteins and 50 µl of eluate was kept aside ('wash'). The protein of interest was then eluted from the column using high imidazole buffer (50 mM Tris-HCl (pH 8), 200 mM NaCl, 500 mM imidazole, 5% glycerol) in 1 ml 'fractions' at a rate of 1 ml per minute over 10-15 minutes using a Frac-950 fraction collector (GE Healthcare). The absorbance at 280 nm was monitored using the Monitor UV-900 (GE healthcare) to identify a peak that indicated in which fractions the protein had eluted. 2X Laemlli buffer was added in an equal concentration with the 'supernatant', 'flow through', 'wash' and 'eluted fractions' and proteins were visualised following electrophoresis in a SDS PA gel. The concentration of purified protein was determined using a Bradford's assay (2.4.3).

2.3.4 Protein storage

For long term storage, eluted protein was dialysed against storage buffer (50 mM Na HEPES (pH 7.8), 5% glycerol, 0.1 mM EDTA, 0.1 mM DTT, 50 mM NaCl). An appropriate length of 3000 MWCO dialysis tube was cut and washed in ddH₂O. Protein fractions were transferred to the tube secured with dialysis clips and dialysed in a volume of storage

buffer that was 250 times larger than the protein fraction, overnight at 4°C with continuous stirring.

2.4 Protein separation and quantification techniques:

(Taken from Helena Spiewak, PhD thesis, 2016, with minor amendments)

2.4.1 SDS PAGE

A 12-15% SDS-PA gel (components described in Table was cast as standard protocol (Laemmli, 1970) using Tetra Cell system (Bio-rad). The resolving gel was poured into the cast, allowed to set followed by adding the stacking gel where a comb was inserted such that no bubbles formed. Once the gel was set, the comb was removed and samples in Laemmli buffer, in a 1:1 ratio, were loaded onto the gel, including 6 µl E.Z-Run pre-stained *Rec* protein ladder (Fisher). The gel was electrophoresed at 120 V for 1 hour in Tris-glycine SDS-PAGE running buffer (25 mM Tris, 192 mM Glycine, 0.1% SDS, pH=8.3)

Table 2.8 SDS-Polyacrylamide gel ingredients

10% Resolving Gel (x4)	12% Resolving Gel (x4)	15% Resolving Gel (x4)	5% Stacking Gel (x4)
ddH ₂ O	8.6	7.1	6.3
40% acrylamide: bis-acrylamide 37.5:1	6	7.5	1.25
1.5 M Tris-HCl (pH 8.8)	5	5	-
1 M Tris-HCl (pH 6.8)	-	-	1.25
10% (w/v) SDS	0.2	0.2	0.1
10% (w/v) ammonium persulfate	0.2	0.2	0.1
TEMED	0.01	0.01	0.005

To visualise proteins, coomassie blue solution (50% (v/v) methanol, 10% (v/v) acetic acid, 2.5g/L Coomassie Brilliant Blue R-250) was added to the gels, heated on full power for 30 seconds in the microwave and stained for 15 minutes by gently shaking. The gels

were destained with destain solution (40% (v/v) methanol, 10% (v/v) acetic acid) to reveal bands of protein.

2.4.2 Western blotting

After SDS-PA gel electrophoresis was completed, the gel was electro-blotted onto a 0.45 µm PVDF membrane (Millipore), using a Mini-trans blot cell (Bio-Rad), at 100 V for 1 hour according to the manufacturer's instructions. Membranes were blocked by incubation with 5% (w/v) skimmed milk powder (Sigma) in Tris-buffered saline (TBS) (8g/l NaCl, 0.2g/l KCl and 3g/l Tris base, pH =7.5) + 0.05% Tween-20 (Fisher) blocking solution for 1 hour at room temperature. They were then incubated with primary antibody diluted in blocking solution at 4°C overnight on a rolling mixer in a plastic tube. Membranes were subjected to four 10 minute washes in TBS + 0.05% Tween-20, then incubated for 1 hour at room temperature with HRP conjugated secondary antibody diluted in blocking solution, followed by four more 10 minute washes in TBS + 0.05% Tween-20, and a final wash in TBS. Fluorescence was then detected using EZ-ECL (Biological Industries) with a ChemiDoc XRS+ System Imager and Image Lab software (Bio-rad).

Antibodies:

Primary antibody	Secondary antibody
Polyclonal rabbit anti-FLAG (Sigma) at 1:5000 dilution	HRP labelled goat anti-rabbit IgG (Vector) at 1:6000 dilution
Monoclonal mouse anti-hexa-histidine tag (Berkeley antibody company) at 1:1000 dilution	HRP labelled rabbit anti-mouse IgG (Thermo Fisher) at 1:4000 dilution

2.4.3 Protein quantification assay

To determine protein concentration after purification the Bradford assay was used employing the 'Bio-Rad protein assay kit-I'. Protein assay dye reagent concentrate (Bio-Rad) was diluted 4 fold in ddH₂O and filter sterilised using a 0.22 µM syringe drive filter unit before performing the assay. Protein standards of 0.2, 0.4, 0.6, 0.8 and 1.0 mg/ml

were prepared in Tris buffer (50 mM Tris-HCl pH 8.0) using a stock BSA solution of 10 mg/ml (Promega). The sample of unknown protein concentration was diluted 1:1 and 1:3 in a final volume of 50 μ l. 20 μ l of protein standard/sample was added to 1 ml protein assay dye reagent in a disposable cuvette (Fisher), mixed and incubated at room temperature for 5 minutes in technical duplicates. The optical density (OD) of the sample at 595 nm was measured using a spectrophotometer blanked with protein assay dye reagent and solvent without protein. A standard curve was plotted of protein concentration standards against OD₅₉₅ using which the concentration of the unknown protein solution was calculated.

2.5 Enzyme assays

2.5.1 β -galactosidase assay

(Taken from K. Agnoli PhD thesis, 2007, with minor amendments)

Each bacterial strain to be assayed was inoculated from a fresh colony in 3 ml of liquid medium supplemented with the appropriate antibiotics in biological triplicates. These were incubated overnight at 37°C with shaking so that the cells were grown to stationary phase. The following day, 50 μ l of the overnight bacterial culture was added to 5 ml (i.e. 1:100 dilution) fresh liquid medium containing appropriate antibiotics and media supplements. The cultures were incubated at 37°C with shaking until cells reached exponential phase, i.e. OD₆₀₀ = 0.35-0.6 where sterile medium was used as blank. The tubes were transferred to ice for at least 30 minutes before carrying out the assay and OD₆₀₀ of each culture was recorded. Depending upon the promoter activity of the plasmids to be assayed, 25-100 μ l of the culture was used for the assay. Z buffer containing β -mercaptoethanol (1 ml – volume of culture added) was added to a clean 13 ml glass tube such that each biological replicate was done in technical duplicates (therefore, each strain = 6 tubes). 30 μ l chloroform was added directly on to the surface of the buffer carefully to make a dome at the bottom of the tube, avoiding the evaporation of chloroform. 30 μ l 0.1% SDS was added to each tube carefully from the side and then appropriate amount of culture was added carefully from the side of the

tube to prevent the disturbance of the chloroform bubble. The tubes were vortexed for 10 seconds for permeabilization of cells. The tubes were immersed in a 30°C water bath for a period of 15 minutes to allow the contents to equilibrate to this temperature before starting the assay. 200 µl of 4 mg/ml ONPG was added to a tube which was then vortexed for 1 second to initiate the reaction. This was repeated at 30 second intervals for each tube of the same strain. When the appropriate yellow colour was achieved (OD₄₂₀ = 0.2-0.8) the reaction was stopped by adding 500 µl of 1 M Na₂CO₃. Reactions of the same strain-plasmid combination were stopped in intervals of 30 seconds so that each biological replicate was assayed for the same length of time. A control tube containing sterile medium and appropriate media supplements was also treated the same way. The absorbance of each reaction was recorded at wavelengths 420 nm and 550 nm. The control tubes were used to set the blank for each of these two ODs.

Calculation of β-galactosidase activity:

$$1 \text{ Miller Unit} = 1000 \times \frac{\text{Abs}_{420} - (1.75 \times \text{Abs}_{550})}{\text{Time (minutes)} \times \text{Volume of cells (ml)} \times \text{Abs}_{600}}$$

where,

Abs₄₂₀: Absorbance of the yellow ortho-nitrophenol

Abs₅₅₀: Light scattering factor for the contribution of the cells to the Abs₄₂₀ value.

Time: Reaction time (stopping time – starting time) in minutes

Volume of cells: Volume of culture assayed in ml

Abs₆₀₀: Cell density

An average value was obtained from technical duplicates for each biological replicate assayed. The mean of the biological triplicate values was calculated for each strain. Each experiment was carried out in experimental replicates.

Z buffer: 16.1 g Na₂HPO₄·7H₂O or 8.53 g Na₂HPO₄ anhydrous, 5.5 g NaH₂PO₄·H₂O, 0.75 g KCl and 0.246 g MgSO₄·7H₂O were dissolved in a final volume of 1000 ml deionized H₂O

and stored at 4°C. 0.27ml β -mercaptoethanol was added per 100 ml of Z buffer on the day the assay was performed.

1M Na_2CO_3 : 10.6 g sodium carbonate was dissolved in 100 ml of ddH₂O and stored at room temperature.

4 mg/ml ONPG: 40 mg ONPG was added to 10 ml Z buffer containing β -mercaptoethanol and vortexed until it was completely dissolved. This was made fresh on the day of the assay.

2.5.2 Catalase assay

The standard catalase assay (Sigma) was followed where the final reaction volume of 3 ml contained 2.90 ml of hydrogen peroxide solution [0.036% (w/w) prepared in phosphate buffer by adding 30% (w/w) hydrogen peroxide], to which 0.1 ml of test enzyme solution (in PBS) was added and immediately mixed. Using phosphate buffer (50 mM potassium phosphate, dibasic, trihydrate pH 7.0) as blank, the time required for the A_{240} to decrease from 0.45 to 0.40 was measured. As a positive control, commercially available catalase (Sigma) was used. The activity was calculated using the formula:

$$\text{Units/ml enzyme} = (3.45) (\text{dilution factor}) / (\text{time}) (0.1)$$

Where, 3.45 = decomposition of 3.45 μM of hydrogen peroxide in a 3.0 ml reaction mixture that causes a decrease of A_{240} from 0.45 to 0.40.

Time = number of minutes required for the decrease of A_{240} from 0.45 to 0.40

0.1 = Amount of enzyme solution added to the final reaction in ml.

2.6 Interaction studies:

2.6.1 BACTH assay (protein-protein interaction)

(Taken from Sayali Haldipurkar, MSc. Thesis, 2012, with minor amendments)

Principle: Interaction between two proteins of interest *in vivo* can be studied using the bacterial two hybrid system by cloning genes encoding the proteins of interest into BACTH vectors pUT18C, pUT18, pKT25 and pKNT25. pUT18C and pUT18 encode the T18 fragment, and pKT25 and pKNT25 encode the T25 fragment which are catalytic domains

of the *Bordetella pertussis* *cyaA*, an adenylate cyclase effector. When physically separate they are not active, but if the proteins to which they are fused are able to interact, it brings them into close proximity leading to functional complementation. As a result, cAMP is synthesized and interacts with *E. coli* CRP, switching on catabolic genes such as those of *mal* and *lac37* operons (Figure 2.2). This property is used for identification on selective media such as MacConkey media to give maroon colonies in cases where the test proteins interact and white colonies where there is no interaction.

Method: The DNA region for making the test gene-*cyaA* fusions was amplified by PCR as described in section 2.2.10.1 by using appropriate primers and cloned into the BACTH vectors so that it was in the same reading frame as the *cyaA* fragment. For fusions to the N-terminus of T18 and T25, the stop codon was omitted from the test gene. BTH101 competent cells were co-transformed with complementing plasmids (pKNT25 or pKT25 derivative in combination with a pUT18 or pUT18C derivative) as shown in Table 2.9 and spread on MacConkey agar containing ampicillin and kanamycin. The plates were inverted and incubated for 72 hours at 30°C.

Table 2.9: Combinations of plasmids used in a BACTH study of the interaction between two proteins A and B.

	pKT25	pKT25-A	pKT25-B	pKNT25	pKNT25-A	pKNT25-B
pUT18	Control	Control	Control	Control	Control	Control
pUT18-A	Control	Self interaction	A-B interaction	Control	Self interaction	A-B interaction
pUT18-B	Control	A-B interaction	Self interaction	Control	A-B interaction	Self interaction
pUT18C	Control	Control	Control	Control	Control	Control
pUT18C-A	Control	Self interaction	A-B interaction	Control	Self interaction	A-B interaction
pUT18C-B	Control	A-B interaction	Self interaction	Control	A-B interaction	Self interaction

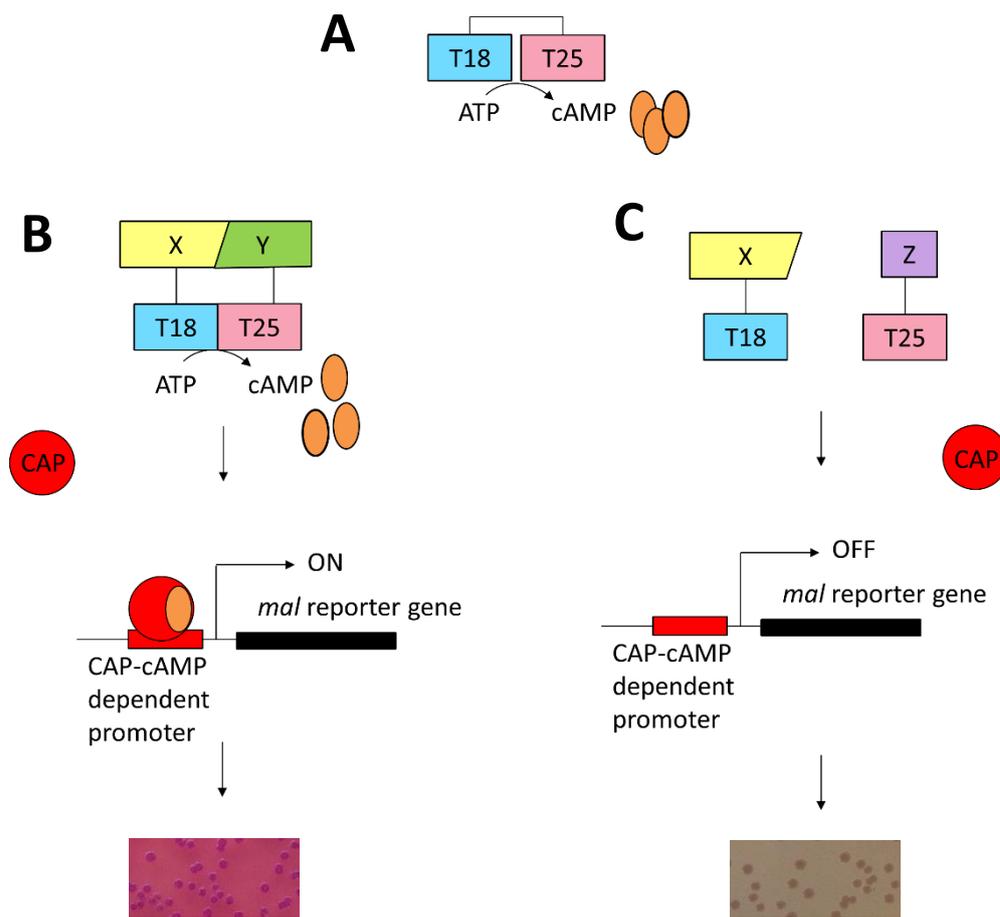


Figure 2.2 Principle of BACTH assay

(A) Functional complementation of the two domains T18 and T25 of adenylate cyclase (CyaA) allows formation of cAMP from ATP. (B) When two interacting proteins X and Y are fused to these two domains, this allows them to be brought in close proximity thus causing formation of cAMP. cAMP can combine with catabolite activator protein (CAP) and switch on the CAP-cAMP-dependent promoter such as the one belonging to the *mal* gene. Thus, cells harbouring these can utilise maltose which gives rise to dark maroon colonies on MacConkeys' agar showing a positive phenotype. (C) In the case where the two fused proteins, X and Z do not interact, T18 and T25 are not brought in close proximity and therefore the cAMP-dependent promoter cannot be switched on as no cAMP is produced. As a result such colonies appear white on MacConkeys' agar representing a negative phenotype.

2.6.2 Chromatin immunoprecipitation and deep sequencing (DNA-protein interaction)

The required bacterial strains were grown overnight at 37°C in LB broth containing appropriate antibiotics. The following day, 50 ml sterile LB was inoculated with the overnight culture, such that the starting OD₆₀₀ was 0.05. When the cultures reached log phase, (OD₆₀₀ = 0.3–0.5), IPTG was added to a final concentration of 1 mM to the cultures to induce expression of the plasmid-encoded σ factor and the cultures were grown at 37°C with shaking for a further 3 hours. Formaldehyde was then added at a final concentration of 1% to allow crosslinking to occur and incubation at 37°C was continued with gentle swirling for 30 minutes. To quench the reaction, glycine was added to a final concentration of 0.36 M and incubated at room temperature for 2 minutes. Cells were then harvested by centrifugation at 4500 rpm for 10 minutes, washed once with 10 ml/g bacterial cells of ice-cold Tris-buffered saline - TBS (10 mM Tris, 150 mM NaCl, pH = 7.4). Cells were washed with 1 ml lysis buffer (50 mM Tris, 150 mM NaCl, 1 mM EDTA, 1% Triton-X100) and lysozyme to a final concentration of 0.2 mg/ml was added and incubated at room temperature for 30 minutes with gentle agitation. Then, phenylmethylsulfonyl fluoride (PMSF) was added at a final concentration of 25 μ g/ml and incubated on ice for 15 minutes. Lysed cells were sonicated using the Jencons Scientific Vibracell sonicator at 10 μ m amplitude at 50% power for 30 seconds, with 30 second intervals on ice until the solution turned clear yellow. This sheared the chromosomal DNA to fragments of 300–800 bps. Insoluble material was removed by centrifugation. The remaining supernatant was applied to 60 μ l anti-FLAG M2 affinity gel (Sigma-Aldrich), which had been washed three times with TBS as described by the manufacturer. The supernatant–anti-FLAG suspension was incubated at 4°C for 2 hours with gentle agitation on a rotating wheel. Resin was collected by centrifugation at 5,000 x *g* for 30 seconds at 4°C, and the supernatant was carefully removed and discarded. The resin was re-suspended in 1000 μ l of ice-cold TBS, and collected by centrifugation at 3,000 x *g* for 30 seconds. The resin was then washed an additional two times with ice-cold TBS. Bound material was eluted from the resin by incubation with 100 μ l TBS containing 150 ng/ μ l FLAG peptide (Sigma-Aldrich) for 30 min at 4°C with gentle

agitation on a rotating wheel. The supernatant was collected without disturbing the resin and transferred to a fresh tube. RNA from the supernatant was removed by treatment with RNase at approximately 4 µg/ml final concentration (ThermoFisher) at 65°C for 6 hours. Protein moieties of the cross-linked complexes were removed by the addition of proteinase K (Sigma-Aldrich) to a final concentration of 200 µg/ml and incubating at 37°C for 2 hours. Released DNA was passed through a clean-up column (section 2.12.9) and eluted in 40 µl of elution buffer provided by the manufacturer. Purified DNA was subjected to library preparation using the Next Ultra II DNA library prep kit (New England Biolabs) as per manufacturers instructions followed by high-throughput sequencing at the 'Microarray and Next Generation Sequencing Core Facility', SiTran, University of Sheffield.

2.7 Phenotypic studies:

2.7.1 CAS assay

One colony of required bacteria was picked using a sterile loop and streaked out on CAS agar. The plates were incubated at 37°C for ~12 hours to allow diffusion of siderophores.

2.7.2 Motility and swarming assays

A single colony was used to inoculate 2 ml of LB and grown overnight aerobically at 37°C. The following day, 5 µl of the overnight culture (OD₆₀₀ adjusted to 1) was spotted on to motility agar or swarming agar plates. The plates were incubated upright at 37°C and the zone of swimming/swarming was measured and photographed the following day. Experiments were carried out in biological, technical and experimental replicates.

2.7.3 Protease assay

The required bacterial strain was aerobically grown overnight in nutrient rich broth (LB/BHI) at 37°C. Then, the OD₆₀₀ was adjusted to 1 and 5 µl of culture was spotted on protease agar. After allowing the droplet to dry the agar plates were incubated overnight at the required temperature and the clear halos surrounding the bacterial

spot were measured the following day. Experiments were carried out in biological and experimental replicates.

2.7.4 Biofilm formation

A single colony was used to inoculate 3 ml of BHI and grown overnight aerobically at 37°C. The following day, 10 µl of the overnight culture ($OD_{600} = 1$) was added to 990 µl BHI in a sterile borosilicate glass tube covered with sterile aluminium foil and placed back in the 37°C shaker incubator. After 24 hours the culture was discarded, the tube washed with ddH₂O twice and 1.3 ml 0.5% crystal violet (CV) was added. This was incubated at room temperature for 15 minutes followed by three ddH₂O washes. The biofilm formed was photographed. As a control, medium without bacterial culture was used. To quantify the formation of biofilm, 1.3 ml of ethanol was added to the tube and incubated for 10 minutes at room temperature. The optical density (OD) of the sample at 570 nm was measured using a spectrophotometer blanked with control sample. Similarly, biofilm formation was photographed and quantified at 48 hours. Each experiment was carried out in biological and experimental replicates.

2.8 Peroxide sensitivity stress tests:

2.8.1 Peroxide sensitivity test in liquid media (survival following 20 minute exposure)

A single colony was used to inoculate 2 ml of LB and grown overnight aerobically at 37°C. The following day, the overnight culture was sub-cultured in the same medium and grown to log phase. 1 ml of the log phase culture that had been adjusted to OD_{600} 0.5 was transferred to a sterile eppendorf tube and H₂O₂ to a final concentration of 10 mM was added. The tube was returned to the 37°C shaker incubator for 20 minutes, and then centrifuged for 3 minutes following which the medium was removed. The bacterial cells were re-suspended in 1 ml fresh LB. The culture was serially diluted and plated on LB agar to measure CFU/ml. Each experiment was carried out in experimental triplicates.

2.8.2 Peroxide sensitivity test in liquid media (monitoring survival time)

A single colony was used to inoculate 2 ml of LB and grown overnight aerobically 37°C. The following day, the overnight culture was sub-cultured in the same medium and grown to log phase. 1 ml of log phase culture that was adjusted to an OD₆₀₀ of 0.5 was removed in a sterile Eppendorf tube and H₂O₂ was added to a final concentration of 7.5 or 10 mM. The tube was returned back to the 37°C shaker incubator and every 5 minutes an aliquot of 10 µl was taken out, serially diluted and 100 µl amounts were plated on LB agar to measure CFU/ml. Each experiment was carried out in experimental triplicates.

2.8.3 Peroxide sensitivity test on solid culture medium (CFU assay)

LB agar was autoclaved and before pouring into Petri dishes, oxidants hydrogen peroxide (1 mM) or sodium hypochlorite (1 mM) or methyl viologen (500 µM) were added. These concentrations were first empirically determined. After the agar solidified and dried, cultures (OD₆₀₀ =1) were serially diluted (10⁻¹ to 10⁻¹²) and spotted on to the agar and allowed to dry. LB agar without oxidants was used as a control. Each experiment was carried out in three biological and experimental replicates.

2.8.4 Zone inhibition assay to test peroxide sensitivity

5 ml of 0.7% LB agar (soft agar) was autoclaved, allowed to cool to approximately 50°C to which 0.1 ml of overnight culture (OD₆₀₀ = 1) was added and overlaid on LB agar. After the soft agar was set and dry, a sterile 6 mm Whatman filter paper disc soaked in 10 µl of 1M hydrogen peroxide or 2M sodium hypochlorite or 250 mM methyl viologen or 100 mM of t-butyl hydroperoxide was placed on top of the soft agar. The concentrations of oxidants used were first empirically determined. As a control, a sterile 6 mm filter disc soaked in 10 µl ddH₂O was used. The plates were incubated at 37°C and the diameter of zone of inhibition was measured the following day. Each experiment was carried out in biological triplicates and over three experimental replicates.

**Chapter 3. Investigation of the role
of the ECF σ factor I35_RS16290
(PrtI)**

3.1 Introduction

I35_RS16290 from *B. cenocepacia* (H111) appears to encode an ECF σ factor. Orthologues of this gene referred to as *prtI* in *Pseudomonas fluorescens* are distributed across the *Burkholderia*, *Bordetella*, *Enterobacter*, *Lysobacter* and the *Pseudomonas* genera in the Proteobacteria phylum with more than 60% similarity; although, some species within this genera such as *P. aeruginosa* lack this σ factor. Figure 3.1 shows the gene organisation of the *prtI*-homologous locus in representative species. *prtI* always appears to be located upstream of a gene that is predicted to encode a trans-membrane anti- σ factor. The *P. fluorescens* orthologue of the anti- σ factor has been referred to as PrtR. Also, the pair of genes predicted to encode cytochromeB561 and catalase appears as a divergently orientated transcriptional unit upstream of *prtI*. These features are conserved and other genes located upstream and downstream of this four-gene cluster seem to vary among different species.

PrtI along with PrtR are reported to have roles in regulating the production of either one or more of the following – protease, cyclic lipopeptide, EPS (and colony morphology factors), lipase and secondary metabolite (GAF) in *pseudomonads*, as described in section 1.6.3.1. (Burger et al., 2000, Liehl et al., 2006, Vodovar et al., 2006, Kimbrel et al., 2010, Okrent et al., 2014, Song et al., 2014, Yang et al., 2014, van den Broek, 2005). Based on the genomic context conservation it is predicted that PrtI has a role in regulation of thiosulphate oxidation (Staron et al., 2009). Current literature suggests that PrtI might be a virulence factor since it appears to regulate genes that form important aspects of pathogenicity. Nevertheless, the environmental signal to which it responds, the target genes, promoter requirements and role of PrtR in the regulation of PrtI remain to be elucidated. The PrtI-regulated proteases present in *Pseudomonas* are not encoded by the *Burkholderia* genome, although these species occupy similar niches. Moreover, there is no current literature investigating the role of PrtI in *Burkholderia*.

To test the hypothesis and to address the aims described in section 1.7, the following experiments were designed:

- (1) To study the role of PrtI by identifying its promoters, two complementary approaches were to be employed: (i) a genome wide scan using chromatin

immunoprecipitation and deep sequencing (ChIP-seq) and (ii) bioinformatics analysis of the non-coding region located upstream of the *prtI* gene.

(2) The promoters identified were to be validated by reporter-fusion analysis.

(3) The DNA sequence requirements for effective promoter utilisation by PrtI were to be studied by carrying out a systematic single base-pair substitution analysis of the identified promoter regions.

(4) A marker-less deletion mutant for the predicted σ factor gene *prtI* and for the anti- σ factor gene *prtR* were to be constructed and their phenotypic effects, particularly in relation to stress response, were to be studied.

(5) The target genes of PrtI were to be studied by mutation studies and/or by analysis of their protein products to further ascertain the role of PrtI.

(6) The interaction between PrtI and PrtR *in vivo* and *in vitro* were to be investigated. The mode of regulation by PrtR was to be studied using a marker-less deletion mutant in *B. cenocepacia* or by expressing the genes in *E. coli*.

(7) The PrtI operon and target operons was to be defined by undertaking a RT-PCR study.

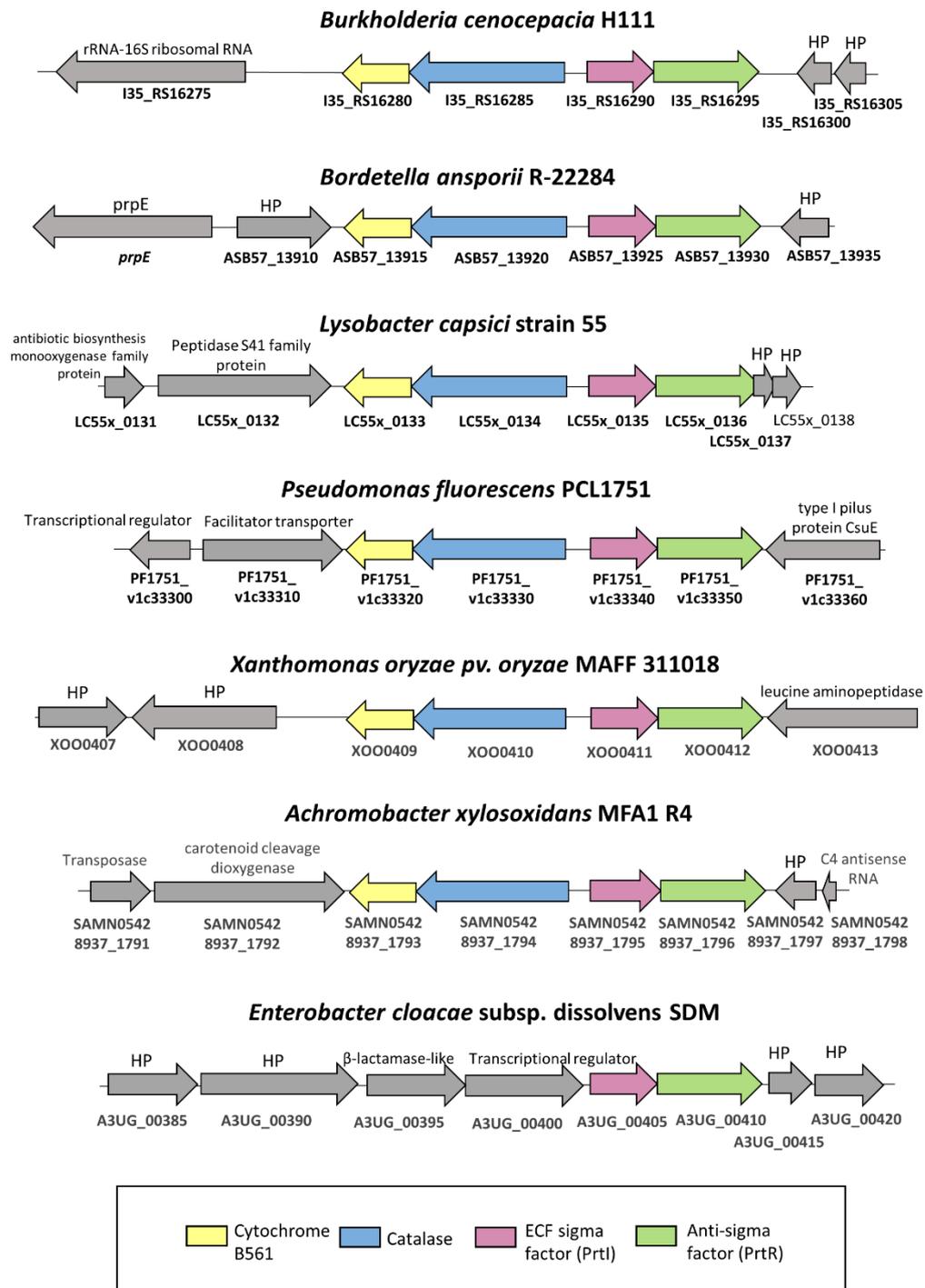


Figure 3.1. Genes present at the *prtI* locus in a variety of Gram-negative bacteria

Gene organisation at the *prtI* locus of *B. cenocepacia* H111 compared to the corresponding locus in other species. Genes that are conserved at the same locus are shown in the same colour represented in the key at the bottom with the predicted proteins encoded by the genes. Genes that are not conserved are shown in grey and their predicted gene products are shown above the corresponding arrow. Each gene locus number is shown below each arrow representing the gene. HP = hypothetical protein.

3.2 Investigation of the role of PrtI by identification of PrtI dependent promoters using ChIP-seq and experimental verification of identified promoters

To study the role of PrtI in stress response of *B. cenocepacia*, the first approach was to identify its target promoter. This would reveal its target genes which in turn would potentially suggest the 'stress' to which PrtI might respond. Hence, a *Burkholderia* genome wide study using chromatin immunoprecipitation and deep sequencing was carried out, where the σ factor PrtI was allowed to interact with its target promoters. Once the promoters were identified, they were to be verified by carrying out reporter-fusion enzyme assays.

3.2.1 Construction of plasmid to express FLAG-tagged PrtI

During normal growth ECF σ factors within cells are generally inactive, as they need an appropriate stress signal for activation. Therefore, it was required to use an inducible broad host-range plasmid such as pSRK-Km that expressed PrtI. Moreover, to facilitate immunoprecipitation of PrtI a FLAG tag was added to the C-terminus. The *prtI* gene (522 bases) ORF was amplified from *B. cenocepacia* strain H111 boiled lysate using KOD polymerase with primers PrtIFLAGFor and PrtIFLAGRev and ligated between *NdeI* and *XbaI* sites of pSRK-Km. The primers PrtIFLAGFor and PrtIFLAGRev were designed such that the C-terminal coding sequence of *prtI* had an additional 24 bases that encoded a FLAG epitope tag (DYKDDDDK) before the stop codon. JM83 cells were transformed with the ligations and using a blue/white selection, white colonies were PCR screened using insert specific primers to identify the 6.322 kbp plasmid (Figure 3.3). The DNA sequence integrity of the clone pSRK-Km-PrtI-FLAG #6 was confirmed by nucleotide sequencing.

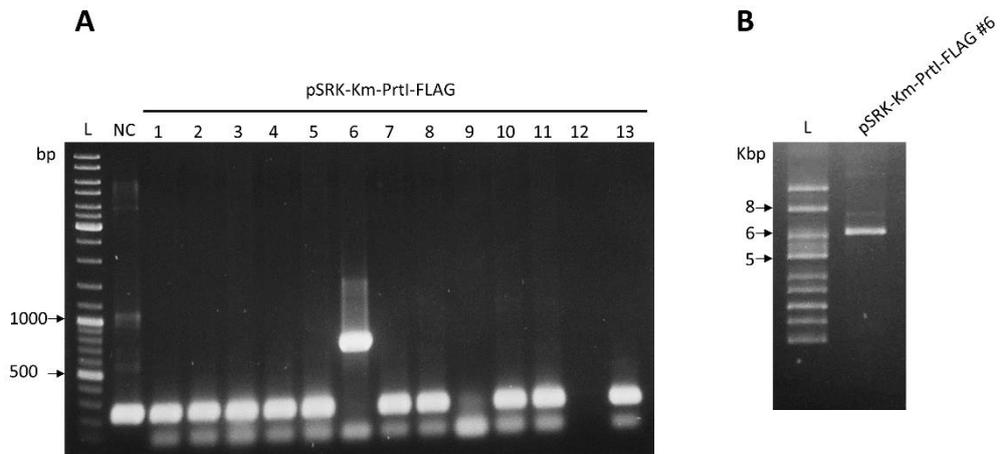


Figure 3.3 Construction of pSRK-Km-PrtI-FLAG

Agarose gel showing (A) PCR screen of clones #1-#13 using vector primers M13for and M13revBACTH to identify pSRK-Km-PrtI-FLAG candidates. Clone 6 gives rise to product of 725 bp consistent with the presence of *prtI*-FLAG; L = GeneRuler DNA mix ladder; NC = pBBR1MCS2. (B) Miniprep of positive clone (6.322 kbps); L = Supercoiled ladder

3.2.2 Validation of antibody specificity

The gene encoding σ factor PrtI was cloned incorporating a FLAG epitope tag at the C-terminal end which would act as bait to immunoprecipitate PrtI in the ChIP-seq experiment. To assess if the protein was being produced, *E. coli* and *B. cenocepacia* cultures containing pSRK-Km-PrtI-FLAG or pSRK-Km were induced and total protein content was visualised on a SDS-PA gel. As shown in Figure 3.4 (A) a protein migrating with an estimated molecular weight of approximately 21 kDa appeared in induced *E. coli* cultures which was the predicted size of PrtI-FLAG. However, in the *B. cenocepacia* lanes, very faint indistinguishable protein bands were observed around 21 kDa. In order to validate that the epitope tag could be recognised when the PrtI protein was expressed, western blotting was carried out (section 2.4.2) using anti-FLAG primary antibody (raised in rabbit) and an anti-rabbit secondary antibody (raised in goat). A protein migrating with a molecular weight consistent with PrtI-FLAG (approximately 21 kDa) was observed for induced cultures of *B. cenocepacia* and *E. coli* while no proteins were observed in control cultures (Figure 3.4B). This confirmed that PrtI-FLAG was expressed from pSRK-Km-PrtI-FLAG under inducing conditions. It was noticeable that the amount of PrtI-FLAG produced in *B. cenocepacia* was substantially less than that produced in *E. coli*.

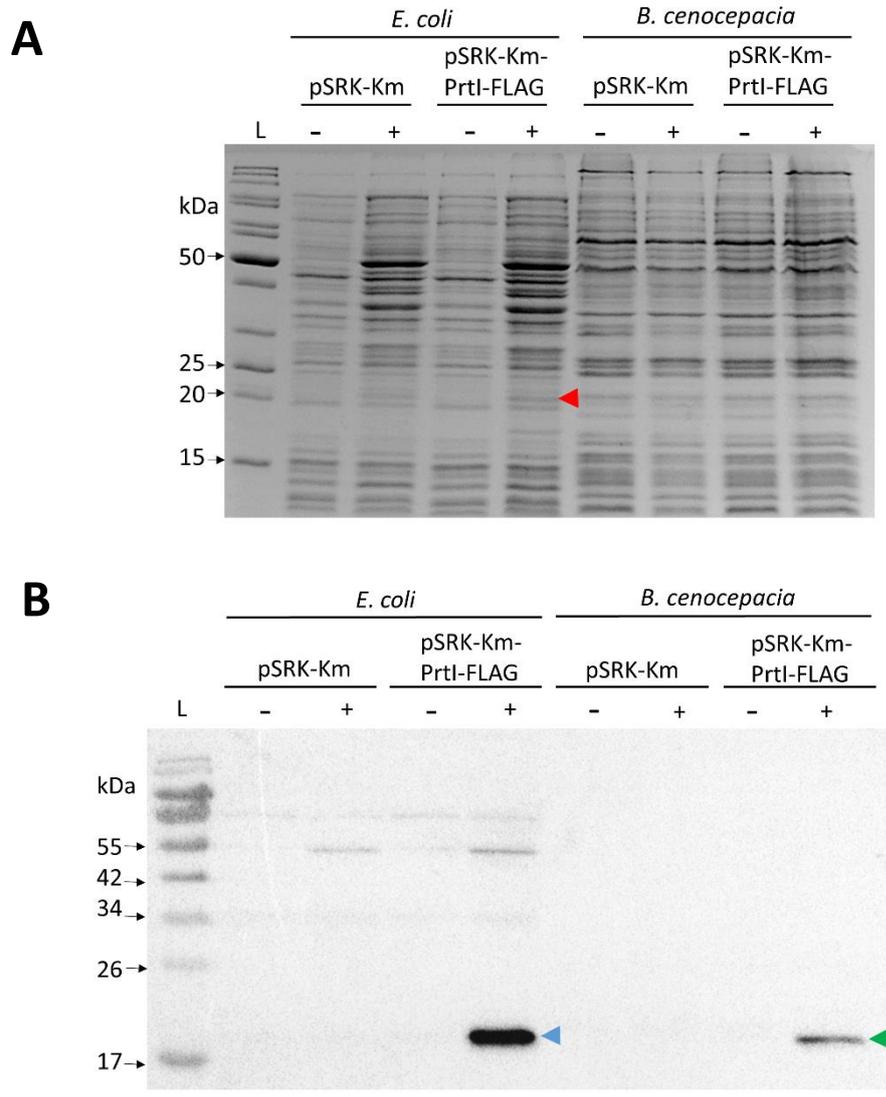


Figure 3.4: Production of Prtl-FLAG in cells containing pSRK-Km-Prtl-FLAG

E. coli MC1061 and *B. cenocepacia* H111 containing pSRK-Km-Prtl-FLAG or pSRK-Km were grown in LB at 37°C and the *lacZ* promoter contained on pSRK-Km was induced by addition of IPTG. (A) Comassie blue stained SDS polyacrylamide gel used to separate total protein from *E. coli* and *B. cenocepacia* containing the indicated plasmids, under non-induced and induced conditions, - = non-induced, + = Induced, L = ladder (EZ-runp). Red arrow corresponds to Prtl-FLAG. (B) Western blot of the same samples from SDS polyacrylamide gel shown in (A) using anti-FLAG antibody, L = ladder (EZ-run pre-stained). Blue and green arrows correspond to Prtl-FLAG in *E. coli* and *B. cenocepacia*, respectively.

3.2.3 Validation of ChIP-seq experimental design by reporter fusion assay: Analysis of putative PrtI regulated promoter activity under induced and non-induced conditions

The ChIP-seq experimental design consisted of cells harbouring the pSRK-Km-PrtI-FLAG plasmid, which when induced by IPTG would express PrtI-FLAG. PrtI-FLAG would be cross-linked with the genomic DNA using formaldehyde, at sites where it binds to its target promoters and would then be immunoprecipitated using an anti-FLAG antibody following purification of genomic DNA. Since incorporation of a FLAG tag might affect the activity of the σ factor, it was necessary to demonstrate that PrtI-FLAG was functional *in vivo*. Often σ factors drive their own promoter and hence it was decided to study the activity of the modified σ factor at the predicted *prtI* promoter using a reporter fusion enzyme assay.

3.2.4 Construction of pKAGd4-P_{prtI}full

The intergenic region between *prtI* and the gene upstream, I35_RS16285, was predicted to contain the *prtI* promoter (P_{prtI}full) and therefore this region of 160 bp was amplified using primers prtIproForPCR and prtIproRevPCR with KOD polymerase and ligated into the pKAGd4 reporter plasmid between the *HindIII* and *BamHI* sites. MC1061 cells were transformed with the products of ligation reactions and resultant transformant colonies were PCR screened using vector specific primers AP10 and AP11 (Figure 3.5). The sequence of positive clones (pKAGd4-P_{prtI}full) was confirmed by nucleotide sequencing.

3.2.5 Analysis of the activity of PrtI-FLAG at the P_{prtI}full promoter

pKAGd4-P_{prtI}full and pSRK-Km-PrtI-FLAG or pSRK-Km were transferred to *E. coli* MC1061 cells and the promoter activity was measured under induced and non-induced conditions using the β -galactosidase assay. As shown in Figure 3.6, the region upstream of *prtI* contained a PrtI-dependent promoter since there was very high β -galactosidase activity in cells containing pSRK-Km-PrtI-FLAG growing under inducing conditions. The activity of this promoter was much lower, although not completely abolished, under

non-inducing conditions in the presence of PrtI-FLAG. Further, no activity was observed in the absence of the σ factor showing that the promoter activity observed was due to the presence of PrtI-FLAG. These results provided validation that PrtI is active under this experimental design and the FLAG epitope tag does not affect its activity. The presence of a promoter in the PrtI dependent I35_RS16285-*prtI* intergenic region also provides a means of validating the ChIP-seq output.

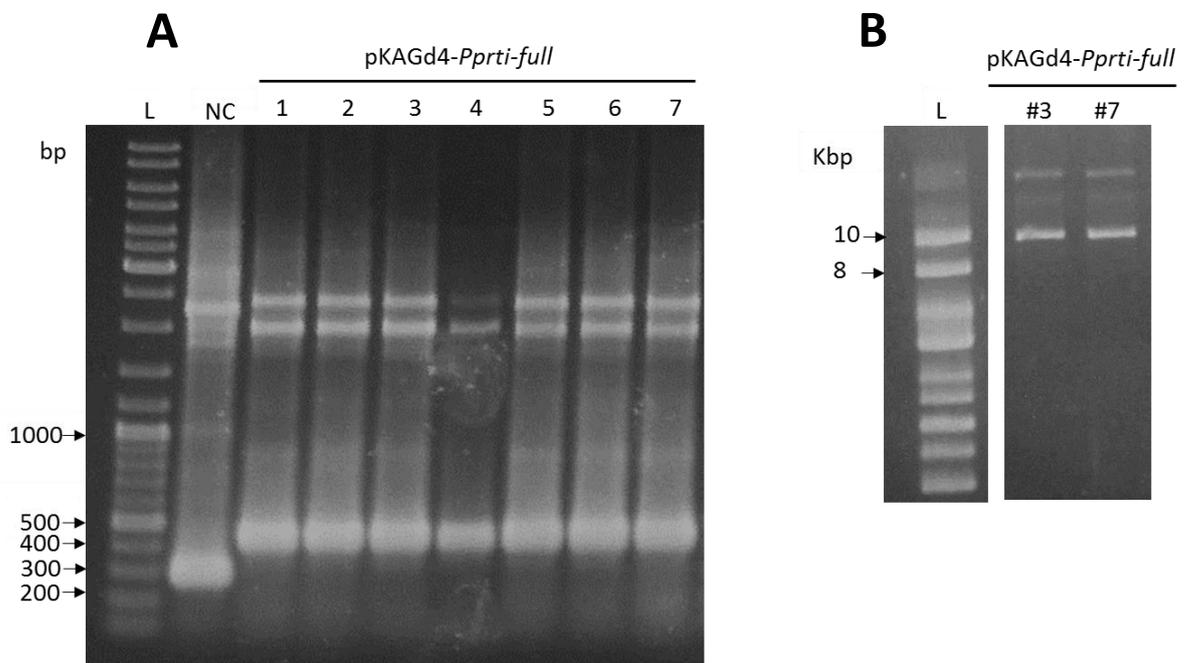


Figure 3.5 Construction of pKAGd4- P_{prtI} -full

Agarose gel of (A) PCR screen of clones #1-#7 using primers AP10 and AP11 to identify pKAGd4- P_{prtI} -full; L = GeneRuler DNA mix ladder; NC = pKAGd4 (282 bp amplicon). All seven candidates gave rise to the expected 442 bp amplicon (B) Miniprep of positive clones #3 and #7 (9.762 kbp); L = Supercoiled ladder

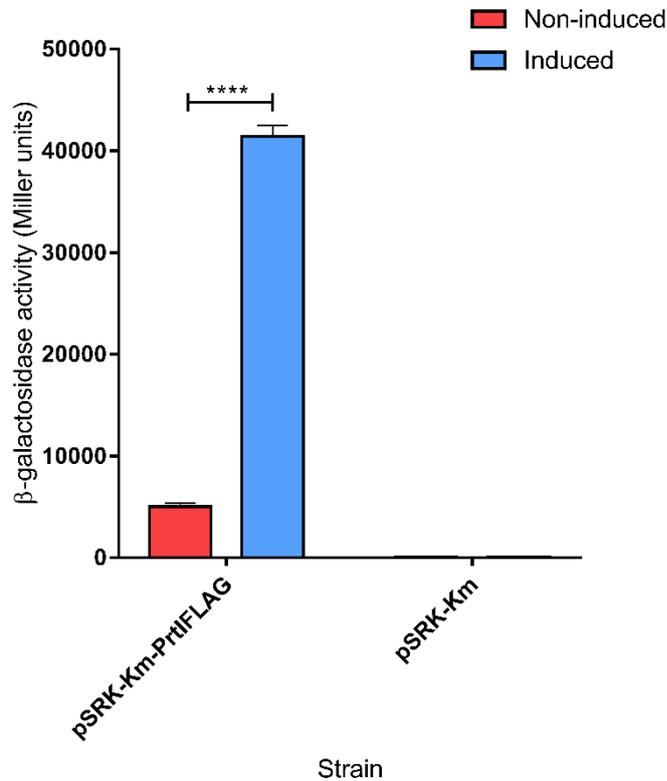


Figure 3.6. Activity of the *prtI* promoter in the presence of Prtl-FLAG

pKAGd4-P_{prtI}full was transferred to *E. coli* MC1061 along with either pSRK-Km or pSRK-Km-Prtl-FLAG as indicated and the promoter activity was measured using the β -galactosidase assay where cultures were grown in LB with kanamycin and chloramphenicol under IPTG induced and non-induced conditions. Activities shown have been obtained by subtracting the activity of same strain with pKAGd4 carried out under same conditions (background activity). Statistical significance was determined by performing a students' t-test, **** = $p < 0.0001$, (n=6).

3.2.6 Optimisation of conditions for chromatin immunoprecipitation

ChIP-seq is a technique used to investigate protein-DNA interactions on a genomic scale, including studying histone modifications and transcription factor targets during different growth conditions. This technique was employed in this study to identify target promoters recognised by σ factor PrtI. Ideally, cells grown under conditions that allow transcription of PrtI should be used to identify its target promoters. However, the exact signal required for PrtI activation is not known as its role remains unknown. Hence, a protocol suitable for identification of the PrtI-dependent promoters was designed by combining methodologies from previous similar studies on ECF σ factors (Markel et al., 2013, Blanka et al., 2014, Giacani et al., 2013).

A ChIP-seq protocol consists of the following main steps as outlined in the schematic representation in Figure 3.7. These are: (1) expressing the protein under suitable conditions such that it can bind to the genome, (2) cross-linking the protein to the genomic DNA, (3) lysing cells and shearing genomic DNA to the appropriate fragment size (approximately 500 bp), (4) immunoprecipitating the protein of interest along with genomic DNA fragments with minimal background noise, (5) isolation of the DNA from the bound protein, and contaminating RNA, (6) library preparation which includes adding adaptors, end correction and DNA amplification, (7) high through-put sequencing using NGS platform, (8) retrieving data, analysing the quality of output, (9) transforming the data output to a suitable format such that it can be used in analysis softwares such as GALAXY and IGV to align sequences back to the *B. cenocepacia* genome and (10) MACS based peak calling and Motif analysis for point source factors.

In addition, it is important to note that since it was decided to use a multicopy plasmid expressing PrtI, the *prtI* ORF will cause a spike in the background noise in the final DNA analysis step and therefore must be ignored while identifying enrichment areas. Along with these main steps to carry out the procedure, ChIP-seq data is acceptable when the antibody is validated and the DNA obtained in step 5 before library preparation is validated for appropriate enrichment as well as purity (Landt et al., 2012). As described in sections 3.2.2 and 3.2.5, the antibody expression was validated by

carrying out a western blotting experiment as well as this experimental design validated by carrying out a reporter fusion enzyme assay.

For the ChIP-seq procedure, genomic DNA shearing by sonication was to be optimised next. To do this two different protocols were tested. Protocol 1 followed the same method as described in Section 2.6.2. In Protocol 2 the lysis buffer was replaced by a lysis-sucrose buffer (10 mM Tris, 50 mM NaCl, 10 mM EDTA, 25 mM sucrose), lysozyme was excluded, TBS buffer was excluded but used IP buffer (50 mM Tris HCl, 150 mM NaCl, 1 mM EDTA, 1 tablet protease inhibitor solution per 50 ml buffer, 0.1% Triton-X-100 solution, 0.1% sodium deoxycholate) where before sonication bacterial cells were resuspended and incubated. In both protocols the growth conditions and crosslinking conditions were the same. As shown in Figure 3.8A, more abundant amounts of DNA recovered after sonication were observed using Protocol 1. Also, interference from RNA was observed using both protocols. Therefore, in the next optimisation step, using protocol 1, after sonication, RNA was digested and DNA-protein un-crosslinking was carried out. For this, 2.5 µl of 4 mg/ml RNase was added per 500 µl sonicated cell lysate and incubated at 65°C for 6 hours. DNA samples were analysed following different periods of total sonication time and it appeared that up to 30 minutes of sonication allowed most amounts of fragmented DNA to be recovered. The frequency of sonication was then increased by sonicating for 30 seconds followed by 30 second intervals on ice instead of one minute intervals, which seemed to give a higher yield of sheared DNA, largely in the 400 - 900 bp size range which lies towards the upper limit of the commonly accepted range for this technique (Figure 3.8C). Also, the effect of different number of wash steps was tested during immunoprecipitation using the commercial anti-FLAG affinity gel and three washes were found to be optimal (data not shown). Thus, the final protocol was standardised, described in detail in section 2.6.2. As a control for this experiment, strain H111 containing empty plasmid pSRK-Km was used, which would act as a background control (termed **CONTROL**). This is also referred to as 'INPUT' in some studies. H111 harbouring pSRK-Km-PrtI-FLAG was the 'test' or commonly called **IP** sample for this experiment. The ChIP-seq experiment was carried out over three experimental replicates which also served as biological replicates thus

giving a total of 6 samples. It is usually acceptable to have at least two experimental replicates.

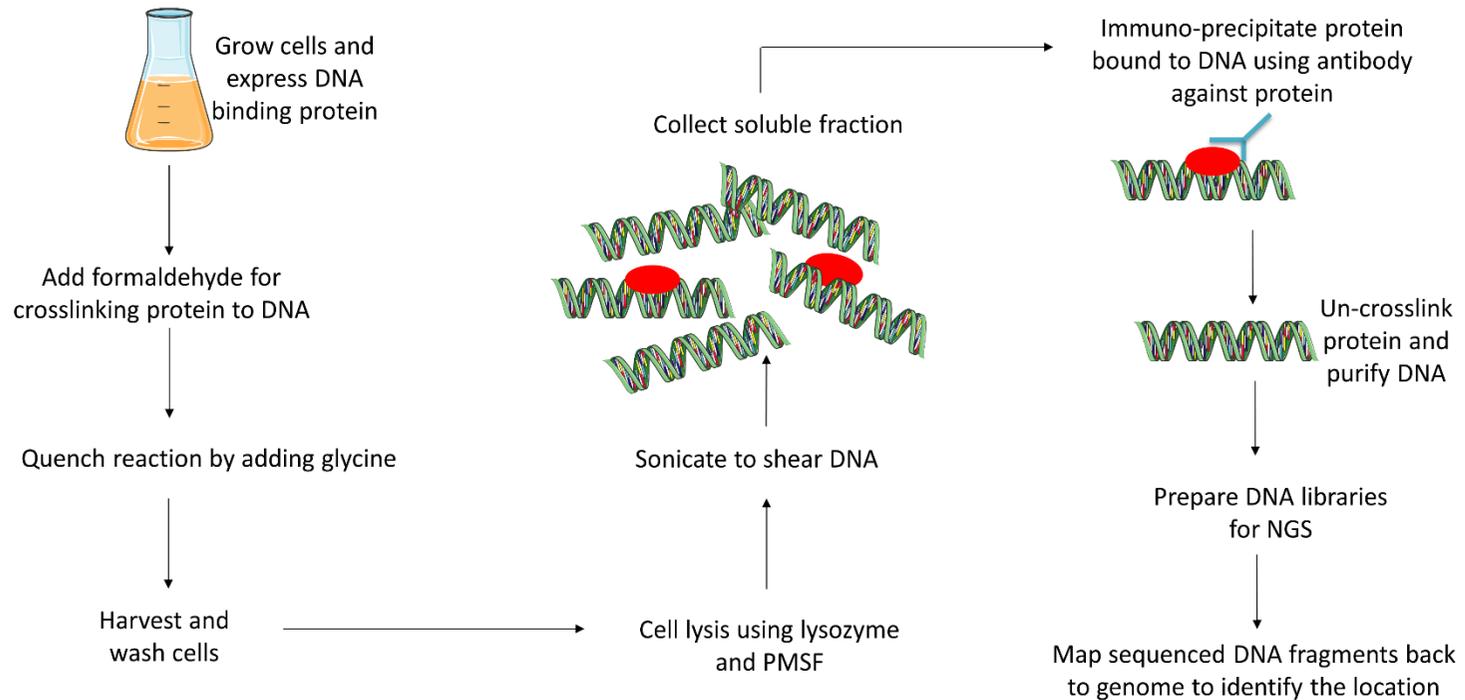


Figure 3.7. ChIP-seq work flow for identification of PrtI binding sites

ChIP-seq is carried out generally by following the steps – (1) Cells are grown to log phase and the protein is expressed, (2) the protein of interest is crosslinked to genomic DNA by adding formaldehyde. The reaction is quenched using glycine, (3) cells are harvested by centrifugation, washed, lysed and DNA is sheared by sonication. (4) Immunoprecipitation is carried out using antibody against the DNA binding protein. (5) The co-precipitated DNA was purified and (6) libraries are prepared, (7) sequenced, (8) checked for their quality. (9) The sequences are finally matched back to the genome to reveal where the protein had bound on the genome.

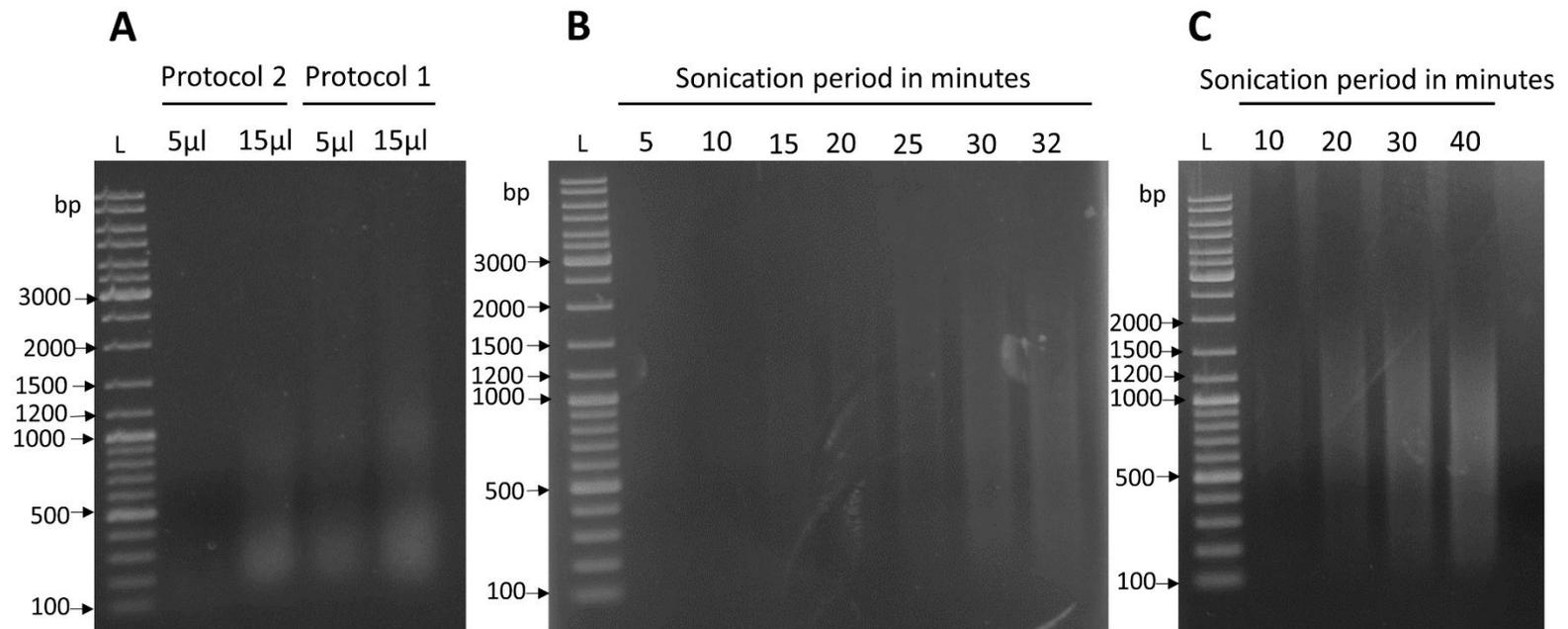


Figure 3.8. Optimisation of cell lysis and shearing genomic DNA by sonication

(A) Testing two different cell lysis protocols: 5 or 15 µl lysed cell suspension obtained using protocol 1 or protocol 2 were electrophoresed in a 0.8% agarose gel (B) Optimisation of sonication period in protocol 1: 20 µl lysed cell suspension obtained by multiple 30 second bursts of sonication with one minute intervals on ice for the indicated total time duration were electrophoresed in a 0.8% agarose gel. (C) Optimisation of sonication period in protocol 1 with decreased intervals: 20 µl lysed cell suspension obtained by multiple 30 second sonication bursts with 30 second intervals on ice for the indicated total time duration were electrophoresed in a 0.8% agarose gel.

3.2.7 PCR validation of ChIP efficiency

After immunoprecipitating FLAG-tagged protein bound to DNA, the DNA-protein complex was un-crosslinked. The protein was digested using proteinase K and the contaminating RNA was also removed from the samples by RNase treatment. The DNA was further purified using a PCR purification kit (Thermo) and the purity of the DNA sample was checked by measuring the absorbance using a nanodrop ($260/280 = 1.8-2.2$). The concentration was found to be between 6-8 ng/ μ l.

To validate for appropriate DNA enrichment in the IP sample as compared to the CONTROL, a standard PCR was carried out using primers prtIproForPCR and prtIproRevPCR (These are positive control primers. The term control should not be confused with CONTROL sample). They are used for amplifying P_{prtI} as explained in section 3.2.4 where DNA obtained from the IP sample or the control sample was used as the template. The rationale for this experimental design was that if the DNA enrichment took place in this region, one can assume that the ChIP-seq experiment has successfully worked since PrtI has recognised its own promoter, as demonstrated earlier in section 3.2.5. One can also assume that it might have enriched other unknown promoters, which was the aim of undertaking a genome wide analysis.

As negative controls for this experiment two pairs of primers were employed: pFlrSfor and pFlrSrev amplified a product of 193 bp, the promoter region of a putative iron starvation ECF σ factor FlrS (J. Mohanlal and M. Thomas, unpublished results). pIOTVfor and pIOTVrev amplified a product of 388 bp, the experimentally determined divergent highly active σ^{70} -dependent *tssJ* and *tagM* promoters, involved in expression of the type VI secretion system (S. Shastri, 2011). Since PrtI should not bind to regions amplified by these two pairs of primers, no or little amplification should be obtained when the IP sample is used as template; these negative control primers should not be confused with the CONTROL sample. Little or no amplification was expected to be obtained when using CONTROL sample as a template with any of the above pairs of primers.

Approximately 8 μ l was removed from the PCR after every 9 cycles to observe rate of PCR amplification. As observed in Figure 3.9A, with primer pair prtIproForPCR

and prtIproRevPCR, the PCR product from the IP template was more abundant after 18 cycles as compared to that from the CONTROL template. However, using pFlrSfor and pFlrSrev and pIOTVfor and pIOTVrev negative control primer pairs, even though there was amplification, the product yield obtained using IP or CONTROL as template was similar. This suggested the amplification occurred from low level contaminating DNA that was non-specifically removed in the IP step. It is important to also compare the yield of PCR products produced by all three primer pairs at cycle no. 18 using IP DNA template. There is a greater amount of product generated by the positive control primers as compared to the negative control primers. However, comparing PCR products among all three primer pairs for cycle no. 18 using control DNA template, it is obvious that as the yields appear to be similar, it must be due to the presence of DNA that non-specifically bound to the anti-FLAG resin. This background is accounted for in the future data analysis steps.

To test if the primer pairs had similar amplification efficiency, making it valid to compare products between primer pairs, a PCR was carried out as described above including an additional positive template control of boiled H111 lysate. As shown in Figure 3.10, primer pairs prtIproForPCR and prtIproRevPCR, and pIOTVfor and pIOTVrev generate products with a similar efficiency when using H111 DNA as template. However, DNA from the IP sample using negative control primers pIOTVfor and pIOTVrev gave a poor yield of product, like the CONTROL template. However, positive control primers prtIproForPCR and prtIproRevPCR used with IP DNA as template gave a more abundant product like the H111 template, suggesting an enrichment of the expected DNA region as well as providing evidence that amplification by the negative control primer pairs must be due to contaminating genomic DNA arising from non-specific binding to the anti-FLAG resin.

The enrichment of PrtI-specific DNA could be verified by carrying out qPCR (quantitative PCR) using a housekeeping gene as a control. However, as the results obtained using the standard PCR were very convincing it was decided to go ahead with the next step of library preparation.

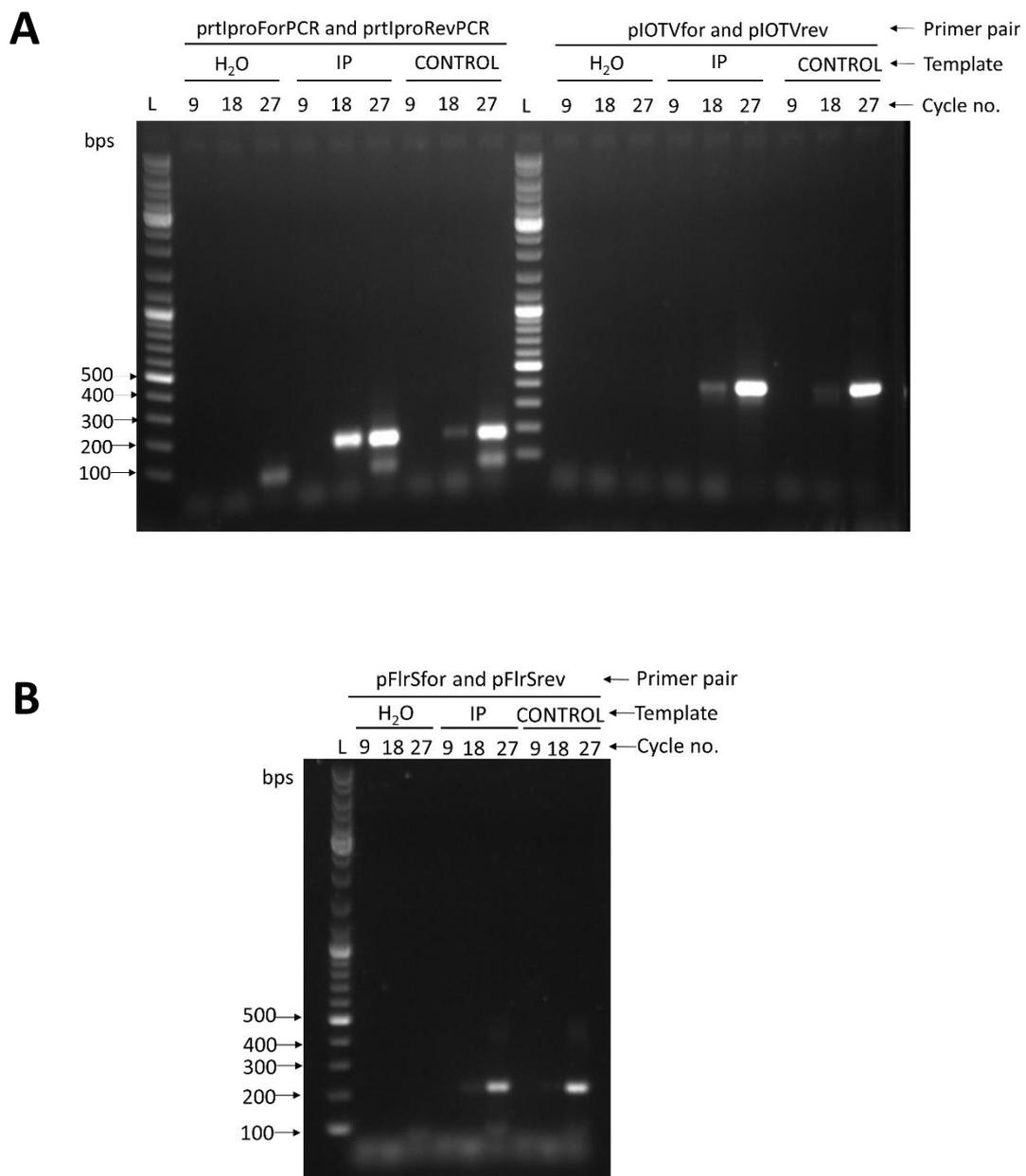


Figure 3.9. Validation of IP and CONTROL samples using PCR

(A) Agarose gel containing samples taken after 9, 18 or 27 cycles of PCR using either no template, IP template or CONTROL template with either positive control primers prtIproForPCR and prtIproRevPCR (160 bp product) or negative control primers pIOTfor and pIOTrev (388 bp product) as indicated. (B) Agarose gel containing samples taken after 9, 18 or 27 cycles of PCR using either no template, IP template or CONTROL template with negative control primers pFlrSfor and pFlrSrev (193 bp product).

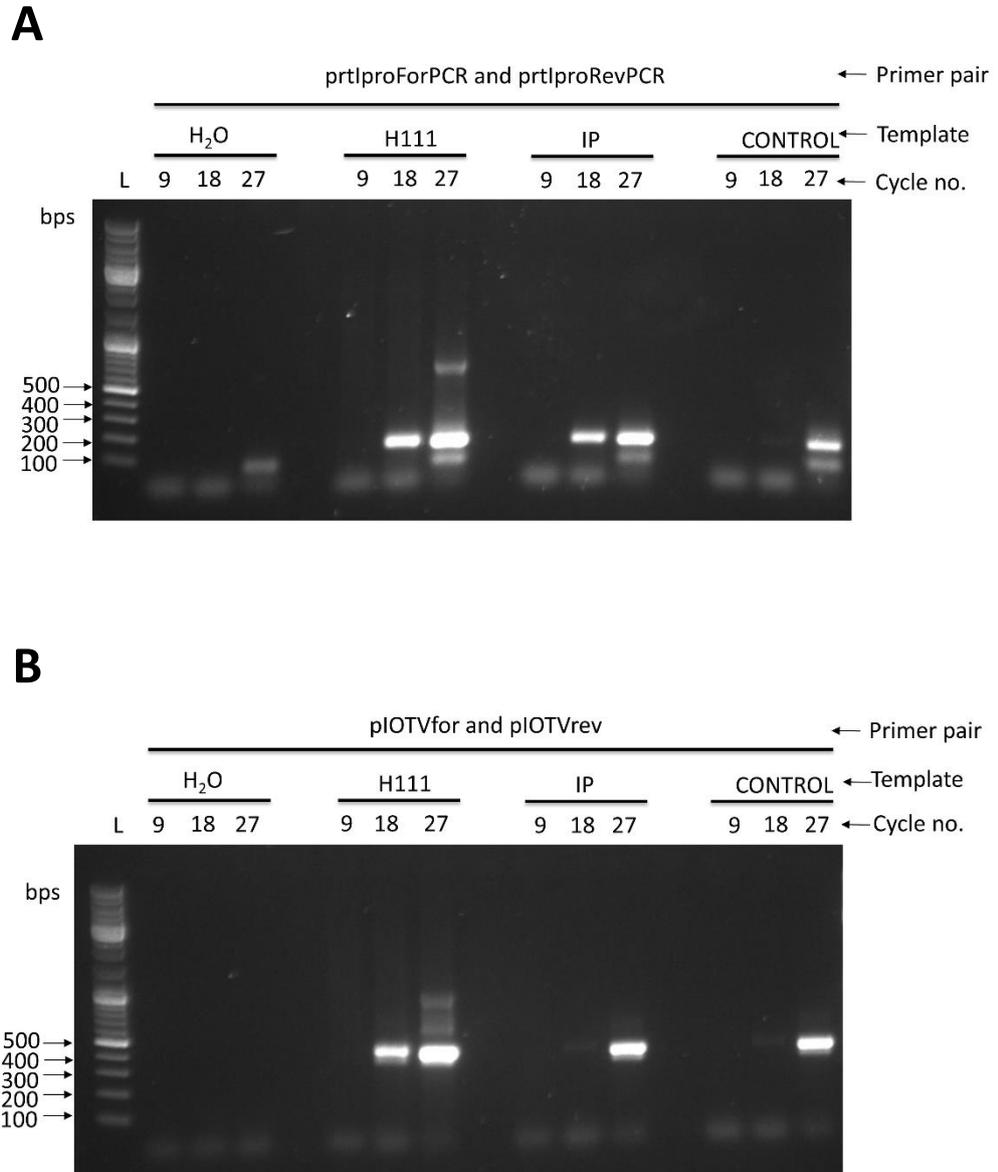


Figure 3.10. Validation of IP and CONTROL samples using PCR with additional positive control
 (A) Agarose gel containing samples taken after 9, 18 or 27 cycles of PCR using either no template, H111 boiled lysate, IP template or CONTROL template with positive control primers prtIproForPCR and prtIproRevPCR (160 bp product). (B) Agarose gel containing samples taken after 9, 18 or 27 cycles of PCR using either no template, H111 boiled lysate, IP template or CONTROL template with negative control primers pIOTfor and pIOTrev (388 bp product).

3.2.8 Quality control of DNA samples library preparation

The DNA fragment size distribution and quality following sonication and ChIP was assessed using the Agilent high sensitivity bioanalyser with the 2100 Expert software, version 1.03. As shown in Figure 3.11A most DNA fragments are concentrated between 600-1000 bp. All pairs of IP and CONTROL samples have a similar distribution, suggesting that the experiment was carried out with in a consistent manner. Also, the number of washes was increased in the second and third pair of experimental replicates, so the samples are less concentrated. A representative image of a detailed sample assessment has been shown for one of the samples (Figure 3.11B). Range of DNA fragments size is 200-1000 bp with average size being 559 bp. The total DNA concentration was 4252.99 $\mu\text{g}/\mu\text{l}$.

For sequencing the samples, they needed to be modified to enable them to be read by the sequencing platform. This was carried out in the DNA library preparation procedure and employed the NEB UltraII DNA library kit for the Illumina Solexa platform. The steps involved in this protocol included end repair by phosphorylation of 5' ends and dA tailing. After this, adaptor ligation and 'U excision' was carried out. Each sample had a unique index number which was used to recognise the readings from the sequencing reaction. A size selection step then followed which selected DNA fragments of appropriate size (average 500 bp) to be amplified by PCR to concentrate the product. A final purification step was carried out after which samples were again checked for their concentration and size using the Agilent high sensitivity DNA bioanalyser with the 2100 Expert software, version 1.03. After DNA library preparation, all samples were found to have a normal distribution of DNA fragments with an average size of 500 bps (Figure 3.12).

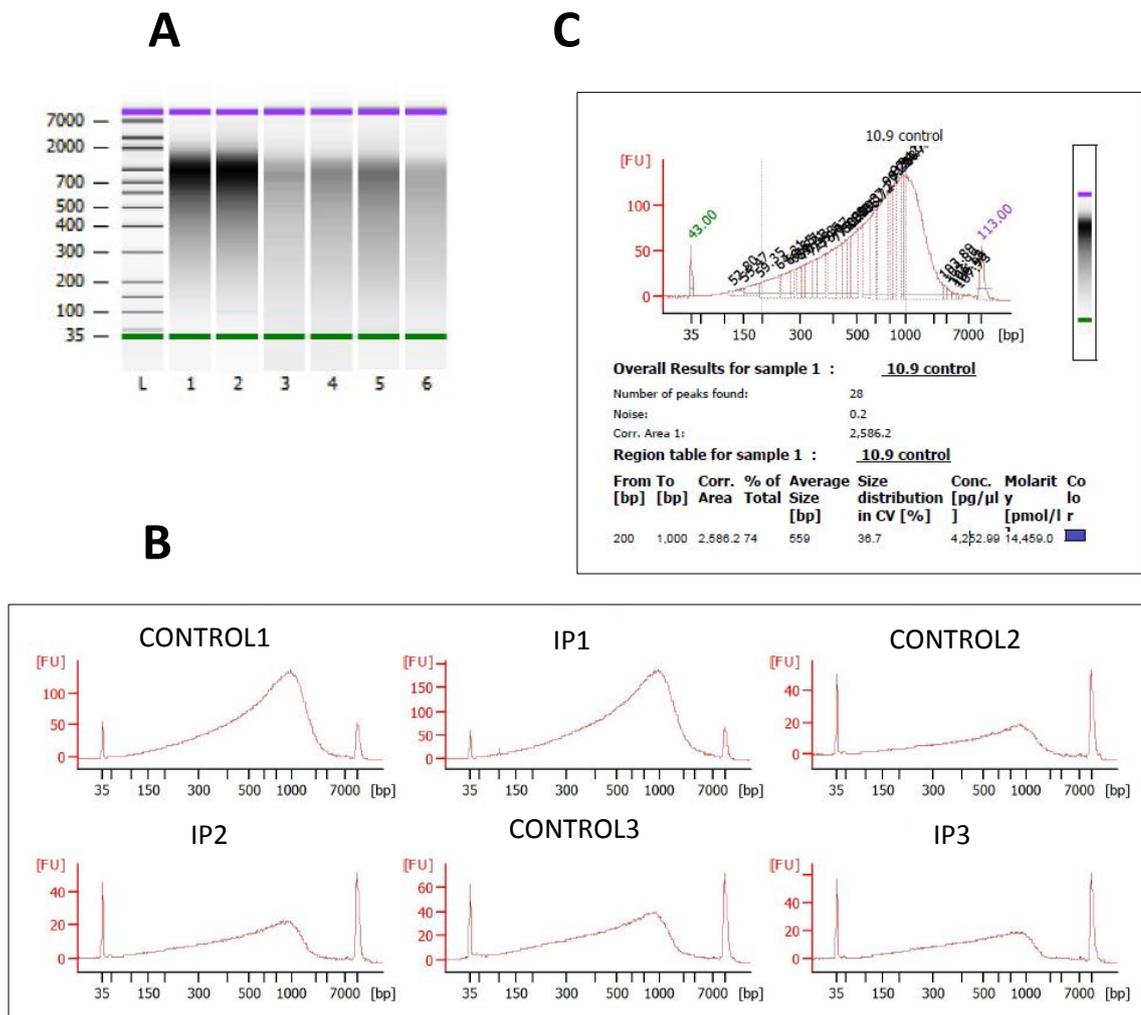


Figure 3.11. Quality assessment of samples before library preparation

(A) Agarose gel electrophoresis run summary showing distribution of DNA fragments: L, ladder (sizes shown in bp); lanes 1 and 2, CONTROL and IP samples from experiment 1; 3 and 4, CONTROL and IP samples from experiment 2; 5 and 6, CONTROL and IP samples from experiment 3. (B) Size distribution of DNA fragments: CONTROL1 and IP1 = CONTROL and IP samples from experiment 1, CONTROL2 and IP2 = CONTROL and IP samples from experiment 2, CONTROL3 and IP3 = CONTROL and IP samples from experiment 3. (C) Electropherogram showing detailed sample analysis summary of CONTROL sample from experiment 1, main parameters being concentration and average size of DNA fragments. X-axis shows DNA size in bp, Y-axis shows signal intensity.

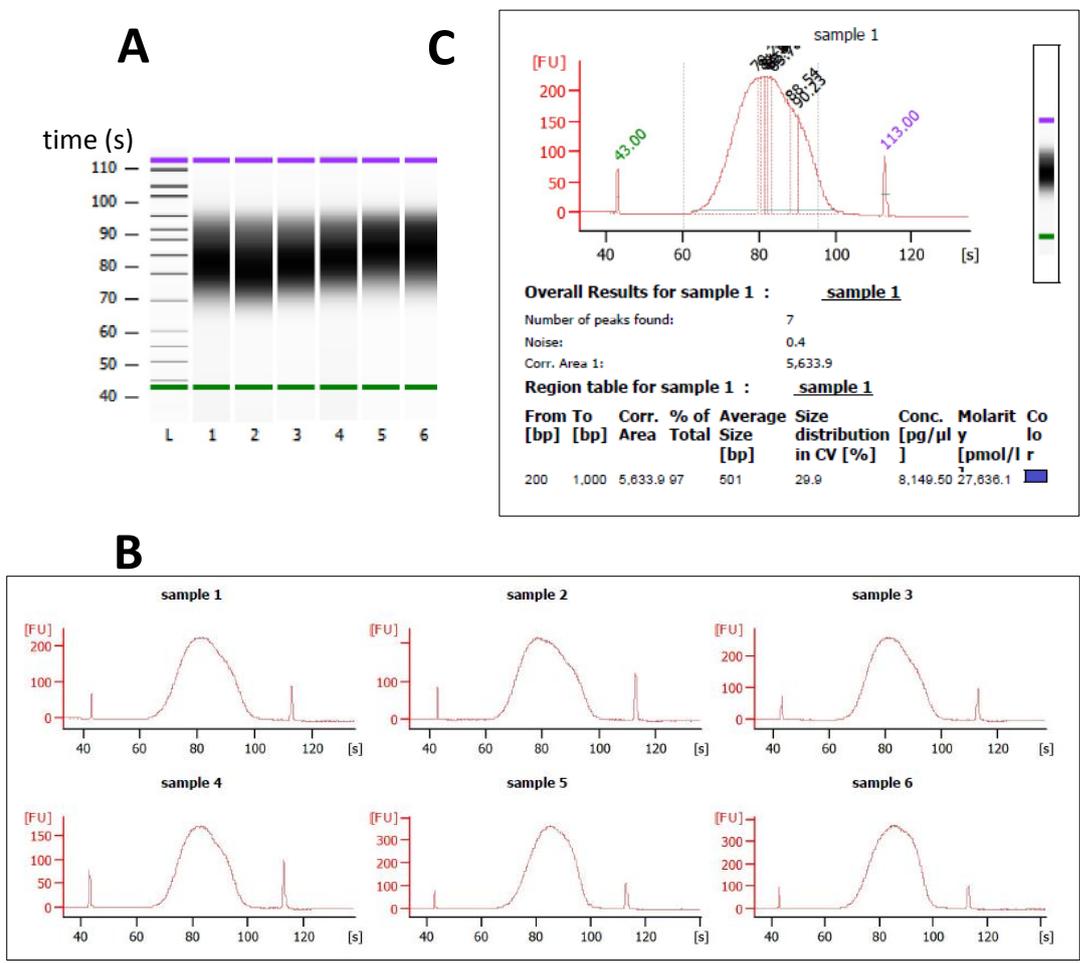


Figure 3.12. Quality assessment of samples after library preparation

(A) Agarose gel electrophoresis run summary showing distribution of DNA fragments: L, ladder; lanes 1 and 2, CONTROL and IP samples from experiment 1; 3 and 4, CONTROL and IP samples from experiment 2; 5 and 6, CONTROL and IP samples from experiment 3. (B) Size distribution of DNA fragments: sample 1 and sample 2 = CONTROL and IP samples from experiment 1, sample 3 and sample 4 = CONTROL and IP samples from experiment 2, sample 5 and sample 6 = CONTROL and IP samples from experiment 3. (C) Electropherogram showing detailed sample analysis summary of CONTROL sample from experiment 1, main parameters being concentration and average size of DNA fragments. X-axis shows time in seconds, Y-axis shows signal intensity.

3.2.9 Data analysis of ChIP-seq output: Quality assessment of sequencing output

Sequencing of the DNA library was carried out using an Illumina HiScan SQ at the Microarray and Next Generation Sequencing Core Facility in SITraN, University of Sheffield. The technician retrieved the data and converted it to a format which could be used for downstream processes. Using the software Sequence Analysis Viewer (SAV), the data files were analysed to check their quality. In Figure 3.13, the left hand panel showing vertical coloured bars that represent the 'flow cell chart' with colour coded quality metrics of the samples that were loaded on to the sequencer. The bar with numbers to its right represents the indicated colours in these tiles.

'Data by cycle' shows the quality of metrics of 4 (A, T, G, C) nucleotides in that sample during the run, i.e. in each 'cycle' of a sequencing reaction. 'Data by lane' shows quality metrics per lane where one lane contained 6 samples. The Q-score distribution plots help in identifying the number of reads based on a pre-determined quality score. Only reads up to a certain quality score threshold (30) are included in the final analysis (green bars). A Q-score heat map shows the Q-score distribution based on the cycles. 'Reads mapped to index ID' allows one to check how much quantity of each sample has been read. This is carried out by the sequencer – therefore, some samples get more reads than the other, although this is an unbiased random experimental occurrence. Here 6 indexes correspond to 6 samples, for example 'ATCACG' is the unique adaptor for sample no 1 as shown in Figure 3.14. The dots on the plot indicate that each sample has been read. Various features on SAV allow the investigator to study quality metrics like signal to noise ratio, error rate, % phasing, outliers, etc. based on tiles, clusters and lanes. However, detailed explanations on these will not be given here since it is beyond the scope of this project. Therefore, only relevant plots have been shown in this thesis. Overall, with the assistance of the trained technician at the Microarray and Next Generation Sequencing Core Facility I was able to confirm that all sequences were read appropriately and were found to be of acceptable quality for data analysis.

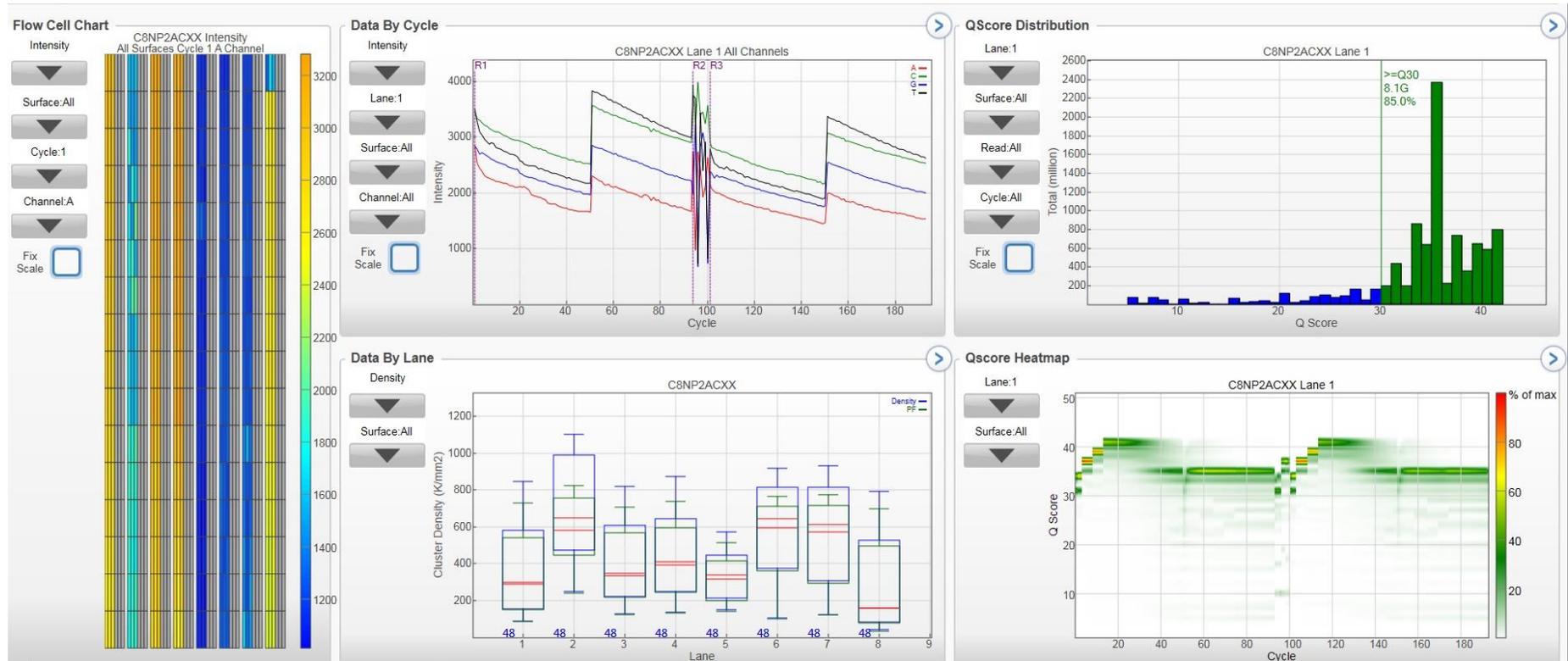


Figure 3.13. Quality metrics of reads from sequencer using SAV

The panel on the left represents the flow cell chart showing colour coded intensity (concentration) of all samples. Top left graph shows sequencing data by cycle plot. Bottom left plot shows sequencing data by lane. Top right plot show Q-score distribution (quality threshold) of samples from lane 1 (6 samples), with the blue bars representing the outliers which were excluded in the final data analysis. Green bars represent data up to the quality threshold, which were accepted in the final data analysis.

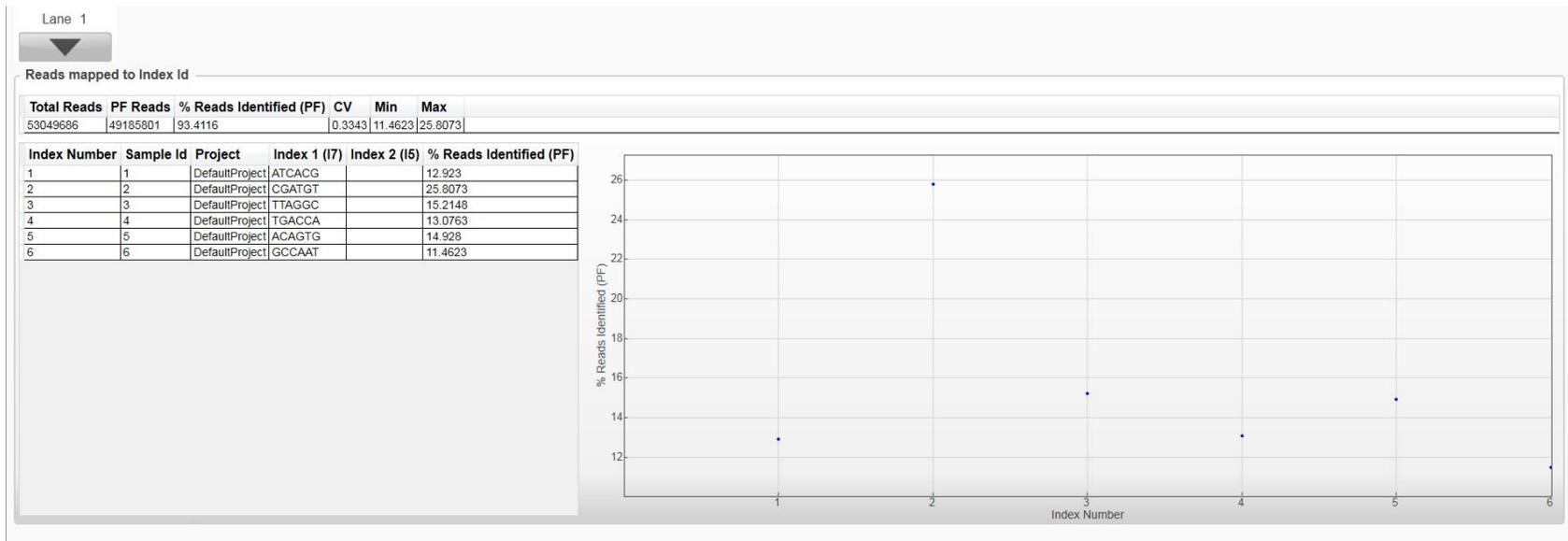


Figure 3.14. Quality metrics of reads mapped to index using SAV

Plot of number of reads identified per index where each index is associated with one sample.

As there were six samples (IP and CONTROL, 3 experimental repeats), which were read in forward and reverse orientation by the sequencer, there were a total of 12 data files. The quality analysis for each of these files was found to be acceptable based on the SAV analysis. Next these files (FASTQ format) were loaded on to the open source data analysis software GALAXY. The function 'fastQC' on GALAXY allows one to test the quality of their files based on some pre-determined parameters. Although SAV was used, the quality control parameters were also checked on GALAXY. Though all samples passed most of the quality control checkpoints on GALAXY, some parameters were not up to the mark. However, it must be noted that these are pre-determined by the programme and they may not necessarily match the requirements of all experimental designs. To choose only best quality data for further sequence analysis, the function called 'Fastgroomer' on GALAXY can be used. Data was corrected by 'groomer' by filtering out those sequences that did not meet the expected quality metrics. Figures 3.15 to 3.20 shows representative images of only one data file – forward run of CONTROL sample from experiment no. 1. The quality control data for other samples are shown in Appendix II. A green tick at the title of the parameter suggests it meets the expected quality, a yellow exclamation mark suggests that it may not fully meet the quality requirements and a red cross represents an error.

Firstly, **basic statistics** gives an overview of the file including information such as GC content, length, file name and type and total sequences. **Per base sequence quality** gives a box and whisker plot showing the quality of sequences for each base. Like the Q-score distribution using SAV, this has to meet a minimum threshold, where it was 20 for GALAXY and 30 for SAV. Next one can also look at **Per sequence quality scores**, where the distribution of means of quality scores of all sequences based on means of each base quality is shown. As seen in Figure 3.15, most of this data was universally of high quality. **Per base sequence content** shows distribution of the 4 bases – A, T, G, and C. In Figure 3.16 it can be seen that for this parameter a yellow exclamation mark was received. This is due to the *Burkholderia cenocepacia* genome being very GC rich (approximately 67% GC content), the sequences tend to have a higher GC content compared to sequences from many other genomes. However, it is important that these lines are mostly parallel

to the X-axis since this suggests it is an unbiased library sample. **Per sequence GC content** across all sequences should ideally be a normal distribution as represented by the blue line. The red line shows the actual distribution for the sample. **Per base N content** shows if there are any uncalled bases in the library. **Sequence length distribution** is a plot of the length of all the sequences. Figure 3.18 shows all sequences in this library are 93 bases long. **Sequence Duplication Levels** shows how unique each of the sequence is in the complete DNA sequence library. Most of the sequences have a sequence duplication level of '1' meaning they appear only once. **Over represented sequences** are also checked for by inspecting if any of the sequences make up more than 0.1% of the library. Figure 3.19 shows there were no over represented sequences in this library. **Adaptor content** shows if there is excessive contamination from adaptor sequences. **Per tile sequence quality** represents the quality based on each tile of the flow cells. **Kmer content** also assesses uniqueness of sequences, but of shorter-length (pentanucleotide) sequences. This did not seem to meet the expected standard, as seen in Figure 3.20. However, this should not affect the data analysis since there were no over represented sequences nor uncalled bases as well as an absence of contamination by adaptors. Also, using the 'Fastgroomer' function on GALAXY trimmed the over-represented Kmer content.

✓ Basic Statistics

Measure	Value
Filename	1_ATCACG_L001_R1_001.fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	6356264
Sequences flagged as poor quality	0
Sequence length	93
%GC	65

✓ Per base sequence quality

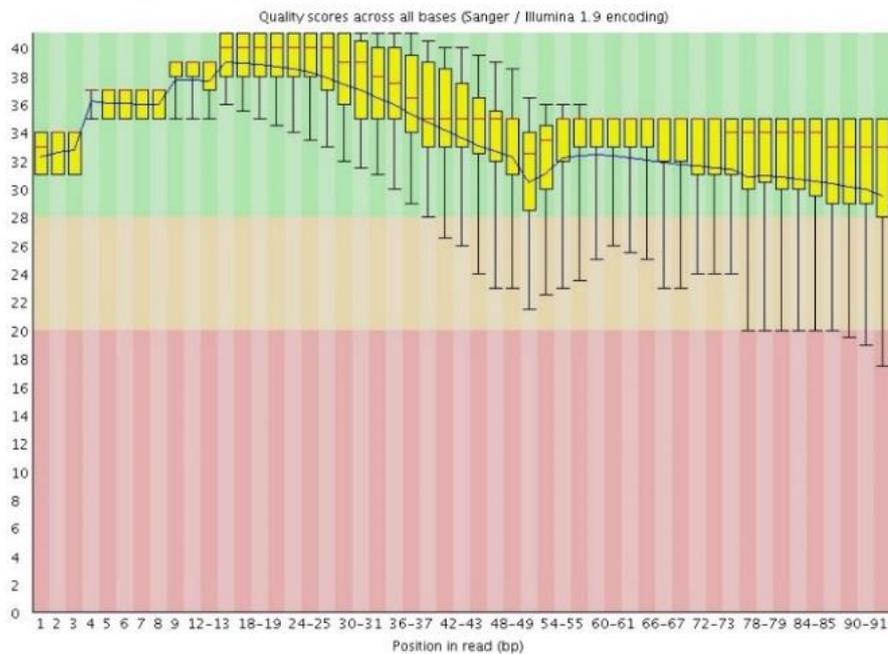
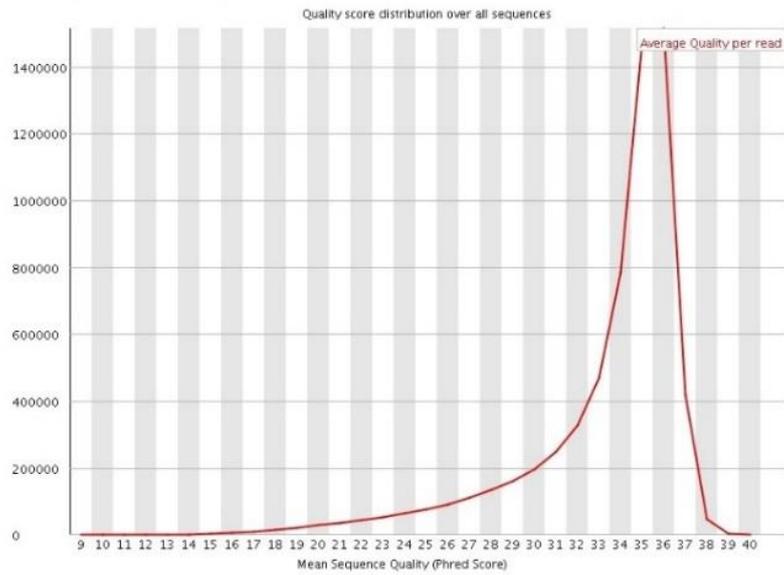


Figure 3.15. Quality metrics of sequenced data using GALAXY (1)

Basic statistics and per base sequence quality of sample no. 1 (CONTROL), forward run

✔ Per sequence quality scores



⚠ Per base sequence content

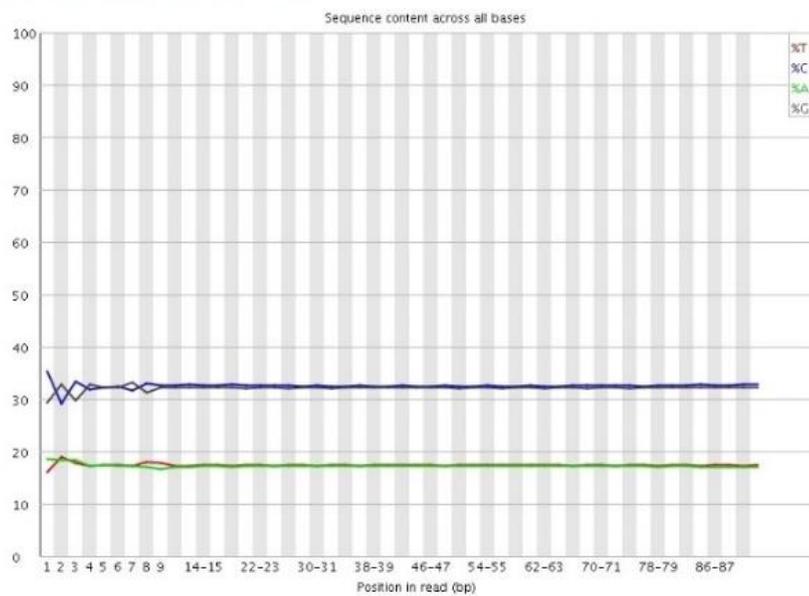
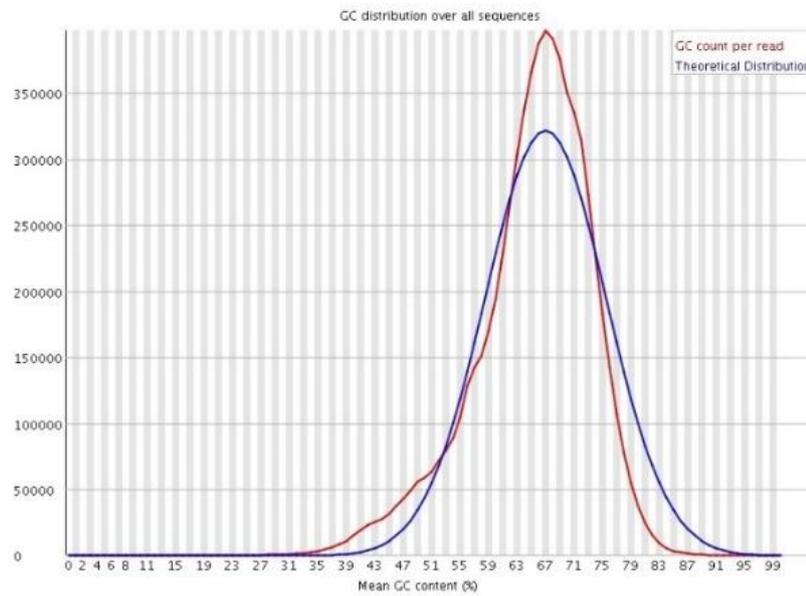


Figure 3.16. Quality metrics of sequenced data using GALAXY (2)

Per sequence quality scores and per base sequence content of sample no. 1 (CONTROL), forward run

🚩 Per sequence GC content



✅ Per base N content

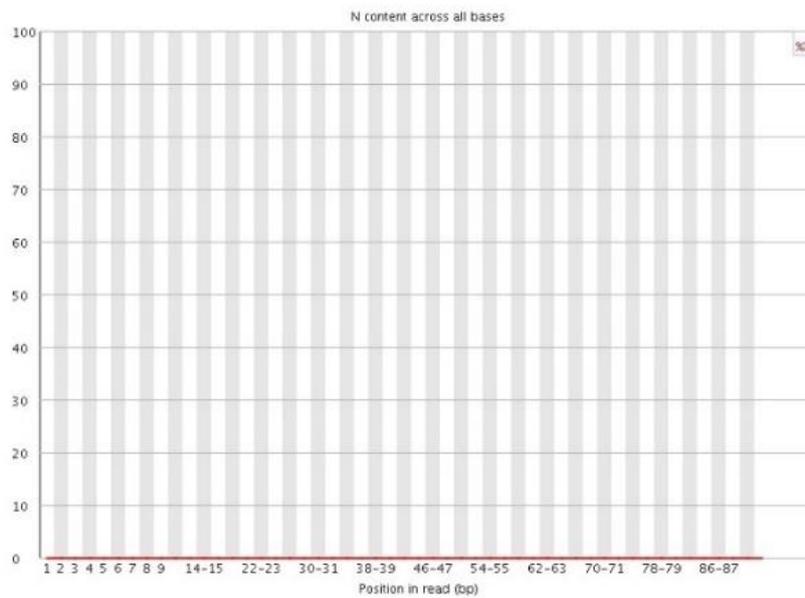
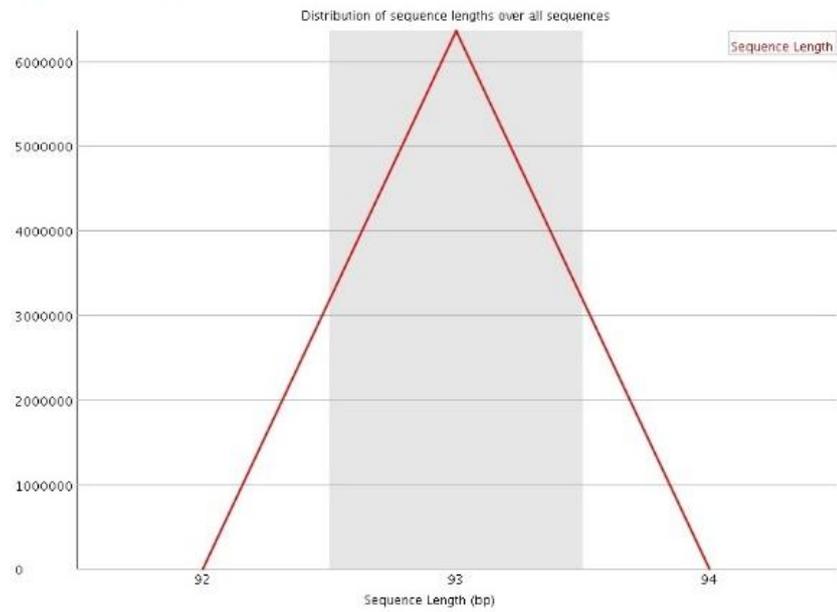


Figure 3.17. Quality metrics of sequenced data using GALAXY (3)

Per sequence GC content and per base N content of sample no. 1 (CONTROL), forward run

Sequence Length Distribution



Sequence Duplication Levels

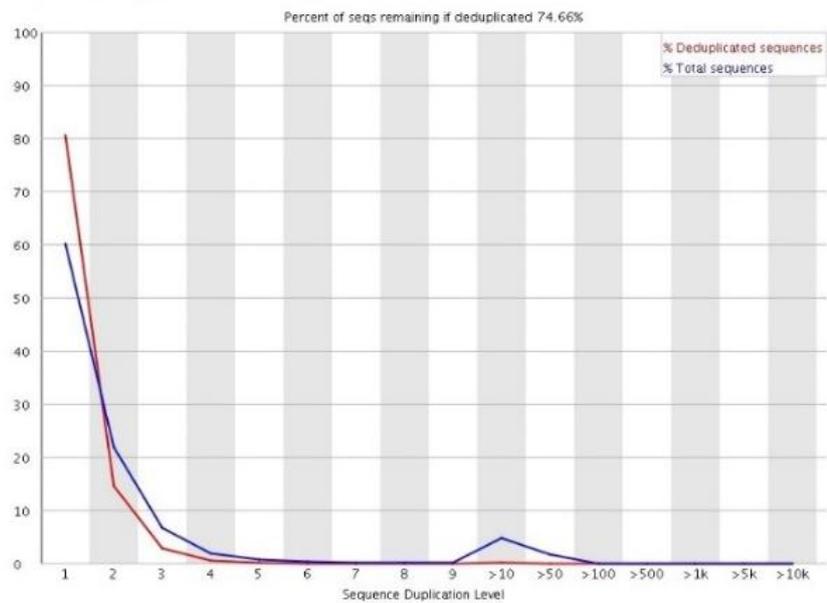


Figure 3.18. Quality metrics of sequenced data using GALAXY (4)
Sequence length distribution and sequence duplication levels of sample no. 1 (CONTROL), forward run

✔ **Overrepresented sequences**
No overrepresented sequences

✔ **Adapter Content**

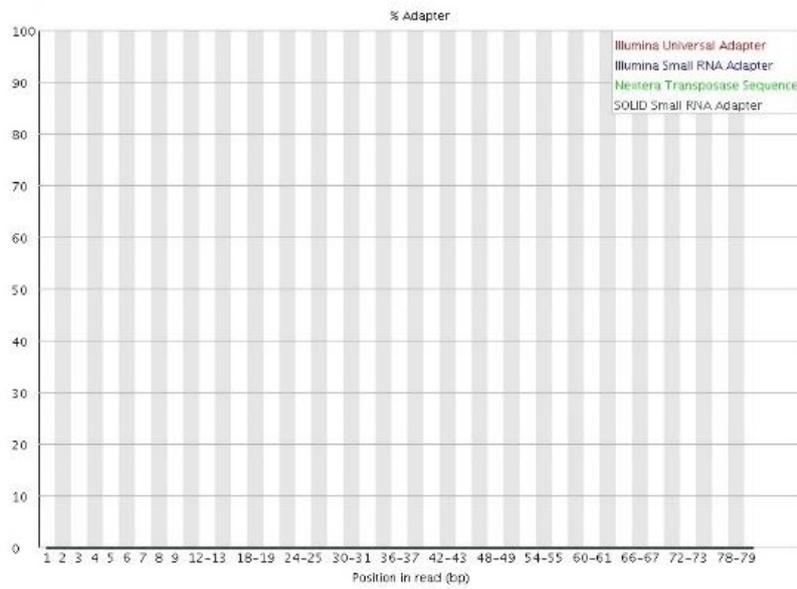
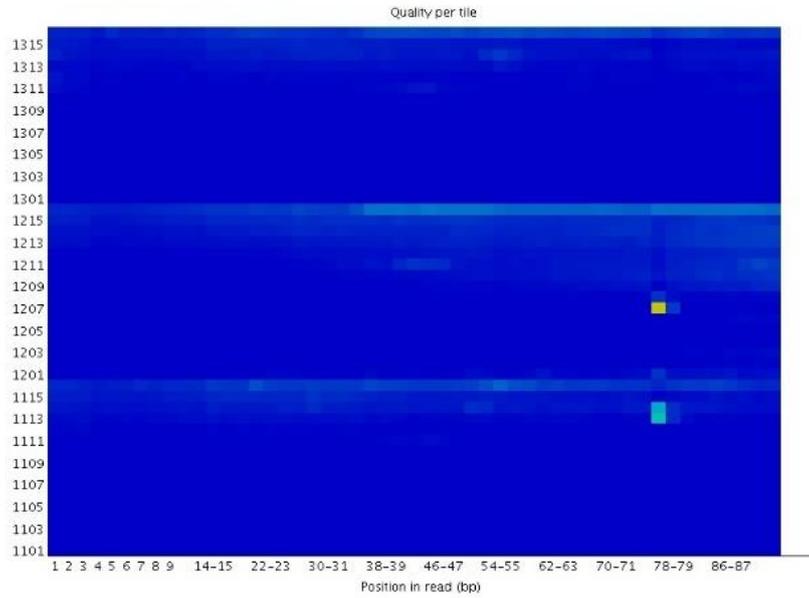


Figure 3.19. Quality metrics of sequenced data using GALAXY (5)
Overrepresented sequences and adaptor content, of sample no. 1 (CONTROL), forward run

Per tile sequence quality



Kmer Content

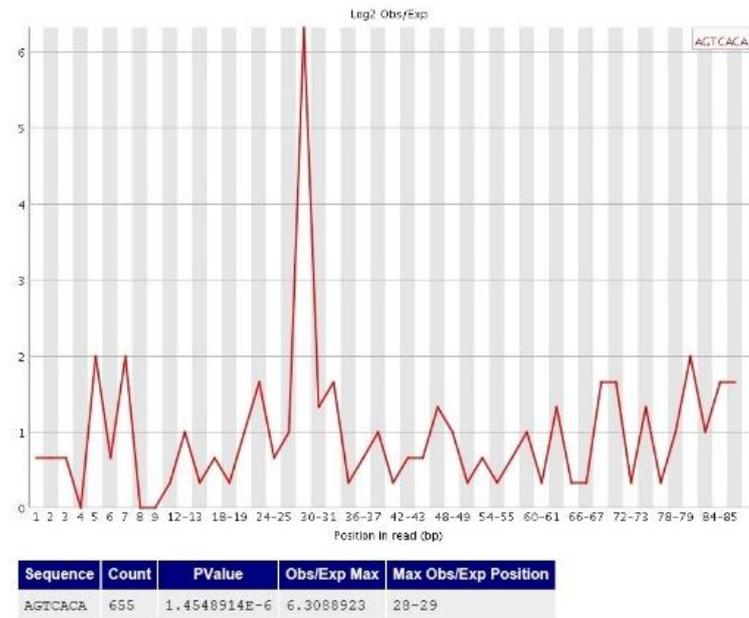


Figure 3.20. Quality metrics of sequenced data using GALAXY (6)
Per tile sequence quality and Kmer content of sample no. 1 (CONTROL), forward run

3.2.10 Mapping DNA sequence reads using Bowtie to the *Burkholderia cenocepacia* genome to reveal PrtI binding sites

Once the FASTQ files of the forward and reverse run of one sample were uploaded on GALAXY, fastQC analysis and fastgroomer was carried out. Then, the *B. cenocepacia* H111 chromosome sequences (NCBI GenBank accession no. HG938370, HG938371, and HG938372) in FASTA format was uploaded on GALAXY through Filezilla, a cross-platform file transfer protocol (FTP) software which assists uploading large data files. The groomed FASTQ uploaded files were mapped with the H111 genome using the 'Burrows-Wheeler Aligner (BWA) for Illumina' tool. Here, the files were in 'Sequence Alignment map (SAM)' format. Using the SAM to BAM conversion tool, the files were converted to 'Binary Alignment map (BAM)' format. BAM files can be downloaded and stored and are compatible with almost all genome visualisation softwares. Sequence metrics were studied using the Picard tool on GALAXY and duplicate reads were removed as required. Integrative Genomics Viewer (IGV) was used to map reads to the *Burkholderia cenocepacia* genome. IGV is a free and interactive genome visualisation software by Broad Institute of MIT and Harvard (United States) dedicated to map micro-array, NGS based datasets and genomic annotations. *B. cenocepacia* H111 genome (NCBI GenBank accession no. HG938370, HG938371, and HG938372) in fasta format was uploaded on IGV and the BAM files were mapped to the genome. The number of sequences aligning with a particular region on the genome were visualised, thus giving peaks where a higher than average number of sequences align. Typically, due to background DNA, some alignment is seen throughout the genome. However, where true enrichment has occurred a sharp peak can be observed. In this way, data from the remaining 5 samples was used to map on to the *B. cenocepacia* genome. Using samples 1 and 2, as the CONTROL and IP pair, at least a 100-fold enrichment in the region between PrtI and the upstream gene (I35_RS16285) was observed in IP as compared to CONTROL (Figure 3.21). Strong enrichment was also observed at the region corresponding to the *prtI* ORF. However, as mentioned before, this was expected due to the fact that pSRK-Km-PrtI-FLAG was present the cells. Therefore, the sequence between the translation start and stop codons of *prtI* should be considered as

background noise. True enrichment has occurred for the region located upstream from the start codon of *prtI* as this region was not present on the plasmid. The N-terminal coding sequence of *prtI* was most abundantly represented and the abundance tapered off towards the C-terminal coding end of the gene as the peak is a cumulative effect formed due to the combination of true enrichment peak + background contamination by plasmid pSRK-Km-PrtI-FLAG. In samples 3 & 4 (experiment 2) and 5 & 6 (experiment 3), at least a 40-fold increase was observed for IP sample compared to its respective CONTROL (Figure 3.22, Figure 3.23). This is because the number of wash steps was increased for these pairs of samples. Finally, a manual scan around the genome did not reveal any other apparent enrichments. This may suggest that PrtI has only one binding region on the genome.

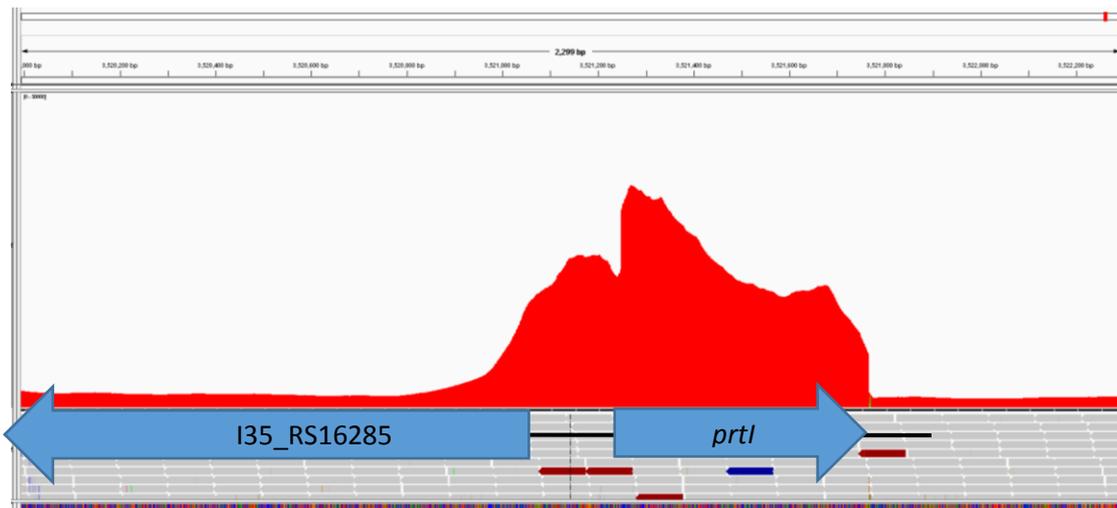


Figure 3.21. Mapping DNA sequence reads to the *B. cenocepacia* genome using IGV(1)
 Top panel represents CONTROL (experiment no.1) reads mapped to the H111 genome. Bottom panel represents IP (experiment no.1) reads mapped to the genome. Numbers on top of each panel represent base co-ordinates on chromosome 1. Sequence counts are shown in red. The location of *prtI* (I35_RS16290) and I35_RS16285 are shown in blue block arrows.

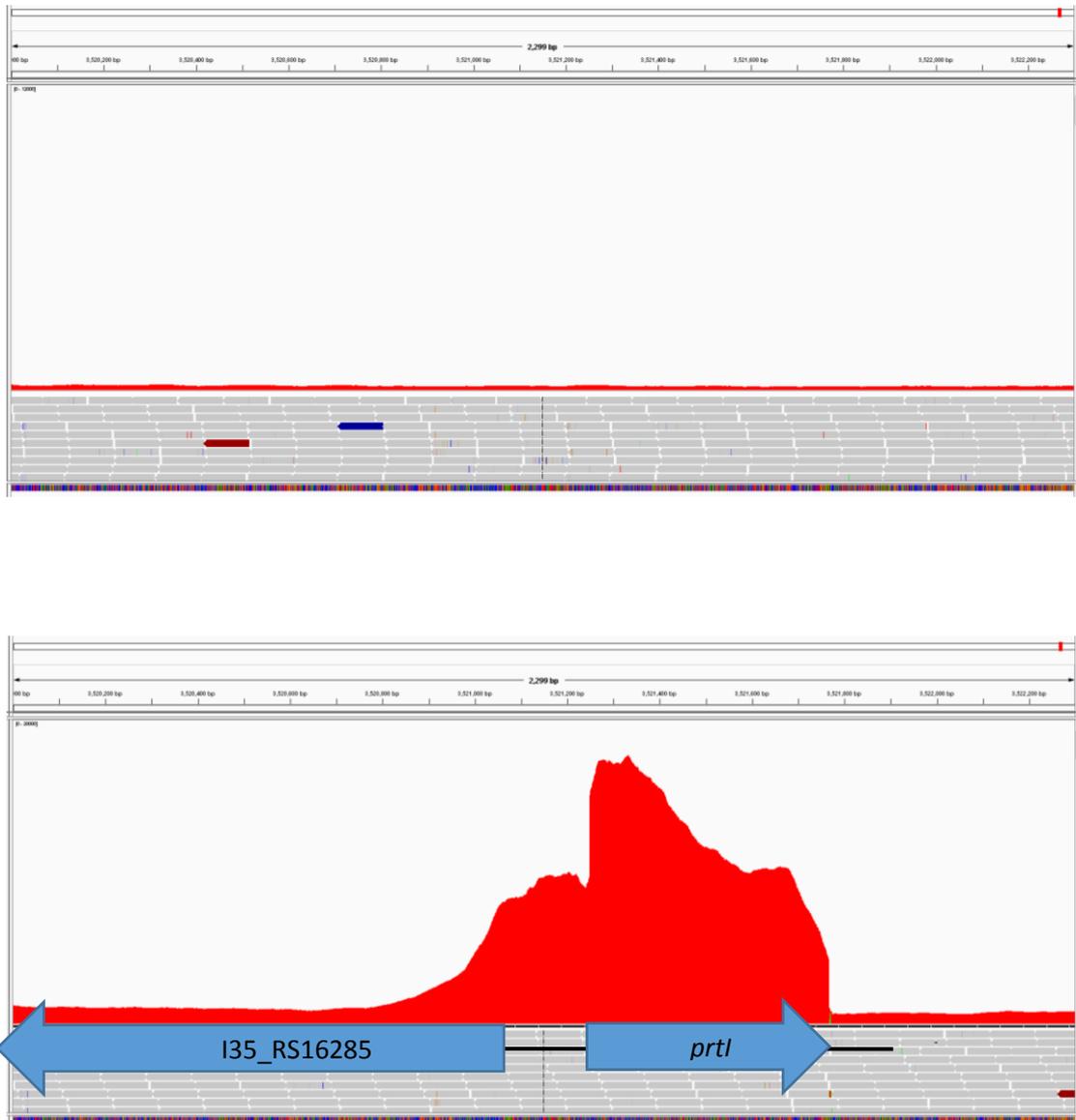


Figure 3.22. Mapping DNA sequence reads to the *B. cenocepacia* genome using IGV(2)
 Top panel represents CONTROL (experiment no.2) reads mapped to the H111 genome. Bottom panel represents IP (experiment no.2) reads mapped to the genome. Numbers on top of each panel represent base co-ordinates on chromosome 1. Sequence counts are shown in red. The location of *prtI* (I35_RS16290) and I35_RS16285 are shown in blue block arrows.

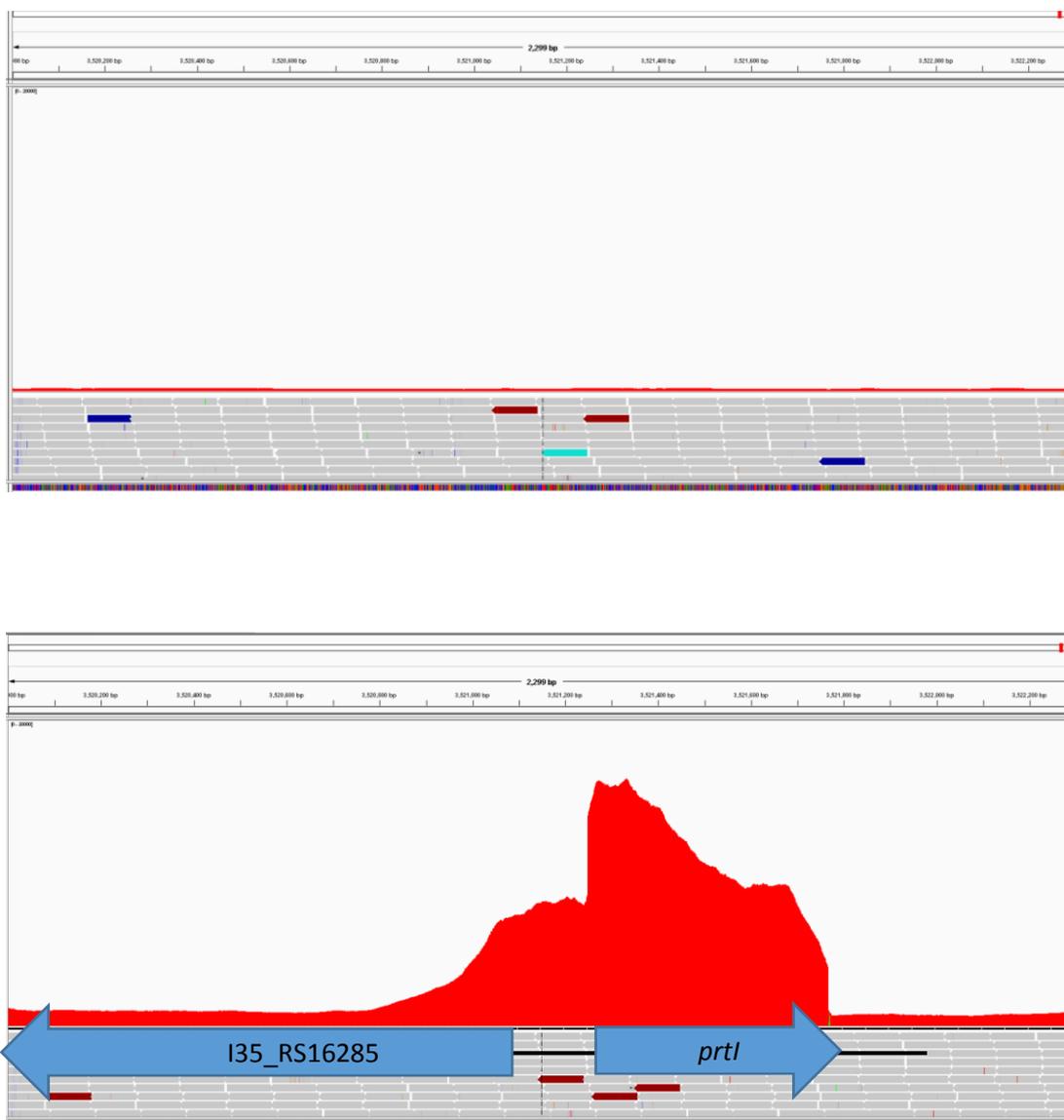


Figure 3.23. Mapping DNA sequence reads to the *B. cenocepacia* genome using IGV(3)
 Top panel represents CONTROL (experiment no.3) reads mapped to the H111 genome. Bottom panel represents IP (experiment no.3) reads mapped to the genome. Numbers on top of each panel represent base co-ordinates on chromosome 1. Sequence counts are shown in red. The location of *prtI* (I35_RS16290) and I35_RS16285 are shown in blue block arrows.

3.2.11 Peak calling and motif identification

MACS is a tool on GALAXY that allows peak calling by comparing IP to CONTROL and identifying a pre-determined 'm'-fold enrichment. The m value can be set to 10-40 depending on experimental requirements. This is particularly useful when studying large genomes such as the human genome. Once peaks are called, each of those DNA regions can be studied to identify a common motif within them using tools such as R-SAT or MEME-Chip. They help identify most conserved bases and are particularly useful to identify promoter motifs.

However, PrtI was observed to bind at only one genomic location (section 3.2.10), further bioinformatics analysis was not carried out and it was decided to study this promoter region by carrying out reporter-fusion assays.

3.2.12 *In silico* identification of putative PrtI regulated promoters at the *prtI* locus

Once the PrtI binding region had been identified by the ChIP-seq scan, it was decided to identify conserved motifs by aligning the promoter region with the corresponding region of other species that harbour *prtI* like genes. Figure 3.24A shows an alignment of the intergenic region between *prtI* and the gene located upstream (I35_RS16285) from *B. cenocepacia* with the same region from *B. cepacia*, *B. multivorans*, *B. pseudomallei*, *B. vietnamiensis*, *B. ambifaria* and *P. fluorescens*. Two pairs of conserved motifs were evident which have been highlighted in the Figure 3.24. Each pair was comprised of a GGAATAA and CGTC motif separated by a 14 basepair non-conserved spacer that could be putative -35 and -10 regions, respectively, that are putatively recognised for PrtI. The pair highlighted in green could constitute the *prtI* promoter whereas, the TTATTCC and GACG motifs, which correspond to the reverse complement of GGAATAA and CGTC could be the I35_RS16285 (catalase) promoter, highlighted in blue. Therefore, two divergently arranged PrtI-dependent promoters may be located in the I35_RS16285-*prtI* intergenic region.

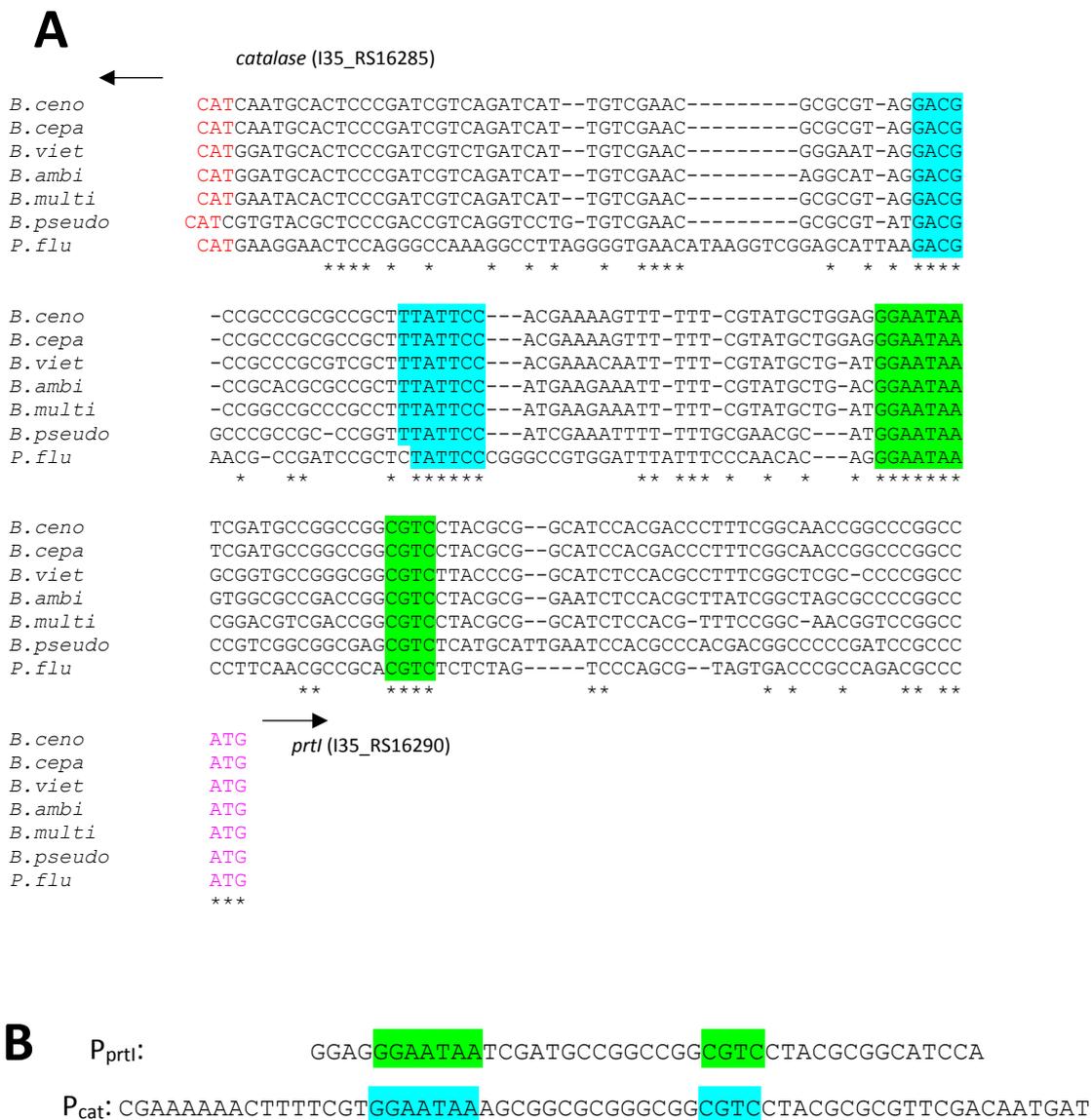


Figure 3.24. Clustal Omega alignment for identification of putative PrtI recognition sequences
 (A) Intergenic region between I35_RS16285 (*catalase*) and *prtI* from *Burkholderia cenocepacia* was aligned with the corresponding region from *Burkholderia cepacia*, *Burkholderia multivorans*, *Burkholderia pseudomallei*, *Burkholderia vietnamiensis*, *Burkholderia ambifaria* and *Pseudomonas fluorescens*. The stop codon of '*catalase*' is in red font while the start codon of *prtI* is in magenta. Bases that are conserved at the corresponding position in all sequences are marked with an asterisk. Putative pairs of divergent -10 and -35 promoter regions are highlighted in blue and green. (B) DNA sequences that were hypothesised to form the *prtI* promoter (P_{prtI}) and *catalase* (I35_RS16285) promoter (P_{cat}) where the putative -10 and -35 regions are highlighted in green and blue as in the alignment above.

3.2.13 Construction of reporter vector pKAGd4 harbouring putative Prtl regulated promoters

To test the possibility that there are a pair of divergent Prtl-dependent promoters located between I35_RS16285 and *prtl* that consist of GGAATAA and CGTC motifs, one DNA fragment containing the putative *prtl* promoter and one DNA fragment containing the *catalase* promoter were cloned into the transcription reporter vector pKAGd4 (sequences shown in Figure 3.24B). To do so, complementary single stranded oligonucleotide pairs 'prtlproFor2 and prtlproRev2' and 'Pcatfor and Pcatrev' were annealed as described in section 2.2.8. Each of the resulting double-stranded oligonucleotides were then ligated with pKAGd4 digested with *HindIII* and *BamHI*. MC1061 cells were transformed with these ligations, the colonies were screened for the presence of the plasmid with the desired insert using GoTaq PCR with primers AP10 and AP11 and the DNA sequence of the insert in the positive clones was confirmed.

3.2.14 Analysis of putative Prtl regulated promoter activity in *E. coli* and *B. cenocepacia*

pSRK-Km-Prtl-FLAG was transferred to *E. coli* MC1061 containing each of the reporter plasmids harbouring the promoter derivative. As a control, empty pSRK-Km was used. Promoter activity was measured by carrying out β -galactosidase enzyme assays on cells growing under conditions for induction of the *lac* promoter on pSRK-Km (addition of IPTG) and under non-inducing conditions. Both DNA fragments were found to contain a highly active promoter in cells growing under inducing conditions in the presence of Prtl-FLAG, where there seemed to be an approximately five to eight-fold increase in activity compared to the non-induced condition (Figure 3.25). Also, it was apparent that there was no promoter activity in the absence of Prtl in case of P_{prtl} and P_{cat}.

To verify if the native, un-tagged *B. cenocepacia* Prtl was capable of switching on the putative promoters under standard laboratory conditions, initially, reporter plasmids pKAGd4 harbouring the promoter derivatives were transferred into *B.*

cennocepacia H1111 by conjugation using S17-1 as a donor strain and the β -galactosidase activity was measured. It was found that promoters P_{prtI} and P_{cat} were inactive under normal growth conditions (i.e. in LB) which suggests that $PrtI$ may be activated only under certain 'stressful' conditions or under an appropriate extra-cytoplasmic signal as is the case for many ECF σ factors (Figure 3.26). pSRK-Km- $PrtI$ -FLAG and pSRK-Km were then transferred to *B. cennocepacia* H1111 containing the reporter plasmids and the β -galactosidase activities were measured. Both promoters now were found to be active in the presence of the plasmid expressing *prtI*, though approximately 15 times less active than in *E. coli*, whereas, there was no activity observed in the absence of $PrtI$. The results support the idea that the conserved GGAATAA(N)₁₄CGTC motifs form the $PrtI$ -dependent promoters.

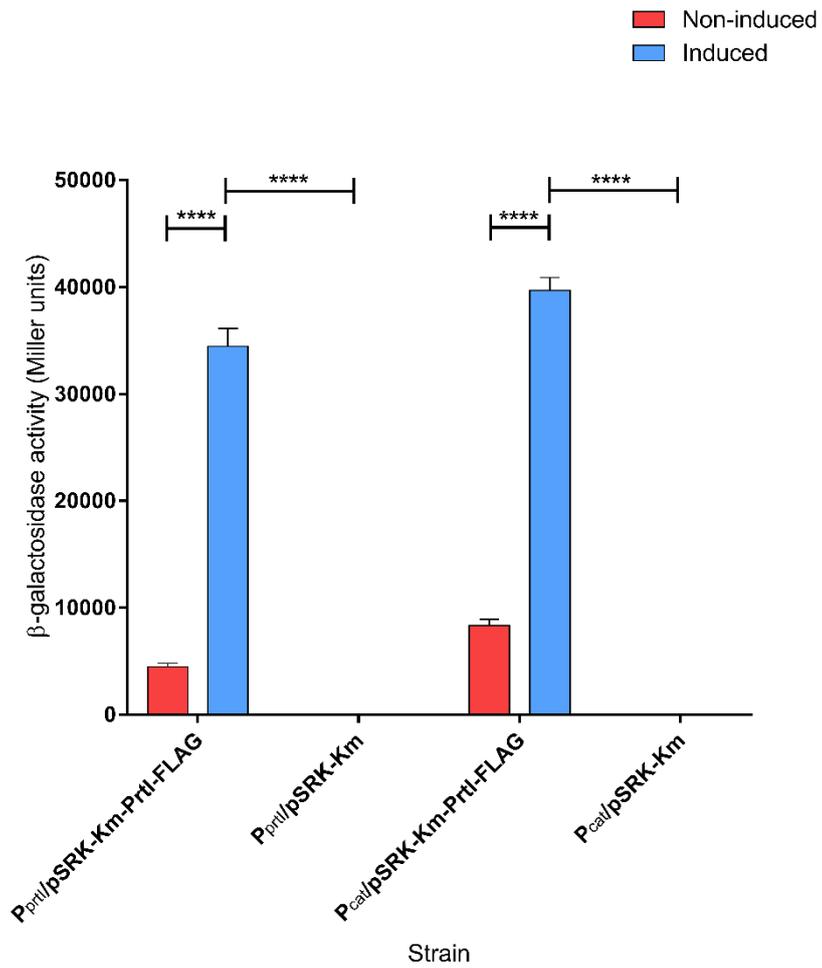


Figure 3.25. Activity of P_{prtI} and P_{cat} in the presence and absence of PrtI-FLAG in *E. coli*
 pKAGd4 containing the indicated promoter derivative were each separately transferred to MC1061 along with either pSRK-Km or pSRK-Km-PrtI-FLAG and promoter activity was measured using the β -galactosidase assay where cultures were grown in LB with kanamycin and chloramphenicol under IPTG induced and induced conditions until log phase. Activities shown were obtained by subtracting activity of the same strain containing pKAGd4 assayed under the same conditions (background controls). Statistical significance was determined by performing a two-way ANOVA, with Tukeys post-test, **** = $p < 0.0001$ (n=6).

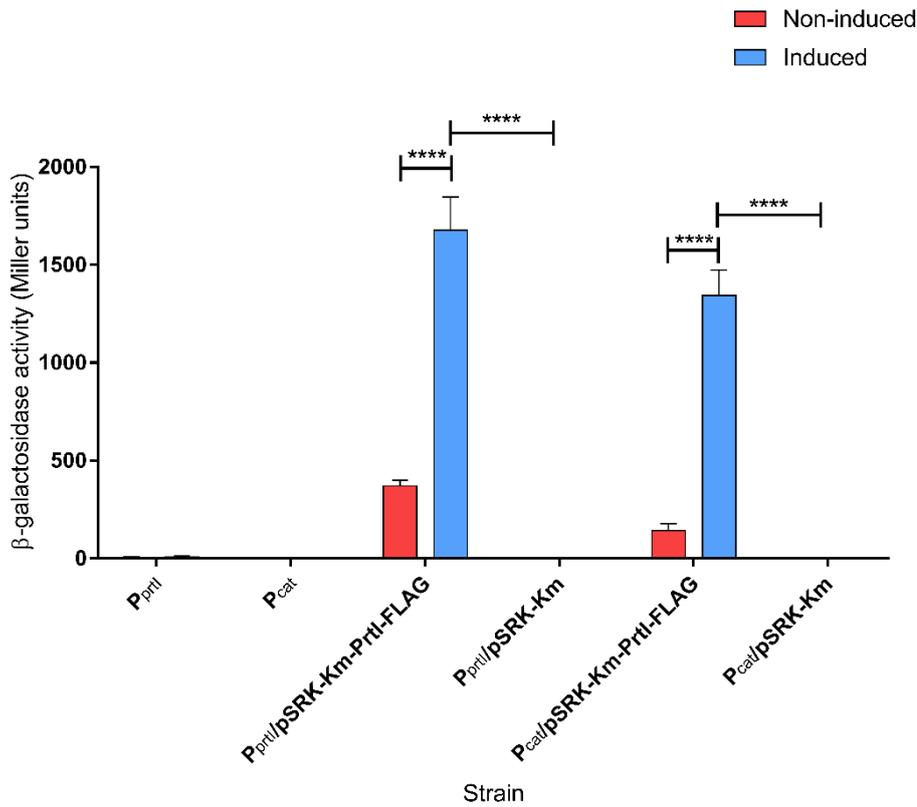


Figure 3.26 Activity of P_{prtI} and P_{cat} in the presence and absence of PrtI-FLAG in *B. cenocepacia*: pKAGd4 containing P_{prtI} or P_{cat}, as indicated, were each separately transferred to H111 either on their own or along with pSRK-Km or pSRK-Km-PrtI-FLAG and the promoter activity was measured using the β -galactosidase assay where cultures were grown in LB with kanamycin and chloramphenicol under IPTG induced and induced conditions. Activities shown were obtained by subtracting the activity of the same strain pKAGd4 assayed under the same conditions (background controls). Statistical significance was determined by performing a two-way ANOVA test, with Tukeys post-test, **** = $p < 0.0001$ (n=6).

3.2.15 Construction of plasmids expressing native PrtI

Plasmids expressing the native (i.e. untagged) form of PrtI were constructed to study the activity of putative PrtI-dependent promoters to verify that the activity obtained with the FLAG tagged σ factor was replicable. The broad host-range plasmid, pBBR2 was used to express PrtI. pBBR2 can constitutively express *prtI* when cloned downstream from the *lacZ* promoter when present in an *E. coli* strain lacking the *lac* repressor, such as MC1061, or in *B. cenocepacia*. Moreover, the constitutive promoter of the kanamycin resistance gene augments the expression of the cloned gene from the *lac* promoter. The *prtI* gene was amplified separately with two different forward primers PrtIforA and PrtIforB, but the same reverse primer, PrtIrev, cut with *Hind*III and *Bam*HI, and ligated to pBBR2 cut with the same enzyme. This generated two pBBR2 derivatives expressed *prtI* that differed with respect to the UTR located upstream of the *prtI* start codon. The PrtIforA primer was designed such that the amplicon contained the putative *prtI* promoter and Shine-Dalgarno sequence giving rise to pBBR2-PrtI while PrtIforB omitted the *prtI* promoter and contained an improved Shine-Dalgarno sequence generating pBBR2-PrtI(b). In both cases the 5' end of the forward primer contained a stop codon which could terminate read-through translation of *lacZ* into the cloned DNA. JM83 cells were transformed with the ligations and then the colonies obtained were screened by using M13for and M13rev primers. Plasmid clones with the expected insert were checked for their overall size (pBBR2-PrtI = 5.667 kb and pBBR2-PrtI(b) = 5.654 kb) by agarose gel electrophoresis and the DNA integrity was confirmed by nucleotide sequencing.

3.2.17 Analysis of the activity of $P_{prtI\text{full}}$, P_{prtI} , and P_{cat} in the presence and absence of native PrtI in *E. coli*

pKAGd4 derivatives containing $P_{prtI\text{full}}$, P_{prtI} , and P_{cat} were transferred to MC1061 cells along with either pBBR2-PrtI or pBBR2-PrtI(b) or empty vector pBBR2. The promoter activity was then measured by performing β -galactosidase assays. All three promoters were significantly active in the presence of either plasmid harbouring *prtI*

whereas strains containing empty pBBR2 vector exhibited only background levels of β -galactosidase activity (Figure 3.27). Except for P_{prtI} , all the promoters were at least two-fold more active in the presence of pBBR2-Prtl as compared to pBBR2-Prtl(b). Overall, these results were consistent with the activities obtained under the conditions used to carry out the ChIP-seq experiment and they confirm conserved motifs GGAATAA(N)₁₄CGTC to be the promoters that Prtl must be recognising. This suggested that Prtl might regulate at least two genes, I35_RS16285, a predicted catalase, along with transcription of the gene itself. This also suggested that Prtl might have role in oxidative stress response, as it regulates a putative catalase.

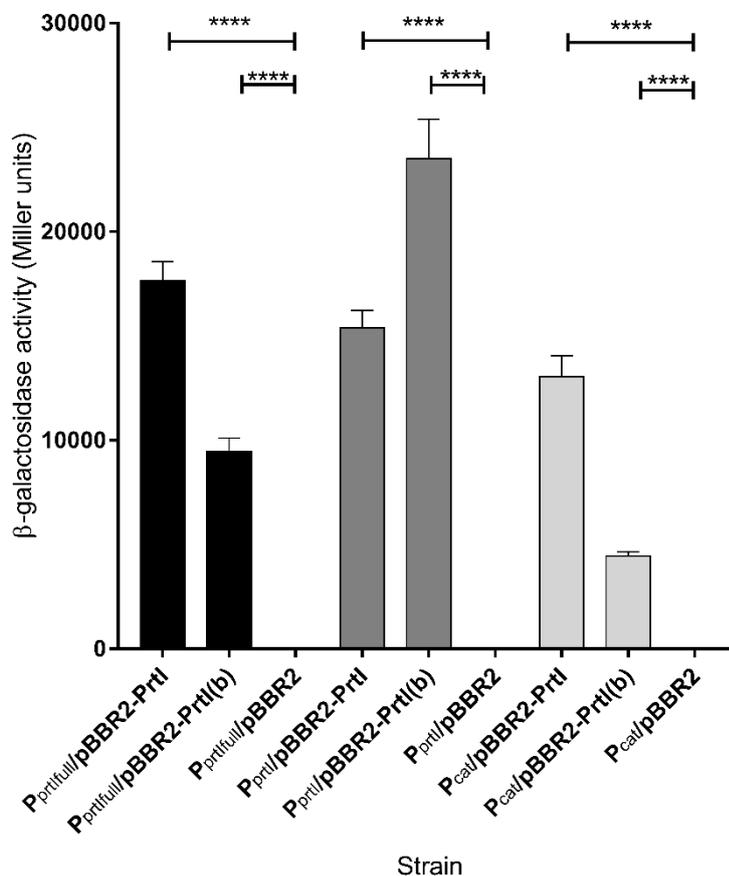


Figure 3.27. Activity of $P_{prtIfull}$, P_{prtI} and P_{cat} in the presence and absence of native Prtl in *E. coli* pKAGd4 containing $P_{prtIfull}$, P_{prtI} and P_{cat} as indicated, were each separately transferred to MC1061 either on their own or along with pBBR2, pBBR2-Prtl or pBBR2-Prtl(b) and promoter activity was measured using the β -galactosidase assay where cultures were grown in LB with kanamycin and chloramphenicol. Activities shown were obtained by subtracting the activity of the same strain containing pKAGd4 assayed under the same conditions (background control). Statistical significance was determined by performing a one-way ANOVA test, with Tukeys post-test, **** = $p < 0.0001$ (n=6).

3.3 Systematic study of the DNA sequence requirements for promoter recognition by P_{prtI}

As a putative σ factor, P_{prtI} initiates transcription by recognising key DNA bases at its target promoters and makes contacts with them. In order to study the specific DNA bases that might be involved in promoter recognition, single nucleotide ‘extreme’ base pair substitutions were introduced at key positions within P_{prtI}. P_{prtI} is a 28 bp derivative that is active in the presence of P_{prtI} suggesting that it might constitute the minimum sequence required for recognition P_{prtI}. The nucleotide sequence alignment of the *catalase-prtI* intergenic region (Figure 3.24) revealed key bases ‘GGAATAA’ and ‘CGTC’ which were conserved across different related species suggesting that they might form the ‘-35’ and ‘-10’ promoter regions. Therefore, the experimental design consisted of making a series of P_{prtI} promoter derivatives with single base substitutions within and adjacent to the predicted -35 and -10 promoter regions and measuring their activity in the presence of P_{prtI}.

3.3.1 Construction of plasmids harbouring single base substituted P_{prtI} promoter derivatives

Initially, 15 derivatives (ds1 to ds15) of P_{prtI} were designed as shown in Figure 3.28A. To make ‘extreme’ substitutions, a purine base was substituted by a pyrimidine that it does not pair with and *vice versa*. Therefore, a guanine was replaced by thymine and *vice versa*; and cytosine was replaced by an adenine and *vice versa*. Up to one base on either side of the predicted -35 and -10 regions were chosen to determine the promoter limits of the core elements. Appropriate complementary oligonucleotides (e.g. ‘Pprtids1for’ and ‘Pprtids1rev’ for ds1) were annealed and ligated into the reporter plasmid pKAGd4 between the *HindIII* and *BamHI* restriction enzyme sites. Positive clones were checked for their sequence integrity.

3.3.2 Analysis of DNA sequence requirements for promoter utilisation by P_{prtI}

Each pKAGd4 derivative harbouring a single base substituted P_{prtI} variant was transferred to *E. coli* MC1061 along with pBBR2-P_{prtI} and the activity was measured by

performing β -galactosidase assays (Figure 3.28B). In the predicted -35 region of GGAATAA, base substitution at each of the first six positions (GGAATA) significantly affected the promoter activity. However, the largest effect on promoter activity occurred when the second G and first A (positions 2 and 3 of the motif) were substituted. Substitution at position 7 of the predicted -35 element also caused a decrease in promoter activity, but the promoter remained very active (75% of WT). When single base substitutions were introduced at the base positions immediately upstream and downstream of the GGAATAA motif there was no effect on promoter activity at the upstream position. But, substitution of the T residue, located adjacent to position 7 of the -35 motif caused an approximate 40% decrease in promoter activity suggesting that it might be an important base even though it is not part of the conserved GGAATAA motif. The next base downstream of the T (the C at position 10 of the promoter), was therefore substituted. The activity of this promoter mutant was measured and was found to be similar to that of the WT. At the predicted -10 promoter region, substitution at the G, T and C bases of the CGTC motif exerted severe reduction in promoter activity, with the T base being extremely important as the promoter activity was completely abolished when it was substituted. However, substitution of the C base at first position of the CGTC motif did not cause a significant reduction in promoter activity. Overall, these results suggest that GGAATA and GTC motifs may constitute the -35 and -10 regions, respectively. The bases G, A, in the -35 region and base T in the -10 region (positions 3, 4 and 25, respectively) might be the most important for efficient promoter utilisation.

A

```

WT:      GGAG GGAATAATCGATGCCGGCCGG CGTC CTACGCGGCATCCA
ds1     GGAT GGAATAATCGATGCCGGCCGG CGTC CTACGCGGCATCCA
ds2     GGAG TGAATAATCGATGCCGGCCGG CGTC CTACGCGGCATCCA
ds3     GGAG GTAATAATCGATGCCGGCCGG CGTC CTACGCGGCATCCA
ds4     GGAG GGCATAATCGATGCCGGCCGG CGTC CTACGCGGCATCCA
ds5     GGAG GGA CTAATCGATGCCGGCCGG CGTC CTACGCGGCATCCA
ds6     GGAG GGAAG AATCGATGCCGGCCGG CGTC CTACGCGGCATCCA
ds7     GGAG GGAAT CATCGATGCCGGCCGG CGTC CTACGCGGCATCCA
ds8     GGAG GGAATAC TCGATGCCGGCCGG CGTC CTACGCGGCATCCA
ds9     GGAG GGAATAAG CGATGCCGGCCGG CGTC CTACGCGGCATCCA
ds9.1   GGAG GGAATAAT A GATGCCGGCCGG CGTC CTACGCGGCATCCA
ds10    GGAG GGAATAATCGATGCCGGCCGG TCGTC CTACGCGGCATCCA
ds11    GGAG GGAATAATCGATGCCGGCCGG A GTC CTACGCGGCATCCA
ds12    GGAG GGAATAATCGATGCCGGCCGG CTC CTACGCGGCATCCA
ds13    GGAG GGAATAATCGATGCCGGCCGG CGGC CTACGCGGCATCCA
ds14    GGAG GGAATAATCGATGCCGGCCGG CGTACTACGCGGCATCCA
ds15    GGAG GGAATAATCGATGCCGGCCGG CGTCA TACGCGGCATCCA
  
```

B

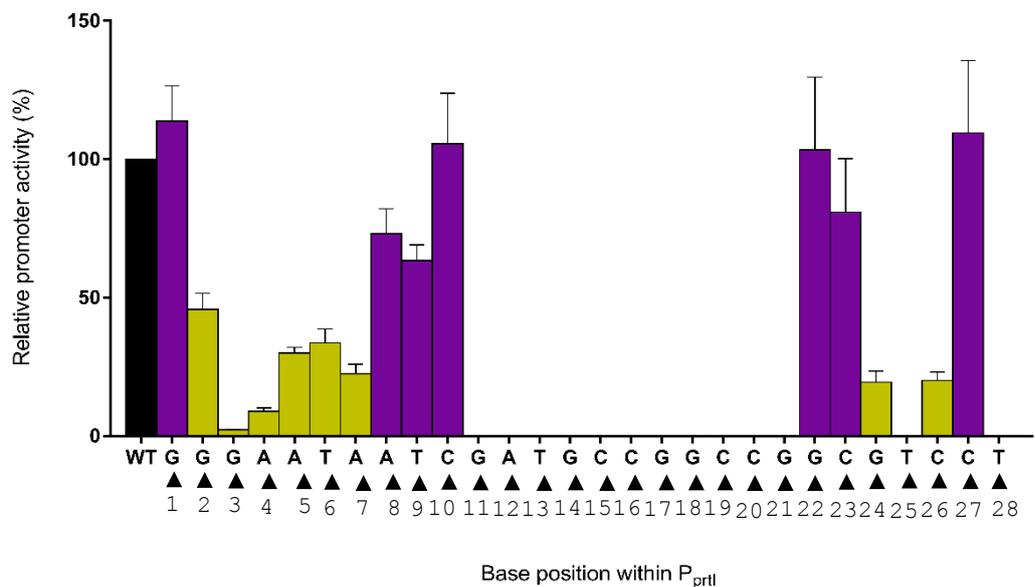


Figure 3.28. Identification of the -35 and -10 promoter elements of a Prtl-dependent promoter: (A) P_{prtl} promoter substitution derivatives to determine the most important bases for recognition by P_{prtl}. Highlighted in green are the predicted -35 or -10 regions. Highlighted in red is the extreme base substitution in each promoter derivative. (B) pKAGd4-P_{prtl} and pKAGd4-P_{prtl}ds1 to pKAGd4-P_{prtl}ds15 were each separately transferred to MC1061 along with pBBR2-PrtI and promoter activity was measured using the β -galactosidase assay where cultures were grown in LB containing kanamycin and chloramphenicol. Activities shown were obtained by subtracting the activity of the same strain carrying pKAGd4 assayed under the same conditions (background control). Numbers under X-axis refer to positions of bases described in the text. The activities are represented as the relative percentage assuming the activity of P_{prtl} was 100% (black bar). Where a substitution causes $\geq 50\%$ reduction in activity relative to the WT promoter, the activity bar is shown in dark yellow.

3.4 FUZZNUC based promoter search in other bacteria

A FUZZNUC software based search using Prosite patterns was carried out on the *Burkholderia cenocepacia* genome to identify bioinformatically whether PrtI-dependent promoters existed anywhere else in the genome. Based on the identification of the -10 and -35 promoter regions, the following Prosite pattern was used in the search:

GGAATAN(16)GTC

Unlike BLAST, this search tool allows use of 'N' or bases that can be replaced, bases that should be included and bases that have to be excluded. However, other than the *prtI* and *catalase* promoters no other relevant matches were obtained. One additional match was identified on chromosome 2, located between two genes encoding a peptidoglycan associated outer membrane protein and a putative dihaem cytochrome C-553. This region did not seem to contain a promoter for either gene as the predicted -35 and -10 motifs were present downstream of the stop codons for both genes. Moreover, this region was not identified as a PrtI binding site in the ChIP-seq screen. As gene clusters that are highly homologous to the *B. cenocepacia* '*cytb561-cat-prtI-prtR*' gene cluster are present in other bacterial genera, including *Pseudomonas* and *Xanthomonas*, FUZZNUC was used to screen for PrtI-dependent promoters. As an example, *Pseudomonas fluorescens* genome was scanned using the Prosite pattern GGAATAN(16)GTC, which gave two expected hits corresponding to the *prtI* and *catalase* promoters, but no additional matches were obtained. Nevertheless, this pattern can be used in other organisms to search for potentially novel promoters.

3.5 Genetic analysis of the role of PrtI

In the second approach to identify the role of PrtI, it was decided to investigate the phenotypic effects of deleting *prtI* and *prtR*. In order to avoid polar effects, in-frame marker-less deletion mutants in *B. cenocepacia* H111 strain were constructed. Based on current literature, PrtI orthologues in other bacterial species have roles in protease regulation, biofilm formation, surface motility etc., as described in section 3.1. Based on the results described in section 3.2 it was evident that PrtI regulated the gene upstream

of *prtI*, i.e. I35_RS16285, which is predicted to encode a catalase enzyme based on BLASTP analysis of the encoded product. Hence, the objective of constructing deletion mutants was particularly to observe the effects of oxidative stress amongst the above mentioned phenotypes.

3.5.1 Construction of a *prtI* in-frame deletion mutant

Four SOE PCR primers PrtI_{delA}, PrtI_{delB}, PrtI_{delC} and PrtI_{delD} were designed as described in section 2.2.10.2 such that 'AB' gave a 525 bp product that included the first six codons of *prtI*, the UTR located upstream of *prtI* and a segment of the catalase gene, while CD generated a 548 bp product that included the last six codons of *prtI*. In a subsequent PCR, using the AB and CD products as template with A and D primers, a final product of 1073 bp was generated, such that 162 codons of the central region of *prtI* (total gene length 174 codons) was deleted. This product was cloned into pEX18-TpTer-*pheS* between *Bam*HI and *Hind*III sites and its integrity was confirmed by DNA sequencing. In this plasmid the *pheS* gene is under control of a promoter that does not allow cells to survive in the presence of chlorophenylalanine. The resulting plasmid was transferred to the *E. coli* SM10 donor strain and subsequently delivered to *B. cenocepacia* H111 by conjugation. As pEX18-TpTer-*pheS* cannot replicate in *B. cenocepacia* selection of exconjugants on medium containing trimethoprim allows for identification of recombinants in which the plasmid has integrated into H111 genome. This gave rise to 'co-integrates' containing both the WT version of *prtI* as well as the deleted version of the gene. This was confirmed by carrying out a PCR screen using vector specific primers pEX18Tpfor and pEX18Tprev. Using the chlorophenylalanine based counter-selection technique, co-integrates were streaked on M9 medium containing chlorophenylalanine. However, several attempts did not give rise to mutants and the selection was concluded to be inefficient due to a possibly leaky promoter. Hence, an alternate mode of selection was used with a modified plasmid called pEX18-TpTer-*Scel-pheS* (Fazli et al., 2015, H. Spiewak, 2015, unpublished). Here, the plasmid contained an I-*Scel* meganuclease restriction site. Therefore, once the PCR fragment AD containing the *prtI* deletion allele and flanking sequences was cloned into pEX18-TpTer-

Scel-*pheS* between *Bam*HI and *Hind*III sites, and its integrity confirmed, the plasmid was transferred into *B. cenocepacia* H111 by conjugation. Putative co-integrates were screened as before using pEX18Tpfor and pEX18Tprev. Next, plasmid pDAI-Scel, which expresses the I-Scel meganuclease, was conjugally transferred to a co-integrate, where the I-Scel enzyme was expressed resulting in the introduction of a double-stranded break at the I-Scel site located at the *prtI* locus in chromosome 1. Repair of the break by homologous recombination gave rise to either WT or *prtI* deletion mutants as shown in Figure 3.29. Colonies that lost the integrated pEX18-TpTer-Scel-*pheS* plasmid were selected for on medium containing tetracycline and were PCR screened using primers Prtloutfor2 and Prtloutrev2 that anneal to genomic sequences outside the region complementary to primers A and D. As a control, a WT colony was used as template which amplified a product of 1.680 kb whereas H111 Δ *prtI* gave rise to a product of 1.194 kb (Figure 3.30A). The mutant was streaked 0.1% chlorophenylalanine media 2-3 times to remove pDAI-I-Scel. The selected mutant was re-streaked 2-3 times on rich medium to purify it and streaked out on M9-glucose agar for medium term bench storage. For carrying out complementation experiments, pBBR2-Prtl constructed in section 3.2.13 was to be used.

3.5.2 Construction of a *prtR* in-frame deletion mutant

Two pairs of SOE PCR primers (*prtRdelA* and *prtRdelB*, and *prtRdelC* and *prtRdelD*) were used separately to amplify genomic H111 DNA using Q5 polymerase to give products of 523 and 553 bp in length that corresponded to regions located upstream and downstream of *prtR*, respectively. In the second PCR where 'AB' and 'CD' were used as template in equimolar concentration with A and D as primers, the final product was 1076 bp, with 239 codons of the central region of *prtR* (total 253 codons) deleted. This DNA fragment was cloned into pEX18-TpTer-Scel-*pheS*-Cm (a derivative of pEX18-TpTer-Scel-*pheS* containing a Cm^R selection marker) between the *Bam*HI and *Hind*III sites and the integrity of the cloned DNA was checked by nucleotide sequencing. pEX18-TpTer-Scel-*pheS*-Cm- Δ *prtR* was transferred to H111 by conjugation using *E. coli* SM10 and recombinants containing chromosomally inserted plasmid co-integrates were

selected for on medium containing chloramphenicol and trimethoprim and confirmed by PCR screening using primers pEX18Tpfor and pEX18Tprev. Following this pDAI-SceI was introduced into a co-integrate strain and ex-conjugants were selected on medium containing tetracycline. These should have lost the chromosomally integrated plasmid due to recombination that repaired the double stranded DNA break caused by the I-SceI enzyme. Colonies where excision of the plasmid had resulted in exchange of the *prrR* deletion allele for the WT allele were screened for by PCR with primers prtRoutfor and prtRoutrev that anneal to genomic sequences outside the region complementary to primers A and D. The efficiency of mutant generation by this technique is usually 1 in 10 in our experience. However, a *prrR* deletion mutant was obtained with an efficiency of 1 in 100. This might suggest that deleting the putative anti- σ factor gene may have made PrtI constitutively active, thereby making the cells 'sick' and leading to formation of small colonies which were overlooked during the screening process. As seen in Figure 3.30B WT (H1111) used as a template gave a product of 2021 bps while H1111- Δ *prrR* used as a template gave a product of 1304 bps.

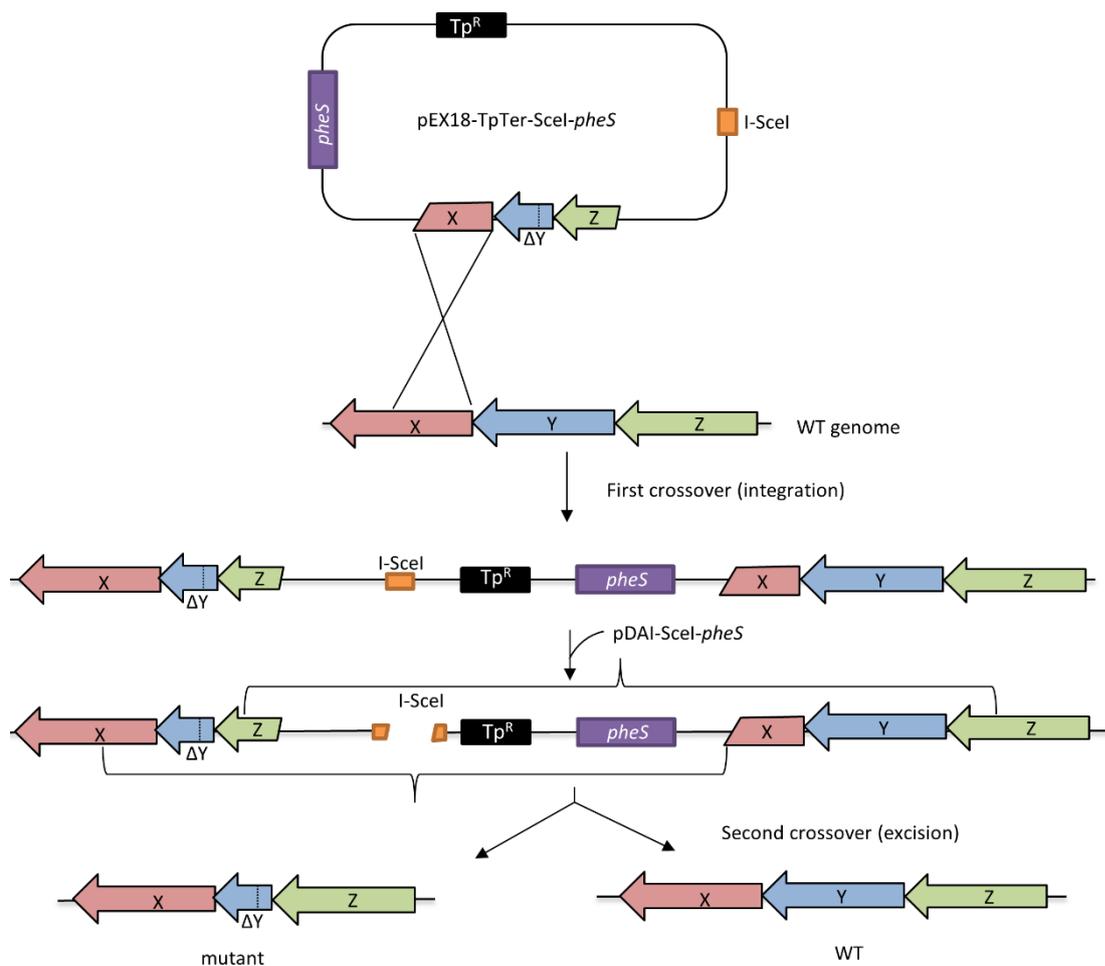


Figure 3.29. Schematic representation of mutant construction using pEX18TpTer-Scel-*pheS*
 SOE PCR was used to amplify a DNA fragment harbouring the deletion gene, ΔY , which was flanked by ≥ 500 bp of DNA on either side (here represented by fragments of genes X and Z) and cloned into pEX18-TpTer-Scel-*pheS*. Following introduction into *B. cenocepacia*, the plasmid can only be maintained by integration into the chromosome by homologous recombination, giving rise to two versions of gene Y – WT and ΔY . After introducing pDAI-Scel-*pheS* into the co-integrate strain, a cut was introduced at the I-SceI site, allowing a second homologous recombination to occur which rescues the bacterium. Depending upon the site of the second recombination event, either WT version or deleted version of the gene was retained.

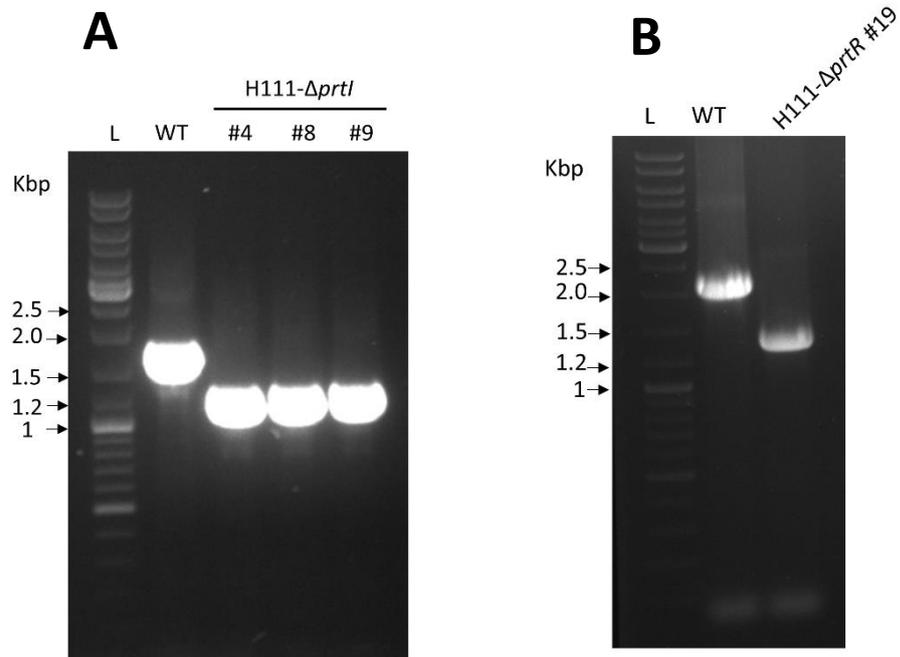


Figure 3.30. PCR screening of candidate H111ΔprtI and H111ΔprtR mutants

(A) PCR screen of candidate H111ΔprtI mutants #4, #8 and #9 using primers Prtloutfor2 and Prtloutrev2; L = GeneRuler DNA mix; WT = H111 boiled lysate used as template. (B) PCR screen of candidate H111ΔprtR #19 using primers PrtRoutfor and PrtRoutrev; L = GeneRuler DNA mix; WT = H111 boiled lysate used as template.

3.5.3 Construction of a plasmid for complementing the $\Delta prtR$ mutant

To make a plasmid that could complement H111- $\Delta prtR$, in case any phenotypic effects associated with deletion of *prtR* were due to polar effects of the lesion or an adventitious event that occurred during construction of the mutant, the WT *prtR* gene was cloned into pBBR2. Using primers PrtRfor and PrtRrev, *prtR* (793 bp) was amplified with its native promoter and ligated to the *Hind*III and *Xba*I sites of pBBR2. The size of the plasmid (5.907 kb) and sequence integrity of the cloned DNA was confirmed and the plasmid (pBBR2-PrtR) was transferred to H111- $\Delta prtR$ by carrying out conjugation using S17-1 as the donor strain.

3.5.4 Effect of deletion of *prtI* or *prtR* on bacterial growth under normal conditions

To assess whether *prtI* or *prtR* were required by cells under standard laboratory growth conditions, liquid cultures were grown and optical density was measured as described in section 2.1.4. Additionally, the efficiency of forming colonies was investigated. As shown in Figure 3.31, *prtI* and *prtR* deletion mutants grew similar to the wild type and also formed colonies with similar efficiency to that of the wild type. However, it was also observed that the $\Delta prtR$ mutant occasionally formed different sized colonies, though this was not consistent. The underlying mechanism attributing to this was not fully investigated. However, it was hypothesised that deletion of the anti- σ factor made cells slightly 'sick' due to PrtI being constitutively active.

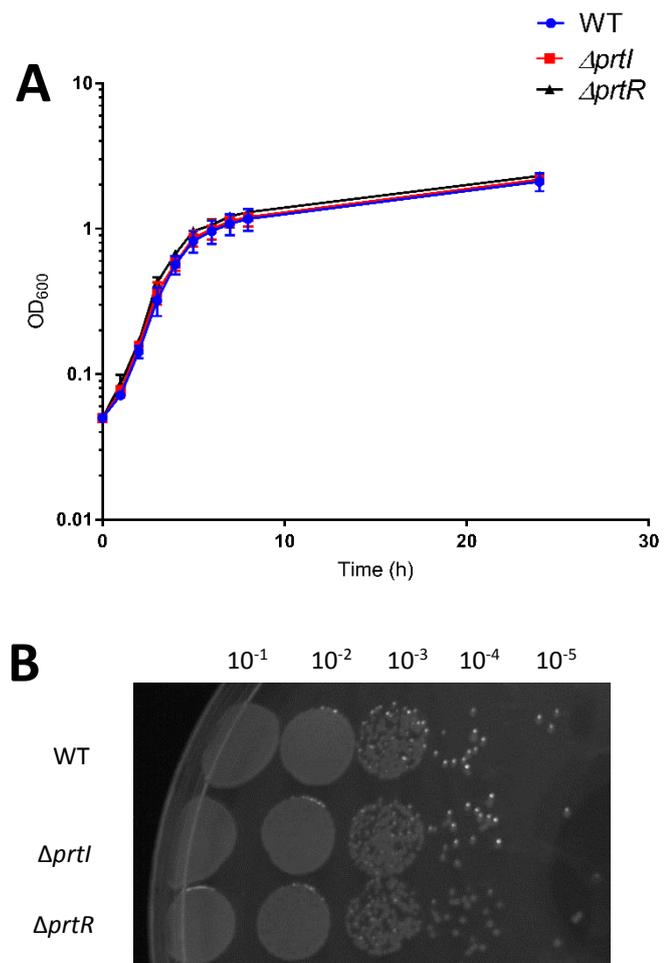


Figure 3.31. Effect of deletion of *prtI* and *prtR* on growth under normal conditions

(A) Growth curves of WT (H111), H111- $\Delta prtI$ and H111- $\Delta prtR$ in LB, n=9. (B) Representative image of serially diluted log phase cultures of WT (H111), H111- $\Delta prtI$ and H111- $\Delta prtR$ grown in LB at 37°C were plated on LB agar.

3.5.5 Effect of deletion of *prtI* or *prtR* on bacterial growth under oxidative stress conditions

The CHIP-seq and reporter plasmid fusion analyses showed that *PrtI* regulates the gene located upstream of *prtI*, i.e. I35_RS16285. I35_RS16285 is predicted to be a catalase based on genome sequence annotation (BLASTP search). Therefore, it was decided to study the effects of *prtI* and *prtR* deletion on the ability of *B. cenocepacia* to respond to oxidative stress. Initially hydrogen peroxide was used as an oxidant and a variety of different experimental designs were tested to optimise the assay.

In the first method, bacteria were grown in LB containing different concentrations of H₂O₂ (0, 2.5, 5 and 10 mM), and the growth was monitored hourly by measuring the OD₆₀₀. As seen in Figure 3.32A there was no significant difference between growth rates for WT and Δ *prtI*. This could be attributed to other antioxidant mechanisms present in cells that help bacteria survive under oxidative stress that might have been constantly released as the bacteria continuous to grow. Subsequently, the rate at which hydrogen peroxide could kill bacteria within a limited time was investigated. This 'killing curve' was carried out according to two methods as described in sections 2.8.1 and 2.8.2. The '20 minute' killing curve was implemented using 10 mM H₂O₂, during which samples of the cultures were serially diluted and spotted on LB agar (section 2.8.1) (data not shown). Killing curves as described in section 2.8.2 were carried out using 5 mM, 7.5 mM and 10 mM H₂O₂ to identify appropriate conditions to generate a killing curve for the WT (data not shown) where 7.5 mM was found to be the optimum concentration. However, deletion of *prtI* did not have significant effect on bacterial survival in liquid culture under oxidative stress as compared to WT (Figure 3.32B).

Further, the techniques that tested the effect of oxidative stress on bacteria growing on an agar based medium were investigated. Two different methods were employed – the first one included the oxidant within the agar, onto which serially diluted bacteria were spotted, as described in section 2.8.3. The optimal concentration was first empirically determined, however, this technique produced variable results giving very high error bars and was decided to not be used (data not shown). In the second approach, bacteria in soft agar were overlaid on standard agar plates onto which a filter

disc, containing the oxidant was placed. During overnight incubation the oxidant diffuses and creates a zone of growth inhibition (section 2.8.4). In addition to hydrogen peroxide, effects of the oxidants sodium hypochlorite, methyl viologen and tert-butyl hydroperoxide were also tested. In each case, the optimal concentration was first empirically determined (data not shown). This technique was found to be reliable as it produced similar results each time and was fairly easy to perform. Using the above technique, it was found that the $\Delta prtI$ and $\Delta prtR$ were similarly sensitive to the oxidants as the WT (Figure 3.33 and 3.34). These results suggest that the putative σ factor-anti σ factor pair, PrtI-PrtR, may not be involved in an oxidative stress response in *B. cenocepacia*.

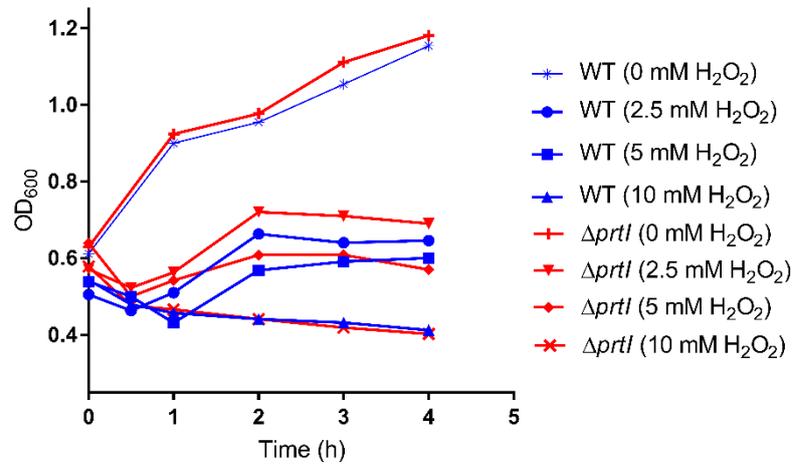
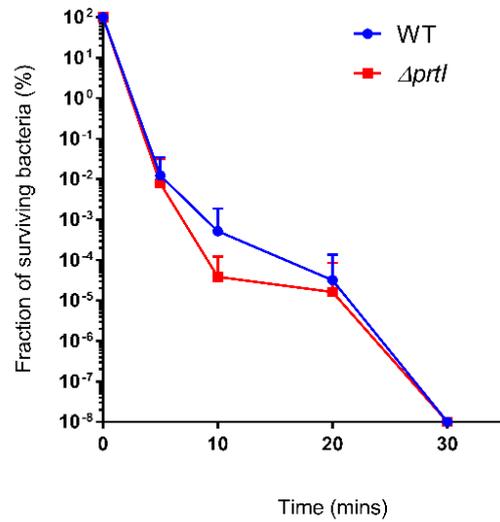
A**B**

Figure 3.32. Effect of peroxide stress on the *B. cenocepacia* $\Delta prtI$ mutant

(A) Representative growth curves of WT (H111) and H111- $\Delta prtI$ in LB containing 0, 2.5, 5 or 10 mM H_2O_2 grown at 37°C aerobically. (B) Killing curve of WT (H111) and H111- $\Delta prtI$ in LB containing 7.5 mM H_2O_2 at 37°C.

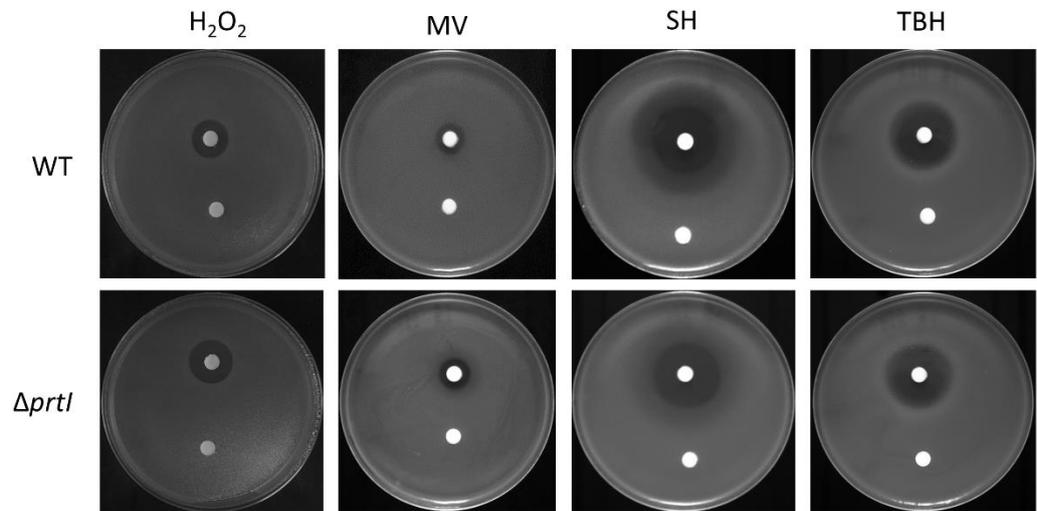
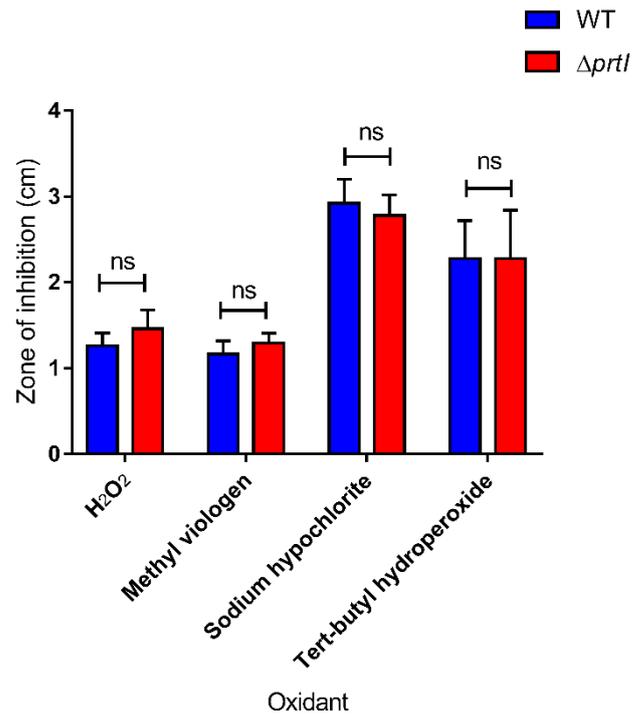
A**B**

Figure 3.33. Effect of oxidative stress due to deletion of *prtI*

(A) Representative images of plates showing a zone of inhibition assay to measure sensitivity to different oxidants. In each case top filter disc contains oxidant and bottom filter disc contains water. Top horizontal panels represent WT (H111), bottom panels represent H111- $\Delta prtI$. Vertical panels from left to right represent oxidants H₂O₂, methyl viologen, sodium hypochlorite and tert-butyl hydroperoxide. (B) Quantification of the zones of inhibition for WT (H111) and H111- $\Delta prtI$. A student's t-test was carried out to determine statistical significance, ns = not significant, (n = 9).

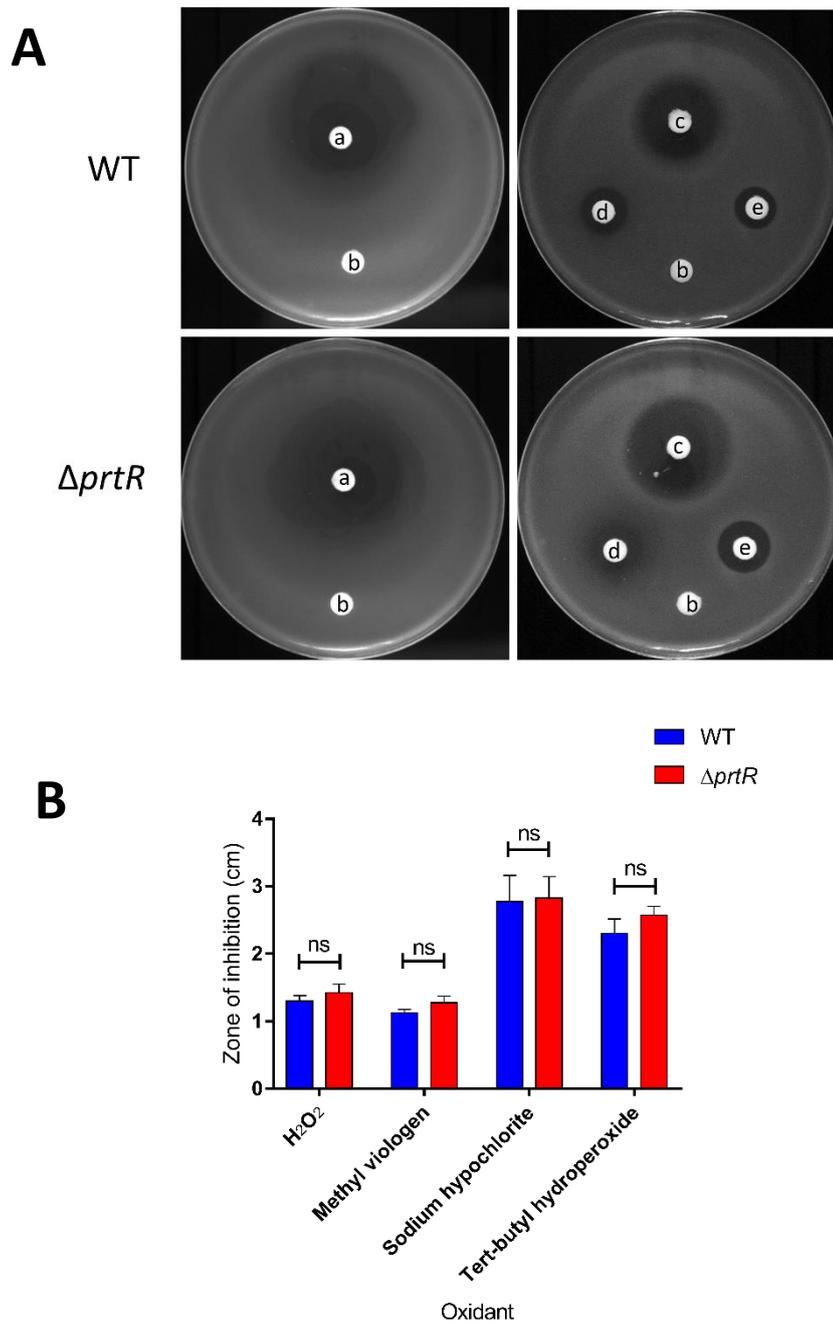


Figure 3.34. Effect of oxidative stress due to deletion of *prtR*

(A) Representative images of plates showing zone of inhibition assay to measure sensitivity to different oxidants. Top horizontal panels represent WT (H111), bottom panels represent H111- $\Delta prtR$ – filter disc contains (a) sodium hypochlorite (b) H₂O (c) tert-butyl hydroperoxide (d) methyl viologen (e) H₂O₂. (B) Quantification of the zones of inhibition for WT (H111) and H111- $\Delta prtR$. A student's t-test was carried out to determine statistical significance, ns = not significant, (n = 5).

3.5.6 Effect of deletion of *prtI* or *prtR* on protease activity

The *P. fluorescens* PrtI orthologue plays a role in extracellular protease production as described before. To assess the role of the PrtI system on protease production in *B. cenocepacia* 5 μ l of bacterial culture ($OD_{600} = 1$) was spotted on d-BHI agar (2.1.2) and incubated at one of the following temperatures - 25, 30, 37 or 42°C overnight to allow formation of clear zones around the growth due to protease activity. No protease activity was seen at 25°C for any strains. H111- Δ *prtI* consistently formed a clear zone of similar size to that of the WT (H111) at 30, 37 and 42°C. However, H111- Δ *prtR* gave slightly smaller zones, although this difference was subtle. This may be due to H111- Δ *prtR* being unstable and forming different sized colonies as seen before. It was obvious that the density of growth of the Δ *prtR* mutant was less than that of the WT and H111- Δ *prtI* which was perhaps the reason for the slightly smaller zones of casein hydrolysis. To determine whether this difference was complementable, pBBR2 and pBBR2-Prtr were each transferred to H111- Δ *prtR* by conjugation using S17-1 as donor cells. However, it was observed that even in the presence of *prtR*, the Δ *prtR* mutant still continued behaving as it did in the absence of *prtR*. Importantly, protease activity was not abolished for either of the mutants at any temperature in contrast to the studies with *Pseudomonas*, suggesting that PrtI and PrtR do not have roles in regulating protease activity in *B. cenocepacia*.

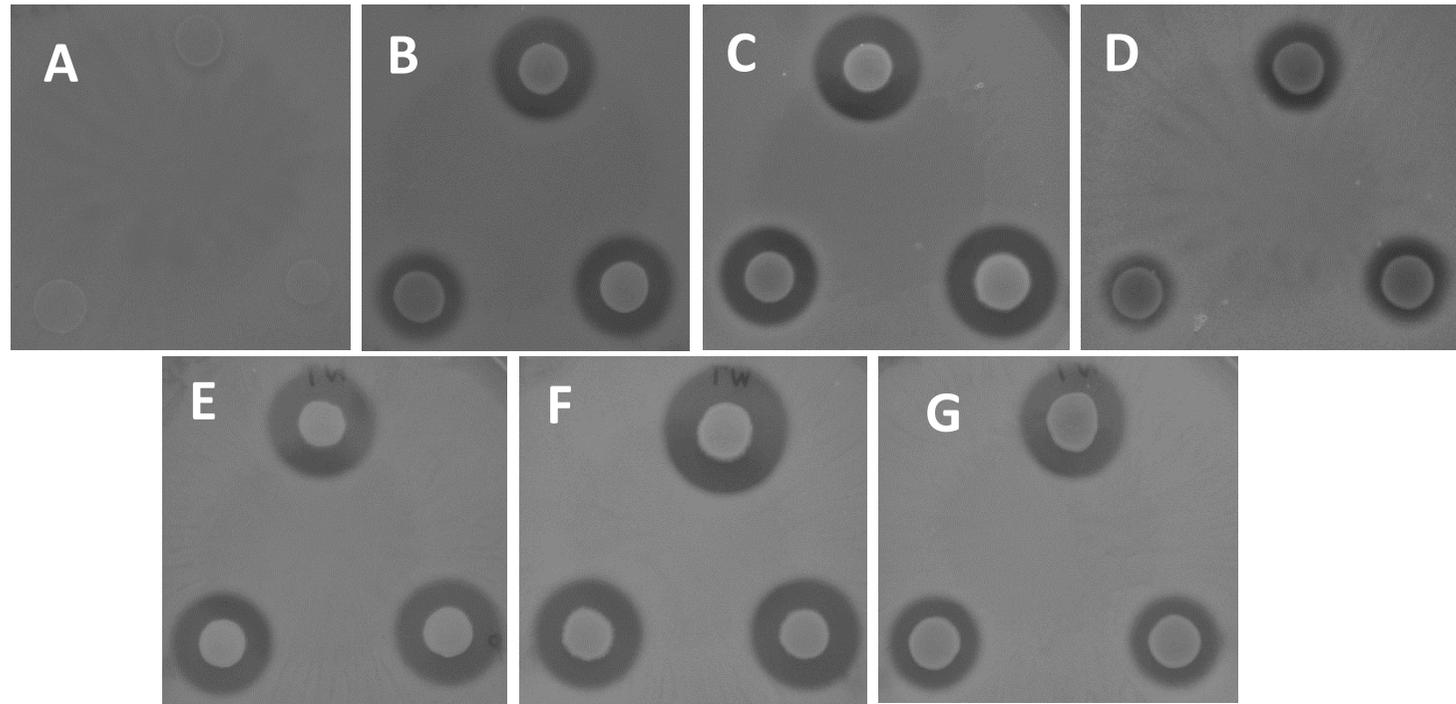


Figure 3.35. Effect of deletion of *prtI* and *prtR* on extracellular protease production by *B. cenocepacia*

5 μ l of bacterial culture ($OD_{600} = 1$) was spotted on d-BHI agar and incubated overnight such that a clear zone of inhibition formed around the growth. Top row with panels A, B, C and D has H111 (WT) on top, H111- Δ *prtI* on bottom right and H111- Δ *prtR* on bottom left. (A) 25°C (B) 30°C (C) 37°C (D) 42°C. Bottom row with panels E, F and G has H111 (WT) on top, H111- Δ *prtR*/pBBR2 on bottom right and H111- Δ *prtR*/pBBR2-PrtR on bottom left. (E) 30°C (F) 37°C (G) 42°C.

3.5.7 Effect of deletion of *prtI* or *prtR* on bacterial motility and swarming

The effect of deletion of *prtI* or *prtR* on motility and swarming were determined as described in section 2.7.3. Deletion of *prtR* seemed to result in small reduction in surface motility and swarming, however this was not statistically significant (Figure 3.36 A-C). This may be related to the same property of H111 Δ *prtR* that gives rise to variable colony sizes. It was overall concluded that both PrtI and PrtR may not have a role in motility.

3.5.8 Effect of deletion of *prtI* or *prtR* on biofilm formation

It was investigated whether deletion of *prtI* or *prtR* affected the ability of bacteria to form biofilms using the technique described in section 2.7.4. WT (H111), H111- Δ *prtI* and H111- Δ *prtR* formed biofilms with similar efficiency as shown in Figure 3.36D. Therefore, it was concluded that deletion of *prtI* and *prtR* may not play a role in biofilm development.

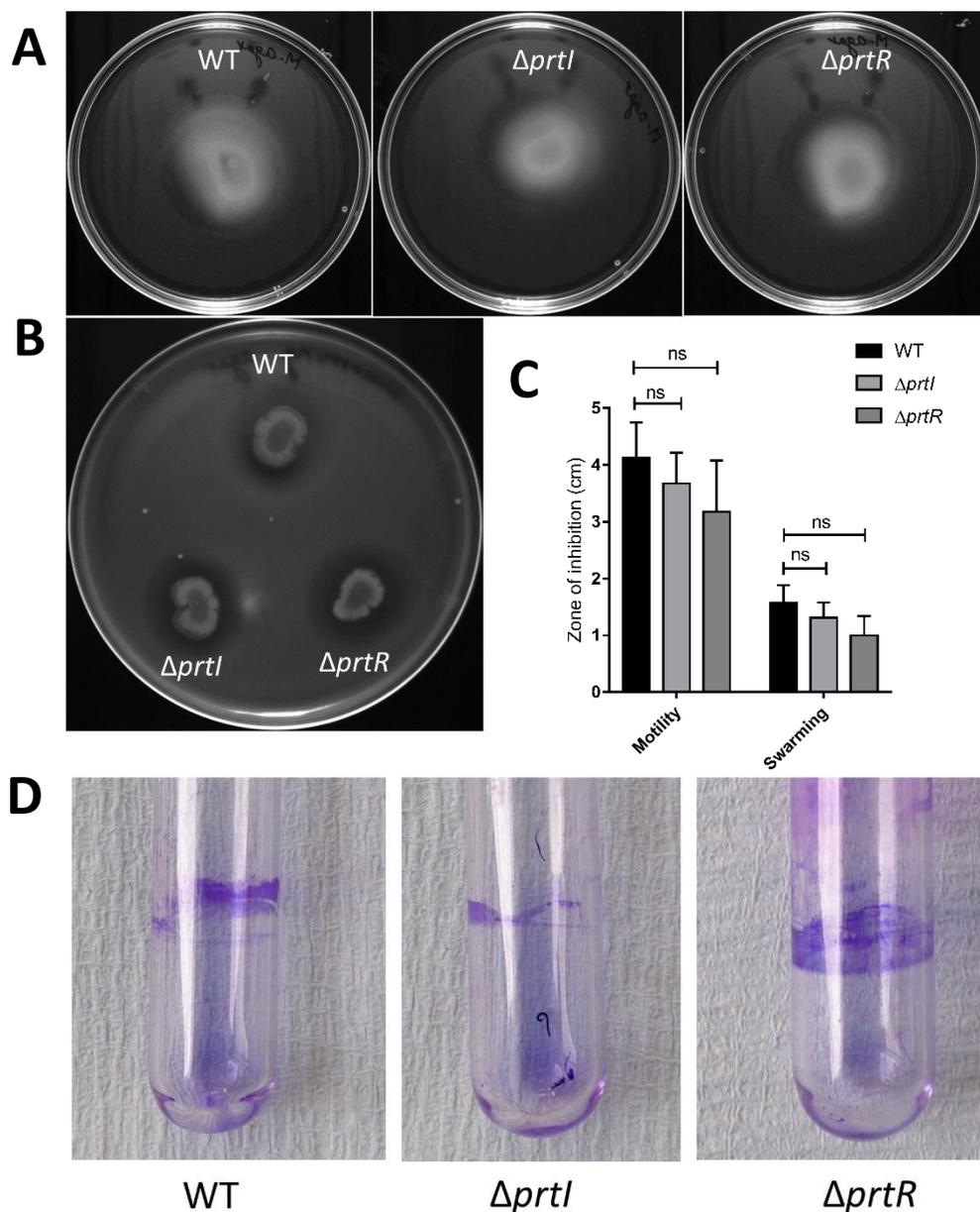


Figure 3.36. Effect of deletion of *prtI* or *prtR* on motility, swarming and biofilm formation: (A) Representative images of motility agar inoculated with WT, H111- $\Delta prtI$ or H111- $\Delta prtR$. (B) Representative image of swarming agar inoculated with WT, H111- $\Delta prtI$ or H111- $\Delta prtR$. (C) Quantification of zones of motility/swarming of the WT, H111- $\Delta prtI$ or H111- $\Delta prtR$. A one-way ANOVA with a Tukey's post test was carried out to determine statistical significance, ns = not significant, (n=6). (D) Representative images of WT, H111- $\Delta prtI$ or H111- $\Delta prtR$ biofilms stained with crystal violet.

3.5.9 Effect of deletion of *prtI* or *prtR* on EPS production

In order to study exopolysaccharide production, a medium which allowed mucoid colonies to form was required. A previously developed 'EPS medium' (K.L. Farmer and M.S. Thomas, unpublished) which was based on Y-glutamate-glycerol M9 minimal salts medium containing glutamic acid as nitrogen source, glycerol, and at a low pH was employed. EPS medium was varied to nature and concentration of components to optimise the production of *B. cenocepacia*. These included nitrogen source, pH and growth temperature as shown in Table 3.1. The effect of the amount of glycerol and phosphate concentration were not tested. Two different strains were tested – H111 and 715j. Increasing pH, increasing concentration of glutamic acid or changing the nitrogen source caused a decrease in production of EPS for both strains. It was found that medium containing 1.69% L-glutamic acid, at pH 6.0 or 6.5 and grown at 30°C was found to be the most effective for development of mucoid colonies for H111 strain whereas for 715j, most mucoidy was obtained when the medium contained 1.69% L-glutamic acid, at pH 6.0 and bacteria were grown at 37°C.

H111, H111- Δ *prtI* and H111- Δ *prtR* were streaked on EPS medium in biological and experimental triplicates. While H111 and H111- Δ *prtI* formed the same type of mucoid colonies, H111- Δ *prtR* seemed to give either mucoid colonies, like the WT, or a mixture of both mucoid and smaller, less mucoid colonies. Sometimes, only smaller less mucoid colonies were formed (Figure 3.37). The instability of the H111- Δ *prtR* phenotype may be the result of the constitutive activity of PrtI which may make the mutant 'sick' as described earlier (section 3.3.4). To assess whether the altered mucoid phenotype was due to the loss of *prtR*, H111- Δ *prtR* containing pBBR2 expressing *prtR* was also streaked on EPS medium.

H111- Δ *prtR*/pBBR2 and H111- Δ *prtR*/pBBR2-PrtR gave rise to mucoid colonies similar to the WT. Since H111- Δ *prtR*/pBBR2 should give the same phenotype as H111- Δ *prtR* it was concluded that the mixed colony phenotype might be caused by secondary site mutations, and that PrtR does not have a direct role in regulating EPS production.

Table 3.1. Optimisation of medium for EPS production by *B. cenocepacia*

Nitrogen source	% nitrogen source	pH						Temperature °C	715j mucoidy	H111 mucoidy
		5.0	5.5	6.0	6.8	7.5	8.5			
L-glutamic acid	0.169	5.0						37	+++	+++
L-glutamic acid	0.169		5.5					37	+++	+++
L-glutamic acid	0.169			6.0				37	++++	++++
L-glutamic acid	0.169				6.8			37	+++	+++
L-glutamic acid	0.169					7.5		37	--	--
L-glutamic acid	0.169						8.5	37	--	--
L-glutamic acid	0.5			6.0				37	--	--
Tryptophan	0.169			6.0				37	--	--
alanine	0.169			6.0				37	++	++
proline	0.169			6.0				37	+	+
lysine	0.169			6.0				37	+	+
ammonium chloride	0.169			6.0				37	--	--
Aspartic acid	0.169			6.0				37	++	++
histidine	0.169			6.0				37	+++	++
L-glutamic acid	0.169				6.5			37	++	++
L-glutamic acid	0.169			6.0				30	+++	++++
L-glutamic acid	0.169				6.5			30	+++	++++
L-glutamic acid	0.169			6.0				25	--	+
L-glutamic acid	0.169				6.5			25	--	+

To make Exopolysacchiride (EPS) producing media the following components to the final concentrations of 0.169 % glutamic acid (sodium salt), 0.3 % Tris base, 0.1 ml MgSO₄.7H₂O (10 % w/v), 0.1 ml CaCl₂.6H₂O (22 % w/v), 0.1 ml K₂HPO₄.3H₂O (22 % w/v) were added to ddH₂O. After adjusting the pH, the final volume was adjusted to 99 ml. A final concentration of 1.5 % agar (Oxoid) was added prior to autoclaving. 50% sterile glycerol solution was added to a final concentration of 1% to the medium and poured into sterile Petri dishes. For optimisation, variations in one or more components were made in the recipe. Each row represents the modification made at a time and the mucoidy obtained for the strains. Key: -- = non mucoid; +, ++, +++, ++++ = increasing amounts of mucoidy.

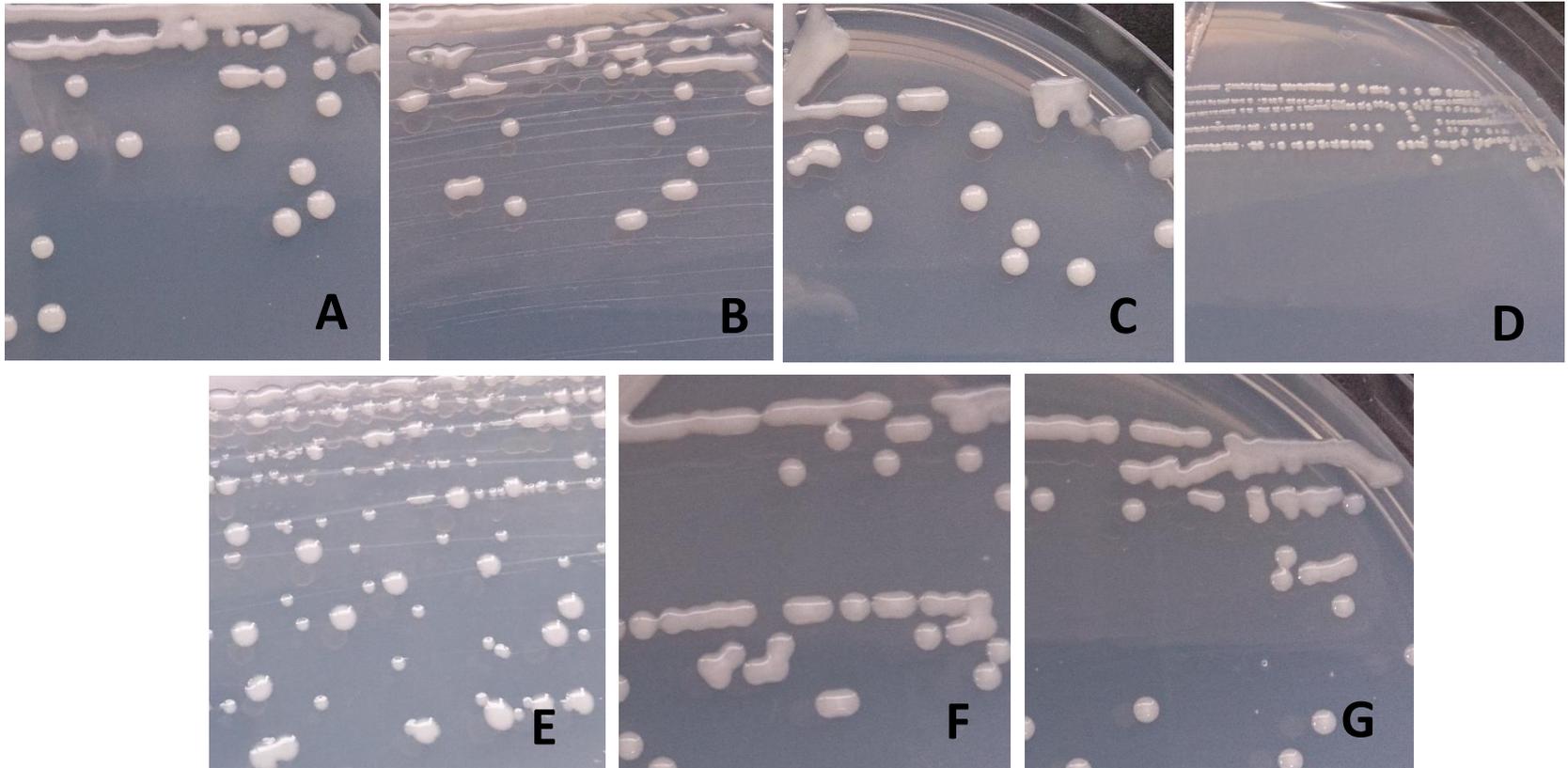


Figure 3.37. Effect of inactivation of *prtI* and *prtR* on EPS production by *B. cenocepacia* on EPS medium (pH 6.0) at 30°C
 (A) H111 (WT) (B) H111- $\Delta prtI$ (C) H111- $\Delta prtR$ forming uniform mucoid colonies. (D) H111- $\Delta prtR$ forming small non-mucoid colonies. (E) H111- $\Delta prtR$ forming mixed sized colonies of variable mucoidy. (F) H111- $\Delta prtR/pBBR2$ (G) H111- $\Delta prtR/pBBR2-PrtR$ forming uniform mucoid colonies.

3.6 Investigation of the role of PrtI by studying the role of its target gene – I35_RS16285

In a complementary approach to identify the role of PrtI, the role of its target gene I35_RS16285, was investigated. This gene is predicted to encode a catalase based on the amino acid sequence of its encoded product. However, deletion mutant studies described in section 3.5 did not provide evidence that PrtI-PrtR have a role in an oxidative stress response. Therefore, it was decided to investigate the role of I35_RS16285 and confirm if it functions as a catalase or not.

3.6.1 Testing the enzyme activity of purified I35_RS16285

To investigate the enzyme activity of I35_RS16285, it was decided to overexpress the protein using the *E. coli* host strain BL21 (DE3) and the pETDuet expression plasmid which can be used to overproduce target proteins with an N-terminal hexa-histidine affinity tag. Thus the expected host strain would be used to produce His₆-I35_RS16285, a 34.88 kDa protein which could be tested for its catalase activity by carrying out an enzyme assay.

3.6.1.1 Construction of plasmid harbouring I35_RS16285

I35_RS16285 was amplified without its predicted sec-dependent signal sequence using primers Catduetfor and Catduetrev with boiled H111 lysate as template and Q5 polymerase. The PCR product of 963 bp digested with *Bgl*III and *Hind*III and ligated to pETduet-1 digested with *Hind*III and *Bam*HI. As *Bgl*III and *Bam*HI generate compatible sticky ends this results in cloning of I35_RS16285 into the MCS located downstream of the first (i.e. upstream) T7 promoter. The ligations were used to transform MC1061 cells, which were spread on LB containing ampicillin. Colonies were PCR screened using I35_RS16285-specific primers and the transformants thought to contain the desired plasmid were checked for their size (6.383 Kb) and sent for sequencing to confirm the integrity of the cloned DNA (pETduet-1- I35_RS16285).

3.6.1.2 Purification of I35_RS16285

Clones that contained the expected *I35_RS16285* insert were transferred to BL21 (λ DE3) cells. To test whether *I35_RS16285* was produced when the T7 promoter on pETDuet- *I35_RS16285* was induced by IPTG, 8 ml cultures were induced with IPTG up to late-log phase and a 1ml aliquot was centrifuged at 13,000 rpm for 3 minutes. Cells were re-suspended in 2X Laemmli buffer and boiled for 10 minutes. After a clearing centrifugation step, 20 μ l of sample was loaded on to an SDS polyacrylamide gel and electrophoresis was carried out. As shown in Figure 3.38A, a protein consistent with the size of the expected protein (approx. 35 kDa) was observed in the induced sample, while it was not present in the non-induced sample, thus strongly suggesting that His₆-*I35_RS16285* was being expressed when induced. Next, to test whether His₆-*I35_RS16285* was soluble, 25 ml cultures were IPTG induced up to late-log phase and cells were harvested at 12,500 rpm for 20 minutes at 4°C. Cells were washed and lysed, and samples were prepared and electrophoresed in a SDS polyacrylamide gel as described in section 2.3.2. As shown in Figure 3.38B, a protein corresponding to approximately 35 kDa size was visible in the whole cell lysate, crude lysate and soluble fraction (supernatant) thus showing that His₆-*I35_RS16285* was largely soluble. Next, His₆-*I35_RS16285* was purified by IMAC as described in section 2.3.3. Here, a 200 ml culture was grown to log phase at 30°C and induced for 3 hours and harvested by centrifugation at 12,500 rpm for 20 minutes at 4°C. Cells were lysed with lysozyme and the supernatant was loaded onto a HisTrap nickel Sepharose column attached to an ÄKTA purifier 10 FPLC purification system. The column was washed in buffer containing a low concentration of imidazole and protein was eluted using a buffer containing a high concentration of imidazole buffer. As shown in Figure 3.38C & E, eluate tube A8 was found to contain most concentrated amount of His₆-*I35_RS16285*. Most of the protein sample thus obtained constituted of His₆-*I35_RS16285* with very mild contamination by other bacterial proteins, based on the visual appearance of the SDS PA gel, although the exact percentage of purified protein in the sample was not measured. To confirm that the purified protein was indeed His₆-*I35_RS16285* a Western blot was carried out using anti-hexa-histidine primary antibody raised in mouse and a secondary anti-mouse

antibody raised in rabbit. As shown in Figure 3.38D purified samples of the protein were tested and confirmed to contain His₆. I35_RS16285 based on its cross-reactivity with the antibody. Using the Bradfords assay (section 2.4.3), the purified protein sample was quantified and was found to contain approximately 0.6 mg/ml total protein.

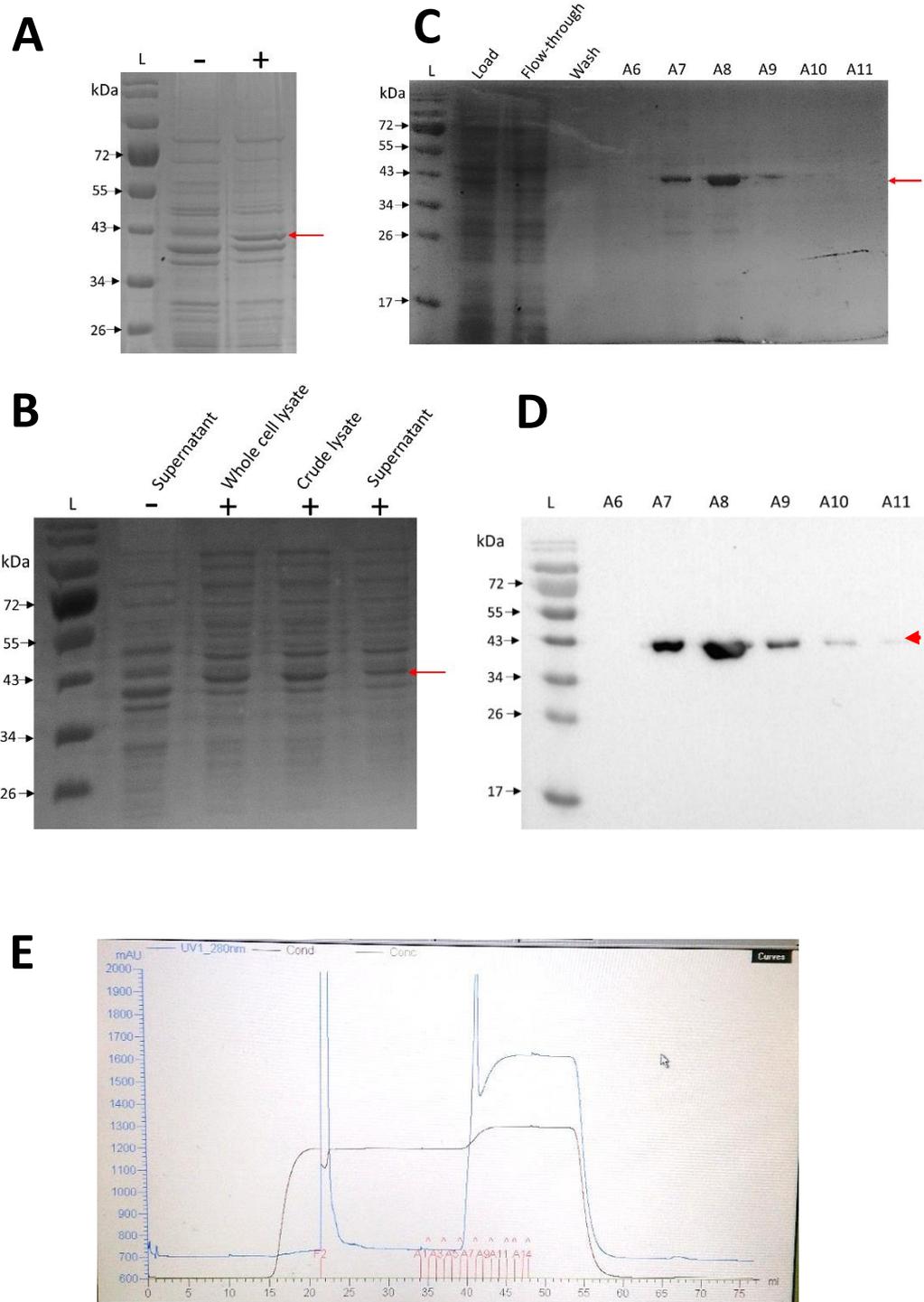


Figure 3.38. Purification of His₆. I35_RS16285

(A) Coomassie blue stained SDS PA gel showing total protein content from uninduced and induced BL21(DE3)/pETDuet-His₆. I35_RS16285 cultures, L = ladder, - = uninduced, + = induced, red arrow indicates location of His₆.I35_RS16285. (B) Coomassie blue stained SDS PA gel showing

samples used to test solubility of His₆.I35_RS16285, red arrow = His₆.I35_RS16285. (C) Coomassie blue stained SDS PA gel showing purification of His₆.I35_RS16285 by IMAC, A6-A11= elution samples, red arrow indicates location of His₆. I35_RS16285 (D) Western blot of same eluate samples from image C using anti-hexa-histidine antibody red arrow indicates His₆. I35_RS16285 (E) A₂₈₀ trace to identify His₆. I35_RS16285 protein content in IMAC eluates. X-axis indicates volume of liquid that is passed through the column (ml); Y-axis indicates absorbance (A₂₈₀). Initial blue peak at 22 minutes corresponds to the wash step to remove unbound protein. Second blue peak at 41 ml corresponds to His₆.I35_RS16285 elution and plateau formed from 46 ml onwards is due to high concentration imidazole buffer.

3.6.2.3 Catalase enzyme assay

Immediately after protein purification, to measure the catalase activity of His₆-I35_RS16285, the procedure described in section 2.5.2 was used where the rate of degradation of H₂O₂ was measured. As a positive control for the experiment, commercially available catalase (haem dependent) with a pre-determined catalase enzyme activity was used. Commercial catalase (from bovine liver) at approximately 100 units/ml was used in the assay which corresponded to 20 µg/ml of protein. Due to batch to batch variation, Sigma was contacted to find the enzymatic activity for this batch which turned out to be 69.2 units/ml for a concentration of 20 µg/ml of protein. As a control for any background catalase/anti-oxidant enzyme activity from proteins expressed by BL21 (DE3), a different *B. cenocepacia* protein, 'BCAS0667.CTD', was used, expressed from a previously constructed plasmid pHLS16 (H. Spiewak & M.S.Thomas, unpublished). It was purified in parallel exactly the same way as I35_RS16285 using the same expression plasmid, same host strain and with a N-terminal hexa-histidine tag. BCAS0667.CTD is the C-terminal domain of BCAS0667, a putative effector of the type VI secretion system and was expected to be devoid of catalase activity. The catalase activity was measured at room temperature (25°C) and 37°C using a thermostatted spectrophotometer. As shown in Figure 3.39B, no catalase enzyme activity was measured even when up to 10 times more I35_RS16285 protein (200 µg/ml) was used as compared to the positive control either at 25°C or at 37°C. These results suggest that I35_RS16285 may not function as a catalase.

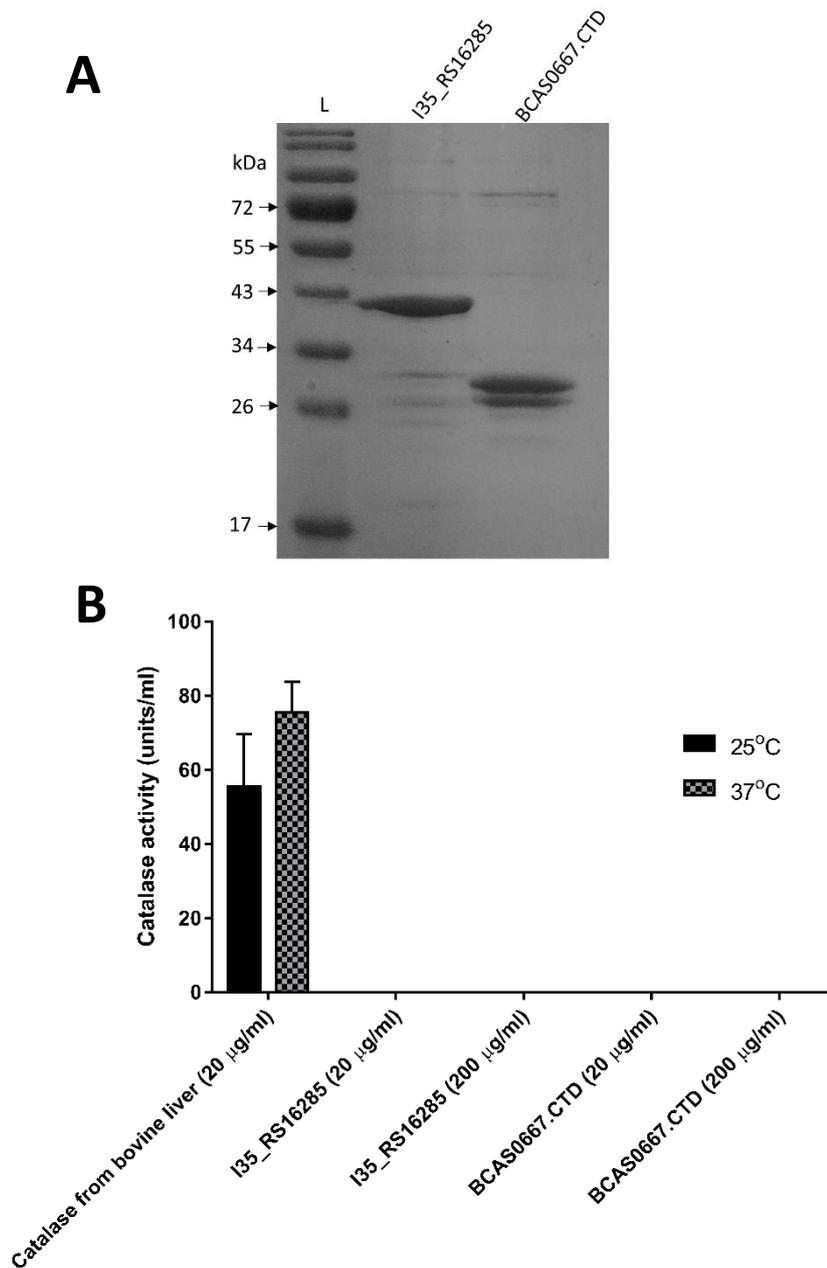


Figure 3.39. Testing the catalase activity of I35_RS16285

(A) Coomassie blue stained SDS PA gel showing purified I35_RS16285 and BCAS0667.CTD indicated by red and blue arrows, respectively. (B) Catalase enzyme activity (ability to degrade H_2O_2) in units/ml was calculated based on the time required for A_{240} to reduce from 0.450 to 0.400 in reaction buffer containing H_2O_2 . n = 9.

3.6.2 Identifying the function of I35_RS16285 by testing for complementation of an *E. coli* catalase deficient mutant

A complementary approach to test catalase activity was used, where I35_RS16285 was cloned into a plasmid that could constitutively express it at physiological levels in an *E. coli* catalase deficient mutant. As I35_RS16285 is predicted to contain a signal sequence, which suggests it might be exported into the periplasm where it might be carry out its target function, this sequence was retained.

3.6.2.1 Construction of plasmid harbouring I35_RS16285

Using primers Cat-compFor and Cat-compRevflag, I35_RS16285 was amplified from a boiled H111 lysate and using Q5 polymerase such that it incorporated the FLAG epitope tag coding sequence before the stop codon. The PCR product of 1153 bp was digested with *Hind*III and *Xba*I was ligated to the constitutive expression vector pBBR1MCS which was digested with the same enzymes. The ligation was then used to transform MC1061 cells which were spread on LB plates containing chloramphenicol. The resulting colonies were PCR screened using I35_RS16285 specific primers and the sequence integrity of the insert in putative recombinants was confirmed by nucleotide sequencing.

3.6.2.2 Complementation of a catalase deficient mutant

In order to test the hypothesis that I35_RS16285 could complement a strain deficient in catalase, the *E. coli* strain LC106, a mutated version of MG1655 which lacked the scavenging enzyme like catalases and NADH dependent peroxidases (*katG*, *katE*, *ahp*) was transformed with pBBR- I35_RS16285 and pBBR2 as a control. Subsequently, a zone of inhibition assay was carried out as described in section 2.8.4, using H₂O₂ as the oxidant. As observed in Figure 3.40A, pBBR- I35_RS16285 did not increase the resistance of LC106 to peroxide stress. Additionally, Western blotting was carried out to test if I35_RS16285 was being expressed using an anti-FLAG primary antibody raised in rabbit and a secondary anti-rabbit antibody raised in goat. As shown in Figure 3.40C, a cross-

reacting protein band migrating with a size consistent with the expected molecular weight of FLAG-tagged I35_RS16285 (40.7 KDa) was detected in total protein sample extracted from LC106/pBBR- I35_RS16285, confirming that I35_RS16285_{FLAG} (FLAG tagged I35_RS16285) was expressed.

Therefore, based on results from the catalase enzyme assay and the complementation experiment it seemed that I35_RS16285, at least on its own, does not act as a catalase.

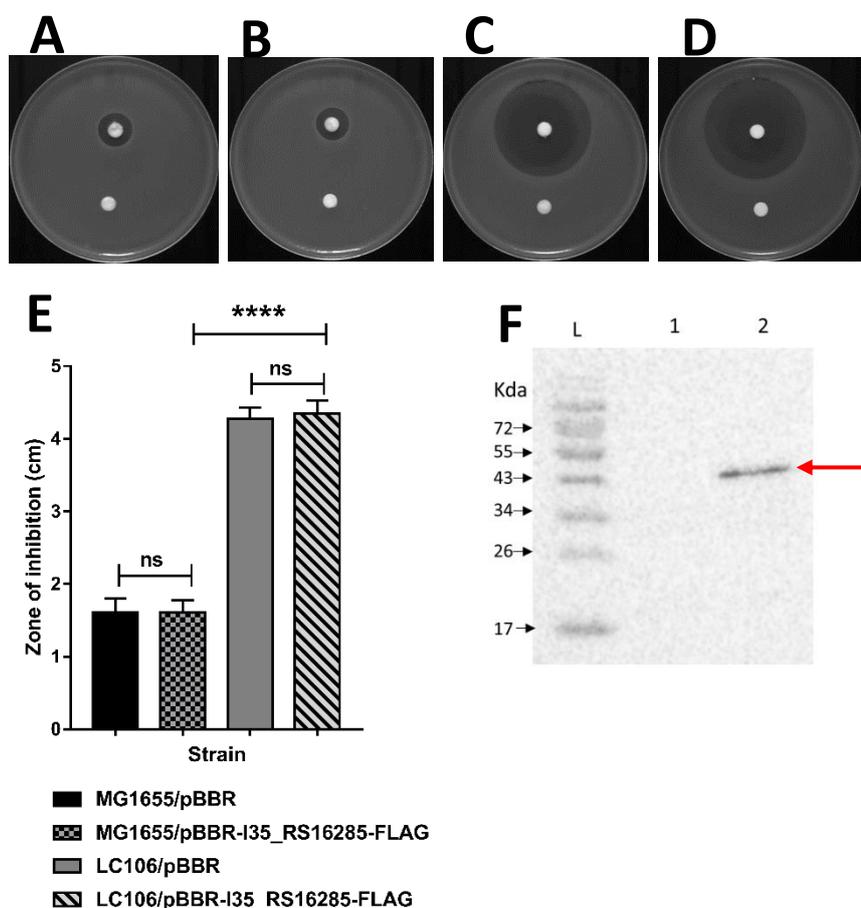


Figure 3.40 Complementation assay on an *E. coli* catalase and NADH-dependent peroxidase deficient mutant with *B. cenocepacia* I35_RS16285

Representative images of plates showing top filter containing oxidant (H₂O₂) with a zone of inhibition around and bottom filter disc containing water (A) MG1655/pBBR1MCS (B) MG1655/pBBR-I35_RS16285-FLAG, (C) LC106/pBBR1MCS, (D) LC106/pBBR-I35_RS16285-FLAG. (E) Quantified results showing diameter of zone of inhibition for pBBR1MCS Vs pBBR-I35_RS16285. A one way ANOVA with a Tukey's post test was carried out to determine statistical significance, ns= not significant, (n=5). (F) Western blot of total protein content obtained from (1) LC106/pBBR1MCS (2) LC106/pBBR-I35_RS16285-FLAG, using anti-FLAG antibody, L=ladder.

3.6.3 Bioinformatic analysis of I35_RS16285

In order to identify the putative role of I35_RS16285, since it did not seem to function like a catalase, I35_RS16285 orthologues were searched using BLASTP. I35_RS16285, a putative haem binding catalase is homologous to SrpA (36 kDa), with 40% similarity, a currently uncharacterised putative sulphur regulated protein from *Synechococcus* species (cyanobacteria). None of the current literature has yet defined its role although it is regulated by sulphate. CysR a protein involved in sulphur regulation and from the Crp family of regulators is reported to regulate *srpA*. Nevertheless, deletion of *srpA* did not affect growth under low or high sulphate concentrations, leading to the speculation whether or not it may be involved in sulphate uptake or metabolism (Nicholson and Laudenbach, 1995). SrpA is predicted to be a monofunctional haem type catalase, and seems to have conserved residues involved in haem binding (Nicholson and Laudenbach, 1995). Catalases act on hydrogen peroxide and decomposes it to give molecular oxygen and water. Catalases are 'haem-containing' and can be monofunctional or bifunctional (catalase-peroxidase). But, non-haem containing catalases (manganese containing) also exist in bacteria. Some proteins like SrpA may have similar folds and haem binding sites as catalase, but there is no evidence showing the presence of catalase enzyme activity. SrpA forms part of a cluster of 13 sulphur regulated proteins, where upstream of *srpA* is present an uncharacterised gene encoding a putative transcriptional regulator (Chen et al., 2008), However this is distinct from PrtI. Therefore, currently it is concluded that this protein may function as a periplasmic catalase, although my experiments suggest I35_RS16285 cannot function as a catalase on its own.

BCAM0931 is a monofunctional haem-type catalase, KatE from *B. cenocepacia* J2315 (Holden et al., 2009). An alignment of KatE with I35_RS16285 and SrpA was carried out. Figure 3.41 shows conserved and similar residues amongst the three proteins. Most of the amino acids forming the haem binding pocket are conserved across the three, especially His69, Phe146, Phe154, Arg145 and Tyr149 of I35_RS16285. However, additional conserved motifs are apparent in addition to the haem binding pocket in an

alignment between I35_RS16285 and SrpA, suggesting they might indeed be orthologues.

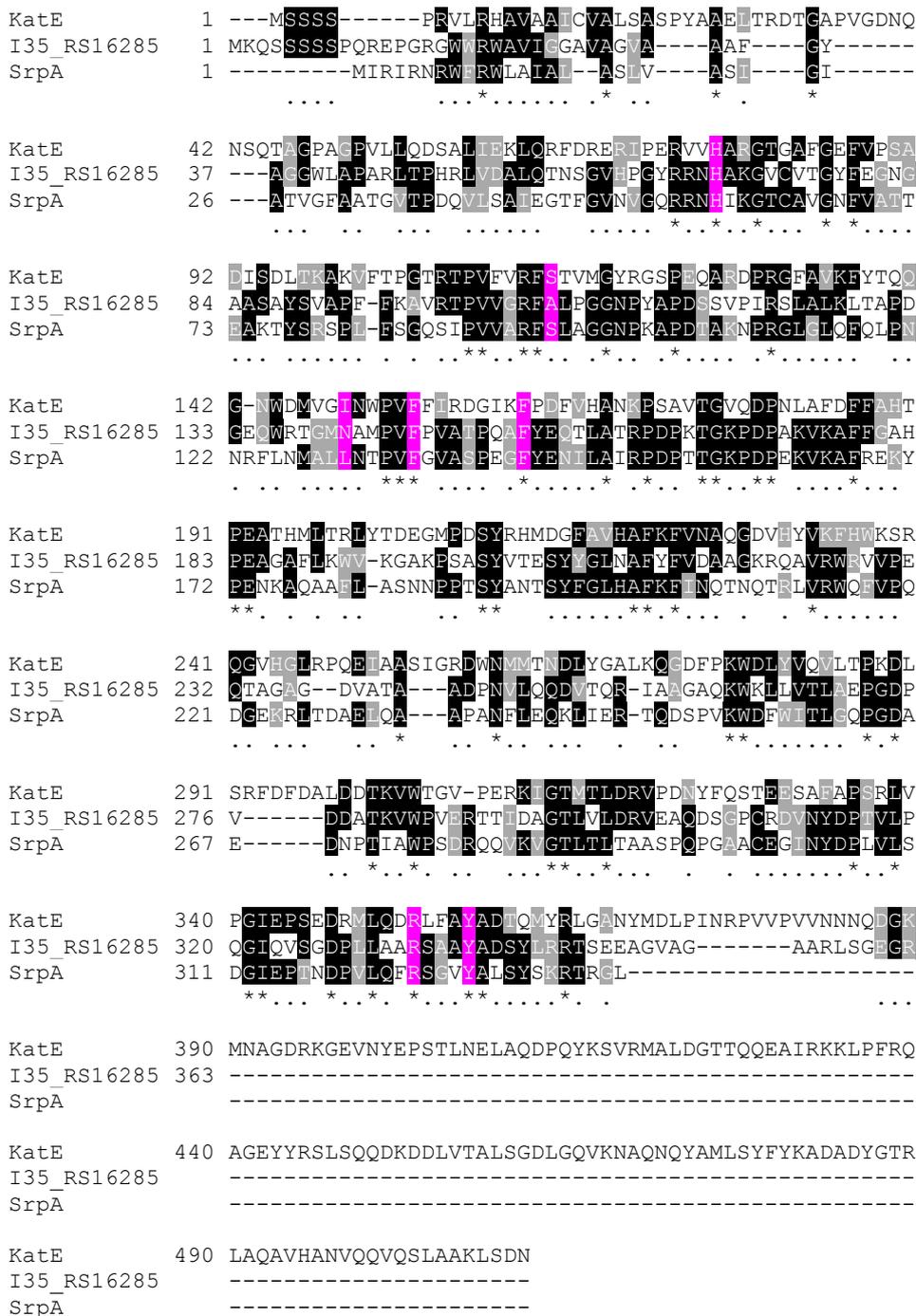


Figure 3.41. Alignment of KatE, I35_RS16285 and SrpA

KatE (a *B. cenocepacia* catalase), I35_RS16285 and SrpA sequences were aligned using clustal omega and shaded using BOXSHADE. Amino acids which are identical at the corresponding position in $\geq 50\%$ of sequences are shown in white font with black highlighting while amino acids which are similar at the corresponding position are shown in white font with grey highlighting. Residues forming the haem binding pocket are highlighted in magenta.

3.7 Regulation of PrtI by the putative anti- σ factor PrtR

Downstream of *prtI* lies I35_RS16295, predicted to encode a putative trans-membrane anti- σ factor. Since the encoded product is similar to PrtR of *Pseudomonas fluorescens* it is referred to as PrtR here. PrtI and PrtR have over-lapping stop and start codons which suggest that they might be translationally coupled. It was therefore hypothesized that, PrtR acts as a conventional anti- σ factor that keeps the σ factor PrtI bound and inactive under normal conditions, but releases it, in response to an appropriate signal to initiate transcription of specific genes. The first objective was to test the interaction between PrtI and PrtR *in vivo*. Further, the activity of PrtI-dependent promoters was studied in *E. coli* in and *B. cenocepacia* in the absence and presence of PrtR.

3.7.1 Construction of plasmids expressing PrtI and PrtR hybrid proteins for two-hybrid analysis

To study protein-protein interactions, the bacterial two-hybrid assay is a useful and robust technique. As described in section 2.6.1, the principle of this technique lies in obtaining a lactose or maltose phenotype produced by cells harbouring pairwise combinations of compatible plasmids that express proteins which may or may not interact. The proteins of interest are fused to either the T18 or T25 domains of *B. pertussis* CyaA (adenylate cyclase) by cloning the genes encoding the proteins into vectors pUT18, pUT18C, pKT25 or pKNT25. If the proteins of interest interact, it allows the two domains T18 and T25 to be brought in close proximity creating a functional adenylate cyclase. This causes production cyclic AMP from ATP. cAMP can combine with catabolite activator protein (CAP) which can switch on genes such as those of the maltose and lactose operons that have CAP-cAMP dependent promoters. If the medium on which these cells are grown contains maltose (i.e. maltose-MacConkeys agar), cells can now utilise it causing a change in the local pH of the medium and therefore the colour of the colony. A strong maltose phenotype gives dark maroon colonies whereas a weak or negative maltose phenotype gives white colonies.

Initially, it was decided to test whether the C-terminal domain of the σ factor (σ_4) interacts with the N-terminal domain of the anti- σ factor. Hence, it was decided to study the interaction between the C-terminal domain of PrtI (PrtI-CTD) with the N-terminal domain of PrtR (PrtR-NTD). Using primers PrtICTDfor and PrtICTDrev, a DNA fragment encoding PrtI-CTD (237 basepairs) was amplified from a boiled H111 lysate and cloned between the *Xba*I and *Acc*65I sites of plasmids pKT25 and pUT18C plasmids. Similarly using primers PrtICTDfor and PrtICTDrev2, a DNA fragment encoding PrtI-CTD was amplified from a boiled H111 lysate and cloned between the same sites of plasmids pKNT25 and pUT18. The colonies expected to contain the plasmid with the insert based on a PCR screen using insert-specific primers, were checked for their size on an agarose gel and their sequence was also confirmed. Similarly, a DNA fragment encoding PrtR-NTD (246 basepairs) was amplified from a boiled H111 lysate using Q5 polymerase with primers PrtRNTDfor and PrtRNTDrev, digested with *Xba*I and *Acc*65I and ligated with pKT25 and pUT18C which were digested with the same enzymes. Similarly, using primers PrtRNTDfor and PrtRNTDrev2, a DNA fragment encoding PrtR-NTD was amplified and cloned into pKNT25 and pUT18 between the *Xba*I and *Acc*65I sites. MC1061 cells were transformed with the ligations and the resulting colonies were screened by carrying out a colony PCR using insert specific primers. The sizes of the plasmids were checked and the integrity of the cloned DNA was also confirmed by nucleotide sequencing.

Unfortunately, due to time constraints it was not possible to construct plasmids to test the proposition that the N-terminal domain of the σ factor (σ_2) interacts with the N-terminal domain of the anti- σ factor.

3.7.2 Analysis of PrtI-PrtR interaction and self-interaction using the bacterial two-hybrid assay

Compatible pairwise combinations of plasmids as shown in Table 2.9 were used to transform BTH101 cells and the transformants were spread on maltose-MacConkey medium. Each such transformation was repeated two more times to give a final experimental replication of 3 ($n' = 3$). The phenotypes were recorded as shown in Table 3.3. The strongest phenotype was observed when T18-PrtI_{CTD} was expressed in the same

cells as PrtR_{NTD}-T25 (i.e. the plasmid combination pUT18C-PrtI_{CTD} and pKNT25-PrtR_{NTD}). A weaker maltose-positive phenotype was also observed for PrtI-PrtR interaction with the combinations (1) PrtI_{CTD} at the N-terminal domain of T18 and PrtR_{NTD} at the N-terminal domain of T25 (2) PrtI_{CTD} at the C-terminal domain of T18 and PrtR_{NTD} at the C-terminal domain of T25 (3) PrtI_{CTD} at the C-terminal domain of T25 and PrtR_{NTD} at the C-terminal domain of T18 and (4) PrtI_{CTD} at the N-terminal domain of T25 and PrtR_{NTD} at the C-terminal domain of T18. Self-interaction by PrtI_{CTD} was also observed, although there has been no evidence so far in the literature of ECF σ factor dimerization. Other combinations and controls gave the negative maltose phenotype. Overall, these results suggested that PrtI and PrtR interact through their C-terminal and N-terminal domains, respectively.

Plasmid 1				Plasmid 2				Result
CTD	N	T18	C	N	T25	C	CTD	+/-
CTD	N	T18	C	N	T25	C	NTD	-
CTD	N	T18	C	CTD	N	T25	C	-
CTD	N	T18	C	NTD	N	T25	C	+/-
NTD	N	T18	C	N	T25	C	CTD	-
NTD	N	T18	C	N	T25	C	NTD	-
NTD	N	T18	C	CTD	N	T25	C	--
NTD	N	T18	C	NTD	N	T25	C	-
N	T18	C	CTD	N	T25	C	CTD	+
N	T18	C	CTD	N	T25	C	NTD	+/-
N	T18	C	CTD	CTD	N	T25	C	+/-
N	T18	C	CTD	NTD	N	T25	C	++
N	T18	C	NTD	N	T25	C	CTD	+/-
N	T18	C	NTD	N	T25	C	NTD	--
N	T18	C	NTD	CTD	N	T25	C	+/-
N	T18	C	NTD	NTD	N	T25	C	-

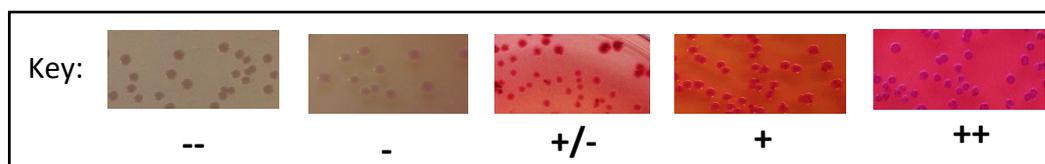


Figure 3.42. Two hybrid analysis of PrtI-PrtR interaction

The combinations of T18 and T25 domains (represented by blue and yellow boxes, respectively) used for each interaction are shown. The NTD of PrtR and CTD of PrtI are shown as orange and green boxes, respectively, fused to either N or C terminal end of the T18 and T25 fragments. The maltose phenotype of each interaction is presented on the right and the key for these results is given in a box below the plasmid combinations.

3.7.3 Effect of PrtR on the activity of PrtI in *E. coli*

To determine whether PrtR has an inhibitory effect on the activity of PrtI, as expected for an anti- σ factor an experiment was designed involving construction of a plasmid expressing *prtI* and *prtR* which would be used in combination with a plasmid containing a reporter fusion to the PrtI-dependent promoter, P_{cat}.

3.7.3.1 Construction of plasmid harbouring prtIR

A DNA fragment encoding 'PrtIR' (beginning from the start codon of PrtI till the stop codon of PrtR) was amplified using primers PrtIforB and PrtRrev with Q5 polymerase, using a boiled H111 lysate as template, giving a product of 1308 bp. pBBR2 and PCR product 'PrtIR' were digested using enzymes *HindIII* and *XbaI* and ligated together. The ligations were used to transform competent MC1061 cells, which were selected on LB agar containing kanamycin. The resultant colonies were screened using insert-specific primers by performing a colony PCR with GoTaq polymerase. Plasmid DNA was prepared for positive colonies and checked for their size which was expected to be 6.422 kb. The integrity of the DNA sequence was confirmed before proceeding further.

3.7.3.2 Analysis of PrtI activity in the presence of PrtR in *E. coli*

The reporter plasmid pKAGd4 harbouring the PrtI-dependent promoter 'P_{cat}' (constructed as described in section 3.2.13) was transferred to *E. coli* cells (MC1061) along with pBBR2-PrtI (constructed as described in section 3.2.15) or pBBR2-PrtIR or pBBR2. The activity of the promoter was measured by carrying out β -galactosidase assays. As shown in Figure 3.43, the presence of PrtR resulted in a severe decrease in the promoter activity rendering it nearly inactive, suggesting that PrtR exerts an inhibitory effect on the PrtI activity. As expected, no activity was obtained for the promoter in the presence of empty plasmid pBBR2.

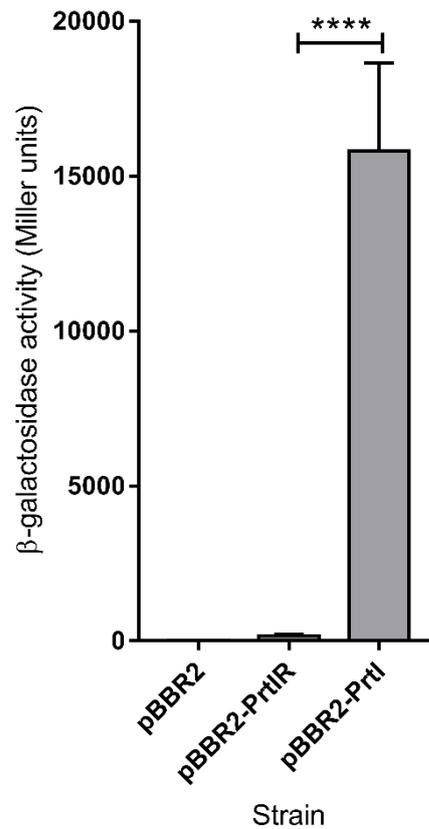


Figure 3.43. Activity of a PrtI-dependent promoter in the absence and presence of PrtR in *E. coli*

pBBR2, pBBR2-PrtI or pBBR2-PrtIR, as indicated, were transferred to MC1061 containing pKAGd4- P_{cat} and the promoter activity was measured using the β -galactosidase assay where cultures were grown in LB with kanamycin and chloramphenicol. Activities shown have been obtained by subtracting the activity of the same strain containing pKAGd4 carried out under same conditions (background control). Statistical significance was determined by performing a one-way ANOVA with a Tukey's post test, **** = $p < 0.0001$, (n=6).

3.7.4 PrtR on the activity of PrtI in *B. cenocepacia*

To further investigate the inhibitory effect of PrtR on the activity of PrtI the activity of P_{cat} was measured in the *prtR* deletion mutant, H111- $\Delta prtR$ described in section 3.5.2 and compared to the activity in the WT (i.e. *prtR*⁺) strain. pKAGd4 empty plasmid and pKAGd4- P_{cat} were conjugally transferred to H111 and H111- $\Delta prtR$ using S17-1 as a donor strain and the activity of the promoter was measured by carrying out β -galactosidase assays. As shown in Figure 3.44, P_{cat} was found to be highly active in H111- $\Delta prtR$ as compared to H111, perhaps due the absence of PrtR, allowing the σ factor PrtI to be freely available to recognise and activate the P_{cat} promoter. These results are consistent with the hypothesis that PrtR is the anti- σ factor for PrtI operating its regulation via an inhibitory mode.

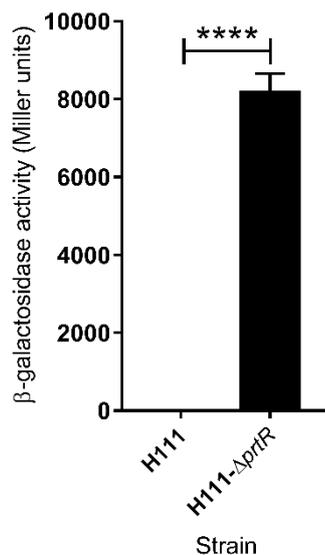


Figure 3.44. Activity of a PrtI-dependent promoter in the absence and presence of PrtR in *B.cenocepacia*

pKAGd4- P_{cat} was transferred to H111 or H111- $\Delta prtR$, as indicated, and the P_{cat} activity was measured using the β -galactosidase assay where cultures were grown in LB with chloramphenicol. Activities shown have been obtained by subtracting the activity of the same strain containing pKAGd4 assayed under the same conditions (background control). Statistical significance was determined by performing a students' t-test, **** = $p < 0.0001$ (n=6).

3.8 Defining the extent of the I35_RS16285 and *prtI* operons

To establish the limits of the two *PrtI*-dependent operons a reverse-transcriptase PCR experiment was designed. It was hypothesized that *prtI* and *prtR* are co-transcribed, driven by promoter P_{prtI} and that I35_RS16285 (putative catalase) and I35_RS16280 (putative cytochrome B561) are co-transcribed, driven by the promoter P_{cat} . To induce expression of *PrtI* for production of *PrtI*-dependent transcripts, pBBR2-*PrtI* was conjugally transferred to H111 (WT) cells by performing a conjugation using S17-1 as donor strain. As a control for this experiment, pBBR2 was also transferred to H111 via conjugation. Although four pairs of primers were designed to define the extent of the I35_RS16285 and *prtI* operons as shown in Figure 3.45A, only two pairs were used in this experiment (the Cat-cytRFor2 and Cat-cytRev primer pair and the *PrtI*RFor and *PrtI*RRev primer pair) due to lack of time. As described in section 2.2.10.2, mRNA from H111/pBBR2 and H111/pBBR2-*PrtI* was extracted using the Machery-Nagel RNA extraction kit. cDNA was prepared separately using reverse primers Cat-cytRev and *PrtI*RRev. Subsequently, standard PCR was carried out to amplify the 'reverse transcripts', using the Cat-cytRFor2 and Cat-cytRev primer pair (which should give rise to a 247 bp PCR product) and the *PrtI*RFor and *PrtI*RRev primer pair (which should give rise to a 273 bp PCR product). The efficiency of these pairs of primers were first tested by carrying out only a standard PCR using boiled H111 lysate as template and were found to work efficiently (data not shown). Following that, RT-PCR was carried out using cDNA as template. mRNA that was used to make the cDNA was also used as template in a separate PCR reaction to control for the presence of contaminating DNA.

As shown in Figure 3.45B, DNA contamination was observed for the control mRNA samples which is apparent as the yield of the PCR product (amplicon) was higher than that obtained for cDNA. A higher amount of PCR product (247 bp) was obtained using H111/pBBR2-*PrtI* cDNA as compared to H111/pBBR2 cDNA with Cat-cytRFor2 and Cat-cytRev primer pair. Similarly, a higher amount of PCR product at 273 bp was obtained using H111/pBBR2-*PrtI* cDNA as compared to H111/pBBR2 cDNA with *PrtI*RFor and *PrtI*RRev primer pair. This suggests that these two pairs of genes may be co-transcribed. However, due to the apparent DNA contamination it is not possible to

conclude whether these PCR products arose from the mRNA transcript-derived cDNA or amplification due to DNA contamination. A few attempts were made by altering and optimising the DNA degradation (treatment with DNase) step but it was not possible to complete this experiment satisfactorily due to lack of time.

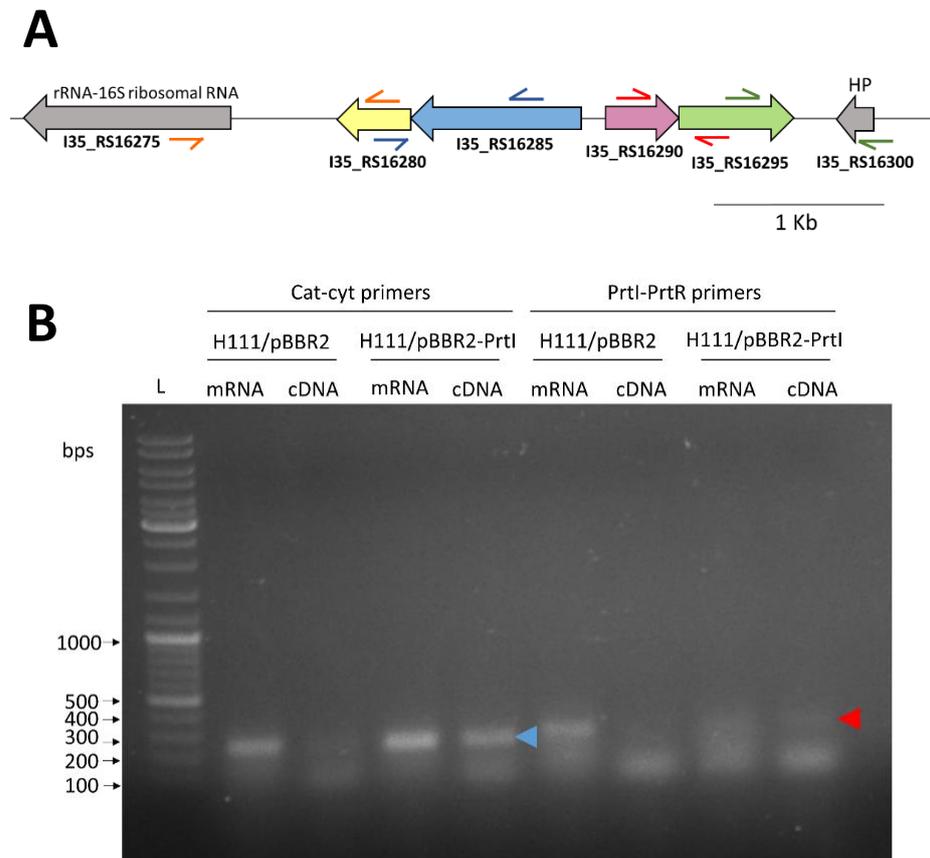


Figure 3.45. Identification of the *I35_RS16285* (*catalase*) and *prtI* operons by reverse transcriptase PCR

(A) Location of primers used in RT-PCR study (not to scale): Approximate genomic locations where the designed primers annealed. Genes encoding putative cytochrome B561, putative catalase, PrtI and PrtR are shown as yellow, blue, pink and green block arrows, respectively. Locations of primers designed to amplify the regions between 16S rRNA (*rrn*) gene and cytochrome B561 encoding gene are indicated by orange arrows, between genes encoding cytochrome B561 and catalase are shown by blue arrows (Cat-cytFor2 and Cat-cytRev), between *prtI* and *prtR* by red arrows (PrtIRFor and PrtIRRev) and between *prtR* and gene encoding hypothetical protein is shown in green arrows. (B) RT-PCR to identify PrtI regulon: mRNA was isolated from H111 containing pBBR1MCS2 or pBBR2-PrtI and cDNA was prepared using either reverse primer – Cat-cytRev or PrtIRRev. Using cDNA or mRNA as template PCR was carried out using the primer pair Cat-cytFor2 and Cat-cytRev (indicated as Cat-cyt primers) that gave a product of 247 bp, or using the primer pair PrtIRFor and PrtIRRev (indicated as PrtI-PrtR primers) that gave a product of 273 bp. Blue arrow indicates expected product obtained using the primer pair Cat-cytFor2 and Cat-cytRev. Red arrow indicates expected product obtained using the primer pair PrtIRFor and PrtIRRev.

3.9 Discussion

This chapter describes the attempt to characterise the putative ECF σ factor I35_RS16290 (PrtI) of *B. cenocepacia*. The role of PrtI in *B. cenocepacia*, its target promoters, target genes and mode of regulation have not yet been defined making this a novel study. These research questions are of importance as the current literature suggests that PrtI is involved in regulating genes which encode significant virulence determinants in the *Pseudomonas* species, although *Burkholderia* does not contain most of these putative target genes. Nevertheless, the molecular mechanisms of this regulation still remain uncertain. Therefore, studying the role of PrtI in the clinically relevant pathogen *B. cenocepacia* might reveal some interesting signal transduction mechanisms involved in the infection process.

3.9.1 PrtI recognises at least two promoters

A genome wide scan was carried out using ChIP-seq to identify *in vivo* the PrtI-dependent promoters on the *B. cenocepacia* genome. Proteins such as transcription factors are point source proteins unlike chromatin marks/histone modifications which are broad source. Therefore, the genome mapping step revealed visually detectable enrichment regions. Also, for proteins which are not generally expressed under standard laboratory conditions an inducible plasmid to express a tagged protein can be used. However, it is recommended by ENCODE to include its native promoter in this instance. As the aim was to identify true binding targets of PrtI, it was important not include the native promoter region in the plasmid construct. Hence, in such a case the same experiment was carried out with 'CONTROL' samples, consisting of cells harbouring the parent plasmid without any insert (Landt et al., 2012). Some studies also recommend using a MOCK sample (or IgG control) which consists of using a control antibody, that targets a non-specific antigen (Park, 2009). However, this is more often used for cases in which the antibody is generated for the particular protein of interest. However, as a commercial antibody (anti-FLAG) was used which meets standard practices and also has been previously used in other ChIP-seq experiments, I decided not to include a MOCK sample (Giacani et al., 2013, Markel et al., 2013, Blanka et al., 2014). Other experimental

designs include ‘mutant vs WT study’, mass spectrometry based validation, immunofluorescence validation, etc. more relevant for eukaryotic models. Validation on the samples before sequencing is often done by carrying out qPCRs. However, the results from standard PCRs showed obvious enrichment at the PrtI promoter. The depth of sequencing and complexity of library are directly related to site discovery (Bailey et al., 2013). A poorly constructed library with high sequencing depth can give more false positives while a well-constructed DNA library with low sequencing depth can give more false negatives. Therefore, in this study the quality of the samples was checked before library preparation, after library preparation and before data analysis to balance quality of the DNA libraries with sequencing depth. Lastly, at least two experimental replicates are commonly accepted to account for experimental variability. In this study I included three experimental replicates which also served as biological replicates to check for consistency in peak calling. To make biological sense from the results obtained in a ChIP-seq study it is important to validate the functionality of promoters by carrying out a different experimental design.

ChIP-seq revealed only one region that was enriched for PrtI binding, which was located between *prtI* and the gene upstream *I35_RS16285*. An *in silico* analysis by aligning this region with other such regions from closely related species revealed two putative PrtI-dependent promoters, based on the conservation of nucleotide sequence motifs that could constitute promoter core elements, one for *prtI* and one for *I35_RS16285*. A short version of the *prtI* and *I35_RS16285* (*‘catalase’*) promoters – P_{prtI}, and P_{cat} were cloned into a reporter plasmid. Their activities were tested in the presence and absence of PrtI and were found to be active in the presence of PrtI in both *E. coli* and *B. cenocepacia*, thus supporting the ChIP-seq results. Moreover the promoters were switched off in *B. cenocepacia* in the absence of a plasmid expressing the σ factor, but switched on in its presence, suggesting that PrtI requires a signal to be active and is inactive under normal growth conditions as in the case of most other ECF σ factors. Therefore, it was concluded that P_{prtI} and P_{cat} contain the minimum promoter elements required for recognition by PrtI and that PrtI regulates transcription of at least two genes – *prtI* itself and *I35_RS16285* (*catalase*).

To study the specific DNA sequence requirements for efficient promoter utilisation by P_{prtI}, a single basepair substitution scan of P_{prtI} was carried out. 16 derivatives of the promoter were studied to define the -10 and -35 regions. The results suggest that, GGAATA possibly forms the -35 region with 'GA' being most important for promoter utilisation. The pair of adenines might play a structural role rather than participating in direct interactions with σ_4 comparable to that of the 'AA motif' in the -35 promoter region of σ^E -dependent promoters in *E. coli* (Lane and Darst, 2006). Substituting the T following the -35 region also caused a decrease in promoter activity. The base at this position is not conserved in other species. It may suggest that though it may not form the -35 region its conformation may allow P_{prtI} to make productive interactions with the -35 region. As the 'CGTC' motif is so well conserved it may perhaps form the -10 region, although experimentally substituting only the G, T and C actually reduced promoter activity. It has been experimentally shown that a DXXR motif present in σ_2 of *Thermoanaerobacter tengcongensis* ECF σ factor specifically recognises CGTC, and this CGTC motif is often conserved in a lot of ECF σ factors (Liu et al., 2012). It was found that substituting the T base within the -10 region completely abolished promoter activity. This may suggest that perhaps T is an important base for σ_2 of P_{prtI} to make efficient contact with the promoter and possibly in opening the double stranded DNA during transcription initiation (Feklistov and Darst, 2011). Recent structural studies suggest that unlike σ_2 of housekeeping σ factor σ_{70} , forming two pockets to flip two bases from the -10 region (position -12 and -7), σ_2 of ECF σ factors can recognise six bases across the non-template ssDNA, when strand opening begins, and flips only one base to fit in its 'pocket' (Campagne et al., 2014, Campagne et al., 2015). As opposed to housekeeping σ factors, alternative σ factors have very specific target genes. They therefore have evolutionarily developed low melting capacity but high specificity (Darst et al., 2014). As compared to the traditional AT rich -10 region (TATAAT) recognised by the housekeeping σ factor, ECF σ factors are proposed to have increased promoter stringency by reducing the AT richness (Koo et al., 2009). A bioinformatics study on promoter analysis across all ECF σ factors shows that in nearly 60 ECF26 σ factors (like P_{prtI}) 'GGAATAA' is the -35 promoter region consistent with my experimental results, but

suggests 'GTTT' as the -10 region with a 15-16 spacer, where the first G and T are most important (Rhodius et al., 2013). My results demonstrate that G and T are indeed very important residues forming the -10 region, but this region does not extend beyond the C after the GT.

In addition to studying the -10 and -35 regions it would be useful to study the contribution of bases in the 14 bp spacer region to promoter recognition. The P_{prtI} spacer is slightly less GC-rich as compared to that of P_{cat}. It may be insightful to discern if this affects promoter activity as well as whether the length of the spacer region is important. To further gain confidence in the experimentally determined -10 and -35 regions, it would be useful to determine the transcription start site (TSS) of the genes.

The identification of the promoter motifs in *B. cenocepacia*, this now provides a tool for identifying the P_{prtI} dependent promoters in other species using FUZZNUC, without having to carry out CHIP-seq.

3.9.2 Deletion of *prtI* or *prtR* did not affect stress responses in *B. cenocepacia*

To construct marker-less deletion mutants in *B. cenocepacia*, a modified pEX18Tp-*pheS* system was initially used (Barrett et al., 2008). However, as the chlorophenylalanine based counter selection did not give rise to any mutants, an I-SceI based approach was adopted (Fazli et al., 2015) involving a further modified pEX18Tp-*pheS* derivative (H. Spiewak and M.S. Thomas, unpublished). The effect of deleting *prtI* or *prtR* on the oxidative stress response was investigated in *B. cenocepacia* since it was identified that P_{prtI} regulates a putative catalase. However, deleting either gene encoding the σ or anti- σ factor did not affect the ability to cope with oxidative stress using oxidants hydrogen peroxide, t-butyl hydroperoxide, sodium hypochlorite or methyl viologen. A study investigating the transcriptional response of biofilms of *B. cenocepacia* J2315 to hydrogen peroxide and sodium hypochlorite showed genes comprising approximately 5% of the genome being upregulated. Two iron starvation ECF σ factors *orbS* and *fIrS* were upregulated along with *BCAL3478* (*prtI*) and *BCAL3477* (*I35_RS16285*) in response to hydrogen peroxide but not sodium hypochlorite (Peeters et al., 2010). As *BCAL3477* transcription was higher when there was longer duration of

H₂O₂ exposure, it is hypothesised that it might not have a direct role in neutralisation of exogenous H₂O₂ and it might have been indirectly activated, perhaps through BCAL3478. However, these findings are not consistent with the results reported in this thesis since deleting *prtI* or *prtR* does not seem to affect oxidative stress response.

Deletion of *prtI* or *prtR* also did not affect protease production, motility, biofilm formation or EPS production. However, it must be noted that Δ *prtR* tended to form a mixed colony phenotype especially when forming mucoid colonies on EPS agar, but this was not consistent or significant. It was hypothesised that this was perhaps an effect of deleting the gene encoding the anti- σ factor rendering the cells 'sick'. A quantitative method of testing exopolysaccharide production is to grow cells in liquid media, extract EPS and weigh the amounts of polysaccharide produced. However, since the WT and mutants gave a similar degree of mucoidy on EPS media, this was not further investigated.

Although the current literature assigns multiple roles to PrtI-PrtR, they all have a common theme of regulating the production of pathogenically important proteins (Burger et al., 2000, Liehl et al., 2006, Vodovar et al., 2006, Kimbrel et al., 2010, Okrent et al., 2014, Song et al., 2014, Yang et al., 2014) van den Broek, 2005). Therefore, it might be useful to study the ability of the Δ *prtI* and Δ *prtR* mutants to infect models such as *Galleria mellonella* or lung cell lines compared to the WT. However, since no phenotypes were obtained on their deletion, and due to lack of time, this experiment was not carried out in this project.

3.9.3 Does PrtI regulate the production of a catalase?

In the final approach for identifying the role of PrtI, the catalase activity of I35_RS16285 was investigated, as no evidence had been provided at this point to show that PrtI-PrtR were involved in oxidative stress response. Here two experiments were carried out – (1) Purification of I35_RS16285 protein to test catalase enzyme activity and (2) complementing a catalase and flavo-enzyme deficient mutant of *E. coli* to determine whether it could help restore its ability to cope with oxidative stress. However, results from both experiments suggested that I35_RS16285, at least on its own, can not

function as a catalase. A useful positive control for testing enzyme activity with purified I35_RS16285 could be using a characterised catalase such as KatE from *B. cenocepacia* expressed with a 6xHis tag to establish that the tag did not affect catalase enzyme function. However, due to time constraints it was decided to only use commercial catalase as a positive control.

A BLAST search using I35_RS16285 revealed its similarity to SrpA (sulphur regulatory protein A) from *Synechococcus* species. The exact role of SrpA is not yet known, though it is present within a cluster of genes (*srpA* -> *srpM*) is proposed to be involved in sulphur regulation, along with *srpC* that encodes a putative chromate/sulphate transporter (Chen et al., 2008). The transcription activator CysR (cyanobacterial sulphur regulated gene) regulates *srpA* and also blocks the ability of cells to use thiocyanate as sole sulphur source (Laudenbach and Grossman, 1991, Nicholson and Laudenbach, 1995). However, deletion of *srpA* did not affect growth under high or low sulphate conditions (Nicholson and Laudenbach, 1995). A novel therapeutic agent (rhodamine based) that blocked peroxiredoxin activity also affected SrpA activity suggesting its role in oxidative stress in *Mycobacterium avium* (Bull et al., 2009). Peroxiredoxins are involved in degrading peroxidases and comprise a part of the anti-oxidant defence mechanism in bacteria (Hofmann et al., 2002).

An alignment of I35_RS16285 with SrpA and I35_RS16285 showed conserved haem binding residues, but this does not imply that SrpA/ I35_RS16285 may act on oxidant decomposition as there are proteins related to haem-dependent catalases that do not function as catalases. Additionally, SrpA and I35_RS16285 (approx. 36 kDa) are much smaller than characteristic catalases which typically contain 4 subunits with each subunit being 62 kDa. SrpA expression has been reported to be downregulated in the presence of H₂O₂ and has been proposed to be a monofunctional catalase, although its catalase activity has been explored but not yet proven in *Xanthomonas* (Zhang et al., 2009, Tondo et al., 2010).

SrpA has an N-terminal signal peptide-like sequence. This could be a signal sequence which targets SrpA to the Sec system and it might get translocated into the periplasm with removal of the signal peptide. It could then remain there or could get

transported out of the cell by the Type 2 secretion system. Alternatively, the hydrophobic sequence could act as a trans-membrane domain (TMD). Preceding this sequence are positively charged arginine residues which could imply that this region of the protein could be present inside the cytoplasm. Therefore, it can be hypothesised that if SrpA is anchored to the cytoplasmic membrane, the catalytic component will be located in the periplasm. Both these conjectures point towards the possibility that the function of SrpA (or catalase) is extracytoplasmic. It might catalyse reactions with nutrients such as thiocyanate which can then enter the cell and be metabolised.

In *B. cenocepacia* *I35_RS16280* is predicted to encode cytochrome B561. It might be involved in donating electrons to *I35_RS16285* in the above process. These two genes *I35_RS16280* and *I35_RS16285* encoding the putative cytochrome B561 and putative catalase respectively, possibly interact and work collaboratively.

Possibly, *I35_RS16285* might need certain modifications for carrying out its function which may be clearer if one could identify what stress PrtI responds to. Also, it is possible that *I35_RS16285* and SrpA might have different functions even though they have a similar haem binding pockets and conserved residues.

3.9.4 PrtR, an inhibitory σ factor regulatory protein

Based on the results obtained in the bacterial two-hybrid studies, the C-terminal domain (σ_4) of the σ factor PrtI interacts with the N-terminal domain of PrtR. It would be useful to next test whether PrtR-NTD is also capable of interacting with PrtI-NTD (σ_2). Since PrtR is a transmembrane protein this may demonstrate how its N-terminal domain exposed towards the cytoplasm keeps PrtI sequestered to the membrane under normal conditions. A positive maltose phenotype was unexpectedly obtained for PrtI-CTD self-interaction. However, there is no evidence in the literature for dimerization or oligomerisation of ECF σ factors. Alternative methods to experimentally verify this protein-protein interaction include pull-down assays or co-immunoprecipitation.

PrtR might regulate PrtI through an inhibitory mode. Experiments in *E. coli* showed loss of PrtI activity in the presence of PrtR and in *B. cenocepacia* it was shown that promoters were active in a *prtR* mutant. These results together suggest that

perhaps PrtR might release PrtI only upon receiving an inducing signal, consistent with the conventional and well established mode of ECF σ factor regulation (Mascher, 2013, Paget, 2015). Additionally, in the phenotypic experiments to study the effect of *prtR* deletion, there was a failure in complementing the mutant. It would be useful to measure the activity of the PrtI-dependent promoter P_{cat} in the *prtR* mutant harbouring pBBR2-PrtR. This would ascertain if the phenotypes that occurred following inactivation of a *prtR* was a result of that lesion rather than an undesired effect on the expression of an adjacent gene or a mutation that occurred elsewhere during the mutagenesis procedure. Also, *prtI* and *prtR* have overlapping stop and start sites, respectively, and they are likely to be co-transcribed. This suggests a functional relationship between the two encoded proteins.

However, the exact signal recognised by PrtR and conformational changes that might be occurring to release PrtI are yet to be studied. Elevated temperature has been suggested to be the inducing signal for PrtR activation in *P. fluorescens* although this was not experimentally determined (Burger et al., 2000). If I35_RS16285 is a SrpA orthologue, and if it is involved in sulphur metabolism, PrtR may be activated by alternative sulphur sources such as thiocyanate. Thiocyanate has been reported to be utilised by the *Bcc*, however its toxicity towards *Burkholderia* species has not yet been fully understood (Vu et al., 2013, Bernier et al., 2016). This raises the question whether thiocyanate concentration is indeed an extra-cytoplasmic stress, and if not, then what would be the requirement for an ECF σ factor (PrtI) to mediate its intracellular concentrations, since in such a case a cytoplasmic sensor would suffice.

3.9.5 The PrtI operon

Based on the gene organisation and experimentally determined location of the PrtI dependent promoters it appears that the PrtI regulon consists of four genes: *prtI* and *prtR* (the σ – anti- σ pair) that are co-transcribed and driven by the P_{prtI} promoter, and, 'catalase' (I35_RS16285) and 'cytochromeB561' (I35_RS16280), co-transcribed and driven by the P_{cat} promoter. RT-PCR results showed that these two pairs of genes might be co-transcribed, however due to DNA contamination in the mRNA samples this cannot

be concluded with conviction. Using an effective DNA degradation step (DNase) to clear the mRNA from contaminating DNA might help while repeating this experiment. Since there is a transcription terminator downstream of *prtR* and the gene upstream of *cytochromeB561* is predicted to be transcribed in the opposite orientation, it is expected that the PrtI operon might only consist of the above mentioned four genes. However, it would be useful to include primers from genes upstream and downstream of the PrtI regulon to determine the PrtI regulon limits, which would also serve as negative controls. It has been previously proposed that PrtI and PrtR can be co-transcribed driven from both a single promoter and separate promoters (Okrent et al., 2014). However, this is an unusual phenomenon for ECF σ -anti- σ factors.

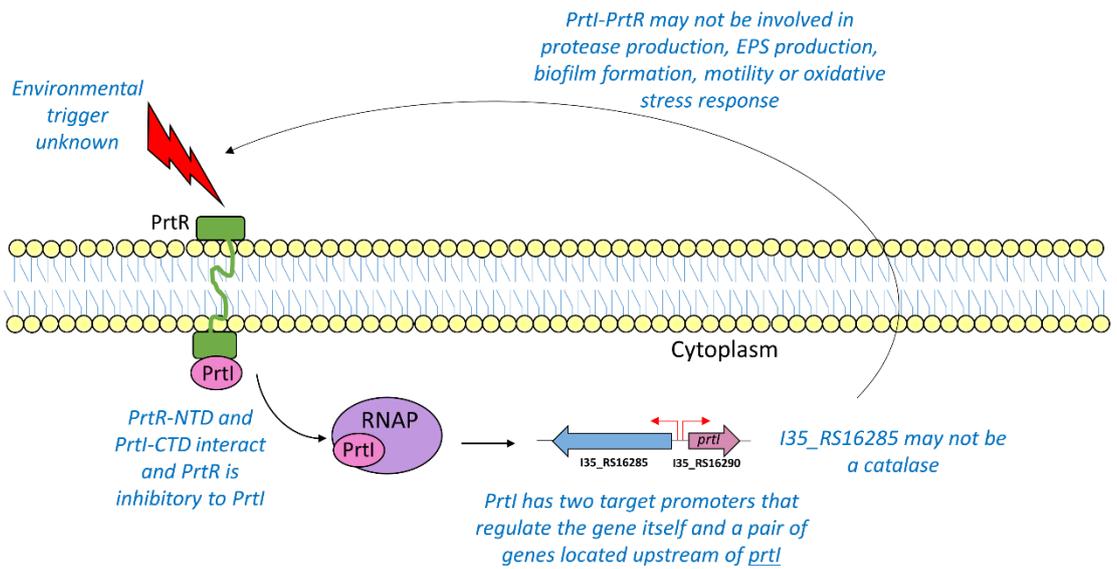


Figure 3.46. Current understanding of the PrtI-PrtR system in *B. cenocepacia*

Summary of the current understanding of the PrtI system based on the results presented in this chapter. PrtR is shown in green, PrtI in pink and core RNAP in purple. The red bent arrows represent approximate locations of PrtI-dependent promoters. While the environmental trigger remains unknown, it has been shown that PrtR is inhibitory to the activity of PrtI and PrtI recognises two promoters, P_{prtI} and P_{cat} .

**Chapter 4. Investigation of the roles
of ECF σ factors I35_RS16330
(Ecf41_{Bc1}), I35_RS29600 (Ecf41_{Bc2})
and I35_RS20320 (Ecf42_{Bc})**

4.1 Introduction

Two currently uncharacterised σ factors of *B. cenocepacia* are homologous to ECF41 type σ factors (I35_RS16330 or Ecf41_{Bc1} and I35_RS29600 or Ecf41_{Bc2}) and one is similar to the ECF42 group (I35_RS29320 or Ecf42). As described in section 1.6.3.2 the target promoters for ECF41 σ factors in *Bacillus licheniformis* and *Rhobacter sphaeroides* have been experimentally tested. The motifs TGTCACA and CGTC are proposed to form the -35 and -10 promoter elements (Wecke et al., 2012). A C-terminal on-board regulatory domain has been identified where a DGGG motif plays the role of an activator (Wecke et al., 2012). However, the exact role of this σ factor remains unknown. Although *M. tuberculosis* SigJ from this group does have a role in oxidative stress response, it may not be equivalently applicable for every member of this group of σ factors (Hu et al., 2004).

Similarly, as described in section 1.6.3.3 although an involvement has been proposed for an ECF42-type σ factor in biofilm formation and drug resistance, its exact role has not been defined and the molecular mechanisms of its activation, that presumably involves the C-terminal TPR domain, and the nature of the stress signal remain to be determined (Tettmann et al., 2014). The promoters recognised by this σ factor are also currently not known. In addition, the involvement of PhnB-type proteins encoded by genes usually present in the vicinity of *ecf42* type genes also need to be researched. Previous studies on other C-terminally extended ECF σ factors have shown that the extension can play a regulatory role (Gomez-Santos et al., 2011, Marcos-Torres et al., 2016). In case of *ecf42_{Bc}* in *B. cenocepacia* a rubrerythrin encoding gene is present downstream which suggests a possible role in oxidative stress response.

Based on the hypotheses and the aims described in section 1.7 the following experiments were designed:

1. *In silico* identification of TGTCACA(N)_{16±1}CGTC promoter motifs near *ecf41_{Bc1}* and *ecf41_{Bc2}* and their experimental verification to identify Ecf41-dependent genes.
2. To investigate the role of Ecf41_{Bc1} and Ecf41_{Bc2} in the oxidative stress response by constructing *ecf41_{Bc1}* and *ecf41_{Bc2}* mutants and comparing the effect of oxidative stress on the WT and *ecf41_{Bc1}* and *ecf41_{Bc2}* mutants.

3. Ascertain the role of the C-terminal extension by studying the activity of truncated derivatives of Ecf41_{Bc1} and Ecf41_{Bc2} at the target promoters.
4. *In silico* identification of putative Ecf42_{Bc}-dependent promoters and their experimental verification by reporter fusion analysis.
5. Study effect of oxidative stress, biofilm formation and antibiotic drug resistance of *ecf42_{Bc}* mutant and of Ecf42_{Bc} target genes if identified in 4.
6. Investigate the role of the C-terminal extension by studying the activity of a C-terminally truncated derivative of Ecf42_{Bc}.

4.2 Genomic context of *ecf41_{Bc1}* and *ecf41_{Bc2}*

The genomic loci of *ecf41_{Bc1}* and *ecf41_{Bc2}* along with those of homologous genes from other representative bacterial species are shown in Figure 4.1. As identified earlier, genes encoding Ecf41 σ factors mostly tend to be located close to genes encoding proteins that belong to one or more of the carboxymuconolactone decarboxylase, epimerase and oxidoreductase families of proteins or hypothetical proteins (Wecke et al., 2012). The gene encoding Ecf41_{Bc1} is present downstream and in the same orientation of a gene that is predicted to encode an alkyl hydroperoxidase D belonging to the carboxymuconolactone decarboxylase family of proteins. The gene encoding Ecf41_{Bc2} is located within a cluster of seven genes transcribed in the same orientation. *ecf41_{Bc2}* is located downstream of a gene encoding a putative diaminopimelate (DAP) epimerase (*dapF*) and upstream of a gene encoding a putative alkyl hydroperoxidase D. The last gene in this cluster is predicted to encode a putative organic hydroperoxide resistance protein belonging to the oxidoreductase family of proteins. Thus, the genomic context of these genes suggests a role in responding to oxidative stress.

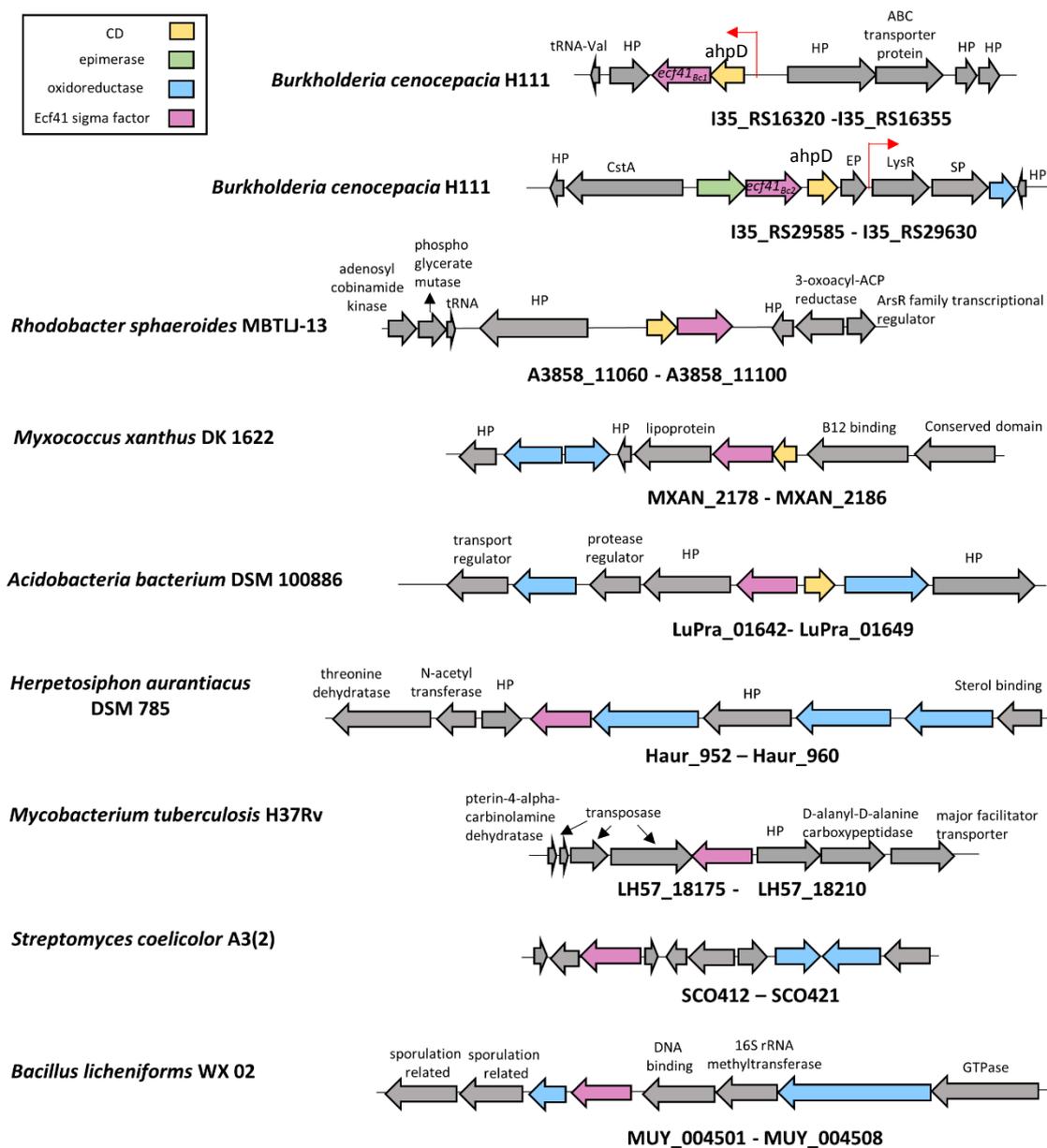


Figure 4.1 Genomic context of ECF41 σ factors

Gene organisation at *ecf41* locus from *B. cenocepacia* H111 compared to the corresponding locus in other species as indicated. *ecf41* genes are shown in pink while genes belonging to the carboxymuconolactone decarboxylase (CD), epimerase and oxidoreductase families are shown in yellow, green and blue, respectively. All other genes are shown in grey and their predicted gene products are shown above the corresponding arrow. Red bent arrows indicate the location of Ecf41_{BC}-dependent promoters. Gene locus numbers are shown below the corresponding gene. HP = hypothetical protein, SP = secreted protein, MP = membrane protein, EP = Exported protein. For a detailed genomic context conservation analysis of ECF41 σ factors across all phyla, refer to Wecke et al., 2012.

An alignment of the amino acid sequence of Ecf41_{Bc1} and Ecf41_{Bc2} with the homologous Ecf41 proteins in some other species is shown in Figure 4.2. It is evident that along with σ_2 and σ_4 , all the aligned proteins share amino acid homology in their C-terminal extension as observed previously (Wecke et al., 2012). Of note is a WLPEP motif located between σ_2 and σ_4 that seems to be highly conserved. The DGGGK motif in the C-terminal extension appears to be fairly conserved. In addition, an 'RNPDKL' motif located at the end of the C-terminal domain also seems to be fairly conserved.

Together, the genomic context, presence of a C-terminal extension as well as the conservation of motifs that lie outside of the σ_2 and σ_4 regions in the putative Ecf41 σ factors of *B. cenocepacia* suggest that they must indeed belong to the ECF41 group of σ factors based on the current characterisation of this group.

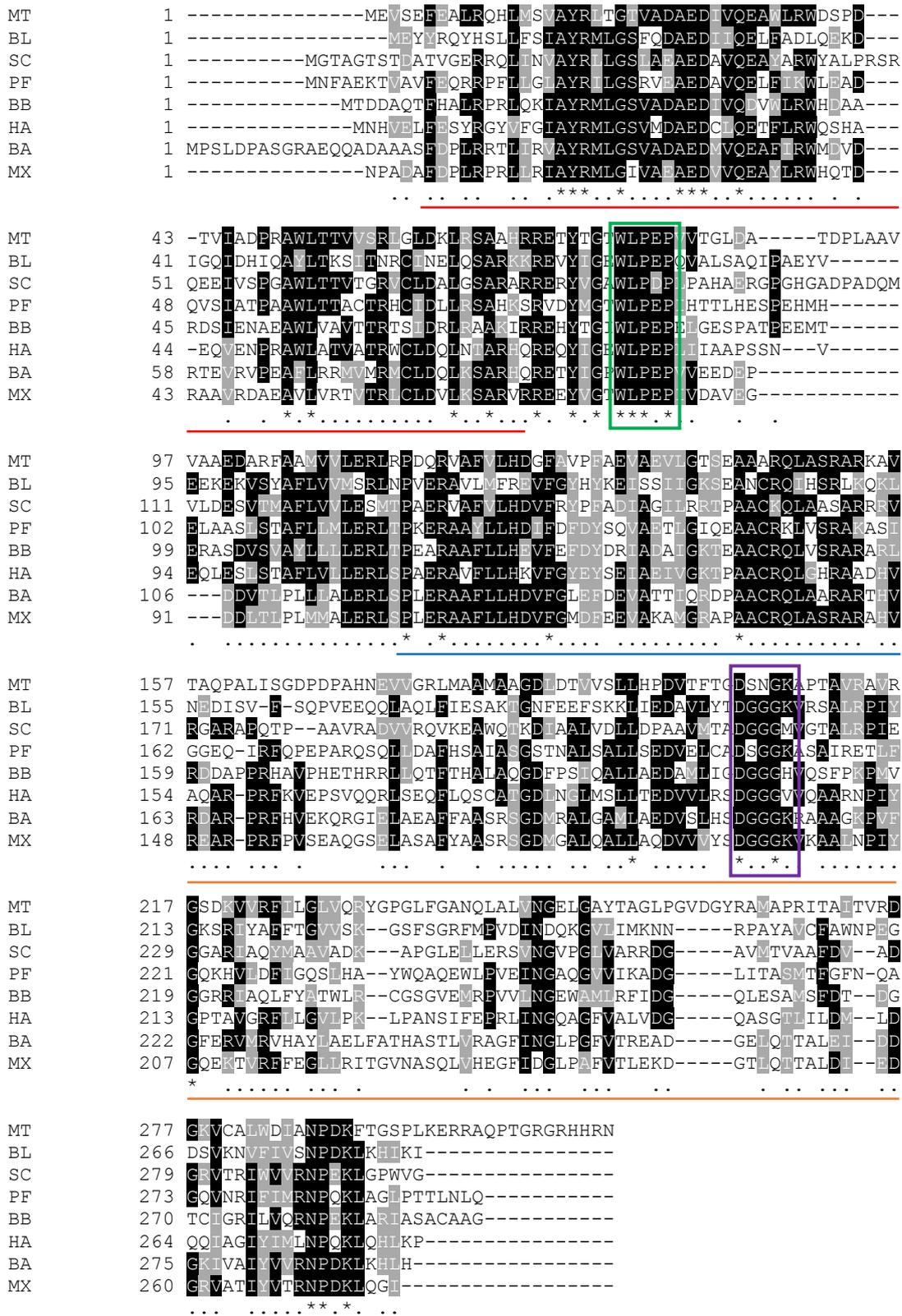


Figure 4.2. Amino acid sequence alignment of Ecf41_{Bc1} and Ecf41_{Bc2} from *B. cenocepacia* H111 with related proteins in other species

Ecf41_{Bc1} (BA) and Ecf41_{Bc2} (BB) amino acid sequence from *B. cenocepacia* H111 was aligned with homologous proteins from *Mycobacterium tuberculosis* H37Rv (SigJ) (MT), *Bacillus*

licheniforms WX02 (BL), *Streptomyces coelicolor* A3(2) (SC), *Pseudomonas fluorescens* Pf01 (PF), *Herpetosiphon aurantiacus* DSM 785 (HA) and *Myxococcus xanthus* DK 1622 (MX) using Clustal Omega and shaded using BOXSHADE. Amino acids which are identical at the corresponding position in $\geq 50\%$ of sequences are shown in white font with black highlighting and asterisks underneath while amino acids which are similar at the corresponding position are shown in white font with grey highlighting with dots underneath. σ_2 , σ_4 and the C-terminal extension are underlined in red, blue, and orange, respectively. Conserved motifs 'WLPEP' and 'DGGGK' are shown in green and purple boxes, respectively.

4.3 Investigation of the role of Ecf41_{Bc1} and Ecf41_{Bc2} by identification target promoters

The objective of this study was to infer the role of Ecf41_{Bc1} and Ecf41_{Bc2} by identifying the genes that have Ecf41-dependent promoters. This hypothesis could then be tested by challenging the σ factor mutant with appropriate stresses. In order to identify Ecf41_{Bc}-targeted genes, it was decided to bioinformatically locate putative promoters within the *B. cenocepacia* genome and test their functionality experimentally.

4.3.1 Identification of Ecf41_{Bc1} and Ecf41_{Bc2}-regulated promoters *in silico*

The putative promoter recognised by Ecf41 σ factors has been previously defined as 'TGTCACA'(N)₁₆'CGTC' (Wecke et al., 2012). This motif was scanned for within the intergenic regions in the vicinity of genes encoding Ecf41_{Bc1} and Ecf41_{Bc2}. In case of *ecf41_{Bc1}*, this motif was found to be present upstream of the putative *ahpD* gene (I35_RS16335). It was speculated that P_{ahpD} drives transcription of *ecf41_{Bc1}* since *ecf41_{Bc1}* is present immediately downstream of *ahpD* (the translation initiation codon of *ecf41_{Bc1}* overlaps the translation termination codon of *ahpD*). Therefore, the DNA sequence between this gene I35_RS16335 and I35_RS16340 that was predicted to contain this promoter was aligned with the corresponding region from closely related species of the *Bcc*. As shown in Figure 4.3A, the TGTCACA and CGTC motifs were found to be conserved along with other bases in the putative promoter region. An 85 bp DNA fragment predicted to contain this promoter, as well as an upstream conserved region of multiple A and T bases, termed as P_{ahpD} is shown in Figure 4.3B.

In the case of *ecf41_{Bc2}*, the putative promoter motif was also found to be present upstream of a gene encoding a LysR-type transcriptional regulator (I35_RS29615). However, the spacer was found to consist of 17 bp rather than 16. This gene is separated from *ecf41_{Bc2}* by two genes and as described earlier is predicted to be transcribed in the same orientation as *ecf41_{Bc2}*. The intergenic region between I35_RS29610 and I35_RS29615 was aligned with the corresponding regions of three other closely related species within the *Bcc*, as shown in Figure 4.4A. Along with the predicted -10 and -35 promoter regions the bases forming the spacer region as well as bases lying outside the promoter regions were found to be fairly conserved. A 104 bp long DNA fragment termed as P_{lysR} possessing the putative -10 and -35 promoter elements and a conserved A+T-rich tract located upstream of the -35 motif is shown in Figure 4.4B.

To determine if TGTCACA(N)₁₆CGTC motifs constituted the *ecf41*_{Bc2} promoter the region upstream of the putative epimerase (*dapF*) present immediately upstream of *ecf41*_{Bc2} (the *ecf41*_{Bc2} start codon overlaps the *dapF* stop codon) or the intergenic region upstream of *ecf41*_{Bc2}, in cases where the gene encoding the epimerase was absent, was aligned with the corresponding region in other species of the *Bcc*. As shown in Figure 4.5, although TGTCACA(N)₁₆CGTC motifs were not present in this region, 'TTGCCC(N)₁₇TACAGT' was present 28 bases upstream of the *dapF* ORF in *B. cenocepacia* and *B. cepacia* which may be speculated to constitute the -35 and -10 σ^{70} -dependent promoter elements. However, these bases were not conserved in *B. multivorans* and *B. dolosa*, perhaps due to the gene encoding the epimerase being absent in these two cases.

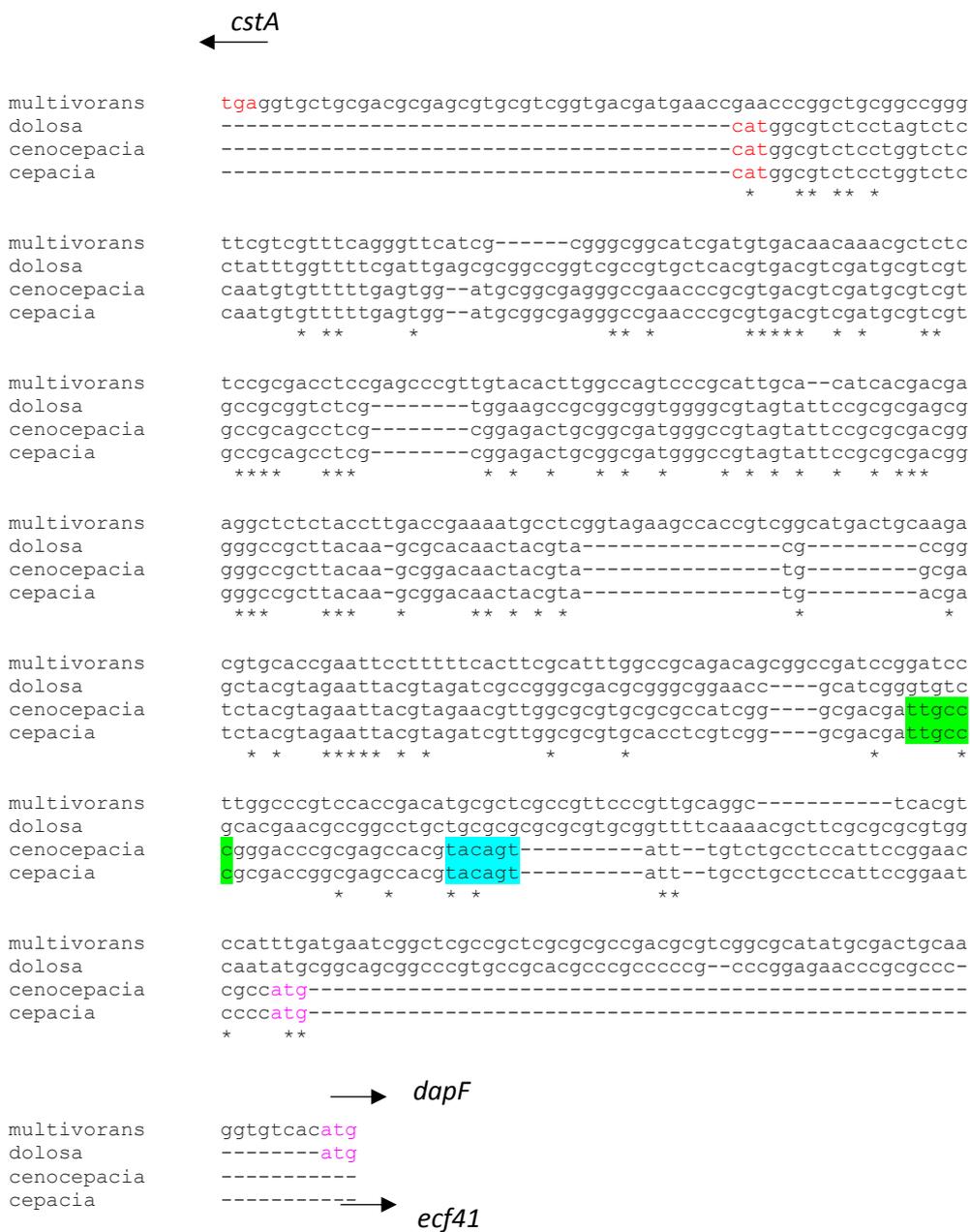


Figure 4.5. Clustal Omega alignment for identification of putative *ecf41*_{Bc2} promoter

The 275 bp intergenic region between I35_RS29590 (putative carbon starvation A (CstA family protein) and I35_RS29595 (putative DAP epimerase, DapF, or in cases where epimerase was absent, Ecf41) from *B. cenocepacia* was aligned with the corresponding region from *B. cepacia*, *B. multivorans* and *B. dolosa* using Clustal-omega. Note that in *B. multivorans* and *B. dolosa* the *dapF* gene is not present upstream of the *ecf41* gene. The stop codons are shown in red and start codons are shown in magenta font. Bases that are conserved at the corresponding position in all sequences are marked with an asterisk. Putative -10 and -35 promoter elements are highlighted in blue and green, respectively.

4.3.2 Construction of plasmids harbouring either of the σ factors Ecf41_{Bc1}/Ecf41_{Bc2} or putative promoter derivatives P_{ahpD}/P_{lysR}

Since the conditions required for *ecf41_{Bc1}* σ factor gene expression were not known, it was decided to use plasmid pBBR1MCS2 (pBBR2) that could allow constitutive expression of genes encoding these two proteins under the control of the *lac* promoter. Using primers BCAL3486for and BCAL3486rev, a 946 bp DNA fragment containing *ecf41_{Bc1}* and its putative Shine-Dalgarno sequence was amplified using KOD polymerase and a boiled H111 lysate as template. Similarly, a DNA fragment consisting of 960 bp containing *ecf41_{Bc2}* and its putative Shine-Dalgarno sequence was amplified using primers BCAM2748for and BCAM2748.bam.rev. Each of these DNA fragments were digested with *Hind*III and *Bam*HI enzymes and ligated separately into pBBR2 also digested with the same enzymes. The ligation reactions were used to transform *E. coli* JM83 cells which were spread on LB agar containing kanamycin, X-gal and IPTG to allow for blue/white colony screening. The transformants thus obtained that gave white colonies were screened by GoTaq PCR using the insert specific primers which were used in the initial amplifications mentioned above. Plasmids present in clones that screened positive were checked for their size in an agarose gel. pBBR2-Ecf41_{Bc1} was expected to be 6.090 kb and pBBR2-Ecf41_{Bc2} was expected to be 6.104 kb. Plasmids of the correct size were then checked for their DNA sequence integrity of the cloned DNA by nucleotide sequencing.

The *lacZ*-reporter plasmid pKAGd4 was used as a vector for measuring Ecf41_{Bc}-promoter dependent activity. In order to do this, complementary oligonucleotides corresponding to segments of P_{ahpD} and P_{lysR} (Figures 4.3B and 4.4B, respectively) were annealed as described in section 2.2.8. Oligonucleotides pBCAL3487for, pBCAL3487for2, pBCAL3487rev and pBCAL3487rev2 were annealed to assemble the P_{ahpD} DNA fragment. Likewise, to assemble the P_{lysR} DNA fragment oligonucleotides pBCAM2751for, pBCAM2751for2, pBCAM2751rev2 and pBCAM2751rev were annealed. Each of these DNA fragments were ligated separately to pKAGd4 that was digested with *Hind*III and *Bam*HI enzymes. *E. coli* MC1061 cells were then transformed with the ligation reactions and the transformants were spread on LB agar containing chloramphenicol and X-gal.

Blue colonies thus obtained were PCR screened using vector-specific primers AP10 and AP11 and GoTaq polymerase. Clones that appeared to contain the desired insert based on the presence of an amplicon of 85 bp for P_{ahpD} and 104 bp for P_{lysR} were confirmed to have the correct DNA sequence by nucleotide sequencing.

4.3.3 Analysis of putative Ecf41_{Bc1} and Ecf41_{Bc2}-dependent promoter activity in the presence of the respective ECF σ factors in *E. coli*

pBBR2 and pBBR2-Ecf41_{Bc1} were transferred to MC1061 cells containing pKAGd4- P_{ahpD} . Similarly, pBBR2 and pBBR2-Ecf41_{Bc2} were transferred to MC1061 cells containing pKAGd4- P_{lysR} . The activities of P_{ahpD} and P_{lysR} were then measured by carrying out the β -galactosidase assay. As shown in Figure 4.6, both promoters were found to be active in the presence of the respective σ factors and inactive in their absence. In particular, P_{lysR} was highly active in the presence of Ecf41_{Bc2}. This provided evidence that these DNA fragments might indeed contain Ecf41-dependent promoters. In addition, the ability of each of these σ factors to utilise the promoter dependent on the other σ factor was also tested, as the promoter sequences were so similar. To test this, pBBR2-Ecf41_{Bc1} was transferred to MC1061 cells containing pKAGd4- P_{lysR} and pBBR2-Ecf41_{Bc2} was transferred to MC1061 cells containing pKAGd4- P_{ahpD} . The activities of each P_{ahpD} and P_{lysR} were then measured by carrying out the β -galactosidase assays. pKAGd4- P_{lysR} was found to be poorly active in the presence of Ecf41_{Bc1}. In contrast, pKAGd4- P_{ahpD} was found to be very active in the presence of Ecf41_{Bc2}. Although this activity was approximately 20% of the activity of P_{lysR} in the presence of Ecf41_{Bc2}, it was still 4 times more active than its Ecf41_{Bc1}-dependent activity.

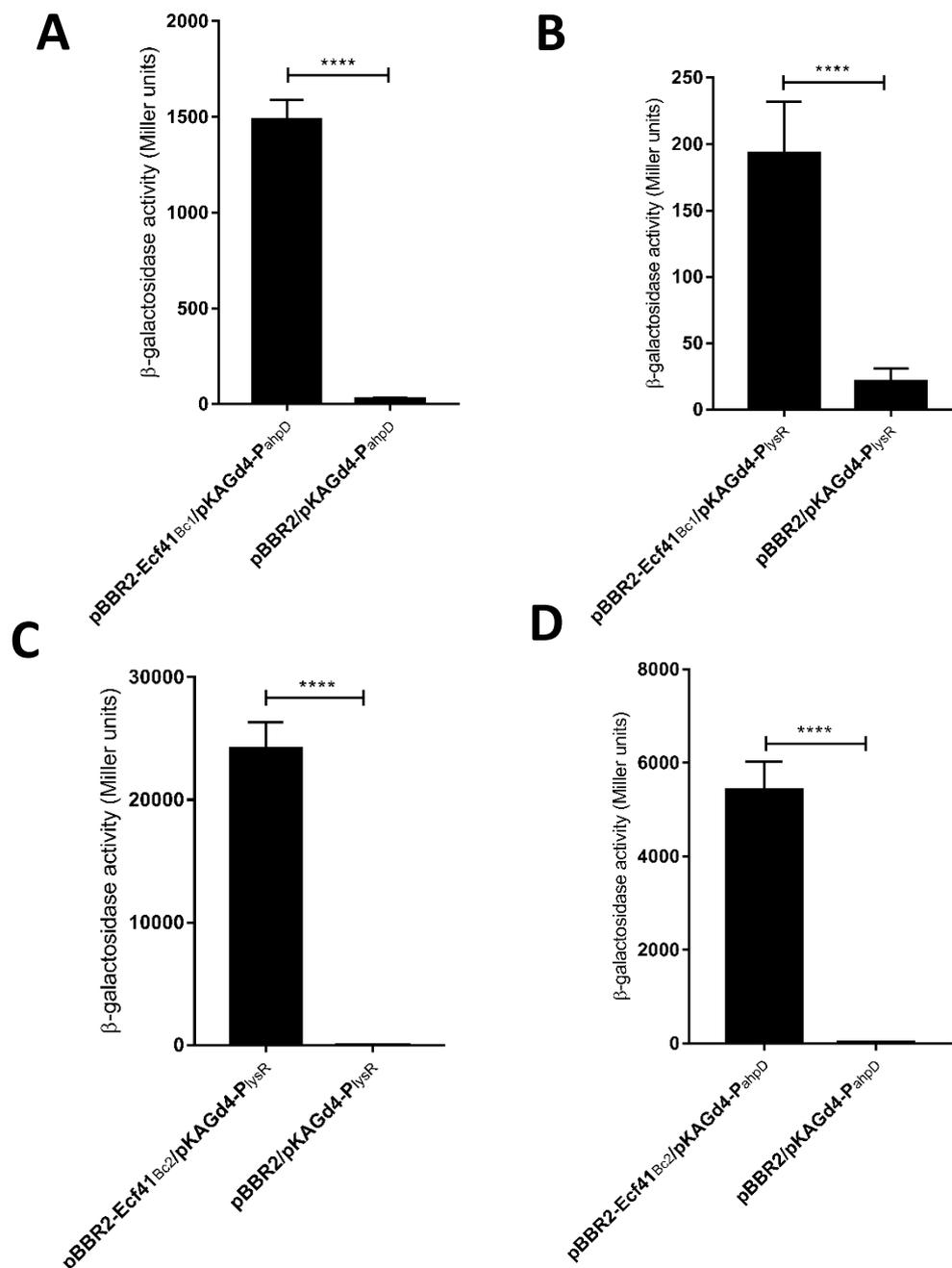


Figure 4.6. Activity of putative Ecf41-dependent promoters in the presence and absence of the ECF41 σ factors in *E. coli*

All possible combinations of pBBR2, pBBR2-Ecf41_{Bc1} and pBBR2-Ecf41_{Bc2} with pKAGd4-P_{ahpD} and pKAGd4-P_{lysR} were introduced into MC1061 as indicated. The promoter activity of each of these was measured using the β -galactosidase assay where cultures were grown in LB with kanamycin and chloramphenicol. All activities shown were obtained by subtracting the activity of the same strain containing pKAGd4 and the corresponding pBBR2 derivative assayed under the same conditions (background control). X-axis represents strain. Statistical significance was determined by performing a students-t test, **** = $p < 0.0001$, (n=6).

4.3.4 Attempted construction of plasmids expressing FLAG-tagged Ecf41_{Bc} derivatives for ChIP-seq

Once the putative promoters were verified experimentally it was decided to scan for additional Ecf41-promoters *in vivo* by carrying out chromatin immunoprecipitation and deep sequencing. It was decided to express the σ factors under the control of an IPTG inducible promoter. Hence the vector pSRK-Km that carries P_{lac} and *lacI* was used. Using primers BCAL3486FLAGFor and BCAL3486FLAGRev a DNA fragment of 664 bp constituting *ecf41_{Bc1}* without its C-terminal extension coding sequence was amplified using KOD polymerase. The C-terminal extension was excluded since it has been reported to play an inhibitory role (Wecke et al., 2012). The reverse primer was designed such that a FLAG epitope tag would be incorporated at the C-terminal end of the encoded truncated Ecf41_{Bc1}. Similarly, *ecf41_{Bc2}* without its C-terminal extension coding sequence but with an added FLAG tag coding sequence was also amplified using BCAM2748FLAGFor and BCAM2748FLAGRev, giving a product of 666 bp. Each of these amplicons were separately digested with *NdeI* and *XbaI* enzymes and ligated into pSRK-Km which was also digested with the same enzymes. Then the ligation reactions were used to transform JM83 cells. Using the blue/white screen, white colonies which were expected to contain the plasmid with desired insert were PCR screened using insert-specific primers used for generating the original amplicons for cloning. The clones that screened positive were checked for their DNA sequence integrity. Although the clones contained the desired insert, nucleotide sequencing revealed presence of multiple mutations in the insert. While this molecular cloning was repeated several times, clones with the expected sequence could not be successfully obtained. Due to lack of time, it was decided to terminate this experiment.

4.3.5 FUZZNUC based identification of additional ECF41-dependent promoters

An *in silico* search can be carried out to identify a particular motif within any given genome using the program FUZZNUC. Therefore, this tool was used to identify whether other potential Ecf41-dependent promoters occur in other regions within the *B. cenocepacia* genome as well as in other species.

4.3.5.1 FUZZNUC based identification of additional ECF41-dependent promoters on the *B. cenocepacia* genome

Using pattern TGTCACA(N)_{16,17}CGTC, one novel hit was obtained on the *B. cenocepacia* H111 genome located on chromosome 1 (1487817-1487844), with the following sequence: **TGTCACA**GCGCAGGTCCGGGATCG**CGTC**. This sequence is located within a potential lipoprotein coding sequence but approximately 250 bp upstream of a LysR-type regulator. Since the putative -35 and -10 regions are separated by 17 bp, it may be a Ecf41_{Bc2}-dependent promoter.

4.3.5.2 FUZZNUC based identification of additional ECF41-dependent promoters in other bacterial genomes

In three other bacterial species belonging to different phyla that encode ECF41 group σ factors, the same Prosite pattern, TGTCACA(N)_{16,17}CGTC, was used to scan the entire genome employing FUZZNUC. In the Actinobacterium *Mycobacterium tuberculosis* H37Rv, 4 hits were obtained. One of these was located approximately 40 bp upstream of the sole ECF41 σ factor encoding gene. The remaining three hits were not present upstream of the start codon of a gene and thus it was concluded that they perhaps do not act as promoters. A scan in *Myxococcus xanthus* DK1622 revealed 7 hits. As might be expected, one of these hits was located upstream of an oxidoreductase encoding gene present near the Ecf41 encoding gene. One hit was located upstream of the start codon of a gene encoding a metal binding protein but was within a gene encoding a sulphate protein. Interestingly, this gene encoding a metal binding protein was present upstream of a gene encoding an uncharacterised ECF σ factor. Another putative novel promoter was identified upstream of a gene encoding a putative glycosyl transferase group 1 family protein. All other hits were concluded not to be promoters since they were not present upstream of a start codon. A scan of the *Bacillus licheniformis* ATCC 14580 genome gave 4 hits. However, they were present upstream of a converging gene and therefore not likely to be promoters.

4.4 Genetic analysis of the role of ECF41 group σ factors in *B. cenocepacia*

Verification of the Ecf41-dependent promoters showed that Ecf41_{Bc1} directs the transcription of a putative *ahpD* while Ecf41_{Bc2} directs the transcription of one or two genes encoding putative LysR-type regulators. The presence of oxidoreductase genes in the purlieu of these σ factor genes suggests their role in an oxidative stress response. Although previous studies have experimentally shown that inactivation of the *ecf41* genes in *Bacillus licheniformis* and *Rhodobacter sphaeroides* did not affect the ability of the bacterium to respond to oxidative stress (Wecke et al., 2012), it was decided to investigate such a role for Ecf41_{Bc1} and Ecf41_{Bc2} in *B. cenocepacia*. To do this, in *B. cenocepacia* H111, it was decided to construct in-frame *ecf41* deletion mutants to avoid polar effects. It was envisaged that loss of one of the *ecf41* genes might be partially or completely compensated for by the other and hence in addition to single mutants (Δ *ecf41*_{Bc1} and Δ *ecf41*_{Bc2}), it was also decided to construct a double deletion mutant (Δ *ecf41*_{Bc1}- Δ *ecf41*_{Bc2}).

4.4.1 Construction of an *ecf41*_{Bc1} mutant

Using SOE PCR primers 3486delA and 3486delB, a 543 bp DNA fragment (AB) that contained the region upstream of *ecf41*_{Bc1} as well as the first six codons of *ecf41*_{Bc1} was amplified using Q5 polymerase and boiled H111 lysate as template. Similarly, using SOE PCR primers 3486delC and 3486delD, a 574 bp DNA fragment (CD) that contained the region downstream of *ecf41*_{Bc1} and last three codons of *ecf41*_{Bc1} was amplified. In a subsequent PCR using Q5 polymerase and using 'AB' and 'CD' PCR products in equimolar concentrations as the template, amplification was carried out using the 3486delA and 3486delD primers. This gave a fusion product of 1.117 kb since primers 3486delB and 3486delC were designed to contain homologous bases. This PCR product (AD), which had thus lost 285 codons of the central region of *ecf41*_{Bc1} (total 294 codons), was digested with *Bam*HI and *Hind*III restriction enzymes was ligated with pEX18-TpTer-Scel-*pheS*-Cm also digested with the same enzymes. Then these ligation reactions were used to transform JM83 cells which were then selected on LB containing chloramphenicol and trimethoprim. A PCR screen was carried out to find clones with the desired insert using

vector-specific primers pEX18Tpfor and pEX18Tprev. Plasmids obtained from transformants that screened for positive were subjected to DNA sequence analysis and the correct plasmid, pEX18-TpTer-Scel-*pheS*-Cm- Δ Ecf41_{Bc1}, was transferred into *B. cenocepacia* H111 conjugally via *E. coli* S17-1 cells. To confirm that the plasmid had integrated into the genome by recombination in the trimethoprim-resistant colonies, thus forming 'co-integrates', a PCR screen was carried out using GoTaq polymerase and vector-specific primers pEXTpfor and pEXTprev. Next, pDAI-I-SceI was introduced into the co-integrate by performing a conjugation using *E. coli* SM10 as donor strain and selecting *B. cenocepacia* exconjugants on LB agar containing tetracycline and chloramphenicol. This plasmid expressed the yeast meganuclease I-SceI which introduced a cut at the I-SceI recognition site contained on the integrated vector, allowing a second recombination event to occur that rescues strains containing the lesion. This recombination can give rise to cells that had either reverted back to WT or to an *ecf41*_{Bc1} mutant (Δ *ecf41*_{Bc1}). Tetracycline-resistant *B. cenocepacia* exconjugants were PCR screened using primers 3486outfor and 3486outrev that anneal to genomic sequences outside the region complementary to primers A and D. WT exconjugants used as a template were expected to give rise to a product of 2.056 kb while Δ *ecf41*_{Bc1} mutants were expected to give rise to a product of 1.201 kb since it had lost 285 codons of *ecf41*_{Bc1}. Figure 4.7A shows a PCR screen of one such mutant compared to the WT. The mutant was streaked 3 times on 0.1% chlorophenylalanine medium to select for loss of pDAI-I-SceI. A general schematic outline of this procedure is shown in Figure 3.29.

4.4.2 Construction of an *ecf41*_{Bc2} mutant

2748delA and 2748delB SOE PCR primers were used to amplify a DNA fragment (AB) of 539 bp containing the region upstream of *ecf41*_{Bc2} and the first five codons of *ecf41*_{Bc2} using the H111 boiled lysate and Q5 polymerase. Likewise, a DNA fragment of 589 bp (CD) containing the last 14 codons of *ecf41*_{Bc2} was amplified using primers 2748delC and 2748delD. Then, using primers 2748delA and 2748delD an amplification was carried out again, this time using AB and CD as template in equimolar concentration. The resulting product of 1128 bp was obtained where 276 codons in the central region

of *ecf41*_{Bc2} (295 codons). This fragment (AD) was subsequently cloned into pEX18-TpTer-Scel-*pheS* between the *Hind*III and *Bam*HI sites. The sequence integrity of this construct was confirmed by nucleotide sequencing. pEX18-TpTer-Scel-*pheS*- Δ Ecf41_{Bc2} was transferred to *B. cenocepacia* H111 by performing a conjugation using *E. coli* S17-1 cells as the donor strain. The integration of the plasmid into the genome of H111 by recombination, thereby forming co-integrates, was confirmed in the exconjugants by carrying out a PCR where vector-specific primers pEX18Tpfor and pEX18Tprev were used. pDAI-I-SceI was then transferred to a co-integrate conjugally using *E. coli* SM10 as donor strain. This allowed the expression of I-SceI that introduced a cut at the I-SceI restriction site contained on the integrated vector. The second recombination event allowed either formation of WT or Δ *ecf41*_{Bc2}. This was confirmed by another PCR screen this time using primers 2748outfor2 and 2748outrev2 that anneal to genomic sequences outside the region complementary to primers A and D. WT when used as a template was expected to give a product of 2.009 kb while Δ *ecf41*_{Bc2} when used as a template was expected to give a product of 1.181 kb since it had lost 276 codons of *ecf41*_{Bc2}. Figure 4.7B shows a PCR screen of one such mutant compared to the WT. The mutant was streaked 3 times on 0.1% chlorophenylalanine medium to select for loss of pDAI-I-SceI. A general schematic outline of this procedure is shown in Figure 3.29.

4.4.3 Construction of an *ecf41*_{Bc1} and *ecf41*_{Bc2} in-frame deletion mutant

To construct a double deletion mutant (Δ *ecf41*_{Bc1}- Δ *ecf41*_{Bc2}), pEX18-TpTer-Scel-*pheS*-Cm- Δ Ecf41_{Bc1} (constructed as described in section 4.4.1) was transferred to *B. cenocepacia* H111 Δ *ecf41*_{Bc2} by carrying out a conjugation where S17-1 cells were used as a donor strain on LB medium containing chloramphenicol and trimethoprim. Then, formation of co-integrates due to the first recombination event was confirmed by PCR screening exconjugants using primers pEX18Tpfor and pEX18Tprev that are vector-specific. Once this co-integration was confirmed, the plasmid pDAI-I-SceI was introduced conjugally through the *E. coli* SM10 donor strain. The I-SceI enzyme thus caused a cut on the restriction site allowing a second recombination event to occur which resulted in viable bacteria. Thus, either H111 Δ *ecf41*_{Bc2} or H111 Δ *ecf41*_{Bc1}- Δ *ecf41*_{Bc2} was obtained

depending on the site of the second recombination event. The formation of the desired mutant was confirmed by PCR screening the colonies using primers 3486outfor and 3486outrev as well as 2748outfor2 and 2748outrev2 as described in section 4.4.1 and 4.4.2. The resulting double mutant would be expected to have lost both 285 codons of *ecf41_{Bc1}* and 276 codons of *ecf41_{Bc2}*. WT used as a template using primers 3486outfor and 3486outrev was expected to give a product of 2.056 kb while $\Delta ecf41_{Bc1}-\Delta ecf41_{Bc2}$ used as a template was expected to give a product of 1.201 kb. WT used as a template using primers 2748outfor2 and 2748outrev2 was expected to give a product of 2.009 kb while $\Delta ecf41_{Bc1}-\Delta ecf41_{Bc2}$ used as a template was expected to give a product of 1.181 kb. Figure 4.7A shows a PCR screen of one such mutant compared to the WT. The mutant was streaked 3 times on 0.1% chlorophenylalanine medium to select for the loss of pDAI-I_{Scel}.

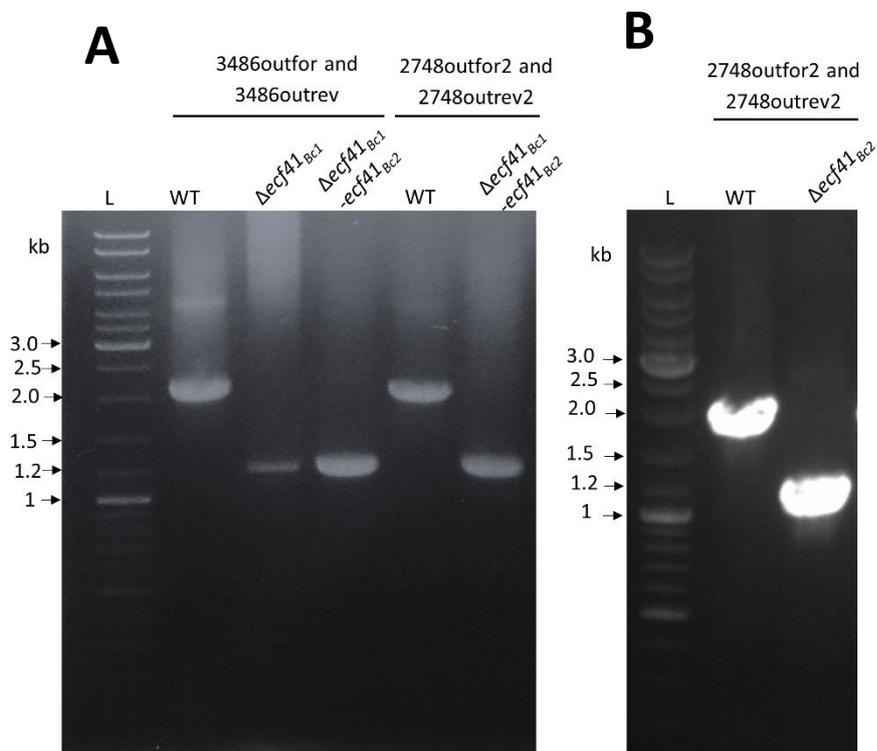


Figure 4.7. PCR screening of candidate H111 $\Delta ecf41_{Bc1}$, H111 $\Delta ecf41_{Bc2}$ and H111 $\Delta ecf41_{Bc1} \Delta ecf41_{Bc2}$ mutants

(A) Agarose gel showing PCR screen of candidate H111 $\Delta ecf41_{Bc1}$ mutant using primers 3486outfor and 3486outrev which anneal at genomic locations lying outside the expected deleted region of *ecf41_{Bc1}*, giving a PCR product of 2.056 kb for WT and 1.201 kb for $\Delta ecf41_{Bc1}$, and a screen of a candidate H111 $\Delta ecf41_{Bc1} \Delta ecf41_{Bc2}$ mutant using primers 2748outfor2 and 2748outrev2 which anneal at genomic locations lying outside the expected deleted region of *ecf41_{Bc2}* giving a PCR product of 2.009 kb for WT and 1.181 kb for $\Delta ecf41_{Bc2}$. L = GeneRuler DNA mix; WT = H111 boiled lysate used as template. (B) Agarose gel showing PCR screen of candidate H111 $\Delta ecf41_{Bc2}$ mutant using primers 2748outfor2 and 2748outrev2. L = GeneRuler DNA mix; WT = H111 boiled lysate used as template.

4.4.4 Effect of deletion of *ecf41* genes on growth of *B. cenocepacia* under normal conditions

To assess whether either or both of the *ecf41* genes were required for normal growth of *B. cenocepacia* under standard laboratory conditions, liquid cultures were grown in LB and the optical density was monitored throughout the growth cycle as described in section 2.1.4. As shown in Figure 4.8A, it was evident that $\Delta ecf41_{Bc1}$, $\Delta ecf41_{Bc2}$ and $\Delta ecf41_{Bc1-\Delta ecf41_{Bc2}}$ cells grew at a similar rate as that of the WT and reached a similar final density. Also, the efficiency to form colonies was tested on LB agar. As shown in Figure 4.8B and C, the $\Delta ecf41_{Bc1}$, $\Delta ecf41_{Bc2}$ and $\Delta ecf41_{Bc1-\Delta ecf41_{Bc2}}$ mutants formed colonies with a similar efficiency as that of WT. It was observed that the $\Delta ecf41_{Bc2}$ mutant occasionally gave rise to a mixed colony phenotype whereby some colonies were as those formed by WT *B. cenocepacia* while the other colonies were significantly smaller. However, this observation was not consistent and was therefore not further investigated. Overall, it was concluded that deleting either or both of the *ecf41* genes did not severely affect bacterial growth under standard lab conditions.

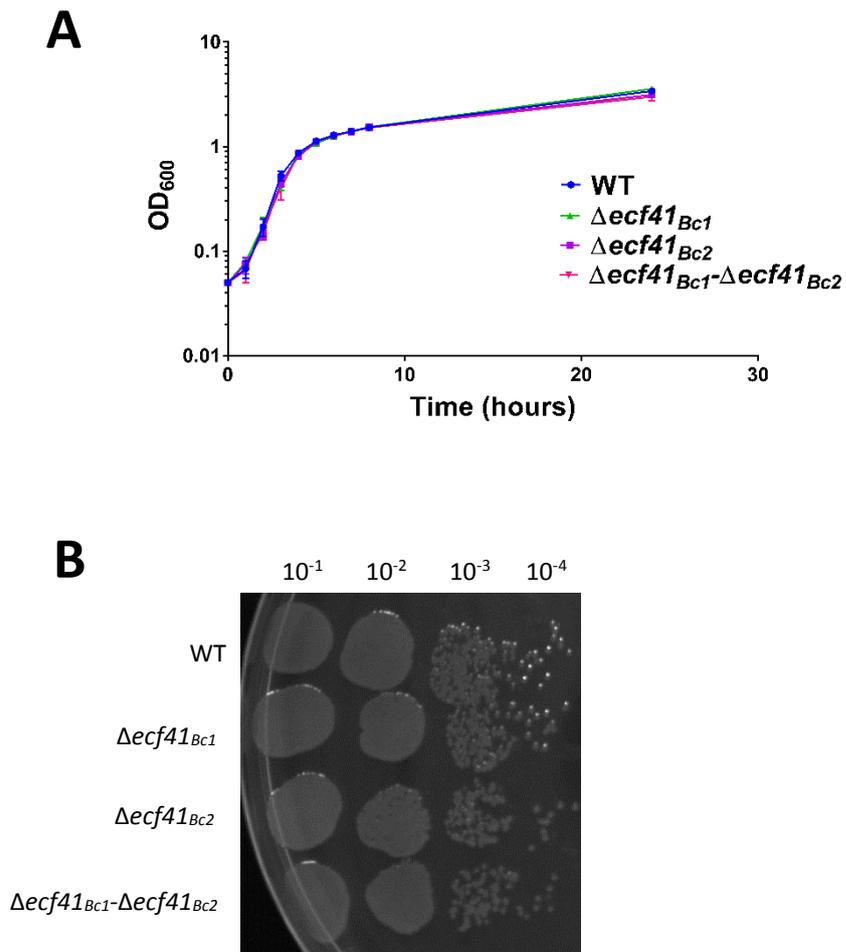


Figure 4.8. Effect of deletion of *ecf41*_{Bc1} and/or *ecf41*_{Bc2} on growth under normal conditions
 (A) Growth curves of WT (H111), H111- $\Delta ecf41_{Bc1}$, H111- $\Delta ecf41_{Bc2}$ and H111- $\Delta ecf41_{Bc1}\text{-}\Delta ecf41_{Bc2}$ in LB, n=9. (B) Representative image showing serially diluted log phase cultures of WT (H111), H111- $\Delta ecf41_{Bc1}$, H111- $\Delta ecf41_{Bc2}$ and H111- $\Delta ecf41_{Bc1}\text{-}\Delta ecf41_{Bc2}$ grown in LB at 37°C and spotted on LB agar.

4.4.5 Effect of deletion of *ecf41* genes on bacterial growth under oxidative stress conditions

The effect of deleting *ecf41* genes on the ability of *B. cenocepacia* cells to respond to oxidative stress was tested by carrying out a filter disc assay as described in section 2.8.4. In this experiment H₂O₂, methyl viologen, tertiary-butyl hydroperoxide and sodium hypochlorite were used as oxidants. It was found that deleting either or both *ecf41* genes did not affect the oxidative stress response of the bacterial cells (Figure 4.9).

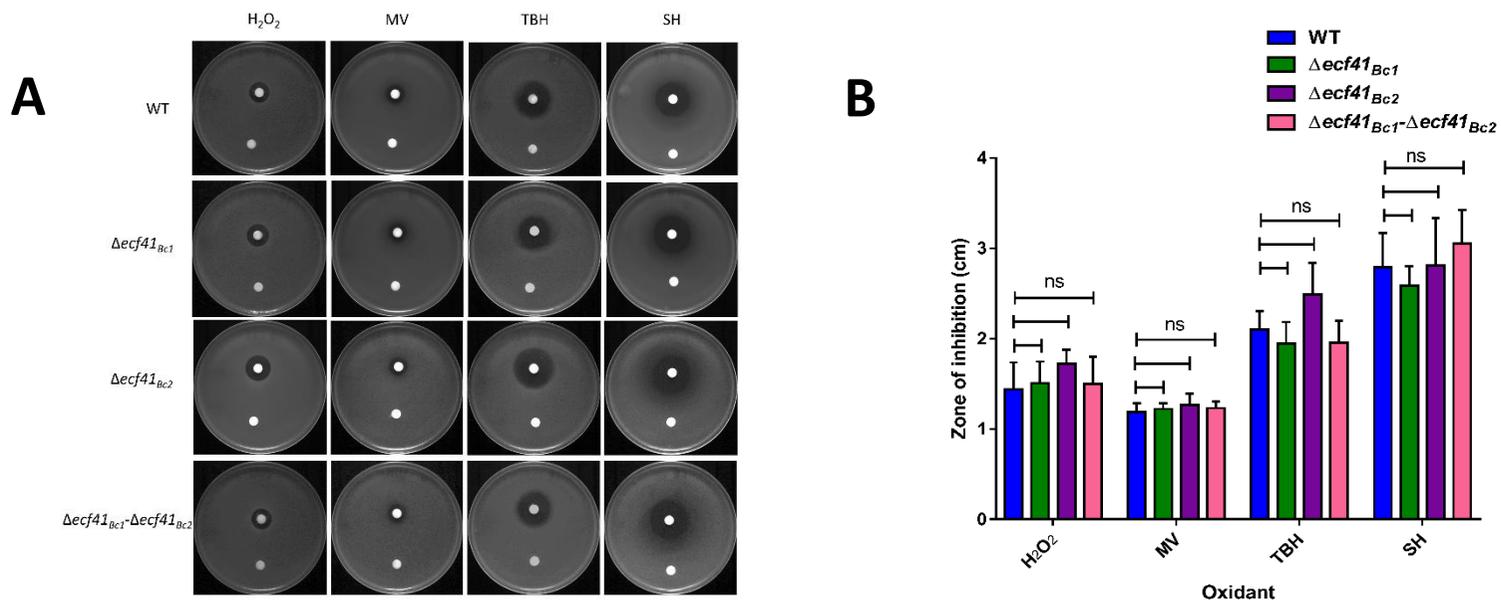


Figure 4.9. Representative images showing effect of oxidative stress due to deletion of *ecf41*_{Bc1} or *ecf41*_{Bc2} or both *ecf41*_{Bc1} and *ecf41*_{Bc2}

(A) Representative images of plates showing a zone of inhibition assay to measure sensitivity to different oxidants. In each case, top filter disc contains oxidant and bottom filter disc contains water. From top to bottom, first set of horizontal panels represent WT (H111), second set of horizontal panels represent H111- $\Delta ecf41_{Bc1}$, third set of horizontal panels represent H111- $\Delta ecf41_{Bc2}$ and last set of horizontal panels at the bottom represent H111- $\Delta ecf41_{Bc1-\Delta ecf41_{Bc2}}$ as indicated. Vertical panels from left to right represent oxidants hydrogen peroxide (H₂O₂), methyl viologen (MV), tertiary-butyl hydroperoxide (TBH) and sodium hypochlorite (SH), respectively, as indicated. (B) Quantification of the zones of inhibition due to oxidants hydrogen peroxide (H₂O₂), methyl viologen (MV), tertiary-butyl hydroperoxide (TBH) and sodium hypochlorite (SH), for WT (H111), H111- $\Delta ecf41_{Bc1}$, H111- $\Delta ecf41_{Bc2}$ and H111- $\Delta ecf41_{Bc1-\Delta ecf41_{Bc2}}$ as indicated. A one-way ANOVA was carried out, with a Tukey's post test to determine statistical significance, ns = not significant, (n = 9).

4.5 Regulatory mechanism of Ecf41_{Bc1} and Ecf41_{Bc2}

The activity of Ecf41 σ factors in *Bacillus licheniformis* and *Rhodobacter sphaeroides* are reported to be regulated through their C-terminal domain and it has been suggested that a 'DGGG' motif present in this domain plays an important role (Wecke et al 2012). Therefore, it was decided to adopt an analogous experimental design to verify if the DGGG motif within the C-terminal domain of Ecf41_{Bc1} and Ecf41_{Bc2} in *B. cenocepacia* plays a regulatory role. It was decided to test three truncated versions of the Ecf41 σ factors in this experimental design: one that contained the C-terminal extension only up to the DGGG motif, second that included the C-terminal extension up to but excluding the DGGG motif and a third derivative that did not contain the C-terminal extension. These derivatives are schematically represented in Figure 4.10.

4.5.1 Construction of plasmids encoding truncated derivatives of Ecf41_{Bc1} and Ecf41_{Bc2}

Depending on the *ecf41_{Bc1}* truncated derivative to be amplified Bcal3486.bam.213.rev, Bcal3486.bam.201.rev or Bcal3486.bam.176.rev reverse primers were used separately with the forward primer BCAL3486for to generate products of 686, 650 or 575 bp, respectively, using H111 boiled lysate as template and KOD polymerase. Similarly, depending on the *ecf41_{Bc2}* truncated derivative to be amplified bcam2748.bam.215.rev, bcam2748.bam.198.rev or bcam2748.bam.175.rev reverse primers were used separately with the forward primer BCAM2748for to generate products of 655, 610 or 541 bp, respectively, using H111 boiled lysate as template and KOD polymerase. Each of these six PCR products were digested with *HindIII* and *BamHI* and separately ligated with pBBR2 that had also been digested with the same enzymes. The ligation reactions were used to transform JM83 cells that were spread on LB agar containing kanamycin, X-gal and IPTG to take advantage of the blue/white screening facility. Plasmids with an insert formed white colonies which were PCR screened using the insert-specific primers which were used in the initial PCR amplification mentioned above. Plasmids obtained from colonies which generated the correct sized PCR product were checked for their sizes by agarose gel electrophoresis. pBBR2-Ecf41_{Bc1}-213 (including C-terminal extension with DGGG motif), pBBR2-Ecf41_{Bc1}-201 (including C-terminal

extension without DGGG motif) and pBBR2-Ecf41_{Bc1-176} (excluding C-terminal extension) were expected to be 5.800, 5.764 and 5.689 kb, respectively. pBBR2-Ecf41_{Bc2-215} (including C-term extension with DGGG motif), pBBR2-Ecf41_{Bc2-198} (including C-terminal extension without DGGG motif) and pBBR2-Ecf41_{Bc2-175} (excluding C-terminal extension) were expected to be 5.769, 5.724 and 5.655 kb, respectively. The clones with the correct plasmid size were checked for their DNA sequence integrity by nucleotide sequencing.

4.5.2 Analysis of putative Ecf41_{Bc1} and Ecf41_{Bc2} regulated promoter activity in the presence of the respective C-terminally truncated ECF σ factors in *E. coli*

pBBR2, pBBR2-Ecf41_{Bc1}, pBBR2-Ecf41_{Bc1-213}, pBBR2-Ecf41_{Bc1-201} and pBBR2-Ecf41_{Bc1-176} were each separately transferred to MC1061 cells containing pKAGd4-P_{ahpD}. Similarly, pBBR2, pBBR2-Ecf41_{Bc2}, pBBR2-Ecf41_{Bc2-215}, pBBR2-Ecf41_{Bc2-198} and pBBR2-Ecf41_{Bc1-175} were each separately transferred to MC1061 cells containing pKAGd4-P_{lysR}. The promoter activities in each of the resulting strains was measured by carrying out the β -galactosidase assay. P_{ahpD} was found to be significantly active in the presence of full-length Ecf41_{Bc1} and Ecf41_{Bc1-201} (including C-terminal extension without DGGG motif) as shown in Figure 4.10A. Interestingly, the activity of P_{ahpD} in the presence of Ecf41_{Bc1-201} was 4 times greater than in the presence of full-length Ecf41_{Bc1}. P_{ahpD} was also weakly active in the presence of Ecf41_{Bc1-213} (including C-terminal extension upto the DGGG motif). In contrast, P_{lysR} was found to be significantly active in the presence of full-length Ecf41_{Bc2} and Ecf41_{Bc2-175} (excluding C-terminal extension) (Figure 4.10B). The activity of P_{lysR} in the presence of Ecf41_{Bc2-175} was found to be comparable to that in the presence of full-length Ecf41_{Bc2} whereas in the presence of Ecf41_{Bc2-198} the activity was only approximately 10% of that in the presence of Ecf41_{Bc2}. Contrary to the results reported for ECF41 σ factors of *Bacillus licheniformis* and *Rhodobacter sphaeroides*, presence of the C-terminal extension upto the DGGG motif on the ectopically expressed σ factor rendered the promoter inactive for both Ecf41_{Bc1} and Ecf41_{Bc2} in *E. coli*.

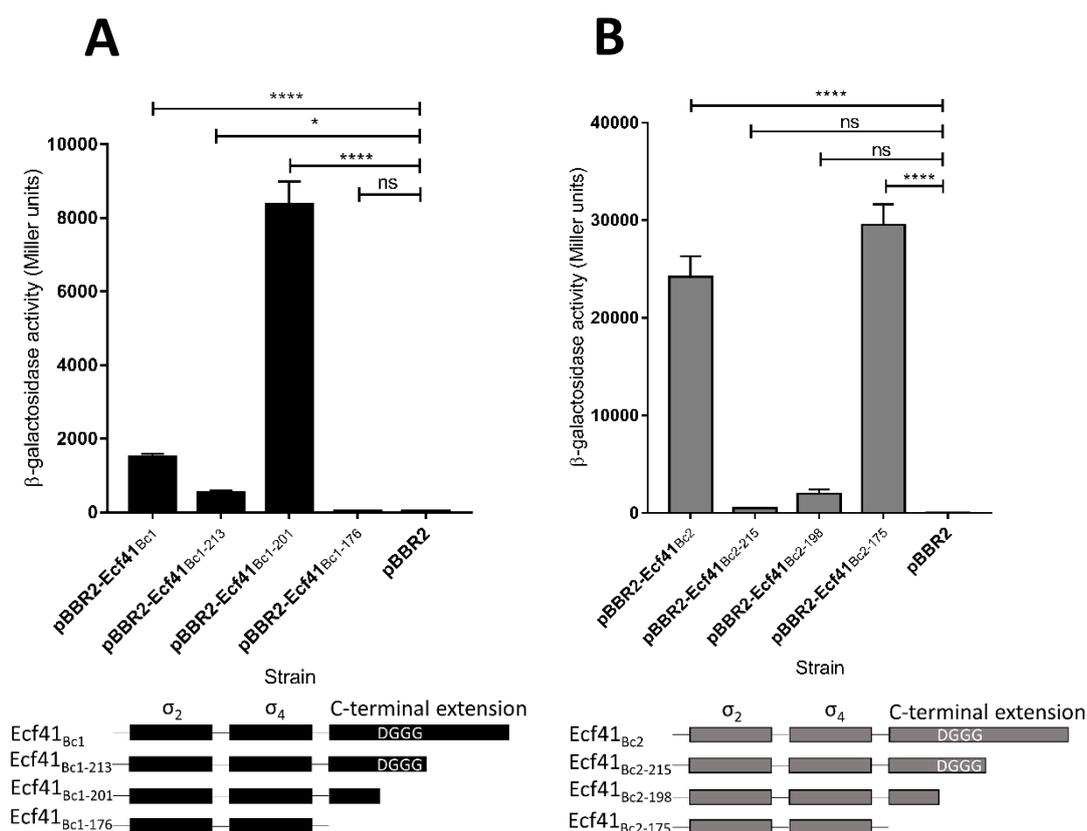


Figure 4.10. ECF41-dependent promoter activity in the presence and absence of full-length and truncated ECF41 σ factor derivatives in *E. coli*.

(A) pKAGd4 containing P_{ahpD} was transferred to MC1061 along with pBBR2-Ecf41_{Bc1}, pBBR2-Ecf41_{Bc1-213}, pBBR2-Ecf41_{Bc1-201}, pBBR2-Ecf41_{Bc1-176} or pBBR2 as indicated. The promoter activity in each of these strains was measured using the β -galactosidase assay where cultures were grown in LB containing kanamycin and chloramphenicol. Schematic representations of the full-length and truncated versions of Ecf41_{Bc1} used in this experiment are given below the plot with the appropriate names. σ_2 and σ_4 are shown as black boxes with linker sequences as black lines. The C-terminal extension is also shown as a black box with the approximate location of the DGGG motif indicated. (B) pKAGd4 containing P_{lySR} was transferred to MC1061 along with pBBR2-Ecf41_{Bc2}, pBBR2-Ecf41_{Bc2-215}, pBBR2-Ecf41_{Bc2-198}, pBBR2-Ecf41_{Bc2-175} or pBBR2 as indicated. The promoter activity in each of these strains was measured using the β -galactosidase assay where cultures were grown in LB containing kanamycin and chloramphenicol. Schematic representations of the full-length and truncated versions of Ecf41_{Bc2} used in this experiment are given below the plot with the appropriate names. σ_2 and σ_4 are shown as grey boxes with linker sequences as black lines. The C-terminal extension is also shown as a grey box with the approximate location of DGGG motif indicated. All activities shown were obtained by subtracting the activity of the same strain containing pKAGd4 assayed under the same conditions (background control). Statistical significance was determined by performing a one way ANOVA, with a Tukey's post test, **** = $p < 0.0001$, * = $p < 0.1$, (n=6).

4.5.3 Analysis of putative Ecf41_{Bc1} and Ecf41_{Bc2} regulated promoter activity in the presence of the respective C-terminally truncated ECF σ factors in *B. cenocepacia*

To verify whether the native *B. cenocepacia* Ecf41 σ factors were capable of switching on the respective putative promoters under standard laboratory conditions, initially, pKAGd4 derivatives harbouring the Ecf41_{Bc}-dependent promoters were transferred into *B. cenocepacia* H111 by conjugation using S17-1 as a donor strain and the β -galactosidase activity was measured. It was found that promoters P_{ahpD} and P_{lysR} were inactive under normal growth conditions (i.e. in LB) which suggests that these σ factors may be induced and/or activated only under certain 'stressful' conditions or under an appropriate extra-cytoplasmic signal, as is the case for many ECF σ factors (Figure 4.11). Next, pBBR2, pBBR2-Ecf41_{Bc1} and pBBR2-Ecf41_{Bc1-201} were transferred to *B. cenocepacia* H111 containing pKAGd4-P_{ahpD} as these derivatives were found cause the promoters to be highly active in *E. coli*. Also, pBBR2, pBBR2-Ecf41_{Bc2} and pBBR2-Ecf41_{Bc2-175} were transferred to *B. cenocepacia* H111 containing pKAGd4-P_{lysR} as these derivatives were found cause the promoters to be highly active in *E. coli*. The β -galactosidase activity in each of these strains was measured. Only P_{lysR} was found to be significantly active in the presence of full-length Ecf41_{Bc2} and pBBR2-Ecf41_{Bc2-175} (excluding C-terminal extension). However, unlike the situation in *E. coli* the activity of P_{lysR} in the presence of Ecf41_{Bc2-175} was found to be ten times lower than that in the presence of full-length Ecf41_{Bc2}. It should be noted that even in the presence of full-length Ecf41_{Bc2}, the activity was ten times less than when present in the *E. coli* host strain.

In the above experimental analysis, it is assumed that the promoter activities have been obtained in the presence of functional truncated derivatives of the σ factors. However, it can be speculated that the deletions introduced might cause the protein to become insoluble, degrade or quickly turn over. In such situations, the promoter activities obtained may not be a true representation of the importance of the length of the respective ECF σ factors in transcription initiation. This can be easily validated by carrying out a Western blot analysis to confirm expression and size of the respective protein.

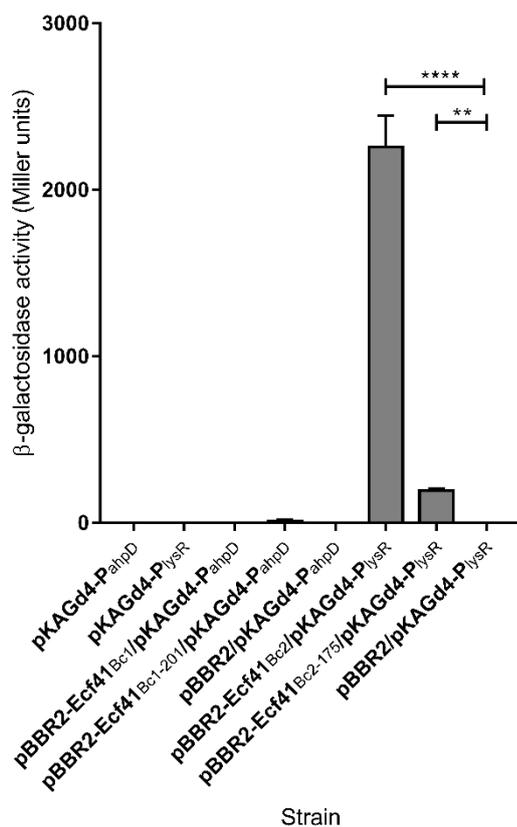


Figure 4.11. ECF41-dependent promoter activity in the presence and absence of full-length and truncated derivatives of ECF41 σ factors in *B. cenocepacia*.

pKAGd4 containing P_{ahpD} was transferred to *B. cenocepacia* H111 along with pBBR2-Ecf41_{Bc1}, pBBR2-Ecf41_{Bc1-201} or pBBR2 as indicated or without a pBBR2 derivative. pKAGd4 containing P_{lysR} was transferred to H111 along with pBBR2-Ecf41_{Bc2}, pBBR2-Ecf41_{Bc2-176} or pBBR2 as indicated. The promoter activity in each of the resulting strains was measured using the β -galactosidase assay where cultures were grown in LB containing kanamycin and chloramphenicol. All activities shown were obtained by subtracting the activity of the same strain containing pKAGd4 assayed under the same conditions (background control). Statistical significance was determined by performing a one way ANOVA, with a Tukey's post test, **** = $p < 0.0001$, (n=6).

4.6 Genomic context of ECF42_{Bc}

ECF42 σ factors appear across five different phyla including Gram-positive and Gram-negative bacteria. The genomic locus of *ecf42_{Bc}* from *B. cenocepacia* is shown in Figure 4.12 along with those of *ecf42* genes in other bacterial species. It is evident that the gene encoding this class of σ factor appears near one or more copies of a putative *phnB* type gene. However, the involvement of this gene in Ecf42 function, if any, remains undefined. In the case of *B. cenocepacia* two such genes occur upstream of *ecf42* and are predicted to be transcribed in the same direction as *ecf42*.

An alignment of the amino acid sequence of Ecf42 with homologous proteins from other species is shown in Figure 4.13. It was apparent that the sequences share a large C-terminal extension of approximately 250 amino acids. As described earlier they tend to contain a TPR (tetratricopeptide repeat) domain at the C-terminus. However, the significance of this domain and the C-terminal extension yet remains unknown.

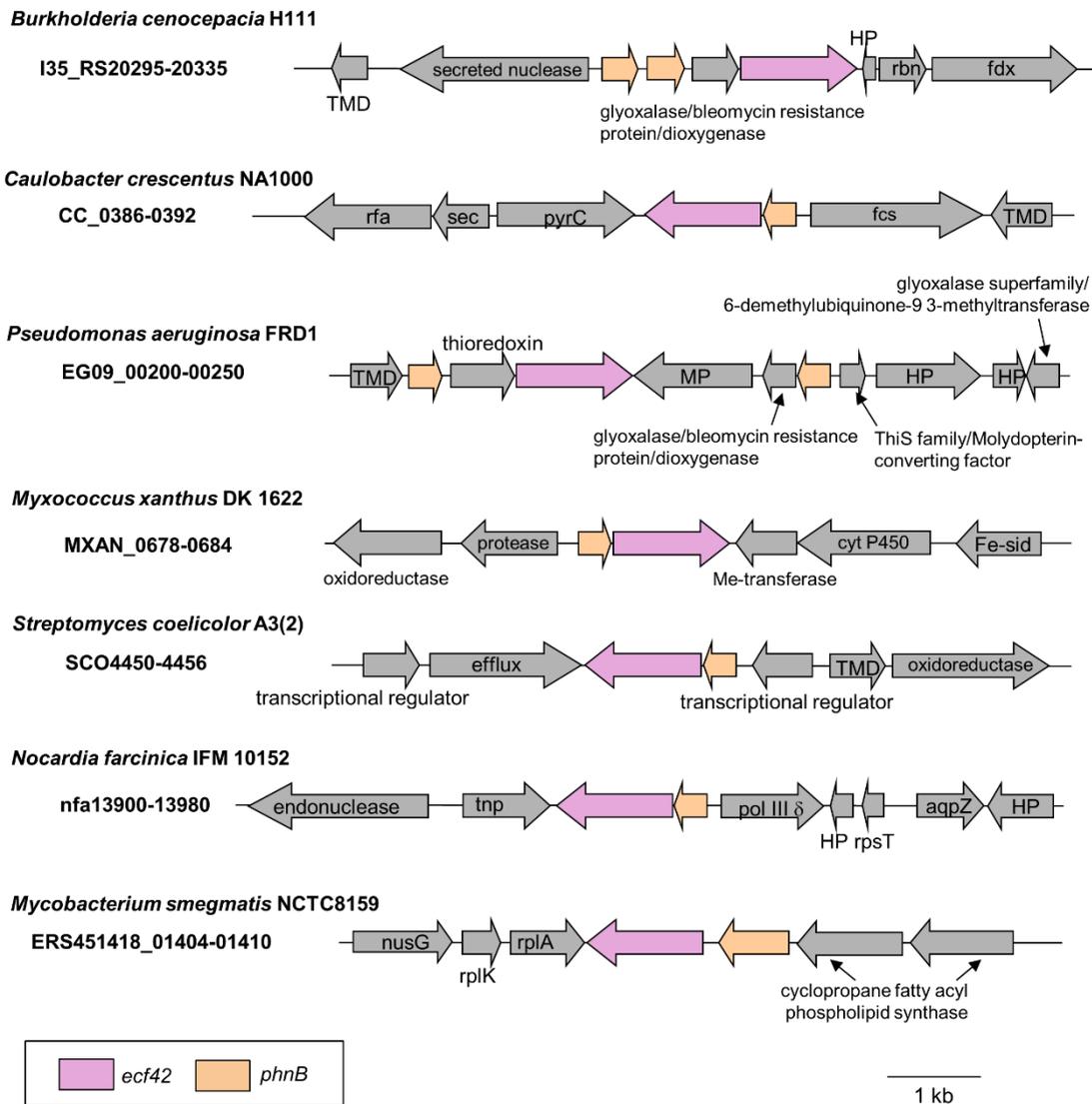


Figure 4.12. Genomic context of *ecf42* σ factor genes

Gene organisation at the *ecf42_{Bc}* locus from *B. cenocepacia* H111 compared to the corresponding locus in other species as indicated. *ecf42* genes are shown in pink while genes belonging to the *phnB* family are shown in orange. All other genes are shown in grey and their predicted gene products are shown near or on the corresponding block arrow. Gene loci numbers are shown on the left of the gene loci. HP = hypothetical protein, MP = membrane protein, TMD = Trans membrane domain protein. Excerpts taken from bioinformatic analysis by M. Thomas (unpublished).

NF 1 MTGADTAAAVAHTARTSYGRIVAVLA-AGCGDLT LAEDT LAEAFEQALRTWPRRGVFN
 CC 1 -----MSPLDQTARRDGGRIVAALA-AAFRDL LAEEAYAEACARAAAVWP-APPELD
 SC 1 -----MEALVRSLSLTPGVLT VLVRRGAEFAAEDAVQDALVEAVRGWP-ADPERD
 MS 1 -----MAELD SVFRREWGPTVATTA-RWSGDL LAEDAVQEAACHAVQAWA-DGLPND
 BC 1 MTHEA-THRAIEAVWRTEAPK I IARAA-RVVRD VGVAEELAQDTLVAALEHWVQGVFDN
 MX 1 --MMD-TSSILRAHREDYGRIVATLIRVLDGDFCAAEVVOQDAFAALNQWPRRGVRE
 .. * .. ** .. * .. * .. *

NF 60 PEGWLLTVARNRQRDEWKSAGYRRSVPLDTEP-----PPVGGVALDDL DADAI PDR
 CC 52 PAAWLYATARRCALDMFRNRVRRRATLERPE-----PQP-SAEVLTSDRLIPDE
 SC 50 PKGWLVTVAWRRELDAAARADTARRRREDRVEE-----EPVPG-----PAPVDVDD
 MS 52 PGAWLTTVARNRDLRRLRRESQRTGREYAA-----V-----YDMTARADTE LHPVRDD
 BC 59 PAAWLMTAVRRALDRVQESLHAAKRDQLGHEMDALEAHVVPDIADALADASDDDIGDD
 MX 58 PEGWLMRVARNAIDRMRRSARLEAVRGEI-----EAVAPSVDDAEWQTPDD
 * .. ** .. * .. * .. * .. *

NF 112 RLALIFVCAHPAAAEARTPLMLQTVLGFDSQVARAFVSPAAAMQORLVRAKRRIRDAR
 CC 103 RLRLIFVCCHPAWAPESRAALTLRLVCGLSVAE IARAFLVAEPAMQORITRAK K IAEAG
 SC 94 TLQLYFLCAHPSL IESAVALTLRAVGGTLTROIARAYLVPEATMAQRISRAKRTVSGVR
 MS 101 QLRLMFTCAHPALDRASQIALTLRLISGLTAAE IARALLOSETAVRORITRAK K IRKAN
 BC 119 LLRLIFTSCHPVLST DARVALTLRLIGGLTTE IARAFITPEPTAQRIVRAKRTAAAH
 MX 106 SLRLIFTCCHPALSPAQVALLRCLCGLTTEQ IARAFITVSPATIAQRIVRAQRKIRTA
 * .. * .. * .. * .. * .. * .. * .. * .. * .. * .. *

NF 172 IPFVVPGRGALAERIPVPLEATYGHALTWREG-----TRSLAGEARYLAVTLATLLAD
 CC 163 VSEFVPGQAWPERMAAVLSTIEVAVARAHEDAAGSG-PHAGYAREMIDLTSLVADLLBG
 SC 154 F-----DRPGDVATVLRVLYLVENEGYSGD-----VDLAAEAVRLTRQLAAS-VD
 MS 161 IPLRVPPAELLPERTPHVLACYSVFTEGYWS TAGPSATRDELCEGVRLLAAELCALLPD
 BC 179 VPFEVPAADARPARLASVLEVIYLVFNEGHAATAGDDWMPALCDEALRLGRVLAQLAPD
 MX 166 IPVVIPEPEALAE RTEGVLHTIYLVFSEGYAATDCGALLRVLDCEALRLARARSLLBG
 .. * .. * .. * .. * .. * .. * .. * .. * .. * .. *

NF 226 EPEAWALAALTTLVAARNS---PADTVFRPLDAQDPADWDPRLITEGEDYLRRAE RPGA
 CC 222 EPEALAFAAALRFAEARRPARIDADGAMIPLAEQDPKRWRNRALISGDMILLRRAAR-LNR
 SC 198 HPEVAGLLALMLLHHARRAARTAPDGSVLP LAEQDRGRWDTASTAEGVRILQAALA-RDR
 MS 221 DRDAHALSALLLHDSRRSTRDHTGALIPLDEQDRTRWDRGRITARGLDQLRLA---EGA
 BC 239 ESEVLGLVALMELQASRMHARVDAQGRPVLLLLQDRSRWDPLLIRRGLAALERATLGGV
 MX 226 DAEVASIRALMLLHHSRRRARVADGGGLVLLDRQDRSLWDGAQRDEGLAELGAALA-MGA
 .. * .. * .. * .. * .. * .. * .. * .. * .. * .. *

NF 283 PGRFOLEAAIHA VHC-ARLTTGDTDWQALRTLYTALMAVAPT LGSRTAHAAVLGRTLSAE
 CC 281 PGFRQLRAALHCAWALRRSLAEPPPWETLLALYDILLDAGDDPFVRLNRAVVLAEVAGPA
 SC 257 LGEFQAQAAIAALHA-DAPTAGE TDWPOIVEWYDELARITDNPVRLNRAVAVGEADGER
 MS 278 TGPYLPQAVIAA VHA-TAPGWQATDWSTICLAYDRLQQADS PVVRLNRAVALATGFRDGH
 BC 299 RGPYALQAA LAACHA-RARQAADTDWAQI VALYDALAEVAPSPVVELNRAVAVGMAFGPA
 MX 285 RGPYTVQAAIAALHA-QAPRAEDTDWAQIAALYARLVVLT PGPVVVALNHAAAVMARGPE
 * .. * .. * .. * .. * .. * .. * .. * .. * .. *

NF 342 DGLCLIDRLPPE--RERIPYHATRADLLARAGRFAEAAAAYDTAATIT-DDPVREVFIR
 CC 341 ALAEVEALDAD-RLEGIPYHAVRADLLARLGRVEEASSAYDRALEMA-PGAERLWLI
 SC 316 AGLAALAE LDGT--LPRRTA---VAAYLHERDGDVDTAARLYAEAAKA-PALAE R DHLT
 MS 337 AGLAALDEVADDPILLARSNTVAVIRADLLRRAGRTGEAIHWYGVALHHAINEPAAAFIR
 BC 358 AALELVDVLRDDPALARVHWLPSVRCDLLAKLGRADEAKLEFRRAADIT-RNERERLILL
 MX 344 HGLSLVDDLEASGRLDYHLIPARADLLRRLGRKDEAASAYRRALAMV-RTA PERRFLE
 * .. * .. * .. * .. * .. * .. * .. * .. * .. *

NF	399	ARCAAVRGSPRR--
CC	399	RRRSFVASAN----
SC	370	RQAARINSRRRGRR
MS	397	RRLAECESATAAN-
BC	417	KRAMDA-----
MX	403	ARLLETLAE-----
	

Figure 4.13. Amino acid sequence alignment of Ecf42_{Bc} from *B. cenocepacia* H111 with related proteins in other species

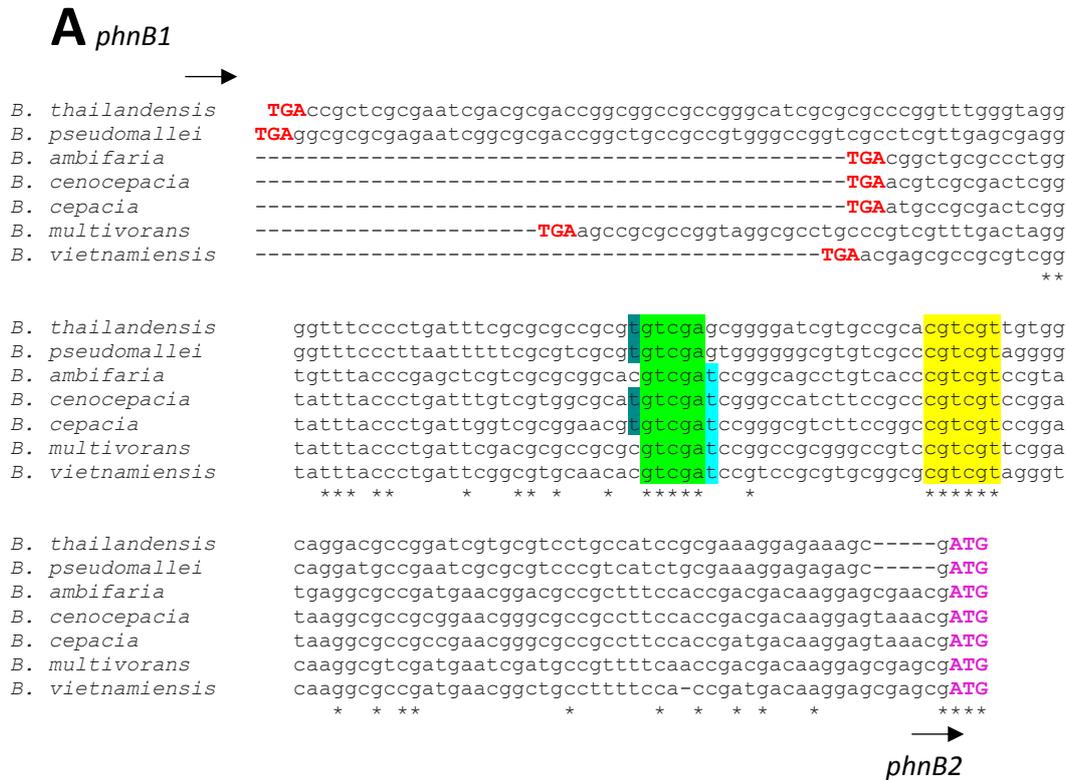
Ecf42_{Bc} amino acid sequence from *B. cenocepacia* H111 (BC) was aligned with homologous proteins from *Caulobacter crescentus* NA1000 (CC), *Myxococcus xanthus* DK 1622 (MX), *Streptomyces coelicolor* A3(2) (SC), *Nocardia farcinica* IFM 10152 (NF) and *Mycobacterium smegmatis* NCTC8159 (MS) using Clustal Omega and shaded using BOXSHADE. Amino acids which are identical at the corresponding position in ≥50% of sequences are shown in white font with black highlighting while amino acids which are similar at the corresponding position are shown in white font with grey highlighting. σ_2 , σ_4 and the C-terminal extension are underlined in red, blue, and orange, respectively. Predicted TPR domain is shown as a green box.

4.7 Investigation of the role of ECF42 by identification of target promoters

The potential role of a putative Ecf42 σ factor has been previously investigated in *Pseudomonas putida* where it was found that deleting this gene improved the ability of the bacterium to form biofilms and resist antibiotics. As ECF σ factors are normally required for stress resistance, it is likely that Ecf42 is involved in increasing the resistance to other stressors. However, the exact stress inducing signal and mode of action by Ecf42 is not yet known. In the first approach to identify the role of Ecf42 in *B. cenocepacia*, it was initially decided to identify the target promoters.

4.7.1 *In silico* identification of ECF42 regulated promoters

Since *ecf42* always tends to appear along with *phnB* genes it was hypothesised that *phnB* genes are dependent on Ecf42 for transcription initiation. Hence in *B. cenocepacia* H111, the region between I35_RS20300 and I35_RS20305 (I35_RS20305 is referred to as *phnB1* in this thesis) (Figure 4.12) was predicted to contain a promoter for *phnB1* and was aligned with corresponding regions from closely related species of the *Bcc*. Similarly the region between *phnB1* and I35_RS20310 (I35_RS20310 is referred to as *phnB2* in this thesis) was aligned with the corresponding region from other closely related species of the *Bcc*. The gene located immediately upstream of *ecf42* (I35_RS20315) encodes for a putative bleomycin resistance protein. This gene could also be potentially regulated by Ecf42_{Bc} and can be speculated to contain the *ecf42*_{Bc} promoter, as many ECF σ factors are auto-regulated. Hence, the region between *phnB2* and I35_RS20315 was also aligned as above. It was observed that a pair of conserved motifs were present in the alignment of the intergenic region between the two *phnB* genes as shown in Figure 4.14A. Here a TGTCGAT motif was found to be fairly conserved, with the central 'GTCGA' being completely conserved. Located 17 bp downstream of this motif is a highly conserved CGTCGT motif. These two motifs were hypothesised to form the -35 and -10 promoter regions of an Ecf42-dependent promoter. A 125 bp long DNA fragment termed as P_{phnB} possessing the putative -10 and -35 promoter elements is shown in Figure 4.14B. Alignments of the other intergenic regions did not show any specific conserved motifs and are therefore not shown here.



B

P_{phnB2} :
 ACGTCGCGACTCGGTATTTACCCTGATTTGTCTGGCGCA **TGTCGAT** CGGGCCATCTTC
 CGCC **CGTCGT** CCGGATAAGGCGCCGCGGAACGGGCGCCGCCTTCCACCGACGACAAGGA
 GTAAACG

Figure 4.14. Clustal Omega alignment for identification of putative Ecf42 recognition sequences

(A) Intergenic region between *I35_RS20310* (putative *phnB1*) and *I35_RS20315* (putative *phnB2*) from *B. cenocepacia* was aligned with the corresponding region from *B. cepacia*, *B. multivorans*, *B. thailandensis*, *B. pseudomallei*, *B. ambifaria* and *B. vietnamiensis*. The stop codon of *I35_RS20310* is shown in red and start codon of *I35_RS20315* is shown in magenta. The arrows represent the direction of transcription. Bases that are conserved at the corresponding position in all sequences are marked with an asterisk. Putative -10 and -35 promoter regions are highlighted in yellow and green, respectively. Within the putative -35 region two 'T' bases are highlighted in dark green and blue, respectively, since they are conserved in only some of the species. (B) DNA sequence that was hypothesised to form the PhnB promoter (P_{phnB2}) where the putative -10 and -35 regions are highlighted in yellow and green, respectively.

4.7.2 Construction of plasmids harbouring ECF42_{Bc} and putative ECF42_{Bc}-dependent promoter

The conditions required for *ecf42_{Bc}* gene expression were not known so it was decided to use the plasmid pBBR2 to allow constitutive expression of *ecf42_{Bc}* under the control of the *lac* promoter. Using primers BCAM0849For and BCAM0849Rev(full), a 1311 bp DNA fragment containing *ecf42* and native Shine-Dalgarno sequence was amplified with KOD polymerase and using a boiled H111 lysate as template, digested with *Hind*III and *Bam*HI enzymes and ligated into pBBR2 also digested with the same enzymes. The ligation reactions were used to transform *E. coli* JM83 cells and spread on LB agar containing kanamycin, X-gal and IPTG to test for blue/white screening. The transformants thus obtained that gave white colonies were screened by GoTaq PCR using the insert specific primers which were used in the initial amplification mentioned above. Transformants clones that screened positive were checked for the size of their plasmids by agarose gel electrophoresis. Plasmids with the desired insert (6.425 kb) was then checked for its DNA sequence integrity by nucleotide sequencing.

The *lacZ*-reporter plasmid pKAGd4 was used as a vector harbouring the putative Ecf42_{Bc}-dependent promoter. In order to do this, complementary oligonucleotides were annealed as described in section 2.2.8. pBCAM0847For, pBCAM0847For2, pBCAM0847Rev, and pBCAM0847Rev2 were annealed to generate a 125 bp DNA fragment containing P_{phnB2} (Figure 4.14B) which was ligated with pKAGd4 that was digested with *Hind*III and *Bam*HI enzymes. *E. coli* MC1061 cells were then transformed with the ligation reaction and the transformants were spread on LB agar containing chloramphenicol and X-gal. Blue colonies thus obtained were PCR screened using vector-specific primers AP10 and AP11 and GoTaq polymerase. Clones that seemed to contain the promoter fragment were confirmed to have the correct DNA sequence by nucleotide sequencing.

4.7.3 Analysis of putative ECF42-dependent promoter activity in the presence of ECF42 in *E. coli* and *B. cenocepacia*

pBBR2 and pBBR2-Ecf42 were separately transferred to *E. coli* MC1061 cells containing pKAGd4-P_{phnB2} and the activity of the putative promoter was measured by performing β -galactosidase assays in LB containing chloramphenicol and kanamycin along with appropriate background control strains. It was found that the putative promoter fragment was inactive in both the presence and absence of Ecf42. In order to test if the σ factor was capable of recognising the putative promoter in *B. cenocepacia* where other necessary regulatory factors maybe present, pKAGd4-P_{phnB2} was conjugally transferred to H111 via S17-1 donor strain and β -galactosidase assays were performed. However, again the promoter was found to be inactive. Then, pBBR2 and pBBR2-Ecf42 were separately transferred to *B. cenocepacia* H111 cells containing pKAGd4-P_{phnB2} and β -galactosidase assays were performed but the promoter was still found to be inactive (data not shown). It was concluded that either the full-length *ecf42* that was used in this experiment is not sufficient for promoter recognition or that the putative promoter may not be Ecf42-dependent.

4.7.4 FUZZNUC based identification of additional putative ECF42-dependent promoters

Although the putative P_{phnB} promoter was found to be inactive in the above experimental set up, it was decided to scan the genome of *B. cenocepacia* and other bacterial species to identify additional sequences that may serve as Ecf42-dependent promoters using FUZZNUC.

4.7.4.1 In silico identification of additional putative ECF42-dependent promoters on the *B. cenocepacia* genome

Using the Prosite pattern TGTCGAT(N)₁₆CGTCGT the *B. cenocepacia* H111 genome was scanned using FUZZNUC. Five novel hits were identified that could behave as putative promoters. These are presented in Table 4.1 showing the sequence, genomic

location and the gene present downstream. Interestingly, one of these promoters is located upstream of a gene encoding 3-demethylubiquinone-9, 3-methyltransferase belonging to the glyoxalase family and another promoter is also located upstream of a gene encoding a protein from the same family. Such type of a gene encoding the same product is also present upstream of *ecf42*_{BC}. Also, one of these putative promoters is located upstream of a putative PhnB encoding gene (I35_RS) not present in the vicinity of *ecf42*_{BC}. Although these promoters could be experimentally verified for their activity, the conditions for *ecf42* to execute its function were not known and hence these were not tested.

Table 4.1. Putative Ecf42-dependent promoters identified in the *B. cenocepacia* H111 genome^a

Gene locus ^b	Sequence ^c	Putative promoter of gene encoding -	Gene locus no.
1649974-1650002, chr 1	TGTCGATTTCCGTCGGCATCGTCGT GT	glyoxalase family protein	I35_RS07730
3308627-3308655, chr 1	TGTCGATCTACACCATCCTGTCCGT GT	gamma glutamyl GABA hydrolase	I35_RS15330
2135408-2135436, chr 1	TGTCGATTTCCCGCTGCGTCGCCGT GT	oxalate/formate antiporter	I35_RS09835
2560197-2560225, chr 2	TGTCGATTCGCGTCGCCGTTCGCGGT GT	PhnB	I35_RS27625
1039151-1039179, chr 2	TGTCGATTTCCGGCGGGCGTCGT GT	3-demethylubiquinone-9, 3-methyltransferase	I35_RS21145

a: Using FUZZNUC with Prosite query TGTCGAT(N)₁₆CGTCGT. b: Location of the corresponding sequence showing gene loci numbers and chromosome numbers. c: Putative -35 and -10 regions are highlighted in green and yellow, respectively.

An alignment of the region located upstream of *phnB* genes in other bacterial species revealed a conserved GTCGA(N)_{17/18}CGAC sequence that may constitute an Ecf42-dependent promoter (Figure 4.15). In some cases an additional matching sequence was divergently orientated with respect to the putative promoter sequence

located upstream of the *phnB* gene. One possibility is that this is the true recognition sequence for Ecf42 group σ factors and that in *B. cenocepacia* the *ecf42* gene could be controlled by a different ECF σ factor that recognises GTCGA(N)₁₇CGTCGT. However, using GTCGA(N)₁₇CGAC Prosite pattern (17 bp spacer was observed in all species except *Ralstonia solanacearum*) to scan for matches on the *B. cenocepacia* H111 genome, 147 hits were obtained on chromosome 1, 149 on chromosome 2 and 58 on chromosome 3. Therefore, it was concluded that this might be a commonly occurring motif within the genome making it difficult to identify putative promoters.

```

PA2      --agcgcgcgaccgt--tcgtccgcgcgcgggctttttt-----ccccgccactgtcg
PA1      -----ttct--gcgcaa--tgacggcccgttctggcccgcctcccgaactgtcg
PA1rev   -----cctc--ctatac--tgccccggccagcgaaaaaacccgcgctcgctgtcg
BC       --aacgtcgcgact--cgggat--ttaccctgatttgt-----cgtggcgcgatgtcg
LS       -----acgact--cccgat--tcccctttcccgattcccggcctcaaccgactgtcg
PB       ----cacatcaat---cggttgaaaaatctcc---ccgaaatcgccggggaggctgtcg
RSrev    -----ccccgcggttacacctgtttacatcttaggcggacggggatgtcg
RS       -gagtgagtgggagccgcgagagggttcgctctctacggg-----tacgtcg
MX       cggggcgggtaac---cgaaaagaaatcccccgtaaaa-----ggcgtcg
MXrev    ----cacgggtgag---cgggagaagagcctccctgcttc-----cgtgtcg
NP       ----cattggcttgccctcgttgttttctttgtgag-----taagtgtcg
RL       ----gcatttcgacagcggcgaaat---tttttgtgacgaacgaa-----gtgtcg

```

```

PA2      aaaggc-gccctcgccgggcgactaccgaacgaacgg-----atcgtcgatccgcacca
PA1      aacggc-gggccggcggatcgacaggcgacggcggttttttcgctggccggggcagtata
PA1rev   atccgc-cggccccgcggttcgaccagtctg-----ggaggcggggccagaacg
BC       atcggg-ccatcttccgcccgctcgtccgga-----taaggcgcgcggaacg
LS       attccc-gcccgcgccattcgtcgcttga-----tgaagccagcccgaccg
PB       atttcg-cgcctcctcgttcgacgtgatggtag-----agcagccggtcgaaccg
RSrev    attcggccaggggtgtcattcgacgtaccctgtag-----agacgcgaacctctcgc
RS       aatgacaccctggccgaatcgacatccccgtcc-----gcctaagatgt-aacaa
MX       attcac-ggaggtttcgttcgaccacggaagca-----gggaggctctt-ctccc
MXrev    aacgaa-acctccgtgaatcgaccgccttttca-----cgggggatttc-ttttc
NP       aaccgc-aattgctcaaatcgacacctgcgcaaaa--aatttttttcattagatacaatcg
RL       aaa-tcgcaaggcagcactcgacaaggcttcat-----cgaaacctgctcgaggag

```

* ** *

```

PA2      ggagcca-----
PA1      ggaggcccgcgcca---
PA1rev   ggccgtcattgcgcgaaa
BC       ggcgccgccttccaccga
LS       ggaccgcgcgggacc--
PB       ggcgaaaacgag-----
RSrev    ggctcccactcactc---
RS       ggtgtaacggcggg---
MX       gctcaccctg-----
MXrev    ggttaccgcggcccg---
NP       agctttac----ctg---
RL       aactgaaa----tg----

```

Figure 4.15. Clustal Omega alignment for identification of putative P_{phnB} promoter

Intergenic region between *I35_RS20310* (putative *phnB1*) and *I35_RS20315* (putative *phnB2*) from *B. cenocepacia* H111 was aligned with the regions from *P. aeruginosa* FRD1 (**PA1** = EG09_00230 - 00235), (**PA1rev** = EG09_00230 - 00235), (**PA2** = EG09_00205 - 00210), *Myxococcus xanthus* DK1622 (**MX** = MXAN_0679 - 0680), (**MXrev** = MXAN_0680 - 0679), *Ralstonia solanacearum* CMR15 (**RS** = CMR15_mp10589 - mp10590), (**RSrev** = CMR15_mp10590 - mp10589), *Lysobacter capsici* 55 (**LS** = LC55x_0258 - 0259), *Nostoc punctiforme* PCC 73102 (**NP** = Npun_F2132 - R2131), *Rhizobium leguminosarum* 3841 (**RL** = RL1476 - RL1477) and *Paludisphaera borealis* PX4 (**PB** = BSF38_03031 - 3032). Only bases relevant to the alignment are shown. Bases that are conserved at the corresponding position in all sequences are marked with an asterisk. Putative -10 and -35 promoter regions are highlighted in blue and green, respectively. Within the putative -10 region one 'T' base is highlighted in yellow since it is conserved only in two species.

4.7.4 *In silico* identification of putative ECF42-dependent promoters in the genome of other bacterial species

A slightly less stringent Prosite pattern was adopted for scanning the genomes of other bacterial species – TGTCGA(N)₁₇CGTC. Here, the last T base of the putative -35 region was replaced by N, thus making the spacer 17 bp long and only the first four bases ‘CGTC’ were retained within the -10 region.

In the Proteobacterium occupying a similar niche as *B. cenocepacia*, *Pseudomonas aeruginosa* PAO1, this scan identified 7 hits. One of them was found to be located within a probable enoyl-CoA hydratase encoding gene, and upstream of a gene encoding a probable transcriptional regulator. All other sequences were either located distantly from the start codon of a gene (≥ 1000 bp) or not present upstream of the start codon of a gene and therefore unlikely to be promoters. A FUZZNUC scan of the *Streptomyces coelicolor* A3(2) genome gave 9 hits with this pattern. Of these, only 3 seemed to be putative promoters based on their location. One was found to be present upstream of a gene encoding a putative TetR family transcriptional regulatory protein located near an *ecf42* gene. The second hit was found upstream of *phnB* type gene not present near an *ecf42* type gene and the third hit was found to be located upstream of a gene encoding a putative integral membrane ATPase. In the *Myxococcus xanthus* DK 1622 genome only 3 hits were obtained. Only one of these seemed likely to constitute a promoter and it was located upstream of a *phnB* type gene, but again, not present near an *ecf42* type gene. However, as the promoter could not be experimentally verified it cannot be established that the hits obtained in other species actually constitute promoters or not.

4.8 Genetic analysis of the role of Ecf42_{Bc}

In *B. cenocepacia* a rubrerythrin encoding gene is present downstream of *ecf42*. Rubrerythrin is reported to be involved in protection against oxidative stress. If this gene forms part of the *ecf42* operon it may suggest that Ecf42 might be involved in oxidative stress response. To test this proposition it was decided to construct an *ecf42* in-frame deletion mutant.

4.8.1 Construction of an *ecf42_{Bc}* in-frame deletion mutant

SOE PCR primers TPRdelA and TPRdelB were used to amplify a DNA fragment 571 bp long (AB) which contained the region upstream of *ecf42_{Bc}* and the first six codons of *ecf42_{Bc}* using H111 boiled lysate and Q5 polymerase. Similarly, using primers TPRdelC and TPRdelD a 563 bp DNA fragment (CD) was amplified which contained the region downstream of *ecf42_{Bc}* and the last seven codons of *ecf42_{Bc}*. Using AB and CD PCR products in an equimolar concentration as the DNA template, and primers TPRdelA and TPRdelD, a third PCR was carried out to generate a product (AD) which was 1134 bp long. This product had lost 410 codons of *ecf42_{Bc}* (total 423 codons) as primers TPRdelB and TPRdelC were designed such that they had homologous bases. Then using enzymes *Hind*III and *Xba*I, restriction digestion of the product AD was carried out and subsequently ligated with pEX18-TpTer-Scel-*pheS* that had been digested with the same enzymes. JM83 cells were transformed with the ligation reactions and spread on LB medium containing trimethoprim, X-gal and IPTG. White colonies were PCR screened using GoTaq polymerase and insert-specific primers TPRdelA and TPRdelD. Plasmid clones that seemed to contain the appropriate insert were then checked for their size and DNA sequence integrity. One such construct, pEX18TpTer-Scel-*pheS*-*Ecf42_{Bc}* was transferred conjugally transferred to *B. cenocepacia* H111 using *E. coli* S17-1 as the donor strain. This allowed the first recombination event to occur which integrated the plasmid into the *B. cenocepacia* genome at the *ecf42* locus thus giving plasmid co-integrates. The formation of co-integrates was confirmed by carrying out a PCR using vector-specific primers pEX18Tpfor and pEX18Tprev. Following that, pDAI-Scel was conjugally transferred to a co-integrate strain using *E. coli* SM10 as the donor. The I-Scel enzyme thus expressed generates a cut at the I-Scel site in the integrated vector allowing selection for another recombination event which rescues the strain and can give rise to either WT or a *ecf42_{Bc}* deletion mutant. The presence of mutants among tetracycline-resistant survivors was checked by carrying out a PCR using primers 0849outfor2 and 0849rev2 that anneal to genomic sequences outside the region complementary to primers A and D. For the WT used this would generate a product of 2.490 kb while Δ *ecf42_{Bc}* would be expected to give rise to a product of 1.260 kb since it had lost 410

codons of *ecf42_{Bc}*. Figure 4.16 shows the result of a PCR screen of one such mutant compared to the WT. The H111 Δ *ecf42_{Bc}* mutant was streaked 3 times on 0.1% chlorophenylalanine medium to allow for loss of pDAI-ISceI. A general schematic outline of this procedure is shown in Figure 3.29.

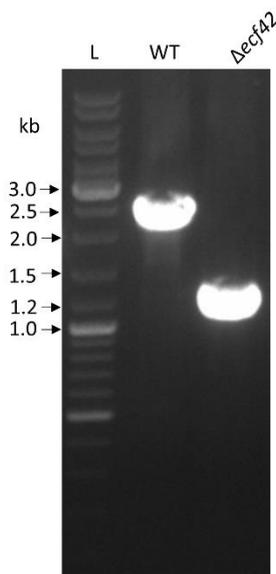


Figure 4.16. PCR screening of a candidate H111 Δ *ecf42_{Bc}* mutant

Agarose gel showing PCR screen of candidate H111 Δ *ecf42_{Bc}* using primers 0849outfor2 and 0849outrev2 which anneal at genomic locations lying outside the expected deleted region of *ecf42* giving a PCR product of 2.490 kb for WT and 1.260 kb for Δ *ecf42*; L = GeneRuler DNA mix; WT = H111 boiled lysate used as template.

4.8.2 Effect of *ecf42_{Bc}* deletion on growth of *B. cenocepacia* in standard lab conditions

To assess whether *ecf42_{Bc}* was required by cells growing under standard laboratory conditions, liquid cultures were grown in LB and the optical density was measured as described in section 2.1.4. As shown in Figure 4.17A, it was evident that Δ *ecf42_{Bc}* grew at a similar rate to that of WT and attained the same final cell density in this medium. Also, the efficiency of colony formation was tested on LB agar. As shown in Figure 4.17B, Δ *ecf42_{Bc}* formed colonies with a similar efficiency as that of WT. It was

concluded that deletion of *ecf42_{Bc}* did not affect bacterial growth under normal conditions.

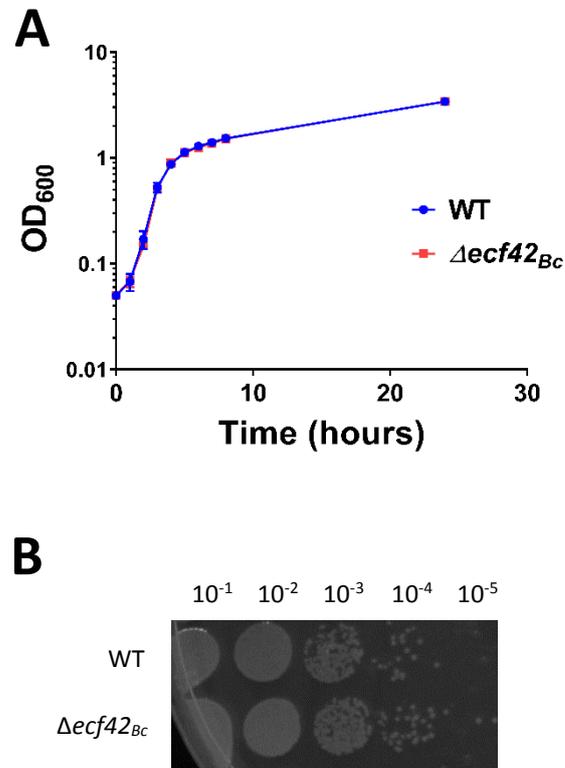


Figure 4.17. Effect of deletion of *ecf42_{Bc}* on growth of *B. cenocepacia* under normal conditions (A) Growth curves of WT (H111) and H111- $\Delta ecf42_{Bc}$ in LB, n=9. (B) Representative image showing serially diluted log phase cultures of WT (H111) and H111- $\Delta ecf42_{Bc}$ grown in LB at 37°C and spotted on LB agar.

4.8.3 Effect of *ecf42_{Bc}* deletion on growth of *B. cenocepacia* under oxidative stress

The effect of deleting *ecf42_{Bc}* on the ability of bacterial cells to respond to oxidative stress was tested by carrying out a filter disc assay as described in section 2.8.4. In this experiment hydrogen peroxide, methyl viologen, tert-butyl hydroperoxide and sodium hypochlorite were used as oxidants. It was found that deleting *ecf42_{Bc}* did not affect the response of *B. cenocepacia* to any of these stressors (Figure 4.18).

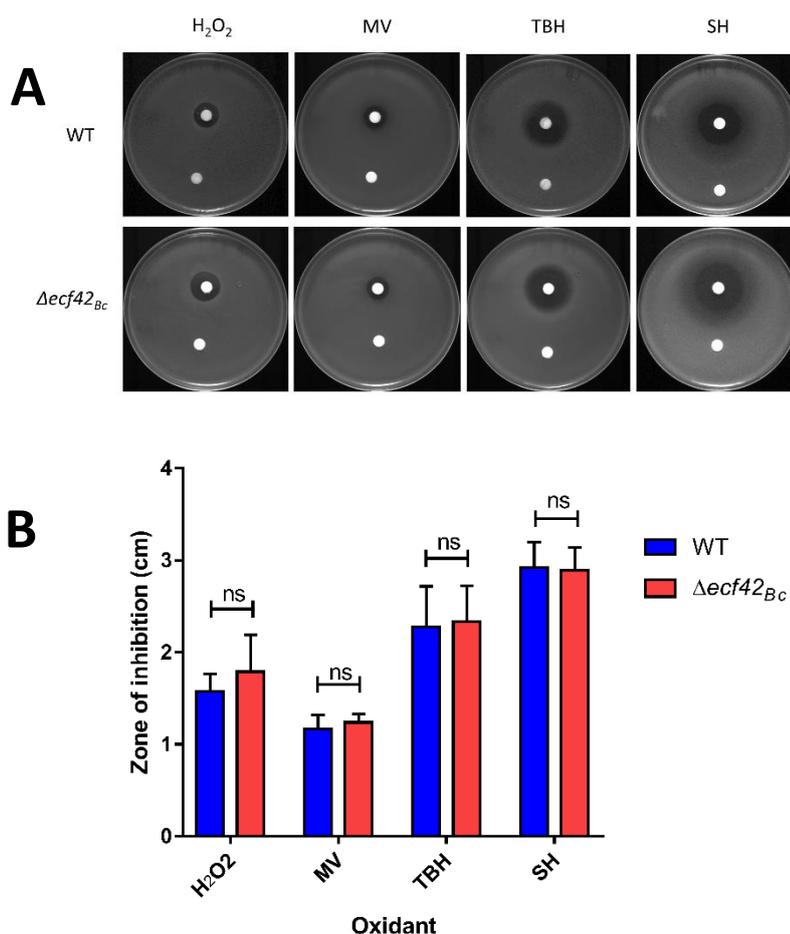


Figure 4.18. Effect of deletion of *ecf42_{Bc}* on the oxidative stress response of *B. cenocepacia* (A) Representative images of plates showing a zone of inhibition assay to measure sensitivity to different oxidants. In each case top filter disc contains oxidant and bottom filter disc contains water. The top row of panels represent WT (H111), bottom row of panels represent H111-Δ*ecf42_{Bc}* exposed to oxidants hydrogen peroxide (H₂O₂), methyl viologen (MV), tert-butyl hydroperoxide (TBH) and sodium hypochlorite (SH) as indicated. (B) Quantification of the zones of inhibition for WT (H111) and H111-Δ*ecf42_{Bc}*. A student's t-test was carried out to determine statistical significance, ns = not significant, (n = 9).

4.9 Investigation of the role of the Ecf42_{Bc} C-terminal domain in the regulation of Ecf42_{Bc}

It was hypothesised that the C-terminal extension of Ecf42_{Bc} in *B. cenocepacia* might be an auto-inhibitory domain based on other examples of auto-regulatory ECF σ factors in current literature (Wecke et al., 2012). To test this, the effect of deleting the C-terminal extension on the ability of Ecf42_{Bc} to utilise P_{phnB} was assessed.

4.9.1 Construction of a plasmid expressing C-terminally truncated Ecf42_{Bc}

A boiled H111 lysate was used as template to amplify a DNA fragment of 567 bp containing region upstream of its start codon as well as the first 175 codons of *ecf42_{Bc}* including the translation initiation region using primers BCAM0849For and BCAM0849w/oCTDRev and using Q5 polymerase. It was digested with *Hind*III and *Bam*HI and ligated into pBBR2 which had been digested with the same enzymes. The ligation reaction was used to transform *E. coli* MC1061 cells and the resulting colonies were screened using the insert-specific primers mentioned above. Colonies containing plasmids with the desired insert (pBBR2-Ecf42_{Bc-175}) were then checked for the size of the plasmid (5.681 kb) and then their sequence integrity was confirmed by nucleotide sequencing.

4.9.2 Analysis of putative Ecf42_{Bc}-dependent promoter activity in the presence of the C-terminally truncated Ecf42_{Bc} in *E. coli* and *B. cenocepacia*

pBBR2 and pBBR2-Ecf42_{Bc-175} were separately transferred to *E. coli* MC1061 cells containing pKAGd4-P_{phnB2} and the activity of the promoter was measured by performing β -galactosidase assays on cells grown in LB containing chloramphenicol and kanamycin along with appropriate background control strains. It was found that the putative promoter was inactive in both the presence and absence of Ecf42_{Bc-175}. In order to test whether the promoter might be active in *B. cenocepacia*, pKAGd4-P_{phnB2} was conjugally transferred to H111 via S17-1 donor strain. Then, through a second conjugation step, using *E. coli* S17-1 donor strain, pBBR2 and pBBR2-Ecf42_{Bc-175} were separately

transferred to *B. cenocepacia* H111 cells containing pKAGd4-P_{phnB2}. However, β -galactosidase measurements showed that the promoter remained inactive in the C-terminally truncated Ecf42_{Bc} derivative (data not shown).

4.10 Discussion

This chapter describes the attempt to identify the roles of three uncharacterised ECF σ factors of *B. cenocepacia* which belong to the ECF41 and ECF42 groups. The proposed ECF41-dependent promoters were experimentally verified in the case of Ecf41_{Bc1} and Ecf41_{Bc2} where the promoter of a putative alkylhydroperoxidase D encoding gene was Ecf41_{Bc1} dependent and the promoter of a putative LysR-type regulator encoding gene was Ecf41_{Bc2} dependent. It was also observed that these σ factors did not seem to play a role in oxidative stress response. Additionally, it was also concluded that the conserved DGGG motif in the C-terminal extension of these σ factors might play an inhibitory role in *B. cenocepacia* as opposed to what has been previously observed in *B. licheniformis* and *R. sphaeroides*.

Although a putative Ecf42_{Bc}-dependent promoter has been identified by a bioinformatics analysis, it could not be experimentally verified. In addition, the role and the regulatory mechanism of Ecf42_{Bc} although explored, could not be ascertained.

4.10.1 Ecf41_{Bc1} and Ecf41_{Bc2}-dependent promoters

It was experimentally shown that Ecf41_{Bc1} drives the expression of a putative alkylhydroperoxidase D gene and probably the *ecf41_{Bc1}* gene. Ecf41_{Bc2} was shown to drive a putative LysR-type transcriptional regulator. The family of LysR-type transcriptional regulators is relatively large where members behave as activators involved in myriad processes including virulence, motility, metabolism and oxidative stress amongst others (Maddocks and Oyston, 2008, Santiago et al., 2015). In *B. cenocepacia*, a LysR-type regulator (ShvR) has been reported to regulate expression of multiple genes involved in virulence (O'Grady et al., 2011). If the Ecf41_{Bc2}-dependent LysR-type transcriptional regulator has a similar role, Ecf41_{Bc2} might be an important

player in the pathogenicity of this bacterium which could be tested in various infection models. An *in silico* scan using the Prosite pattern TGTCACA(N)₁₇CGTC revealed another putative promoter for a gene encoding a LysR-type transcriptional regulator. This putative promoter can be tested to be Ecf41_{Bc}-dependent by carrying out reporter-fusion assays in the future.

Although the putative -35 and -10 regions have been predicted, it would be useful to carry out a systematic single base substitution analysis to experimentally verify these predictions and assess the role of the importance of each base. Also, the studying the role of the bases within the spacer region would provide evidence of how these σ factors make contacts with the promoter. Multiple σ factors within the same bacterium have been reported to be involved in crosstalk directly or indirectly (Schulz et al., 2015). Since Ecf41_{Bc2} could recognise the Ecf41_{Bc1}-dependent promoter and *vice versa*, this may suggest that these two σ factors may be inter-related in carrying out their function.

4.10.2 Role of Ecf41_{Bc1} and Ecf41_{Bc2}

Deletion of *ecf41_{Bc1}*, *ecf41_{Bc2}* individually or together did not affect the ability of the bacterium to respond to oxidative stress by hydrogen peroxide, methyl viologen, sodium hypochlorite or tert-butyl hydroperoxide. These results are in accordance with those previously obtained for *B. licheniformis* and *R. sphaeroides* but in contrast to those obtained for SigJ in *M. tuberculosis* (Hu et al., 2004, Wecke et al., 2012). It might be possible that although these σ factors have a role in responding to oxidative stress, the conditions required to activate them, for example, under specific anaerobic or growth conditions cannot be replicated in the lab. Although *ecf41* mutants have been previously tested for a phenotype under 960 different conditions, their role remains unknown (Wecke et al., 2012). Since ECF σ factors are often virulence determinants in pathogens that assist bacteria in pathogenicity and survival, it could be useful to study the effects of *ecf41* deletion on the ability of the bacterium to infect hosts such as zebrafish or *Galleria* larvae (Seed and Dennis, 2008, Deng et al., 2009).

4.10.3 Regulation of Ecf41Bc activity

The effect of deleting specific regions of the C-terminal extension revealed interesting patterns in Ecf41 activity. Firstly, for both Ecf41_{Bc1} and Ecf41_{Bc2}, deleting the C-terminal extension up to the DGGG motif rendered the σ factors weakly active suggesting that this motif may play an inhibitory role. This observation is contrary to that observed for *B. licheniformis* and *R. sphaeroides* where C-terminally truncated derivatives of Ecf41 where the DGGG motif was retained were highly active. Secondly, contrasting results were obtained for Ecf41_{Bc1} and Ecf41_{Bc2} when the C-terminal domain was completely removed or where the C-terminal endpoint occurred just before the DGGG motif. In the latter case Ecf41_{Bc1} was observed to be highly active (up to 4 times of that of the native σ factor) in contrast to Ecf41_{Bc2} that was completely inactive. On the other hand, completely deleting the C-terminal extension caused Ecf41_{Bc1} to be inactive while Ecf41_{Bc2} was as active as with the full C-terminal extension. Together, these contrasting results make it difficult to propose a general model of regulation for Ecf41 σ factors in *B. cenocepacia*. T. Wecke suggests two plausible mechanisms of regulation – one where the C-terminal extension is normally auto-inhibitory, where a conformational change allows the removal of this domain and activates the σ factor, or where the σ factors form inactive dimers through this C-terminal extension, and when needed form active monomers (Wecke et al., 2012). These hypotheses could be tested by carrying out bacterial two-hybrid assays. Other regulatory proteins that interact with the C-terminal extension might also be present. Some possible models of ECF41 regulation are illustrated in Figure 4.19.

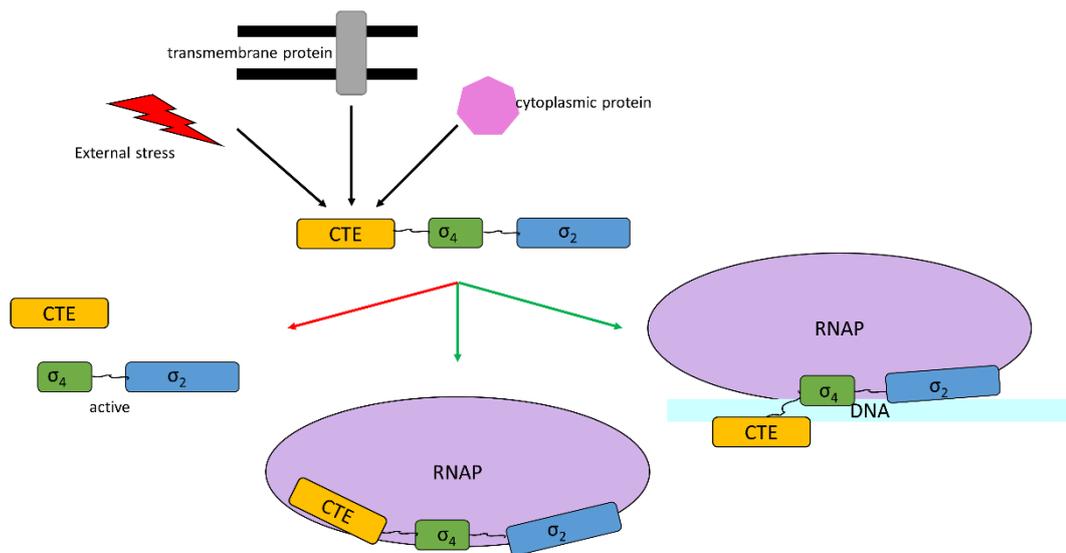


Figure 4.19 Some proposed models of ECF41 regulation by the C-terminal extension (CTE)

An external stress signal (red), or a transmembrane protein or cytoplasmic protein might pass a signal to the C-terminal extension (yellow) that might in turn be cleaved off to render the σ factor active, or might be retained for efficient binding with the core RNAP (purple) or might be required for recognition of additional promoter elements on the DNA (cyan). The red arrow indicates that the extension is inhibitory while the green arrow represents cases where the C-terminal extension is required for effective σ factor activity.

In *B. cenocepacia* the promoters were found to be inactive in the absence of the respective ectopically expressed Ecf41, presumably because the appropriate stress signals were not present. However, even in the presence of the Ecf41_{Bc1} σ factor the respective promoter was still inactive, but the Ecf41_{Bc2}-dependent promoter was active in the presence of Ecf41_{Bc2}.

4.10.4 Role and regulation of Ecf42_{Bc}

A conserved DNA sequence motif, TGTCGAT(N)₁₆CGTCGT, was identified that looked to be a likely candidate for an Ecf42_{Bc}-dependent promoter. However, this could not be experimentally verified in *E. coli* or *B. cenocepacia* as the putative promoter was inactive in the presence of full-length as well as the truncated version of Ecf42_{Bc} that

lacked the C-terminal extension. Either the identified motif did not constitute the promoter or the σ factor requires another regulatory protein that might work interact with the TPR domain to relieve its possible inhibitory function. Since the TPR domain is known to be involved in forming protein-protein interactions, they might form complexes within themselves or with the PhnB-type protein to carry out their function. This interaction could be tested by carrying out pull-down or bacterial two-hybrid assays. However, deletion of the C-terminal extension did not result in Ecf42_{BC}-dependent activity at the putative promoter. Therefore, it is possible that the transcription of *ecf42_{BC}* and the upstream genes is under the control of an alternative ECF σ factor that recognises this sequence. It can also be speculated that Ecf42_{BC} and Ecf42_{BC-175} were not expressed by pBBR2. This could be tested by carrying out a western blot where the σ factor is epitope-tagged.

An in silico scan of the *B. cenocepacia* genome for the presence of this promoter motif revealed five novel hits. If the identified motif does indeed constitute a promoter and if the conditions required to induce promoter activity can be obtained it would be useful to test these novel motifs to determine if they are Ecf42_{BC}-dependent.

Ecf42_{BC} does not seem to have a role in oxidative stress response based on the response of an *ecf42_{BC}* mutant to the oxidants hydrogen peroxide, tert-butyl hydroperoxide, sodium hypochlorite and methyl viologen. The effect of deletion of *ecf42_{BC}* on biofilm formation and drug resistance could not be tested due to time constraints. However, in future, even if such an association is found, it would be important to then ascertain the exact role of Ecf42_{BC} in this pathway as well as the signal that activates *ecf42_{BC}* and ultimately if it affects virulence, since *B. cenocepacia* is an opportunistic pathogen.

Chapter 5. Investigation of the iron starvation ECF σ factor, OrbS

5.1 Introduction

Iron acquisition is an important feature of pathogenic bacteria especially within host tissues in addition to the environment. The ECF σ factor OrbS in *B. cenocepacia* regulates the siderophore (ornibactin) mediated iron capture mechanism. The components and current understanding of this system are explained in detail in section 1.6.3.4. While the role of this σ factor is well studied, some research questions still remain unanswered of which some were addressed in this chapter.

The role of OrbS in iron acquisition through regulating biosynthesis of siderophore ornibactin in *B. cenocepacia* has been previously reported. The OrbS dependent promoters P_{orbH} , P_{orbE} and P_{orbI} were identified and the genes forming transcriptional units within the OrbS operon were also identified using RT-PCR and RNase protection assays (Agnoli et al., 2006). P_{orbH} , P_{orbE} and P_{orbI} were found to be regulated by OrbS in response to iron while P_{orbS} was found to be dependent on the housekeeping σ factor for transcription initiation (Agnoli et al., 2006). The minimal required length of the promoter and transcription start sites were determined for two OrbS-dependent promoters and P_{orbS} (K. Agnoli, 2007).

Additionally a systematic study was carried out by making single base substitutions in P_{orbH} to study the contribution of each individual base to promoter recognition by OrbS and thereby to identify the -10 and -35 promoter regions. Moreover, the length and contribution of the spacer region were also studied (K. Agnoli, 2007). It was also found that there was an unexpected contribution of the bases around the transcription start site (+1 position). These experiments were carried out in the *E. coli* host strain, since assays carried out in *B. cenocepacia* are more time consuming and hazardous. Here OrbS formed a holoenzyme with the native *E. coli* RNA polymerase core enzyme rather than the *B. cenocepacia* RNA polymerase which may have given the unexpected promoter activities when bases near the transcription start site were substituted. These results suggested that 'TAAA' present between positions -33 to -30 with respect to the transcription start site formed the -35 promoter region. Bases 'CGTC' present between positions -12 to -9 with respect to the transcription start site corresponded to the -10 promoter region. Since strong effects on promoter utilisation

were also observed near the transcription start site which are not associated with σ factors it was necessary to carry out these experiments where OrbS interacts with the native *B. cenocepacia* core RNA polymerase.

OrbS is similar to PvdS (33% similarity) although OrbS possesses an N-terminal extension of 29 amino acids which was found to be important for OrbS to retain its functionality (K. Agnoli, 2007). Moreover, the sequence of the core promoter elements that are recognised by PvdS are similar to those located at OrbS-dependent promoter. Therefore, the functionality of PvdS with OrbS dependent promoters and *vice versa* was also tested. While PvdS could recognise OrbS dependent promoters, it was found that OrbS could not recognise PvdS dependent promoters (K. Agnoli, 2007).

Based on the ideal profile of DNA base requirements for effective promoter utilisation by OrbS, an *in silico* analysis for identification of novel OrbS-dependent promoters was carried out that prompted seven candidate promoters (K. Agnoli, 2007). Their functionality however was not tested.

P_{orbS} was shown to be σ^{70} -dependent and regulated by Fur. A Fur box was identified which had a 13 bps match out of the conventional 19 bp recognition site (Agnoli et al., 2006). Since there are no candidate anti- σ factors, the presence of an alternative post-translational mechanism for iron regulation of OrbS activity was explored. An *E. coli fur* mutant was used to measure OrbS activity at P_{orbH} and this promoter was found to be regulated even in the absence of *fur* (K. Agnoli, 2007). Since then a *B. cenocepacia fur* mutant has been constructed (Han and Thomas, 2011, unpublished) but it has not yet been used to test whether a Fur-independent mechanisms exists for iron-dependent regulation of OrbS activity.

Therefore, even though this σ factor has been well characterised, some research questions still remain to be answered. The study on specific promoter requirements by OrbS in *B. cenocepacia*, presence of novel promoters and the investigation of the presence of an alternative regulatory mechanism of OrbS form the crux of this chapter.

Based on the hypotheses and aims described in section 1.7, the following experiments were designed:

- 1) The previously made constructs used to study the contribution of single and multiple bases in the recognition of promoters by OrbS in *E. coli* would be transferred to *B. cenocepacia* and the promoter activity would be tested under inducing conditions. In particular, the effect of sequence alterations in the spacer region and the region around the transcription initiation site, that would not be expected to be specifically recognised by OrbS.
- 2) The activity of putative OrbS-dependent promoters identified by the previous *in silico* analysis would be tested in *E. coli* and *B. cenocepacia*.
- 3) The presence of additional target promoters in the *B. cenocepacia* genome that are recognised by OrbS would be functionally scanned through a ChIP-seq study.
- 4) The activity of OrbS-dependent promoters would be tested in a *B. cenocepacia fur* mutant to test if the OrbS activity is regulated in the absence of *fur*. First, the *fur* mutant was to be phenotypically tested by studying its siderophore production, growth efficiency and effect on known Fur-regulated promoters.
- 5) The target regions of Fur on the *B. cenocepacia* genome would be identified by undertaking a ChIP-seq study.

5.2 Investigation of DNA sequence requirements for OrbS activity

As described before, the OrbS-dependent promoter P_{orbH} was previously examined to determine the important DNA sequence requirements for effective promoter utilisation. To substitute a base, 'extreme' base substitutions were introduced where a purine was replaced with a pyrimidine that the original purine does not pair with and *vice versa*. Since these experiments were carried out in *E. coli* where some unexpected contribution of bases located away from the -35 and -10 promoter region was observed, it was decided to study the constructs in *B. cenocepacia*.

5.2.1 Activity of OrbS dependent promoters under inducing conditions in *B. cenocepacia*

First, the activity of all three OrbS-dependent promoters was verified in *B. cenocepacia*. An alignment of the three promoters shows conserved features such as the -10 (CGTC) and -35 (TAAA) regions, the GC-rich spacer and the A+G-rich region around the TSS (Figure 5.1A). In addition, it also highlights the conserved GC rich and AT rich regions located upstream and downstream of the -35 region, respectively. Previously constructed plasmids pKAGd4-P_{orbHds6}, pKAGd4-P_{orbEds1} and pKAGd4-P_{orblds1} containing the experimentally determined 'minimal' promoters required for effective utilisation by OrbS were separately transferred to *B. cenocepacia* 715j by conjugation using *E. coli* S17-1 and the promoter activity was measured using the β -galactosidase enzyme assay in cells growing under iron limiting (inducing) conditions. All three promoters were found to be active under inducing conditions. However, P_{orblds1} was found to be approximately two fold more active than P_{orbHds6} and approximately ten fold more active than P_{orbEds1} (Figure 5.1B).

A

P_{orbH} : GCGG TAAA AAAAC GCGCCGGCC AAC CGTC TATCAGACAGGAG
 P_{orbE} : GGGC TAAA TATTT GCGCGGCC GTT CGTC TATCAGGGAG
 P_{orbI} : ACGG TAAA AAATC GGCCGCGCC GTT CGTC ACACCAGTGA

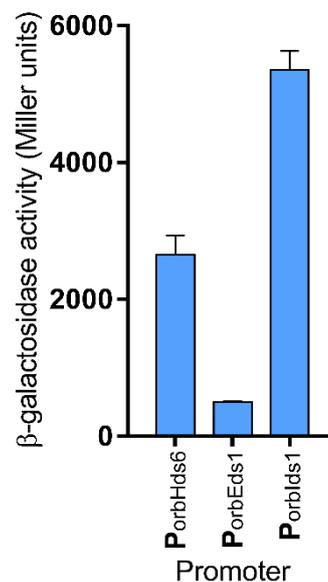
B

Figure 5.1. Promoter activities of OrbS dependent promoters in *B. cenocepacia*:

(A) Alignment of the minimal P_{orbH} , P_{orbE} and P_{orbI} promoter sequences with highlighted conserved promoter features – green showing -10 and -35 promoter regions, yellow showing GC rich spacer region and magenta showing the A+G rich region. (B) pKAGd4- $P_{orbHds6}$ or pKAGd4- $P_{orbEds1}$ or pKAGd4- $P_{orblds1}$ were each separately transferred to *B. cenocepacia* 715j and promoter activity was measured using the β -galactosidase assay where cultures were grown in LB with chloramphenicol and dipyrindyl. Activities shown have been obtained by subtracting activity of same strain containing pKAGd4 assayed under same conditions (background control), n=3.

5.2.2 Analysis of the P_{orbH} promoter by studying the effect of single base substitutions on OrbS-dependent activity in *B. cenocepacia*

As described in section 5.1, the effect of introducing single base extreme substitutions in and adjacent to the putative -10 and -35 promoter regions of the minimal *orbH* promoter (P_{orbHds6}) (K. Agnoli 2007). The locations of the substituted bases in these derivatives which were analysed in *E. coli* are shown in Figure 5.2. Since some unexpected contribution of bases were expected as described before it was necessary to study the activities of these derivatives when OrbS interacts with the *B. cenocepacia* RNA core polymerase to form a holoenzyme.

Therefore, the previously constructed pKAGd4-P_{orbHds6} substitution derivatives were transferred to *B. cenocepacia* 715j by carrying out conjugation using S17-1 cells as the donor strain. The promoter activity under iron starvation conditions was then measured by carrying out β -galactosidase assays. Figure 5.3 shows the promoter activity of each derivative relative to WT activity (P_{orbHds6}). The most detrimental effects of substitution were observed at the 'CGTC' (-12 to -9) and 'TAAA' (-33 to -30) motifs forming the -10 and -35 promoter regions, respectively, consistent with the observations in the previous study. The decrease in promoter activity was more pronounced at the -10 region compared to the -35 region, especially following single base substitution at positions -9, -10 and -11 where promoter activity was completely abolished. The promoter activity was reduced to approximately 6% and 8% of the WT activity by substitution at positions -32 and -33, respectively. Further, substitution at positions -30 and -31 led to a reduction in promoter activity to approximately 12% and 15% of the WT promoter activity. This may suggest that 'T' and 'A' at positions -32 and -33 might play a role in the interaction of OrbS with the -35 promoter region.

Further, substitution at positions -27 and -28 gave rise to a ~ 50% decrease in promoter activity compared to the WT promoter activity suggesting that the adenine bases forming the extended AT track downstream of the -35 promoter region might be important in promoter utilisation. Substitution of the position C -25 with 'A' increased promoter activity by an approximate two fold as compared to WT activity, consistent with the results obtained in *E.coli*. This may be attributed to the additional 'A' further

extending the AT track downstream of the -35 region. Substitutions at positions -14 and -15 also led to a $\geq 50\%$ decrease in promoter activity.

Extreme base substitutions that led to an increase in promoter activity in comparison to WT activity were also seen at position -4 in the A+G rich region, position -13 located upstream of the -10 promoter region, and positions -37 and -35 in the GC rich region upstream of the -35 promoter region in *E. coli* (K. Agnoli 2007). However no such increases in promoter activity were observed in *B. cenocepacia*. Bases that were substituted near the TSS did not lead to significant decrease in promoter activity as compared to WT activity in *B. cenocepacia* inconsistent with the results obtained in *E. coli*. This may suggest that individually bases near the TSS may not be important when OrbS interact with the promoter.

Location	Sequence
WT	GCGG TAAA AAAACGCGCCGGCCAAC <u>CGTC</u> TATCAGACAGGAG
-37	T CCGG TAAA AAAACGCGCCGGCCAAC <u>CGTC</u> TATCAGACAGGAG
-36	G A GG TAAA AAAACGCGCCGGCCAAC <u>CGTC</u> TATCAGACAGGAG
-35	GC T G TAAA AAAACGCGCCGGCCAAC <u>CGTC</u> TATCAGACAGGAG
-34	GCG T TAAA AAAACGCGCCGGCCAAC <u>CGTC</u> TATCAGACAGGAG
-33	GCGG G AAA AAAACGCGCCGGCCAAC <u>CGTC</u> TATCAGACAGGAG
-32	GCGG T C AAA AAAACGCGCCGGCCAAC <u>CGTC</u> TATCAGACAGGAG
-31	GCGG T A C AAA AAAACGCGCCGGCCAAC <u>CGTC</u> TATCAGACAGGAG
-30	GCGG T AA C AAA AAAACGCGCCGGCCAAC <u>CGTC</u> TATCAGACAGGAG
-29	GCGG T AAA C AA AAAACGCGCCGGCCAAC <u>CGTC</u> TATCAGACAGGAG
-28	GCGG T AAA A C AA AAAACGCGCCGGCCAAC <u>CGTC</u> TATCAGACAGGAG
-27	GCGG T AAA AA C A C AAAACGCGCCGGCCAAC <u>CGTC</u> TATCAGACAGGAG
-26	GCGG T AAA AAA C C G AAAACGCGCCGGCCAAC <u>CGTC</u> TATCAGACAGGAG
-25	GCGG T AAA AAA A A G AAAACGCGCCGGCCAAC <u>CGTC</u> TATCAGACAGGAG
-15	GCGG T AAA AAA A C G AAAACGCGCCGGCC <u>CAC</u> <u>CGTC</u> TATCAGACAGGAG
-14	GCGG T AAA AAA A C C G AAAACGCGCCGGCC <u>AC</u> <u>CGTC</u> TATCAGACAGGAG
-13	GCGG T AAA AAA A C A A AAAACGCGCCGGCC <u>AA</u> <u>CGTC</u> TATCAGACAGGAG
-12	GCGG T AAA AAA A C A G AAAACGCGCCGGCC <u>AA</u> <u>CGTC</u> TATCAGACAGGAG
-11	GCGG T AAA AAA A C C T AAAACGCGCCGGCC <u>AA</u> <u>CGTC</u> TATCAGACAGGAG
-10	GCGG T AAA AAA A C C G G AAAACGCGCCGGCC <u>AA</u> <u>CGTC</u> TATCAGACAGGAG
-9	GCGG T AAA AAA A C G T AAAACGCGCCGGCC <u>AA</u> <u>CGTC</u> TATCAGACAGGAG
-8	GCGG T AAA AAA A C G T C AAAACGCGCCGGCC <u>AA</u> <u>CGTC</u> TATCAGACAGGAG
-7	GCGG T AAA AAA A C G T C T AAAACGCGCCGGCC <u>AA</u> <u>CGTC</u> TATCAGACAGGAG
-6	GCGG T AAA AAA A C G T C T A AAAACGCGCCGGCC <u>AA</u> <u>CGTC</u> TATCAGACAGGAG
-5	GCGG T AAA AAA A C G T C T A A AAAACGCGCCGGCC <u>AA</u> <u>CGTC</u> TATCAGACAGGAG
-4	GCGG T AAA AAA A C G T C T A C AAAACGCGCCGGCC <u>AA</u> <u>CGTC</u> TATCAGACAGGAG
-3	GCGG T AAA AAA A C G T C T A C A AAAACGCGCCGGCC <u>AA</u> <u>CGTC</u> TATCAGACAGGAG
-2	GCGG T AAA AAA A C G T C T A C A G AAAACGCGCCGGCC <u>AA</u> <u>CGTC</u> TATCAGACAGGAG
-1	GCGG T AAA AAA A C G T C T A C A G A AAAACGCGCCGGCC <u>AA</u> <u>CGTC</u> TATCAGACAGGAG
+1	GCGG T AAA AAA A C G T C T A C A G A C AAAACGCGCCGGCC <u>AA</u> <u>CGTC</u> TATCAGACAGGAG
+2	GCGG T AAA AAA A C G T C T A C A G A C A AAAACGCGCCGGCC <u>AA</u> <u>CGTC</u> TATCAGACAGGAG
+3	GCGG T AAA AAA A C G T C T A C A G A C A G AAAACGCGCCGGCC <u>AA</u> <u>CGTC</u> TATCAGACAGGAG
+4	GCGG T AAA AAA A C G T C T A C A G A C A G G AAAACGCGCCGGCC <u>AA</u> <u>CGTC</u> TATCAGACAGGAG
+5	GCGG T AAA AAA A C G T C T A C A G A C A G G A AAAACGCGCCGGCC <u>AA</u> <u>CGTC</u> TATCAGACAGGAG

Figure 5.2. P_{orbHds6} substitution derivatives:

The sequence of the derivatives used to study the contribution of individual bases to P_{orbH} activity is shown with the location of the substitution with respect to the transcription start site (underlined) shown on the left. The substitution is highlighted in red whereas the -10 and -35 regions are highlighted in green.

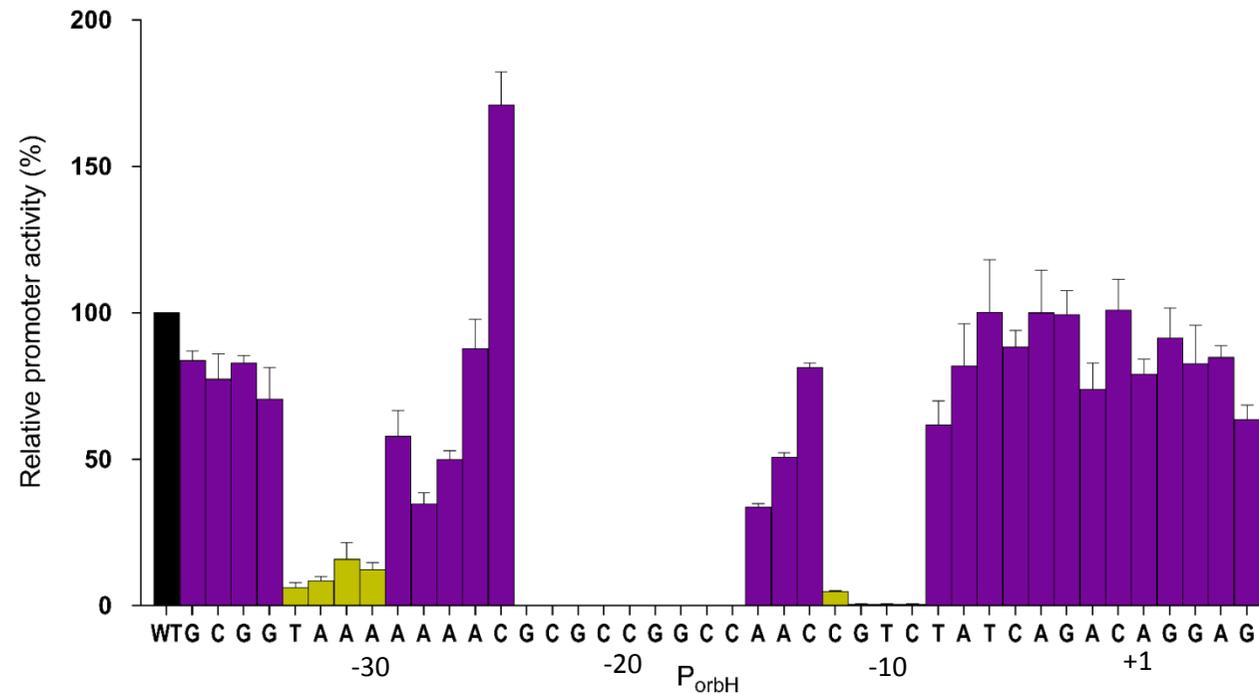


Figure 5.3. Effect of single base substitutions on P_{orbH} activity:

pKAGd4- $P_{orbHds6}$ (WT) or pKAGd4- $P_{orbHds6}$ -substitutions were each separately transferred to *B. cenocepacia* 715j and the promoter activity was measured using the β -galactosidase assay where cultures were grown in LB with chloramphenicol and dipyrindyl at 37°C. Activities shown have been obtained by subtracting the activity of the same strain harbouring pKAGd4 assayed under same conditions (background control), n=3. The relative activity (%) of the promoter in comparison to pKAGd4- $P_{orbHds6}$ (WT) activity is presented. Single base substitutions were not introduced at positions -24 to -16. Bars in dark yellow represent the relative activities of the experimentally determined -10 and -35 regions 'CGTC' and 'TAAA' respectively. Bp co-ordinates are given below the bases (+1 corresponds to the transcription start site).

5.2.3 Analysis of GC rich spacer region in *B. cenocepacia*

P_{orbH} has a 17 bp spacer in between the -10 and -35 promoter regions which has a conserved 9 bp GC-rich block. To explore the role of this spacer, a number of promoter variants containing multiple substitutions within the spacer sequence were previously made and cloned into pKAGd4 by K. Agnoli (Figure 5.4A). To assess the importance of the GC-rich region extreme substitutions were introduced in all or alternate bases from positions -24 to -16. To assess the length of the spacer four promoter variants containing alternations within the spacer were made – two containing nucleotide insertions and two containing nucleotide deletions as shown in Figure 5.4A.

I transformed *E. coli* S17-1 cells with these six plasmid constructs following which they were introduced into *B. cenocepacia* 715j by conjugation. The activities of the promoters were measured under iron starvation conditions and compared to that of pKAGd4- $P_{orbHds6}$ (WT) (Figure 5.4B). Substituting all nine guanine and cytosine bases within the spacer region reduced the activity of the promoter by approximately 90% in *B. cenocepacia*. This may be due to a change in DNA conformation induced by the extended AT-rich region affecting promoter utilisation. However, even substituting alternate bases within the GC-rich tract caused a drop in activity by approximately 30%. In contrast, substituting the GC bases did not have detrimental effects on promoter activity in *E. coli* (K. Agnoli 2007). A marked decrease in promoter activity was observed when the spacer length was increased by two bases and when the spacer length was reduced by one or two bases in *B. cenocepacia*. This mimicked the result observed in *E. coli* (K. Agnoli 2007).

Together these results suggest that the GC richness along with the length of the spacer is extremely important for effective promoter utilisation by OrbS in *B. cenocepacia*.

A

Promoter	Sequence
WT	TAAA AAAACGCGCCGGCCAAC CGTC
GCS1	TAAA AAAAC TATAATTAAAAC CGTC
GCS2	TAAA AAAACGAGACTGACAAC CGTC
GCS+1	TAAA AAAACGCGGCCGGCCAAC CGTC
GCS+2	TAAA AAAACGCGGGCCGGCCAAC CGTC
GCS-1	TAAA AAAACGCGCCGGCCAAC CGTC
GCS-2	TAAA AAAACGCGCCGGCCAAC CGTC

B

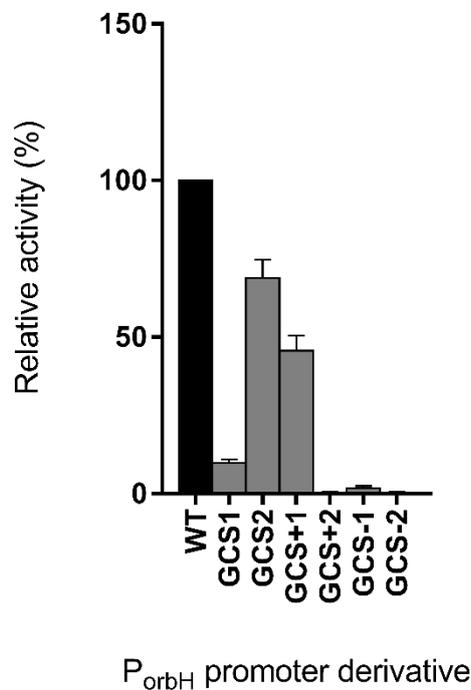


Figure 5.4. Effect of alteration of the GC-rich spacer region of P_{orbH}

(A) Sequences of P_{orbH} altered derivatives in which the rich spacer region has been modified. The sequence of P_{orbH} from the -35 region to the -10 region (positions -33 to -9) shown along with the corresponding regions of the six promoter variants. Alterations are highlighted in red and the promoter -10 and -35 regions are highlighted in green. (B) pKAGd4-P_{orbHds6} (WT) or pKAGd4-P_{orbHds}-substitutions were each separately transferred to *B. cenocepacia* 715j and the promoter activity was measured using the β -galactosidase assay where cultures were grown in LB with chloramphenicol and dipyrindyl. Activities shown have been obtained following subtraction of the activity of the same strain containing pKAGd4 assayed under same conditions (background control), n=3. The relative activity (%) of each promoter in comparison to pKAGd4-P_{orbHds6} (WT) activity is presented.

5.2.4 Analysis of the A+G rich transcription start site region in *B. cenocepacia*

The bases near the transcription start site of all three OrbS-dependent promoters are rich in their AG content. To study the contribution of these bases, all bases in the 9 bp AG block were either deleted or replaced with extreme base substitutions. Additionally, the bases were replaced by extreme substitutions in blocks of three to assess the contribution of each these blocks as shown in Figure 5.5A. These promoter derivatives were made and cloned into pKAGd4 by K. Agnoli (K. Agnoli 2007). I transferred these plasmids to *B. cenocepacia* 715j by carrying out conjugations using S17-1 donor cells. The promoter activity was measured by carrying out β -galactosidase assays in cells growing under iron starvation conditions. The relative activity of each promoter variant with respect to pKAGd4-P_{OrbHds6} (WT) activity is shown in Figure 5.5B.

It was apparent that when all the adenines and guanines in the AG rich block were deleted or replaced, the promoter activity was completely abolished in *B. cenocepacia*. Substituting the bases in blocks of three only severely affected the promoter activity when the substituted triad included the transcription start site. In contrast, promoter activity was not adversely affected in *E. coli* by these substitutions except when all nine bases in the AG rich block were substituted (K. Agnoli, 2007). Although severe loss in promoter activity was seen when the three bases around the TSS were substituted, there were negligible effects on promoter activity when each base was individually substituted.

A

Promoter	Sequence
WT	CGTC TATCAGACAGGAG GGATCC TAA
1	CGTC TATC GGATCC TAA
AGB1	CGTC TATC CTCACTTCT GGATCC TAA
AGBtri1	CGTC TATC CTC CAGGAG GGATCC TAA
AGBtri2	CGTC TATCAGA ACT GAG GGATCC TAA
AGBtri3	CGTC TATCAGACAG TCT GGATCC TAA

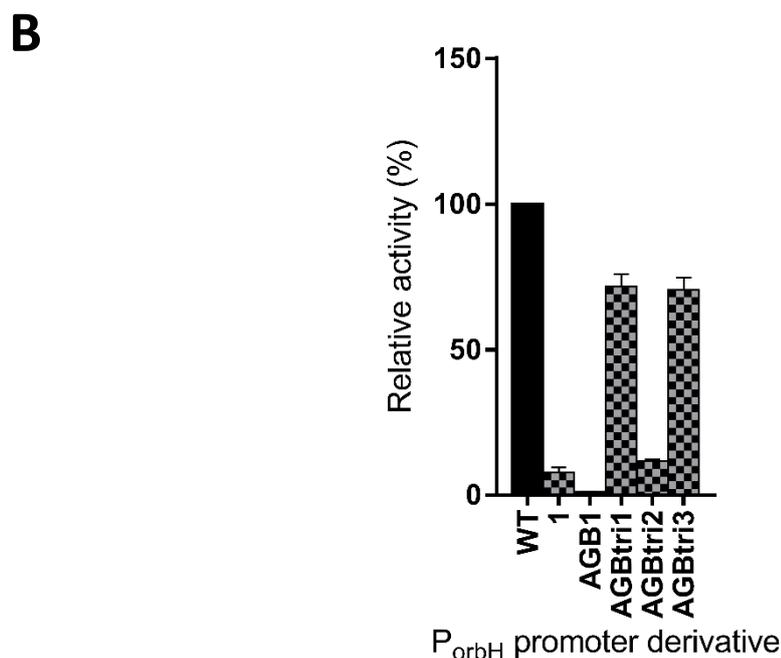


Figure 5.5. Effect of alterations to the A+G-rich region at the TSS on P_{orbH} activity:

(A) Sequences of P_{orbH} derivatives in which region near the transcription start site (base at position +1 underlined) has been modified. The sequence of P_{orbH} from the -10 region up to the vector sequence is shown along with the corresponding region of the six promoter variants. Alterations are highlighted in red and -10 region is highlighted in green. BamHI site is highlighted in blue and stop codon is in red font. (B) pKAGd4-P_{orbHds6} (WT) or pKAGd4-P_{orbHds}-substitutions were each separately transferred to *B. cenocepacia* 715j and the promoter activity was measured using the β -galactosidase assay where cultures were grown in LB with chloramphenicol and dipyrindyl. Activities shown have been obtained following subtraction of the activity of the same strain containing pKAGd4 assayed under same conditions (background control), n=3. The relative activity (%) of each promoter in comparison to pKAGd4-P_{orbHds6} (WT) activity is presented.

5.2.5 Re-construction of one P_{orbH} variant promoter and its analysis in *E. coli*

The DNA sequence of each previously made construct used in this project was checked for its integrity prior to analysis in *B. cenocepacia*. This analysis revealed that for one of the single base substitution derivatives P_{orbHds19}, the base at position -5 was mutated in addition to the desired substitution (C replaced by A). Hence, I annealed oligonucleotides that had the correct P_{orbHds19} sequence such that *HindIII* and *BamHI* compatible overhangs were created. This was ligated to pKAGd4 between the *HindIII* and *BamHI* sites and its sequence identity was confirmed. This construct was transferred to *E. coli* MC1061 and the promoter activity was measured in cells growing under iron deficient conditions and compared to that of WT (P_{orbHds6}).

Also, I re-assayed three P_{orbH} variant promoters in *E. coli* MC1061 which had radically different promoter activities in *B. cenocepacia* as compared to those obtained in previous *E. coli* experiments. These were – P_{orbHds39}, which contains a single base substitution at position -37 where G was replaced by T, AGBtri2 and GCS1. Figure 5.6 shows that promoter activity obtained with the re-constructed mutant did not drastically affect promoter activity as compared to WT activity, suggesting that this base may not be important in promoter recognition. The remaining three promoter constructs that had strikingly different promoter activities in *B. cenocepacia* as compared to the previous *E. coli* results, were consistent with the results obtained in the *E. coli* experiments. This confirms that activities obtained in *B. cenocepacia* were unique to the native *B. cenocepacia* OrbS promoter requirements.

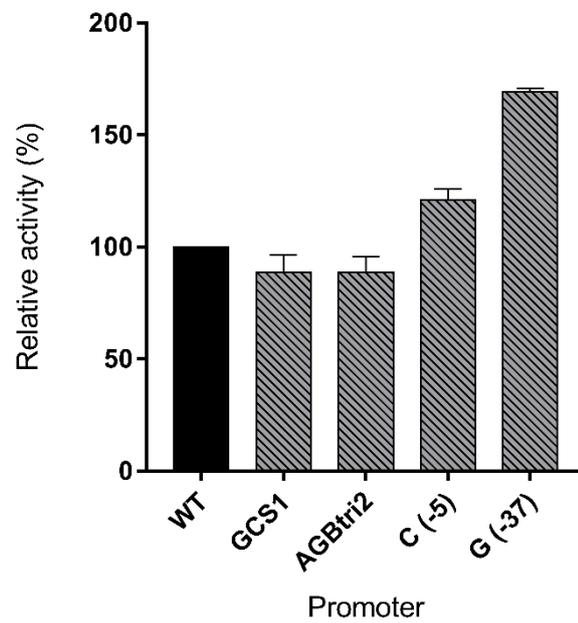


Figure 5.6. Promoter activity of selected P_{orbH} variant promoters in *E. coli*

pKAGd4- $P_{orbHds6}$ (WT) or pKAGd4- P_{orbHds} -substitutions were each separately transferred to *E. coli* MC1061 and promoter activity was measured using the β -galactosidase assay where cultures were grown in LB with chloramphenicol and dipyrindyl. Activities shown have been obtained by following subtraction of the activity of the same strain containing pKAGd4 assayed under same conditions (background control), n=3. The relative activity (%) of each promoter in comparison to pKAGd4- $P_{orbHds6}$ (WT) activity is presented. The names of the promoter derivatives or the base with the location of the substitution in brackets is given on the X-axis.

5.2.6 Testing the functionality of *in silico* determined putative OrbS-dependent promoters P_{fpr} and P_{ureA}

OrbS has 33% similarity to PvdS from *P. aeruginosa*. PvdS, in addition to regulating genes involved in iron uptake, regulates important virulence-associated genes such as PrpL and exotoxin A, and is regulated by factors involved in oxidative stress, biofilm formation and nutrient uptake (Llamas et al., 2014). Hence, it was hypothesized that OrbS might also regulate genes in addition to P_{orbH}, P_{orbE} and P_{orbI} on the *B. cenocepacia* genome. Therefore, an *in silico* analysis was previously carried out to identify potential OrbS-dependent promoters across the *B. cenocepacia* J2315 genome using the FUZZNUC programme (K. Agnoli, 2007). Using the consensus sequence 'TAAA(A/T)₄(N)₁₃CGTC' as the query revealed 7 putative promoters – CODP1 to CODP7 (candidate OrbS dependent promoters), that had conserved sequence features used by OrbS for promoter recognition. CODP1, present on chromosome 1 and located upstream of a gene encoding a putative ferridoxin NADP reductase (involved in electron transfer) also has an A+G rich region around the TSS consistent with known OrbS-dependent promoters (K. Agnoli, 2007). Another putative promoter CODP4, also located on chromosome 1, is present upstream of genes involved in urease production. Ureases are involved in hydrolysis of urea leading to a by product (ammonia) that is an easily assimilable form of nitrogen in most bacteria (K. Agnoli, 2007). This promoter also possesses an A+G rich region near the OrbS-dependent promoters. The functionality of the two putative OrbS-dependent promoters CODP1 (termed P_{fpr}) and CODP4 (termed P_{ureA}) were tested in the presence of OrbS in *B. cenocepacia*. However, the functionality of one putative promoter CODP1 (termed P_{fpr}) was tested in *E. coli* first.

5.2.6.1 Construction of pKAGd4-P_{fpr}

Oligonucleotides Pfprds1for and Pfprds1rev were annealed to give a 44 bp double stranded DNA fragment that contained the putative P_{fpr} promoter and ligated with pKAGd4 that had been digested with *HindIII* and *BamHI*. MC1061 cells were transformed with the ligated products and subsequently spread on LB agar containing

chloramphenicol and X-gal. Colonies were PCR screened using vector-specific primers AP10 and AP11. Transformant colonies containing plasmids that gave rise to a slightly larger amplicon were used to make overnight cultures, from which the plasmid DNA was extracted by miniprep and the sequence of the inserted DNA was determined for confirming its integrity.

5.2.6.2 OrbS-dependent activity of P_{fpr} in *E. coli*

Initially, P_{fpr} , being a strong candidate OrbS-dependent promoter, was first tested in *E. coli*. pKAGd4- P_{fpr} was transferred to *E. coli* MC1061 along with pBBR2 or pBBR2-orbS and the promoter activity was measured under iron deficient conditions. The OrbS-dependent promoter activity of P_{fpr} relative to the analogous P_{orbH} construct, $P_{orbHds6}$, is presented in Figure 5.7. As is evident, P_{fpr} was specifically utilised by OrbS, as in the presence of OrbS the promoter was approximately 50% as active as $P_{orbHds6}$, whereas in the absence of the σ factor the activity was approximately 4% of the activity of $P_{orbHds6}$. Hence it was concluded that P_{fpr} might be OrbS dependent.

A

$P_{orbHds6}$ GCGG TAAAAAAA CGCGCCGGCCAAC CGTCTATCAGACAGGAG
 P_{fpr} AGGG TAAAATAA CGAGTTCCTCAAAC CGTCTGTCATCGAGAAG

B

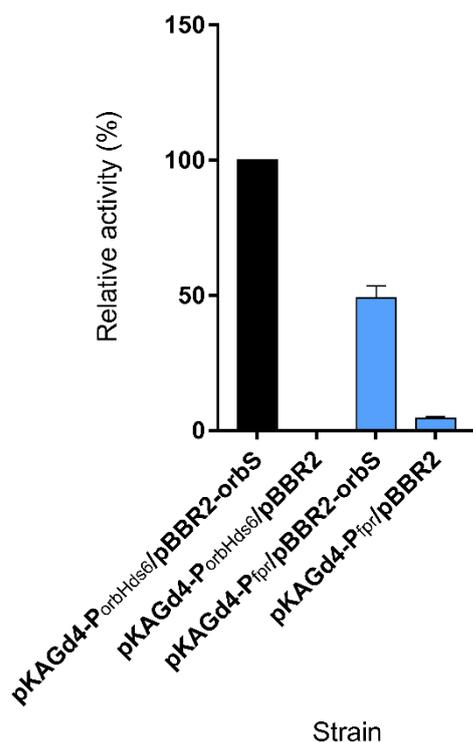


Figure 5.7. Promoter activity of putative promoter P_{fpr} in *E. coli*:

(A) DNA sequence alignment of $P_{orbHds6}$ with P_{fpr} . Conserved features are highlighted. -10 and -35 promoter regions are highlighted in green, extended AT tract downstream of -35 promoter region is highlighted in yellow, A+G rich region near transcription start site is highlighted in magenta. (B) pKAGd4- $P_{orbHds6}$ or pKAGd4- P_{fpr} were each separately transferred to *E. coli* MC1061, along with pBBR2-orbS or pBBR2 and promoter activity was measured using the β -galactosidase assay where cultures were grown in LB containing chloramphenicol, kanamycin and dipyrldyl. Activities shown have been obtained by following subtraction of the activity of the same strain containing pKAGd4 assayed under the same conditions (background control), n = 3. The relative activity (%) of the promoter in comparison to pKAGd4- $P_{orbHds6}$ activity (100%) is presented.

5.2.6.3 OrbS-dependent activity of P_{fpr} and P_{ureA} in *B. cenocepacia*

Since P_{fpr} seemed to be OrbS dependent in *E. coli* it was decided to investigate the activity in *B. cenocepacia*. pKAGd4-P_{fpr} (described in 5.2.6.1) and pKAGd4-P_{ureA} (which was previously constructed by K. Agnoli, unpublished) were transferred to *B. cenocepacia* 715j and the otherwise isogenic *orbS* mutant (715j-*orbS*::Tp). The promoter activity was then measured in cells growing in low and high iron conditions. As a control for iron-dependent regulation, the activity of P_{orbHds6} was also measured. As expected, iron-dependent promoter activity was observed for P_{orbHds6} in the WT strain where an approximate 8 fold increase in activity was seen in iron deficient conditions as compared to iron-replete conditions (Figure 5.8). No P_{orbHds6} activity was observed in the *orbS* mutant, also as expected. P_{fpr} was about 4 times less active as the P_{orbHds6} under low iron conditions and in the presence of OrbS in the *B. cenocepacia* WT strain. However, surprisingly P_{fpr} was insensitive to the iron concentration in the medium and was equally active in the presence and absence of OrbS. Contrastingly, P_{ureA} was inactive under iron replete and deficient conditions in both the presence and absence of OrbS. These results suggest that P_{ureA} does not have OrbS-dependent promoter activity. P_{fpr} seems to be served by another σ factor in *B. cenocepacia*.

A

P_{orbHds6} GCGG **TAAAAAAA** CGCGCCGGCCAAC **CGTC** TATCAGAC **CAGGAG**
P_{fpr} AGGG **TAAAATAA** CGAGTTCCTCAAAC **CGTC** TGTCATCGAGAAG
P_{ureA} CCGA **TAAAAATTA** CGTTTTCTGAAAC **CGTC** TGCCTTCATGAAA

B

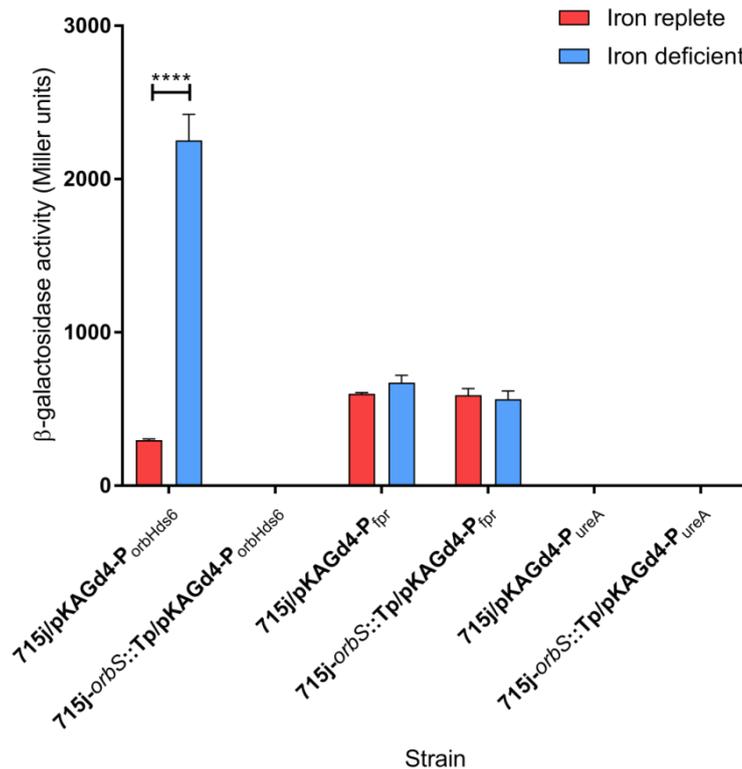


Figure 5.8. Promoter activity of putative promoters P_{fpr} and P_{ureA} in *B. cenocepacia*

(A) DNA sequence alignment of P_{orbHds6} with P_{fpr} and P_{ureA}. Conserved features are highlighted. -10 and -35 promoter regions are highlighted in green, extended AT tract downstream of -35 promoter region is highlighted in yellow, A+G rich region near transcription start site is highlighted in magenta. (B) pKAGd4-P_{orbHds6} or pKAGd4-P_{fpr} or pKAGd4-P_{ureA} were each separately transferred to *B.cenocepacia* 715j or *B.cenocepacia* *orsS* mutant 715j-*orsS*::Tp and promoter activity was measured using the β -galactosidase assay where cultures were grown in LB containing chloramphenicol, kanamycin and dipyrindyl for iron deficient conditions or ferric chloride for iron replete conditions. Activities shown have been obtained by following subtraction of the activity of the same strain containing pKAGd4 assayed under same conditions (background control). The relative activity (%) of the promoter in comparison to pKAGd4-P_{orbHds6} activity is presented. Statistical significance was tested using two-way ANOVA with a Tukeys' post-test (n=3), ****= $p < 0.0001$, ns=not significant.

5.2.6.4 Identification of OrbS binding regions on the *B. cenocepacia* genome

As a complementary approach to identify any additional OrbS-dependent promoters on the *B. cenocepacia* genome it was decided to employ CHIP-seq. To do this, it was decided to use the IPTG-inducible plasmid pSRK-Km where *orbS* encoding the FLAG epitope at the C-terminus of the σ factor could be cloned. For this, primers OrbSfor and OrbSFLAGrev were used to amplify the *orbS* ORF using H1111 boiled lysate and KOD polymerase. The FLAG tag coding sequence was incorporated in the reverse primer before the stop codon. The amplified product (660 bp *orbS* + 24 bp FLAG tag coding sequence) was electrophoresed in an agarose gel, excised with a scalpel and purified. Both pSRK-Km and the amplified product were digested with *NdeI* and *XbaI*, combined and ligated, after which JM83 cells were transformed with the ligations. Cloning an insert into pSRK-Km interrupted the *lacZ* gene giving rise to white colonies, as opposed to blue colonies with no insert. White colonies were PCR screened using GoTaq Polymerase and insert-specific primers OrbSfor and OrbSFLAGrev. Clones that seemed to have an insert were used to inoculate 2 ml of LB to make overnight cultures, which were used to extract plasmid DNA by performing the plasmid miniprep technique. Plasmid clones containing the desired insert were expected to be 6.460 kb (5.776 kb + 684 bp insert). However when they were subjected to DNA sequence analysis, multiple mutations were obtained in the region corresponding to the amplification primers, i.e. at the FLAG tag or 5' end of the *orbS* ORF. Several further attempts were made to clone the *orbS* amplicon, including testing re-ordered primers, but with no success. Unfortunately, due to time constraints it was decided to abort this experiment.

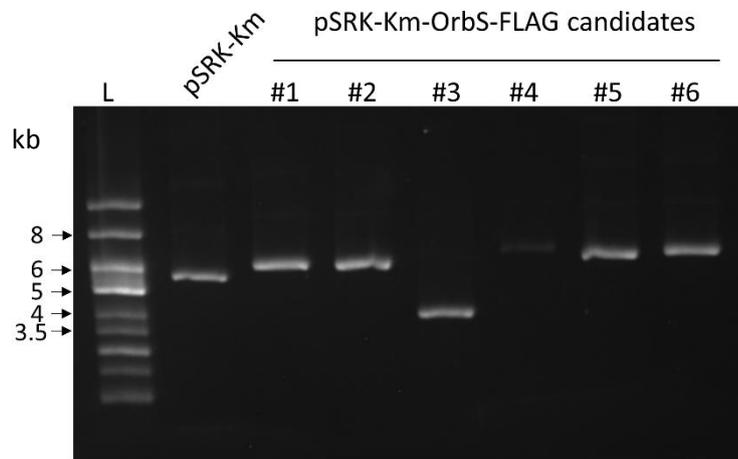


Figure 5.9. Cloning FLAG tagged *orbS* into pSRK-Km

Agarose gel showing candidate clones containing pSRK-Km with *orbS* insert that has a FLAG tag coding sequence incorporated at the C-terminal coding end (positive clone = 6.460 kb), L=supercoiled ladder

5.3 Studying the regulatory mechanisms for the activity of OrbS using a *B. cenocepacia fur* mutant

The activity of P_{orbS} has been previously demonstrated to be Fur-dependent (Agnoli *et al.*, 2006). P_{orbS} activity measured in QC1732, an *E. coli fur* mutant, under both iron deficient and replete conditions showed that it was no longer iron regulated. When this *fur* mutant was complemented with a plasmid bearing functional *B. cenocepacia fur*, the activity of P_{orbS} was found to be iron regulated (Agnoli *et al.*, 2006). These observations suggested that *orbS* is regulated at the transcriptional level by Fur. Additionally, an OrbS-dependent promoter ($P_{orbHds6}$) was found to be partially iron regulated in QC1732 (i.e. in the absence of *fur*) in the presence of *orbS*. This suggested that maybe a mechanism for regulating the activity of OrbS that is independent of Fur and independent of an anti- σ factor. However, since these studies were carried out in *E. coli*, it does not provide firm evidence for the presence of an alternative regulatory mechanism and the hypothesis was needed to be tested in the native host *B. cenocepacia*.

In this project, the existence of an additional mechanism for regulating OrbS activity was investigated in *B. cenocepacia*. The absence of an obvious anti- σ factor and a cognate TBDT implied that in addition to Fur, an alternative regulatory system of OrbS activity may exist which could be tested by studying the activity of OrbS-dependent promoters in the absence of *fur* (Figure 5.10).

A previously made *B. cenocepacia fur* insertion mutant, 715j-*fur*::Tp was used for this purpose. It was constructed by transferring the suicide plasmid pSHAFT-*fur*::Tp into *B. cenocepacia* 715j and selecting for double recombinants, thus replacing the functional WT *fur* by the *fur*::Tp allele (Y. Han and M. Thomas, 2011, unpublished). Although PCR screening was consistent with successful allelic replacement, it was necessary to verify the integrity of the 715j-*fur*::Tp mutant in order to confirm that it behaved as expected (i.e. a *fur* null mutant had been generated), as well as demonstrate that any effects observed were solely due to inactivated *fur* and not due to polar effects. Therefore, in this project, this mutant was phenotypically characterised first.

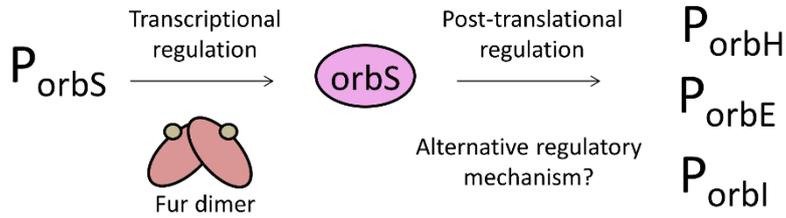
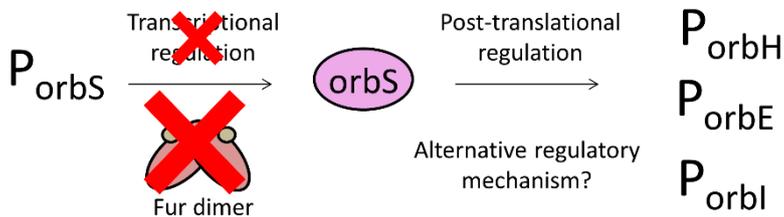
A**WT*****fur::Tp***

Figure 5.10. Experimental design to test for the presence of an alternate mechanism for regulating OrbS activity other than transcriptional regulation by Fur

(A) In *B. cenocepacia* WT it is established that Fur transcriptionally regulates the *orbS* promoter which in turn regulates the *orbH*, *orbE* and *orbI* promoters. However, it is speculated that another regulatory mechanism also exists that regulates OrbS possibly at the post-translational level. (B) In this experimental design, the known OrbS-dependent promoter (P_{orbH}) will be tested for iron regulation to examine the effect on OrbS activity in the absence of Fur.

5.3.1 Phenotypic characterisation of a *B. cenocepacia fur* mutant

Phenotypic characterisation of 715j*fur*::Tp was carried out by testing the total siderophore production, growth efficiency in a variety of liquid and agar media as well as testing the loss of Fur function by measuring the promoter activity of known Fur-dependent promoters.

5.3.1.1 Effect of mutation of *fur* on siderophore production

In order to study the effect on siderophore production of the *fur* insertion mutation, CAS agar medium containing 10 μM and 60 μM Fe^{3+} , which served as iron-limiting and iron-replete conditions, respectively, were used. CAS/HDTMA in the medium forms complexes with iron, causing the appearance of the medium to be blue. Iron capture by siderophore causes a change in colour from blue to yellow or violet which is visually detectable (Koedam et al., 1994). Ornibactin appears as yellowish-orange and pyochelin appears as violet zones around the bacterial growth on this medium, indicating production and diffusion of these siderophores.

The composition of the media and assay conditions were optimised. In order to prepare CAS agar medium, the CAS reagent is added to minimal medium. M9 minimal medium and Y-medium were selected as candidate basal minimal media. Visualisation of siderophore production was found to be most effectively observed using CAS-Y-glutamate medium at pH 6.8 and was hence used for the assay (data not shown).

From previous observations it was expected that strain 715j would produce pyochelin but negligible amounts of ornibactin, under iron-replete conditions, due to the effect of repression of *orbS* promoter by functional Fur. Alternatively, 715j-*fur*::Tp lacks this repression by Fur, hence an over-production of ornibactin was expected. To test for complementation, functional copies of the *fur* gene were introduced in 715j-*fur*::Tp in the form of pBBR-*fur*2 and pBBR2-*fur*3. This was expected to restore the WT phenotype. Additionally, 715j-*fur*::Tp containing only the empty plasmid pBBR or pBBR2 were used to control for any effects exerted by the presence of plasmids.

Figure 5.11 shows representative images of CAS agar plates with indicated strains streaked from a single colony. As expected, under iron limiting conditions (A and

B), there was a less obvious difference between siderophore production by 715j and 715j-*fur*::Tp. Although the size of the yellowish-orange zones are similar, the colour was observed to be more intense for 715j-*fur*::Tp. At a higher iron concentration (C and D), repression of siderophore production was observed for 715j, whereas, 715j-*fur*::Tp continued to over-produce ornibactin as expected. When complemented with a functional *fur* gene, 715j-*fur*::Tp behaved as 715j, shutting off ornibactin production and only producing pyochelin under iron replete conditions.

As the complementation by pBBR2-*fur*3 visually appeared more evident (B and D), it was decided to be used for further complementation experiments.

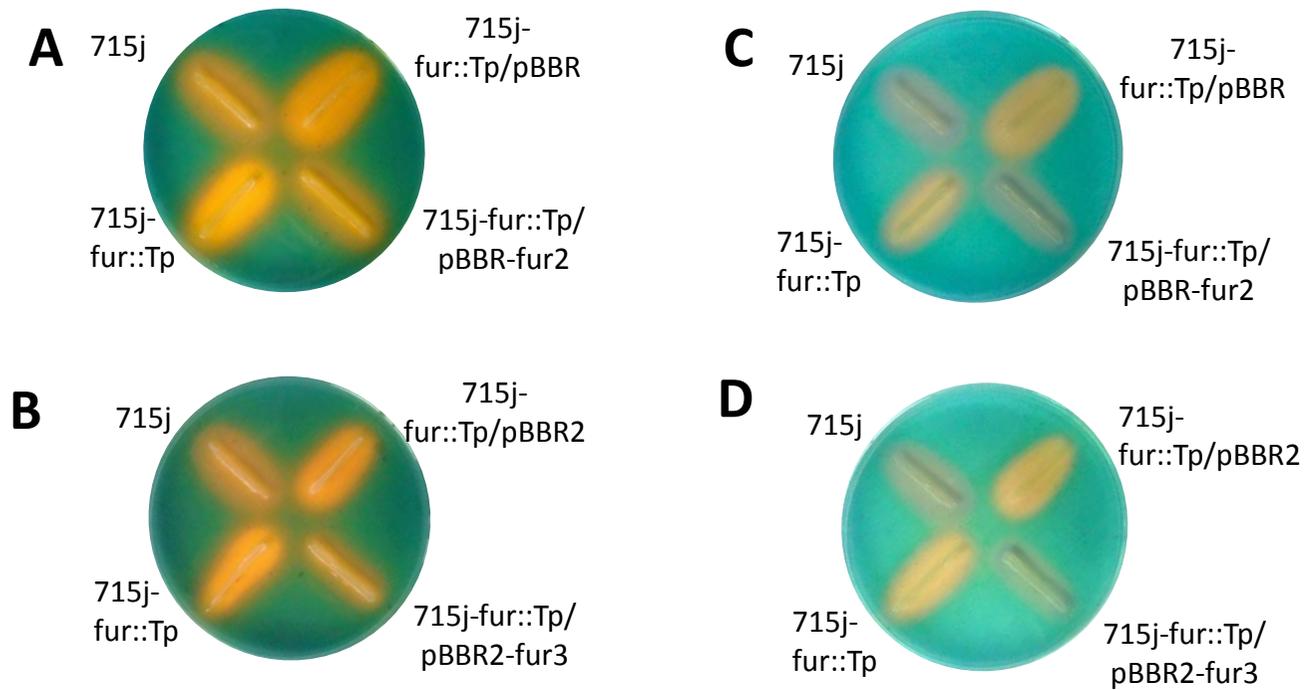


Figure 5.11 CAS plate analysis of siderophore production by *B. cenocepacia* WT and *fur::Tp* mutant:

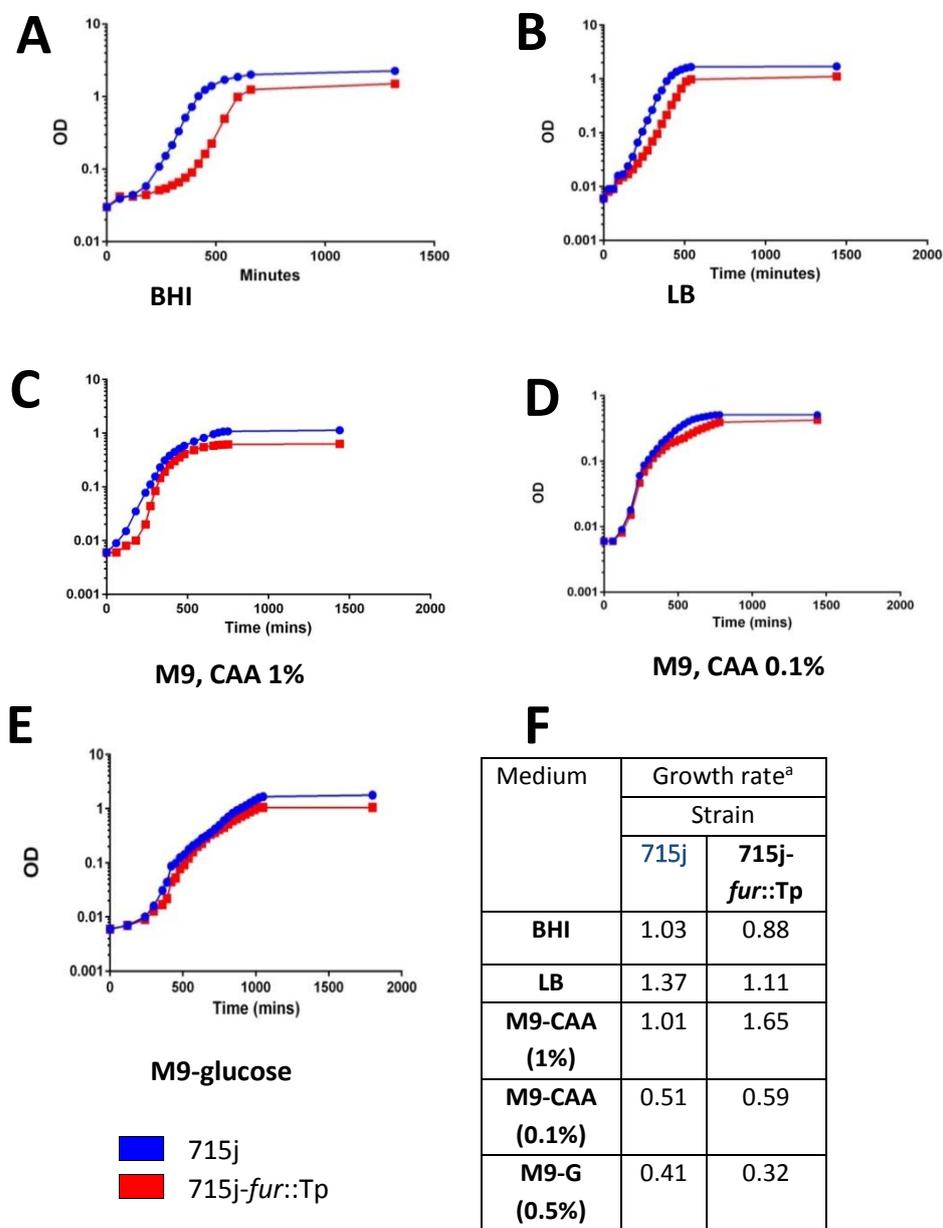
Representative images of *B. cenocepacia* 715j (WT) and *fur::Tp* mutant strains streaked on Y glutamate based CAS plates to show siderophore production. (A) and (B) Standard CAS plates containing 10 μM FeCl_3 ; (C) and (D) High iron CAS plates containing 60 μM FeCl_3 . Yellow-orange halos show ornibactin while purple halos show pyochelin production.

5.3.1.2 Effect of mutation of *fur* on growth in liquid medium

To analyse the effect of inactivation of *fur* on growth, the growth rate (doubling/hr) in nutrient rich and poor media were determined. Two nutrient rich media – LB and BHI were used while three minimal media were used – M9 (glucose) and M9 (supplemented with 1% or 0.1% casamino acids). Growth rate analysis in broth culture revealed that in all types of media, 715j-*fur*::Tp reached a lower final OD₆₀₀ at stationary phase. Representative growth curves and table from figure 5.12 show that in nutrient rich BHI broth 715j-*fur*::Tp grew at a similar rate as compared to 715j although the initial lag phase was longer for 715j-*fur*::Tp. However, once at log phase, 715j-*fur*::Tp and 715j grew at similar rates in minimal media.

5.3.1.3 Effect of mutation of *fur* on colony size and efficiency of plating

To study the effect of loss of *fur* on the ability to form colonies, the efficiency of plating was determined on LB, BHI, M9 (glucose 0.5%) and M9 (supplemented with 1% or 0.1% Casamino acids) agar media as described in section 2.1.5. Representative CFU/ml and the colony size are presented in Figure 5.13 and Table 3.2. The colony forming ability was severely reduced in the *fur* mutant as compared to the WT. Thus, 715j-*fur*::Tp gave rise to approximately 10 fold fewer colonies compared to 715j in both nutritionally rich and poor media. However, in rich media the size of the colonies appeared similar for both strains while on minimal media, especially M9 (casamino acids 0.1%) and M9 (glucose 0.5%), the size of the colonies of 715j-*fur*::Tp were much smaller.



a: Growth rate = doubling/hour

Figure 5.12. Growth rate analysis of *B. cenocepacia* WT Vs *fur::Tp* mutant:

Representative growth curves carried out at 37°C in (A) BHI (B) LB (C) M9, CAA (1%) (D) M9, CAA (0.1%) and (E) M9, glucose (0.5%) where Y-axis represents OD₆₀₀ and X-axis represents time in minutes. (F) Table showing growth rates of the indicated strains calculated as described in Section 2.1.4 using the curves shown in A, B, C, D and E.

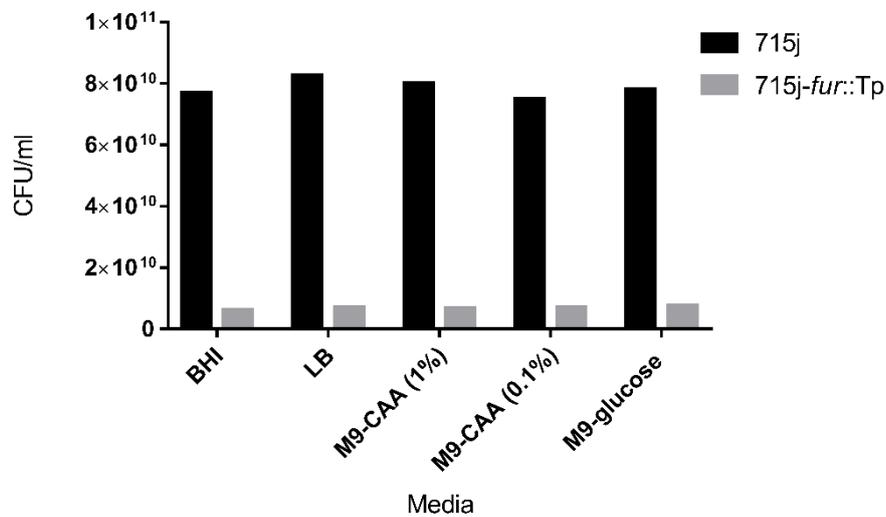


Figure 5.13. Efficiency of plating of *B. cenocepacia* WT and *fur::Tp* mutant: Representative CFU/ml at 37°C on BHI, LB, M9, CAA (1%), M9, CAA (0.1%) and M9, glucose (0.5%).

Table 3.2: Colony size comparison of 715j (WT) and 715j-*fur::Tp*

Strain	Medium				
	BHI	LB	M9-CAA (1%)	M9-CAA (0.1%)	M9-glucose (0.5%)
715j	++++	+++	+++	+++	+++
715j- <i>fur::Tp</i>	+++	+++	+++	++	++

Colony size key: ++++: ~3 mm; +++: ~1-2 mm; ++: ≤0.5 mm

5.3.1.4 Effect of mutation of *fur* on iron-regulation of Fur-regulated promoters

Finally, the *fur* mutant was also characterized by measuring the activity of known *B. cenocepacia* Fur-regulated promoters in cells growing under iron limiting and iron-replete conditions. For this purpose, previously constructed pKAGd4 derivatives containing the Fur-dependent promoters P_{orbS}, P_{firS} and P_{bfd}, fused to *lacZ* gene were transferred into S17-1 donor cells and then transferred to 715j and 715j-*fur*::Tp by conjugation. It was expected that the Fur-dependent promoters would show iron dependent regulation due to the regulatory effect of Fur in 715j (WT). Conversely, in 715j-*fur*::Tp, the promoters were expected to lose iron regulation.

However, unfortunately, conjugations between S17-1 containing pKAGd4-P_{orbS} and 715j-*fur*::Tp did not give rise to any ex-conjugants even after several attempts. Therefore the donor strain SM10 was used; however, ex-conjugants were still not obtained. This phenomenon was not further investigated due to time restrictions. As an alternative, this experiment was carried out using a previously constructed plasmid pKAGd4 containing a shorter version of P_{orbS} called pKAGd4-P_{orbS-69} that lacked sequences upstream of position -69 relative to the *orbS* start codon. In pKAGd4-P_{orbS-69} the core promoter elements are present along with the Fur box, with a further 34 bp located upstream of the -35 sequence (Agnoli *et al.*, 2006). pKAGd4-P_{orbS-69} was transferred to 715j and 715j-*fur*::Tp via S17-1 donor cells by carrying out conjugations. Using the β -galactosidase assay the activity of pKAGd4-P_{orbS} was measured in 715j host strain under both low and high iron conditions while activities of pKAGd4-P_{orbS-69}, pKAGd4-P_{firS} and pKAGd4-P_{bfd} were measured in 715j and 715j-*fur*::Tp hosts under both low and high iron conditions.

Figure 5.14 shows that the promoters were iron regulated in 715j as expected, where promoter activities increased by approximately 10, 8, 4.5 and 3 fold under low iron conditions compared to high iron conditions for P_{orbS}, P_{orbS-69}, P_{firS} and P_{bfd}, respectively. An approximate three-fold increase as compared to that in 715j was observed for P_{orbS} and P_{firS} activity under high iron conditions, in 715j-*fur*::Tp whereas an approximate 1.5 fold increase was observed in P_{bfd} activity. Such an increase in the

activity of a Fur-regulated promoter would be expected under high iron conditions in a *fur* mutant.

However, unexpectedly it was also evident that the promoters were still partially iron regulated in the *fur* mutant as the promoter activities were significantly higher in cells growing under low and high iron conditions.

The maintenance of some degree of iron regulation in the *fur* mutant raised the question as to whether these observations were an effect of making an insertion mutant, or whether insertion of the T_p^R cassette had not completely abolished Fur activity. Therefore, it was decided to perform a complementation experiment with the *fur* mutant by using pBBR2-*fur*3. Also, since these results were not in accordance with those obtained in *E. coli*, it was decided to be further investigated in *E. coli* as described in section 5.4 of this chapter.

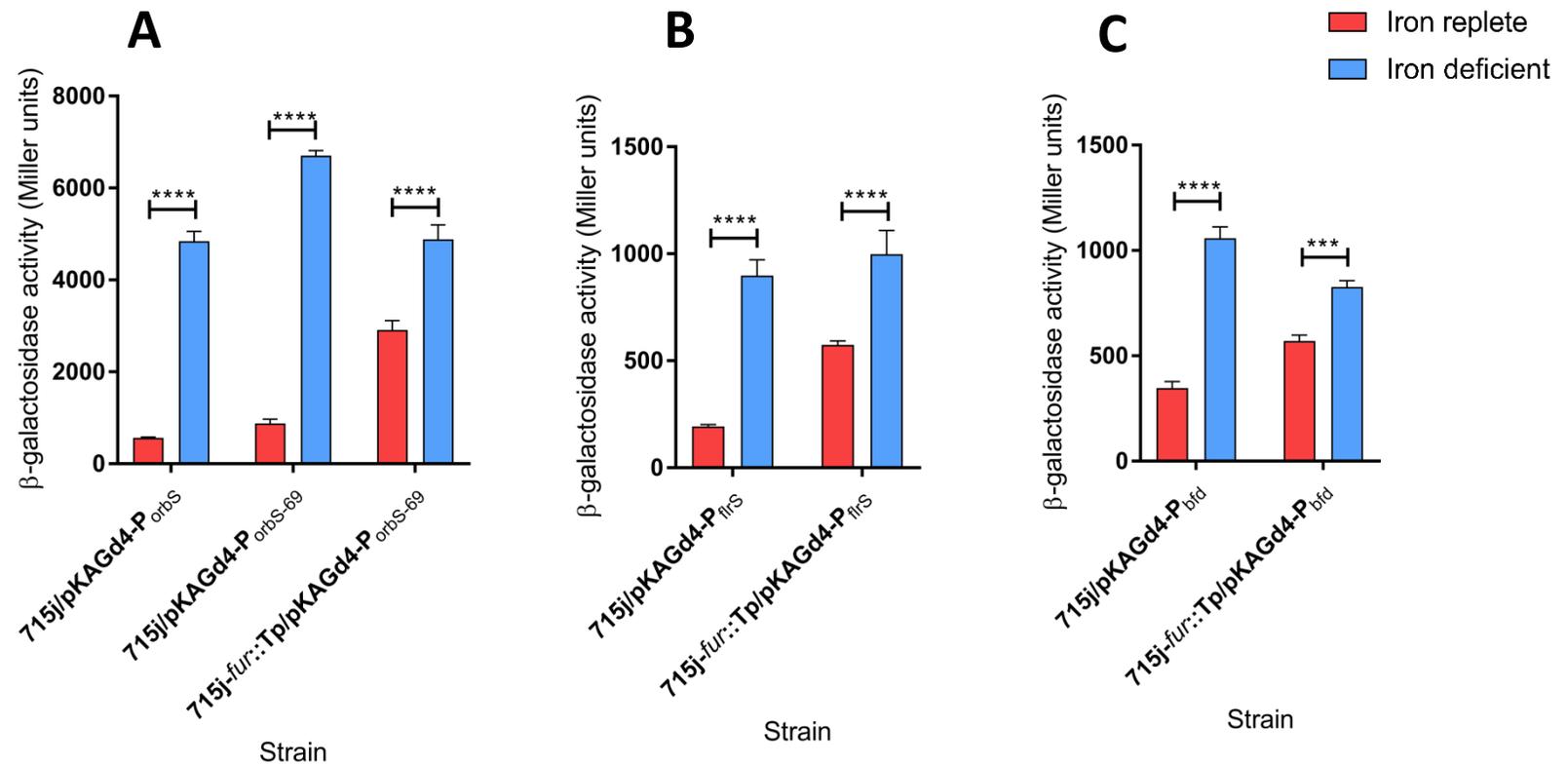


Figure 5.14. Activity of iron dependent Fur regulated promoters P_{orbs} , $P_{orbs-69}$, P_{flrS} and P_{bfd} in *B. cenocepacia* 715j and a *fur* mutant 715j-*fur::Tp* (A) pKAGd4- P_{orbs} and $P_{orbs-69}$, (B) pKAGd4- P_{flrS} and (C) pKAGd4- P_{bfd} were each separately transferred to *B. cenocepacia* 715j or *fur* mutant 715j-*fur::Tp* and promoter activity was measured using the β -galactosidase assay where cultures were grown in LB containing chloramphenicol and dipyrldyl for iron deficient conditions or ferric chloride for iron replete conditions. Activities shown were obtained by subtracting the activity of the same strain containing pKAGd4 assayed under same conditions (background control). Statistical significance was tested using two-way ANOVA with a Tukeys' post-test, ($n=3$), ****= $p<0.0001$, ***= $p<0.001$.

5.3.2 Complementation of the *fur* mutant in the investigation of Fur regulated promoters

pBBR2 and pBBR2-*fur3* were each transferred to 715j and 715j-*fur*::Tp containing pKAGd4-P_{orbS-69} or pKAGd4-P_{firs} and the promoter activities of the Fur-regulated promoters were measured during growth under low and high iron conditions by carrying out β -galactosidase assays. The rationale for performing these experiments was to establish whether the partial iron regulation of these promoters observed in the *fur* mutant could be complemented to give a degree of iron regulation similar to that observed in the WT *B. cenocepacia* strain.

Figure 5.15 shows that introducing copies of *fur* in the form of pBBR2-*fur3* restored the ability of the *fur* mutant to regulate iron at a similar level to WT. The promoter activities for both P_{orbS-69} and pKAGd4-P_{firs} were found to be the same under high iron conditions. In 715j-*fur*::Tp/pKAGd4-P_{orbS-69} complemented with *fur* an approximate seven fold increase in iron starvation conditions was obtained as compared to iron rich conditions similar to WT. Likewise, an approximate three-fold increase in iron starvation conditions compared to iron rich conditions was seen for 715j-*fur*::Tp/pKAGd4-P_{firs}/pBBR2-*fur3* similar to 715j/pKAGd4-P_{firs}/pBBR2 (WT). Also, the presence of empty plasmid did not change the partial iron regulation phenomenon seen for 715j-*fur*::Tp, demonstrating that pBBR2 did not cause any effects on its own.

This complementation experiment demonstrated that the less pronounced degree of iron regulation that was observed in the insertion mutant of *fur* was due to an effect of the Tp^R insertion on *fur* activity, and not a polar effect or the adventitious occurrence of another mutation at a different location while constructing the mutant. The insertion might have completely inactivated or partially inactivated the *fur* gene. Hence to further investigate this research question it would be ideal to study a *fur* deletion mutant further discussed in section 5.6.3.

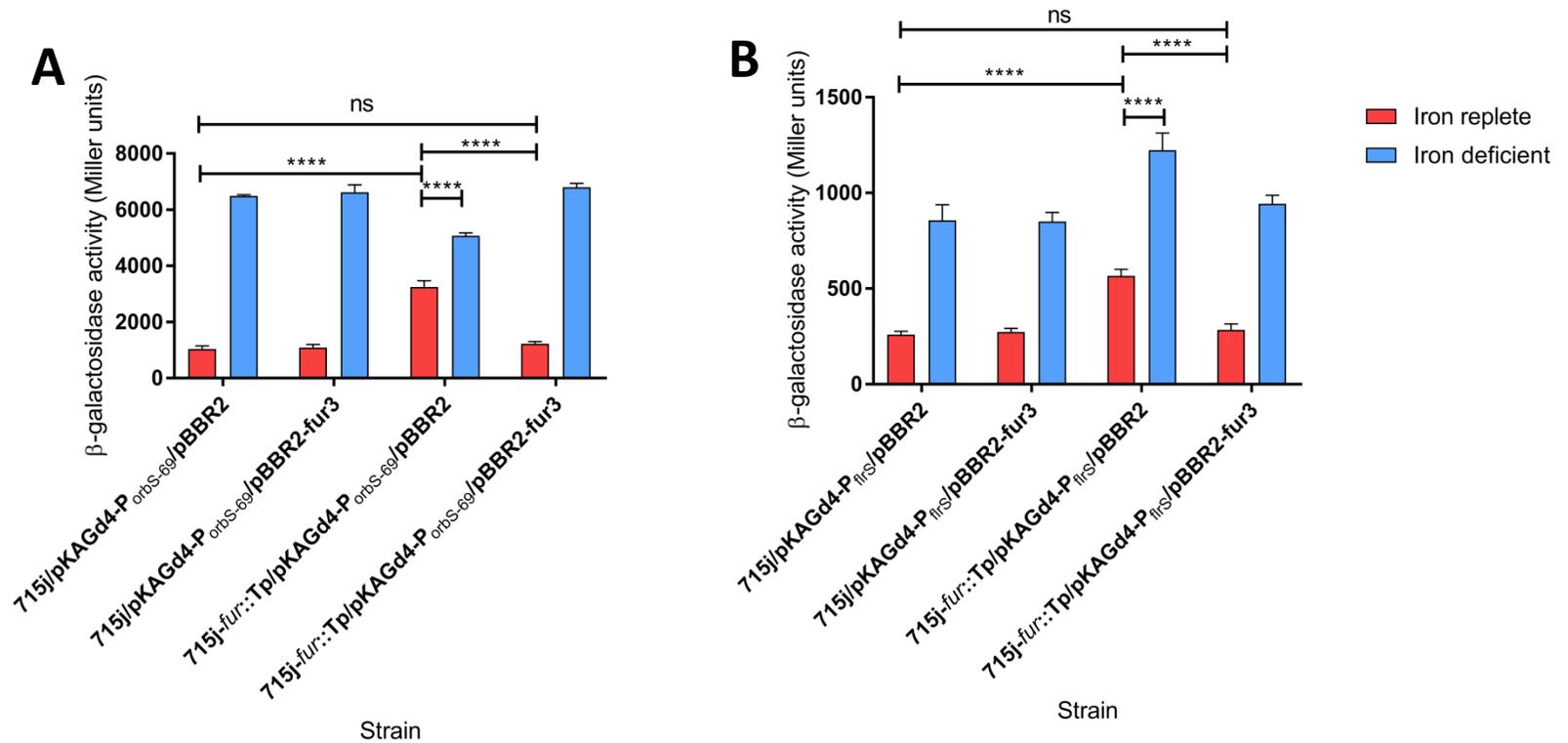


Figure 5.15. Complementation of the iron-regulation defect of a *B. cenocepacia fur*

(A) pKAGd4-P_{OrbS-69} and (B) pKAGd4-P_{flrS} were each separately transferred to *B. cenocepacia* 715j (WT) or *fur* mutant 715j-*fur::Tp* along with either pBBR2 or pBBR2-*fur3* and promoter activities were measured using the β -galactosidase assay where cultures were grown in LB containing chloramphenicol and kanamycin and dipyrindyl to achieve iron deficient conditions or ferric chloride for iron replete conditions. Activities shown have been obtained by following subtraction of the activity of the same strain containing pKAGd4 assayed under same conditions (background control). Statistical significance was tested using two-way ANOVA with a Tukeys' post-test, (n=3), ****= $p < 0.0001$

5.3.3 Iron dependent regulation of OrbS activity in the absence of Fur

Despite the failure to obtain a complete loss of iron-dependent regulation of Fur-regulated promoter in the *B. cenocepacia fur* mutant, it was decided to assess the activity of two forms of the OrbS-dependent promoter P_{orbH} (full length promoter and the 45 bp minimal derivative P_{orbHds6}) to determine if it revealed any interesting patterns of iron regulation. For this purpose, previously constructed plasmids pKAGd4-P_{orbH} and pKAGd4-P_{orbHds6} were transferred to 715j and 715j-*fur*::Tp by conjugation using S17-1 donor cells. The activities of these promoters were subsequently measured in cells growing under iron starvation and iron rich conditions. It was observed that, as with the *orbS* promoter, the OrbS-dependent promoters were iron regulated in the WT strain with an approximate 6 fold difference in activities under low and high iron conditions, whereas the degree of iron regulation was decreased to approximately 2.5 fold for P_{orbHds6} and approximately 4 fold for the full-length promoter in the *fur* mutant (Figure 5.16). This decrease in iron regulation of P_{orbH} is likely to be a consequence of the decrease in iron regulation of *orbS* transcription. However, due to residual iron regulation of *orbS* in the *fur* mutant, it was not possible to draw a conclusion regarding the presence of a mechanism for post-translational regulation of OrbS activity.

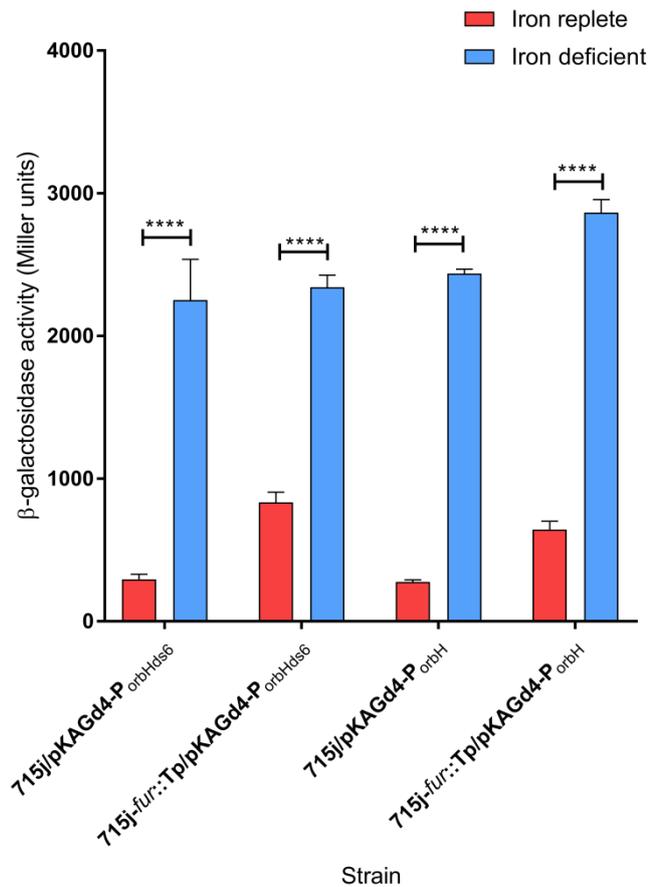


Figure 5.16. Iron-dependent activity of the OrbS-regulated promoter P_{OrbH} in *B. cenocepacia* WT and *fur* mutant strains

pKAGd4 containing either of the two variants of the P_{OrbH} promoter, (A) P_{OrbHds6} or (B) P_{OrbH} (full-length) were each transferred to *B. cenocepacia* 715j or the *fur* mutant 715j-*fur*::Tp and promoter activity was measured using the β-galactosidase assay where cultures were grown in LB containing chloramphenicol and dipyriddy for iron deficient conditions or ferric chloride for iron replete conditions. Activities shown have been obtained by following subtraction of the activity of the same strain containing pKAGd4 assayed under same conditions (background control). Statistical significance was tested using two-way ANOVA with a Tukeys' post-test, (n=3), ****= $p < 0.0001$

5.3.4 Analysis of the activity of iron-independent promoters in response to iron in *B. cenocepacia*

Finally, it was important to demonstrate that using ferric chloride and dipyriddy to obtain iron replete and deficient conditions, respectively, did not affect the promoter activities obtained and that the observed regulation was specific to iron-regulated promoters. Therefore, the responses of two different known iron independent promoters, P_{tac} and P_{cysI} were analysed in the conditions. Previously constructed plasmids pKAGd4- P_{tac} and pKAGd4- P_{cysI} were each separately transferred to 715j and 715j-*fur*::Tp by conjugation using S17-1 donor cells. A second conjugation was then carried out to transfer pBBR2 and pBBR2-*fur3* to each of these strains, again by using S17-1 donor cells. Then, in cells growing under both iron deficient and iron replete conditions, promoter activities were measured using the β -galactosidase assay. As is evident from Figure 5.17, both P_{tac} and P_{cysI} had similar activities under iron deficient and iron replete conditions. This validated that the experimental conditions used above were appropriate for demonstrating accurate iron regulation for iron dependent promoters.

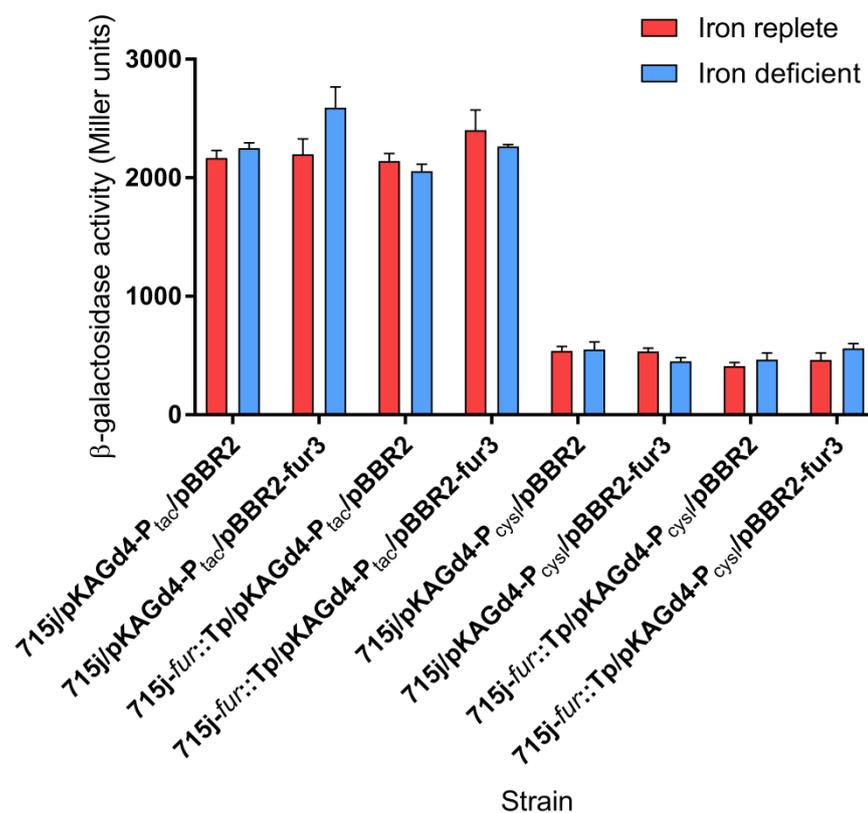


Figure 5.17. Activity of iron independent promoters P_{tac} and P_{cysl} in response to iron in *B. cenocepacia* in the presence and absence of *fur*

pKAGd4 containing P_{tac} or P_{cysl} were each transferred to *B. cenocepacia* 715j (WT) or *fur* mutant 715j-*fur*::Tp containing either pBBR2 or pBBR2-*fur*3 and promoter activity was measured using the β -galactosidase assay where cultures were grown in LB containing chloramphenicol, kanmycin and dipyrindyl to achieve iron deficient conditions or ferric chloride to achieve iron replete conditions. Activities shown have been obtained by following subtraction of the activity of the same strain containing pKAGd4 assayed under the same conditions (background control).

5.4 Investigation of iron-dependent regulation of OrbS activity using an *E. coli fur* mutant

As contrasting results were obtained for iron regulation of P_{orbS} in a *B. cenocepacia fur* mutant as compared to the *E. coli fur* mutant, it was decided to verify the results previously obtained in *E. coli*. As mentioned before, *orbS* promoter activity was measured in the *E. coli fur* mutant QC1732 and found not to be regulated by iron whereas introduction of a plasmid encoding the *B. cenocepacia fur* restored WT behaviour (Agnoli et al., 2007). It was decided to repeat the experiment to ascertain whether in fact there was a small degree of iron regulation of P_{orbS} in the *E. coli fur* mutant. Moreover, unlike the previously carried out experiment, here it was decided to use the WT host strain QC771 as a control. Here, the OrbS activity in presence and absence of the *E. coli fur* could then be assessed by comparing QC1732 to QC771 (WT).

5.4.1 Activity of Fur regulated promoter in an *E. coli fur* mutant

To establish that *E. coli* Fur could regulate P_{orbS} in response to iron and that in an *E. coli fur* mutant, the activity of P_{orbS} could no longer be regulated; the previously constructed plasmid pKAGd4- P_{orbS} was transferred to *E. coli* WT (QC771) and the otherwise isogenic *E. coli fur* mutant (QC1732). The activity of P_{orbS} was then measured under both iron deficient and iron replete conditions in both host strains. In QC771, the activity of P_{orbS} was approximately 20 fold higher when growing in iron deficient conditions as compared to iron replete conditions. In contrast, in QC1732, the difference between P_{orbS} activity in iron deficient versus iron replete conditions was not significant (Figure 5.18). Notably, the activity of P_{orbS} under both conditions was higher in the *fur* mutant than in the WT strain. These results support the previous observation that *E. coli fur* can regulate the activity of P_{orbS} in response to iron. Therefore these host strains serve as apt models to study the activity of OrbS-dependent promoters in the quest to identify iron-dependent regulatory mechanisms other than Fur.

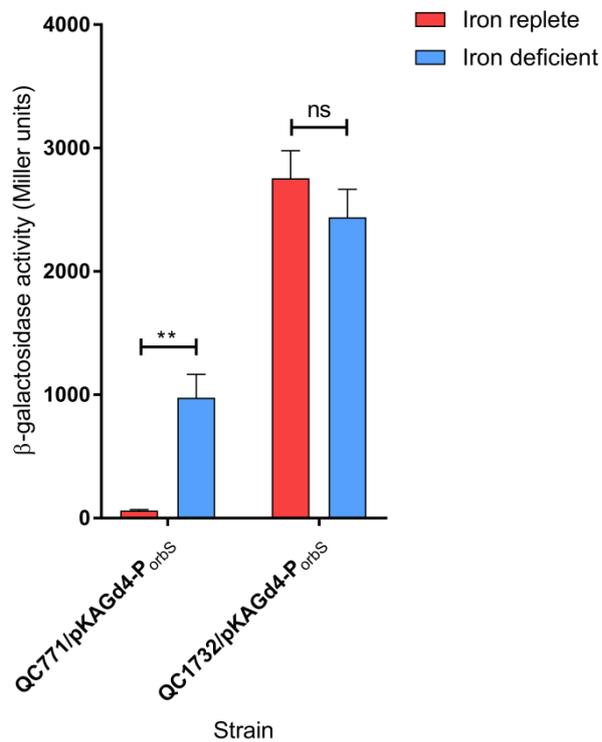


Figure 5.18. Activity of the Fur regulated P_{orbS} promoter in response to iron in *E. coli* WT and *fur* mutant strains

pKAGd4- P_{orbS} was transferred to *E. coli* WT (QC771) or *E. coli fur* mutant (QC1732) strains, and promoter activity was measured using the β -galactosidase assay where cultures were grown in LB containing chloramphenicol, and dipyriddyI for iron deficient conditions or ferric chloride for iron replete conditions. Activities shown have been obtained by following subtraction of the activity of the same strain containing pKAGd4 assayed under same conditions (background control). Statistical significance was tested using two-way ANOVA with a Tukeys' post-test (n=3), **= $p < 0.01$, ns=not significant

5.4.2 Activity of an OrbS-regulated promoter in an *E. coli fur* mutant

Next, a previously made OrbS-dependent promoter construct pKAGd4-P_{orbHds6} was transferred to QC771 (WT) and QC1732 (*fur* null mutant) cells along with either pBBR5 or pBBR5-orbS. The promoter activity was measured in cells growing under iron deficient and iron rich conditions. Figure 5.19 illustrates that in both QC771 (WT) and QC1732 iron dependent-regulation of P_{orbH} occurred. This demonstrates that in QC1732, despite the absence of *fur*, OrbS activity continues to be regulated by iron availability perhaps indicating a self-regulatory system. It is important to note that while in QC771 there was an approximate 3 fold increase in promoter activity under iron deficient conditions compared to iron replete conditions, in QC1732 there was a smaller 1.5 fold increase in promoter activity comparing the same two conditions. However, this difference was still statistically significant. Further, if an additional iron-responsive regulatory mechanism exists, it might function to full potential in combination with Fur to exert a strong iron dependent regulation on OrbS.

These conclusions based on these results were consistent with those seen in previously in *E. coli*. However, based on results obtained in *B. cenocepacia* enough evidence could not be provided to elucidate the means of Orbs activity regulation.

Investigation on the regulation of OrbS was continued further by another member in the laboratory and their results with the current understanding of this system are further described in the 'discussion' (section 5.6) of this chapter.

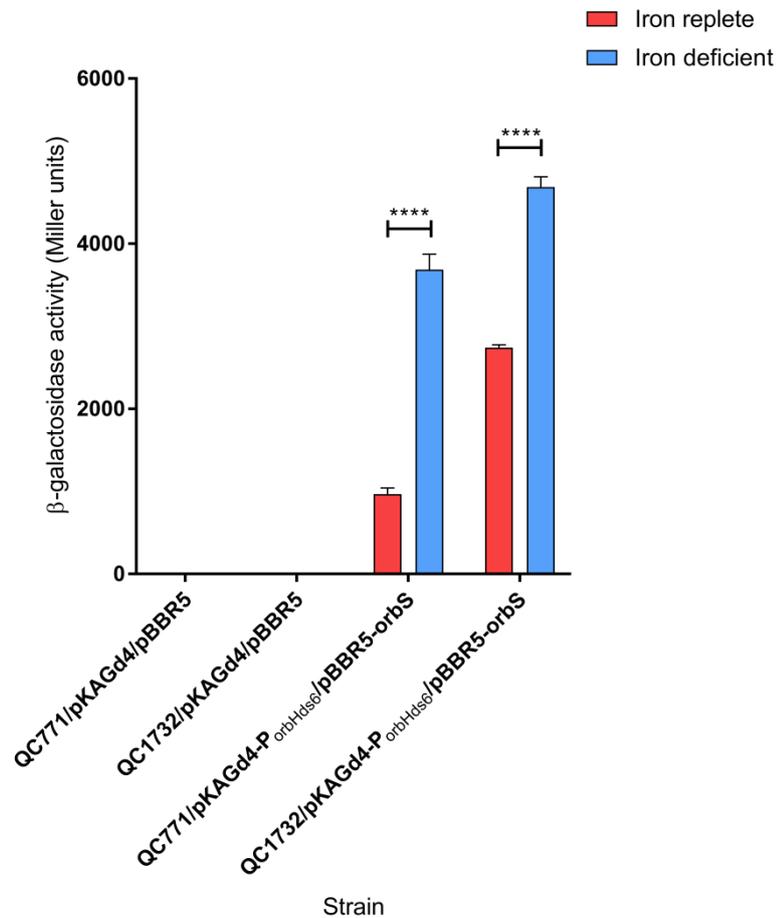


Figure 5.19. Activity of the OrbS-dependent promoter $P_{OrbHds6}$ in *E. coli* WT and *fur* mutant strains containing OrbS in response to iron:

$pKAGd4-P_{OrbHds6}$ was transferred to *E. coli* WT (QC771) or *E. coli fur* mutant (QC1732) strains along with either pBBR5 or pBBR5-orbS, and promoter activity was measured using the β -galactosidase assay where cultures were grown in LB containing chloramphenicol, gentamycin, and dipyriddy for iron deficient conditions or ferric chloride for iron replete conditions. Activities shown have been obtained by following subtraction of the activity of the same strain containing pKAGd4 assayed under the same conditions (background control). Statistical significance was tested using two-way ANOVA with a Tukeys' post-test ($n=3$), ****= $p<0.0001$, ns=not significant.

5.4.3 Construction of an iron-independent promoter and its activity in *E. coli*

To assess that the experimental conditions were reliable to study the iron dependent promoters it was important to test the activity of an iron independent promoter. Initially, QC771 and QC1732 were transformed with previously constructed plasmid pKAGd4-P_{tac} harbouring the iron independent *tac* promoter that was shown to be iron-insensitive in *B. cenocepacia* (Section 5.3.4). However, no activity was obtained. This construct was sequenced to check its DNA integrity from multiple sources (previously made glycerol stocks) but was found to reveal mixed unreadable sequences each time. Using previously annealed oligonucleotides a 128 bp DNA fragment containing P_{tac} was ligated to the *Hind*III and *Bam*HI sites in pKAGd4 that was digested with the same enzymes. MC1061 cells were transformed with the ligation products that should give rise to normal sized blue colonies. However, it gave rise to extremely small and deep blue colonies, perhaps suggesting the promoter was so active that it was exerting toxic effects or destabilising plasmid replication. This molecular cloning was repeated several times but yielded similar results each time.

A smaller version of P_{tac} (56 bp) was then designed containing -35 and -10 promoter regions with only 11 bp of the *lac* operator. Single-stranded oligonucleotides P_{tac}shortfor and P_{tac}shortrev were annealed and ligated into pKAGd4 between the sites *Hind*III and *Bam*HI and the ligation products were used to transform MC1061. However, many transformants gave rise to small and dark blue colonies as obtained with the longer P_{tac} derivative. Nucleotide sequencing of the obtained plasmids from larger colonies revealed that some mutants were obtained during this cloning process. One mutant had a deletion in the -10 region of the promoter that likely reduced its activity (sequence shown in Figure 5.20A). Hence, this clone (named pKAGd4-P_{tac2}) was used to test the assay conditions. pKAGd4-P_{tac2} was transferred to QC771 and QC1732 and the promoter activity was measured by performing β -galactosidase assays in cells growing under iron deficient and iron replete conditions. As shown in Figure 5.20, the promoter activity was found to be the same in both strains and in medium containing both iron concentrations. This implies that assay conditions used to study iron regulation were appropriate and reliable.

A

CCGTTCTGGATAATGTTTTTTTGCGCCGACATCATAACGGTTCTGGCAAATATTCTGAAATGAGCTG **TTGA**
CAATTAATCATCGGCTCG **TATAA** **-**GTGTGGAATTGTGAGCGGATAACAATTTACACA

B

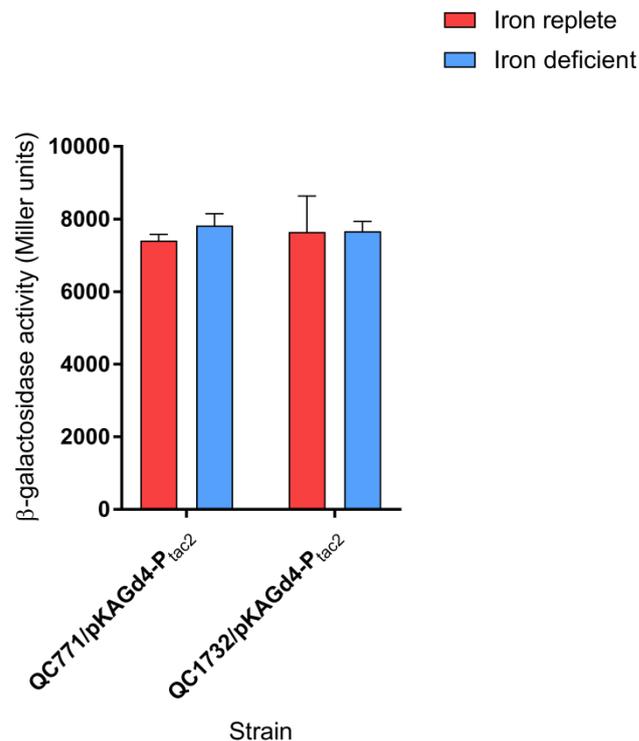


Figure 5.20. Activity of the variant *tac* promoter, P_{tac2}, in response to iron in *E. coli* WT and *fur* mutant strains.

(A) Sequence of P_{tac2} where the -35 promoter element, -10 promoter element and the lac operator are highlighted in green, blue and grey, respectively. The location of the deletion is highlighted in red. pKAGd4-P_{tac2} was transferred to *E. coli* WT (QC771) or *E. coli fur* mutant (QC1732) strains and promoter activity was measured using the β -galactosidase assay where cultures were grown in LB containing chloramphenicol and dipyrindyl to achieve iron deficient conditions or ferric chloride to achieve iron replete conditions. Activities shown have been obtained by following subtraction of the activity of the same strain containing pKAGd4 assayed under the same conditions (background control).

5.5 Identification of Fur binding regions on the *B. cenocepacia* genome

Fur is the master regulator of iron-regulated gene expression and is known to have multiple regulatory targets. These have been explored in other species by undertaking a genome-wide scan based on chromatin immunoprecipitation and deep sequencing, ChIP-seq (Butcher et al., 2011, Davies et al., 2011). In an attempt to employ a ChIP-seq based approach to study the targets of Fur in *B. cenocepacia*, an allelic replacement of WT *fur* with *fur*-FLAG was attempted to construct a strain that produces a C-terminal FLAG-tagged Fur repressor. Using PCR primers FurFLAGa and FurFLAGb, the region upstream of the *fur* stop codon (947 bp) was amplified (product 'ab') using H111 boiled lysate and KOD polymerase. Using primers FurFLAGc and FurFLAGd, the region downstream of the stop codon (545 bp) was similarly amplified (product 'cd'). In a subsequent splice-overlap extension PCR, primers FurFLAGa and FurFLAGd were used with template 'ab' and 'cd'. Since FurFLAGb and FurFLAGc were designed to contain homologous bases (FLAG tag coding sequence + stop codon), a fusion product was obtained such that the region consisting of the *fur* ORF with a FLAG tag coding sequence preceding the stop codon and with the region downstream of *fur* were amplified as a whole giving a 1.519 kb product (i.e 947 bp 'ab' + 24 bp 'FLAG' coding sequence + 3 bp 'stop codon' + 545 bp 'cd') (Figure 5.21). Optimisation of PCR conditions was carried out to give a concentrated product and was found to work most optimally when 'ab' and 'cd' were present in equimolar concentrations. This product was digested with BamHI and XbaI and ligated with pEX18TpTer-*pheS* digested with the same enzymes. JM83 cells were then transformed with the ligations and spread on M9 glycerol CAA-X-gal-IPTG medium. Transformant colonies were white when an insert was present since it disrupted the *lacZ* gene fragment while empty pEX18TpTer-*pheS* gave rise to blue colonies. A PCR screen on white colonies using insert specific primers FurFLAGa and FurFLAGd did not give any positive clones. When the size of the plasmid present in these clones was checked DNA of the expected size (4.912 kb) was not obtained. After several failed attempts at the cloning procedure, it was decided not to further investigate it due to time restrictions.

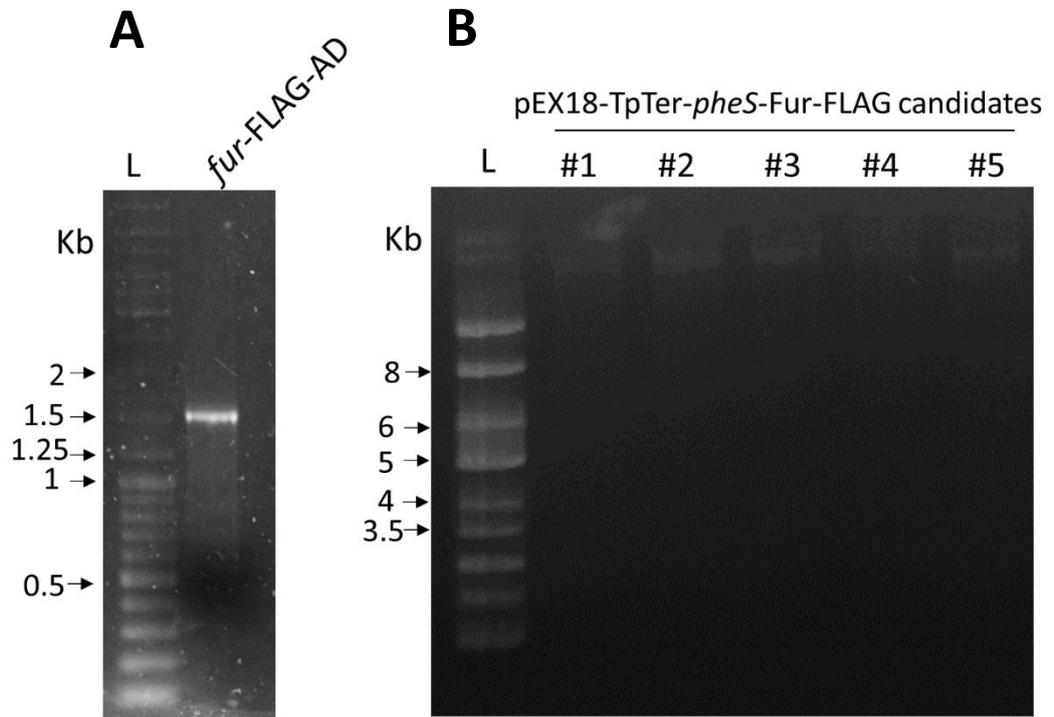


Figure 5.21. Cloning an SOE-PCR product encoding FLAG-tagged Fur into pEX18-TpTer-*pheS*: (A) SOE PCR product containing *fur* (*fur*-FLAG-AD) and a C-terminal FLAG coding sequence made using primers FurFLAGa and FurFLAGd (B) Miniprep of candidate clones containing '*fur*-FLAG', L = supercoiled ladder.

5.6 Discussion

The promoter sequence requirements of the iron starvation ECF σ factor OrbS in *B. cenocepacia* and the presence of an alternate regulatory mechanism for OrbS were explored in this project. Interesting observations with regards to contribution of individual and multiple bases within the OrbS-dependent promoter were made discussed below and it was concluded that OrbS may have one promoter target in addition to the three ornibactin promoters. However, results from this series of experiments did not provide sufficient evidence to conclude whether or not an alternative mechanism exists, other than Fur for regulating OrbS activity in *B. cenocepacia*.

5.6.1 DNA sequence requirements for effective promoter utilisation by OrbS

Currently there are three known OrbS-dependent promoters – P_{OrbH}, P_{OrbI} and P_{OrbE}. In *B. cenocepacia* under iron starvation conditions it was found that P_{OrbI} was approximately twice as active as P_{OrbH} and ten times more active compared to P_{OrbE}. P_{OrbE} drives the transcription of only one gene unlike P_{OrbH} and P_{OrbI} that drive the transcription of several genes. However, P_{OrbS} also drives the same genes as P_{OrbH} (Agnoli *et. al.*, 2006).

Previously, the -10 and -35 promoter regions of the OrbS-dependent promoter P_{OrbH} were experimentally determined in the *E. coli* host strain MC1061 (Agnoli, 2007). In this project, experiments were carried out in a *B. cenocepacia* strain mainly to address the role of bases located outside of the -10 and -35 regions that would not be expected to be contacted by OrbS. As the former experiments were carried out in a heterologous system the interactions between the *E. coli* core RNAP and the promoter DNA may not necessarily apply to the *B. cenocepacia* enzyme.

The results in *B. cenocepacia* confirmed that the -10 and -35 promoter regions are located from positions -12 to -9 (CGTC) and -33 to -30 (TAAA) with respect to the experimentally determined transcription start site. Distinct from the AT rich -10 promoter region that group 1 σ factors utilise which is comparatively flexible in terms of base sequence, ECF σ factors have developed a strict promoter recognition mechanism with reduced AT richness (Koo *et al.*, 2009). Experimental evidence suggests that σ^2 of

many ECF σ factors recognize 'CGTC' through a 'DXXR' motif, a characteristic of many alternative σ factors (Liu et al., 2012). OrbS has the 'DACR' sequence within region 2.4 of σ^2 that may be speculated to play a similar role. Results reported in this thesis reveal that binding of OrbS to the -10 region may be stronger and hence more important than at the -35 region since substitution of those bases completely abolished promoter activity. Moreover, it seemed that the cytosine base at position -12 of the -10 promoter region, and the adenine bases at positions -30 and -31 of the -35 region were slightly less important compared to the other bases in those respective promoter regions. This pair of adenines might play a structural role rather than formation of strong hydrogen bonds with σ^4 similar to that of the 'AA motif' in the -35 promoter region of *E. coli* RpoE-dependent promoters interacting with σ^E (Lane and Darst, 2006). The effects of substituting bases at the -10 and -35 promoter elements were consistent with those obtained in the previous *E. coli* experiments. However, the bases near the TSS, individually, were not found to be important for promoter utilisation contrasting to the results obtained in *E. coli*.

Investigation of the role of the spacer region separating the -10 and -35 regions revealed that the length of this region was important for OrbS-dependent transcription. This is evident from the fact that removing one or two bases or addition of two bases completely abolished promoter activity. This demonstrates the importance of the spacing between the -10 and -35 regions where σ_2 and σ_4 of OrbS interact. A steep reduction in promoter activity due to substitution of the GC tract suggests that the conformation of this nine base pair region is important. The substitution of alternate bases in this tract reduced the promoter activity by approximately 30%, suggesting that the presence of remaining GCs might compensate for the conformational change. In previous experiments performed in *E. coli* such profound effects of substituting G and C bases in the spacer of OrbS-dependent promoter were not seen. Experimental evidence demonstrates that the length and sequence of the spacer region, which were not previously thought to be of much importance, contribute majorly in σ^{70} promoter utilization (Singh et al., 2011, Sztiller-Sikorska et al., 2011). Moreover, it has been

reported that RNAP interacts with specific bases of the spacer through its β' domain (Yuzenkova et al., 2011).

Results from previous experiments analysing the role of the A+G rich region at the TSS in *E. coli* showed that substitution of bases in groups of three did not affect promoter activity to a great extent. In contrast, the results from these experiments in *B. cenocepacia* showed that multiple basepair substitutions that included the TSS base, reduced the promoter activity by up to 90%. However, individual substitutions around this region did not alter the promoter activity to a great extent. Hence, rather than the individual bases itself, the conformation around the start site might be important for proficient promoter utilization or a combination of the two.

In contrast to previous experiments performed with *E.coli*, there were a few substitutions that resulted in a substantial increase in the promoter activity. Only one substitution, a C to A base change that extended the AT tract, increased promoter activity in *B. cenocepacia* perhaps due to a favourable conformational change in this region. Also, it was observed that GC rich region upstream of the -35 promoter region was not critical for promoter utilisation at least not at the individual base level.

In conclusion, the analysis of the effect of single and multiple base pair substitutions in *B. cenocepacia* confirmed various features of the OrbS-P_{orbH} interaction. OrbS requires -10 and -35 promoter regions each of which are 4 bases in length and separated by 17 bases. A conserved 9 base pair GC-rich tract in the spacer and a 9 base pair A+G rich region around the TSS seem to be important for PorbH to adopt an appropriate conformation it to be utilized efficiently by OrbS. Furthermore, a role for the region around the experimentally determined TSS is also evident in promoter utilization. Again, the conformation and not the nature of the individual bases might be important for promoter utilisation by OrbS.

5.6.2 OrbS may have any additional promoter targets

Some iron starvation σ factors have been reported to regulate transcription of multiple promoters other than those directly involved in iron acquisition (Wilderman et al., 2001, Llamas et al., 2014). Therefore, the functionality of two promoters, P_{fpr} and P_{ureA} , identified based on an *in silico* analysis, was tested in *E. coli* and *B. cenocepacia*. These two putative promoters are present upstream of genes that might be important for iron acquisition or virulence and have conserved residues that are important for efficient utilisation of promoters by OrbS (K. Agnoli, 2007). Although P_{fpr} seemed to be regulated by OrbS under inducing conditions in *E. coli*, this was not observed in *B. cenocepacia*. Further, P_{ureA} was found to be completely inactive in *B. cenocepacia*. Therefore, of the two candidate promoters P_{fpr} seems to be OrbS independent while P_{ureA} may not be a promoter at all. A genome-wide scan using ChIP-seq would reveal presence of novel OrbS-dependent promoters. However, due to issues with molecular cloning before carrying out the immunoprecipitation, and time restrictions, the experiment could not be carried out.

5.6.3 Regulation of OrbS activity

The ferric uptake regulator (Fur) is a master regulator that acts in response to iron concentration (Fillat, 2014). *orbS* transcription has been previously established to be regulated by Fur. However, unlike conventional iron starvation ECF σ factors, OrbS does not have an anti- σ factor that regulates its activity (Agnoli et al., 2006). Therefore, one of the research objectives was to assess the presence of an alternative mechanism for regulating OrbS activity. To address this, a previously constructed *B. cenocepacia fur* mutant (715j-*fur*::Tp) was used. Characterization of *fur* mutants is primarily carried out by testing the ability to over-produce siderophore detected using CAS agar (Ebanks et al., 2013, Smith et al., 2013). The construction of a *Burkholderia fur* mutant was first reported in *Burkholderia multivorans* and was characterized by testing its sensitivity to reactive oxygen species, carbon assimilation abilities and the intracellular iron concentrations (Yuhara et al., 2008, Kimura et al., 2012). In this project 715j-*fur*::Tp was characterized by testing its ability to produce siderophore using the CAS agar plate

assay, testing its efficiency of plating and analysis of growth rates compared to WT and finally by comparing the activity of Fur-regulated promoters under iron deficient and iron-replete conditions in WT and *fur* mutant strains. The CAS agar plate assay results demonstrated that 715j-*fur*::Tp had lost regulation by Fur and hence over-produced ornibactin. This phenotype was complementable when *fur* was introduced to the mutant *in trans*. Loss of functional *fur* reduced the ability of the *fur* mutant to produce colonies by an approximate ten-fold on in nutritionally rich and poor media compared to WT. In broth culture it could grow at similar rate as the WT but the most pronounced differences were observed in BHI and M9 (CAA 1%) media. In all types of media, the *fur* mutant had a longer lag before reaching log phase and reached final stationary phase at a lower optical density compared to WT.

Measurement of activities of Fur-regulated promoters under low and high iron concentrations in WT and *fur* mutant strains indicated that the mutant was behaving only partially as expected. Under high iron conditions although the activity of the Fur-regulated promoters in the *fur* mutant was higher than that in WT, they were not derepressed to levels observed in iron starvation conditions. The introduction of *fur* in the form of pBBR2-*fur*3 in the mutant, however, confirmed restoration of WT phenotype in the mutant.

Taken together, the results presented in section 5.3 point towards two plausible explanations. First, there exists, unique to *Burkholderia cenocepacia*, an unknown transcriptional regulatory component that downregulates transcription of iron-regulated transcription of iron-regulated genes. Hence, if this component works in combination with *fur*, it might not be able to carry out its regulation to the extent to that of WT when *fur* is absent, explaining the partial regulation. Second, the *fur* mutant itself retains some activity despite disruption of the *fur* gene by Tp^R cassette. These two conjectures could possibly be resolved by studying the Fur and OrbS-dependent promoter activities in a deletion mutant of *fur* (Yuhara et al., 2008). The *fur* mutant used in this study had a 649 bp Tp^R-cassette inserted between codon 72 and codon 73. Although such a large insertion would be expected to inactivate the functionality of the protein, the DNA binding helix and one metal binding site still could be functional within

the downstream 119 codons. As this insertion mutant may retain some Fur function construction of a *fur* deletion mutant was required. To this end a *B. cenocepacia fur* insertion mutant was constructed in which the 119 fur codons located downstream of the insertion site were removed and in this mutant the *fur* activity was demonstrated to be completely lost (A. Butt and M.S.Thomas, unpublished). This eliminates the first conjecture that an alternative transcriptional regulator of OrbS exists and hence the alternative regulator, if present, must be post-translational. However, as it remains, the data presented in this thesis was not sufficient to support or discard the hypothesis of the presence of an additional mechanism for regulating OrbS activity in response to iron availability.

Unfortunately, it was not possible to transfer the transcriptional reporter plasmid pKAGd4 containing the full-length P_{orbS} into the *fur* mutant. One possible reason for this could be that the derepressed P_{orbS} might lead to excessive β -galactosidase activity which could be toxic to cells or it may destabilise plasmid replication. One can test the first possibility by cloning the promoter P_{orbS} into the plasmid in the opposite orientation. Another alternative would be to use a modified/different *lacZ*-reporter plasmid that specifies lower levels of β -galactosidase activity. However, it was decided to use a shorter version of the *orbS* promoter which was still iron dependent.

The promoter activities of the OrbS-dependent promoter P_{orbH} in the full-length form and as a minimal derivative, $P_{orbHds6}$, were measured in 715j-*fur*::Tp and 715j. Results from this experiment showed iron dependent regulation even in the absence of *fur*. But, since the *fur* mutant retains a degree of iron regulation of Fur-dependent promoters it is difficult to conclude whether this regulation was due to background Fur activity or whether an alternative regulatory mechanism of OrbS exists.

Based on previous experiments carried out in our group, one possible mechanism for regulating OrbS activity is that OrbS itself can sense the intracellular iron concentration. There are four cysteine residues at the C-terminal end of OrbS that are conserved in other closely related iron starvation σ factors of *Burkholderia* such as MbaS from *B. pseudomallei* (K. Agnoli and M. S. Thomas, unpublished). The sulphur of the thiol groups is capable of complexing with metal ions such as iron (White and Dillingham,

2012). Experiments with *orbS* and *fur* deletion mutants that carry a mutant copy of *orbS* encoding cysteine to alanine substitutions in the C-terminal domain of OrbS suggest that these cysteine residues act like a fine tuning regulatory mechanism working in combination with Fur which is the major regulatory component (A. Butt and M. S. Thomas, unpublished). Evidence in literature shows that thiol groups in proteins are required to specifically sense metals, such as those from the MerR family that use cysteine residues within the C-terminal metal binding loops (Changela et al., 2003, Watanabe et al., 2008, Waldron et al., 2009). The transcriptional activator, WhiB7 and the FNR (Fumarate nitrate reduction) transcriptional regulator contain conserved cysteine residues that are used in the formation of iron sulphur clusters to carry out their regulatory functions (den Hengst and Buttner, 2008, Burian et al., 2013, Crack et al., 2013, Bush et al., 2016). Auto-regulation by ECF σ factors using a fused on-board regulator at their C-terminal end has been previously reported (Thakur et al., 2007, Gomez-Santos et al., 2011, Wecke et al., 2012, Mascher, 2013). Studies on the copper regulatory ECF σ factor, CorE, and cadmium and zinc regulatory ECF σ factor, CorE2, from *Myxococcus xanthus* provide evidence that metal detecting σ factors also bind metal ions through a cysteine-rich domain. CorE, like OrbS, does not seem to have a cognate anti- σ factor, but has a C-terminal extension with 6 cysteine residues. It recognises the redox state of copper through at least 4 cysteine residues on this C-terminal domain and initiates transcription of genes involved in copper homeostasis (Gomez-Santos et al., 2011). Different numbers and arrangements of the cysteine residues may be used selectively for detecting different metals (Marcos-Torres et al., 2016).

Alternatively, as direct binding of iron to the C-terminal cysteine thiols of OrbS has not yet been demonstrated, one can also speculate the presence of an altogether different iron regulatory protein.

Chapter 6. Final discussion

6.1 General discussion

The primary aims of this project were to identify the roles of four, currently uncharacterised putative ECF σ factors in *B. cenocepacia* and to further characterise the promoter recognition requirements and regulation mechanisms of the iron-starvation σ factor, OrbS. The approach taken to identify the roles of the uncharacterised ECF σ factors was two-fold: (1) identification of the σ -dependent promoters that would provide an insight into the role of the σ factor and (2) identification of the phenotypes associated with the deletion of the σ factor (or anti- σ factor) encoding gene.

Of the four previously uncharacterised ECF σ factors, the promoters recognised by three were identified in this study. Using ChIP-seq and reporter-fusion assays the PrtI-dependent promoters P_{prtI} and P_{cat} (putative catalase promoter) were identified. Additionally, the -10 and -35 promoter elements were experimentally determined as 'GTC' and 'GGAATA', respectively, for P_{prtI} . Ecf41_{Bc1} was found to recognise P_{ahpD} (promoter for a putative alkyl hydroperoxidase D encoding gene) while Ecf41_{Bc2} recognised P_{lysR} [promoter of a putative LysR-type transcriptional regulator (LTTR)]. However, the putative Ecf42_{Bc}-dependent promoter (P_{phnB2}) was found to be inactive in the presence of both the full-length and C-terminally truncated Ecf42_{Bc}. Hence it cannot be concluded to be Ecf42_{Bc}-dependent. If the identified Ecf42_{Bc}-dependent promoter does not have the required conditions to be active, on the experimental setup used here, one limitation is that the bioinformatics scan that identified more promoter motifs on the *B. cenocepacia* genome cannot be experimentally tested.

These results provide evidence that PrtI regulates a gene encoding a putative catalase and Ecf41_{Bc1} regulates one encoding a putative alkylhydroperoxidase D suggesting a role for these σ factors in oxidative stress responses. Ecf41_{Bc2} regulates a putative LTTR and although the Ecf42_{Bc}-dependent promoter could not be ascertained, the genes encoding both of these σ factors are present in the vicinity of genes encoding proteins are also likely to be involved in oxidative stress responses as described in Chapter 4. Moreover, two genes that are immediately present downstream of the LTTR encoding gene are present in the same orientation are speculated to be co-transcribed along with LTTR and thus can be Ecf41_{Bc2}-dependent. One of these, *I35_RS29625*

encodes a putative peroxiredoxin (OcmC). This gene (*BCAM2753* in *B. cenocepacia* J2315) was found to be upregulated in bacterial cells growing in CF sputum based model of infection (Drevinek et al., 2008). Additionally, one of the ECF41 σ factors in *Mycobacterium tuberculosis*, SigJ, has been shown to be involved in hydrogen peroxide resistance. Reactive oxygen species (ROS) such as hydrogen peroxide (H_2O_2), superoxide ($O_2^{\cdot-}$) and hydroxyl radicals ($OH\cdot$) are produced as a result of biochemical reactions during metabolic processes and are toxic, as they can damage cell components and biomolecules like DNA and proteins (Imlay, 2015). Therefore, bacteria need to scavenge the ROS produced inside cells in addition to those present in their local environment, including the oxidative burst generated by host immune cells during infections (Grant and Hung, 2013, Nathan and Cunningham-Bussel, 2013). Oxidative stress is one of the commonly known extra-cytoplasmic stresses that ECF σ factors respond to and multiple ECF σ factors have been experimentally shown to be involved in such stress responses (Martin et al., 1994, Paget et al., 1998, Alvarez-Martinez et al., 2007, Flannagan and Valvano, 2008).

Thus, each of the σ -factor and anti- σ factor mutants, as well as the double *ecf41_{BC1}-ecf41_{BC2}* deletion mutant, constructed in this project were tested for their ability to withstand oxidative stress. However, they were found to be as equally sensitive as the WT strain to oxidative stressors. Taking this further, the possible role of the PrtI-dependent *I35_RS16285* gene product as a catalase was explored. However, no such activity was observed with the purified enzyme *in vitro*. In addition, PrtI did not seem to have a role in protease production, biofilm formation, motility and exopolysaccharide production unlike PrtI proteins in *Pseudomonas spp.* Thus, the exact stress that each of the four novel σ factors responds to remains unknown.

Bacterial two-hybrid assays were carried out to demonstrate that PrtR interacts with PrtI through their N- and C-terminal (σ_4) domains, respectively. Additionally, PrtR was demonstrated to be inhibitory to PrtI activity. Anti- σ factors usually have an anti- σ domain (ASD) which interacts with the ECF σ factor and keeps it inactive. ASD can belong to class I or class II. PrtR is predicted to contain an ASD-I domain and therefore can be speculated to interact with σ_2 and σ_4 of PrtI such that it remains inactive under non-

inducing conditions similar to the situation when RseA binds SigE in *E. coli* (Campbell et al., 2003). A more recent mode of σ factor activity regulation by a fused on-board domain is proposed in ECF41 σ factors where a DGGG motif is required for the activity of the σ factor (Wecke et al., 2012). However, results presented in Chapter 4 show that this DGGG motif may be inhibitory in case of Ecf41_{Bc1} and Ecf41_{Bc2}. Ecf41_{Bc1} and Ecf41_{Bc2} showed different activities when different lengths of the C-terminal extension were deleted in the experimental design where they were ectopically expressed. Therefore, a homogeneous model for the regulatory role of the C-terminal extension in the Ecf41 σ factors of *B. cenocepacia* still cannot be concluded.

Finally, as presented in Chapter 5, evidence from experiments performed in *B. cenocepacia* unequivocally showed the DNA sequence requirements for efficient promoter utilisation by OrbS. TAAA and CGTC present at positions -33 to -30 and -12 to -9 with respect to the transcription start site of P_{OrbH} form the -10 and -35 promoter elements that OrbS recognises confirming the previous results in *E. coli*. In addition, OrbS requires a GC-rich 17 bp spacer whose length is critical and the conformation of bases at the transcription start site are important for efficient promoter utilisation by OrbS. Disrupting the *fur* gene in *B. cenocepacia* by insertion of an antibiotic resistance cassette did not result in complete abolition of iron regulation of Fur-regulated promoters. It is concluded that the disrupted *fur* gene retains partial activity, because when the bases present downstream of the insertion were later removed a total loss of *fur* activity was observed (A. Butt and M. S. Thomas, unpublished). Hence, currently the evidence is insufficient to support or discard the hypothesis of the presence of a post-translational regulator for OrbS, based on the experiments presented herein.

6.2 Future perspectives

Even though experimental evidence presented in this thesis suggests that the identified promoters were active in the presence of the respective σ factor, a fundamental experiment to demonstrate that each of these ECF σ factors can associate with the core RNAP and promote transcription would show that they function as σ factors. This can be achieved by allowing the purified σ to form a holoenzyme with core

RNAP followed by electrophoretic mobility shift assay (EMSA) or by carrying out *in vitro* transcription assays with the respective σ -dependent promoter. Recently, such an *in vitro* transcription assay has been carried out to demonstrate that PrtI can carry out transcription using the PrtI-dependent promoters identified in this study (A. Butt and M.S. Thomas, unpublished).

The effects of single/multiple base substitutions on the activities promoters recognised by Ecf41_{Bc1} and Ecf41_{Bc2} can be assayed that will reveal how these σ factors interact with their promoters, and identify the role of specific bases within the promoter. A putative target promoter of Ecf41_{Bc1} (for a putative LTTR gene) identified by the *in silico* analysis needs to be tested experimentally in the presence of Ecf41_{Bc}. The transcription start sites of each of the σ -dependent genes also can be determined to confirm the precise locations of the promoter elements recognised by the respective σ factors. Further, RT-PCR can be used to define the extent of the operons governed by each of the σ -factors. While a few unsuccessful attempts were made to define the PrtI operon by RT-PCR, the appropriate DNase steps need to be optimised to remove DNA contamination.

The role of thiocyanate in the PrtI system would be worth testing since the product of the gene regulated by PrtI, I35_RS16285 is homologous to SrpA which in turn is regulated by CysB, a regulatory protein involved in thiocyanate uptake in *Synechococcus*. Thiocyanate can be used as an alternative sulphur source but it is also toxic, so PrtI could be involved in sulphur assimilation or thiocyanate resistance/detoxification (Stafford and Callely, 1969, Vu et al., 2013). Studying the structure of the CTD of PrtR might provide an insight into the nature of the molecule that might bind and induce this system.

The effect of deleting *ecf42_{Bc}* on biofilm formation and antibiotic drug resistance should be tested as its homologue in *P. putida* has been shown to be involved in these phenotypes (Tettmann et al., 2014). A high-throughput approach to scan for a phenotype of the σ -factor and *anti- σ* factor mutants, as well as the double *ecf41_{Bc1}-ecf41_{Bc2}* deletion mutant can be adopted using a phenotype microarray that tests for hundreds of characteristics including nutrient utilisation, stress sensitivity, antibiotic

sensitivity etc. (Bochner, 2003). Additionally, the effect of deleting the σ -factor and *anti- σ factor* mutants, as well as the double *ecf41_{Bc1}-ecf41_{Bc2}* deletion mutant on the virulence of *B. cenocepacia* would determine if they are required for pathogenicity. Examples of appropriate hosts to study *B. cenocepacia* infections include *Galleria wax* moth larvae, *Caenorhabditis elegans*, zebrafish and murine models amongst others (Kothe et al., 2003, Tomich et al., 2003, Seed and Dennis, 2008, Deng et al., 2009). Additionally, macrophage cell lines can be used as *B. cenocepacia* has been previously shown to interact with immune cells of the host system in formation of vacuoles (BcCV) to delay the host response in clearing infections (Keith et al., 2009).

OrbS activity has been shown to be regulated via its C-terminal domain that functions as an on-board regulatory module (A. Butt and M. S. Thomas, unpublished). Studies on *orbS* and *fur* deletion mutants have indicated that while *orbS* is primarily regulated by Fur, a fine tuning regulatory mechanism consisting of four cysteine residues at the C-terminal domain of OrbS is present that responds directly to iron concentration (A. Butt and M. S. Thomas, unpublished). It would be insightful to demonstrate that OrbS can co-ordinate with Fe²⁺ to form iron-sulphur clusters using mass spectrometry (LC-MS) or reconstituting OrbS with iron and measuring the absorbance using UV-spectrometry.

To conclude, the work described in this thesis lays the foundation for further study of these previously uncharacterised ECF σ factors of *B. cenocepacia*. In the wider perspective, it would be ideal to fully characterise all the ECF σ factors of *B. cenocepacia* to define their regulons, roles and regulation. Ultimately, it would be useful to determine if their functions overlap, if the σ -dependent promoters are orthogonal or whether any cross-talk occurs. This will be useful to gain a fundamental understanding of the complexities of signal transduction mechanisms occurring inside bacterial cells. As an example, using an integrated approach in *Mycobacterium tuberculosis*, a study has revealed interesting features such as the hierarchy in the regulatory network dynamics and internal connectivity of σ factors of *M. tuberculosis* (Chauhan et al., 2016).

A more application-based purpose of carrying forward this research would be in studying any potential contribution of the ECF σ factors to the pathogenicity (i.e. their significance as virulence determinants) in *B. cenocepacia*, which is known to cause

persistent and problematic infections in CF patients. This will help determine if ECF σ factors are effective anti-microbial targets to develop better drugs in the future.

References

- AGNOLI, K., LOWE, C. A., FARMER, K. L., HUSNAIN, S. I. & THOMAS, M. S. 2006. The ornibactin biosynthesis and transport genes of *Burkholderia cenocepacia* are regulated by an extracytoplasmic function factor which is a part of the fur regulon. *Journal of Bacteriology*, 188, 3631-3644.
- AGNOLI, K., PhD thesis, 2007, University of Sheffield.
- AHMED, E. & HOLMSTROM, S. J. M. 2014. Siderophores in environmental research: roles and applications. *Microbial Biotechnology*, 7, 196-208.
- ALVAREZ-MARTINEZ, C. E., LOURENCO, R. F., BALDINI, R. L., LAUB, M. T. & GOMES, S. L. 2007. The ECF sigma factor sigma(T) is involved in osmotic and oxidative stress responses in *Caulobacter crescentus*. *Molecular Microbiology*, 66, 1240-1255.
- AUBERT, D. F., XU, H., YANG, J. L., SHI, X. Y., GAO, W. Q., LI, L., BISARO, F., CHEN, S., VALVANO, M. A. & SHAO, F. 2016. A *Burkholderia* Type VI Effector Deamidates Rho GTPases to Activate the Pyrin Inflammasome and Trigger Inflammation. *Cell Host & Microbe*, 19, 664-674.
- BAILEY, T., KRAJEWSKI, P., LADUNGA, I., LEFEBVRE, C., LI, Q. H., LIU, T., MADRIGAL, P., TASLIM, C. & ZHANG, J. 2013. Practical Guidelines for the Comprehensive Analysis of ChIP-seq Data. *Plos Computational Biology*, 9.
- BALASUBRAMANIAN, D., KONG, K. F., JAYAWARDENA, S. R., LEAL, S. M., SAUTTER, R. T. & MATHEE, K. 2011. Co-regulation of beta-lactam resistance, alginate production and quorum sensing in *Pseudomonas aeruginosa*. *Journal of Medical Microbiology*, 60, 147-156.
- BANIN, E., VASIL, M. L. & GREENBERG, E. P. 2005. Iron and *Pseudomonas aeruginosa* biofilm formation. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 11076-11081.
- BAR-NAHUM, G. & NUDLER, E. 2001. Isolation and characterization of sigma(70)-retaining transcription elongation complexes from *Escherichia coli*. *Cell*, 106, 443-451.
- BARINOVA, N., KUZNEDELOV, K., SEVERINOV, K. & KULBACHINSKIY, A. 2008. Structural modules of RNA polymerase required for transcription from promoters containing downstream basal promoter element GGGA. *Journal of Biological Chemistry*, 283, 22482-22489.
- BARRETT, A. R., KANG, Y., INAMASU, K. S., SON, M. S., VUKOVICH, J. M. & HOANG, T. T. 2008. Genetic tools for allelic replacement in *Burkholderia* species. *Applied and Environmental Microbiology*, 74, 4498-4508.
- BARRY, S. M. & CHALLIS, G. L. 2009. Recent advances in siderophore biosynthesis. *Current Opinion in Chemical Biology*, 13, 205-215.
- BASHYAM, M. D. & HASNAIN, S. E. 2004. The extracytoplasmic function sigma factors: role in bacterial pathogenesis. *Infection Genetics and Evolution*, 4, 301-308.
- BASTIAANSEN, K. C., OTERO-ASMAN, J. R., LUIRINK, J., BITTER, W. & LLAMAS, M. A. 2015a. Processing of cell-surface signalling anti-sigma factors prior to signal recognition is a conserved autoproteolytic mechanism that produces two functional domains. *Environmental Microbiology*, 17, 3263-3277.
- BASTIAANSEN, K. C., VAN ULSEN, P., WIJTMANS, M., BITTER, W. & LLAMAS, M. A. 2015b. Self-cleavage of the *Pseudomonas aeruginosa* Cell-surface Signaling Anti-sigma Factor FoxR Occurs through an N-O Acyl Rearrangement. *Journal of Biological Chemistry*, 290, 12237-12246.
- BECKMAN, W. & LESSIE, T. G. 1979. RESPONSE OF PSEUDOMONAS-CEPACIA TO BETA-LACTAM ANTIBIOTICS - UTILIZATION OF PENICILLIN-G AS THE CARBON SOURCE. *Journal of Bacteriology*, 140, 1126-1128.
- BERNIER, S. P., WORKENTINE, M. L., LI, X., MAGARVEY, N. A., O'TOOLE, G. A. & SURETTE, M. G.

2016. Cyanide Toxicity to Burkholderia cenocepacia Is Modulated by Polymicrobial Communities and Environmental Factors. *Frontiers in Microbiology*, 7.
- BIVILLE, F., CWERMAN, H., LETOFFE, S., ROSSI, M. S., DROUET, V., GHIGO, J. M. & WANDERSMAN, C. 2004. Haemophore-mediated signalling in *Serratia marcescens*: a new mode of regulation for an extra cytoplasmic function (ECF) sigma factor involved in haem acquisition. *Molecular Microbiology*, 53, 1267-1277.
- BLATCH, G. L. & LASSLE, M. 1999. The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *Bioessays*, 21, 932-939.
- BOCHKAREVA, A. & ZENKIN, N. 2013. The Sigma(70) region 1.2 regulates promoter escape by unwinding DNA downstream of the transcription start site. *Nucleic Acids Research*, 41, 4565-4572.
- BOCHNER, B. R. 2003. New technologies to assess genotype-phenotype relationships. *Nature Reviews Genetics*, 4, 309-314.
- BOOR, K. J. 2006. Bacterial stress responses: What doesn't kill them can make them stronger. *Plos Biology*, 4, 18-20.
- BRAUN, V. 2003. Iron uptake by *Escherichia coli*. *Frontiers in Bioscience*, 8, S1409-S1421.
- BRAUN, V. & MAHREN, S. 2005. Transmembrane transcriptional control (surface signalling) of the *Escherichia coli* Fec type. *Fems Microbiology Reviews*, 29, 673-684.
- BRAUN, V., MAHREN, S. & OGIERMAN, M. 2003. Regulation of the FecI-type ECF sigma factor by transmembrane signalling. *Current Opinion in Microbiology*, 6, 173-180.
- BRAUN, V., MAHREN, S. & SAUTER, A. 2005. Gene regulation by transmembrane signaling. *Biomaterials*, 18, 507-517.
- BROOKS, B. E. & BUCHANAN, S. K. 2008. Signaling mechanisms for activation of extracytoplasmic function (ECF) sigma factors. *Biochimica Et Biophysica Acta-Biomembranes*, 1778, 1930-1945.
- BRYK, R., LIMA, C. D., ERDJUMENT-BROMAGE, H., TEMPST, P. & NATHAN, C. 2002. Metabolic enzymes of mycobacteria linked to antioxidant defense by a thioredoxin-like protein. *Science*, 295, 1073-1077.
- BULL, T. J., LINEDALE, R., HINDS, J. & HERMON-TAYLOR, J. 2009. A rhodanine agent active against non-replicating intracellular *Mycobacterium avium* subspecies paratuberculosis. *Gut Pathogens*, 1.
- BURGER, M., WOODS, R. G., MCCARTHY, C. & BEACHAM, I. R. 2000. Temperature regulation of protease in *Pseudomonas fluorescens* LS107d2 by an ECF sigma factor and a transmembrane activator. *Microbiology-Uk*, 146, 3149-3155.
- BURGESS, R. R. & ANTHONY, L. 2001. How sigma docks to RNA polymerase and what sigma does. *Current Opinion in Microbiology*, 4, 126-131.
- BURIAN, J., YIM, G., HSING, M., AXERIO-CILIES, P., CHERKASOV, A., SPIEGELMAN, G. B. & THOMPSON, C. J. 2013. The mycobacterial antibiotic resistance determinant WhiB7 acts as a transcriptional activator by binding the primary sigma factor SigA (RpoV). *Nucleic Acids Research*, 41, 10062-10076.
- BURKHOLDER, W. H. 1950. SOUR SKIN, A BACTERIAL ROT OF ONION BULBS. *Phytopathology*, 40, 115-117.
- BUSBY, S. J. W. 2009. More pieces in the promoter jigsaw: recognition of -10 regions by alternative sigma factors. *Molecular Microbiology*, 72, 809-811.
- BUSH, M. J., CHANDRA, G., BIBB, M. J., FINDLAY, K. C. & BUTTNER, M. J. 2016. Genome-Wide Chromatin Immunoprecipitation Sequencing Analysis Shows that WhiB Is a Transcription Factor That Cocontrols Its Regulon with WhiA To Initiate Developmental Cell Division in *Streptomyces*. *Mbio*, 7.
- BUTCHER, B. G., BRONSTEIN, P. A., MYERS, C. R., STODGHILL, P. V., BOLTON, J. J., MARKEL, E. J., FILIATRAULT, M. J., SWINGLE, B., GABALLA, A., HELMANN, J. D., SCHNEIDER, D. J. &

- CARTINHOOR, S. W. 2011. Characterization of the Fur Regulon in *Pseudomonas syringae* pv. tomato DC3000. *Journal of Bacteriology*, 193, 4598-4611.
- BUTCHER, B. G., MASCHER, T. & HELMANN, J. D. 2008. Environmental sensing and the role of extracytoplasmic function sigma factors. *Bacterial Physiology: A Molecular Approach*, 233-261.
- BYLUND, J., BURGESS, L. A., CESCUTTI, P., ERNST, R. K. & SPEERT, D. P. 2006. Exopolysaccharides from *Burkholderia cenocepacia* inhibit neutrophil chemotaxis and scavenge reactive oxygen species. *Journal of Biological Chemistry*, 281, 2526-2532.
- CALLACI, S., HEYDUK, E. & HEYDUK, T. 1999. Core RNA polymerase from E-coli induces a major change in the domain arrangement of the sigma(70) subunit. *Molecular Cell*, 3, 229-238.
- CAMPAGNE, S., ALLAIN, F. H. T. & VORHOLT, J. A. 2015. Extra Cytoplasmic Function sigma factors, recent structural insights into promoter recognition and regulation. *Current Opinion in Structural Biology*, 30, 71-78.
- CAMPAGNE, S., MARSH, M. E., CAPITANI, G., VORHOLT, J. A. & ALLAIN, F. H. T. 2014. Structural basis for -10 promoter element melting by environmentally induced sigma factors. *Nature Structural & Molecular Biology*, 21, 269-276.
- CAMPBELL, E. A., MUZZIN, O., CHLENOV, M., SUN, J. L., OLSON, C. A., WEINMAN, O., TRESTER-ZEDLITZ, M. L. & DARST, S. A. 2002. Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Molecular Cell*, 9, 527-539.
- CAMPBELL, E. A., TUPY, J. L., GRUBER, T. M., WANG, S., SHARP, M. M., GROSS, C. A. & DARST, S. A. 2003. Crystal structure of *Escherichia coli* sigma(E) with the cytoplasmic domain of its anti-sigma RseA. *Molecular Cell*, 11, 1067-1078.
- CARAHER, E., REYNOLDS, G., MURPHY, P., MCCLEAN, S. & CALLAGHAN, M. 2007. Comparison of antibiotic susceptibility of *Burkholderia cepacia* complex organisms when grown planktonically or as biofilm in vitro. *European Journal of Clinical Microbiology & Infectious Diseases*, 26, 213-216.
- CASADABAN, M. J. & COHEN, S. N. 1979. LACTOSE GENES FUSED TO EXOGENOUS PROMOTERS IN ONE-STEP USING A MU-LAC BACTERIOPHAGE - INVIVO PROBE FOR TRANSCRIPTIONAL CONTROL SEQUENCES. *Proceedings of the National Academy of Sciences of the United States of America*, 76.
- CELIA, H., NOINAJ, N., ZAKHAROV, S. D., BORDIGNON, E., BOTOS, I., SANTAMARIA, M., BARNARD, T. J., CRAMER, W. A., LLOUBES, R. & BUCHANAN, S. K. 2016. Structural insight into the role of the Ton complex in energy transduction. *Nature*, 538, 60-+.
- CERVENY, L., STRASKOVA, A., DANKOVA, V., HARTLOVA, A., CECKOVA, M., STAUD, F. & STULIK, J. 2013. Tetratricopeptide Repeat Motifs in the World of Bacterial Pathogens: Role in Virulence Mechanisms. *Infection and Immunity*, 81, 629-635.
- CHANGELA, A., CHEN, K., XUE, Y., HOLSCHEN, J., OUTTEN, C. E., O'HALLORAN, T. V. & MONDRAGON, A. 2003. Molecular basis of metal-ion selectivity and zeptomolar sensitivity by CueR. *Science*, 301, 1383-1387.
- CHARALABOUS, P., RISK, J. M., JENKINS, R., BIRSS, A. J., HART, C. A. & SMALLEY, J. W. 2007. Characterization of a bifunctional catalase-peroxidase of *Burkholderia cenocepacia*. *Fems Immunology and Medical Microbiology*, 50, 37-44.
- CHAUHAN, R., RAVI, J., DATTA, P., CHEN, T. L., SCHNAPPINGER, D., BASSLER, K. E., BALAZSI, G. & GENNARO, M. L. 2016. Reconstruction and topological characterization of the sigma factor regulatory network of *Mycobacterium tuberculosis*. *Nature Communications*, 7.
- CHEN, J., DARST, S. A. & THIRUMALAI, D. 2010. Promoter melting triggered by bacterial RNA polymerase occurs in three steps. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 12523-12528.
- CHEN, Y., HOLTMAN, C. K., MAGNUSON, R. D., YODERIAN, P. A. & GOLDEN, S. S. 2008. The

- complete sequence and functional analysis of pANL, the large plasmid of the unicellular freshwater cyanobacterium *Synechococcus elongatus* PCC 7942. *Plasmid*, 59, 176-192.
- CHOBY, J. E. & SKAAR, E. P. 2016. Heme Synthesis and Acquisition in Bacterial Pathogens. *Journal of Molecular Biology*, 428, 3408-3428.
- COENYE, T. & VANDAMME, P. 2003. Diversity and significance of Burkholderia species occupying diverse ecological niches. *Environmental Microbiology*, 5, 719-729.
- CONDON, B. J., OIDE, S., GIBSON, D. M., KRASNOFF, S. B. & TURGEON, B. G. 2014. Reductive Iron Assimilation and Intracellular Siderophores Assist Extracellular Siderophore-Driven Iron Homeostasis and Virulence. *Molecular Plant-Microbe Interactions*, 27, 793-808.
- CONWAY, B. A. D., CHU, K. K., BYLUND, J., ALTMAN, E. & SPEERT, D. P. 2004. Production of exopolysaccharide by Burkholderia cenocepacia results in altered cell-surface interactions and altered bacterial clearance in mice. *Journal of Infectious Diseases*, 190, 957-966.
- CORNELIS, P., MATTHIJS, S. & VAN OEFFELEN, L. 2009. Iron uptake regulation in Pseudomonas aeruginosa. *Biometals*, 22, 15-22.
- CRACK, J. C., STAPLETON, M. R., GREEN, J., THOMSON, A. J. & LE BRUN, N. E. 2013. Mechanism of 4Fe-4S (Cys)₄ Cluster Nitrosylation Is Conserved among NO-responsive Regulators. *Journal of Biological Chemistry*, 288, 11492-11502.
- CROSA, J. H. & WALSH, C. T. 2002. Genetics and assembly line enzymology of siderophore biosynthesis in bacteria. *Microbiology and Molecular Biology Reviews*, 66, 223-+.
- CROUCH, M. L., BECKER, L. A., BANG, I. S., TANABE, H., OUELLETTE, A. J. & FANG, F. C. 2005. The alternative sigma factor sigma(E) is required for resistance of Salmonella enterica serovar Typhimurium to anti-microbial peptides. *Molecular Microbiology*, 56, 789-799.
- CUNHA, M. V., SOUSA, S. A., LEITAO, J. H., MOREIRA, L. M., VIDEIRA, P. A. & SA-CORREIA, I. 2004. Studies on the involvement of the exopolysaccharide produced by cystic fibrosis-associated isolates of the Burkholderia cepacia complex in biofilm formation and in persistence of respiratory infections. *Journal of Clinical Microbiology*, 42, 3052-3058.
- D'AES, J., DE MAEYER, K., PAUWELYN, E. & HOFTE, M. 2010. Biosurfactants in plant-Pseudomonas interactions and their importance to biocontrol. *Environmental Microbiology Reports*, 2, 359-372.
- D'ANDREA, L. D. & REGAN, L. 2003. TPR proteins: the versatile helix. *Trends in Biochemical Sciences*, 28, 655-662.
- DARLING, P., CHAN, M., COX, A. D. & SOKOL, P. A. 1998. Siderophore production by cystic fibrosis isolates of Burkholderia cepacia. *Infection and Immunity*, 66.
- DARST, S. A., FEKLISTOV, A. & GROSS, C. A. 2014. Promoter melting by an alternative sigma, one base at a time. *Nature Structural & Molecular Biology*, 21, 350-351.
- DAVIES, B. W., BOGARD, R. W. & MEKALANOS, J. J. 2011. Mapping the regulon of Vibrio cholerae ferric uptake regulator expands its known network of gene regulation. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 12467-12472.
- DAVIS, C. A., BINGMAN, C. A., LANDICK, R., RECORD, M. T., JR. & SAECKER, R. M. 2007. Real-time footprinting of DNA in the first kinetically significant intermediate in open complex formation by Escherichia coli RNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 7833-7838.
- DAVIS, M. C., SMITH, L. K. & MACLELLAN, S. R. 2016. The atypical two-subunit sigma factor from Bacillus subtilis is regulated by an integral membrane protein and acid stress. *Microbiology-Sgm*, 162, 398-407.
- DE BRUIJN, I., DE KOCK, M. J. D., YANG, M., DE WAARD, P., VAN BEEK, T. A. & RAAIJMAKERS, J. M. 2007. Genome-based discovery, structure prediction and functional analysis of

- cyclic lipopeptide antibiotics in *Pseudomonas* species. *Molecular Microbiology*, 63, 417-428.
- DEN HENGST, C. D. & BUTTNER, M. J. 2008. Redox control in actinobacteria. *Biochimica Et Biophysica Acta-General Subjects*, 1780, 1201-1216.
- DENG, Y. Y., BOON, C., EBERL, L. & ZHANG, L. H. 2009. Differential Modulation of Burkholderia cenocepacia Virulence and Energy Metabolism by the Quorum-Sensing Signal BDSF and Its Synthase. *Journal of Bacteriology*, 191, 7270-7278.
- DENG, Y. Y., WU, J. E., EBERL, L. & ZHANG, L. H. 2010. Structural and Functional Characterization of Diffusible Signal Factor Family Quorum-Sensing Signals Produced by Members of the Burkholderia cepacia Complex. *Applied and Environmental Microbiology*, 76, 4675-4683.
- DEPOORTER, E., BULL, M. J., PEETERS, C., COENYE, T., VANDAMME, P. & MAHENTHIRALINGAM, E. 2016. Burkholderia: an update on taxonomy and biotechnological potential as antibiotic producers. *Applied Microbiology and Biotechnology*, 100, 5215-5229.
- DOBRTSA, A. P. & SAMADPOUR, M. 2016. Transfer of eleven species of the genus Burkholderia to the genus Paraburkholderia and proposal of Caballeronia gen. nov to accommodate twelve species of the genera Burkholderia and Paraburkholderia. *International Journal of Systematic and Evolutionary Microbiology*, 66, 2836-2846.
- DREVINEK, P., HOLDEN, M. T. G., GE, Z. P., JONES, A. M., KETCHELL, I., GILL, R. T. & MAHENTHIRALINGAM, E. 2008. Gene expression changes linked to antimicrobial resistance, oxidative stress, iron depletion and retained motility are observed when Burkholderia cenocepacia grows in cystic fibrosis sputum. *Bmc Infectious Diseases*, 8.
- DREVINEK, P. & MAHENTHIRALINGAM, E. 2010. Burkholderia cenocepacia in cystic fibrosis: epidemiology and molecular mechanisms of virulence. *Clinical Microbiology and Infection*, 16, 821-830.
- EBANKS, R. O., GOGUEN, M., KNICKLE, L., DACANAY, A., LESLIE, A., ROSS, N. W. & PINTO, D. M. 2013. Analysis of a ferric uptake regulator (Fur) knockout mutant in *Aeromonas salmonicida* subsp *salmonicida*. *Veterinary Microbiology*, 162, 831-841.
- EBERL, L. 2006. Quorum sensing in the genus Burkholderia. *International Journal of Medical Microbiology*, 296, 103-110.
- ENZ, S., MAHREN, S., STROEHER, U. H. & BRAUN, V. 2000. Surface signaling in ferric citrate transport gene induction: Interaction of the FecA, FecR, and FecI regulatory proteins. *Journal of Bacteriology*, 182, 637-646.
- ESCOLAR, L., PEREZ-MARTIN, J. & DE LORENZO, V. 1998. Binding of the Fur (ferric uptake regulator) repressor of *Escherichia coli* to arrays of the GATAAT sequence. *Journal of Molecular Biology*, 283, 537-547.
- FAZLI, M., HARRISON, J. J., GAMBINO, M., GIVSKOV, M. & TOLKER-NIELSEN, T. 2015. In-Frame and Unmarked Gene Deletions in Burkholderia cenocepacia via an Allelic Exchange System Compatible with Gateway Technology. *Applied and Environmental Microbiology*, 81, 3623-3630.
- FAZLI, M., MCCARTHY, Y., GIVSKOV, M., RYAN, R. P. & TOLKER-NIELSEN, T. 2013. The exopolysaccharide gene cluster Bcam1330-Bcam1341 is involved in Burkholderia cenocepacia biofilm formation, and its expression is regulated by c-di-GMP and Bcam1349. *Microbiologyopen*, 2, 105-122.
- FEKLISTOV, A. 2013. RNA polymerase: in search of promoters. *Annals of the New York Academy of Sciences*, 1293, 25-32.
- FEKLISTOV, A. & DARST, S. A. 2011. Structural Basis for Promoter-10 Element Recognition by the Bacterial RNA Polymerase sigma Subunit. *Cell*, 147, 1257-1269.
- FEKLISTOV, A., SHARON, B. D., DARST, S. A. & GROSS, C. A. 2014. Bacterial Sigma Factors: A Historical, Structural, and Genomic Perspective. *Annual Review of Microbiology*, Vol

- 68, 68, 357-376.
- FERGUSON, A. D., CHAKRABORTY, R., SMITH, B. S., ESSER, L., VAN DER HELM, D. & DEISENHOFER, J. 2002. Structural basis of gating by the outer membrane transporter FecA. *Science*, 295, 1715-1719.
- FERREIRA, A. S., LEITAO, J. H., SILVA, I. N., PINHEIRO, P. F., SOUSA, S. A., RAMOS, C. G. & MOREIRA, L. M. 2010. Distribution of Cepacian Biosynthesis Genes among Environmental and Clinical Burkholderia Strains and Role of Cepacian Exopolysaccharide in Resistance to Stress Conditions. *Applied and Environmental Microbiology*, 76, 441-450.
- FILLAT, M. F. 2014. The FUR (ferric uptake regulator) superfamily: Diversity and versatility of key transcriptional regulators. *Archives of Biochemistry and Biophysics*, 546, 41-52.
- FIROVED, A. M., BOUCHER, J. C. & DERETIC, V. 2002. Global genomic analysis of AlgU (sigma(E))-dependent promoters (Sigmulon) in *Pseudomonas aeruginosa* and implications for inflammatory processes in cystic fibrosis. *Journal of Bacteriology*, 184, 1057-1064.
- FIROVED, A. M. & DERETIC, V. 2003. Microarray analysis of global gene expression in mucoid *Pseudomonas aeruginosa*. *Journal of Bacteriology*, 185, 1071-1081.
- FLANNAGAN, R. S. & VALVANO, M. A. 2008. Burkholderia cenocepacia requires RpoE for growth under stress conditions and delay of phagolysosomal fusion in macrophages. *Microbiology-Sgm*, 154, 643-653.
- FLYNN, J. M., LEVCHENKO, I., SAUER, R. T. & BAKER, T. A. 2004. Modulating substrate choice: the SspB adaptor delivers a regulator of the extracytoplasmic-stress response to the AAA plus protease ClpXP for degradation. *Genes & Development*, 18, 2292-2301.
- FRANCEZ-CHARLOT, A., FRUNZKE, J., REICHEN, C., EBNETER, J. Z., GOURION, B. & VORHOLT, J. A. 2009. Sigma factor mimicry involved in regulation of general stress response. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 3467-3472.
- FRANKE, J., ISHIDA, K., ISHIDA-ITO, M. & HERTWECK, C. 2013. Nitro versus Hydroxamate in Siderophores of Pathogenic Bacteria: Effect of Missing Hydroxylamine Protection in Malleobactin Biosynthesis. *Angewandte Chemie-International Edition*, 52, 8271-8275.
- GALLI, F., BATTISTONI, A., GAMBARI, R., POMPELLA, A., BRAGONZI, A., PILOLLI, F., IULIANO, L., PIRODDI, M., DECHECCHI, M. C. & CABRINI, G. 2012. Oxidative stress and antioxidant therapy in cystic fibrosis. *Biochimica Et Biophysica Acta-Molecular Basis of Disease*, 1822, 690-713.
- GASKELL, A. A., CRACK, J. C., KELEMEN, G. H., HUTCHINGS, M. I. & LE BRUN, N. E. 2007. RsmA is an anti-sigma factor that modulates its activity through a 2Fe-2S cluster cofactor. *Journal of Biological Chemistry*, 282, 31812-31820.
- GIACANI, L., DENISENKO, O., TOMPA, M. & CENTURION-LARA, A. 2013. Identification of the *Treponema pallidum* subsp *pallidum* TP0092 (RpoE) Regulon and Its Implications for Pathogen Persistence in the Host and Syphilis Pathogenesis. *Journal of Bacteriology*, 195, 896-907.
- GOLDMANN, D. A. & KLINGER, J. D. 1986. PSEUDOMONAS-CEPACIA - BIOLOGY, MECHANISMS OF VIRULENCE, EPIDEMIOLOGY. *Journal of Pediatrics*, 108, 806-812.
- GOMEZ-SANTOS, N., PEREZ, J., CELESTINA SANCHEZ-SUTIL, M., MORALEDA-MUNOZ, A. & MUNOZ-DORADO, J. 2011. CorE from *Myxococcus xanthus* Is a Copper-Dependent RNA Polymerase Sigma Factor. *Plos Genetics*, 7.
- GOURSE, R. L., ROSS, W. & GAAL, T. 2000. UPs and downs in bacterial transcription initiation: the role of the alpha subunit of RNA polymerase in promoter recognition. *Molecular Microbiology*, 37, 687-695.
- GRANT, S. S. & HUNG, D. T. 2013. Persistent bacterial infections, antibiotic tolerance, and the

- oxidative stress response. *Virulence*, 4, 273-283.
- GRIGOROVA, I. L., CHABA, R., ZHONG, H. J., ALBA, B. M., RHODIUS, V., HERMAN, C. & GROSS, C. A. 2004. Fine-tuning of the Escherichia coli sigma(E) envelope stress response relies on multiple mechanisms to inhibit signal-independent proteolysis of the transmembrane anti-sigma factor, RseA. *Genes & Development*, 18, 2686-2697.
- GRUBER, T. M. & GROSS, C. A. 2003. Multiple sigma subunits and the partitioning of bacterial transcription space. *Annual Review of Microbiology*, 57, 441-466.
- GUZINA, J. & DJORDJEVIC, M. 2017. Mix-and- matching as a promoter recognition mechanism by ECF σ factors. *BMC Evolutionary Biology*, 17, 1-14.
- HALDIPURKAR, S., MSc. thesis, 2012, University of Sheffield.
- HALGREN, A., MASELKO, M., AZEVEDO, M., MILLS, D., ARMSTRONG, D. & BANOWETZ, G. 2013. Genetics of germination-arrest factor (GAF) production by Pseudomonas fluorescens WH6: identification of a gene cluster essential for GAF biosynthesis. *Microbiology-Sgm*, 159, 36-45.
- HANTKE, K. 2001. Iron and metal regulation in bacteria. *Current Opinion in Microbiology*, 4, 172-177.
- HASTIE, J. L., WILLIAMS, K. B. & ELLERMEIER, C. D. 2013. The Activity of sigma(V), an Extracytoplasmic Function sigma Factor of Bacillus subtilis, Is Controlled by Regulated Proteolysis of the Anti-sigma Factor RsiV. *Journal of Bacteriology*, 195, 3135-3144.
- HAUGEN, S. P., BERKMEN, M. B., ROSS, W., GAAL, T., WARD, C. & GOURSE, R. L. 2006. rRNA promoter regulation by nonoptimal binding of a region 1.2: An additional recognition element for RNA polymerase. *Cell*, 125, 1069-1082.
- HAUGEN, S. P., ROSS, W. & GOURSE, R. L. 2008a. Advances in bacterial promoter recognition and its control by factors that do not bind DNA. *Nat Rev Microbiol*, 6, 507-19.
- HAUGEN, S. P., ROSS, W., MANRIQUE, M. & GOURSE, R. L. 2008b. Fine structure of the promoter-sigma region 1.2 interaction. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 3292-3297.
- HAYDEN, J. D. & ADES, S. E. 2008. The Extracytoplasmic Stress Factor, sigma(E), Is Required to Maintain Cell Envelope Integrity in Escherichia coli. *Plos One*, 3.
- HEINRICH, J. & WIEGERT, T. 2009. Regulated intramembrane proteolysis in the control of extracytoplasmic function sigma factors. *Research in Microbiology*, 160, 696-703.
- HELMANN, J. D. 1999. Anti-sigma factors. *Current Opinion in Microbiology*, 2, 135-141.
- HELMANN, J. D. 2002. The extracytoplasmic function (ECF) sigma factors. *Advances in Microbial Physiology, Vol 46*, 46, 47-110.
- HELMANN, J. D. 2016. Bacillus subtilis extracytoplasmic function (ECF) sigma factors and defense of the cell envelope. *Current Opinion in Microbiology*, 30, 122-132.
- HEO, L., CHO, Y.-B., LEE, M. S., ROE, J.-H. & SEOK, C. 2013. Alternative zinc-binding sites explain the redox sensitivity of zinc-containing anti-sigma factors. *Proteins-Structure Function and Bioinformatics*, 81, 1644-1652.
- HERASIMENKA, Y., CESCUTTI, P., IMPALLOMENE, G., CAMPANA, S., TACCETTI, G., RAVENNI, N., ZANETTI, F. & RIZZO, R. 2007. Exopolysaccharides produced by clinical strains belonging to the Burkholderia cepacia complex. *Journal of Cystic Fibrosis*, 6, 145-152.
- HO, T. D. & ELLERMEIER, C. D. 2012. Extra cytoplasmic function sigma factor activation. *Current Opinion in Microbiology*, 15, 182-188.
- HOFMANN, B., HECHT, H. J. & FLOHE, L. 2002. Peroxiredoxins. *Biological Chemistry*, 383, 347-364.
- HOLDEN, M. T. G., SETH-SMITH, H. M. B., CROSSMAN, L. C., SEBAIHIA, M., BENTLEY, S. D., CERDENO-TARRAGA, A. M., THOMSON, N. R., BASON, N., QUAIL, M. A., SHARP, S., CHEREVACH, I., CHURCHER, C., GOODHEAD, I., HAUSER, H., HOLROYD, N., MUNGALL, K., SCOTT, P., WALKER, D., WHITE, B., ROSE, H., IVERSEN, P., MIL-HOMENS, D., ROCHA,

- E. P. C., FIALHO, A. M., BALDWIN, A., DOWSON, C., BARRELL, B. G., GOVAN, J. R., VANDAMME, P., HART, C. A., MAHENTHIRALINGAM, E. & PARKHILL, J. 2009. The Genome of *Burkholderia cenocepacia* J2315, an Epidemic Pathogen of Cystic Fibrosis Patients. *Journal of Bacteriology*, 191, 261-277.
- HONG, H. J., PAGET, M. S. B. & BUTTNER, M. J. 2002. A signal transduction system in *Streptomyces coelicolor* that activates the expression of a putative cell wall glycan operon in response to vancomycin and other cell wall-specific antibiotics. *Molecular Microbiology*, 44, 1199-1211.
- HOOK-BARNARD, I. G. & HINTON, D. M. 2007. Transcription initiation by mix and match elements: flexibility for polymerase binding to bacterial promoters. *Gene Regul Syst Bio*, 1, 275-93.
- HOOK-BARNARD, I. G. & HINTON, D. M. 2009. The promoter spacer influences transcription initiation via sigma(70) region 1.1 of *Escherichia coli* RNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 737-742.
- HORSLEY, A., JONES, A. M. & LORD, R. 2016. Antibiotic treatment for *Burkholderia cepacia* complex in people with cystic fibrosis experiencing a pulmonary exacerbation. *Cochrane Database of Systematic Reviews*.
- HU, Y. M., KENDALL, S., STOKER, N. G. & COATES, A. R. M. 2004. The *Mycobacterium tuberculosis sigJ* gene controls sensitivity of the bacterium to hydrogen peroxide. *Fems Microbiology Letters*, 237, 415-423.
- HUMPHREYS, S., STEVENSON, A., BACON, A., WEINHARDT, A. B. & ROBERTS, M. 1999. The alternative sigma factor, sigma(E), is critically important for the virulence of *Salmonella typhimurium*. *Infection and Immunity*, 67, 1560-1568.
- IMLAY, J. A. 2015. Diagnosing oxidative stress in bacteria: not as easy as you might think. *Current Opinion in Microbiology*, 24, 124-131.
- ISLES, A., MACLUSKY, I., COREY, M., GOLD, R., PROBER, C., FLEMING, P. & LEVISON, H. 1984. PSEUDOMONAS CEPACIA INFECTION IN CYSTIC-FIBROSIS - AN EMERGING PROBLEM. *Journal of Pediatrics*, 104, 206-210.
- JOGLER, C., WALDMANN, J., HUANG, X., JOGLER, M., GLOECKNER, F. O., MASCHER, T. & KOLTER, R. 2012. Identification of Proteins Likely To Be Involved in Morphogenesis, Cell Division, and Signal Transduction in Planctomycetes by Comparative Genomics. *Journal of Bacteriology*, 194, 6419-6430.
- JOHNSON, W. M., TYLER, S. D. & ROZEE, K. R. 1994. LINKAGE ANALYSIS OF GEOGRAPHIC AND CLINICAL CLUSTERS IN PSEUDOMONAS-CEPACIA INFECTIONS BY MULTILOCUS ENZYME ELECTROPHORESIS AND RIBOTYPING. *Journal of Clinical Microbiology*, 32, 924-930.
- JOHNSTON, R. B. 2001. Clinical aspects of chronic granulomatous disease. *Current Opinion in Hematology*, 8, 17-22.
- KADOWAKI, T., YUKITAKE, H., NAITO, M., SATO, K., KIKUCHI, Y., KONDO, Y., SHOJI, M. & NAKAYAMA, K. 2016. A two-component system regulates gene expression of the type IX secretion component proteins via an ECF sigma factor. *Scientific Reports*, 6.
- KAPANIDIS, A. N., MARGEAT, E., HO, S. O., KORTKHONJIA, E., WEISS, S. & EBRIGHT, R. H. 2006. Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science*, 314, 1144-1147.
- KAPPLER, U., FRIEDRICH, C. G., TRUPER, H. G. & DAHL, C. 2001. Evidence for two pathways of thiosulfate oxidation in *Starkeya novella* (formerly *Thiobacillus novellus*). *Archives of Microbiology*, 175, 102-111.
- KARIMOVA, G., ULLMANN, A. & LADANT, D. 2001. Protein-protein interaction between *Bacillus stearothermophilus* tyrosyl-tRNA synthetase subdomains revealed by a bacterial two-hybrid system. *Journal of Molecular Microbiology and Biotechnology*, 3.
- KARPEN, M. E. & DEHASETH, P. L. 2015. Base Flipping in Open Complex Formation at Bacterial

Promoters. *Biomolecules*, 5, 668-678.

- KAZMIERCZAK, M. J., WIEDMANN, M. & BOOR, K. J. 2005. Alternative sigma factors and their roles in bacterial virulence. *Microbiology and Molecular Biology Reviews*, 69, 527-+.
- KEITH, K. E., HYNES, D. W., SHOLDICE, J. E. & VALVANO, M. A. 2009. Delayed association of the NADPH oxidase complex with macrophage vacuoles containing the opportunistic pathogen *Burkholderia cenocepacia*. *Microbiology-Sgm*, 155, 1004-1015.
- KEITH, K. E., KILLIP, L., HE, P. Q., MORAN, G. R. & VALVANO, M. A. 2007. *Burkholderia cenocepacia* C5424 produces a pigment with antioxidant properties using a homogentisate intermediate. *Journal of Bacteriology*, 189, 9057-9065.
- KEITH, K. E. & VALVANO, M. A. 2007. Characterization of SodC, a periplasmic superoxide dismutase from *Burkholderia cenocepacia*. *Infection and Immunity*, 75, 2451-2460.
- KIM, K. Y., PARK, J. K. & PARK, S. 2016. In *Streptomyces coelicolor* SigR, methionine at the-35 element interacting region 4 confers the-31 'adenine base selectivity. *Biochemical and Biophysical Research Communications*, 470, 257-262.
- KIMBREL, J. A., GIVAN, S. A., HALGREN, A. B., CREASON, A. L., MILLS, D. I., BANOWETZ, G. M., ARMSTRONG, D. J. & CHANG, J. H. 2010. An improved, high-quality draft genome sequence of the Germination-Arrest Factor-producing *Pseudomonas fluorescens* WH6. *Bmc Genomics*, 11.
- KIMURA, A., YUHARA, S., OHTSUBO, Y., NAGATA, Y. & TSUDA, M. 2012. Suppression of pleiotropic phenotypes of a *Burkholderia multivorans* fur mutant by oxyR mutation. *Microbiology-Sgm*, 158, 1284-1293.
- KOEDAM, N., WITTOUCK, E., GABALLA, A., GILLIS, A., HOFTE, M. & CORNELIS, P. 1994. DETECTION AND DIFFERENTIATION OF MICROBIAL SIDEROPHORES BY ISOELECTRIC-FOCUSING AND CHROME AZUROL-S OVERLAY. *Biometals*, 7.
- KONTUR, W. S., CAPP, M. W., GRIES, T. J., SAECKER, R. M. & RECORD, M. T., JR. 2010. Probing DNA Binding, DNA Opening, and Assembly of a Downstream Clamp/Jaw in *Escherichia coli* RNA Polymerase-lambda P-R Promoter Complexes Using Salt and the Physiological Anion Glutamate. *Biochemistry*, 49, 4361-4373.
- KOO, B.-M., RHODIUS, V. A., NONAKA, G., DEHASETH, P. L. & GROSS, C. A. 2009. Reduced capacity of alternative ss to melt promoters ensures stringent promoter recognition. *Genes & Development*, 23, 2426-2436.
- KOSHKIN, A., NUNN, C. M., DJORDJEVIC, S. & DE MONTELLANO, P. R. O. 2003. The mechanism of mycobacterium tuberculosis alkylhydroperoxidase AhpD as defined by mutagenesis, crystallography, and kinetics. *Journal of Biological Chemistry*, 278, 29502-29508.
- KOTHE, M., ANTL, M., HUBER, B., STOECKER, K., EBRECHT, D., STEINMETZ, I. & EBERL, L. 2003. Killing of *Caenorhabditis elegans* by *Burkholderia cepacia* is controlled by the cep quorum-sensing system. *Cellular Microbiology*, 5, 343-351.
- KOVACH, M. E., ELZER, P. H., HILL, D. S., ROBERTSON, G. T., FARRIS, M. A., ROOP, R. M. & PETERSON, K. M. 1995. 4 NEW DERIVATIVES OF THE BROAD-HOST-RANGE CLONING VECTOR PBBR1MCS, CARRYING DIFFERENT ANTIBIOTIC-RESISTANCE CASSETTES. *Gene*, 166, 175-176.
- KOVACH, M. E., PHILLIPS, R. W., ELZER, P. H., ROOP, R. M. & PETERSON, K. M. 1994. PBBR1MCS - A BROAD-HOST-RANGE CLONING VECTOR. *Biotechniques*, 16.
- KULESUS, R. R., DIAZ-PEREZ, K., SLECHTA, E. S., ETO, D. S. & MULVEY, M. A. 2008. Impact of the RNA chaperone Hfq on the fitness and virulence potential of uropathogenic *Escherichia coli*. *Infection and Immunity*, 76, 3019-3026.
- KUMMERLI, R., SCHIESSL, K. T., WALDVOGEL, T., MCNEILL, K. & ACKERMANN, M. 2014. Habitat structure and the evolution of diffusible siderophores in bacteria. *Ecology Letters*, 17, 1536-1544.
- KUZNEDELOV, K., MINAKHIN, L., NIEDZIELA-MAJKA, A., DOVE, S. L., ROGULJA, D., NICKELS, B.

- E., HOCHSCHILD, A., HEYDUK, T. & SEVERINOV, K. 2002. A role for interaction of the RNA polymerase flap domain with the sigma subunit in promoter recognition. *Science*, 295, 855-857.
- LANDT, S. G., MARINOV, G. K., KUNDAJE, A., KHERADPOUR, P., PAULI, F., BATZOGLOU, S., BERNSTEIN, B. E., BICKEL, P., BROWN, J. B., CAYTING, P., CHEN, Y. W., DESALVO, G., EPSTEIN, C., FISHER-AYLOR, K. I., EUSKIRCHEN, G., GERSTEIN, M., GERTZ, J., HARTEMINK, A. J., HOFFMAN, M. M., IYER, V. R., JUNG, Y. L., KARMAKAR, S., KELLIS, M., KHARCHENKO, P. V., LI, Q. H., LIU, T., LIU, X. S., MA, L. J., MILOSAVLJEVIC, A., MYERS, R. M., PARK, P. J., PAZIN, M. J., PERRY, M. D., RAHA, D., REDDY, T. E., ROZOWSKY, J., SHORESH, N., SIDOW, A., SLATTERY, M., STAMATOYANNOPOULOS, J. A., TOLSTORUKOV, M. Y., WHITE, K. P., XI, S., FARNHAM, P. J., LIEB, J. D., WOLD, B. J. & SNYDER, M. 2012. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Research*, 22, 1813-1831.
- LANE, W. J. & DARST, S. A. 2006. The structural basis for promoter-35 element recognition by the group IV sigma factors. *Plos Biology*, 4, 1491-1500.
- LARDI, M., AGUILAR, C., PEDRIOLI, A., OMASITS, U., SUPPIGER, A., CARCAMO-OYARCE, G., SCHMID, N., AHRENS, C. H., EBERL, L. & PESSI, G. 2015. sigma(54)-Dependent Response to Nitrogen Limitation and Virulence in Burkholderia cenocepacia Strain H111. *Applied and Environmental Microbiology*, 81, 4077-4089.
- LAUDENBACH, D. E. & GROSSMAN, A. R. 1991. CHARACTERIZATION AND MUTAGENESIS OF SULFUR-REGULATED GENES IN A CYANOBACTERIUM - EVIDENCE FOR FUNCTION IN SULFATE TRANSPORT. *Journal of Bacteriology*, 173, 2739-2750.
- LEFEBRE, M. D. & VALVANO, M. A. 2001. In vitro resistance of Burkholderia cepacia complex isolates to reactive oxygen species in relation to catalase and superoxide dismutase production. *Microbiology-Sgm*, 147, 97-109.
- LEONI, L., ORSI, N., DE LORENZO, V. & VISCA, P. 2000. Functional analysis of PvdS, an iron starvation sigma factor of Pseudomonas aeruginosa. *Journal of Bacteriology*, 182, 1481-1491.
- LEWENZA, S., CONWAY, B., GREENBERG, E. P. & SOKOL, P. A. 1999. Quorum sensing in Burkholderia cepacia: Identification of the LuxRI homologs CepRI. *Journal of Bacteriology*, 181, 748-756.
- LI, W., STEVENSON, C. E. M., BURTON, N., JAKIMOWICZ, P., PAGET, M. S. B., BUTTNER, M. J., LAWSON, D. M. & KLEANTHOS, C. 2002. Identification and structure of the anti-sigma factor-binding domain of the disulphide-stress regulated sigma factor sigma(R) from Streptomyces coelicolor. *Journal of Molecular Biology*, 323, 225-236.
- LIAO, C. H. & MCCALLUS, D. E. 1998. Biochemical and genetic characterization of an extracellular protease from Pseudomonas fluorescens CY091. *Applied and Environmental Microbiology*, 64, 914-921.
- LIEBERMAN, T. D., FLETT, K. B., YELIN, I., MARTIN, T. R., MCADAM, A. J., PRIEBE, G. P. & KISHONY, R. 2014. Genetic variation of a bacterial pathogen within individuals with cystic fibrosis provides a record of selective pressures. *Nature Genetics*, 46, 82-+.
- LIEHL, P., BOCCARD, F., BLIGHT, M., VODOVAR, N. & LEMAITRE, B. 2006. Prevalence of Local Immune Response against Oral Infection in a Drosophila/Pseudomonas Infection Model. *Figshare*, 1.
- LIPUMA, J. J. 2010. The Changing Microbial Epidemiology in Cystic Fibrosis. *Clinical Microbiology Reviews*, 23, 299-+.
- LIU, J., LI, J., WU, Z., PEI, H., ZHOU, J. & XIANG, H. 2012. Identification and Characterization of the Cognate Anti-Sigma Factor and Specific Promoter Elements of a T. tengcongensis ECF Sigma Factor. *Plos One*, 7.
- LIU, X., BUSHNELL, D. A. & KORNBERG, R. D. 2011. Lock and Key to Transcription: sigma-DNA

- Interaction. *Cell*, 147, 1218-1219.
- LLAMAS, M. A., IMPERI, F., VISCA, P. & LAMONT, I. L. 2014. Cell-surface signaling in *Pseudomonas*: stress responses, iron transport, and pathogenicity. *Fems Microbiology Reviews*, 38, 569-597.
- LLAMAS, M. A., VAN DER SAR, A., CHU, B. C. H., SPARRIUS, M., VOGEL, H. J. & BITTER, W. 2009. A Novel Extracytoplasmic Function (ECF) Sigma Factor Regulates Virulence in *Pseudomonas aeruginosa*. *Plos Pathogens*, 5.
- LONETTO, M.A., BROWN, K.L., RUDD, K.E. & BUTTNER, M.J. (1994) Analysis Of The *Streptomyces-Coelicolor* Sige Gene Reveals The Existence Of A Subfamily Of Eubacterial Rna-Polymerase Sigma-Factors Involved In The Regulation Of Extracytoplasmic Functions. *Proceedings of the National Academy of Sciences of the United States of America*, 91, 7573-7577.
- LOURENCO, R. F., KOHLER, C. & GOMES, S. L. 2011. A two-component system, an anti-sigma factor and two paralogous ECF sigma factors are involved in the control of general stress response in *Caulobacter crescentus*. *Molecular Microbiology*, 80, 1598-1612.
- LOUTET, S. A. & VALVANO, M. A. 2010. A Decade of *Burkholderia cenocepacia* Virulence Determinant Research. *Infection and Immunity*, 78, 4088-4100.
- LUTTER, E., LEWENZA, S., DENNIS, J. J., VISSER, M. B. & SOKOL, P. A. 2001. Distribution of quorum-sensing genes in the *Burkholderia cepacia* complex. *Infection and Immunity*, 69, 4661-4666.
- MACLELLAN, S. R., GUARIGLIA-OROPEZA, V., GABALLA, A. & HELMANN, J. D. 2009. A two-subunit bacterial sigma-factor activates transcription in *Bacillus subtilis*. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 21323-21328.
- MADDOCKS, S. E. & OYSTON, P. C. F. 2008. Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology-Sgm*, 154, 3609-3623.
- MAHENTHIRALINGAM, E., URBAN, T. A. & GOLDBERG, J. B. 2005. The multifarious, multireplicon *Burkholderia cepacia* complex. *Nature Reviews Microbiology*, 3, 144-156.
- MAHREN, S., ENZ, S. & BRAUN, V. 2002. Functional interaction of region 4 of the extracytoplasmic function sigma factor fecI with the cytoplasmic portion of the FecR transmembrane protein of the *Escherichia coli* ferric citrate transport system. *Journal of Bacteriology*, 184, 3704-3711.
- MALOTT, R. J., BALDWIN, A., MAHENTHIRALINGAM, E. & SOKOL, P. A. 2005. Characterization of the cciIR quorum-sensing system in *Burkholderia cenocepacia*. *Infection and Immunity*, 73, 4982-4992.
- MALOTT, R. J., O'GRADY, E. P., TOLLER, J., INHULSEN, S., EBERL, L. & SOKOL, P. A. 2009. A *Burkholderia cenocepacia* Orphan LuxR Homolog Is Involved in Quorum-Sensing Regulation. *Journal of Bacteriology*, 191, 2447-2460.
- MARCOS-TORRES, F. J., PEREZ, J., GOMEZ-SANTOS, N., MORALEDA-MUNOZ, A. & MUNOZ-DORADO, J. 2016. In depth analysis of the mechanism of action of metal-dependent sigma factors: characterization of Core2 from *Myxococcus xanthus*. *Nucleic Acids Research*, 44, 5571-5584.
- MARKEL, E., BUTCHER, B. G., MYERS, C. R., STODGHILL, P., CARTINHOOR, S. & SWINGLE, B. 2013. Regulons of Three *Pseudomonas syringae* pv. tomato DC3000 Iron Starvation Sigma Factors. *Applied and Environmental Microbiology*, 79, 725-727.
- MARKEL, E., STODGHILL, P., BAO, Z. M., MYERS, C. R. & SWINGLE, B. 2016. AlgU Controls Expression of Virulence Genes in *Pseudomonas syringae* pv. tomato DC3000. *Journal of Bacteriology*, 198, 2330-2344.
- MARTIN, D. W., SCHURR, M. J., YU, H. & DERETIC, V. 1994. ANALYSIS OF PROMOTERS CONTROLLED BY THE PUTATIVE SIGMA-FACTOR ALGU REGULATING CONVERSION TO MUCOIDY IN *PSEUDOMONAS-AERUGINOSA* - RELATIONSHIP TO SIGMA(E) AND

- STRESS-RESPONSE. *Journal of Bacteriology*, 176, 6688-6696.
- MARTINEZ-BUENO, M. A., TOBES, R., REY, M. & RAMOS, J. L. 2002. Detection of multiple extracytoplasmic function (ECF) sigma factors in the genome of *Pseudomonas putida* KT2440 and their counterparts in *Pseudomonas aeruginosa* PA01. *Environmental Microbiology*, 4, 842-855.
- MASCHER, T. 2013. Signaling diversity and evolution of extracytoplasmic function (ECF) sigma factors. *Current Opinion in Microbiology*, 16, 148-155.
- MATHEW, A., EBERL, L. & CARLIER, A. L. 2014. A novel siderophore-independent strategy of iron uptake in the genus *Burkholderia*. *Molecular Microbiology*, 91, 805-820.
- MATHEW, A., JENUL, C., CARLIER, A. L. & EBERL, L. 2016. The role of siderophores in metal homeostasis of members of the genus *Burkholderia*. *Environmental Microbiology Reports*, 8, 103-109.
- MENARD, A., SANTOS, P., GRAINDORGE, A. & COURNOYER, B. 2007. Architecture of *Burkholderia cepacia* complex sigma(70) gene family: evidence of alternative primary and clade-specific factors, and genomic instability. *Bmc Genomics*, 8, 16.
- MESSIAEN, A. S., NELIS, H. & COENYE, T. 2014. Investigating the role of matrix components in protection of *Burkholderia cepacia* complex biofilms against tobramycin. *Journal of Cystic Fibrosis*, 13, 56-62.
- MIETHKE, M. & MARAHIEL, M. A. 2007. Siderophore-based iron acquisition and pathogen control. *Microbiology and Molecular Biology Reviews*, 71, 413-+.
- MILLS, S. A. & MARLETTA, M. A. 2005. Metal binding characteristics and role of iron oxidation in the ferric uptake regulator from *Escherichia coli*. *Biochemistry*, 44, 13553-13559.
- MIROPOLSKAYA, N., IGNATOV, A., BASS, I., ZHILINA, E., PUPOV, D. & KULBACHINSKIY, A. 2012. Distinct Functions of Regions 1.1 and 1.2 of RNA Polymerase sigma Subunits from *Escherichia coli* and *Thermus aquaticus* in Transcription Initiation. *Journal of Biological Chemistry*, 287, 23779-23789.
- MISSIAKAS, D. & RAINA, S. 1998. The extracytoplasmic function sigma factors: role and regulation. *Molecular Microbiology*, 28, 1059-1066.
- MITCHELL, J. E., ZHENG, D. L., BUSBY, S. J. W. & MINCHIN, S. D. 2003. Identification and analysis of 'extended-10' promoters in *Escherichia coli*. *Nucleic Acids Research*, 31, 4689-4695.
- MUKHOPADHYAY, J., KAPANIDIS, A. N., MEKLER, V., KORTKHONJIA, E., EBRIGHT, Y. W. & EBRIGHT, R. H. 2001. Translocation of sigma(70) with RNA polymerase during transcription: Fluorescence resonance energy transfer assay for movement relative to DNA. *Cell*, 106, 453-463.
- MURAKAMI, K. S. 2015. Structural Biology of Bacterial RNA Polymerase. *Biomolecules*, 5, 848-864.
- MURAKAMI, K. S., MASUDA, S., CAMPBELL, E. A., MUZZIN, O. & DARST, S. A. 2002. Structural basis of transcription initiation: An RNA polymerase holoenzyme-DNA complex. *Science*, 296, 1285-1290.
- MURPHY, M. P. & CARAHER, E. 2015. Residence in biofilms allows *Burkholderia cepacia* complex (Bcc) bacteria to evade the antimicrobial activities of neutrophil-like dHL60 cells. *Pathogens and Disease*, 73.
- NATHAN, C. & CUNNINGHAM-BUSSEL, A. 2013. Beyond oxidative stress: an immunologist's guide to reactive oxygen species. *Nature Reviews Immunology*, 13, 349-361.
- NEILANDS, J. B. 1995. SIDEROPHORES - STRUCTURE AND FUNCTION OF MICROBIAL IRON TRANSPORT COMPOUNDS. *Journal of Biological Chemistry*, 270, 26723-26726.
- NICHOLSON, M. L. & LAUDENBACH, D. E. 1995. GENES ENCODED ON A CYANOBACTERIAL PLASMID ARE TRANSCRIPTIONALLY REGULATED BY SULFUR AVAILABILITY AND CYSR. *Journal of Bacteriology*, 177, 2143-2150.

- NOINAJ, N., GUILLIER, M., BARNARD, T. J. & BUCHANAN, S. K. 2010. TonB-Dependent Transporters: Regulation, Structure, and Function. *Annual Review of Microbiology, Vol 64, 2010, 64*, 43-60.
- O'GRADY, E. P., NGUYEN, D. T., WEISSKOPF, L., EBERL, L. & SOKOL, P. A. 2011. The Burkholderia cenocepacia LysR-Type Transcriptional Regulator ShvR Influences Expression of Quorum-Sensing, Protease, Type II Secretion, and *afc* Genes. *Journal of Bacteriology, 193*, 163-176.
- OCHSNER, U. A. & VASIL, M. L. 1996. Gene repression by the ferric uptake regulator in *Pseudomonas aeruginosa*: Cycle selection of iron-regulated genes. *Proceedings of the National Academy of Sciences of the United States of America, 93*, 4409-4414.
- OCHSNER, U. A., WILDERMAN, P. J., VASIL, A. I. & VASIL, M. L. 2002. GeneChip((R)) expression analysis of the iron starvation response in *Pseudomonas aeruginosa*: identification of novel pyoverdine biosynthesis genes. *Molecular Microbiology, 45*, 1277-1287.
- OGURA, M. & ASAI, K. 2016. Glucose Induces ECF Sigma Factor Genes, *sigX* and *sigM*, Independent of Cognate Anti-sigma Factors through Acetylation of CshA in *Bacillus subtilis*. *Frontiers in Microbiology, 7*.
- OKRENT, R. A., HALGREN, A. B., AZEVEDO, M. D., CHANG, J. H., MILLS, D. I., MASELKO, M., ARMSTRONG, D. J., BANOWETZ, G. M. & TRIPPE, K. M. 2014. Negative regulation of germination-arrest factor production in *Pseudomonas fluorescens* WH6 by a putative extracytoplasmic function sigma factor. *Microbiology-Sgm, 160*, 2432-2442.
- OLSEN, I. 2015. Biofilm-specific antibiotic tolerance and resistance. *European Journal of Clinical Microbiology & Infectious Diseases, 34*, 877-886.
- OSTERBERG, S., DEL PESO-SANTOS, T. & SHINGLER, V. 2011. Regulation of Alternative Sigma Factor Use. *Annual Review of Microbiology, Vol 65, 65*, 37-55.
- OTANI, H., HIGO, A., NANAMIYA, H., HORINOUCI, S. & OHNISHI, Y. 2013. An alternative sigma factor governs the principal sigma factor in *Streptomyces griseus*. *Molecular Microbiology, 87*, 1223-1236.
- PABO, C. O. & SAUER, R. T. 1992. TRANSCRIPTION FACTORS - STRUCTURAL FAMILIES AND PRINCIPLES OF DNA RECOGNITION. *Annual Review of Biochemistry, 61*, 1053-1095.
- PAGET, M. S. 2015. Bacterial Sigma Factors and Anti-Sigma Factors: Structure, Function and Distribution. *Biomolecules, 5*, 1245-1265.
- PAGET, M. S. B., BAE, J. B., HAHN, M. Y., LI, W., KLEANTHOS, C., ROE, J. H. & BUTTNER, M. J. 2001. Mutational analysis of RsrA, a zinc-binding anti-sigma factor with a thiol-disulphide redox switch. *Molecular Microbiology, 39*, 1036-1047.
- PAGET, M. S. B. & HELMANN, J. D. 2003. Protein family review - The sigma(70) family of sigma factors. *Genome Biology, 4*.
- PAGET, M. S. B., KANG, J. G., ROE, J. H. & BUTTNER, M. J. 1998. sigma(R), an RNA polymerase sigma factor that modulates expression of the thioredoxin system in response to oxidative stress in *Streptomyces coelicolor* A3(2). *Embo Journal, 17*, 5776-5782.
- PARK, C. S., WU, F. Y. H. & WU, C. W. 1982. MOLECULAR MECHANISM OF PROMOTER SELECTION IN GENE-TRANSCRIPTION .2. KINETIC EVIDENCE FOR PROMOTER SEARCH BY A ONE-DIMENSIONAL DIFFUSION OF RNA-POLYMERASE MOLECULE ALONG THE DNA-TEMPLATE. *Journal of Biological Chemistry, 257*, 6950-6956.
- PARK, P. J. 2009. CHIP-seq: advantages and challenges of a maturing technology. *Nature Reviews Genetics, 10*, 669-680.
- PEETERS, E., NELIS, H. J. & COENYE, T. 2009. In vitro activity of ceftazidime, ciprofloxacin, meropenem, minocycline, tobramycin and trimethoprim/sulfamethoxazole against planktonic and sessile Burkholderia cepacia complex bacteria. *Journal of Antimicrobial Chemotherapy, 64*, 801-809.
- PEETERS, E., SASS, A., MAHENTHIRALINGAM, E., NELIS, H. & COENYE, T. 2010. Transcriptional

- response of Burkholderia cenocepacia J2315 sessile cells to treatments with high doses of hydrogen peroxide and sodium hypochlorite. *Bmc Genomics*, 11.
- PITT, T. L., KAUFMANN, M. E., PATEL, P. S., BENGE, L. C. A., GASKIN, S. & LIVERMORE, D. M. 1996. Type characterisation and antibiotic susceptibility of Burkholderia (Pseudomonas) cepacia isolates from patients with cystic fibrosis in the United Kingdom and the Republic of Ireland. *Journal of Medical Microbiology*, 44, 203-210.
- POHL, E., HALLER, J. C., MIJOVILOVICH, A., MEYER-KLAUCKE, W., GARMAN, E. & VASIL, M. L. 2003. Architecture of a protein central to iron homeostasis: crystal structure and spectroscopic analysis of the ferric uptake regulator. *Molecular Microbiology*, 47, 903-915.
- POTVIN, E., SANSCHAGRIN, F. & LEVESQUE, R. C. 2008. Sigma factors in Pseudomonas aeruginosa. *Fems Microbiology Reviews*, 32, 38-55.
- PUPOV, D., KUZIN, I., BASS, I. & KULBACHINSKIY, A. 2014. Distinct functions of the RNA polymerase Sigma subunit region 3.2 in RNA priming and promoter escape. *Nucleic Acids Research*, 42, 4494-4504.
- PUPOV, D., MIROPOLSKAYA, N., SEVOSTYANOVA, A., BASS, I., ARTSIMOVITCH, I. & KULBACHINSKIY, A. 2010. Multiple roles of the RNA polymerase beta' SW2 region in transcription initiation, promoter escape, and RNA elongation. *Nucleic Acids Research*, 38, 5784-+.
- RAAIJMAKERS, J. M., DE BRUIJN, I. & DE KOCK, M. J. D. 2006. Cyclic lipopeptide production by plant-associated Pseudomonas spp.: Diversity, activity, biosynthesis, and regulation. *Molecular Plant-Microbe Interactions*, 19, 699-710.
- RAFFAELLE, M., KANIN, E. I., VOGT, J., BURGESS, R. R. & ANSARI, A. Z. 2005. Holoenzyme switching and stochastic release of sigma factors from RNA polymerase in vivo. *Molecular Cell*, 20, 357-366.
- RAIVIO, T. L. 2005. Envelope stress responses and Gram-negative bacterial pathogenesis. *Molecular Microbiology*, 56, 1119-1128.
- RAJASEKAR, K. V., ZDANOWSKI, K., YAN, J., HOPPER, J. T. S., FRANCIS, M. L. R., SEEPERSAD, C., SHARP, C., PECQUEUR, L., WERNER, J. M., ROBINSON, C. V., MOHAMMED, S., POTTS, J. R. & KLEANTHOUS, C. 2016. The anti-sigma factor RsrA responds to oxidative stress by burying its hydrophobic core. *Nature Communications*, 7.
- RATLEDGE, C. & DOVER, L. G. 2000. Iron metabolism in pathogenic bacteria. *Annual Review of Microbiology*, 54, 881-941.
- RAVEL, J. & CORNELIS, P. 2003. Genomics of pyoverdine-mediated iron uptake in pseudomonads. *Trends in Microbiology*, 11, 195-200.
- REDFIELD, R. J. 2002. Is quorum sensing a side effect of diffusion sensing? *Trends in Microbiology*, 10, 365-370.
- REVYAKIN, A., LIU, C., EBRIGHT, R. H. & STRICK, T. R. 2006. Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. *Science*, 314, 1139-1143.
- RHODIUS, V. A., MUTALIK, V. K. & GROSS, C. A. 2012. Predicting the strength of UP-elements and full-length E-coli sigma(E) promoters. *Nucleic Acids Research*, 40, 2907-2924.
- RHODIUS, V. A., SEGALL-SHAPIRO, T. H., SHARON, B. D., GHODASARA, A., ORLOVA, E., TABAKH, H., BURKHARDT, D. H., CLANCY, K., PETERSON, T. C., GROSS, C. A. & VOIGT, C. A. 2013. Design of orthogonal genetic switches based on a crosstalk map of rs, anti-rs, and promoters. *Molecular Systems Biology*, 9.
- RICCHETTI, M., METZGER, W. & HEUMANN, H. 1988. ONE-DIMENSIONAL DIFFUSION OF ESCHERICHIA-COLI DNA-DEPENDENT RNA-POLYMERASE - A MECHANISM TO FACILITATE PROMOTER LOCATION. *Proceedings of the National Academy of Sciences of the United States of America*, 85, 4610-4614.
- ROMLING, U., WINGENDER, J., MULLER, H. & TUMMLER, B. 1994. A MAJOR PSEUDOMONAS-

- AERUGINOSA CLONE COMMON TO PATIENTS AND AQUATIC HABITATS. *Applied and Environmental Microbiology*, 60.
- ROWLEY, G., SPECTOR, M., KORMANEC, J. & ROBERTS, M. 2006. Pushing the envelope: extracytoplasmic stress responses in bacterial pathogens. *Nature Reviews Microbiology*, 4, 383-394.
- RUFF, E. F., RECORD, M. T. & ARTSIMOVITCH, I. 2015. Initial Events in Bacterial Transcription Initiation. *Biomolecules*, 5, 1035-1062.
- RYAN, G. T., WEI, Y. P. & WINANS, S. C. 2013. A LuxR-type repressor of Burkholderia cenocepacia inhibits transcription via antiactivation and is inactivated by its cognate acylhomoserine lactone. *Molecular Microbiology*, 87, 94-111.
- SAECKER, R. M., RECORD, M. T. & DEHASETH, P. L. 2011. Mechanism of Bacterial Transcription Initiation: RNA Polymerase - Promoter Binding, Isomerization to Initiation-Competent Open Complexes, and Initiation of RNA Synthesis. *Journal of Molecular Biology*, 412, 754-771.
- SAHA, R., SAHA, N., DONOFRIO, R. S. & BESTERVELT, L. L. 2013. Microbial siderophores: a mini review. *Journal of Basic Microbiology*, 53, 303-317.
- SAKATA-SOGAWA, K. & SHIMAMOTO, N. 2004. RNA polymerase can track a DNA groove during promoter search. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 14731-14735.
- SALDIAS, M. S., LAMOTHE, J., WU, R. & VALVANO, M. A. 2008. Burkholderia cenocepacia requires the RpoN sigma factor for biofilm formation and intracellular trafficking within macrophages. *Infection and Immunity*, 76, 1059-1067.
- SANTIAGO, A. S., SANTOS, C. A., MENDES, J. S., TOLEDO, M. A. S., BELOTI, L. L., SOUZA, A. A. & SOUZA, A. P. 2015. Characterization of the LysR-type transcriptional regulator YcjZ-like from Xylella fastidiosa overexpressed in Escherichia coli. *Protein Expression and Purification*, 113, 72-78.
- SASS, A., MARCHBANK, A., TULLIS, E., LIPUMA, J. J. & MAHENTHIRALINGAM, E. 2011. Spontaneous and evolutionary changes in the antibiotic resistance of Burkholderia cenocepacia observed by global gene expression analysis. *Bmc Genomics*, 12.
- SASS, A. M., VAN ACKER, H., FORSTNER, K. U., VAN NIEUWERBURGH, F., DEFORCE, D., VOGEL, J. & COENYE, T. 2015. Genome-wide transcription start site profiling in biofilm-grown Burkholderia cenocepacia J2315. *Bmc Genomics*, 16.
- SCHAIBLE, M. E. & KAUFMANN, S. H. E. 2004. Iron and microbial infection. *Nature Reviews Microbiology*, 2, 946-953.
- SCHULZ, S., ECKWEILER, D., BIELECKA, A., NICOLAI, T., FRANKE, R., DOTSCHE, A., HORNISCHER, K., BRUCHMANN, S., DUVEL, J. & HAUSSLER, S. 2015. Elucidation of Sigma Factor-Associated Networks in Pseudomonas aeruginosa Reveals a Modular Architecture with Limited and Function-Specific Crosstalk. *Plos Pathogens*, 11.
- SCHWAB, U., ABDULLAH, L. H., PERLMUTT, O. S., ALBERT, D., DAVIS, C. W., ARNOLD, R. R., YANKASKAS, J. R., GILLIGAN, P., NEUBAUER, H., RANDELL, S. H. & BOUCHER, R. C. 2014. Localization of Burkholderia cepacia Complex Bacteria in Cystic Fibrosis Lungs and Interactions with Pseudomonas aeruginosa in Hypoxic Mucus. *Infection and Immunity*, 82, 4729-4745.
- SCHWARTZ, E. C., SHEKHTMAN, A., DUTTA, K., PRATT, M. R., COWBURN, D., DARST, S. & MUIR, T. W. 2008. A Full-Length Group 1 Bacterial Sigma Factor Adopts a Compact Structure Incompatible with DNA Binding. *Chemistry & Biology*, 15, 1091-1103.
- SEEVER, L. C. & IMLAY, J. A. 2004. Are respiratory enzymes the primary sources of intracellular hydrogen peroxide? *Journal of Biological Chemistry*, 279, 48742-48750.
- SEED, K. D. & DENNIS, J. J. 2008. Development of Galleria mellonella as an alternative infection model for the Burkholderia cepacia complex. *Infection and Immunity*, 76, 1267-1275.

- SHASTRI, S., PhD thesis, 2016, University of Sheffield.
- SILAR, R., HOLATKO, J., RUCKA, L., RAPOPORT, A., DOSTALOVA, H., KADERABKOVA, P., NESVERA, J. & PATEK, M. 2016. Use of In Vitro Transcription System for Analysis of *Corynebacterium glutamicum* Promoters Recognized by Two Sigma Factors. *Current Microbiology*, 73, 401-408.
- SIMON, R., PRIEFER, U. & PUHLER, A. 1983. A BROAD HOST RANGE MOBILIZATION SYSTEM FOR INVIVO GENETIC-ENGINEERING - TRANSPOSON MUTAGENESIS IN GRAM-NEGATIVE BACTERIA. *Bio-Technology*, 1.
- SINGH, S. S., TYPAS, A., HENGGE, R. & GRAINGER, D. C. 2011. Escherichia coli Sigma(70) senses sequence and conformation of the promoter spacer region. *Nucleic Acids Research*, 39, 5109-5118.
- SMITH, C. L., WEISS, B. L., AKSOY, S. & RUNYEN-JANECKY, L. J. 2013. Characterization of the Achromobactin Iron Acquisition Operon in *Sodalis glossinidius*. *Applied and Environmental Microbiology*, 79, 2872-2881.
- SOKOL, P. A., DARLING, P., LEWENZA, S., CORBETT, C. R. & KOOI, C. D. 2000. Identification of a siderophore receptor required for ferric ornibactin uptake in *Burkholderia cepacia*. *Infection and Immunity*, 68, 6554-6560.
- SOKOL, P. A., SAJJAN, U., VISSER, M. B., GINGUES, S., FORSTNER, J. & KOOI, C. 2003. The CepIR quorum-sensing system contributes to the virulence of *Burkholderia cenocepacia* respiratory infections. *Microbiology-Sgm*, 149, 3649-3658.
- SONG, C., AUNDY, K., VAN DE MORTEL, J. & RAAIJMAKERS, J. M. 2014. Discovery of new regulatory genes of lipopeptide biosynthesis in *Pseudomonas fluorescens*. *Fems Microbiology Letters*, 356, 166-175.
- SOUSA, S. A., FELICIANO, J. R., PITA, T., GUERREIRO, S. I. & LEITAO, J. H. 2017. *Burkholderia cepacia* Complex Regulation of Virulence Gene Expression: A Review. *Genes (Basel)*, 8.
- SOUSA, S. A., ULRICH, M., BRAGONZI, A., BURKE, M., WORLITZSCH, D., LEITAO, J. H., MEISNER, C., EBERL, L., SA-CORREIA, I. & DORING, G. 2007. Virulence of *Burkholderia cepacia* complex strains in gp91(phox-/-) mice. *Cellular Microbiology*, 9, 2817-2825.
- SPIEWAK, H., PhD thesis, 2016, University of Sheffield.
- STAFFORD, D. A. & CALLELY, A. G. 1969. UTILIZATION OF THIOCYANATE BY A HETEROTROPHIC BACTERIUM. *Journal of General Microbiology*, 55, 285-&.
- STARON, A., SOFIA, H. J., DIETRICH, S., ULRICH, L. E., LIESEGANG, H. & MASCHER, T. 2009. The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF) sigma factor protein family. *Molecular Microbiology*, 74, 557-581.
- STIEFEL, A., MAHREN, S., OCHS, M., SCHINDLER, P. T., ENZ, S. & BRAUN, V. 2001. Control of the ferric citrate transport system of *Escherichia coli*: Mutations in region 2.1 of the FecI extracytoplasmic-function sigma factor suppress mutations in the FecR transmembrane regulatory protein. *Journal of Bacteriology*, 183, 162-170.
- STOCKWELL, S. B., REUTIMANN, L. & GUERINOT, M. L. 2012. A Role for *Bradyrhizobium japonicum* ECF16 Sigma Factor EcfS in the Formation of a Functional Symbiosis with Soybean. *Molecular Plant-Microbe Interactions*, 25, 119-128.
- STUDIER, F. W. & MOFFATT, B. A. 1986. USE OF BACTERIOPHAGE-T7 RNA-POLYMERASE TO DIRECT SELECTIVE HIGH-LEVEL EXPRESSION OF CLONED GENES. *Journal of Molecular Biology*, 189, 113-130.
- SUBRAMONI, S. & SOKOL, P. A. 2012. Quorum sensing systems influence *Burkholderia cenocepacia* virulence. *Future Microbiology*, 7, 1373-1387.
- SUPPIGER, A., SCHMID, N., AGUILAR, C., PESSI, G. & EBERL, L. 2013. Two quorum sensing systems control biofilm formation and virulence in members of the *Burkholderia cepacia* complex. *Virulence*, 4, 400-409.
- SZTILLER-SIKORSKA, M., HEYDUK, E. & HEYDUK, T. 2011. Promoter spacer DNA plays an active

- role in integrating the functional consequences of RNA polymerase contacts with-10 and-35 promoter elements. *Biophysical Chemistry*, 159, 73-81.
- SZTUKOWSKA, M., BUGNO, M., POTEMPA, J., TRAVIS, J. & KURTZ, D. M. 2002. Role of rubrerythrin in the oxidative stress response of *Porphyromonas gingivalis*. *Molecular Microbiology*, 44, 479-488.
- TETTMANN, B., DOETSCH, A., ARMANT, O., FJELL, C. D. & OVERHAGE, J. 2014. Knockout of Extracytoplasmic Function Sigma Factor ECF-10 Affects Stress Resistance and Biofilm Formation in *Pseudomonas putida* KT2440. *Applied and Environmental Microbiology*, 80, 4911-4919.
- THAKUR, K. G., JOSHI, A. M. & GOPAL, B. 2007. Structural and biophysical studies on two promoter recognition domains of the extra-cytoplasmic function cr factor from *Mycobacterium tuberculosis*. *Journal of Biological Chemistry*, 282, 4711-4718.
- THOMAS, M. S. 2007. Iron acquisition mechanisms of the *Burkholderia cepacia* complex. *Biometals*, 20, 431-452.
- TOMICH, M., GRIFFITH, A., HERFST, C. A., BURNS, J. L. & MOHR, C. D. 2003. Attenuated virulence of a *Burkholderia cepacia* type III secretion mutant in a murine model of infection. *Infection and Immunity*, 71, 1405-1415.
- TOMLIN, K. L., COLL, O. P. & CERI, H. 2001. Interspecies biofilms of *Pseudomonas aeruginosa* and *Burkholderia cepacia*. *Canadian Journal of Microbiology*, 47, 949-954.
- TONDO, M. L., PETROCELLI, S., OTTADO, J. & ORELLANO, E. G. 2010. The Monofunctional Catalase KatE of *Xanthomonas axonopodis* pv. *citri* Is Required for Full Virulence in Citrus Plants. *Plos One*, 5.
- TREPREAU, J., DE ROSNY, E., DUBOC, C., SARRET, G., PETIT-HARTLEIN, I., MAILLARD, A. P., IMBERTY, A., PROUX, O. & COVES, J. 2011a. Spectroscopic Characterization of the Metal-Binding Sites in the Periplasmic Metal-Sensor Domain of CnrX from *Cupriavidus metallidurans* CH34. *Biochemistry*, 50, 9036-9045.
- TREPREAU, J., GIRARD, E., MAILLARD, A. P., DE ROSNY, E., PETIT-HAERTLEIN, I., KAHN, R. & COVES, J. 2011b. Structural Basis for Metal Sensing by CnrX. *Journal of Molecular Biology*, 408, 766-779.
- TU, N., LIMA, A., BANDEALI, Z. & ANDERSON, B. 2016. Characterization of the general stress response in *Bartonella henselae*. *Microbial Pathogenesis*, 92, 1-10.
- TYRRELL, J., WHELAN, N., WRIGHT, C., SA-CORREIA, I., MCCLEAN, S., THOMAS, M. & CALLAGHAN, M. 2015. Investigation of the multifaceted iron acquisition strategies of *Burkholderia cenocepacia*. *Biometals*, 28, 367-380.
- UEHLINGER, S., SCHWAGER, S., BERNIER, S. P., RIEDEL, K., NGUYEN, D. T., SOKOL, P. A. & EBERL, L. 2009. Identification of Specific and Universal Virulence Factors in *Burkholderia cenocepacia* Strains by Using Multiple Infection Hosts. *Infection and Immunity*, 77, 4102-4110.
- VALLET-GELY, I., NOVIKOV, A., AUGUSTO, L., LIEHL, P., BOLBACH, G., PECHY-TARR, M., COSSON, P., KEEL, C., CAROFF, M. & LEMAITRE, B. 2010. Association of Hemolytic Activity of *Pseudomonas entomophila*, a Versatile Soil Bacterium, with Cyclic Lipopeptide Production. *Applied and Environmental Microbiology*, 76, 910-921.
- VAN ACKER, H., SASS, A., BAZZINI, S., DE ROY, K., UDINE, C., MESSIAEN, T., RICCARDI, G., BOON, N., NELIS, H. J., MAHENTHIRALINGAM, E. & COENYE, T. 2013. Biofilm-Grown *Burkholderia cepacia* Complex Cells Survive Antibiotic Treatment by Avoiding Production of Reactive Oxygen Species. *Plos One*, 8.
- VAN ACKER, H., SASS, A., DHONDT, I., NELIS, H. J. & COENYE, T. 2014. Involvement of toxin-antitoxin modules in *Burkholderia cenocepacia* biofilm persistence. *Pathogens and Disease*, 71, 326-335.
- VAN DEN BROEK, D., PhD thesis, 2005, Leiden University, The Netherlands.

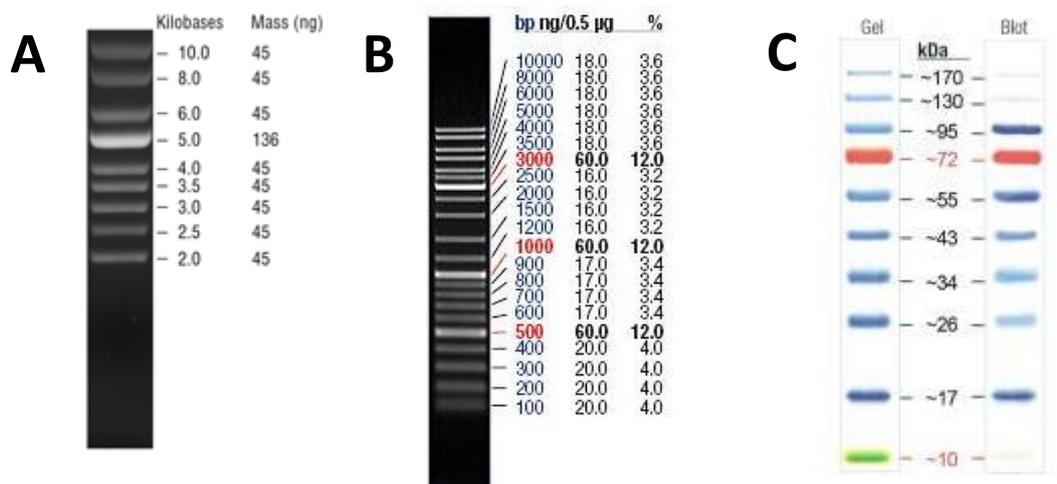
- VASSYLYEV, D. G., SEKINE, S., LAPTENKO, O., LEE, J., VASSYLYEVA, M. N., BORUKHOV, S. & YOKOYAMA, S. 2002. Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 angstrom resolution. *Nature*, 417, 712-719.
- VENTURI, V., FRISCINA, A., BERTANI, I., DEVESCOVI, G. & AGUILAR, C. 2004. Quorum sensing in the Burkholderia cepacia complex. *Research in Microbiology*, 155, 238-244.
- VISCA, P., IMPERI, F. & LAMONT, I. L. 2007. Pyoverdine siderophores: from biogenesis to biosignificance. *Trends in Microbiology*, 15, 22-30.
- VISCA, P., LEONI, L., WILSON, M. J. & LAMONT, I. L. 2002. Iron transport and regulation, cell signalling and genomics: lessons from Escherichia coli and Pseudomonas. *Molecular Microbiology*, 45, 1177-1190.
- VISSER, M. B., MAJUMDAR, S., HANI, E. & SOKOL, P. A. 2004. Importance of the ornibactin and pyochelin siderophore transport systems in Burkholderia cenocepacia lung infections. *Infection and Immunity*, 72, 2850-2857.
- VODOVAR, N., VALLENET, D., CRUVEILLER, S., ROUY, Z., BARBE, V., ACOSTA, C., CATTOLICO, L., JUBIN, C., LAJUS, A., SEGURENS, B., VACHERIE, B., WINCKER, P., WEISSENBACH, J., LEMAITRE, B., MEDIGUE, C. & BOCCARD, F. 2006. Complete genome sequence of the entomopathogenic and metabolically versatile soil bacterium Pseudomonas entomophila. *Nature Biotechnology*, 24, 673-679.
- VU, H. P., MU, A. & MOREAU, J. W. 2013. Biodegradation of thiocyanate by a novel strain of Burkholderia phytofirmans from soil contaminated by gold mine tailings. *Letters in Applied Microbiology*, 57.
- WALDRON, K. J., RUTHERFORD, J. C., FORD, D. & ROBINSON, N. J. 2009. Metalloproteins and metal sensing. *Nature*, 460, 823-830.
- WALSH, N. P., ALBA, B. M., BOSE, B., GROSS, C. A. & SAUER, R. T. 2003. OMP peptide signals initiate the envelope-stress response by activating DegS protease via relief of inhibition mediated by its PDZ domain. *Cell*, 113, 61-71.
- WANDERSMAN, C. & DELEPELAIRE, P. 2004. Bacterial iron sources: From siderophores to hemophores. *Annual Review of Microbiology*, 58, 611-647.
- WANG, F., REDDING, S., FINKELSTEIN, I. J., GORMAN, J., REICHMAN, D. R. & GREENE, E. C. 2013. The promoter-search mechanism of Escherichia coli RNA polymerase is dominated by three-dimensional diffusion. *Nature Structural & Molecular Biology*, 20, 174-181.
- WATANABE, S., KITA, A., KOBAYASHI, K. & MIKI, K. 2008. Crystal structure of the 2Fe-2S oxidative-stress sensor SoxR bound to DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 4121-4126.
- WECKE, T., HALANG, P., STARON, A., DUFOUR, Y. S., DONOHUE, T. J. & MASCHER, T. 2012. Extracytoplasmic function sigma factors of the widely distributed group ECF41 contain a fused regulatory domain. *MicrobiologyOpen*, 1, 194-213.
- WHITBY, P. W., VANWAGONER, T. M., SPRINGER, J. M., MORTON, D. J., SEALE, T. W. & STULL, T. L. 2006. Burkholderia cenocepacia utilizes ferritin as an iron source. *Journal of Medical Microbiology*, 55, 661-668.
- WHITE, M. F. & DILLINGHAM, M. S. 2012. Iron-sulphur clusters in nucleic acid processing enzymes. *Current Opinion in Structural Biology*, 22, 94-100.
- WILDERMAN, P. J., VASIL, A. I., JOHNSON, Z., WILSON, M. J., CUNLIFFE, H. E., LAMONT, I. L. & VASIL, M. L. 2001. Characterization of an endoprotease (PrpL) encoded by a PvdS-regulated gene in Pseudomonas aeruginosa. *Infection and Immunity*, 69, 5385-5394.
- WILKEN, C., KITZING, K., KURZBAUER, R., EHRMANN, M. & CLAUSEN, T. 2004. Crystal structure of the DegS stress sensor: How a PDZ domain recognizes misfolded protein and activates a protease. *Cell*, 117, 483-494.
- WOODS, E. C. & MCBRIDE, S. M. 2017. Regulation of antimicrobial resistance by

- extracytoplasmic function (ECF) sigma factors. *Microbes Infect.*
- WOSTEN, M. 1998. Eubacterial sigma-factors. *Fems Microbiology Reviews*, 22, 127-150.
- XIAO, R. & KISAALITA, W. S. 1997. Iron acquisition from transferrin and lactoferrin by *Pseudomonas aeruginosa* pyoverdinin. *Microbiology-Uk*, 143, 2509-2515.
- YABUUCHI, E., KOSAKO, Y., OYAIZU, H., YANO, I., HOTTA, H., HASHIMOTO, Y., EZAKI, T. & ARAKAWA, M. 1992. PROPOSAL OF BURKHOLDERIA GEN-NOV AND TRANSFER OF 7 SPECIES OF THE GENUS PSEUDOMONAS HOMOLOGY GROUP-II TO THE NEW GENUS, WITH THE TYPE SPECIES BURKHOLDERIA-CEPACIA (PALLERONI AND HOLMES 1981) COMB-NOV. *Microbiology and Immunology*, 36, 1251-1275.
- YANG, M.-M., WEN, S.-S., MAVRODI, D. V., MAVRODI, O. V., VON WETTSTEIN, D., THOMASHOW, L. S., GUO, J.-H. & WELLER, D. M. 2014. Biological Control of Wheat Root Diseases by the CLP-Producing Strain *Pseudomonas fluorescens* HC1-07. *Phytopathology*, 104, 248-256.
- YANISCHPERRON, C., VIEIRA, J. & MESSING, J. 1985. IMPROVED M13 PHAGE CLONING VECTORS AND HOST STRAINS - NUCLEOTIDE-SEQUENCES OF THE M13MP18 AND PUC19 VECTORS. *Gene*, 33.
- YEATS, C., BENTLEY, S. & BATEMAN, A. 2003. New knowledge from old: In silico discovery of novel protein domains in *Streptomyces coelicolor*. *Bmc Microbiology*, 3.
- YOO, J. S., OH, G. S., RYOO, S. & ROE, J. H. 2016. Induction of a stable sigma factor SigR by translation-inhibiting antibiotics confers resistance to antibiotics. *Scientific Reports*, 6.
- YOUNG, B. A., ANTHONY, L. C., GRUBER, T. M., ARTHUR, T. M., HEYDUK, E., LU, C. Z., SHARP, M. M., HEYDUK, T., BURGESS, R. R. & GROSS, C. A. 2001. A coiled-coil from the RNA polymerase beta ' subunit allosterically induces selective nontemplate strand binding by sigma(70). *Cell*, 105, 935-944.
- YUHARA, S., KOMATSU, H., GOTO, H., OHTSUBO, Y., NAGATA, Y. & TSUDA, M. 2008. Pleiotropic roles of iron-responsive transcriptional regulator Fur in *Burkholderia multivorans*. *Microbiology-Sgm*, 154, 1763-1774.
- YUZENKOVA, Y., TADIGOTLA, V. R., SEVERINOV, K. & ZENKIN, N. 2011. A new basal promoter element recognized by RNA polymerase core enzyme. *Embo Journal*, 30, 3766-3775.
- ZDANOWSKI, K., DOUGHTY, P., JAKIMOWICZ, P., O'HARA, L., BUTTNER, M. J., PAGET, M. S. B. & KLEANTHOUS, C. 2006. Assignment of the zinc ligands in RsrA, a redox-sensing ZAS protein from *Streptomyces coelicolor*. *Biochemistry*, 45, 8294-8300.
- ZHANG, J., GAO, J., WU, M., CHEN, H. & HE, C. 2009. [Roles of OxyR_{oo}, a transcriptional regulator of *Xanthomonas oryzae* pv. *oryzae* in regulation of detoxification of hydrogen peroxide]. *Wei Sheng Wu Xue Bao*, 49, 874-9.
- ZHANG, Y., FENG, Y., CHATTERJEE, S., TUSKE, S., HO, M. X., ARNOLD, E. & EBRIGHT, R. H. 2012. Structural Basis of Transcription Initiation. *Science*, 338, 1076-1080.
- ZLOSNIK, J. E. A., HIRD, T. J., FRAENKEL, M. C., MOREIRA, L. M., HENRY, D. A. & SPEERT, D. P. 2008. Differential mucoid exopolysaccharide production by members of the *Burkholderia cepacia* complex. *Journal of Clinical Microbiology*, 46, 1470-1473.

Appendices

Appendix I:

DNA and protein ladders



Images of ladders by the respective manufacturers used in this study

(A) Supercoiled DNA ladder (NEB) with the sizes (in kbp) and concentration indicated on the right. (B) GeneRuler DNA ladder mix (Thermo Fisher Scientific) with the sizes (in bp) and concentration indicated on the right. (C) EZ-Run Prestained Rec Protein Ladder (Fisher) as run on a (4-20%) SDS-PA gel or blot with sizes in kDa as indicated.

Appendix II:

Quality control reports of samples used in the CHIP-seq study carried out in GALAXY:

Each figure shows the summary of the parameters analysed for that run. A green tick indicates that the sample meets quality requirements, yellow exclamation indicates that there may be errors and might need 'data grooming' while a red cross indicates that the data does not meet expected quality requirements set by the website.

1) Sample 1, reverse run

FastQC Report

Summary

-  [Basic Statistics](#)
-  [Per base sequence quality](#)
-  [Per tile sequence quality](#)
-  [Per sequence quality scores](#)
-  [Per base sequence content](#)
-  [Per sequence GC content](#)
-  [Per base N content](#)
-  [Sequence Length Distribution](#)
-  [Sequence Duplication Levels](#)
-  [Overrepresented sequences](#)
-  [Adapter Content](#)
-  [Kmer Content](#)

Basic Statistics

Measure	Value
Filename	1_ATCACG_L001_R2_001.fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	6356264
Sequences flagged as poor quality	0
Sequence length	93
%GC	65

2) Sample 2, forward run

FastQC Report **Basic Statistics**

Summary

- [Basic Statistics](#)
- [Per base sequence quality](#)
- [Per tile sequence quality](#)
- [Per sequence quality scores](#)
- [Per base sequence content](#)
- [Per sequence GC content](#)
- [Per base N content](#)
- [Sequence Length Distribution](#)
- [Sequence Duplication Levels](#)
- [Overrepresented sequences](#)
- [Adapter Content](#)
- [Kmer Content](#)

Measure	Value
Filename	2_CGATGT_L001_R1_001.fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	12693529
Sequences flagged as poor quality	0
Sequence length	93
%GC	65

3) Sample 2, reverse run

FastQC Report **Basic Statistics**

Summary

- [Basic Statistics](#)
- [Per base sequence quality](#)
- [Per tile sequence quality](#)
- [Per sequence quality scores](#)
- [Per base sequence content](#)
- [Per sequence GC content](#)
- [Per base N content](#)
- [Sequence Length Distribution](#)
- [Sequence Duplication Levels](#)
- [Overrepresented sequences](#)
- [Adapter Content](#)
- [Kmer Content](#)

Measure	Value
Filename	2_CGATGT_L001_R2_001.fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	12693529
Sequences flagged as poor quality	0
Sequence length	93
%GC	65

4) Sample 3, forward run

5_ACAGTG_L001_R1_001.fastq FastQC Report

 FastQC Report
 Fri 17 Jun 2016
 5_ACAGTG_L001_R1_001.fastq

Basic Statistics

Measure	Value
Filename	5_ACAGTG_L001_R1_001.fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	7342459
Sequences flagged as poor quality	0
Sequence length	93
%GC	65

Summary

-  [Basic Statistics](#)
-  [Per base sequence quality](#)
-  [Per tile sequence quality](#)
-  [Per sequence quality scores](#)
-  [Per base sequence content](#)
-  [Per sequence GC content](#)
-  [Per base N content](#)
-  [Sequence Length Distribution](#)
-  [Sequence Duplication Levels](#)
-  [Overrepresented sequences](#)
-  [Adapter Content](#)
-  [Kmer Content](#)

5) Sample 3, reverse run

5_ACAGTG_L001_R2_001.fastq FastQC Report

 FastQC Report
 Fri 17 Jun 2016
 5_ACAGTG_L001_R2_001.fastq

Basic Statistics

Measure	Value
Filename	5_ACAGTG_L001_R2_001.fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	7342459
Sequences flagged as poor quality	0
Sequence length	93
%GC	65

Summary

-  [Basic Statistics](#)
-  [Per base sequence quality](#)
-  [Per tile sequence quality](#)
-  [Per sequence quality scores](#)
-  [Per base sequence content](#)
-  [Per sequence GC content](#)
-  [Per base N content](#)
-  [Sequence Length Distribution](#)
-  [Sequence Duplication Levels](#)
-  [Overrepresented sequences](#)
-  [Adapter Content](#)
-  [Kmer Content](#)

6) Sample 4, forward run

6_GCCAAT_L001_R1_001.fastq FastQC Report
 FastQC Report
 Sun 19 Jun 2016
 6_GCCAAT_L001_R1_001.fastq

 **Basic Statistics**

Summary

-  [Basic Statistics](#)
-  [Per base sequence quality](#)
-  [Per tile sequence quality](#)
-  [Per sequence quality scores](#)
-  [Per base sequence content](#)
-  [Per sequence GC content](#)
-  [Per base N content](#)
-  [Sequence Length Distribution](#)
-  [Sequence Duplication Levels](#)
-  [Overrepresented sequences](#)
-  [Adapter Content](#)
-  [Kmer Content](#)

Measure	Value
Filename	6_GCCAAT_L001_R1_001.fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	5637829
Sequences flagged as poor quality	0
Sequence length	93
%GC	65

7) Sample 4, reverse run

6_GCCAAT_L001_R2_001.fastq FastQC Report
 FastQC Report
 Sun 19 Jun 2016
 6_GCCAAT_L001_R2_001.fastq

 **Basic Statistics**

Summary

-  [Basic Statistics](#)
-  [Per base sequence quality](#)
-  [Per tile sequence quality](#)
-  [Per sequence quality scores](#)
-  [Per base sequence content](#)
-  [Per sequence GC content](#)
-  [Per base N content](#)
-  [Sequence Length Distribution](#)
-  [Sequence Duplication Levels](#)
-  [Overrepresented sequences](#)
-  [Adapter Content](#)
-  [Kmer Content](#)

Measure	Value
Filename	6_GCCAAT_L001_R2_001.fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	5637829
Sequences flagged as poor quality	0
Sequence length	93
%GC	65

8) Sample 5, forward run

FastQC Report

Summary

-  [Basic Statistics](#)
-  [Per base sequence quality](#)
-  [Per tile sequence quality](#)
-  [Per sequence quality scores](#)
-  [Per base sequence content](#)
-  [Per sequence GC content](#)
-  [Per base N content](#)
-  [Sequence Length Distribution](#)
-  [Sequence Duplication Levels](#)
-  [Overrepresented sequences](#)
-  [Adapter Content](#)
-  [Kmer Content](#)

Basic Statistics

Measure	Value
Filename	3_TTAGGC_L001_R1_001.fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	7483517
Sequences flagged as poor quality	0
Sequence length	93
%GC	65

9) Sample 5, reverse run

FastQC Report

Summary

-  [Basic Statistics](#)
-  [Per base sequence quality](#)
-  [Per tile sequence quality](#)
-  [Per sequence quality scores](#)
-  [Per base sequence content](#)
-  [Per sequence GC content](#)
-  [Per base N content](#)
-  [Sequence Length Distribution](#)
-  [Sequence Duplication Levels](#)
-  [Overrepresented sequences](#)
-  [Adapter Content](#)
-  [Kmer Content](#)

Basic Statistics

Measure	Value
Filename	3_TTAGGC_L001_R2_001.fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	7483517
Sequences flagged as poor quality	0
Sequence length	93
%GC	65

10) Sample 6, forward run



FastQC Report

Summary

-  [Basic Statistics](#)
-  [Per base sequence quality](#)
-  [Per tile sequence quality](#)
-  [Per sequence quality scores](#)
-  [Per base sequence content](#)
-  [Per sequence GC content](#)
-  [Per base N content](#)
-  [Sequence Length Distribution](#)
-  [Sequence Duplication Levels](#)
-  [Overrepresented sequences](#)
-  [Adapter Content](#)
-  [Kmer Content](#)

 **Basic Statistics**

Measure	Value
Filename	4_TGACCA_L001_R1_001.fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	6431663
Sequences flagged as poor quality	0
Sequence length	93
%GC	65

11) Sample 6, reverse run



FastQC Report

Summary

-  [Basic Statistics](#)
-  [Per base sequence quality](#)
-  [Per tile sequence quality](#)
-  [Per sequence quality scores](#)
-  [Per base sequence content](#)
-  [Per sequence GC content](#)
-  [Per base N content](#)
-  [Sequence Length Distribution](#)
-  [Sequence Duplication Levels](#)
-  [Overrepresented sequences](#)
-  [Adapter Content](#)
-  [Kmer Content](#)

 **Basic Statistics**

Measure	Value
Filename	4_TGACCA_L001_R2_001.fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	6431663
Sequences flagged as poor quality	0
Sequence length	93
%GC	65