University of Nottingham
UK | CHINA | MALAYSIA

Exploring the structure, regulation, and function of the surface tethered *Pseudomonas aeruginosa* virulence factor, AaaA, and its role in maintaining chronic wound infections

---

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

Claire S. Laxton, MRes.

October 2022

Declaration

I declare that, unless otherwise acknowledged, the work presented in this thesis is my own. None of this work has been submitted in whole or in part for another degree at the University of Nottingham or any other institute of learning.

- Claire S. Laxton, October 2022

COVID-19 Impact Statement

This project was impacted by lab closures and restricted working hours for the majority of 2020-21 due to the COVID-19 pandemic. Delays in reagent and equipment deliveries and restricted access to microscopy facilities further limited productivity during this time. Therefore, though steps were taken to mitigate the impact of the pandemic, including the development of an alternative infection model, not all aspects of this project could be fully realised in time for submission.

Abstract

Pseudomonas aeruginosa is a leading cause of bacterial wound infections and is associated with disproportionately high mortality in burn patients, and morbidity in immunosuppressed and diabetic people, due to its ability to establish chronic wound infections. In this study, an *in vitro* synthetic wound model was used to examine the role of arginine-specific aminopeptidase of *P. aeruginosa* A (AaaA), a highly conserved, surface-tethered autotransporter which is known to be important for virulence during murine chronic wound infections. AaaA has potential as an antimicrobial drug or vaccine target due to its accessibility and immunogenicity, yet its mechanism of action is unclear. It is known that AaaA cleaves N-terminal arginine from peptides, which serve as a nutrient source in the oxygen and nutrient-limited environments of chronic wounds. Additionally, arginine can act as a signalling molecule in biofilm regulation, and it not yet clear what role AaaA may play in this regard. This study aimed firstly, to elucidate the structure and conservation of AaaA using both bioinformatic tools and by purifying AaaA for crystallography, and secondly to probe the gene expression and function of AaaA in a synthetic chronic wound (SCW) model.

In this study, AaaA was found to be highly conserved, with zero missense mutations in its active site residues in over 3000 *P. aeruginosa* genomes. A few, likely non-deleterious amino acid substitutions were identified to be common in almost all genomes except for PAO1, highlighting the deviation of this lab strain from other *P. aeruginosa* isolates. Some progress was made towards purifying a truncated version of AaaA, which did not contain the membrane-localising β-barrel, with some evidence that it exists as both a monomer and an SDS-resistant dimer, though further verification is needed.

Using transcriptional reporters and enzymatic AaaA activity assays, this study showed that AaaA is preferentially upregulated in the SCW, compared to in planktonic culture. AaaA was also shown to confer a modest but significant survival advantage in the SCW, as seen previously *in vivo*. This highlights the usefulness of this more disease-relevant model in studying

important virulence factors which have marginal phenotypes in planktonic conditions. Analysis using RT-qPCR showed that *rpoS*, but not arginine metabolism genes, was upregulated ~3-fold in an *aaaA* mutant, suggesting that loss of *aaaA* leads to an increased starvation response. Transcriptomics by RNA-Seq identified further quorum-sensing and RpoS-repressed genes involved in phenazine and alkyl-quinolone production, as well as possible chronic-infection-specific virulence factors, which were downregulated in the *aaaA* mutant, possibly as a result of increased RpoS expression. This is the first time a link between AaaA and RpoS has been demonstrated, both of which have likely distinct roles in nutrient foraging, potentially on either axis of acute or chronic infection phenotypes. Further study, particularly using proteomics, is required to better understand this relationship and how it relates to the growth differences seen in the *aaaA* mutant in the SCW.

Using a colorimetric quantification assay, no difference in extracellular arginine between wild-type and *aaaA* mutant-infected SCWs was detected. However, arginine levels in uninfected SCWs were ~3-fold higher, suggesting that arginine is being utilised in the SCW. A higher resolution approach, such mass spectrometry is required to detect differences related to AaaA. Finally, the Spytag-Spycatcher tagging system was used to localise AaaA in *P. aeruginosa* at the single cell level, with the potential that this technology could be used to localise AaaA in a multi-species biofilm in the future.

In summary, this study has created a number of tools for studying AaaA in chronic wound infections, including validating the use of the SCW. It has also uncovered new links between AaaA, RpoS and quorum sensing, to create an updated model of AaaA regulation in a chronic wound environment.

Acknowledgments

Firstly, I would like to thank my principal supervisor, Prof. Kim Hardie, for their support and guidance. Thanks also to my second and third supervisors, Dr Stephan Heeb (cloning king) and Dr Jeni Luckett (imaging queen). Thank you to the Wellcome Trust for generously funding my PhD. Thank you also to all my pre-PhD supervisors; Brenda, Kirsty and Tim, for the ways you supported and encouraged me to pursue science and made a PhD feel possible.

Thank you to the Hardie Group and everyone in C69, 75 and 81 for all the good times, the coffee breaks, lunches, pubs and parties. Thanks to Dr Ana da Silva for training me in the synthetic wound model and answering my 1000 bioinformatics questions, and to Marina Sedar and Dr Jack Leo, for introducing me to 'team Spycatcher'. Thanks to the UoN SLIM team, especially Dr Robert Markus, who is unrivalled in both his microscopy expertise and kindness. Thanks to Drs Phil Bardelang, Ewan Murray, Will Richardson and Tom Warwick for their endless help with all things proteins; to Drs Maria Papangeli, Danni Scales, Carol Paiva, Fran Smith and Dimitra Panagiotopoulou for all things bacterial genetics; to Izzy Blower for all things Western-blotting; to Dr Dean Walsh for all things imaging; and to Dr Sam Fenn for just, all things, and cheerfully sharing your expansive knowledge and beautiful singing voice with me and C75. Thank you to Bethan Roberts, for her many contributions to this work, her enthusiasm for science, and for making me a better teacher. A special mention goes out to the HelloFresh Recipes That Work cookbook for keeping me well fed- they really do work.

Thank you to Margherita Sosio and her wonderful team at Naicons, for graciously hosting me in Milan for a summer placement in 2019, and to the Singapore Centre for Environmental Life Sciences and Engineering for their generous scholarship to attend the SCELSE 2019 Biofilms summer school and learn from *the* masters. Between the two experiences, I learnt and ate more than I could have imagined, and gained my own international Biofilm in Nora, Josefien, Jimmy, Patricia, Andreas, and 'Mama' Sifat. Thank you also to the

vast team at Alderley Park Lighthouse Lab, for keeping me sane, and useful while the world was shuttered in 2020, you somehow helped me grow in a time of darkness.

Thank you to my Princesses, as well as my housemate Paulina, Wellcome-mates Scott and Emma, and office-mates in C57 and C51 for offering endless distractions, listening to my exploits and populating the JA with me. Beyond Nottingham, thank you to Flic and Emma, my Allmoo crew, my pal Lucy and all my other amazing friends, both inside and outside of science, who've walked this journey with me. I'm excited to see all the ways we'll change our little corners of the world.

A very special thank you goes to my family and to Kelly; you are my biggest and best supporters and cheerleaders. There aren't really words to say how grateful I am for you, for all you've given me, and how happy you make me. The next adventure is both exciting and daunting, but I leave Nottingham with a fuller heart than I ever imagined, and the best teammate I could ever have hoped for.

"And darlin' it was good, never looking down, and right there where we stood, was Holy Ground."

- Taylor Swift

Dedication

During this PhD, the COVID-19 pandemic ripped through the world, tearing lives, families and countries apart. I consider myself extremely lucky to have stayed safe, healthy, and with the support I needed to complete my PhD. To those who weren't this lucky, those still fighting to get back on track, and those who wrung themselves out to keep us all going, I want to dedicate this work to you. We cannot replace what we lost, but we can take what we've learnt to re-build a better, more equitable world.

Table of Contents

Declaration **2**

Abstract **3**

Acknowledgments **5**

Table of Contents **7**

List of figures *12*

List of tables *15*

List of abbreviations *16*

Chapter 1. Introduction to *Pseudomonas aeruginosa* in chronic wounds, autotransporters and AaaA **18**

1.1 Introduction to Pseudomonas aeruginosa *18*

1.1.1 Biofilms and chronic wounds 19

1.1.2 Transcriptional regulation in *P. aeruginosa* 21

1.1.3 Protein transport and secretion 27

1.2 Arginine utilisation by P. aeruginosa *31*

1.2.1 The ArgR regulon 32

1.2.2 Arginine sensing 32

1.2.3 Arginine import 33

1.2.4 Arginine catabolism 35

1.2.5 Role of arginine in *P. aeruginosa* signalling and biofilms 39

1.2.6 Arginine as a substrate for the host immune response 40

1.2.7 Potential for arginine in clinical intervention 41

1.3 Introduction to AaaA *42*

1.3.1 AaaA and arginine in chronic infections 44

1.4 PhD aims and objectives *46*

Chapter 2. Methods **47**

2.1 Strains, plasmids, and primers *47*

2.1.1 Strains used in this study 47

2.1.2 Plasmids used in this study 48

2.1.3 Primers used in this study 50

2.2 Bacterial media, growth, and storage conditions 52

2.3 SDS-PAGE and Western blotting 53

2.3.1 SDS-PAGE 53

2.3.2 Western blotting 53

2.4 Protein overexpression and purification 54

2.4.2 *aaaAt3* overexpression 54

2.4.3 Cell lysis 55

2.4.4 Column chromatography purification 56

2.5 Bacterial cloning 57

2.5.1 Polymerase chain reactions 57

2.5.2 Site directed mutagenesis 58

2.5.3 Restriction digest cloning 59

2.5.4 Transformation and conjugation 60

2.5.5 Sanger sequencing 62

2.5.6 Whole Genome Sequencing (WGS) 62

2.6 miniCTX-lux transcriptional reporter assays 63

2.7 L-arginine-p-nitroanilide cleavage (AaaA activity) assay 64

2.8 The synthetic chronic wound model (SCW) 65

2.8.1 Preparation of the SCW model 65

2.8.2 Colony forming unit determination from SCW samples 66

2.9 Examining gene expression in the SCW 67

2.9.1 For RT-qPCR 67

2.9.2 For RNA-Sequencing 69

2.10 Microscopy 73

2.10.1 Live/Dead staining of the SCW 73

2.10.2 Single cell super resolution structured illumination imaging 73

2.11 Colorimetric arginine quantification assay 74

2.12 Spycatcher binding assay 74

2.13 Bioinformatic analysis and modelling 75

2.14 Software and Data analysis 76

2.14.1 Data management and statistical testing 76

Chapter 3. Using bioinformatics to examine the structure and conservation of AaaA 78

3.1 Background on AaaA structure 78

3.2 Chapter Aims 81

3.3 Generating a predictive model of full-length AaaA 81

3.4 AaaA is highly conserved across P. aeruginosa strains 84

3.4.1 Examination of the PanX genome database shows high AaaA conservation 84

3.5 There are some common and consistent deviations from PA0328 90

3.6 Using site-directed mutagenesis to probe the effect of variations on AaaA activity 101

3.6.1 D580E may have a detrimental effect on AaaA activity 101

3.7 Discussion 102

3.7.1 Previous work suggests the mutations in the *aaaA* sequence only partially impact AaaA activity 104

Chapter 4. Progress toward elucidating the AaaA protein structure 107

4.1 Previous work to purify AaaA 107

4.2 Chapter aims 108

4.3 Over-production and purification of AaaAt3 109

4.3.1 Induction of AaaAt3 with 0.1-0.3 mM IPTG at 22˚C for 5hrs gives a good yield 109

4.3.2 Lysis method selection 110

4.3.3 A portion of AaaAt3 is soluble, but it cannot be efficiently purified by metal ion affinity chromatography 111

4.3.4 Anionic exchange can be used to partially purify AaaAt3 from crude lysate 114

4.3.5 Size exclusion chromatography can further purify active, soluble AaaAt3 117

4.4 Discussion 119

4.4.1 Alternative protein tags for affinity chromatography of AaaA 119

4.4.2 Optimisation of strain background 119

4.4.3 Possibilities for over-production of a near-native AaaA construct 120

4.4.4 Conclusions 122

Chapter 5. Probing *aaaA* expression and regulation in a synthetic chronic wound model 123

5.1 The state of current knowledge regarding aaaA regulation 123

5.1.1 Previous interrogation of the *aaaA* promoter found potential *aaaA* regulators 123

5.1.2 Overview of experimentally determined *aaaA* regulators 125

5.2 Chapter aims 137

5.3 Construction of ΔaaaA mutants in PAO1L and a chronic wound isolate 138

5.3.1 The PAO1N subline and its drawbacks 138

5.3.2 Successful construction of an in-frame *aaaA* deletion in PAO1L and CW2-BI1 140

5.4 Development of the synthetic chronic wound model to study AaaA 152

5.4.1 Introduction to the synthetic chronic wound model 152

5.4.2 AaaA confers a growth advantage in 24 hr SCW biofilms 154

5.5 Examining aaaA transcription and translation in the SCW 156

5.5.1 *aaaA* promoter activation is higher in the SCW than in LB, but previously observed autorepression was not replicated 156

5.5.2 AaaA activity is higher when grown in SCW than planktonically 160

5.6 Development of RT-qPCR for monitoring aaaA expression in the SCW 162

5.6.1 Primer design and validation 162

5.6.2 *aaaA* is transcribed in the SCW at a low level, and is greatest at 16 hrs 164

5.6.3 Arginine-related genes are not differentially expressed in the *aaaA* mutant-infected SCW, but *rpoS* is upregulated 165

5.7 Using RNA-Seq to further study SCW infection transcriptomics in relation to aaaA 167

5.7.1 Introduction and methodology 167

5.7.2 RNA-Seq QC and mapping results 169

5.7.3 Differential expression analysis supports RT-qPCR findings, except differential expression of *rpoS* was not apparent 172

5.7.4 Genes related to phenazine and PQS biosynthesis, biofilm formation and oxidative stress response are downregulated in the Δ*aaaA* mutant 173

5.7.5 Genes related to alternative nutrient metabolism and uptake are upregulated in the Δ*aaaA* mutant 179

5.7.6 Downregulation of *phz*, *chiC* and *tes* genes also occurs in arginine-supplemented Δ*aaaA* mutant, along with upregulation of nitric oxide reductase pathways 182

5.7.7 Other notable findings from RNA-Seq 183

5.8 Discussion 185

5.8.1 AaaA is produced and relevant in the synthetic chronic wound 185

5.8.2 There is no evidence that AaaA directly or indirectly modulates arginine metabolism gene expression 186

5.8.3 A newly uncovered relationship between RpoS and AaaA 187

5.8.4 Potential clues from RNA-Seq link AaaA to phenazine biosynthesis and quorum sensing 191
5.8.5 Ethanol oxidation could be a potential compensatory pathway to arginine starvation 195
5.8.6 Conclusions 197

Chapter 6. Probing AaaA localisation and function in a synthetic chronic wound model 200

6.1 Chapter aims 200

6.2 Localising AaaA using the Spytag-Spycatcher system 200
6.2.1 Previous work to localise AaaA using microscopy 200
6.2.2 Background on the Spytag-Spycatcher system 201
6.2.3 Design of a Spytagged-AaaA construct 203
6.2.4 Use of the pJM220 vector for rhamnose induction of Spytagged AaaA in *P. aeruginosa* 206
6.2.5 Construction of Spytagged AaaA 206
6.2.6 The pJM220 vector can be used to tightly control AaaA expression in a concentration-dependant manner 209
6.2.7 Spytagged AaaA version 4 is active and accessible to Spycatcher binding 210
6.2.8 Spytagged AaaA can be visualised on the surface of single cells 213

6.3 Testing the arginine liberation hypothesis for AaaA in SCW 216
6.3.1 There is no difference in free arginine between wild-type and Δ*aaaA*-infected SCW, but arginine levels are reduced by infection 216

6.4 Discussion 218
6.4.1 AaaA can be localised at the single cell level using the Spytag-Spycatcher system 218
6.4.2 AaaA did not observably alter free arginine levels in the SCW, however arginine utilisation was likely occurring 220
6.4.3 Alternative approaches to examining AaaA function 221

Chapter 7. Final remarks 224

References 228

Supplementary Data 266

List of figures

Figure 1.1: Summary of major TCS in P. aeruginosa. 22

Figure 1.2: Schematic illustrating the role of sigma factors in transcription. 24

Figure 1.3: Summary of the major P. aeruginosa secretion systems. 28

Figure 1.4: Schematic of different roles of arginine in P. aeruginosa infection. 32

Figure 1.5: Summary of the four arginine catabolic pathways in P. aeruginosa. 36

Figure 1.6: Illustration of the L-arginine-p-nitroanilide cleavage assay. 43

Figure 3.1: Map of AaaA primary structure. 80

Figure 3.2: Predicted 3D model of AaaA. 82

Figure 3.3: Comparison of modelling techniques for predicting a 3D model of AaaA. 83

Figure 3.4: Tree of relatedness for AaaA created using the PanX database of 153 P. aeruginosa isolates. 85

Figure 3.5: Alignment of AaaA amino acid sequence in 153 P. aeruginosa isolates in the PanX genome database. 89

Figure 3.6: Cladogram showing representatives of all the unique complete sequences for AaaA in P. aeruginosa in the Pseudomonas.com database. 92

Figure 3.7: Map of AaaA primary structure with predicted active site residues and identified amino acid variations. 95

Figure 3.8: Amino acid variations in AaaA mapped onto the 3D model. 96

Figure 3.9: Alignment of the 50 unique AaaA amino acid sequences identified in the 3748 genomes containing a complete AaaA sequence in the Pseudomonas.com database. 98

Figure 3.10: Consensus_aaaA based on common variations. 99

Figure 3.11: Predicted 3D alignment of P0328 and Consensus_aaaA model. 100

Figure 3.12: Whole cell AaaA activity assay revealed D580E variation had a minor deleterious effect upon AaaA activity. 102

Figure 3.13: Variations in AaaA from clinical isolates have stochastic effects on activity. 105

Figure 3.14: Variations in AaaA in CF isolates and their effect on AaaA activity and P_{aaaA} activation. 106

Figure 4.1: AaaAt3 induction optimisation. 109

Figure 4.2: Yield of active AaaAt3 in soluble and insoluble fractions of over-produced AaaAt3, from three lysis methods. 110

Figure 4.3: AaaAt3 does not appear to bind to the Ni-affinity chromatography column. 111

Figure 4.4: L-arginine-p-nitroanilide cleavage (AaaA activity) assay of AaaAt3 purification by Co-TALON resin chromatography shows AaaAt3 does not bind well to Co-TALON resin. 112

Figure 4.5: Immunoblot to detect AaaA using α-his, HRP-conjugated antibody against whole cell inductions. 112

Figure 4.6: AaaAt3 with 8xHis only slightly improves binding to the HisTrap column. 113

Figure 4.7: AaaAt3 with an 10x His tag also does not bind better to a HisTrap Ni column. 113

Figure 4.8: AaaAt3 can be partially purified from crude lysate by anionic exchange. 116
Figure 4.9: Anionic exchange fractions can be further purified by size exclusion chromatography. 118
Figure 4.10: Illustration of proposed AaaA-tev construct and purification methodology. 121
Figure 5.1: Schematic of the aaaA promoter and its putative transcriptional regulator binding sites identified by Spencer (2018). 124
Figure 5.2: Illustration of miniCTX::P_{aaaA}-lux transcriptional reporter assay. 125
Figure 5.3: Summary of known and putative regulators of aaaA. 136
Figure 5.4: Schematic of cloning strategy for creation of suicide vector for making an in-frame aaaA deletion. 141
Figure 5.5: Cloning for construction of pEX18Gm::ΔaaaA suicide vector, using restriction-based cloning. 142
Figure 5.6: Colony PCR of PAOL ΔaaaA candidate post-sucrose selection. 143
Figure 5.7 Amino acid alignment between PA0328 (1) and aaaA from CW2-BI1 (2). 149
Figure 5.8: AaaA activity is lost in PAO1L ΔaaaA, but intrinsically low in CW2-BI1, with greater non-specific activity in CW2-BI1 ΔaaaA, which is unchanged by complementation. 150
Figure 5.9: AaaA activity varies between P. aeruginosa chronic wound isolates. 151
Figure 5.10: The SCW model used in this study, infected with PAO1L wild type and ΔaaaA. 153
Figure 5.11: Colony forming units provide evidence that aaaA confers a survival advantage at 24 hrs in the SCW. 154
Figure 5.12: Live/Dead staining shows aaaA confers a survival advantage at 24 hrs in the SCW. 155
Figure 5.13: Activation of both P_{aaaA} and P_{kan} is higher in the SCW and P_{aaaA} auto-repression is only present in CW2-BI1. 157
Figure 5.14: Use of a lid instead of a gas-permeable seal makes a negligible difference to P_{aaaA} or P_{kan} lux expression. 159
Figure 5.15: Comparison PAO1L and CW2-BI1 AaaA activity in planktonic and synthetic chronic wound contexts. 161
Figure 5.16: Standard curves for selected primer sets on PAO1L gDNA. 164
Figure 5.17: RT-qPCR of SCW for relative quantification of aaaA over time shows aaaA is most highly transcribed at 16 hrs. 165
Figure 5.18: RT-qPCR results comparing PAO1L and PAO1L ΔaaaA in the SCW after 16 hrs show rpoS is upregulated in the mutant. 166
Figure 5.19: Pipeline for RNA-Seq analysis. 169
Figure 5.20: Pearson correlation between RNA-Seq samples. 172
Figure 5.21: Differential expression of genes in PAO1L ΔaaaA compared with PAO1L. 175
Figure 5.22: Dot plot showing the most significantly enriched KEGG pathways for genes downregulated in the ΔaaaA mutant compared to wild-type. 176
Figure 5.23: Dot plot showing the most significantly enriched KEGG pathways for genes upregulated in the ΔaaaA mutant compared to wild-type. 177

Figure 5.24: Differential expression of genes in PAO1L ΔaaaA supplemented with 100 µM L-arginine, compared with unsupplemented PAO1L. 182
Figure 5.25: Amino acid alignment of AaaA (1, top) with PaAP (2, bottom). 189
Figure 5.26: Potential poorly conserved binding site for RpoS exist upstream of aaaA. 191
Figure 5.27: Summary of the known and putative aaaA regulators as well as potential downstream effects of aaaA expression uncovered in this study. 199
Figure 6.1: Illustration of the Spytag-Spycatcher system for labelling AaaA. 202
Figure 6.2 Design of Spytag-containing AaaA constructs. 205
Figure 6.3: Cloning strategy for the construction of Spytagged aaaA and insertion into the pJM220 rhamnose-inducible mini Tn7 vector. 208
Figure 6.4: AaaA activity can be tightly induced in the pJM220-aaaA construct. 210
Figure 6.5: AaaA activity assay in rhamnose-induced Spytagged AaaA shows both AaaA-St4 and -St5, but not -St6 have similar activity levels to native AaaA (non-Spytagged) induction. 212
Figure 6.6: Spytag-Spycatcher binding assay shows that only aaaA-St4 binds Spycatcher-mCherry more than the non-Spytagged AaaA. 213
Figure 6.7: Structured illumination microscopy captures Spytagged AaaA on the surface of single cells. 214
Figure 6.8: Spytagged-AaaA can be detected by structured illumination microscopy on the surface of fixed cells. 215
Figure 6.9: Colorimetric arginine quantification assays to compare arginine levels in the SCW between PAO1L wild-type (WT) and ΔaaaA. 217
S-Figure 0.1: Addition of $ZnCl_2$ to fractions of AaaAt3H10 following HisTrap chromatography yielded no difference in activity. 266
S-Figure 0.2: Alignment of sequenced inserts mapped onto pEX18Gm-ΔaaaA vector. 266
S-Figure 0.3: Alignments of sequenced PCR products from PAO1L ΔaaaA candidates C7 and C13, spanning the region around PA0328, mapped onto the PAO1 genome. 267
S-Figure 0.4: Activation of promotors of Kan and aaaA genes in the SCW vs LB over 19 hrs. 268
S-Figure 0.5: Alignments of sequenced products show successful construction of Spytagged AaaA. 281

List of tables

Table 2.1: List of strains used in this study. 47

Table 2.2: List of plasmids used in this study. 48

Table 2.3: List of primers used in this study. 50

Table 2.4: List of media used in this study. 52

Table 3.1: PROVEAN scores for all amino acid substitutions present in AaaA in the 153 PanX genomes. 86

Table 3.2: Details of the 15 isolates from the 153 P. aeruginosa genomes on the PanX database with a potentially deleterious L170Q substitution in AaaA. 87

Table 3.3: PROVEAN scores for all amino acid substitutions identified in AaaA in the 3748 genomes containing a complete AaaA sequence in the Pseudomonas.com database. 93

Table 3.4: The prevalence of each of the 3 substitutions identified in the aaaA sequence which are predicted to be deleterious. 94

Table 4.1: Details of the existing AaaA truncated constructs created by Drs Paredes Osses and Spencer. 108

Table 5.1: List of transcriptional factors with putative binding sites in the aaaA promoter. 124

Table 5.2: Genes previously shown to alter P_{aaaA}-lux activation. 127

Table 5.3: Genes previously shown to effect AaaA enzymatic activity. 128

Table 5.4: List of non-synonymous variations identified in PAO1L ΔaaaA compared with the PAO1L reference genome. 146

Table 5.5: Amino acid variations in AaaA in clinical wound isolates. 152

Table 5.6: Genes selected for RT-qPCR analysis and their primer binding sites. 163

Table 5.7: QC of RNA-Seq samples. 170

Table 5.8: Mapping of RNA-Seq samples to the reference PAO1 genome. 171

Table 6.1: Summary of location choices for 6 Spytag construct candidates. 203

S-Table 0.1: List of all differentially expressed genes in PAO1L ΔaaaA compared with PAO1L in the 16-hr SCW ($p<0.05$). 269

S-Table 0.2: List of all differentially expressed genes in PAO1L + 100 µM L-arginine compared with PAO1L in the 16-hr SCW. 274

S-Table 0.3: List of all differentially expressed genes in PAO1L ΔaaaA + 100 µM L-arginine compared with PAO1L. 277

List of abbreviations

$3OC_{12}$-HSL	N-3-oxo-dodecanoyl-homoserine lactone
ADC	arginine decarboxylase
ADH	arginine dehydrogenase
ADI	arginine deiminase
AHL	N-acylated homoserine lactones
AMR	antimicrobial resistance
AQ	alkyl-quinolone
ASM	artificial sputum media
AST	arginine succinyltransferase
ATA	arginine transaminase
ATP	adenosine triphosphate
AUC	area under the curve
BAM	β-barrel assembly machinery
BLAST	basic local alignment search
c-di-GMP	bis-(3′-5′)-cyclic diguanosine monophosphate
C_4-HSL	N-butanoyl-homoserine lactone
cAMP	3',5'-cyclic adenosine monophosphate
cDNA	complementary DNA
CF	cystic fibrosis
CI	confidence interval
CIP	alkaline phosphatase, calf intestinal
cOTC	catabolic ornithine carbamoyltransferase
DNA	deoxyribose nucleic acid
EBP	enhancer binding protein
eDNA	extra-cellular DNA
EPS	extra-cellular polymeric substances
FDR	false discovery rate
FPKM	Fragments Per Kilobase of transcript per Million base pairs
gDNA	genomic DNA
HHQ	2-heptyl-4-quinolone
IgA/G	immuno-globulin A/G
IL	interleukin
IM	inner-membrane
INDEL	insertion/deletion
IPTG	isopropyl-β-D-thiogalactoside
LB	Lysogeny broth
Log2FC	Log2 fold change
LPS	lipopolysaccharide
MDR	multi-drug resistance
MMP	Minimal media P
mRNA	messenger RNA
NAD	nicotinamide adenine dinucleotide
NAD-GDH	NAD-dependant glutamate dehydrogenase
NADP	nicotinamide adenine dinucleotide phosphate
NF	nuclease-free
NO	nitric oxide
NOS	nitric oxide synthase

OM	outer membrane
Omp	outer membrane protein
OMV	outer membrane vesicle
ORF	open reading frame
PBS	phosphate buffered saline
PBST	PBS, 0.5% Tween 20
PCR	polymerase chain reaction
PDE	phosphodiesterase
PN	L-arginine-p-nitroanilide
PNK	polynucleotide kinase
PQS	*Pseudomonas* quinolone signal (hydroxy-4(1H)-quinolone)
QS	quorum sensing
RCF	relative centrifugal force
RFU	relative fluorescence unit
RGD	arginyl-glycyl-aspartic acid
RLU	relative light units
RMSD	root-mean square deviation
RNA	ribose nucleic acid
RPM	revolutions per minute
RT	reverse transcriptase
RT-qPCR	reverse transcriptase, real-time quantitative PCR
SCW	synthetic chronic wound
SD	standard deviation
SDM	site-directed mutagenesis
SDS	sodium dodecyl sulphate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
ShD	Shine-Dalgarno sequence
SMA	styrene maleic acid
SMALP	SMA lipid particles
SNP	single nucleotide polymorphism
St	Spytag
SWF	synthetic wound fluid
TAT	twin-arginine transporter
TCA	tricarboxylic acid
TCS	two-component system
TDNN	time delay neural network
TNF-α	tumour necrosis factor alpha
UTR	Untranslated region
WGS	whole genome sequencing

Chapter 1. Introduction to *Pseudomonas aeruginosa* in chronic wounds, autotransporters and AaaA

1.1 Introduction to *Pseudomonas aeruginosa*

Pseudomonas aeruginosa is a Gram-negative bacterial species, and a World Health Organisation-designated high-priority 'ESKAPE' pathogen, which can cause serious, nosocomial, multi-drug resistance (MDR) infections (Santajit & Indrawattana, 2016). *P. aeruginosa* infections particularly affect immunocompromised individuals, such as the very young and old, and those with cancer and diabetes. *P. aeruginosa* is also a major hospital-acquired pathogen, chiefly responsible for hospital-acquired respiratory tract infections, as well as being associated with the greatest sepsis-related mortality rate in burn patients (Norbury *et al.*, 2016). Individuals with cystic fibrosis (CF) are also especially vulnerable to *P. aeruginosa* respiratory infections, which can dramatically worsen their overall prognoses, as once established, clearance is almost impossible (Henry *et al.*, 1992). In the context of chronic wounds, which affect 1-2% of the population, infection with *P. aeruginosa*, particularly of diabetic ulcers, delays healing and leads to increased morbidity, which is increased further when co-infecting with *Staphylococcus aureus* (Serra *et al.*, 2015).

Clinical *P. aeruginosa* isolates tend to have high levels of antimicrobial resistance (AMR), due to intrinsic tolerance and resistance mechanisms and acquisition of AMR genes by horizontal gene transfer, which can further limit treatment outcomes in chronic infections (Horcajada *et al.*, 2019). Infection with MDR *P. aeruginosa* is associated with a 2-fold increase in all-cause mortality in hospitalised patients, as well as increased length of hospital stay (Nathwani *et al.*, 2014). One major tolerance mechanism which contributes to AMR is the proclivity for *P. aeruginosa* to form and persist within complex

biofilms, which provide protection from antimicrobials and host immune factors (Hogardt & Heesemann, 2010).

1.1.1 Biofilms and chronic wounds

1.1.1.1 The biofilm matrix

A biofilm is defined as a multicellular community, bound to a solid surface or in self-adhered aggregates, in which microorganisms live embedded in a self-produced extra-cellular matrix made of extra-cellular polymeric substances (EPS) which have structural and protective roles (Flemming *et al.*, 2007). Though bacterial cells can live individually, biofilms are considered to be the prevailing mode of life for micro-organisms in the environment (Flemming & Wingender, 2010). The EPS matrix acts as a physical scaffold which attaches cells together, and to substrata, and is the basis for the 'emergent properties' of biofilms. These emergent properties include social cooperation, resource capture, horizontal gene transfer, and enhanced survival following exposure to antimicrobials, and create a lifestyle substantially different to that of free-living bacterial cells (Flemming *et al.*, 2016; Stalder & Top, 2016)

EPS mainly consists of extracellular DNA (eDNA), proteins, and exopolysaccharides. eDNA is found in biofilms of almost all species, and while initially thought to be incorporated from lysed cells, it has been shown to form defined, stable filamentous networks (Böckelmann *et al.*, 2006). This heavily influences the structural architecture of the biofilm, through interactions with other EPS components, as well as potentially propagating cell-cell communication (Campoccia *et al.*, 2021; Flemming *et al.*, 2007).

In *P. aeruginosa*, the primary exopolysaccharides are Psl, Pel and alginate. Each is expressed at different stages of biofilm development and confer different properties upon it (Cendra & Torrents, 2021). For example, chronic CF *P. aeruginosa* infections are typified by a switch from non-mucoid to mucoid phenotype, in which alginate is overproduced, although Psl still plays an essential role (Martin *et al.*, 1993; Ma *et al.*, 2012). The matrix adhesin CdrA also binds to Psl, and this complex stabilises the matrix, to allow bacteria

to form robust, protease resistant aggregates (Reichhardt *et al.*, 2018). CdrA, Psl and Pel are all dependant on the second messenger bis-(3′-5′)-cyclic diguanosine monophosphate (c-di-GMP) which is discussed later in 1.1.2.4 (Borlee *et al.*, 2010).

P. aeruginosa additionally produces two soluble lectins, LecA and LecB, which both bind to other matrix components to stabilise the matrix and enhance biofilm formation. LecA crosslinks with sugars, particularly galactose, present in matrix components, and LecB binds to Psl (Diggle *et al.*, 2006; Passos da Silva *et al.*, 2019). Interestingly, isopropyl-β-D-thiogalactoside (IPTG), commonly used as an inducer for overexpression vectors, was shown to inhibit biofilm formation via competitive inhibition of LecA (Diggle *et al.*, 2006). Thus, induction systems using IPTG are less suitable for studying *P. aeruginosa* biofilms.

1.1.1.2 Chronic wound biofilms

Chronic wounds are defined as wounds which do not progress through the healing stages, but are arrested in a predominantly inflammatory phase (Agren *et al.*, 2000). A number of bacterial species are frequently isolated from chronic wounds, including *S. aureus*, *Enterococcus faecalis*, *P. aeruginosa*, coagulase-negative staphylococci, *Proteus* species and other anaerobic bacteria (Gjødsbøl *et al.*, 2006). *S. aureus* and *P. aeruginosa* are the most common chronic wound colonisers, and are both associated with delayed wound healing, with the presence of *P. aeruginosa* also associated with enlarged wounds (Gjødsbøl *et al.*, 2006; Halbert *et al.*, 1992; Madsen *et al.*, 1996).

Contrary to the common conception of biofilms forming lawns or ‘mushroom’ structures attached to a surface, chronic wound biofilms are characterised by the formation of microcolony aggregates, embedded in EPS, in a non-random distribution (Kirketerp-Møller *et al.*, 2008). Interestingly, using imaging techniques, *P. aeruginosa* has been found in samples which were negative by culture-based detection techniques, suggesting that the actual prevalence of *P. aeruginosa* in chronic wounds is much higher than previously predicted (Kirketerp-Møller *et al.*, 2008). Cornforth et al. (2018) showed that the

transcriptomes of *P. aeruginosa* in *in vivo* and *in vitro* models, including a range of chronic and acute wound models can vary substantially, highlighting the importance of using a disease and stage-relevant model when studying chronic wound biofilms.

1.1.2 Transcriptional regulation in *P. aeruginosa*

The first *P. aeruginosa* genome sequence revealed a staggering 10% of its 5570 genes were predicted to encode regulators of other genes (Stover *et al.*, 2000). The following section details a few of the many gene regulation pathways utilised by *P. aeruginosa* to co-ordinate this vast, environmentally-responsive repertoire of metabolic pathways, virulence factors, motility mechanisms and biofilm life cycle stages. As discussed below, these systems are often highly interlinked, depending on a series of hierarchies and system cross-regulation.

1.1.2.1 Two component systems

Two component systems (TCS) are heavily relied upon by *P. aeruginosa* and other bacteria to respond directly to environmental cues. They consist of two components: 1) a membrane-bound histidine kinase which responds to external stimuli by adenosine triphosphate (ATP)-dependant autophosphorylation, and transfers its phosphoryl group to 2), a cognate response regulator which, when activated by phosphorylation, acts as a transcription factor to alter gene expression (Stock *et al.*, 2000). *P. aeruginosa* encodes a huge range of TCSs, many of which are highlighted in **Figure 1.1**, with individual TCSs described later, where pertinent.

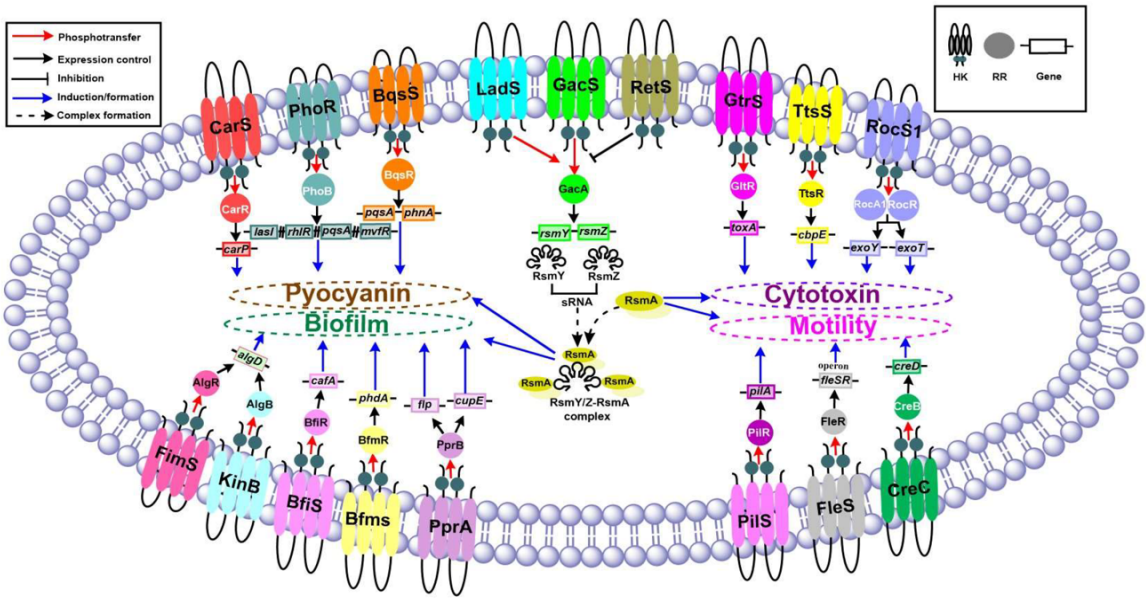


Figure 1.1: Summary of major TCS in P. aeruginosa. *Abbreviations: HK- histidine kinase, RR- response regulator. Taken from* Sultan et al. (2021).

1.1.2.2 Quorum sensing

A further layer of gene regulation is controlled by quorum sensing (QS), the phenomena of bacteria sensing and responding to their population density via self-produced small diffusible molecules. In *P. aeruginosa*, these are primarily via the *las* and *rhl* systems. Each system consists of a signal synthase, LasI or RhlI which produce the diffusible signals known as *N*-acylated homoserine lactones (AHLs). LasI produces *N*-3-oxo-dodecanoyl-homoserine lactone ($3OC_{12}$-HSL), and RhlI produces the *N*-butanoyl-homoserine lactone (C_4-HSL). Each signal binds to its cognate signal receptor, LasR or RhlR, which then induce transcription of a range of genes, including *lasI* and *rhlI*, in a positive feedback loop known as autoinduction (Schuster & Greenberg, 2006). LasR-LasI also hierarchically activates RhlR-RhlI (Latifi *et al.*, 1996).

P. aeruginosa also encodes a third QS system based on the production of a large variety of alkyl-quinolone (AQ) signal molecules (Heeb *et al.*, 2011). This is controlled by PqsR (also known as MvfR) which binds to the *Pseudomonas* quinolone signal, hydroxy-4(1H)-quinolone, (PQS), as well as its precursor 2-heptyl-4-quinolone (HHQ) to drive expression of genes involved in virulence and biofilm formation (Pesci *et al.*, 1999). PqsR also induces the *pqsABCDE* operon to produce HHQ, which is converted to PQS by the monooxygenase PqsH, although the latter step is controlled by LasR (Déziel *et*

al., 2004). PqsR positively regulates itself as well as *lasR, rhlR* and *rhlI* (McKnight *et al.*, 2000). Therefore, the three QS systems are co-operatively regulated by their AQ and AHL products, and work in a hierarchical fashion to control a number of QS virulence factors including toxins and secreted enzymes including pyocyanin, elastase, lectins and rhamnolipids, and genes involved in biofilm formation and dispersal (Lee & Zhang, 2015).

1.1.2.3 Sigma factors

Sigma (σ) factors are bacterial transcriptional initiation factors. As illustrated in **Figure 1.2**, these proteins form a complex with RNA polymerase, to create the RNA polymerase holoenzyme, which then binds to a σ-specific recognition site in the gene promoter, at -10 and -35 bp upstream of the open reading frame (ORF) (Gruber & Gross, 2003). Binding at the -10 site, which is AT rich, initiates melting of the double stranded DNA to form a small single stranded 'bubble' called the open complex. The now single-stranded AT rich -10 region can then enter the RNA polymerase active site for transcription (Abril *et al.*, 2020).

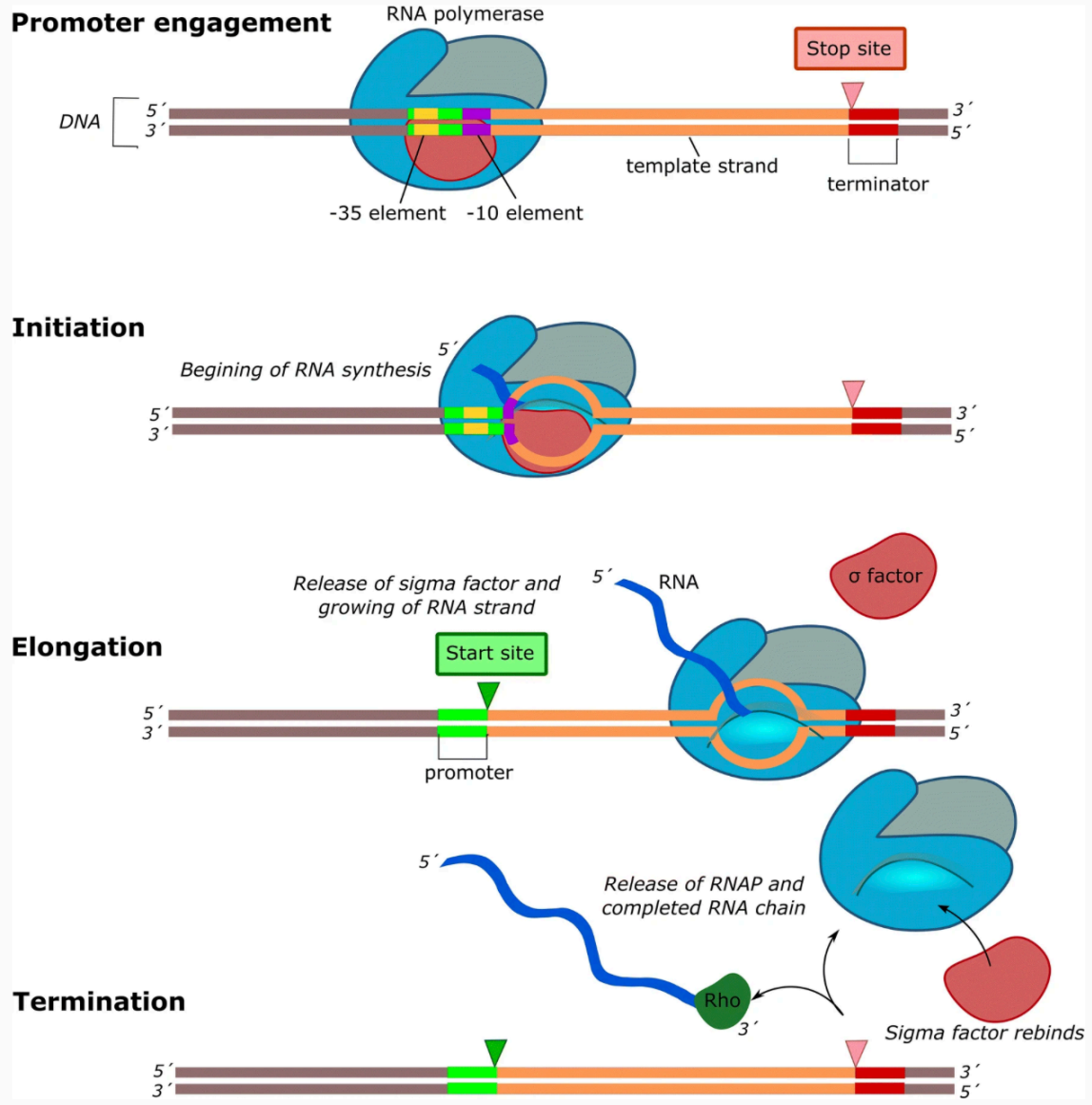


Figure 1.2: Schematic illustrating the role of sigma factors in transcription. *The sigma factor (red protein) which is reversibly bound to RNA polymerase (blue protein) recognises and 'engages' the -10 (purple) and -35 (yellow) element in the gene promoter (green). Binding at the -10 site initiates melting of the double stranded DNA to form a transcription bubble of single stranded DNA, which can enter the RNA polymerase active site, initiating transcription of RNA (blue strand) at the gene start site. During elongation, the sigma factor dissociates, and Rho factor (green protein) moves along the growing RNA strand following the RNA polymerase. When the transcriptional terminator (stop site) is reached, Rho unwinds the DNA-RNA hybrid within the transcriptional bubble, and Rho, RNA polymerase and the new RNA are released. Taken from* Abril et al. (2020)

This mechanism allows for efficient, global control of the transcription of hundreds of genes via a small group of proteins. Sigma factors were originally described in *Escherichia coli* (*E. coli*) and their numerical designation originates from their molecular weight in kDa, e.g., sigma factor 70 (σ^{70}, RpoD) is 70 kDa (Helmann & Chamberlin, 1988). *P. aeruginosa* encodes 24 putative

sigma factors, the principal, RpoD, and a number of 'alternative' sigma factors, which have been reviewed extensively elsewhere (Potvin *et al.*, 2008). Those most pertinent to this study, RpoD, RpoS and RpoN are described briefly here.

1.1.2.3.1 σ^{70}, RpoD

RpoD in *P. aeruginosa* shares extensive homology with *E. coli* RpoD, the 'major' or principal sigma factor, which controls the expression of housekeeping genes during exponential growth (Fujita *et al.*, 1994; Tanaka & Takahashi, 1991).

1.1.2.3.2 σ^{s}, RpoS

In *P. aeruginosa*, RpoS is involved in co-ordinating responses to starvation and cellular stress during stationary phase growth, such as heat, low pH, high osmolarity, hydrogen peroxide and ethanol. However, it is not as important for these functions as its *E. coli* ortholog (Jørgensen *et al.*, 1999). RpoS is also closely linked with the QS systems in *P. aeruginosa*, where it has been shown to be both induced by RhlR, in response to C_4-HSL, and act as a repressor of RhlI (Latifi *et al.*, 1996; Whiteley *et al.*, 2000). Therefore, deletion of *rpoS* leads to reduced production of exotoxin A and slightly reduced LasA and LasB, but increased production of phenazines and pyocyanin (Suh *et al.*, 1999). RpoS is also positively regulated by the TCS GacS-GacA in *P. fluorescens* (Whistler *et al.*, 1998).

1.1.2.3.3 σ^{54}, RpoN

The primary role for RpoN was initially thought to be in co-ordinating nitrogen assimilation, however evidence has since emerged for its role in virulence phenotypes, cell motility, pili formation, nutrient transport, mucoid phenotypes and cell-cell signalling (Hendrickson *et al.*, 2001; Potvin *et al.*, 2008). RpoN also interacts with QS systems by positively regulating RhlI, and in turn increasing C_4-HSL, to drive elastase and pyocyanin production (Thompson *et al.*, 2003).

Due to differences in its structure, unlike RpoS and RpoD, RpoN does not automatically transition into the 'open complex' upon binding to its recognition site, but rather requires energy from the hydrolysis of ATP by transcriptional activators known as enhancer binding proteins (EBPs), which

bind 80 to 150 bp upstream of the promoter (Bush & Dixon, 2012; Schumacher *et al.*, 2004). As the EBP binding sites are further away from the RpoN binding site, an additional integration host factor (IHF) is also required to bend the DNA and form DNA loops which allow interaction between the RpoN holoenzyme and the EBPs (Carmona & Magasanik, 1996; Rippe *et al.*, 1997). Proteins containing this σ^{54} ATP-binding region include flagella and pilin biosynthesis genes *fleQ*, *fleR* and *pilR* as well as *algB*, which is involved in alginate exopolysaccharide production (Dasgupta *et al.*, 2003; Ma *et al.*, 1998; Mattick *et al.*, 1996).

1.1.2.4 Second messengers: c-di-GMP and cAMP

C-di-GMP is an environmental signal transduction molecule involved in a range of cellular processes in *P. aeruginosa.* Generally, low c-di-GMP concentrations are associated with the planktonic lifestyle and biofilm dispersal, whereas high c-di-GMP levels are associated with biofilm formation and sessility (Hengge, 2009).

While c-di-GMP generally upregulates sessility and biofilm related genes, another second messenger, 3',5'-cyclic adenosine monophosphate (cAMP), has the inverse role of upregulating acute phase virulence factors (Smith *et al.*, 2004). This is mediated by the global cAMP-dependant transcription factor Vfr, which upregulates the type III secretion system, type IV pili components (both described in 1.1.3.2), twitching motility, and endotoxin A, elastase, pyocyanin and other QS-related virulence factors (Albus *et al.*, 1997; Beatson *et al.*, 2002; West *et al.*, 1994; Wolfgang *et al.*, 2003). Vfr also downregulates FleQ and flagella biogenesis (Dasgupta *et al.*, 2002).

C-di-GMP and cAMP also exist on a feedback axis with each other; cAMP-Vfr induces the production of the phosphodiesterases (PDEs) DipA, RbdA and particularly BifA, which degrade c-di-GMP (Almblad *et al.*, 2019; Fuchs *et al.*, 2010). Likewise, high levels of c-di-GMP reduce cAMP levels and thus, Vfr-dependent transcription of acute virulence factors (Almblad *et al.*, 2015). The mechanism for this is unclear, but is not due to activation of cAMP PDEs, or inhibition of cAMP-producing adenylate cyclases (Almblad *et al.*, 2015).

There is evidence that this inverse relationship is in response to external stimuli. For example, exogenous bicarbonate triggered an increase in the concentration of cAMP, and decrease in c-di-GMP in a concentration-dependant manner (Ruksakiet *et al.*, 2021). This suggests an internal feedback axis which reduces the level of one in response to an increase in the other (Almblad *et al.*, 2019). However, this relationship is complex, and more work is needed to fully understand it.

1.1.3 Protein transport and secretion

Like other Gram-negative bacteria, *P. aeruginosa* has a complex double-envelope consisting of an inner and outer membrane (IM and OM), separated by a thin peptidoglycan layer and aqueous periplasm (Malinverni & Silhavy, 2009). The OM is asymmetrical, with a phospholipid inner leaflet, and an outer leaflet packed with lipopolysaccharide (LPS) which provides a formidable barrier against antibiotics and immune factors (Okuda *et al.*, 2016). Proteins also contribute to the heterogeneity of the OM and may span the OM or can be peripherally localized to one side. Proteins localized to the OM have various roles, including as porins, efflux pumps, adhesins, toxins, enzymes, and mediators of motility (Henderson *et al.*, 2004).

Secretion to the exterior of the bacterium therefore involves translocation across two membranes, which is achieved by several types of secretion system, depending on function. Of these, *P. aeruginosa* is known to possess 5 out of the 9 currently characterised: Type I, II, III, V and VI, as summarised in **Figure 1.3**, and extensively reviewed elsewhere (Bleves *et al.*, 2010; Green & Mecsas, 2016). The type-II and type-V secretion systems (T2SS and T5SS) both require a two-step process, where exoproteins first cross the IM. IM translocation, discussed in 1.1.3.1, is typically either Sec- (general secretion) or TAT-mediated (twin arginine transporter) or otherwise via more specialized mechanisms reviewed elsewhere (Thanassi & Hultgren, 2000).

The T2SS is used to secrete a number of important virulence factors discussed in 1.1.2, such as proteases like elastase (LasB) and PaAP, phosphatases such as LapA and phospholipases. These proteins, unlike those

secreted by the T5SS, are secreted already folded (Nivaskumar & Francetic, 2014).

Proteins secreted by T5SS, also referred to as auto-transporters, can be further divided into subtypes (a-f), which differ in their secretion mechanisms and protein architecture. The 'classical' type Va (T5aSS) are discussed in detail in 1.1.3.2 (Meuskens *et al.*, 2019). Importantly, members of the T5SS require assistance from the β-barrel assembly machinery (BAM) complex, of which BamA is member of the Omp85 family of outer membrane proteins (Omps), for insertion into the OM (Knowles *et al.*, 2009; Leo *et al.*, 2012). The other secretion systems export exoproteins in a single step, to either the extracellular milieu, as with type I (T1SS), or into a target host cell as with types III and VI (T3SS and T6SS) (Bleves *et al.*, 2010).

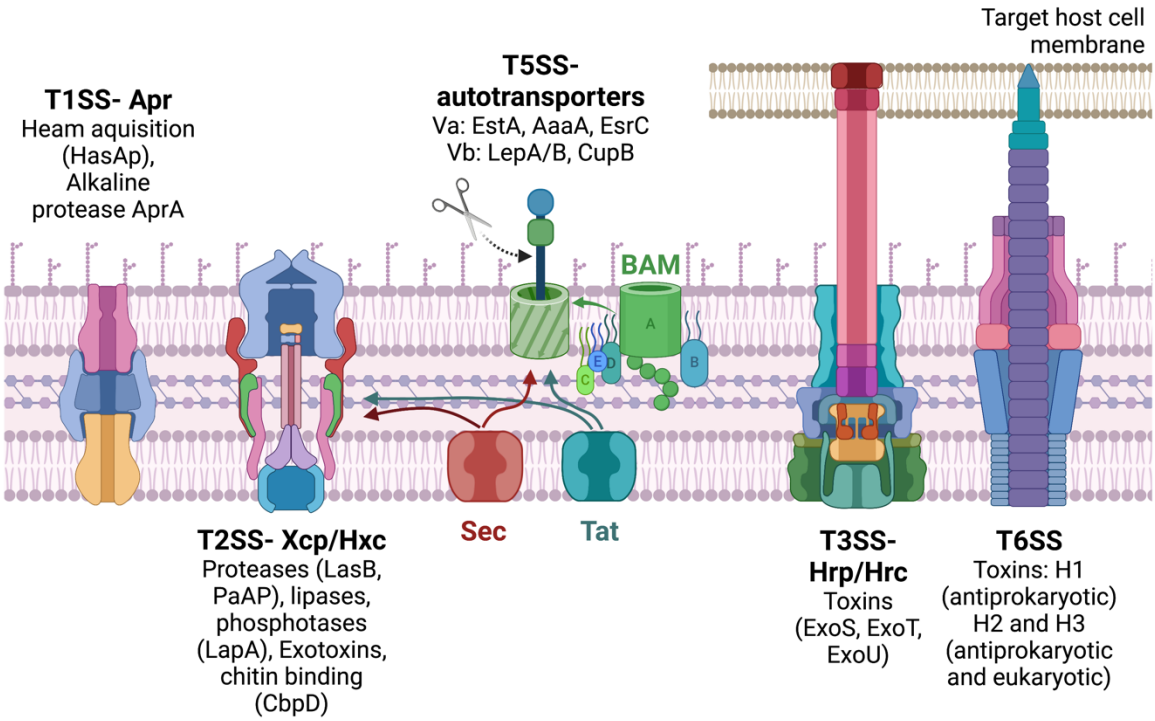


Figure 1.3: Summary of the major* P. aeruginosa *secretion systems. *Scissors represent the optional cleavage of T5SS passenger domains for release. Abbreviations: T1-6SS- Type 1-6 secretion system, Sec- general secretion machinery, TAT- twin arginine transporter, BAM- β-barrel assembly complex. Based on* (Bleves *et al.*, 2010; Pena *et al.*, 2019)*, created using Biorender.*

1.1.3.1 Secretion across the IM via Sec and TAT

The Sec translocase is comprised of 6 polypeptides which make up the SecYEGDF-YajC complex. Within this, SecB is a chaperone which delivers preproteins to SecA, an ATPase that drives protein movement into and across

the membrane. It does this via cycles of ATP binding and hydrolysis, which are coupled to SecA membrane insertion and deinsertion, respectively. This promotes the stepwise translocation of the hydrophilic regions of the preprotein across the membrane. SecY, SecE and SecG form a hetero-trimeric complex (SecYEG) which forms the channel through which the polypeptides move. SecD and SecF are also integral membrane proteins and are required for efficient protein export *in vivo*. They have periplasmic domains and along with SecG and YajC, improve the efficiency of translocation by regulating the insertion/deinsertion cycles (Dalbey & Kuhn, 2000; Mori & Ito, 2001).

The twin arginine transporter (TAT) system also plays a significant role in transporting, for example, co-factor bound globular proteins, or already tightly folded proteins. TAT substrates are recognised by their signal peptides, which contain a twin-arginine motif before a hydrophobic region (Robinson & Bolhuis, 2001). The machinery for TAT differs substantially from Sec, particularly in that it uses proton-motive force, rather than ATP, to power translocation.

1.1.3.2 Introduction to autotransporters (T5aSS)

The term 'autotransporter' was first used by Meyer and colleagues (1995) when reporting the *Neisseria meningitidis* IgA (Immunoglobulin A) protease, to describe a polypeptide which had all the necessary information required within itself for translocation to the OM or 'auto display'. Classical (type Va) autotransporters have a tripartite structure, containing an N-terminal signal peptide domain, a passenger domain carrying the functional protein, and a β-barrel, for OM insertion. Autotransporters are firstly localised to the periplasm via Sec, which recognises the signal peptide. Translocation occurs across the IM, although in a manner independent of a cytosolic energy source, and the signal peptide is cleaved off in the process (Leo *et al.*, 2012; Natale *et al.*, 2008; Thanassi *et al.*, 2005). Once in the periplasm, various autotransporters have been known to associate with chaperone proteins (such as the *E. coli* autotransporter EspP with the chaperones SurA and DegP), mostly to prevent premature folding or misfolding (Leo *et al.*, 2012).

As mentioned above, though it was initially thought that autotransporters could insert themselves into the OM unassisted, it is now clear that the Omp85-BAM complex is crucial for autotransporter biogenesis, and indeed required for the insertion of almost all OM-β-barrel proteins (Knowles *et al.*, 2009; Leo *et al.*, 2012). Depletion of both SurA and to a lesser extent, BamB have been associated with significantly lower levels of autotransporters in the OM (Klein *et al.*, 2019). Many autotransporters have been shown to localise to the bacterial cell poles, such as IcsA and SepA of *Shigella flexneri* and BrkA of *Bordetella pertussis*, and this is dependent on LPS (Jain *et al.*, 2006). However, the autotransporter Antigen 43 of *E. coli* appears to cover the whole cell with no preference for the poles (Danese *et al.*, 2000; Kjaergaard *et al.*, 2000).

There are two major mechanistic theories regarding autotransporter translocation of the OM. The classical (hairpin) model assumes that the passenger domain C-terminus forms a hairpin which inserts into a β-barrel-formed pore, through which the passenger domain moves, folding sequentially on the outside of the cell (Junker *et al.*, 2006; Leo *et al.*, 2012; Peterson *et al.*, 2010). The action of sequential folding provides the energy to pull the passenger domain through the narrow barrel and prevent backsliding, as described by the Brownian ratchet model. This is supported by the existing partial crystal structures of β-barrel and passenger domains, which reveal 12-stranded β-barrels which are too narrow to accommodate a folded protein, and elongated passenger domains composed largely of β-helix. These features are uncommon, yet would be key in a sequential-folding mechanism as described by the hairpin model (van den Berg, 2010). In the first full crystal structure of the *P. aeruginosa* autotransporter EstA, the long α-linker which connects the passenger domain and the periplasmic side of the β-barrel is clearly located within the β-barrel lumen- strongly supporting this model (van den Berg, 2010).

However, it was observed that an already stably-folded passenger domain could still be secreted via an autotransporter, challenging the idea that folding occurs during secretion (Skillman *et al.*, 2005). This lends weight to an

alternative ‘Omp85’ model in which passenger translocation, as well as β-barrel insertion, are both mediated by the BAM complex. Many other aspects of autotransporter trafficking and surface expression/secretion are still unclear, and subject to debate (Leyton *et al.*, 2012). In reality, due to the structural and functional diversity of Type Va autotransporters, it is likely that both mechanisms are utilised, depending on the specific autotransporter structure.

It is becoming increasingly clear that autotransporters are important virulence factors in many species, with wide-ranging functions. For example, the previously mentioned Ag43 is involved in agglutination, while *N. meningitidis* NalP processes other autotransporters and catalyses the release of lactoferrin binding protein (Danese *et al.*, 2000; Kjaergaard *et al.*, 2000; Roussel-Jazédé *et al.*, 2010; van der Woude & Henderson, 2008).

The *P. aeruginosa* genome encodes a handful of type Va autotransporters, the best characterised of which is the previously mentioned EstA (PA5112) (Ramos, 2004 pg. 763). EstA has been shown to alter the level of extracellular rhamnolipids, a cytosolic surfactant which can kill host immune cells, in a QS-dependant manner (Jensen *et al.*, 2007; Wilhelm *et al.*, 2007). EstA also modulates twitching, swimming and swarming motility and can influence the formation and architecture of biofilms (Wilhelm *et al.*, 2007). The second classical autotransporter is a serine protease named EprS (PA3535) which is thought to activate host inflammatory responses as well as exerting pleiotropic effects on *P. aeruginosa* pathogenesis (Kida *et al.*, 2013, 2015). The third is AaaA (PA0328)- the topic of this study- a surface-tethered aminopeptidase with potential roles in arginine utilisation, which is described in further detail in 1.3 (Li *et al.*, 2011).

1.2 Arginine utilisation by *P. aeruginosa*

P. aeruginosa exhibits extensive metabolic flexibility and can utilise arginine as a sole nitrogen and carbon source via multiple catabolic pathways. Moreover, as summarised in **Figure 1.4**, arginine may have roles in cell signalling, both for bacteria and host immune cells, as well as being a substrate for the host immune system. This study examines the role of AaaA and seeks

to understand its function in the context of arginine acquisition and utilisation. This section discusses the components involved in arginine sensing, uptake, and metabolism by *P. aeruginosa*, as well the mechanisms by which arginine can serve as a signalling molecule and host immune component.

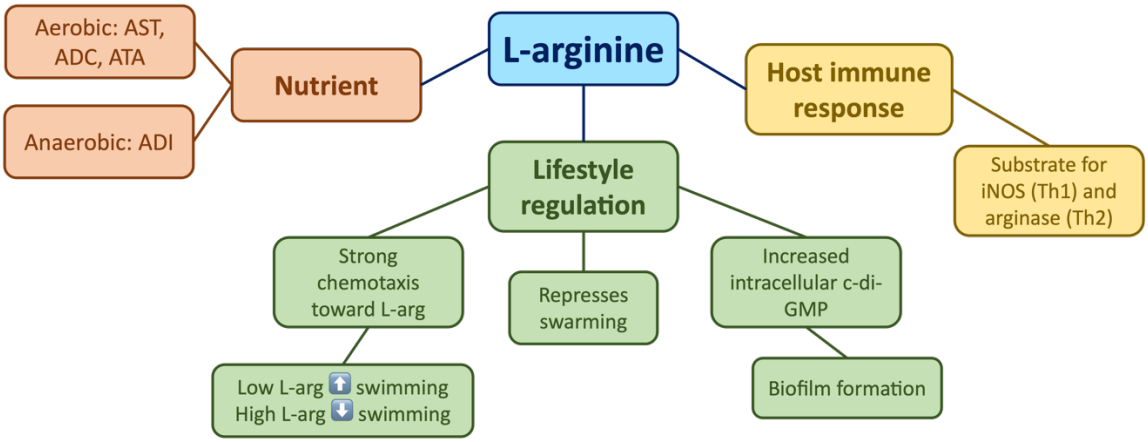


Figure 1.4: Schematic of different roles of arginine in* P. aeruginosa *infection.

1.2.1 The ArgR regulon

ArgR is a cytoplasmic DNA binding protein, part of the Arac/XylS family, is the master regulator for cellular responses to arginine (Gallegos *et al.*, 1993). It is encoded in the *aotJQMOP-argR* operon for arginine and ornithine uptake and regulation, upstream of the *aruCFGDB* operon, which codes for arginine catabolism genes (see: 1.2.4.1). ArgR is auto-induced when exogenous arginine is available (Nishijyo *et al.*, 1998). It represses arginine biosynthesis genes *argF*, *argG*, and *carAB*, as well as *gdhA* and *gltBD* genes involved in glutamate biosynthesis and induces *aotJQMOP-argR, arcDABC* and *aruCFGDBE* operons which encode arginine response regulation, import, and metabolism, as described later, as well as *gdhB* (Hashim *et al.*, 2004; Lu *et al.*, 2004; Park *et al.*, 1997). Similarly functioning regulators which respond to arginine are present in other bacteria, including *Streptococcus pneumoniae,* where it functions to co-ordinate arginine uptake (Kloosterman & Kuipers, 2011).

1.2.2 Arginine sensing

RmcA is redox-sensing c-di-GMP PDE, localised to the cytoplasmic membrane, which decreases c-di-GMP levels in response to L-arginine,

among other signals, and thus downregulates biofilm formation (Okegbe *et al.*, 2017; Paiardini *et al.*, 2018). RmcA was shown to be upregulated two-fold in a 48 hr biofilm compared with a 4 hr planktonic culture (Dötsch *et al.*, 2012). RmcA and a related protein, MorA, are therefore thought to be involved in nutrient sensing and co-ordinating cellular responses to adapt to low nutrient conditions in mature biofilms, by reducing c-di-GMP levels as well as possibly by physically interacting with PelD (Katharios-Lanwermeyer *et al.*, 2021).

1.2.3 Arginine import

1.2.3.1 Arginine uptake across the outer-membrane

1.2.3.1.1 Outer membrane porins

It is thought that uptake of arginine and other hydrophilic molecules is facilitated by channels in the OM. Most Gram-negative bacteria have non-specific porins, for example the OM Porins (Omps) of *E. coli*, which facilitate general diffusion of a wide range of molecules across the OM (Nikaido, 2003). *P. aeruginosa* instead expresses more specific channel proteins, still contentiously referred to as porins, despite their relative selectivity, or 'slow porins'. For simplicity, these channels will also be referred to as porins here. Their low level of expression, and substrate specificity likely contributes to the comparably low outer-membrane permeability of *Pseudomonas spp.,* including to antibiotics (Yoshimura & Nikaido, 1982). The major, and most extensively studied, *P. aeruginosa* OM porin is OprF, which has structural, adhesion and signalling roles, while its role in nutrient import is debated (Chevalier *et al.*, 2017). Diffusion of glucose is attributed to OprB and OprB2 and phosphate/pyrophosphate uptake is by OprP and OprO. OprG may be involved in iron uptake, and OprH is thought to have structural roles. A further nineteen porins belong to the so-called OprD (Occ) family, with the two subfamilies, OccD and OccK (Chevalier *et al.*, 2017). The best studied regarding arginine transport is OprD, a member of the OccD subfamily.

1.2.3.1.2 OprD and arginine import

OprD possesses a binding site specific for the basic amino acids arginine and lysine and dipeptides containing these residues, as well as

carbapenem antibiotics (Eren *et al.*, 2012; Huang & Hancock, 1993; Tamber *et al.*, 2006; Trias & Nikaido, 1990). OprD is constitutively expressed at a low level, but is induced by arginine in an ArgR-dependant manner, as is its ortholog OpdP (Ochs *et al.*, 1999; Tamber & Hancock, 2006). Uptake of radioactive arginine was slower in the *oprD*/*opdP* double mutant, but not completely ablated, indicating further complexity and redundancy in arginine import across the OM (Tamber & Hancock, 2006).

Molecular dynamics simulations of arginine transport through OprD, by Samsudin et al. (2019) show that arginine first associates with LPS in the OM via hydrogen bonds with LPS sugar headgroups and salt bridges with the phosphate groups until it encounters OprD. Arginine is then transferred to the putative binding site on the external loops of OprD, before proceeding through the pore via interactions between its carboxyl group and the OprD 'basic ladder'- an arrangement of basic acidic amino acids along the length of the lumen- surrounded by water molecules, until it is released into the periplasm. This study crucially also showed large energy barriers would generally not permit arginine to spontaneously transverse the hydrocarbon core of the OM.

However, doubt was cast on the need for substrate specific porins for arginine uptake by Ude et al. (2021). They created a PA14 Δ*40* mutant, in which all 40 known or predicted OM porins in *P. aeruginosa* were deleted and found that arginine uptake was almost completely unaffected. The Δ*40* mutant did show restricted growth on glutamate or succinate, suggesting porins are essential for import of molecules with two or more carboxylate groups. Taken together, this suggests that either arginine is indeed able to cross the OM unaided, or more likely, that there is an alternative yet undiscovered pathway for arginine uptake across the OM.

1.2.3.2 Arginine uptake across the inner membrane

Arginine is transported across the IM via two transporters, encoded by *arcD* and *aotJ*, on the *arcDABC* and *aotJQMOP-argR* operons, respectively. ArcD encodes a 53 kDa arginine-ornithine exchanger, which is thought to primarily deliver arginine to the arginine deiminase (ADI) pathway, which is encoded by *arcABC* and is discussed in 1.2.4.4, under anaerobic conditions

(Bourdineaud *et al.*, 1993; Lüthi *et al.*, 1990; Verhoogt *et al.*, 1992). The *aot* operon similarly encodes an arginine ornithine exchanger with similarities to arginine importers in other enteric bacteria (Nishijyo *et al.*, 1998). Being part of the *aotJQMOP-argR* operon, this transport mechanism is induced by arginine, via ArgR, as well as being repressed by succinate (Nishijyo *et al.*, 1998).

1.2.4 Arginine catabolism

It has been known for over 60 years that *P. aeruginosa* can utilise arginine for energy, and work to elucidate the pathways and roles for arginine utilisation have continued since (Thornley, 1960). Arginine catabolism occurs via four distinct pathways, which were elegantly summarised by Scribani Rossi et al. (2022), shown in **Figure 1.5**. These are: the arginine succinyltransferase (AST) pathway, the arginine dehydrogenase, or transaminase (ADH, or ATA) pathway, the arginine decarboxylase (ADC) pathway and the previously mentioned ADI pathway. The AST, ADH/ATA and ADC pathways operate under aerobic conditions, and ultimately produce succinate which can feed into the tricarboxylic acid (TCA) cycle, whereas the ADI pathway operates under anaerobic conditions and produces ATP directly, though less efficiently. *P. aeruginosa* also encodes a D-arginine racemisation pathway to convert D-arginine to L-arginine, primarily to feed into the AST and ADH pathways (Haas *et al.*, 1984; Li & Lu, 2009). Detailed figures depicting the biochemical stages of each pathway and their involved genes can be found in Spencer (2018).

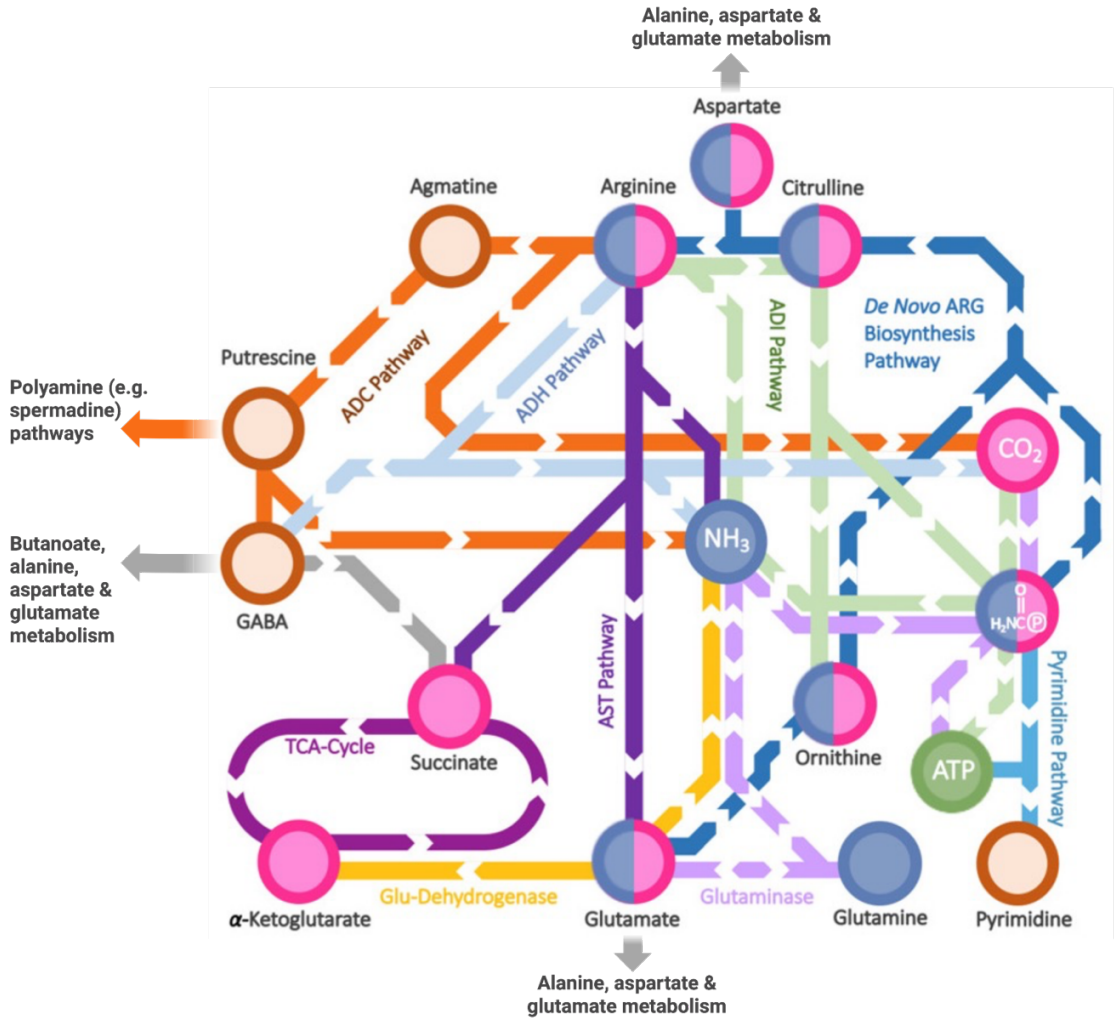


***Figure 1.5: Summary of the four arginine catabolic pathways in* P. aeruginosa.** *Nodes represent the most relevant metabolite(s) and those coloured with magenta, and blue represent carbon and nitrogen source metabolites, respectively. Different coloured lines and their corresponding text represent different metabolic categories. Green nodes are representative of energy production and rust orange nodes represent for other N metabolites. Abbreviations: 2NHCO℗: carbamoyl phosphate. Adapted from* Scribani Rossi et al. (2022).

1.2.4.1 The AST pathway

The AST pathway is the preferred pathway for arginine utilisation (*aru*) under aerobic conditions (Itoh, 1997). In *P. aeruginosa*, it consists of five arginine catabolic enzymes, encoded by the *aruCFGDB* operon, and separately transcribed *aruE*, which, initiated by succinyl-coenzyme A-dependent succinylation, convert L-arginine into succinate and L-glutamate. Each molecule of L-arginine is converted into one molecule each of succinate L-glutamate, and one molecule of co-enzyme A, three molecules of ammonia and one of carbon dioxide are concurrently released, as well as the reduction

of one NAD^+ (nicotinamide adenine dinucleotide) to NADH. L-glutamate is then converted by the NAD+ dependent glutamate dehydrogenase (NAD-GDH, encoded by *gdhB*) into 2-ketoglutarate, also by reduction of a further NAD+ to NADH (Lu & Abdelal, 2001). This 2-ketoglutarate, along with succinate can then be fed into the TCA cycle to produce energy.

Both the *aruCFGDBE* operon and *gdhB* are upregulated by ArgR, in response to arginine (Lu *et al.*, 2004). NAD-GDH enzymatic activity is also subject to allosteric control by arginine and citrate, which act as positive and negative effectors, respectively, by influencing NAD-GDH affinity for glutamate (Lu & Abdelal, 2001). Interestingly, all genes in the *aruCFGDB* operon were essential for normal *P. aeruginosa* virulence in the a *C. elegans* infection model (Feinbaum *et al.*, 2012).

1.2.4.2 The ADH/ATA pathway

Induction of the ADH/ATA pathway is not by ArgR, but by the TCS AruRS, in response to arginine (Yang & Lu, 2007). Thus, this pathway occurs in the absence of the AST pathway, producing succinate via the conversion of pyruvate to L-alanine, and 2-ketoglutamtate to L-glutamate and the subsequence release of urea, carbon dioxide and reduction of two NAD+ to NADH. It was originally called the ADH (arginine dehydrogenase) pathways as it was thought that the L-arginine α-amino group was removed by oxidative deamination by a dehydrogenase or oxidase. However, it has since been argued that removal is via transamination, and transfer to pyruvate resulting in conversion to L-alanine, thus it was dubbed the arginine transaminase (ATA) pathway instead. The L-alanine is then recycled to pyruvate by the alanine dehydrogenase, DadA (Yang & Lu, 2007).

The AST pathway is more efficient than the ATA pathway, both in the binding affinities of their operational enzymes for L-arginine, and in their gene organisation. The genes for the AST pathway are encoded by one operon and one additional gene (*aru* operon and *gdhB*), whereas the genes for the ATA pathway are encoded on various regulatory modules, and are regulated in response to a number of intermediate compounds in polyamine metabolism (Yang & Lu, 2007).

1.2.4.3 The ADC pathway

The primary function of the ADC pathway is thought to be for putrescine biosynthesis, by converting L-arginine into agmatine, and then putrescine. However, these can then be further catabolised to succinate as a carbon and nitrogen source (Mercenier *et al.*, 1980a; Nakada & Itoh, 2003). SpeA is an arginine decarboxylase encoded by *speA*, which converts arginine to agmatine, which is converted into putrescene by agmatine deiminase and N-carbamoylputrescine amidohydrolase, encoded by the *aguAB* operon. Induction of *aguAB* is repressed by the AguR regulator, which itself is antagonised by binding of agmatine and N-carbamoylputrescine which act as anti-repressors to induce *aguAB* (Nakada *et al.*, 2001).

Putrescine can then also be converted to spermidine and other polyamines which have wide-ranging roles in cell proliferation, gene expression and cell-cell communication in both prokaryotes and eukaryotes (Lenis *et al.*, 2017). The initial product of the ADC pathway, agmatine, has also been shown to upregulate inflammatory immune responses in mice (Paulson *et al.*, 2014).

1.2.4.4 Anaerobic arginine catabolism: the ADI pathway

P. aeruginosa can survive and even slowly grow on arginine (30-40 mM) as the sole carbon source in an anoxic environment (Glasser *et al.*, 2014; Vander Wauven *et al.*, 1984). Under conditions of low oxygen and nitrate, *P. aeruginosa* instead metabolises arginine via the ADI pathway, by coupling the conversion of arginine to ornithine, carbon dioxide, and ammonia, to the production of ATP from adenosine diphosphate (ADP) (Verhoogt *et al.*, 1992). The induction of the deiminase pathway is thought to mobilize intra- and extracellular reserves of arginine, which is used as a source of ATP in the absence of respiration (Mercenier *et al.*, 1980b). This ability to derive ATP from arginine for fermentative growth is unique to *P. aeruginosa*, when compared to other Pseudomonas species, such as *P. fluorescens*, *P. putida* and *P. mendocina* (Vander Wauven *et al.*, 1984).

The ADI pathway is encoded by the *arcDABC* operon, which is induced by the anaerobic regulator ANR, in response to low-oxygen conditions, along

with augmentation by ArgR (Galimand *et al.*, 1991; Lu *et al.*, 2004). L-arginine is first imported by the arginine-ornithine antiporter (ArcD, discussed in 1.2.3.2), then converted into citrulline and ornithine by arginine deiminase (ArcA). Citrulline is then converted to carbamoyl phosphate by the catabolic ornithine carbamoyltransferase (cOTC) with the addition of phosphate, while the by-product ornithine can be exported via the arginine-ornithine antiporter or converted to L-glutamate by the arginosuccinate synthase pathway. Carbamate kinase then catalyses the generation of 1 molecule of ATP per molecule of carbamoyl phosphate (Lüthi *et al.*, 1990; Mercenier *et al.*, 1980b).

Though arginine is not essential for ADI induction, its presence increased the maximum enzyme levels twofold, whereas glucose exerts catabolite repression of the ADI pathway (Mercenier *et al.*, 1980b). The NarX-NarL TCS, which controls denitrification, was shown to repress the *arc* operon in response to nitrate, suggesting that in anaerobic conditions, where nitrate is available, denitrification, which is more efficient, is preferred over arginine fermentation (Benkert *et al.*, 2008).

1.2.5 Role of arginine in *P. aeruginosa* signalling and biofilms

Besides being a nutrient source in anaerobic or microaerobic conditions, arginine may also be an environmental cue for sessility. It represses swarming and increases intracellular concentrations of c-di-GMP, the signal for biofilm formation and sessile lifestyle described in 1.1.2.4, and this is dependent on the DGCs SadC and RoeA (Barrientos-Moreno *et al.*, 2020; Bernier *et al.*, 2011). Arginine metabolism genes have also been shown to be important for the formation of robust biofilms, and addition of exogenous arginine promotes biofilm formation (Liu *et al.*, 2022; Müsken *et al.*, 2010). There is also a potential additional role for the AQ QS systems in arginine-induced swarming repression, as HHQ was responsible for repression of swarming on arginine-swarming media, and this was independent of c-di-GMP (Ha *et al.*, 2011).

However, arginine sensing by RmcA may lead to a reduction in c-di-GMP, suggesting that this occurs on a concentration-dependant axis. Everett and colleagues (2017) showed that swimming motility is slightly increased at lower arginine concentrations (100 mM), but significantly reduced at higher L-arginine concentrations (250-750 mM). *P. aeruginosa* also showed strong chemotaxis toward L-arginine (double that of L-serine). In the context of burn wound environment, this suggested that reduced L-arginine levels in burns could trigger chemotaxis away from the burn site, along an L-arginine concentration gradient, possibly encouraging *P. aeruginosa* dissemination and sepsis (Everett *et al.*, 2017). Moreover, L-ornithine, the catabolic product of L-arginine via the ADI pathway, can also act as a chemoattractant, suggesting that this could be part of a nutrient scavenging function (Dhodary *et al.*, 2022).

Everett and colleagues (2017) also found that in a murine burn-infection model, addition of ~600 mM L-arginine (1 g/kg) alongside the infection inoculation led to a dramatic reduction in *P. aeruginosa* cells at the distal wound site at 18-hr post-burn, and even more so 36-hr post burn, compared with a phosphate-buffered saline (PBS) control. Addition of L-arginine to the wound also reduced systemic spread to the spleen 100-fold, and provided striking protection from clinical sepsis (Everett *et al.*, 2017). Use of a flagella-deficient *flgK* mutant in this model confirmed that the reduced spread of *P. aeruginosa* upon L-arginine addition is due to decreased swimming motility. This suggests that addition of a high concentration of L-arginine discourages dissemination, potentially driving a preference for sessility. Taken together, these results indicate that arginine is a key signalling agent in determining *P. aeruginosa* biofilm formation and motility/sessility switching, with important implications for wound infections.

1.2.6 Arginine as a substrate for the host immune response

1.2.6.1 iNOS and Arginase

Nitric oxide (NO) synthases (NOS) are homodimeric proteins that use the essential co-factors NADPH, reduced flavins, heme-bound iron and 6(R) 5,6,7,8-tetrahydrobiopterin to catalyse the oxidation of L-arginine to release

citrulline and NO, an antimicrobial, especially effective against intracellular pathogens (Das *et al.*, 2010; Frank *et al.*, 2002; Wu & Morris, 1998). Inducible NOS (iNOS) is induced in host cells by LPS, as well as a number of type-1 cytokines. The availability of arginine, especially extracellular arginine, is the rate-limiting factor in NO synthesis, and thus it competes with arginase for arginine. Therefore, as well as inducing iNOS, type 1 cytokines also inhibit arginase (Munder *et al.*, 1999).

Arginase is a cytosolic, trimeric, binuclear manganese metalloenzyme that catalyses the hydrolysis of L-arginine to urea and ornithine (Das *et al.*, 2010). Arginase production is induced by cAMP, LPS, and type 2 cytokines. Its role is to reduce arginine levels to restrict bacterial access to arginine as a nutrient and reduce inflammation. These type 2 cytokines also inhibit NOS, and are associated with increase humoral immunity, tissue repair, and allergic and anti-parasitic responses (Munder *et al.*, 1999). Polyamines produced in the arginase pathway can themselves downregulate pro-inflammatory cytokine release. In the wound context, it is thus thought that iNOS expression is important in early wound healing, by creating a more inflammatory, cytotoxic and antimicrobial environment. As healing progresses, arginase activity then becomes more important, to create a favourable environment for fibroblast replication and collagen production. In mice with lungs infected with *P. aeruginosa,* infection resulted in increased iNOS and arginase activity in the lungs, but only arginase activity was increased in other organs including the trachea and liver. This suggests an imbalance of iNOS and arginase, induced by *P. aeruginosa* infection across organ systems (Grasemann *et al.*, 2014).

1.2.7 Potential for arginine in clinical intervention

As mentioned in 1.2.5, and noted by Everett et al. (2017), it is possible that addition of exogenous arginine may have therapeutic benefits in infected wounds. A systematic review found some evidence that arginine supplementation could improve wound healing, although there was not enough studies with high quality clinical data to draw clear conclusions (Naziat, 2020).

Another potential therapeutic approach is in directly modulating the host arginine utilisation, to push the immune response in a desired direction. Due to its competition with iNOS for L-arginine, arginase inhibiting drugs have been developed to aid in treating *P. aeruginosa* infection in CF. One such drug, CB-280, significantly improved lung function and reduced *P. aeruginosa* CFUs in a CF mouse model, and is now in Phase II clinical trials (Boas *et al.*, 2021).

1.3 Introduction to AaaA

As mentioned in 1.1.3.2, AaaA is one of three classical autotransporters encoded in the *P. aeruginosa* genome. It is structurally related to the M28 family of peptidases, with 100% conservation in predicted active site catalytic and ligand-binding residues between AaaA and other M28 members (Luckett *et al.*, 2012). It is thought to tetrahedrally coordinate two zinc ligands, though this requires further supporting evidence. Functional studies have determined that AaaA is an aminopeptidase, which cleaves N-terminal arginine from peptides, and that *aaaA* is essential for survival and growth on minimal media with the Arginyl-glycyl-aspartic acid (RGD) tripeptide.

The PAO1 Δ*aaaA* knock-out mutant, unlike its wild-type counterpart, is also unable to release L-arginine from L-arginine-p-nitroanilide, which produces a colour change upon cleavage from colourless to yellow that can be measured by reading absorbance at 370-410 nm (**Figure 1.6**). The L-arginine-p-nitroanilide cleavage assay is now a key phenotypic assay for assessing AaaA activity, and is used throughout this study, as described in 2.7.

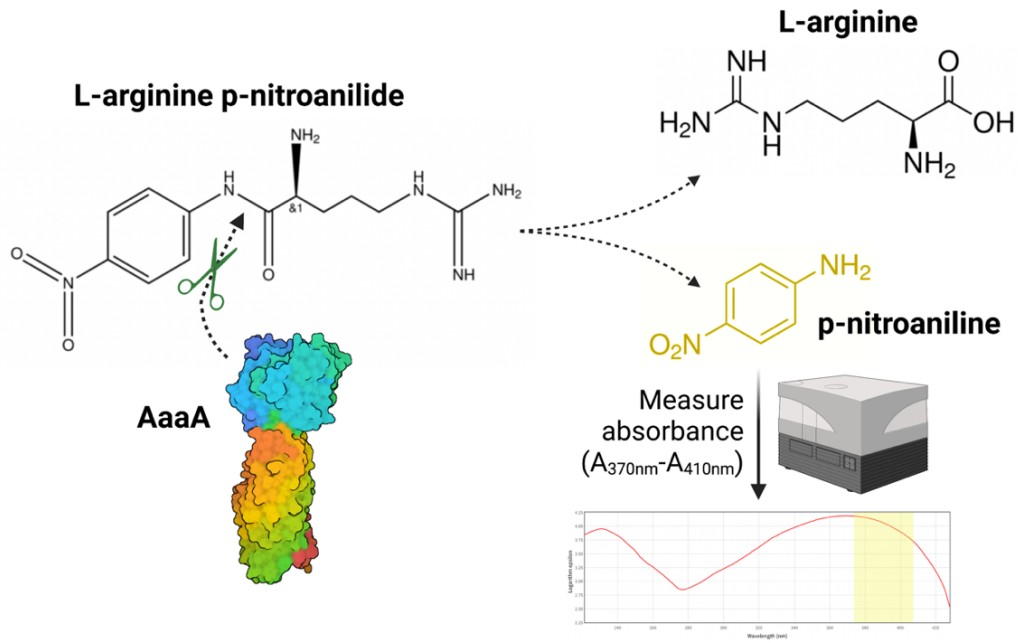


Figure 1.6: Illustration of the L-arginine-p-nitroanilide cleavage assay. *The peptide bond in L-arginine p-nitroanilide is cleaved by AaaA to release L-arginine and p-nitroaniline, causing a colour change from colourless to yellow that can be detected by measuring absorbance at 370nm-410nm. Spectra for p-nitroaniline taken from* NIST (2021), structure for AaaA taken from AlphaFold Protein Structure Database (2022), chemical structures taken from Sigma-Aldrich (2022). Created using BioRender.com.

A PAO1 Δ*aaaA* knock-out mutant also showed reduced respiration when dipeptides with an N-terminal arginine were the sole nitrogen source, except for Arg-Arg/Arg-Lys. This effect was most pronounced with Arg-Val and Arg-Ile, suggesting a level of substrate specificity. However, Arg-His, Arg-Thr, Arg-Asn, Arg-Cys, Arg-Pro and Arg-Gly were not tested, so their potential specificity has not yet been assessed (Luckett *et al.*, 2012).

Bacterial aminopeptidases in general often have important roles in virulence. For example PhpA is a homolog to the *E. coli* leucine-specific aminopeptidase PepA, and is thought to be involved in AlgD repression (Woolwine & Wozniak, 1999). PhpA (PA3831) was named as so to avoid confusion with the existing *P. aeruginosa* PepA (Hauser *et al.*, 1998). PaAP (PA2939), previously known as PepB, is a leucine-specific aminopeptidase likely to be important during proteolytic growth, as discussed in 5.8.3.1 (Cahan *et al.*, 2001). Therefore, given the roles for both autotransporters and aminopeptidases as important virulence factors, it remains likely that AaaA is similarly important.

Indeed, AaaA has been shown to be important for *P. aeruginosa* in maintaining chronic, but not acute, wound infections in mice (Luckett *et al.*, 2012). Because of its ability to release local extracellular arginine from peptides, AaaA may be advantageous for survival in this, and other nutrient-poor, and microaerobic environments, such as within wound biofilms. In this context, AaaA could release arginine to be catabolised for energy, as discussed in 1.2.4, or to serve as an important signalling molecule, as discussed in section 1.2.5. Increased local arginine concentrations may also have immunomodulatory roles, discussed in section 1.2.6, which could hinder wound healing, though these mechanisms have not been fully elucidated.

Being surface-exposed, immunogenic and highly-conserved, *aaaA* has potential as an anti-*P. aeruginosa* drug target, as well as being in the top five of *P. aeruginosa* conjugate vaccine candidates (Bianconi *et al.*, 2018). More generally, bacterial autotransporters are emerging as important virulence factors across a wide range of species, therefore, elucidating the structure and function of AaaA would add novel contributions to this field, and potentially uncover previously unknown metabolic or virulence mechanisms in *P. aeruginosa*.

Other species have similar arginine utilisation machinery, such as *Por. gingavalis*, which uses an essential Arginine-specific gingipain (Rgp) to release arginine as a nutrient as the bacterium cannot metabolise sugars (Kadowaki *et al.*, 2000). Rgp also has roles in the processing and maturation of other surface proteins such as *fimA* fimbrilin, hemagglutinins, and the haemoglobin receptor protein, all of which are important in establishing and maintaining colonies (Grenier *et al.*, 2001).

1.3.1 AaaA and arginine in chronic infections

Proteomics has identified AaaA as the 17th most abundant protein present in the extracellular matrix of biofilms as well as being abundantly enriched in biofilm outer membrane vesicles (OMVs) (Toyofuku *et al.*, 2012; Couto *et al.*, 2015; Park *et al.*, 2015). These studies also found arginine deiminase (ArcA) to be highly abundant in both the biofilm matrix and OMVs,

suggesting the role of ArcA and AaaA may be linked, supporting the hypothesis that AaaA releases arginine destined for the ADI pathway. As well as AaaA, several other exoenzymes involved in macromolecule degradation were detected in the biofilm matrix, including PasP (PA0423), alkaline metalloproteinase AprA (PA1249), chitinase ChiC (PA2300), aminopeptidase PaAP, elastases LasA (PA1871) and LasB (PA3724), and Protease IV (PA4175). This lends support to the idea of an upregulation of a contingent of matrix proteins responsible for host tissue degradation, potentially for nutrient acquisition.

As mentioned above, *in vivo*, PAO1 and PAO1 Δ*aaaA* were equally able to establish an acute infection in a mouse burn wound model, but PAO1 Δ*aaaA* showed significantly reduced survival in a mouse chronic wound infection at 2- and 8-days post infection. Complementation with chromosomally stable *aaaA* rescued the chronic infection phenotype. At 2-days post infection, mice colonised by PAO1 Δ*aaaA* also had lower proinflammatory tumour necrosis factor (TNF)-α, Interleukin (IL)-1α, keratinocyte-derived cytokine and cyclooxygenase-2, suggesting AaaA is proinflammatory in the acute stage (Luckett *et al.*, 2012). During the acute stage (day 2) of infection, mice infected with PAO1 Δ*aaaA* also show reduced expression of iNOS and Arginase I and II than those infected with wild-type PAO1. However, by the chronic stage (8-days), iNOS expression rose in mice infected with PAO1 Δ*aaaA*, whilst Arg I expression fell compared with wild-type and *aaaA* complemented strains. Luckett *et al* (2012) also showed generally lower bacterial loads and increased fibroblast infiltration of the wound site in PAO1 Δ*aaaA*, suggestive of a greater Type 2 healing response. Taken together, these results suggest that in the chronic stage of infection, AaaA acts to disrupt the normal balance of arginine-regulated host immune responses, potentially by steering the host away from the 'alternative', healing-associated arginase responses (Luckett *et al.*, 2012).

Preliminary data from a murine lung infection model showed PAO1 Δ*aaaA* also appeared to be cleared more quickly than wild type and, unlike the wild type, did not produce black spots on the lungs, which is indicative of a chronic infection (Spencer, 2018). This suggested that Δ*aaaA* mutants were

forming a more acute, rather than chronic infection. However, there weren't significant differences in either colony forming unit (CFU) recovery or the levels of inflammatory markers IL-10 and TNF-α in this study, suggesting a complicated picture which requires further study. A simple explanation is that AaaA is not as relevant in a chronic lung infection context as it is in a chronic wound, which is possible given that they are very different niches, although this is yet to be determined.

1.4 PhD aims and objectives

This study aimed to build on previous work exploring the structure and function of AaaA, and apply this to a synthetic chronic wound, so as to study AaaA under conditions where it is likely to be most relevant to disease. Therefore, this study had two main aims: Firstly, to elucidate the structure of AaaA, using bioinformatic tools and protein purification for crystallography. Gaining structural insight into AaaA could then help to guide the design of potential of drugs or vaccines which target AaaA. Secondly, this study sought to examine the mechanism underlying the advantage conferred by AaaA in a chronic wound infection, by probing its regulation, localisation and function in a synthetic chronic wound model.

Chapter 2. Methods

2.1 Strains, plasmids, and primers

2.1.1 Strains used in this study

Table 2.1: List of strains used in this study.

Name	*Description*	*Source*
***E. coli* DH5α**	Cloning strain. Genotype: F' *endA1 hsdR17(rk-,mk-) supE44 thi- 1 recA1 gyrA (Nalr) relA1Δ(lacIZYA- argF) U169 deoR (ϕ 80dLac Δ(lacZ)M15).*	(Hanahan, 1983), supplied by New England Biolabs UK Ltd.(Hitchin)
***E. coli* Top10**	Cloning strain. Genotype: F- *mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 Δ lacX74 recA1 araD139 Δ(araleu)7697 galU galK rpsL (StrR) endA1 nupG*	Fisher Scientific UK Ltd. (Loughborough)
E. coli* S17-1 λ-*pir	Cloning strain, used for conjugations. Genotype: *Thi, pro hsdR-, hsdμ+, recA; RP4 2- Tc::Mu- Kn::Tn7*	(Simon *et al.*, 1983)
***E. coli* BL21(DE3)**	For protein overexpression via the T7 promoter. Genotype: *fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*	New England Biolabs UK Ltd. (Hitchin)
***E. coli* NiCO21(DE3)**	BL21(DE3) derived, engineered to minimize production of metal binding proteins which could contaminate IMAC fractions. Genotype: *can::CBD fhuA2 [lon] ompT gal (λ DE3) [dcm] arnA::CBD slyD::CBD glmS6Ala ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*	
***E. coli* C41**	*E.coli* BL21(DE3) derived, engineered to be more resistant to overexpression of toxic proteins. Genotype: F – *ompT hsdSB (rB- mB-) gal dcm (DE3)*	Merck Life Science UK Ltd (Dorset)
PAO1N	*P. aeruginosa* PAO1 Nottingham sub-strain	(Dubern *et al.*, 2022)
PAO1N Δ*aaaA*	PAO1N with an in-frame deletion of *aaaA*	(Luckett *et al.*, 2012)
PAO1N miniCTX::P_{kan}-lux	PAO1N with chromosomally integrated miniCTX::P_{kan}-lux	(Paredes-Osses, 2014)
PAO1N miniCTX::P_{aaaA}-lux	PAO1N with chromosomally integrated miniCTX::P_{aaaA}-lux	
PAO1N Δ*aaaA* miniCTX::P_{kan}-lux	PAO1N Δ*aaaA* with chromosomally integrated miniCTX::P_{kan}-lux	

Name	*Description*	*Source*
PAO1N Δ*aaaA* miniCTX::P_{kan}-lux	PAO1N Δ*aaaA* with chromosomally integrated miniCTX::P_{aaaA}-lux	(Paredes-Osses, 2014)
PAO1L	*P. aeruginosa* PAO1 Lausanne sub-strain	(Dubern *et al.*, 2022)
PAO1L Δ*aaaA*	PAO1L with an in-frame deletion of *aaaA*	This study
PAO1L miniCTX::P_{kan}-lux	PAO1L with chromosomally integrated miniCTX::P_{kan}-lux	
PAO1L miniCTX::P_{kan}-lux	PAO1L with chromosomally integrated miniCTX::P_{aaaA}-lux	
PAO1L Δ*aaaA* miniCTX::P_{kan}-lux	PAO1L Δ*aaaA* with chromosomally integrated miniCTX::P_{kan}-lux	
PAO1L Δ*aaaA* miniCTX::P_{kan}-lux	PAO1L Δ*aaaA* with chromosomally integrated miniCTX::P_{aaaA}-lux	
PAO1L Δ*aaaA* pJM220::*aaaA*	PAO1L Δ*aaaA* with chromosomally integrated pJM220::*aaaA*	
PAO1L Δ*aaaA* pJM220::*aaaA*-St4-6	PAO1L Δ*aaaA* with chromosomally integrated pJM220::*aaaA*, with the Spytag, St (version 4, 5 or 6)	
CW2-BI1	*P. aeruginosa* isolated from a chronic wound (bone isolate)	(da Silva, 2018)
CW2-BI1 Δ*aaaA*	*P. aeruginosa* CW2-BI1 with an in-frame deletion of *aaaA*	This study

2.1.2 Plasmids used in this study

Table 2.2: List of plasmids used in this study.

Name	*Description*	*Source*
pBlueScript II KS(+)	Standard cloning vector (phagemid excised from λZAPII). The f1 (+) orientation allows rescue of sense strand ssDNA. N.B. pBluescript II KS(+) and pBluescript II SK(+) differ by the orientation of the MCS. Ampicillin resistance (100 µg/mL in *E. coli*). GenBank: X52327.	(Alting-Mees & Short, 1989)
pBlueScript*::aaaA*	pBlueScript II KS (+) with *aaaA* inserted using *Bam*HI and *Eco*RI. Ampicillin resistance (100 µg/mL in *E. coli*).	(Luckett *et al.*, 2012)
pET-21a(+)	Bacterial vector for inducible expression of N-terminally T7-tagged proteins. Ampicillin resistance (100 µg/mL in *E. coli*).	(Novagen, 2021)
pET21a::*aaaA*	pET-21a(+) containing full length *aaaA* under the control of an IPTG-inducible T7 promoter. Generated by restriction cloning from pBlueScript::*aaaA* using *Nde*I and EcoRI. Ampicillin resistance (100 µg/mL in *E. coli*).	(Luckett *et al.*, 2012)
pET21a::*aaaAt3*	pET21a::*aaaA* with *aaaA* edited to exclude the auto-transporter β-barrel domain for periplasmic localisation (*aaaA* truncation, version 3)	(Spencer, 2018)

Name	***Description***	***Source***
pET21a::*aaaAt3-H8/H10*	pET21a::*aaaAt3* with the hexa-histidine tag extended to 8 or 10 histidine	This study
pET21a::*aaaA*-D580E	pET21a::*aaaA* with *aaaA* edited to create the D580E mutation.	This study
miniCTX-1	TetR, oriT, pMB1-derived origin of replication, a modified ϕCTX integrase (int) gene, multiple cloning site (MCS) flanked by T4 transcriptional termination sequences (Ω elements), and the ϕCTX attachment site. This vector can be integrated to att sites in the chromosome. Tetracycline resistance (10 µg/mL in *E. coli*, 50-125 µg/mL in *P. aeruginosa*). GenBank: AF140576.1	(Becher & Schweizer, 2000)
miniCTX*::aaaA*	miniCTX-1 with *aaaA* ORF plus promoter region (500 bp upstream) inserted by restriction digest using *Not*I and *Eco*RV. Tetracycline resistance (10 µg/mL in *E. coli*, 50-125 µg/mL in *P. aeruginosa*).	(Paredes-Osses, 2014)
miniCTX*::P_{aaaA}-lux*	miniCTX-lux backbone containing the *aaaA* promoter region (312 bp upstream, plus the first four codons in the PA0328 ORF) inserted by restriction digest using *Xho*I and EcoRI. Tetracycline resistance (10 µg/mL in *E. coli*, 50-125 µg/mL in *P. aeruginosa*). Backbone is GenBank: AF251497	(Paredes-Osses, 2014)
miniCTX*::P_{kan}-lux (pSC356)*	miniCTX-lux (as above) containing the constitutively active Kanamycin promoter. Tetracycline resistance (10 µg/mL in *E. coli*, 50-125 µg/mL in *P. aeruginosa*).	(Chen, 2014)
pEX18-Gm	*oriT*$^{+}$ *sacB*$^{+}$, suicide vector, gentamicin resistance (10 µg/mL in *E. coli*, 20 µg/mL in *P. aeruginosa*). Genbank: AF047518	(Hoang *et al.*, 1998)
pEX18-Gm-Δ*aaaA*	pEX18-Gm vector containing 618 bp upstream plus the first and last 3 codons of the *aaaA* ORF plus 617 bp downstream, inserted by restriction digest using *Xba*I, BamHI and EcoRI. Used to create an *aaaA* in-frame deletion in *P. aeruginosa* using sucrose selection. Gentamicin resistance (10 µg/mL in *E. coli*, 20 µg/mL in *P. aeruginosa*).	This study
pJM220	pUC18T-miniTn7T-gm-rhaSR-PrhaBAD. Ampicillin/gentamicin resistance (100 µg/mL ampicillin in *E. coli*, 20 µg/mL gentamicin in *P. aeruginosa*). GenBank: KX777256	(Meisner & Goldberg, 2016)
pUX-BF13	Mini-Tn7 helper vector to assist with chromosomal insertion of pJM220 (and other Tn7 insertion vectors). Ampicillin resistance (100 µg/mL in *E. coli*).	(Bao *et al.*, 1991)
pJM220*::aaaA*	pJM220 vector containing the ORF for *aaaA* from PA0328.	This study
pJM220*::AaaA-St4, 5 and 6*	Spytagged versions of *aaaA* constructed by SDM	This study

Name	*Description*	*Source*
pIBA3-SpyCatcher-mCherry	For production of SpyCatcher002 protein fused N-terminally to mCherry. Includes N-terminal His tag for purification and C-terminal StrepII tag for detection. Ampicillin resistance (100 µg/mL in *E. coli*)	Dr Jack Leo, Unpublished, based on (Chauhan *et al.*, 2019)

2.1.3 Primers used in this study

Table 2.3: List of primers used in this study. *Lower case indicates bases which do not bind to the template. Inserted restriction sites are underlined.*

Purpose	*Name*	*Sequence*	*Source*
Sequencing	T7-promoter	TAATACGACTCACTATAGG	Standard primers (Eurofins Genomics , 2020)
	T3-promoter	AATTAACCCTCACTAAAGG	
	M13F	GTTGTAAAACGACGGCCAG	
	M13R	CACACAGGAAACAGCTATGAC	
	U+Δ*aaaA*+D-Fw	TGACTATCCACGCCCAAACT	This study
	U+Δ*aaaA*+D-Rv	TGGTCACCAACGGCTTGT	
	D*aaaA*-midSeq	GAACACGGCACTTTCCTCTG	
	pet21a_seqR	TTAGCAGCCGGATCTCAG	
	pet21a_seqF	GATGTCGGCGATATAGGC	
	P_{aaaA}-Esteban-F	TATCTCGAGAGCTGAGGCCTGGCGGC	(Paredes-Osses, 2014)
	P_{aaaA}-Esteban-R	TATGAATTCGAACACGGCACTTTCCTCTG	
	ctx-P_{aaaA}-R	TCCTGAGTATTATTTCACAAATA	
SDM of AaaA	*aaaA*-SDM_D580E _F	CGGACGCCTGgaaGACATGGACC	This study
	aaaA-SDM_D580E _R	TCCGCCAGCTCGCGGACC	
Construction of pEX18Gm-ΔaaaA	Δ*aaaA*-Rx-UpFrag-F	atattctagaAGGCCATCGAGTACATCAGCCG	
	Δ*aaaA*-Rx-UpFrag-R	atatggatccTTCTGAGCGGGCCAGGGG	
	Δ*aaaA*-Rx-DnFrag-F	atatggatccGAACACGGCACTTTCCTCTGTGTC	
	Δ*aaaA*-Rx-DnFrag-R	atatgaattcATCTGAAGAAAGCGAAAGACGGCC	
RT-qPCR	rpoD-F	GGGCGAAGAAGGAAATGGTC	(Savli *et al.*, 2003)
	rpoD-R	CAGGTGGCGTAGGTGGAGAA	
	rpoS-F	CTCCCCGGGCAACTCCAAAAG	
	rpoS-R	CGATCATCCGCTTCCGACCAG	
	Dani-qPCR-*aaaA*7-F	GTCACCGGCGACAAGATGTA	(Spencer, 2018)
	Dani-qPCR-*aaaA*7-R	CAGGATCTGTTCGCGGTAGG	
	rmcA_F	TACTGCATGAGCGAAAGCC	This study
	rmcA_R	ACACCGAAGCGAACTTCTG	
	argR_F	AATATCGAGGAGCCACTGA	
	argR_R	GTACGCGGTTCAGGTATTG	
	arcD_F	CCGCTGTTCCTGTTCATC	
	arcD_R	CGCACCTGGTTCATCAC	

Purpose	*Name*	*Sequence*	*Source*
	aruF_F	CGGCGAGGAAAGCTATTT	This study
	aruF_R	GTTGCGGAAGCTGTAGAA	
	speA_F	ATCGAGCGCCTCGAATA	
	speA_R	CGATGATGTTCTCCACCAC	
	aruH_F	CCGACTTCACTCAACGTATC	
	aruH_R	CGACAGCAGGAGGATTTC	
	arcA_F	CAAACTGCGCAAAGTGATG	
	arcA_R	GTCGTCGAACAGCAACTC	
Spytag v1-6 construction (HPLC purified)	STv1_F	gatgcgtataaacgttataaaGGCCAGGAGCGCATCGAGC	
	STv1_R	caccatcacaatggtcggcacGAGGGCGTTGGTCAGTACTTCC	
	STv2_F	gatgcgtataaacgttataaaGGCTTCGGCTACCAGACC	
	STv2_R	caccatcacaatggtcggcacCAGGCGCGACTGCATCAG	
	STv3_F	gatgcgtataaacgttataaaGGCAACCGCAGTTCGCAG	
	STv3_R	caccatcacaatggtcggcacGGCCCAGGTGAAGTCCTG	
	STv4_F	gatgcgtataaacgttataaaTCCAACAGCGTGTCCAACC	
	STv4_R	caccatcacaatggtcggcacGCCGGCATGGGCGTACAT	
	STv5_F	gatgcgtataaacgttataaaCTGAACGCCGAATACCCGGC	
	STv5_R	caccatcacaatggtcggcacGCCGGGGTTGGTGAACAAC	
	STv6_F	gatgcgtataaacgttataaaGGGGAGACCCTGGAGCGC	
	STv6_R	caccatcacaatggtcggcacGGCGTATTCCCCGTATTGATAGGC	
Hifi to insert Spytagged aaaA into* miniCTX-*1* *(HPLC purified)	Hifi-MiniCTX_F	TGAGCGGATATCAAGCTTATC	
	Hifi-MiniCTX_R	GCACTTTCCTCTGTGTCG	
	Hifi-miniCTX-*aaaA*-F	accgacacagaggaaagtgcGTGTTCAAACCATTAGCTGTCG	
	Hifi-miniCTX-*aaaA*-R	ataagcttgatatccgctcaGAACTGCCAGTTCACCCC	
Hifi to insert Spytagged aaaA into* pJM220 *(HPLC purified)	Hifi- pJM220_F	CTCGAGAAGCTTGGGCCC	
	Hifi- pJM220_R	GAATTCCTGCAGAGCACTAGTTG	
	HiFi-pJM220-*aaaA*_F	ctagtgctctgcaggaattcGTGTTCAAACCATTAGCTGTCG	
	HiFi-pJM220-*aaaA*_R	ccgggcccaagcttctcgagTCAGAACTGCCAGTTCACCCC	
Extension of AaaAt3 6xHis tag	Fw-*aaaA*t3+2xHis	GCACTGACCCGCCTGCACGACCACCACCATCATCACCACCACCATTAAGAATTCGAGCTCCG	
	Fw-*aaaA*t3+4xHis	GCACTGACCCGCCTGCACGACCACCACCACCATCATCATCACCACCACCATTAAGAATTCGAGCTCCG	
	Rv-*aaaA*t3+His	CTGATGCTGGCGCTGCAACTGGTCTTCC	

2.2 Bacterial media, growth, and storage conditions

2.2.1.1 Media

In general, Lysogeny Broth (LB) was used for all routine culture of all strains, with addition of antibiotics for the maintenance of plasmids as detailed in **Table 2.2**, which were prepared and stored according to the manufacturer's instructions. Recipes for all media used in this study are listed in **Table 2.4**.

Table 2.4: List of media used in this study.

Media	*Use*	*Preparation*
Lysogeny Broth (LB)	Routine culture	Made and autoclaved in-house using 10 g/L Tryptone (Oxoid), 5 g/L Yeast Extract (Oxoid), 5 g/L NaCl (Merck), dissolved in deionised water, adjusted to pH 7 and autoclaved. 15 g/L Bacto Agar (Difco), was added to make LB agar.
Pseudomonas isolation agar (PIA)	Streaking out *P. aeruginosa* from cryo-stocks, or isolation of *P. aeruginosa* from a mixed culture	Dissolved BD Difco™ Pseudomonas Isolation Agar (peptone 20 g/L, $MgCl_2$ 1.4 g/L, K_2SO_4 10 g/L, Irgasan™ 25 mg/L, agar 13.6 g/L) in deionised water, autoclaved, then added 2% (v/v) filter sterilised glycerol.
Minimal Media P (MMP)	L-arginine-p-nitroanilide cleavage (AaaA activity) assay	10.4 mM Na_2HPO_4, 4.8 mM KH_2PO_4, 1.7 mM $MgSO_4$, 6.6 µM $FeSO_4$. Components dissolved in deionised water then filter sterilised.
Synthetic wound fluid	For use in the synthetic wound chronic model	50% foetal bovine serum, 50% peptone water. Details in 2.8.1

2.2.1.2 Bacterial growth

Routine growth of *E. coli* and *P. aeruginosa* in liquid culture was in LB at 37˚C, 200 revolutions per minute (RPM), or statically overnight at 37 ˚C for growth on agar plates. Unless otherwise stated, 'overnight' bacterial cultures were made by adding a colony from a freshly streaked plate to 5 mL LB plus any required antibiotics (described in **Table 2.2**), and incubating at 37 ˚C, 200 RPM for 16-18 hrs. Bacterial growth in liquid media was assessed by its optical density, the absorbance at 600 nm (OD_{600nm}) using a Jenway 6705 visible spectrophotometer (Fisher Scientific).

2.2.1.3 Bacterial storage

Bacterial cryostocks were made of each strain by adding 500 µL fresh overnight culture and 500 µL filter sterilised 50% glycerol to a sterile 2 mL

cryotube and vortexing for 5 s. All bacterial cryostocks were stored at -80 ˚C and handled exclusively on ice.

2.3 SDS-PAGE and Western blotting

2.3.1 SDS-PAGE

Samples were resuspended in either 5x or 2x sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) sample buffer, and 2.5 x the loading volume was added to thin walled-PCR tubes which were boiled for 10 mins at 96˚C in a thermocycler. Sample (volume specified in relevant figure legend) was loaded onto either a 12% Tris-glycine SDS-PAGE gel (made in-house using Bio-Rad short glass plates and Sigma reagents according to the manufacturer's instructions), or a pre-cast 4-20% gradient Tris-glycine SDS-PAGE gel (Invitrogen). Electrophoresis was undertaken using either the Mini-PROTEAN Tetra Vertical Electrophoresis Cell (Bio-Rad) or the XCell SureLock™ Mini-Cell (Invitorgen) depending on the gel. Gels were run at 90-100 V / 15 mins then 120-200 V / 45-60 mins. Gels were stained with Coomassie blue (made in house: 0.1% Coomassie Blue R250 in 10% acetic acid, 50% methanol) for 30 mins with gentle agitation, then de-stained (made in house: 10% acetic acid, 50% methanol) for at least 45 min or overnight with gentle agitation. Gels were rinsed in distilled water and imaged using a Gel Doc™ XR+ System (Bio-Rad).

2.3.2 Western blotting

SDS-PAGE gels for Western immunoblotting were prepared as in 2.3.1, however instead of staining, gels were transferred onto a nitrocellulose membrane either at 150 mA for 60 mins or 20 V overnight at 4 ˚C. Membrane was then blocked for 1 hr at room temperature, or at 4˚C overnight with 5% skimmed milk (Oxoid, U.K.) in PBST (PBS, 0.5% Tween 20). The membrane was then incubated with primary antibody for 1 hr at room temperature with gentle agitation, washed for 5 mins in PBST thrice, and where applicable, incubated with secondary antibody as with the primary. Antibodies used include HRP-Anti-6X His tag® antibody (Abcam, Cambridge UK), which does not require a secondary antibody, and rabbit α-AaaA Luckett et al. (2012)

following pre-adsorption with a bacterial lysate as described by Hardie et al. (2003), with secondary α-rabbit IgG-HRP (Sigma). Antibodies and concentrations used are specified in their relevant figure legends. Following extensive final washing (3 x 5 min, 2 x 15 min, 3 x 5 min) with PBST, 1 mL of 1:1 Pierce™ ECL Western Blotting Substrate, or ECL Prime (Amersham, UK) was added to the membrane for 1 min, or 5 mins with ECL Prime. Substrate was blotted off and the membrane transferred to a dark room, where it was exposed to X-ray Hyperfilm™ (Amersham, UK) for 1 min – 1 hr, and the film was developed and fixed according to the manufacturer's instructions, allowed to dry, and imaged using the Gel Doc™ XR+ System.

2.4 Protein overexpression and purification

2.4.1.1 A note about normalisation of sample preps from over-produced truncated AaaA fractions

For consistency, all samples which were taken during protein over-production studies in Chapter 4 were normalised by volume throughout to ensure the final volume of all samples, for both SDS-PAGE gels and in the AaaA activity assay, was equivalent to 1 mL of original culture (or 2 mL for resolubilised pellets, to account for possible loss in dialysis) at OD_{600nm} 1.0.

2.4.2 *aaaAt3* overexpression

Overnight cultures of the pET21a*::aaaAt3* strain were grown according to 2.2. Over-day cultures were made using a starting OD_{600nm} 0.05 of overnight culture, in 1 L Terrific Broth (24g yeast extract, 12g tryptone, 8 mL glycerol, 0.17 M KH_2PO_4, 0.72 M K_2HPO_4) split into 2x 500 mL each in 2 L baffled flasks. Over-days were incubated at 37 ˚C until cells reached an OD_{600nm} = 0.7-0.9, at which point, they were induced with 0.3 mM IPTG and grown for a further 4 hrs at 30 ˚C, 200 RPM, unless otherwise stated. Cells were spun at 2160 RCF / 25 mins, pellet washed with Lysis buffer (50 mM Tris, 10 mM NaCl, pH 8.0), collected by spinning at 2000 RCF / 10 min and frozen at -22 ˚C.

2.4.3 Cell lysis

2.4.3.1 Spheroplast lysis

All fractions from each lysis method, including the resolubilised pellets, were aliquoted into 1.5 mL Eppendorf tubes, flash frozen in dry ice and stored at -20˚C. For spheroplast lysis, pellets were resuspended in 1/25th original culture volume of ice cold spheroplast buffer A (10 mM Tris, pH 7.4, 20% sucrose, 125 µg/mL lysozyme) followed by addition of an equal volume of spheroplast buffer B (10 mM Tris, pH 7.4, 1 mM or 0.5 mM EDTA) and shaken on ice for 30 mins. Then 1/250th original culture volume of 1M $MgSO_4$ was added, and samples were spun at 3000 RCF / 3 mins. The supernatant, the periplasm fraction, was aspirated. To isolate the soluble cytoplasmic contents, pellets were resuspended in 1/125th original culture volume 10 mM Tris, pH 7.4, which was then freeze thawed 3x using dry ice, and then 37˚C, 100 µL $MgCl_2$ and 3 µL 25 U/µL Benzonase™ was added, samples were spun at 3000 RCF /15mins and supernatant (cytoplasmic contents) aspirated off.

2.4.3.2 Sonication lysis

For sonication, samples were resuspended in 5 mL sonication buffer (50 mM Tris, pH 7.4, 0.1 mg/mL lysozyme, ½ cOmplete™ EDTA-free protease inhibitor tablet) and probe-sonicated (15 mA, 15 s on/15 s off, 10 mins). Samples were spun at 6000 RCF /15 min and supernatants (lysate) aspirated off from pellets.

2.4.3.3 French press lysis

For French press lysis, pellets were thawed and resuspended in 1/40th original culture volume in 50 mM Tris, pH 7.4, 40 µL 10 mg/mL lysozyme, 40 µL 1M $MgCl_2$ (10 mM) and 3 µL (25 U/µL) Benzonase™, incubated on ice / 15 min and then lysed using a FRENCH PRESS G-M® Model 11 (Glen Mills Inc. Clifton NJ, USA) with a 40K manual fill-cell at 14,000 psi with 4-6 passages until no longer viscous. The lysate was collected and spun first at 1000 RCF /10 min to collect a loose cell debris pellet, and the supernatant spun again 6000 RCF /15 min to separate the insoluble (tight) pellet from the soluble lysate.

2.4.4 Column chromatography purification

2.4.4.1 Metal ion affinity chromatography

Cell lysates were filtered through a 0.22 µm pore filter and loaded using a bench-top peristaltic pump at 1 mL/min onto a pre-charged 5 mL Nickle iMAC HisTrap™ column (GE Healthcare) which had been equilibrated in HisTrap Buffer A (50 mM Tris, 200 mM NaCl, 10 mM imidazole, pH 8.0). Column was then washed with 5 column volumes of HisTrap buffer A (wash). Fractions were then eluted either into a single elution, using the bench-top peristaltic pump, with 5 column volumes of HisTrap buffer B (50 mM Tris, 200 mM NaCl, 250 mM imidazole, pH 8.0), or the column was transferred to an AKTA Pura (GE Healthcare) and 1 mL fractions were eluted on a gradient from 0-100% HisTrap buffer B. All fractions of interest were stored at 4 ˚C.

In the case of biding Co TALON resin in a gravity flow column, lysate was incubated for at least one hour with the resin at 4˚C, with gentle agitation. Column was washed with 2 x 15 mL buffer A, eluted with 10 mL buffer B and washed again with 10 mL buffer A (buffer A and B in this case contained 300 mM NaCl).

2.4.4.2 Anionic exchange chromatography

The periplasmic cell fraction was prepared as in 2.4.3.1, filtered through a 0.22 µm pore filter and the pH adjusted to 8.0. The fraction was then loaded onto HiTrap Resource Q™ column equilibrated in HiTrap Q buffer A (50 mM Tris, 10 mM NaCl, 10% glycerol pH 8.0) using peristaltic pump on bench at 1 mL/min. 1 mL fractions were eluted on a gradient from 0-100% HiTrap Q buffer B (50 mM Tris, 1 M NaCl, 10% glycerol pH 8.0) using an AKTA Pura (GE Healthcare). Fractions were immediately transferred to 4 ˚C for storage after elution.

2.4.4.3 Size exclusion chromatography

Fractions from 2.4.4.2 containing AaaAt3 (determined by AaaA activity assay and SDS-PAGE gel) were pooled and concentrated in a Vivaspin® 20, 10,000 MWCO spin column, with spinning at 10,000 RCF at 4 ˚C until volume was <1 mL. Sample was directly loaded onto a Superdex™ 75 10/300 column

pre-equilibrated in 50 mM Tris, 200 mM NaCl, 10% glycerol, pH 8.0, using an AKTA Pura (GE Healthcare). 1 mL fractions were then eluted and transferred to 4 ˚C. There was an unavoidable delay of ~72 hrs between anionic exchange and size exclusion chromatography, during which samples were stored at 4 ˚C.

2.5 Bacterial cloning

All cloning reagents were from New England Biosciences (U.K.) and used according to the manufacturer's instructions unless otherwise stated.

2.5.1 Polymerase chain reactions

Primers (listed in **Table 2.3**) were synthesised by Merck Life Science UK Ltd (Dorset) and delivered either in Tris-HCl buffer (10 mM, pH 8.0) as a 100 µM stock, or as lyophilised products which were reconstituted to 100 µM in nuclease-free (NF) water (Qiagen Ltd, Manchester, UK). Axygen™ thin wall polypropylene PCR tubes (Fisher Scientific UK Ltd, Loughborough) and either a Biometra T3 or Biometra TProfessional (with 96-well gradient block) thermocycler was used for thermal cycling.

2.5.1.1 Cloning PCRs

Unless otherwise specified, for all high-fidelity PCRs for cloning, Q5 high fidelity 2X master mix or the Q5 high fidelity DNA polymerase kit with high GC enhancer (NEB UK Ltd, Hitchin) was used. Total reaction volume was 25 µL and thermocycling conditions were according to the manufacturer's instructions. In general, a gradient thermocycler ranging from 50˚C to 72˚C was initially used to conduct 5-10 identical PCR reactions, varying the annealing step temperature across this range, to determine the optimal T_m for each PCR. If a product amplified successfully at multiple T_m's, the highest temperature was selected to minimise non-specific products. The chosen T_m will be specified in the relevant results chapter.

2.5.1.2 Colony PCRs

For colony PCRs, 20 µL OneTaq HotStart (NEB) or GoTaq Green Hot-Start (Promega UK, Southampton) was used according to manufacturer's instructions. A spike of a colony was taken from a plate and added directly to

the PCR tube using a 1 µL loop, which was spun against the wall of the tube to disperse the colony in the master mix. Thermocycling conditions were according to the manufacturer's instructions, except that the initial denaturation step lasted for 10 mins, to boil samples and release gDNA, as well as to activate the Hot-start polymerase. In cases where this resulted in poor amplification, colonies were instead dispersed into 100 µL PBS and boiled at 98℃ for 10 mins in a heat block, and 1 µL of this was used as the template.

2.5.1.3 Agarose gel electrophoresis

To visualise amplicons, 6x loading dye was added and DNA were then loaded, along with 5 µL of a size-appropriate ladder (2-log, 2kb+ or 100 bp) onto 0.8-1% agarose gel containing 0.01 % (v/v) of SYBR® Safe (Life Technologies), in 1X TAE buffer (40 mM Tris base, pH 8.0, 50 mM EDTA, 0.1142 % (v/v) glacial acetic acid). Gels were electrophorized using a horizontal gel apparatus (Bio-rad) at 45 V for 40-60 mins, then imaged using a Gel Doc™ XR+ System (Bio-rad, UK).

2.5.2 Site directed mutagenesis

Site directed mutagenesis (SDM) was used to create the AaaA-D580E (3.6), Spytagged AaaA (6.2.5) and AaaAt3 extended-His-tag (4.3.3.2) constructs. Primers for inverse PCRs were designed using NEB-base changer, with some alterations using SnapGene to limit multiple binding sites. Cloning PCRs were completed as specified in 2.5.1.1. Details of specific primers, templates and T_m will be covered in the relevant results chapters. Successful PCR products were selected based on agarose gel inspection and cleaned using the Monarch® PCR clean up kit, or gel extraction kit (NEB UK Ltd, Hitchin) and DNA quality and quantity was checked using a NanoDrop ND-1000 spectrophotometer (Nanodrop Technologies).

The 3' ends of the linear amplicons were phosphorylated by polynucleotide kinase (PNK) reaction then ligated. PNK reactions were made up in a total volume of 50 µL, in thin-walled PCR tubes containing 1 µL T4 PNK, 5 µL T4 DNA Ligase Buffer (10x) and 100 ng PCR product. The reaction was

incubated in a thermocycler at 37°C for 60 minutes and then heat inactivated at 65°C for 20 minutes. 1 µL T4 DNA Ligase, plus a further 5 µL T4 DNA Ligase Buffer (10x) was added directly to the tube, which was incubated again at 16°C for 16 hrs, then heat inactivated at 65°C for 10 minutes.

2.5.3 Restriction digest cloning

2.5.3.1 Double restriction digest

Double restriction digests for the construction of pEX18-Gm-Δ*aaaA* (5.3.2.1) were done in a 20 µL reaction volume including at least 1 µg of target DNA and 0.5 µL of each restriction enzyme (all NEB HF versions, 20,000 U/mL). The vector digestion reaction also contained 10 units (1 µL of 10,000 U/mL) Calf Intestinal Alkaline Phosphatase (CIP) to prevent recircularization. The reactions were incubated in a thermocycler at 37°C for 1 hr, then to improve efficiency, an additional 0.5 µL of each restriction enzyme was added and reaction incubated for a further 1 hr, then heat inactivated at 65°C for 20 min. Following digestion, 4 µL 6x loading dye was added and the full volume of product was run on a 0.8% agarose gel. DNA were visualised on a UV transilluminator, excised using a scalpel and extracted using the Monarch™ gel extraction kit, according to manufacturer's instructions. DNA quality and quantity was checked using a NanoDrop ND-1000 spectrophotometer (Nanodrop Technologies).

2.5.3.2 Ligation

Ligation reactions were made in a total volume of 20 µL using 1 µL T4 DNA Ligase (NEB, 400,000 U/mL) and 20-100 ng of the insert and vector at a range of ratios (1:2, 1:3, 2:3 of vector: insert). The reactions were incubated in a thermocycler at 16°C for 16 hrs, then heat inactivated at 65°C for 10 minutes. The product was dialysed by dispensing the full volume on a MF-Millipore™ 0.22 µM MCE Membrane (Merck Life Science UK Ltd, Dorset), floating on NF-water and incubating for 5-15 mins. The ligation product was then aspirated and immediately transformed into bacteria and any remaining product stored at -22°C.

2.5.4 Transformation and conjugation

2.5.4.1 Creation of competent cells

Electrocompetent *E. coli* was made in-house by first growing overnight cultures as described in 2.2. Over-day cultures were made using the overnight cultures, inoculated at a starting OD_{600nm} = 0.05 into 100-250 mL LB in flasks (with a total capacity volume 5-fold that of the media), and incubated at 37˚C, 200 RPM. Cells were grown to mid-log phase (OD_{600nm} = 0.6-0.8) then harvested by centrifugation at 8000 relative centrifugal force (RCF)/ 5 min/ 4˚C and washed thrice in $1/10^{th}$ the original culture volume of filter-sterilised, ice-cold 10% glycerol. Cells were pelleted and finally re-suspended in $1/50^{th}$ culture volume in 10% glycerol and either used immediately or stored at -80˚C. Once thawed cells were never re-frozen. Where chemically competent cells were used, they were purchased from NEB UK Ltd (Hitchin).

Electrocompetent *P. aeruginosa* was made in-house by growing overnight cultures of recipient cells at in 5 mL LB at 42˚C with 200 RPM shaking for 16-18 hrs. Growth at 42˚C assists in inhibiting endonucleases in *P. aeruginosa* that degrade foreign DNA, thus improving conjugation efficiency (Sana *et al.*, 2014). Therefore, overnight cultures of all recipient *P. aeruginosa,* for transformation and conjugation, were grown at 42˚C. Cells were harvested by centrifugation (8000 RCF/ 5 min) and washed thrice in 5 mL filter-sterilised 300 mM sucrose at room temperature. Cells were pelleted and finally re-suspended in 100 µL culture volume in 300 mM sucrose and used immediately.

2.5.4.2 Electroporation transformation

1-5 µL plasmid DNA (~100 ng/µL), purified either by dialysis or using a mini-prep extraction kit, was added to pre-chilled 2 mm electroporation cuvettes. Added to this was 50-100 µL electrocompetent cells (specified in the relevant results chapter, either purchased commercially or made in-house as described in 2.5.4.1) and the mixture was incubated on ice for 15 mins. Cells were electroporated using a Bio-rad Micropulser at 12.5 kV/cm as previously described (Sambrook & Russell, 2001). Pre-warmed LB or Super Optimal broth with Catabolite repression (SOC) media (NEB UK Ltd, Hitchin) was immediately added directly to the cuvette to a final volume of 1 mL, and cells were

recovered at 37°C, either statically for 1 hr if *E. coli* or with shaking at 200 RPM for 2-3 hrs if *P. aeruginosa*. Following recovery, 100 µL cells was spread onto pre-warmed LB plates containing selective antibiotics (specified in **Table 2.2**). To maximise recovery the remaining 900 µL was pelleted, resuspended in 100 µL LB, and this volume was also plated as above.

Plates were incubated overnight at 37°C, or for a maximum of 48 hrs to allow colonies to appear. After the incubation period, 4-8 colonies were sub-cultured onto further selection plates, and incubated at 37°C overnight. In most cases, the loop used to subculture was then also used to spike a 20 µL master mix for a colony PCR (as described in 2.5.1.2) using relevant primers, as specified in results chapters, to identify clones with successful insertion.

2.5.4.3 Chemical transformation

1-10 µL plasmid DNA (not necessarily purified) was added to 50-100 µL chemically competent cells, as specified in the relevant results chapters, and incubated on ice for 15-30 mins. Cells were then heat-shocked at 42°C using a water bath or heat block for 30 s, then quickly transferred back to ice and incubated for 5-15 mins. Cells were then recovered, plated, and colonies screened for as described above.

2.5.4.4 Blue/white screening

Where stated in the relevant results chapters, blue/white screening was also employed to confirm insertion of a desired gene into a plasmid via disruption of the LacZ gene. In this case, in addition to selective antibiotics, 32 µL of 125 mM IPTG + 75 mg/mL X-Gal was spread onto selection plates prior to addition of cells, and plates dried under a flame or in a hood for 15-30 mins. Transformed cells were then spread on the plates as above. Following incubation, white colonies, which were indicative of successful disruption of LacZ, were picked and screened as above.

2.5.4.5 Conjugation

In all conjugations performed here, *E. coli* S17-1 *λ-pir* was used as a donor strain, and the recipient strain was *P. aeruginosa*, details of which will be specified in the relevant chapters. Overnight cultures of donor and recipient

strains were grown, as in 2.2, but at 37˚C and 42˚C, respectively (see 2.5.4.1). Cells were then pelleted (8000 RCF /5 min /4 ˚C), carefully washed once in LB and pelleted again. Each pellet was resuspended in 50 µL LB, the two suspensions mixed, dropped onto an LB plate and allowed to dry for 10-15 min under a Bunsen flame. Plate was incubated at 30˚C for at least 6 hrs, then all bacteria were scraped up using a 10 µL loop and resuspended in 1 mL LB. 100 µL and 900 µL of this suspension was added to PIA plates with the required antibiotics to select for insertion positive *P. aeruginosa* strains and incubated and screened as described previously in 2.5.4.2.

2.5.5 Sanger sequencing

Sanger sequencing of both plasmids and PCR products was undertaken externally by both Source Bioscience (Nottingham, UK) and Eurofins Genomics (Ebersberg, Germany). Fresh plasmid DNA was extracted from 1-2 mL overnight bacterial cultures (prepared as in 2.2) using either the Monarch® Plasmid Miniprep (NEB UK Ltd, Hitchin) or GenElute™ Plasmid Miniprep (Merck Life Science UK Ltd, Dorset) kits. For sequencing of PCR products, product was cleaned using either Monarch® PCR clean up, or gel extraction kit (NEB UK Ltd, Hitchin). DNA quantity and quality was checked using a NanoDrop ND-1000 spectrophotometer (Nanodrop Technologies) and adjusted according to the sample preparation instructions in NF-water. Sequences were analysed and aligned to their respective reference using Snapgene (Insightful Science, 2022).

2.5.6 Whole Genome Sequencing (WGS)

To confirm successful *aaaA* knock out and assess any possible secondary mutations in PAO1L Δ*aaaA*, a bacterial pellet containing cells equivalent to OD_{600nm} 12.0 was shipped to MicrobesNG (University of Birmingham, UK) for DNA extraction, library prep and Illumina sequencing. The resulting raw reads were assembled with PATRIC, using parameters set to automatic (Wattam *et al.*, 2017). The genome was annotated using RAST tool kit (RASTtk) and assigned the unique genome identifier 287.25427 (Brettin *et al.*, 2015).

Two alignment tools were used, BWA-mem strict, which is specifically useful when the two genomes are predicted to be very similar, as in this case, and Bowtie2 (Li, 2014; Langmead & Salzberg, 2012). Two SNP callers, Freebayes and SAMtools were then used on each alignment (Marth *et al.*, 1999; Li *et al.*, 2009b). Both SNP callers were run with the default parameters and the raw SNPs filtered by SNP quality (QUAL>10) and read depth (DP>5) to keep only the high-quality SNPs. SnpEff was then used to annotate variants based on their genomic locations and predict coding effects (Cingolani *et al.*, 2012). The resulting all.var.tsv files generated by each of the four methods were collated into one file and only the entries which conformed to the following criteria were retained:

1) The var_frac was ≥ 0.95, var_cov was ≥ 2 and quality score > 20
2) SNP called by Freebayes was identified in both BWA-mem strict and SAMtools alignment

Var_cov refers to the average read depth of the variant and Var_frac refers to the fraction of the variant read depth among the all the reads that cover this region. These both contribute to the Score, which is a quality score assigned by the variant caller tool used, either Freebayes or SAMtools. The higher each of these number are, the higher the quality of the predictions.

2.6 miniCTX-lux transcriptional reporter assays

The miniCTX-lux transcriptional reporter assays were done in planktonic and SCW conditions, in the same 96-well plate, in a matched pairs fashion. First, overnight cultures of bacteria containing the miniCTX-*lux* plasmid fused to P_{aaaA} and/or P_{kan} were grown from three different colonies (biological triplicate) from the same plate. Overnights were then subcultured into over-day cultures at a starting OD_{600nm} 0.05, grown until mid-exponential phase (0.5-0.8) then pelleted by centrifuging at ~8000 RCF / 2 min and washed in fresh LB once and normalised to OD_{600nm} 0.1. The over-days were then split in two, and diluted 1000-fold into either LB (planktonic condition) or SWF (SCW). The SCW condition was set up in 100 µL final volumes across one half of the plate as described in 2.8.1, while the LB condition was set up on the other half of the

plate, adding the equivalent volume of bacteria into a final volume of 100 µL LB.

Plates were sealed with a 4titude Gas Permeable, optically clear, 96-well adhesive seal (except when stated otherwise) and incubated in a TECAN Infinite® F200Pro plate reader. The plate was incubated statically at 37 ˚C for 18 hrs, reading luminescence (in relative light units, RLU) and OD_{600nm} every 20 min using the TECAN iControl application. Data were analysed using Microsoft Excel®, where luminescence was OD_{600nm}-normalised by dividing RLU by the corresponding OD_{600nm} for each data point.

2.7 L-arginine-p-nitroanilide cleavage (AaaA activity) assay

As depicted in **Figure 1.6**, AaaA enzymatic activity was assayed by observing the cleavage of L-arginine from L-arginine-p-nitroanilide by either a whole cell suspension or lysate, which correlates with a colour change from colourless to yellow at absorbance at 370-410 nm (Luckett *et al.*, 2012). For whole cell assays, cells were grown and induced as described in the relevant results chapters. In general, 1 mL of culture at OD_{600nm} 1.0 was harvested, pelleted (8000 RCF /2 mins /4 ˚C) and washed thrice in Minimal Media P (MMP, recipe in **Table 2.4**). The final OD_{600nm} of whole cell suspension added to the experiment depended on the expected level of AaaA production and is specified in the relevant figure legend.

Depending on the experiment, 150 µL of whole cell suspension, equivalently volume-adjusted cell lysates (prepared as described in 2.4.3) or protein fractions (obtained as described in 2.4.4) were added in triplicate to a clear 96-well plate along with 150 µL of L-arginine p-nitroanilide dihydrochloride (Merck Life Science UK Ltd, Dorset) to a final well concentration 0.5 mM, and placed in TECAN Infinite® F200Pro or SPARKS plate reader at 37˚C for 6-24 hours, with a 5 s orbital shaking step followed by absorbance readings at both 370-410 nm (yellow) and 600 nm (OD_{600nm}) every 5 - 20 minutes, depending on the experiment and plate-reader available.

For the activity assay in SCW conditions described in 5.5.2, the SCW was prepared in 96-well plates, as described in 5.5.1, and after a 24 or 48 hr incubation, 100 µL collagenase (500 µg/mL) was added to each well and the plate incubated at 37 ˚C for 1 hr. Samples were then transferred to Eppendorf tubes using a pipette, vortexed for 5 s and incubated at 37 ˚C for a further 30 min. Each sample was washed twice in 1 mL PBS followed by resuspension in 1 mL MMP and 100 µL added to the activity assay as described above.

2.8 The synthetic chronic wound model (SCW)

'Synthetic chronic wound' (SCW) media refers to the final media used in the synthetic wound model, following collagen and bacterial suspension (where relevant) addition and polymerisation. Synthetic wound fluid (SWF) refers to the liquid media designed to simulate wound exudate (recipe in **Table 2.4**), which was the major constituent of the SCW, along with collagen.

2.8.1 Preparation of the SCW model

The SCW model here is from da Silva (2018), and was made using Corning™ Collagen I, High Concentration, Rat Tail (Fisher Scientific, Loughborough, catalogue number: 10224442) and synthetic wound fluid (SWF). SWF was prepared in batches with 50% peptone water (Oxoid) which was made up according to the manufacturer's instructions, autoclaved then added to 50% foetal bovine serum and filter sterilised. Once prepared, the SWF was frozen in 10 mL aliquots, and freeze-thaw was avoided.

The combination of the SCW and SWF, plus acetic acid and sodium hydroxide, created semi-solid 3D matrix which could be polymerised in either 8-well chambers for microscopy, 96/24 well plates, or on glass slides in custom-cut silicon moulds (photographed in **Figure 5.10**). The silicon moulds were cut from a 5 mm thick sheet of silicon using a 16 mm diameter punch tool and glued to sterile glass slides with silicon glue. Glued slides were left to dry for at least an hour then autoclaved.

Bacteria for inoculation into the SCW were prepared by first growing overnight cultures in LB. Each biological replicate came from a different colony from the same plate. Overnight cultures were incubated at 37 ˚C for 16-18 hrs

then inoculated at 1% v/v into a further 5 mL LB for over-day culture, which was incubated as before for 8 hrs to allow cells to enter mid-stationary phase. After incubation, bacteria were normalised to OD_{600nm} 0.1 in 1 mL maximum recovery diluent (Oxoid, UK) then diluted 1000-fold into 1 mL SWF.

The following volumes are for a 200 µL SCW plug, however volumes varied, and this will be stated in the relevant methods section. Addition of a 10-15% error margin is also recommended when preparing the media, as collagen is viscous and prone to pipetting error. The SCW was prepared on ice, under a sterile Bunsen flame, by carefully combining 40 µL collagen (final concentration of 2 mg/mL, adjusted by volume if the stock concentration was not exactly 10 mg/mL) and 20 µL ice cold acetic acid 0.1% in a 1.5 mL Eppendorf tube. Where the SCW was inoculated with bacteria, 100 µL of the 1000-fold diluted bacterial inoculum (diluted from the over-day) was added followed by 20 µL of SWF.

To initiate polymerisation of the collagen, 20 µL ice cold sodium hydroxide 0.2 M was finally added, carefully mixed with a pipette, avoiding air bubble creation, then aliquoted into the mould or well. The mixture would begin to polymerise at room temperature within one minute, thus each biological repeat was made one at a time. The SCW was then moved to a humid chamber, a water-tight box containing tissue soaked in sterile water, as well as four bijou tubes containing 4 mL sterile water (photographed in **Figure 5.10a**), which had been UV-sterilised for 20 mins. Once all samples had been prepared, this was transferred to a 37 ˚C standing incubator.

2.8.2 Colony forming unit determination from SCW samples

After preparation and incubation of the SCW, samples harvested by removing the silicon mould with a sterile scalpel, and transferring the collagen plug whole by tipping (with assistance of the scalpel if needed) into sterile 2 mL FastPrep™ tubes containing 6x 2 mm diameter ceramic beads plus 400 µL sterile PBS. Used silicon moulds were soaked in 70% ethanol overnight, dried, autoclaved and re-used. Samples were homogenised with one 10 s cycle at 4.0 m/s using the FastPrep-24 5G (MP Biomedicals™). Tubes were then

transferred to a sonicating water bath, disrupted for 5 min, then briefly vortexed. 20 µL of each sample was added to each well of a 96-well plate containing 180 µL maximum recovery diluent (Oxoid). A multichannel pipette was used to make 1:10 serial dilutions in each row of 96-well plates to a final dilution of 10^{-11}. 10 µL of each dilution was plated in triplicate onto PIA plates and incubated at 22 ˚C for 18-36 hrs, until clear colonies had appeared. Plates were then photographed over a light box, the photos viewed on a computer and colonies were counted from the lowest dilution that had between 30-150 colonies visible.

2.9 Examining gene expression in the SCW

2.9.1 For RT-qPCR

2.9.1.1 RNA extraction

Firstly 200 µL SCW (slide mounted) plugs were prepared according to 2.8.1 and biological triplicate plugs were inoculated with either PAO1L or PAO1L Δ*aaaA* and incubated in a humid box at 37˚C for 16-48 hrs. Following incubation, each silicon mould was removed with a scalpel and the whole plug tipped into a 1.5 mL Eppendorf containing 400 µL RNA-Protect™ (Qiagen), vortexed for 10 s then incubated at room temperature for 5 min, as per the manufacturer's instructions. Tubes were spun at 5000 xg /10 mins, the supernatant gently removed by pouring, then the inverted tube dabbed gently on a paper towel for 5 s to dry, and immediately frozen at -80 °C.

For RNA extraction, the RNeasy mini extraction kit (Qiagen) was used following protocol 5 for mechanical and enzymatic lysis according to the manufacturer's instructions (QIAGEN, 2020 pg. 34). Previously frozen RNAprotect® pellets were thawed on ice and each resuspended in 100 µL TE buffer (30 mM Tris-HCl, 1 mM EDTA, pH 8.0) containing 15 mg/mL lysozyme. 10µL proteinase K was added and the mixture incubated at 22 ˚C, 200 RPM for 30 mins. Following addition of 700 µL buffer RLT, the mixture was then transferred to UV-sterilised 2 mL FastPrep™ safe-lock tubes containing 0.6g lysis beads (150–600 µm diameter) and homogenised with two 30 s cycles at 6.0 m/s using the FastPrep-24 5G (MP Biomedicals™). Following lysis, tubes

were centrifuged at max. speed for 1 min, supernatant added to 590 µL 80% ethanol and mixed well. RNA was then purified using the RNeasy mini spin column kit according to protocol 7 (QIAGEN, 2020 pg. 40) and eluted in 10 µL nuclease-free water. The Turbo DNA free kit (Invitrogen) was used to remove genomic DNA according to manufacturer's protocol for rigorous DNase treatment and a NanoDrop ND-1000 spectrophotometer (Nanodrop Technologies) was used to quantify RNA and check 260/280 for signs of impurities.

2.9.1.2 Reverse transcription to cDNA

The Promega GoScript™ RT (reverse transcriptase) kit with random primers was used to create cDNA according to manufacturer's protocol, including the optional RNA denaturation step. 200 ng of RNA was added to each reaction (final volume 20 µL), and alongside each RT reaction, a parallel RT negative control reaction was done, where no RT enzyme was added.

2.9.1.3 RT-qPCR

2.9.1.3.1 Standard curves for primer efficiency

Firstly, gDNA was extracted from PAO1L using the GenElute™ Bacterial Genomic DNA Kit according to the manufacturer's instructions and eluted into 50 µL NF-water (Merck Life Science UK Ltd, Dorset). A working stock of 10 ng/µL gDNA was made in NF-water and stored in aliquots at -20 ˚C. To generate the standard curve, a 5 x 1:10 dilution series was made fresh each time, by adding 10 µL working gDNA stock to 90 µL NF-water, vortexing briefly then spinning in a benchtop microfuge for ~10 s.

For the RT-qPCR, all reagents were sourced from Applied Biosystems, UK. Primer-specific master mixes for a final well volume of 25 µL per reaction were prepared. Each master mix contained 0.625 µL of each forward and reverse primer (final concentration 250 nM), 12.5 µL PowerSYBR master mix (Applied Biosystems) and 10.25 µL NF-water per reaction plus 10% error volume, which was mixed by pipetting then added to MicroAmp™ Optical 96-Well Reaction Plates. Finally, 1 µL of each gDNA standard curve dilution (final well concentration: 400 pg/µL to 0.04 pg/µL = ~6.2 copies) was added to respective wells and the plate was sealed with MicroAmp™ Optical

Adhesive Film. 'No template' controls (where NF-water was used instead of template) were including on every plate for each primer set. All RT-qPCR reactions were run on the same ABI7500 Fast qPCR machine, within the same calibration time-frame. Thermocycling included an initial 10 min 95 ˚C denaturation step followed by 40 cycles of 15 s at 95 ˚C and 1 min at 60 ˚C, followed by a melt curve. Standard curves and reaction efficiency were calculated using the ABI7500 software.

2.9.1.3.2 Relative quantification using ΔΔCt

For quantifying relative gene expression, 0.5 µL cDNA generated in 2.9.1.2 was added to qPCR master mix, which was prepared and added to the qPCR plate as in 2.9.1.3.1, except with 10.75 µL NF-water per reaction. 'No template' controls for each primer set, and a *rpoD* RT negative control for each cDNA sample were added to check for the absence of contaminants or gDNA, respectively. Primer concentrations were 250 nM, with the exception of *aaaA* primers in 5.6.2 which were 300 nM.

Thermocycling was run as in 2.9.1.3.1. Expected primer efficacies were adjusted using the ABI7500 software, according to the efficiencies measured by the standard curve experiments for each primer set. At the end of the run, negative controls were inspected to ensure there was no amplification- where there was amplification in negative controls, it was considered acceptable if Cts were ≥10 greater than sample Cts. Melt curves were also run to assess for the presence of primer-dimers which could explain this. ΔΔCt and relative quantity (RQ) were calculated using the ABI7500 software, using *rpoD* as the endogenous control.

2.9.2 For RNA-Sequencing

2.9.2.1 Sample preparation

For RNA-Seq, a total of 12 SCW samples were prepared as described in 2.8.1. Six each were infected with 8-hour over day cultures of PAO1L and PAO1L Δ*aaaA* from different colonies of the same plate, plus two non-infected controls. The 12 samples were randomised and assigned to one of two humid boxes, each containing six samples plus one non-infected control. The

preparation of the samples for each box was staggered by one hour at every stage to allow timings to remain precisely equivalent between all samples. Once the SCW was prepared, samples were incubated for 16 hrs in their respective humid box at 37 ˚C, in the same incubator. After 16 hrs incubation, samples were removed and 50 µL of either SWF alone, or SWF containing 100 µM L-arginine was added to each SCW sample, which were then returned to the box and incubated for a further 1 hr. Next, moulds were removed using a sterile scalpel and samples transferred by tipping into 400 µL TRI-Reagent (Invitrogen), vortexed for 10 s then couriered to Novogene on dry ice the same day. Extraction was completed by their partners at Qiagen, and library prep, sequencing and data analysis completed by Novogene. Sample quality control was done by assessing RNA integrity and quantity using gel electrophoresis and Bioanalyzer (Agilent).

2.9.2.2 RNA-Seq library prep and analysis

Analysis was completed by Novogene, using the following methods, supplied by Novogene Co. Ltd. (2022). Firstly, ribosomal RNA was removed from total RNA, followed by ethanol precipitation. After fragmentation, the first strand cDNA was synthesized using random hexamer primers. During the second strand cDNA synthesis, dUTPs were replaced with dTTPs in the reaction buffer. The directional library was ready after end repair, A-tailing, adapter ligation, size selection, USER enzyme digestion, amplification, and purification.

2.9.2.2.1 Workflow of library construction

The library was checked with Qubit and real-time PCR for quantification and bioanalyzer for size distribution detection. Quantified libraries were pooled and sequenced on Illumina platforms, according to effective library concentration and data amount required.

2.9.2.2.2 Clustering and sequencing

The clustering of the index-coded samples was performed according to the manufacturer's instructions. After cluster generation, the library preparations were sequenced on an Illumina platform and paired-end reads were generated.

2.9.2.2.3 Data Analysis Quality control

Raw data (raw reads) of FASTQ format were firstly processed through fastp. In this step, clean data (clean reads) were obtained by trimming reads containing adapters and removing poly-N sequences and reads with low quality from raw data according to the following criteria:

1. Reads with sequencing adaptor contamination
2. Reads where uncertain nucleotides (designated N) constitute more than 10 percent of either read.
3. Reads where low-quality nucleotides (Base Quality less than 5) constitute more than 50 percent of the read.

At the same time, Q20, Q30 and GC content of the clean data were calculated. All the downstream analyses were based on the clean data with high quality.

2.9.2.2.4 Reads mapping to the reference genome

The PAO1 genome and gene model annotation files were downloaded from Pseudomonas genome database website. Bowtie2 was used for both indexing and aligning clean reads to the reference genome (Langmead & Salzberg, 2012).

2.9.2.2.5 Novel gene and gene structure analysis

Rockhopper was used to identify novel genes, operon and transcription start sites, and the upstream 700bp sequence of transcription start sites were extracted to predict promoters using TDNN (Time-Delay Neural Network) (Waibel *et al.*, 1989).

2.9.2.2.6 Quantification of gene expression level

FeatureCounts was used to count the read numbers mapped to each gene. The FPKM, the expected number of Fragments Per Kilobase of transcript sequence per Million base pairs sequenced of each gene was calculated based on the length of the gene and reads count mapped to this gene.

2.9.2.2.7 Differential expression analysis

Differential expression analysis of biological groups was performed using the DESeq2 R package (Love *et al.*, 2014). DESeq2 provide statistical routines for determining differential expression in digital gene expression data

using a model based on the negative binomial distribution. The resulting P-values were also adjusted using the Benjamini and Hochberg's approach for controlling the false discovery rate to produce a 'padj' value.

2.9.2.2.8 GO and KEGG enrichment analysis of differentially expressed genes

Gene Ontology (GO) and KEGG enrichment analysis of differentially expressed genes was implemented by the clusterProfiler R package, in which gene length bias was corrected. GO terms with corrected P-value < 0.05 were considered significantly enriched by differential expressed genes.

2.9.2.2.9 Untranslated region prediction

5' and 3' untranslated region (UTR) sequences were extracted according to the Transcription Start Site (Transcription terminal Site) and Translation start site (Translation terminal site). RBSfinder (Suzek *et al.*, 2001) and TransTermHP (Kingsford *et al.*, 2007) were then used to predict the Shine Dalgarno (ShD) sequence and terminator sequence respectively.

2.9.2.2.10 Analysis of ncRNA

IntaRNA was used to predict sRNA targets (Busch *et al.*, 2008) and RNAfold was then used to predict RNA secondary structures (Hofacker & Stadler, 2006).

2.9.2.2.11 Mutation analysis

Firstly, Picard tools and Samtools were used to sort, mark duplicated reads and reorder the bam alignment results of each sample (Li *et al.*, 2009b). Then the tool HaplotypeCaller in GATK software was used to perform variant discovery, including single nucleotide polymorphisms (SNPs) and insertion/deletions (INDELs) (McKenna *et al.*, 2010). Raw VCF files were filtered with GATK standard filter method and other parameters (cluster: 3; WindowSize: 35; QD < 2.0 or FS > 60.0). Finally, SnpEff was used to annotate variants based on their genomic locations and predicts coding effects (Cingolani *et al.*, 2012).

2.10 Microscopy

2.10.1 Live/Dead staining of the SCW

First, 200 µL of SCW was prepared and inoculated as described in 2.8.1 and added to µ-Slide 8 Well high Glass Bottom chambers (Ibidi, Gräfelfing, Germany). Biological triplicate wells inoculated with either PAO1L or PAO1L Δ*aaaA* with the remaining two wells being non-infected controls. The chamber was incubated in a humid box at 37˚C for 24 hrs. Following incubation, 50 µL of SYTO-13 (25µM) followed by 50 µL propidium iodide (1 mg/mL) (Invitrogen) was added to each well and incubated in the dark at 37˚C for 30 mins. Duplicate points on each well were imaged using a Ziess laser scanning 710 confocal microscope, at magnification 40x, using channels for SYTO-9 (laser 5, gain 800) and PI (laser 5, gain 750). Images were exported as .lsm files and Fiji was used to first split the red and green channels and calculate the intensity of each from separate Z projects (max intensity). The % intensity of live (green channel, SYTO) and dead (red channel, PI) cells were calculated using the following equations:

$$\%\ \text{Live} = \frac{\text{SYTO intensity}}{\text{SYTO intensity} + \text{PI intensity}} \times 100$$

$$\%\ \text{Dead} = 100 - \%\ \text{Live}$$

Fiji was used to finally merge channels to generate a 3D-project (along X-axis, brightest point) for each z-stack and add scale bars.

2.10.2 Single cell super resolution structured illumination imaging

Overnight cultures of PAO1L Δ*aaaA::*pJM220-*aaaA* and PAO1L Δ*aaaA::*pJM220-*aaaA-St4* were inoculated at OD_{600nm} 0.05 into 500 mL flasks containing 100 mL LB supplemented with 0.005% rhamnose. Flasks were incubated at 37˚C, 200 RPM for 3 hrs, with final OD_{600nm} between 1.1-1.2. 500 µL cells at OD_{600nm} 1.0 were pelleted by spinning at 8000 RCF / 2 min, washed once, and resuspended in 500 µL PBS. 20µL ~7 mg/mL Spycatcher-mCherry was added, tubes were vortexed briefly and incubated for 45 min at 22 ˚C in

dark. Following this, cells were washed 3x in 1 mL PBS, spinning at 13000 RCF /1 min. Cells were finally resuspended in 50 µL PBS, and 2 µL added to a coverslip, which was pressed onto a slide and imaged by super-resolution microscopy using Zeiss Elyra PS.1 microscope in SIM mode, with the Plan-Apochromat 63x/1.4 Oil DIC M27 objective at λ610 nm emission.

2.11 Colorimetric arginine quantification assay

Slide mounted SCW plugs were prepared according to 2.8.1 with three biological replicates for each group: not infected, infected with PAO1L or PAO1L Δ*aaaA*. Plugs were harvested after 16 hrs of incubation at 37˚C and transferred by tipping into sterile, pre-weighed 2 mL FastPrep SafeLock™ tubes each containing six 2 mm ceramic beads. Tubes were weighed again, and 700 µL arginine assay buffer (supplied) / 100 mg sample was added. Samples were homogenised as in 2.8.2 then spun at 10,000 RCF /10 mins/ 4 ˚C. Supernatant was filtered using Vivaspin® 20, 10,000 MWCO spin columns, with spinning at 10,000 RCF/ 20 mins at 4 ˚C. 40 µL of filtrate was added to the colorimetric arginine quantification assay, in duplicate, along with standards, and the assay was performed and analysed according to the manufacturer's instructions (Abcam, 2019a).

2.12 Spycatcher binding assay

PAO1L Δ*aaaA* containing either pJM220:*aaaA*, pJM220*:aaaA-St4*, *St5* or *St6* were grown and induced with 0.01% rhamnose, as described in 2.10.2. 500 µL cells at OD_{600nm} 1.0 were harvested and washed in PBS then resuspended in 500 µL PBS. Cells were then incubated with 20µL ~7 mg/mL Spycatcher mCherry for 1 hr at 22˚C in the dark. Cells were washed thrice in 200 µL PBS then 100 µL was added in duplicate to a black, flat bottomed 96-well plate, and fluorescence λ587/610 read.

2.13 Bioinformatic analysis and modelling

2.13.1.1 DIAMOND BLASTP interrogation of AaaA sequence conservation

The sequence of AaaA from PA0328 was used to run a DIAMOND BLASTP search at www.pseudomonas.com using stringent (E-value cutoff: 10^{-12}; Default Sensitivity, Query Coverage: 100%, Identity Cutoff: 95%) parameters on all complete *Pseudomonas aeruginosa* genomes. The search produced 3,748 hits having between 0 and 27 mismatches. A temporary fasta file was produced with the hits obtained, and unique sequences were extracted using the following Perl script, RemoveRep2.pl:

```
#!/usr/bin/perl
use strict;
use Bio::SeqIO;
my %unique;
my $file   = "both.fasta";
my $seqio  = Bio::SeqIO->new(-file => $file, -format => "fasta");
my $outseq = Bio::SeqIO->new(-file => ">$file.out", -format => "fasta");

while(my $seqs = $seqio->next_seq) {
  my $id  = $seqs->display_id;
  my $seq = $seqs->seq;
  unless(exists($unique{$seq})) {
    $outseq->write_seq($seqs);
    $unique{$seq} +=1;
  }
}
```

The resulting final fasta file contained the resulting 50 unique sequences. Occurrences of each of these sequences in the temporary file were counted (using the batch CLI command grep -wc "sequence" temp.fast >> output.txt). Clustal Omega was run with the sequences in the final fasta file to produce a cladogram, which was then further manually edited to add the sequence counts.

2.13.1.2 Other bioinformatic analyses

Bioinformatic analysis of whole genome sequences was completed using PATRIC online bioinformatic resource, which provides a black box package for genome assembly, annotation, alignment and SNP calling as

stated in the relevant results chapters (Davis *et al.*, 2020). Other online tools used in this work including PROVEAN, AlignmentViewer and the PanX Genome database, with more detail given in the relevant results chapters (Choi *et al.*, 2012; Ding *et al.*, 2018; Reguant *et al.*, 2018). Snapgene was used to simulate cloning experiments, design workflows and primers and to align, visualise and present genetic sequences and vectors (Insightful Science, 2022).

3D models of AaaA were generated using iTasser, CiTasser and AlphaFold2, with manipulations including SDM, colouring and truncation done using the PyMOL Molecular Graphics System, Version 2.0 (Schrödinger, LLC), as stated in the relevant results chapters (Roy *et al.*, 2010; Varadi *et al.*, 2022; Yang & Zhang, 2015; Zheng *et al.*, 2021).

2.14 Software and Data analysis

2.14.1 Data management and statistical testing

Data were collected, compiled and analysed using Microsoft Excel (2018 or later) and GraphPad Prism 9 (Motulsky, 1995). GraphPad was used to create graphs and conduct analysis, including Area Under the Curve and Standard Curve calculations and statistical comparisons (detailed in 2.14.1.1)

RT-qPCR data was collected and analysed using the Applied Biosystems 7500 Real-Time PCR System and software (ThermoFisher Scientific, UK) using default analysis parameters and controls as stated in the relevant results chapter.

2.14.1.1 Statistical tests

All data sets were first tested for normality using the Shapiro-Wilk normality test, as this is specialised for data where n <50 (Ghasemi & Zahediasl, 2012; Nahm, 2016). For comparing between two groups of continuous data, if normally distributed, an un-paired t-test was used, and if not normally distributed, a Mann-Whitney test was used. For comparing more than two groups of continuous data, if normally distributed, unless otherwise stated, equal standard deviations were not assumed, and differences were tested for using One-Way Brown-Forsythe and Welch's ANOVA test with Dunnett T3 multiple comparisons. If not normally distributed, differences were

tested for using the Kruskal-Wallis test with multiple comparisons. Specific groups compared and other details are stated in the relevant figure legends.

Chapter 3. Using bioinformatics to examine the structure and conservation of AaaA

3.1 Background on AaaA structure

AaaA is structurally related to the M28 family of peptidases, with 100% conservation in predicted active site catalytic and ligand-binding residues between AaaA and other M28 members (Luckett *et al.*, 2012). M28 peptidases can be amino or carboxy-specific, and tetrahedrally coordinate two zinc ligands using residues including histidine, aspartic acid and glutamic acid to create their catalytic pocket (Chevrier *et al.*, 1996; Luckett *et al.*, 2012). **Figure 3.1** shows the primary sequence, along with the predicted structural domains of AaaA, including the N-terminal signal peptide, M28 peptidase and membrane-spanning auto-transporter β-barrel. Predictions were made of the signal peptide using SignalP (Armenteros *et al.*, 2019) and other domains using Pfam (Mistry *et al.*, 2021).

A BLAST search of the MEROPS peptidase database predicted that AaaA is a member of the subfamily M28.021, sharing similarities with *E. coli* isozyme activating protein (IAP) (Darch, 2009). Within the M28.005 subgroup, AaaA shares 64% amino acid similarity with *Pectobacterium carotovorum* ECA2163, throughout the protein, with 100% conservation of the predicted catalytic active sites (Darch, 2009; Luckett *et al.*, 2012). There is also absolute conservation of predicted active site residues with *Streptomyces griseus* aminopeptidase S, for which a crystal structure has been obtained (Gilboa *et al.*, 2001).

Site directed mutagenesis (SDM) studies also identified several residues which were predicted to form the active site essential for aminopeptidase activity, based on their similarity to other M28 peptidases in the MEROPS database. These, highlighted in **Figure 3.1**, were H100, D102, D115, E147 and E148 (Luckett *et al.*, 2012). Mutation of each of these residues to alanine

caused either complete loss of enzymatic activity in the case of D102, D115 and D176, or partial loss in the case of H100, E147 and E149. Fractionation confirmed that mutants D115A, E147A and E149A all localised to the OM, whereas mutation of D102 and D176 caused a total loss of detectable AaaA, and localisation was unclear in the H100A mutant (Darch, 2009; Paredes-Osses, 2014).

Functional studies have confirmed empirically that the M28 passenger domain of AaaA is an aminopeptidase, which cleaves N-terminal arginine from peptides. The *P. aeruginosa* PAO1 Δ*aaaA* knock-out mutant showed reduced respiration when dipeptides with an N-terminal arginine were the sole nitrogen source, except for Arg-Arg/Arg-Ly. This effect was most pronounced with Arg-Val and Arg-Ile, suggesting a level of substrate specificity (Luckett *et al.*, 2012). Trypsin degradation studies also showed that AaaA, which was vulnerable to trypsin degradation, is tethered to the cell surface, rather than secreted (Luckett *et al.*, 2012)

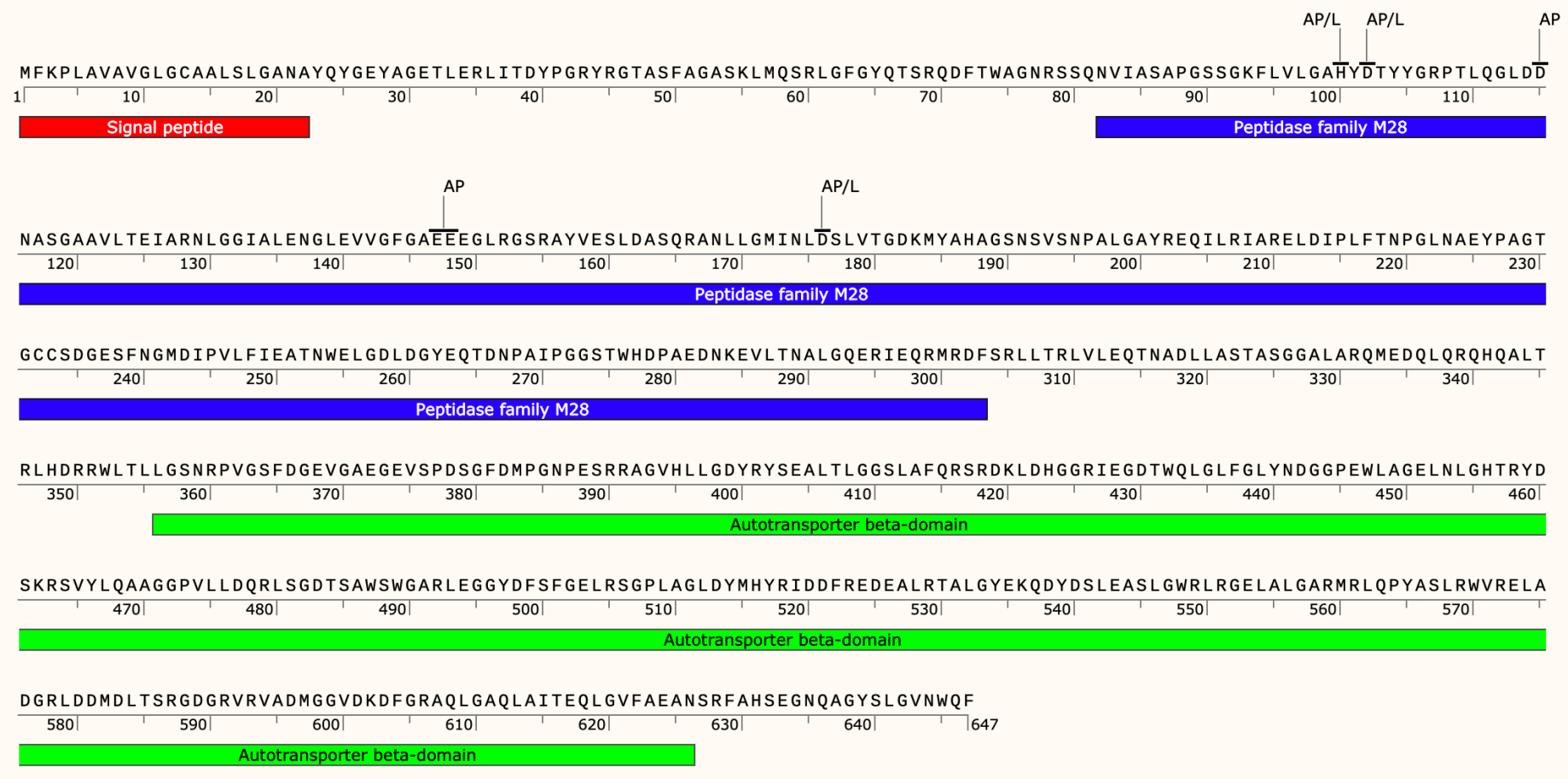


Figure 3.1: Map of AaaA primary structure. *Predicted signal peptide, M28 peptidase and autotransporter β-domain are highlighted in red, blue, and green, respectively. Predicted signal peptide cleavage site between A22-Y23 is marked CS, and residues contributing to the active site are marked AP (aminopeptidase) or AP/L if they are also essential for localisation. Generated using Snapgene.*

3.2 Chapter Aims

As AaaA is a potential vaccine target, it is essential to properly understand its structure and conservation, to evaluate its usefulness for this purpose, as well as glean insights into its function (Bianconi *et al.*, 2018). Experimental work to elucidate the structure of AaaA is described in Chapter 4, while this chapter focused on examining AaaA conservation by:

1) Using bioinformatic tools to interrogate the structure and conservation of AaaA, to better understand the likelihood and consequence of mutations.
2) Probing 'of interest' amino acid variations in AaaA using site directed mutagenesis and evaluate their potential impact on AaaA production and activity.

3.3 Generating a predictive model of full-length AaaA

Modelling the 3D predicted structure of membrane proteins, particularly autotransporters, has in the past been notoriously difficult, due to their complex motifs, including membrane spanning β-barrel domains. However recent advances have made these predictions much more reliable, with the introduction of the C-iTasser platform which incorporates machine learning and contact mapping into the existing iTasser methodology (Roy *et al.*, 2010; Yang *et al.*, 2015; Yang & Zhang, 2015; Zheng *et al.*, 2021). Using CiTasser, it was possible to create the first predictive 3D model of full-length AaaA, shown in **Figure 3.2**, which retains the structural features predicted in each domain.

In 2020, 3D protein modelling was further revolutionised by DeepMind's AlphaFold, and later AlphaFold2, which outperformed every other methodology at the Critical Assessment of Techniques for Protein Structure Prediction (CASP-14) meeting in 2020 (Callaway, 2020; Jumper *et al.*, 2021; Varadi *et al.*, 2022). AlphaFold has since become the platform of choice for 3D structure prediction for many scientists. Therefore, to compare the 3D models for AaaA predicted by C-iTasser and AlphaFold2, the model for AaaA was obtained from the AlphaFold Protein Structure Database (AlphaFold Protein

Structure Database, 2022) and the two M28 domains were aligned using PyMOL (**Figure 3.3**). There was generally good alignment, with an RMSD of 1.29 Å. The RMSD, or root-mean square deviation, is the most commonly used quantitative measure of the similarity between two superimposed atomic coordinates, with a smaller RMSD indicative of closer structural similarity (Kufareva & Abagyan, 2012). The low RMSD between the two M28 domains, plus the near-perfect alignment of the active site residues suggests that the two models are in general agreement.

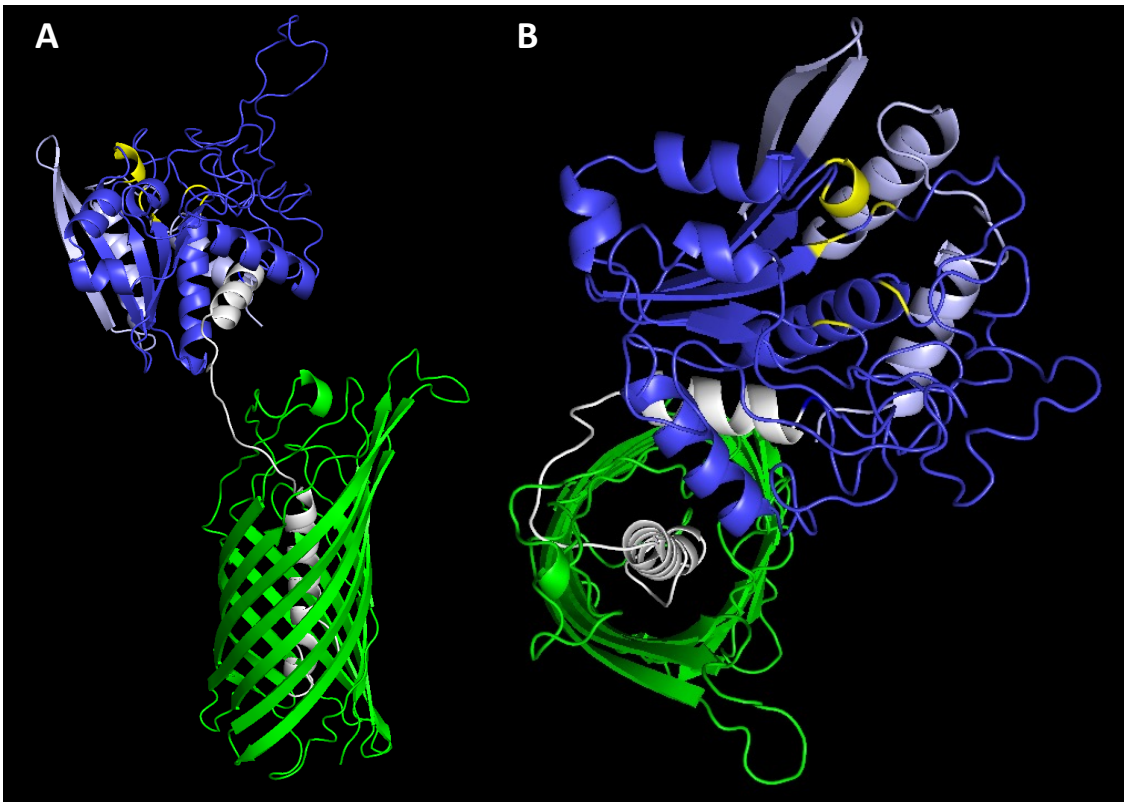


Figure 3.2: Predicted 3D model of AaaA. Panel A *shows a side view,* ***Panel B*** *a top view. The M28 peptidase and membrane spanning autotransporter β-barrel domains are shown in blue and green, respectively. Linker region between cleaved signal peptide domain (not shown) and M28 domain in light blue and other linker regions shown in white. The residues contributing to the active site or localisation are in yellow (H100, D102, D115, E147-E149 and D176). Generated using C-iTasser (model 1)* (Roy *et al.*, 2010; Yang *et al.*, 2015; Yang & Zhang, 2015; Zheng *et al.*, 2021)*, C-score= 0.37, Estimated TM-score = 13.3 ±4.1Å, Estimated RMSD = 7.0 ±4.1Å. Coloured using PyMOL.*

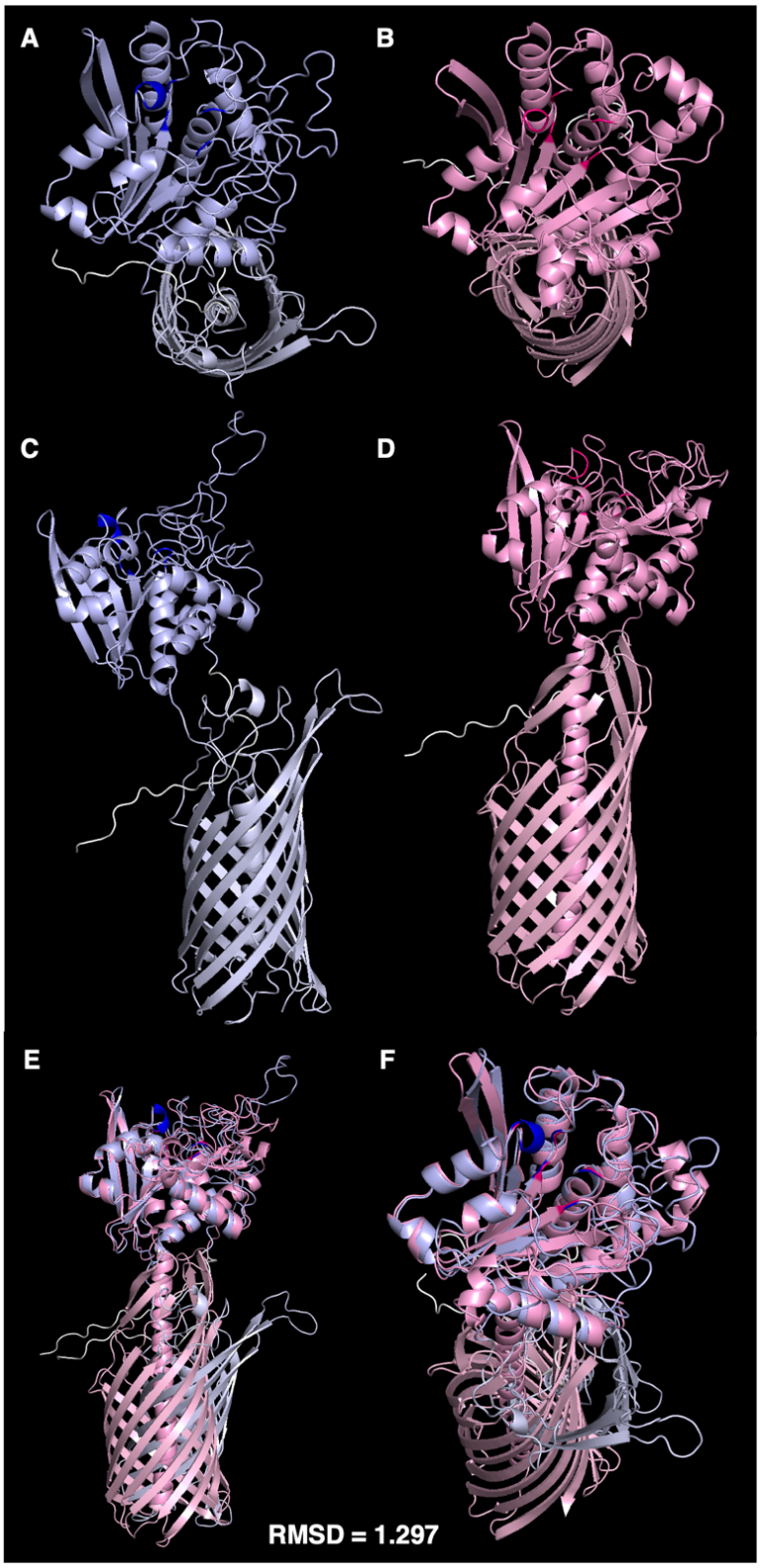


Figure 3.3: Comparison of modelling techniques for predicting a 3D model of AaaA. *Models by CiTasser and AlphaFold are shown respectively, in blue (**A**: top view, **C**: side view) and pink (**B**: top view, **D**: side view). Active site residues are highlighted in the darker shade of blue and pink, respectively. Though there is some disagreement on the autotransporter domain, major features in the M28 domains are retained in both models, which align well, with near-perfect active site alignment, as shown in **panels E** (top side view) and **F** (top view). Models aligned and coloured using PyMOL.*

3.4 AaaA is highly conserved across *P. aeruginosa* strains

One of the major reasons for the interest in AaaA is because of its high level of genetic conservation within *P. aeruginosa*. Not only does this high conservation point to the likelihood that the gene is essential for *P. aeruginosa*, but, as mentioned above, puts AaaA in the top 10 potential candidates for inclusion in a *P. aeruginosa* conjugate vaccine (Bianconi *et al.*, 2018). To investigate this further, conservation of the AaaA primary structure was probed using both the PanX genome database of 153 genomes, and then the Pseudomonas.com database of 3748 genomes.

3.4.1 Examination of the PanX genome database shows high AaaA conservation

PanX is an online tool which, at the time of this work being undertaken, allowed genes from the genomes of 153 *P. aeruginosa* isolates to be searched and aligned to PAO1 (Ding *et al.*, 2018). Within the 153 isolates, AaaA (PA0328 in the reference PAO1) had >99% protein sequence identity in all but 4 isolates, which form a distinct clade, as seen in the gene tree in **Figure 3.4**. 2 of these were CR1 strains, isolated from the environment, 1 was a MDR PA7 strain, a non-respiratory human isolate from Argentina and 1 was an AR0356 strain, from the FDA/CDC Antimicrobial Resistant Isolate Bank.

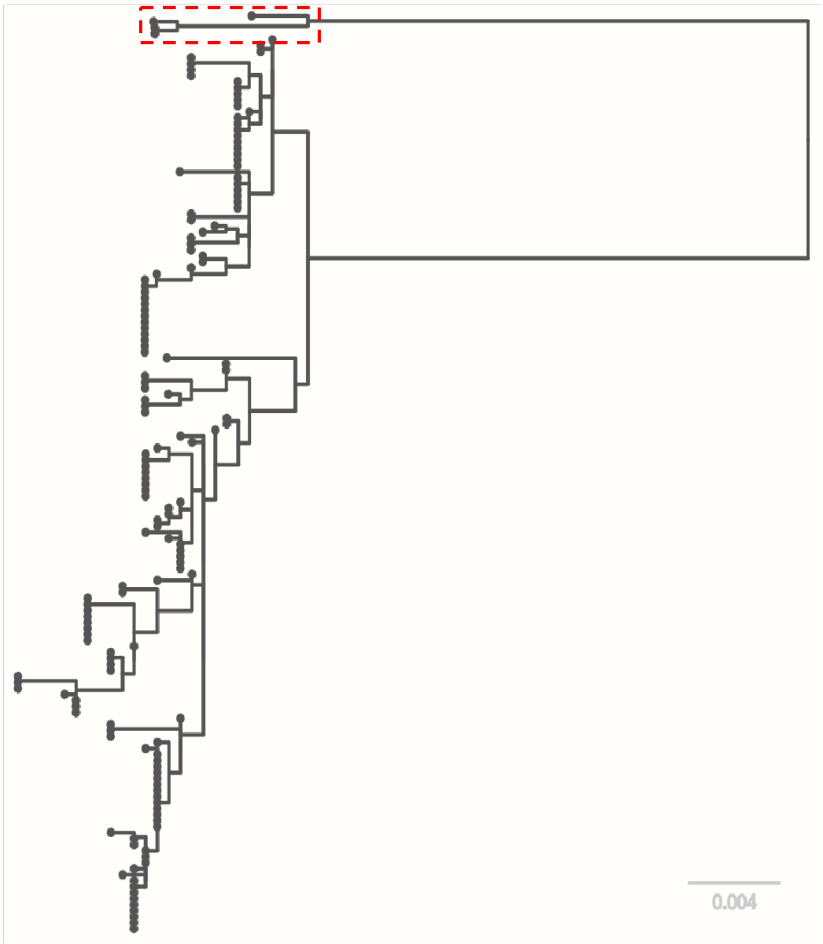


Figure 3.4: Tree of relatedness for AaaA created using the PanX database of 153* P. aeruginosa *isolates. *Amino acid sequence identity is >99% for all but four isolates: two CR1, one PA7 and one AR0356, which form a distinct clade (dashed red box).*

All amino acid variations found in AaaA in the 153 PanX database genomes were run through PROVEAN (Protein Variation Effect Analyzer). PROVEAN is an online tool which uses BLAST searches to predict the likelihood that a given amino acid substitution/deletion will have a deleterious effect on the protein, depending on how well a query sequence aligns with all other homologous or related sequences (Choi *et al.*, 2012).

Table 3.1 shows the PROVEAN scores calculated from the AaaA multiple sequence alignments generated from the PanX database. Only two substitutions are predicted to be deleterious, Y65H, which occurs in one isolate, and L170Q, which occurs in 15 (9.8%) of the isolates. All but two of these strains were isolated from human infection sites, as summarised in **Table 3.2**. Neither of the two potentially deleterious substitutions were present in 4 outlying strains which formed a distinct clade, shown in **Figure 3.4**.

Table 3.1: PROVEAN scores for all amino acid substitutions present in AaaA in the 153 PanX genomes.

Variant	PROVEAN score	Prediction (cut-off= -2.5)	Variant	PROVEAN score	Prediction (cut-off= -2.5)
G10S	0.113	Neutral	T407I	-0.931	Neutral
L18S	0.237	Neutral	G410A	0.256	Neutral
S54T	-0.800	Neutral	N441H	-0.290	Neutral
K55E	1.250	Neutral	E446Q	-0.037	Neutral
Y65H	-4.255	Deleterious (1 occurrence)	S464P	-0.399	Neutral
F94V	0.754	Neutral	S488P	-0.380	Neutral
R128H	-1.132	Neutral	S506N	-0.267	Neutral
G132D	0.251	Neutral	H516R	0.174	Neutral
A167T	-0.465	Neutral	V571I	-0.067	Neutral
L170Q	-3.936	Deleterious (15 occurrences)	D576N	-1.000	Neutral
K183R	-0.548	Neutral	D580E	-0.339	Neutral
S194A	-0.223	Neutral	M582I	0.160	Neutral
N195S	-0.125	Neutral	S586N	-0.267	Neutral
F217L	0.311	Neutral	V592I	-0.080	Neutral
F217S	1.228	Neutral	V594G	-0.262	Neutral
A228T	-0.821	Neutral	A595T	-0.083	Neutral
N267D	-1.978	Neutral	M597L	-0.000	Neutral
A280T	0.349	Neutral	M597I	0.033	Neutral
T288N	0.686	Neutral	G599D	0.250	Neutral
N289G	-2.069	Neutral	A607V	-0.433	Neutral
L291F	0.048	Neutral	L609I	0.008	Neutral
E297D	-1.008	Neutral	L609F	-0.563	Neutral
Q314L	3.197	Neutral	A611T	-0.117	Neutral
N359S	0.105	Neutral	A611V	0.300	Neutral
G363A	-0.656	Neutral	I615V	0.033	Neutral
E368K	-0.611	Neutral	V621I	0.008	Neutral
N386D	0.239	Neutral	G638S	-0.433	Neutral
G393S	-0.156	Neutral	N644S	-0.267	Neutral
H395Y	-0.333	Neutral			

Table 3.2: Details of the 15 isolates from the 153* P. aeruginosa *genomes on the PanX database with a potentially deleterious L170Q substitution in AaaA. *Compiled by tracing NCBI reference sequences given by PanX back to NCBI entries. Abbreviation: ad-cp, autotransporter_domain-containing_protein.*

NCBI Reference (genome_locus tag)	Strain name	Source
Nz_cp020603_e613_rs01925-1- ad-cp	E6130952	Sputum isolate from patient with respiratory failure
Nz_cp028162-c8257_rs01910-1-ad-cp	Mrsn12280	Wound isolate
Nz_cp027166-csb94_rs06455-1-ad-cp	Ar0357	FDA/CDC antimicrobial resistant isolate bank isolate
Nc_017549-ncgm2_rs29055-1-ad-cp	Ncgm2.s1	Hospital urinary tract infection outbreak isolate
Nz_ap014646-ncgm1984_rs01675-1-ad-cp	Ncgm 1984	Urinary catheter isolate
Nz_cp032761-pa268_rs02130-1-ad-cp	268	Patient with left ventricular assist device
Nz_cp033843-egy29_rs10970-1-ad-cp	Fdaargos_501	FDA-ARGOS database, 'ucc isolate'
Nz_cp025056-cwi20_rs01725-1-ad-cp	Pb367	Tracheal aspirate isolate
Nz_ap014622-ncgm1900_rs01675-1-ad-cp	Ncgm 1900	Urinary catheter isolate
Nz_cp027171-csb91_rs23735-1-ad-cp	Ar0354	FDA/CDC antimicrobial resistant isolate bank isolate
Nz_cp033833-egy26_rs17660-1-ad-cp	Fdaargos_571	FDA-ARGOS database, 'ucc isolate'
Nz_cp025055-cwi21_rs01725-1-ad-cp	Pb350	Sputum isolate
Nz_cp029605-ctt40_rs01990-1-ad-cp	24pae112	Blood isolate from patient with sepsis
Nz_cp028917-c5b40_rs01690-1-ad-cp	Jb2	Soil isolate
Nc_023019-u769_rs01670-1-ad-cp	Mtb-1	T-hch contaminated soil isolate

AlignmentViewer (Reguant *et al.*, 2018) was used to visualise the distribution of amino acid mismatches of AaaA in the database, as shown in **Figure 3.5**. Interestingly, it was noted first that 13 of the 15 L170Q mutants all had the exact same protein sequence for AaaA. As illustrated in Figure 4, Y65H is in the linker region between the signal peptide and passenger domain, and L170Q is present in the passenger domain. Neither of these residues are predicted to be important in the active site (Luckett *et al.*, 2012). However, L170Q does sit in the passenger domain between predicted active site residues E147-149 and D176. The chemical differences of the two amino acids,

leucine being hydrophobic, and glutamine being hydrophilic, as well as their proximity to the active site, may produce a significant conformational effect (illustrated later in **Figure 3.8**), which warrants investigation. Considering this, overall, these data still support the findings of Bianconi *et al*. (2018), that AaaA is genetically and structurally highly conserved, a conservation that suggests an indispensable role in the fitness of *P. aeruginosa*.

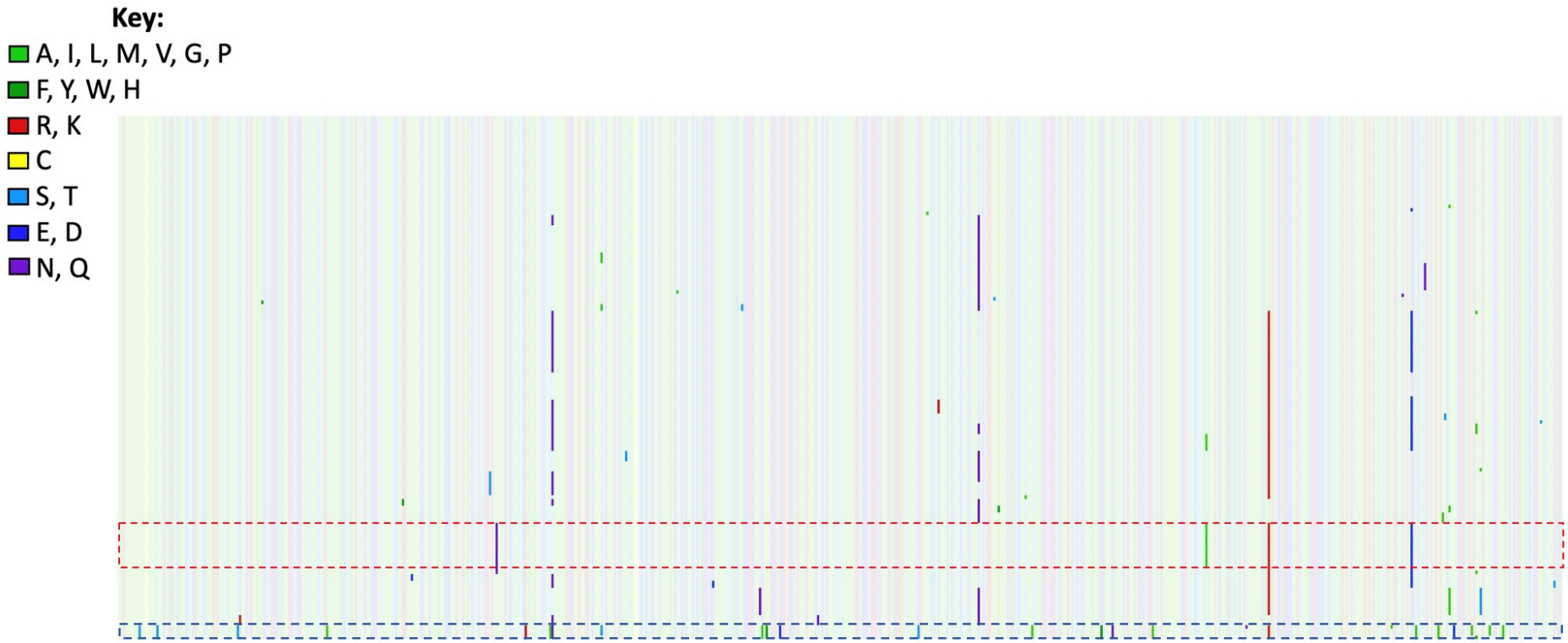


Figure 3.5: Alignment of AaaA amino acid sequence in 153* P. aeruginosa *isolates in the PanX genome database. *Strains are stacked along the vertical axis, with the amino acid sequences reading left to right. Mismatches to the consensus are highlighted with bold colouring. The 4 sequences which form a distinct clade are highlighted by a blue dashed box and the 13 homologous sequences containing the L170Q mutation are highlighted by a red dashed box. (Colour scheme- hydrophobic: light green, aromatic: dark green, N/Q: purple, S/T: light blue, negative charge: dark blue, positive charge: red, C: yellow). Created using AlignmentViewer and annotated using Microsoft PowerPoint (2020).*

3.5 There are some common and consistent deviations from PA0328

Noting the occurrence of clusters of homologous AaaA sequences which deviated from that of strain PAO1, the full Pseudomonas.com database of genomes was probed using DIAMOND BLASTP to search stringently for missense variations, as described in 2.13.1.1. Briefly, all complete sequences of AaaA (coverage 100%, identity 95%) were searched for amino acid mismatches and a cladogram was created containing all unique sequences. As shown in **Figure 3.6**, there are 50 unique sequences within the 3748 genomes containing an intact sequence for AaaA, out of the 4669 in the Pseudomonas.com database. Of these, the sequence for AaaA in PAO1 (PA0328) occurs in just 48 genomes (1.2% of all sequences). This is markedly lower than the occurrence of other sequences, such as BUH64_RS12720, which occurs in 505 of the genomes, over 10 times more frequently, or the sequence from PA14, which occurs in 379 genomes. The sequence for the 13 homologues with the L170Q variation, found in the PanX genome database (**Figure 3.5**), was also identified as NT18_RS24260, which occurred in 235 genomes.

Most notably, there are no amino acid mismatches in the five residues known to contribute to the active site or AaaA surface expression in the entire population (**Figure 3.9**). This highlights the extent of conservation of AaaA structure and implies its role as essential to the cell.

The second most notable finding is that there are 6 variations which were far more common among the genomes in the database than the PA0328 sequence. These are K55E, N195S, Q314L, H516R, N386D and D580E. Indeed, two of these mismatches, K55E and Q314L, occur in 98.6% of the genomes searched, thus the K55 and Q314 isoforms are almost unique to PAO1, suggesting that E55 and L314 should be considered the true ‘wild type’ isoforms.

The PROVEAN prediction software was again used to predict if any mismatches identified in the list of 50 unique sequences were likely to be

deleterious (**Table 3**). None of the common mismatches from PA0328 listed above, also illustrated in **Figure 2.4**, were predicted to be deleterious. A single AaaA sequence with an additional potentially deleterious mismatch, A251V was identified, with the sequence being otherwise identical to BUF29_RS22415, which occurs in ~13% of the genomes. Interestingly, when searched in the larger database, the L170Q mutation occurred in a total of 266 genomes (7%), indicating it is relatively prevalent (**Table 4**).

The six most common substitutions and three potentially deleterious substitutions are both highlighted in **Figure 3.7**, which shows their location in the primary amino acid sequence, and in **Figure 3.8** which shows their location in the 3D structure of AaaA.

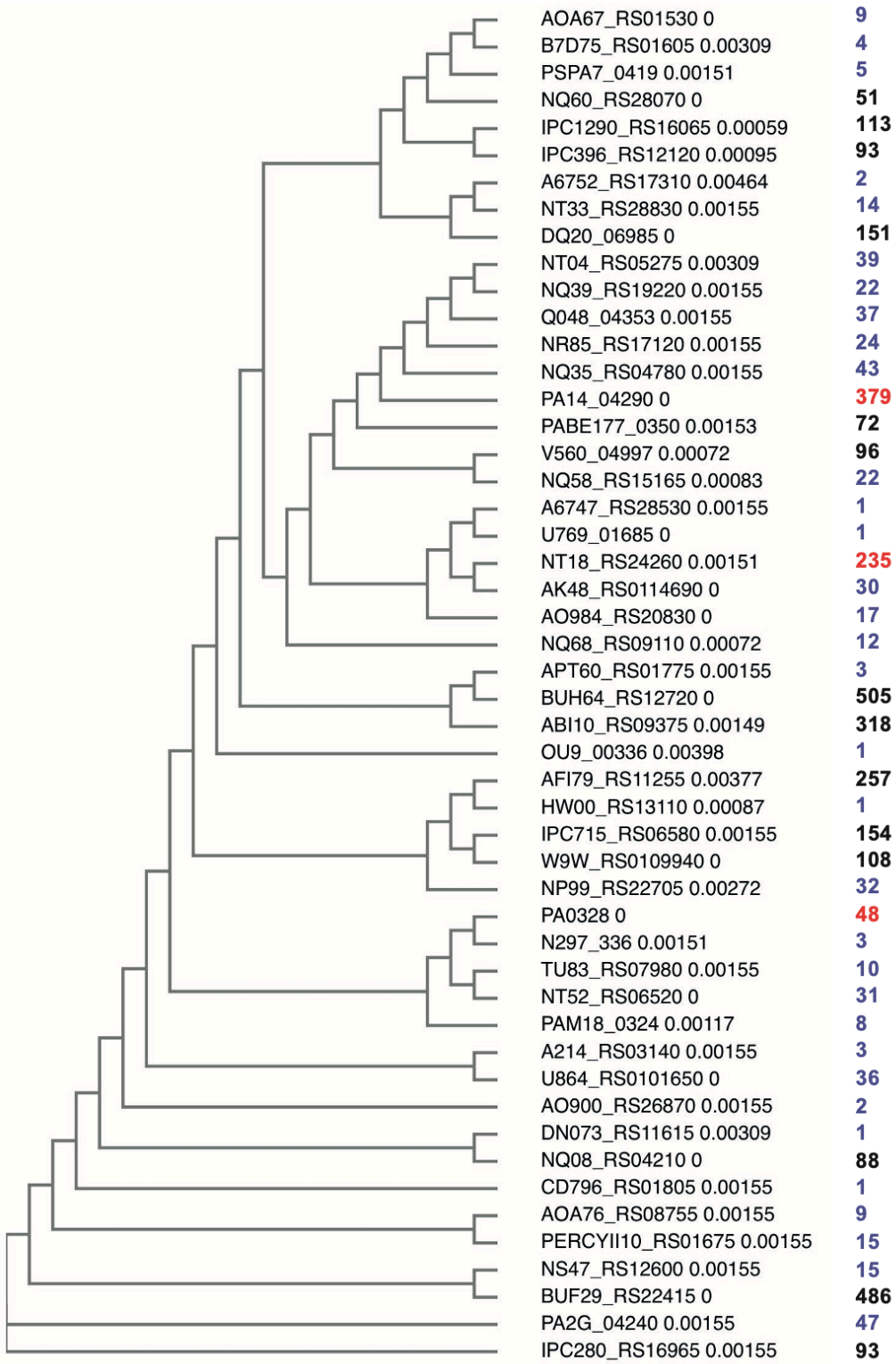


Figure 3.6: Cladogram showing representatives of all the unique complete sequences for AaaA in* P. aeruginosa *in the Pseudomonas.com database. *Numbers immediately following the sequence ID indicate the evolutionary distance between the sequence and its diverging node. Numbers aligned to the right of the figure indicate the frequency of occurrence of each sequence, out of 3748 genomes. Those which occur more and less commonly than PA0328 are shown in black, and blue respectively. The sequence for AaaA found in PAO1 and PA14 are shown in red, as well as N18_RS24260, which formed the distinct clade in* ***Figure 3.4****.*

Table 3.3: PROVEAN scores for all amino acid substitutions identified in AaaA in the 3748 genomes containing a complete AaaA sequence in the Pseudomonas.com database*. PA0328 was used as the reference gene, and predictions regarding deleterious substations were made using 228 fasta sequences and respective E-values, in 30 clusters. Variations that are either high in occurrence, or potentially deleterious are highlighted pink and yellow, respectively.*

Variant	PROVEAN score	Prediction (cutoff= -2.5)	Variant	PROVEAN score	Prediction (cutoff= -2.5)
M1V	-1.401	Neutral	T407I	-0.931	Neutral
G10S	0.113	Neutral	G410A	0.256	Neutral
L18S	0.237	Neutral	N441H	-0.290	Neutral
S54T	-0.800	Neutral	E446Q	-0.037	Neutral
K55E	1.250	Neutral (high occurrence)	S464P	-0.399	Neutral
Y65H	-4.255	Deleterious (15 occurrences)	S488P	-0.380	Neutral
A75T	0.850	Neutral	R505Q	-0.381	Neutral
F94V	0.754	Neutral	S506N	-0.267	Neutral
G98S	-2.219	Neutral	H516R	0.174	Neutral (high occurrence)
R128H	-1.132	Neutral	A558P	-0.247	Neutral
G132D	0.251	Neutral	V571I	-0.067	Neutral
A167T	-0.465	Neutral	D576N	-1.000	Neutral
L170Q	-3.936	Deleterious (266 occurrences)	D580E	-0.339	Neutral (moderate occurrence)
K183R	-0.548	Neutral	D581A	-0.433	Neutral
S194A	-0.223	Neutral	M582I	0.160	Neutral
N195S	-0.125	Neutral (high occurrence)	T585S	-0.550	Neutral
F217L	0.311	Neutral	S586N	-0.267	Neutral
F217S	1.228	Neutral	V592I	-0.080	Neutral
A228T	-0.821	Neutral	V594G	-0.262	Neutral
D243N	0.644	Neutral	A595T	-0.083	Neutral
A251V	-3.727	Deleterious (1 occurrence)	M597L	-0.000	Neutral
N267D	-1.978	Neutral	M597I	0.033	Neutral
A280T	0.349	Neutral	G599D	0.250	Neutral
T288N	0.686	Neutral	K602R	-0.489	Neutral
N289G	-2.069	Neutral	A607V	-0.433	Neutral
L291F	0.048	Neutral	L609I	0.008	Neutral
E297D	-1.008	Neutral	L609F	-0.563	Neutral
Q314L	3.197	Neutral (high occurrence)	A611T	-0.117	Neutral
N359S	0.105	Neutral	A611V	0.300	Neutral
G363A	-0.656	Neutral	I615V	0.033	Neutral
E368K	-0.611	Neutral	V621I	0.008	Neutral
N386D	0.239	Neutral (moderate occurrence)	G638S	-0.433	Neutral
G393S	-0.156	Neutral	N644S	-0.267	Neutral
H395Y	-0.333	Neutral			

Table 3.4: The prevalence of each of the 3 substitutions identified in the aaaA sequence which are predicted to be deleterious.

Mutation	Strains present in	Prevalence in database
Y65H		Total= 15 (0.4%)
	NS47_RS12600	15
L170Q		Total= 266 (7.1%)
	AK48_RS0114690	30
	U769_01685	1
	NT18_RS24260	235
	A6747_RS28530	1
A251V		Total= 1 (0.02%)
	CD796_RS01805	1

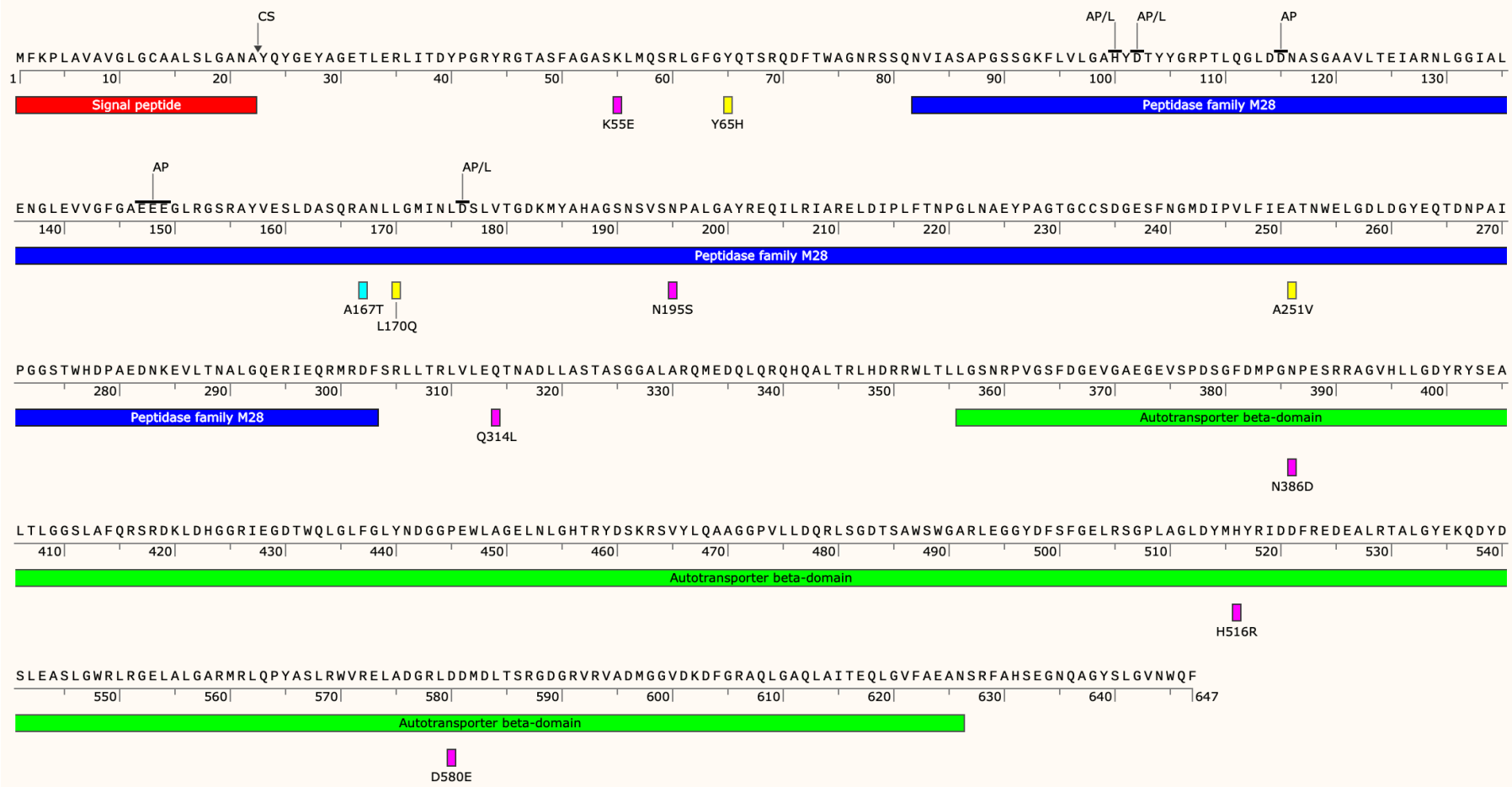


Figure 3.7: Map of AaaA primary structure with predicted active site residues and identified amino acid variations. *Predicted signal peptide, M28 peptidase and autotransporter β-domain are highlighted in red, blue, and green, respectively along with the predicted signal peptide cleavage site (CS). Residues essential for aminopeptidase activity or localisation are labelled AP/L. Locations of either potentially deleterious, or common substitutions are labelled in pink and yellow, respectively. A167T, present in CW2-BI1 (see 5.3.2.5), is marked in cyan.*

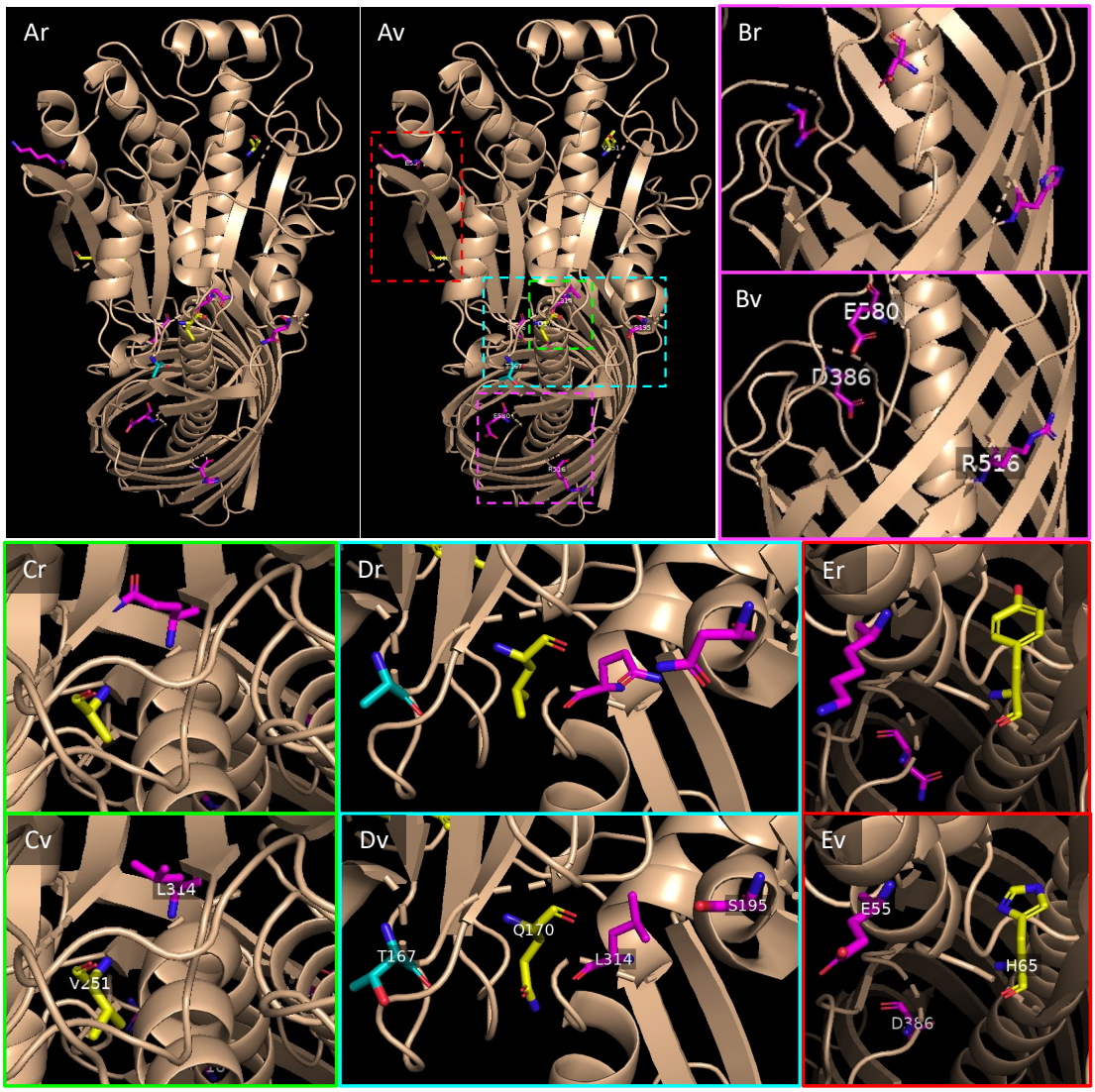


Figure 3.8: Amino acid variations in AaaA mapped onto the 3D model. Panel A-E *shows the 3D model with the highlighted amino acids as in the PAO1 reference (**r**) as its variant (**v**) from various angles, as indicated by the coloured dashed boxes in panel Av (**B** pink, **C** green, **D** cyan, **E** red). The organic chemical (stick) structure of amino acids which vary are shown and coloured according to **Figure 3.7** (pink predicted to be neutral, yellow predicted to be deleterious), with oxygen shown in red and nitrogen in dark blue. The AaaA 3D model was taken from the* AlphaFold Protein Structure Database (2022)*. Colouring and mutagenesis simulation was done using PyMOL.*

The fasta file containing the 50 unique sequences was also used to create a multiple sequence alignment, with P0328 as reference, using Snapgene, and viewed using AlignmentViewer (**Figure 3.9**). The multiple sequence alignment was also used to build a consensus model for AaaA (**Figure 3.10**) based on the most common variations among the population. However, due to computational limitations, this did not account for the

frequency of occurrence for each isoform in the Pseudomonas.com database, with each of the 50 isoforms being weighted equally. If this were accounted for, it is possible that the consensus would contain a few more variations from the PA0328 sequence.

Next, this consensus model was fed into CiTasser to generate a 3D model (**Figure 3.11**) which was aligned onto the model for AaaA (PA0328) in **Figure 3.2**. The consensus model has greater uncertainty, and more flexible loop regions, but the general structure of the M28 domain is retained. When the two M28 domains are aligned, as in **Figure 3.1**, the alignment of the active sites is good, with an RMSD of 0.338 Å, indicating that structure and activity should be similar to PA0328 in the isoforms of AaaA generally present in the real-world *P. aeruginosa* population.

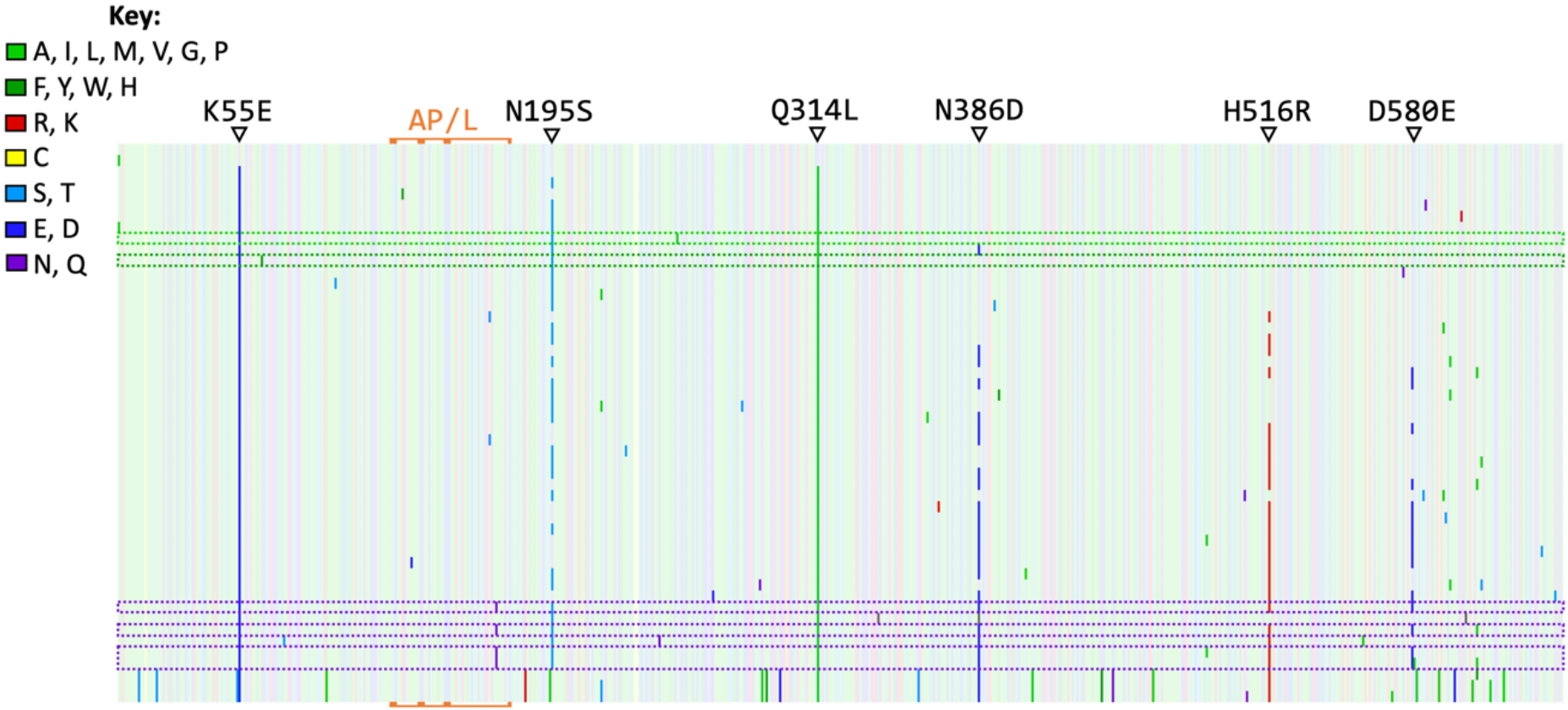


Figure 3.9: Alignment of the 50 unique AaaA amino acid sequences identified in the 3748 genomes containing a complete AaaA sequence in the Pseudomonas.com database. *Mismatches to the reference sequence (PA0328 from PAO1, top sequence) are highlighted with bold colouring. Sequences containing a mismatch predicted using PROVEAN to be deleterious are highlighted with dashed boxes as follows: Y65H- dark green, L170Q- purple, A251V- light green. The six most common mismatches are also labelled, K55E, N195S, Q314L, N386D, H516R and D580E. The residues (H100, D102, D115, E147 and E148) essential for aminopeptidase activity and/or AaaA surface localisation (AP/L) are highlighted in orange, showing there are no mismatches in any of these residues. Amino acid key colour scheme as in* ***Figure 3.5****. Created using AlignmentViewer and annotated using Microsoft PowerPoint (2020).*

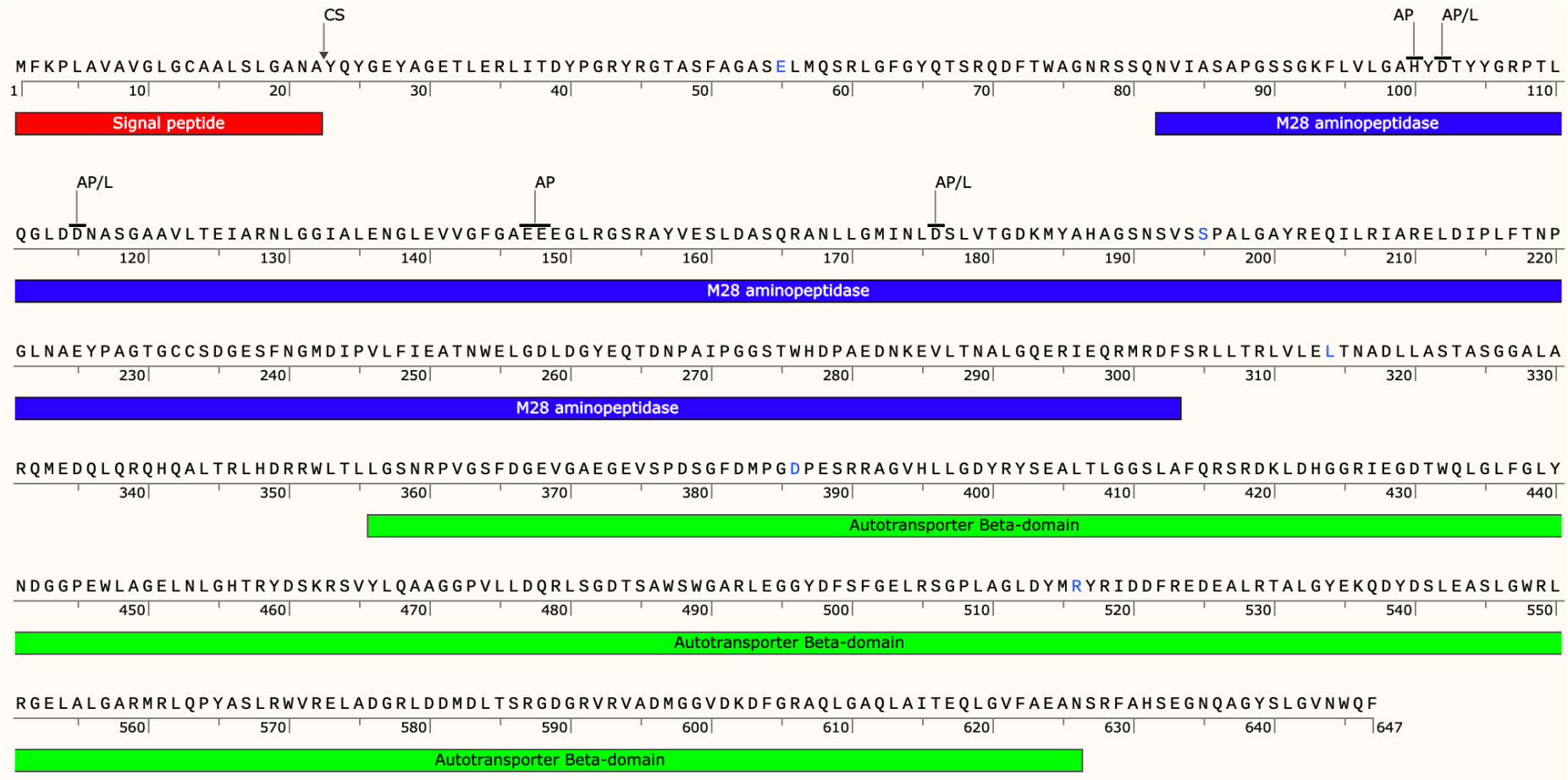


Figure 3.10: Consensus_aaaA based on common variations. *Predicted signal peptide, M28 peptidase and autotransporter β-domain are highlighted in red, blue, and green, respectively along with the predicted signal peptide cleavage site (CS). Residues essential for aminopeptidase activity or localisation are labelled AP/L and amino acid variations from PAO1 are in blue text.*

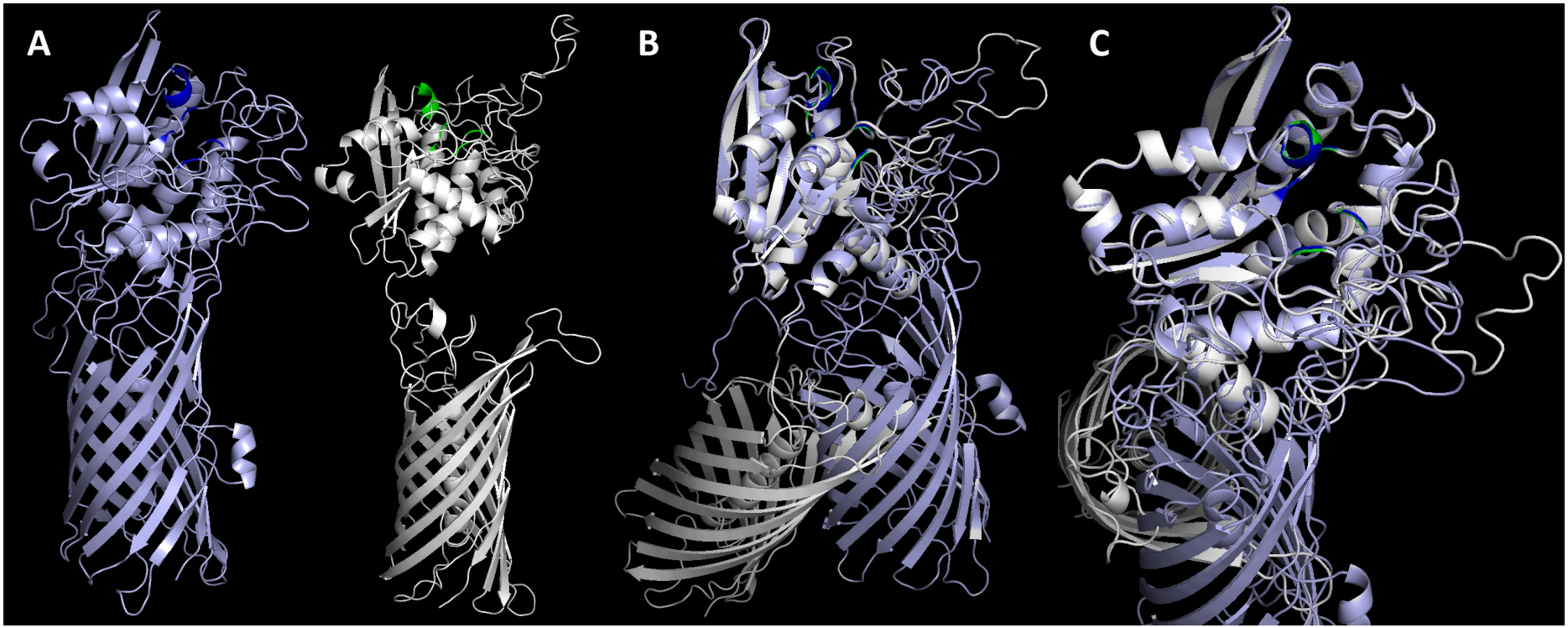


Figure 3.11: Predicted 3D alignment of P0328 and Consensus_aaaA model. Panel A *shows the C-iTasser model 2 for consensus_aaaA (light blue, active site residues dark blue) C score= -1.87 and model 1 for AaaA PA0328 (white, active site residues green) C score= 0.37. General structures are retained between the two molecules, including the β-strand motif in the M28 domain.* ***Panel B*** *and* ***Panel C*** *show a side view and top view, respectively, when the two models are aligned by their M28 domain. Alignment of key structural features is good, with near exact alignment between active site residues (RMSD = 0.338 Å). Models were generated by CiTasser, aligned and coloured using PyMOL.*

3.6 Using site-directed mutagenesis to probe the effect of variations on AaaA activity

Work has previously been carried out to sequence *aaaA* in various clinical isolates and characterise the activity and expression profile of these, as will be discussed in 3.7.1. However, in order to better understand the individual effect on AaaA activity of each of the common and potentially deleterious amino acid substitutions found in this chapter, single amino-acid substituted mutants were constructed using SDM. Primers were designed for inverse PCRs using pBlueScript*::aaaA* as a template to create mutants carrying the Y65H, L170Q, A251V, K55E, N195S, Q314L, N386D, H516R and D580E variations as described in 2.5.2.

3.6.1 D580E may have a detrimental effect on AaaA activity

Despite multiple primer redesigns, many of the inverse PCR reactions could not be optimised to produce a clean amplicon. Detailed discussion of the cloning and optimisation completed for these SDM constructions can be found in the 1st year annual PhD review of Bethan Roberts who assisted with this work (Bethan Roberts and Kim Hardie, University of Nottingham, unpublished, 2022).

However, the pET21a*::aaaA-D580E* construct was successfully completed, and its effect on AaaA activity is shown in **Figure 3.12**. D580E had a moderate to mild impact on over-produced AaaA activity over 16 hrs. A simple linear regression (line) of AaaA, AaaA-D580 and AaaA-uninduced gives slopes of 0.123 (R^2= 0.9337, mean AUC= 21.32), 0.093 (R^2= 0.9864, mean AUC= 16.55) and 0.070 (R2= 0.9896, mean AUC= 12.63), respectively. These data indicate that in the conditions tested, overproduced AaaA-D580 has ~75-77% the activity of overexpressed wild-type AaaA.

The D580E substitution was present in 1031 out of the 3748 sequences included in bioinformatic analysis, meaning it occurs in 27.5% of all genomes analysed. This suggests it may play a minor role, possible in translocation, due to its location in the β-barrel of AaaA. Both aspartate and glutamate are negatively charged, hydrophilic amino acids, with the only structural difference being that aspartate has one less carbon in its backbone than glutamate. It is

possible that the extra length increases flexibility, though this is speculative. It was also predicted by PROVEAN analysis to be a neutral substitution. Further study of other common or potentially deleterious amino acid substitutions in AaaA could provide clues to the contribution of these residues to AaaA structure and function.

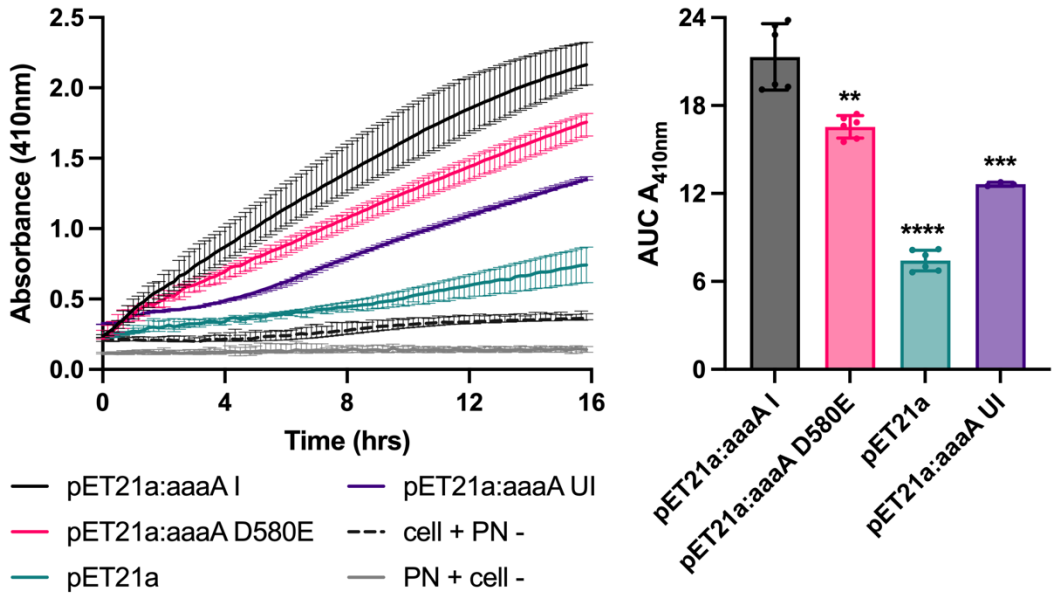


Figure 3.12: Whole cell AaaA activity assay revealed D580E variation had a minor deleterious effect upon AaaA activity. E. coli *carrying* pET21a *containing either wild-type* aaaA*,* aaaA-D580E*, or neither, was grown in selective LB to* OD_{600nm} *0.5-0.9 and transferred to 22°C for 5 hours following the addition of 0.3 mM IPTG (except for* pET21a:aaaA *UI- uninduced). Following washing, a final* OD_{600nm} *of 0.5 was added to the L-arginine-p-nitroanilide cleavage assay.* ***Panel A*** *shows absorbance at 410nm over 16 hrs (n=6, 2 biological x 3 technical, except for UI, where n=3, median ±95% CI).* ***Panel B*** *shows the mean (±SD) area under the curve (baseline Y=0) for each. Significant differences were determined by Brown-Forsythe and Welch's ANOVA, with Dunnett's T3 multiple comparisons comparing each group to induced wild-type (*pET21a*:aaaA I) (***p<0.0005). cell+ PN- control contains cells from the* pET21a*:aaaA I without L-arginine p-nitroanilide and PN+ cell- control contains L-arginine p-nitroanilide only (no cells). Data generated by Bethan Roberts.*

3.7 Discussion

In this chapter, bioinformatic tools were used to create a predictive 3D model of full length AaaA and probe its conservation across *P. aeruginosa* strains. Comparison of 3D models created using CiTasser and AlphaFold2 show similar structures, and good alignment between the M28 domains.

Analysis of over 3000 AaaA sequences from genomes in Pseudomonas.com database found that AaaA is highly conserved. This is

supported by the finding that there are no single amino acid substitutions in the five amino acids previously shown to be essential for AaaA activity and localisation, H100, D102, D115, E147 and E148 (Luckett *et al.*, 2012). A caveat to this is that the stringent criteria used here only searched for single polymorphisms and excluded indels and other larger mutations or genomes with poor coverage of the gene. It is therefore not possible to know whether these findings extend to the 921 excluded genomes, which did not contain an intact *aaaA* sequence, either due to indels or poor sequencing coverage.

The other notable finding here was that, compared to the PAO1 reference sequence, there were several very commonly occurring single amino acid variations. These include K55E, N195S, Q314L, N386D, H516R and D580E, as well as L170Q, which is predicted to be potentially deleterious and was present in ~7% of the genomes analysed. This work highlights the importance of acknowledging that the reference PAO1 sequence is often not perfectly representative of wider *P. aeruginosa* populations.

In this chapter, PROVEAN was used to predict whether amino acid substitutions identified in AaaA were likely to be deleterious. This method uses sequence homology of close and distantly related proteins to make these predictions. This, in essence, relies on the assumption that polymorphisms that are deleterious are selected against and removed over time, while advantageous or neutral polymorphisms are retained. This technique has been used previously to predict the effect of mutations in *Salmonella typhimurium* with 72% accuracy in 5 different proteins encoded by genes involved in growth on arabinose (Lind *et al.*, 2017). However, a broader assessment of the tool found that its effectiveness was variable and depended on various factors, including the species and genes in question. For example, it could make more useful predictions for *Chlamydomonas reinhardtii* genes than for the *Saccharomyces cerevisiae* genes (Sandell & Sharp, 2022). These predictions therefore require further data from complimentary bioinformatic tools or experimental evidence.

Previous work has examined *aaaA* sequences, and the effect of different clusters of mutations, in clinical *P. aeruginosa* isolates. However, prior to this

work, a systematic approach examining each individual mutation had not yet been taken. Therefore, experimental work was undertaken in this study to systematically determine the effect of these amino acid substitutions on AaaA activity. This study found that the D580E substitution (present in 27.5% of AaaA sequences analysed) caused a moderate reduction in AaaA activity when over-produced in *E. coli*. Further work to complete the SDMs of the other common variations and create a 'consensus' mutant as determined in **Figure 3.11** would aid in elucidating the potential cumulative effects of these common mutations and give a clearer understanding of the true 'wild-type' AaaA activity levels *in vivo*.

3.7.1 Previous work suggests the mutations in the *aaaA* sequence only partially impact AaaA activity

Paredes-Osses (2014) previously sequenced *aaaA* orthologues from six clinical isolates. A number of amino acid variations, also reported here, were found and their enzymatic activity was compared to that of wild type AaaA from PAO1 (**Figure 3.13**). Clinical strain (CL) 09, which carried among others, the D580E variation, appeared to have much lower AaaA activity in their study than PAO1. However, AaaA in CL09 and CL49 had the same mutation profile, indicating they shared a AaaA sequence, yet activity of AaaA in CL49 was close to PAO1 levels. Moreover, the variation profile of AaaA in CL09 was not as extensive as CL10, which had all the CL09 variations, plus N195S, A595T and the potentially deleterious L170Q, yet retained activity levels equivalent to PAO1.

Using a luciferase-transcriptional reporter assay, described further in 5.1.2.1, Paredes-Osses (2014) noted that CL10 appeared to have a greater level of *aaaA* promoter (P_{aaaA}) activation than either CL09 or PAO1, indicating *aaaA* is upregulated in CL10. This suggested that any defect in the protein may be counteracted by higher gene expression levels. As discussed in 5.1.2.2, P_{aaaA} activation is also increased in the *aaaA* knock-out mutant, supporting the idea of a compensatory mechanism. Another isolate, CL14, which has 3 of the same variations present in CL09, had a maximum activity level similar to the wild type, though the rate of activity was approximately halved, suggesting perhaps a

lower level of AaaA at the cell surface. Taken together, this suggests that variations in AaaA primary structure may be only one of many factors affecting AaaA activity levels, which could possibly depend more on *aaaA* expression and AaaA secretion, or at least on the accumulation of a specific set of variations.

A

Clinical Strain	E34D	K55E	S68N	L170Q	N195S	G197T	V311L	Q314L	N359S	N386D	H516R	M560I	D580E	A595T	Total
01	X	X	X			X	X	X	X	X	X				9
09		X						X		X	X		X		5
10		X		X	X			X		X	X		X	X	8
14		X			X			X							3
33		X						X			X	X			4
49		X						X		X	X		X		5

B

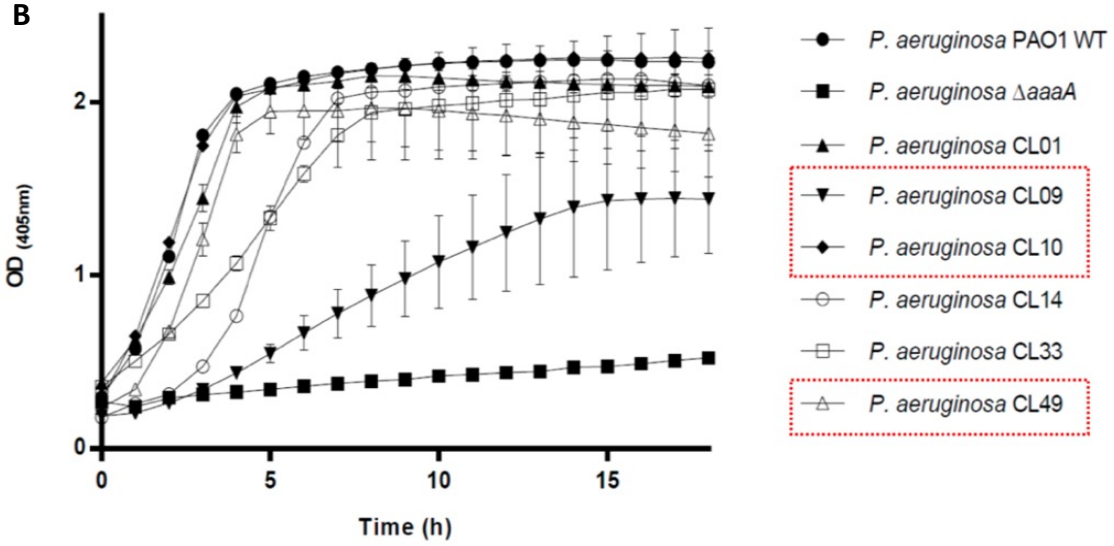


Figure 3.13: Variations in AaaA from clinical isolates have stochastic effects on activity. Panel A *is a table showing amino acid variations present in AaaA from six clinical isolates. Strains CL09, CL10 and CL49 all carry the D580E variation, and are highlighted.* ***Panel B*** *shows whole-cell L-arginine-p-nitroanilide cleavage (AaaA activity) assays exhibiting varying levels of AaaA activity between the isolates. Strains CL09, 10 and 49 are displayed as inverted black triangles, black diamonds, and white triangles, respectively, and are highlighted. Adapted from* Paredes-Osses (2014).

Spencer (2018) also screened 10 *P. aeruginosa* strains isolated from CF patients, shown in **Figure 3.14**, and found that while none carried the D580E substitution, all but one carried the K55E substitution and all but four carried

the Q314L and N195S substitutions. All isolates had lower AaaA activity than PAO1N, however CF484 and CF570 had higher P_{aaaA} activation, potentially also pointing toward compensation.

Strain	K55E	N195S	I270V	A228T	N240I	255-6 EL->GT	Q314L
CF129	X	X					X
CF098	X	X					X
CF983	X	X					X
CF971	X	X					X
CF326	X	X					X
CF393		X	X		X		
CF030	X			X		X	
CF892	X			X		X	
CF484	X	N195C	X		X		
CF570	X						X

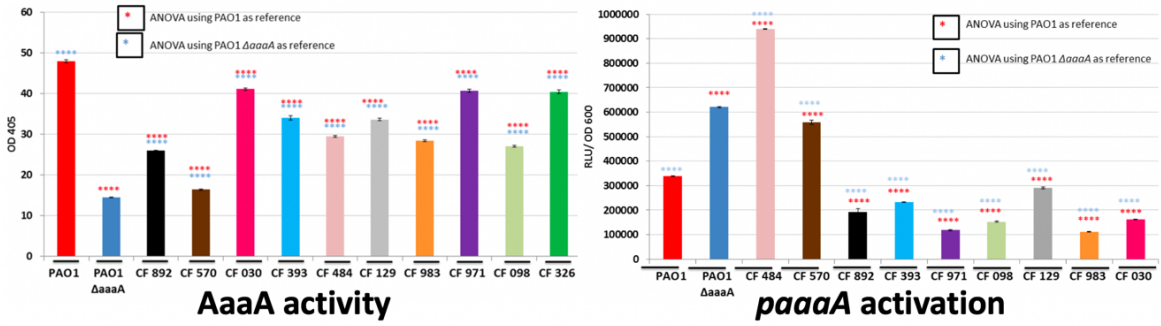


Figure 3.14: Variations in AaaA in CF isolates and their effect on AaaA activity and P_{aaaA} activation. Top *panel is a table showing amino acid variations present in AaaA from ten clinical isolates. Rows are coloured according to their respective bars in the L-arginine-p-nitroanilide (AaaA activity) assay and lux-P_{aaaA} transcriptional reporter assays* ***Below****. There is no clear link between any of the amino acid variations and either activity or P_{aaaA} activations. Adapted from* Spencer (2018)*.*

Like Paredes-Osses (2014), it was not possible for Spencer (2018) to draw a clear link between any one variation, or even set of variations, and a major change in AaaA activity or P_{aaaA} activation. This points to a complicated picture regarding AaaA, which cannot be explained just by amino acid variations, and more likely relates to transcriptional or post-translation regulation of AaaA, which will be explored in later chapters.

Chapter 4. Progress toward elucidating the AaaA protein structure

4.1 Previous work to purify AaaA

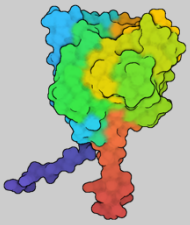
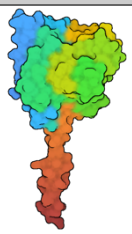
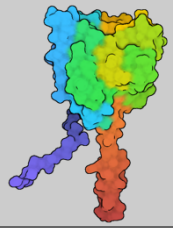
As described in 3.1, previous work has established the enzymatic function of AaaA as an arginine-specific aminopeptidase and provided evidence for the essentiality of certain residues in activity and localisation (Darch, 2009; Luckett *et al.*, 2012; Paredes-Osses, 2014). However, many questions remain. For example, AaaA is predicted to co-ordinate zinc for its metalloprotease activity, yet this has not been proven experimentally, and no experimentally determined structure is yet available. One aim of this thesis, and the subject of this chapter, is to address this by over-producing and purifying the M28 peptidase domain of AaaA for biophysical and structural studies including crystallography. Obtaining purified AaaA would also allow advancement of work by Paredes-Osses (2014) to identify AaaA inhibitors, which could be used as a novel anti-virulence strategy.

Because AaaA in its native form is tethered to the OM, truncated constructs were developed by Paredes-Osses (2014) and Spencer (2018) which exclude the β-barrel motif, and added a hexa-histidine tag at the C-terminus for metal-affinity chromatography-based purification. Three constructs have so far been created, as detailed in **Table 4.1**. Each was cloned into a pET21a (+) vector and transformed into *E. coli* BL21(DE3). Retention of the signal peptide in AaaAt1 and AaaAt3 means they should localise to the periplasm, via translocation through the Sec machinery. Removal of the signal peptide in AaaAt2 prevents translocation and thus localises it to the cytoplasm.

Solubility proved an issue, particularly for AaaAt1. Spencer (2018) had found that active AaaAt1 could be recovered from the resolubilised pellet fraction, as lysate yields were low, yet this could not be taken forward as it could not be recovered in the expected fractions by size exclusion chromatography, suggesting it was forming aggregates. The cytoplasm-localised AaaAt2 formed insoluble, inactive aggregates, and was toxic to the

cell. The latest AaaAt3 construct is an extended version of AaaAt1, containing 12 additional amino acids from the α-helical linker between the passenger domain and β-barrel of the protein, with the hope of improved solubility. Initial studies by Spencer (2018) suggested AaaAt3 may have promising activity and solubility, thus this is the construct focused on in this chapter for purification.

Table 4.1: Details of the existing AaaA truncated constructs created by Drs Paredes Osses and Spencer. *Predicted size values calculated by Spencer (2018) using ProteinCalculator. Model from* AlphaFold Protein Structure Database (2022) *and truncated using PyMOL.*

	Selected residues	Predicted cellular localisation	Predicted size (KDa)	Model	Notes
AaaAt1	1-337 (337)	periplasm	36.05		Active, largely insoluble
AaaAt2	22-349 (327)	cytoplasm	35.48		Toxic to cell and inactive
AaaAt3	1-349 (349)	periplasm	38.35		Soluble and active

4.2 Chapter aims

As a continuation of the ongoing effort to resolve the structure of AaaA, this chapter describes approaches to overproducing and purifying AaaAt3. This would allow for downstream probing of the structure and function of AaaA, including co-crystallography with inhibitors identified by Paredes-Osses (2014). Moreover, AaaA has potential as a vaccine target candidate, and the ability to overproduce immunogenic fragments of AaaA would aid the development of a conjugate vaccine against *P. aeruginosa* (Bianconi *et al.*, 2018).

4.3 Over-production and purification of AaaAt3

First, the induction conditions for overproduction of high yields of AaaAt3 were optimised. Next, different cell lysis methods were compared by examining the yield of enzymatically active, soluble AaaAt3 by SDS-PAGE analysis and AaaA-activity assays of each lysis fraction. Finally, AaaAt3 purification was optimised by testing different chromatography methods, including by metal affinity, anionic exchange and size exclusion.

4.3.1 Induction of AaaAt3 with 0.1-0.3 mM IPTG at 22˚C for 5hrs gives a good yield

To determine optimal induction conditions, comparative IPTG inductions were performed, where either induction time, temperature or IPTG concentration was varied (**Figure 4.1**). it was determined that 5-8 hr induction at 22˚C was optimal, with a concentration of IPTG between 0.1-0.3 mM.

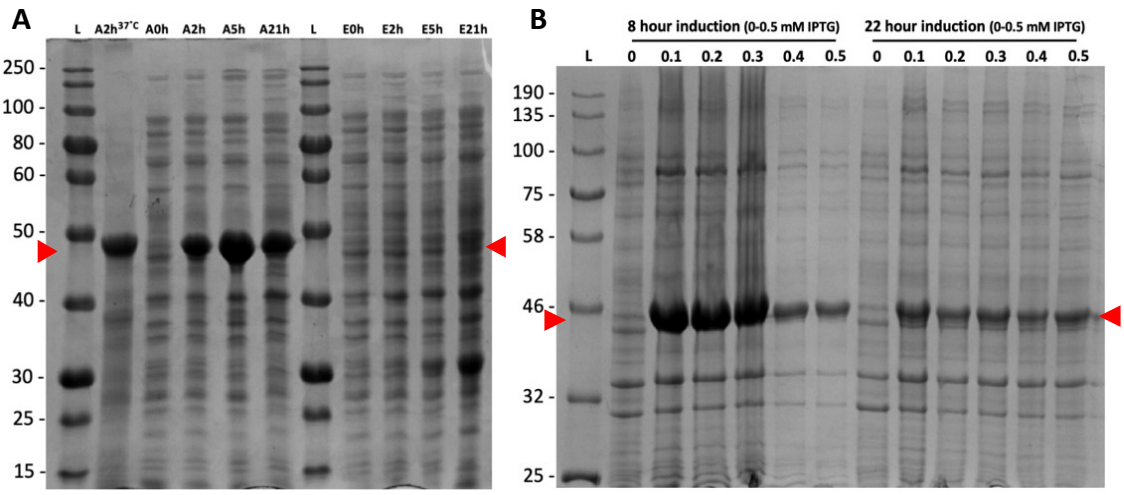


Figure 4.1: AaaAt3 induction optimisation. Panel A *shows an SDS-PAGE gel of whole cell pellets pre- and post-induction, with varying induction times. 100 mL cultures of* E. coli *BL21 (DE3)* pET21a::aaaAt3 *(A) and empty* pET21a *vector (E) were grown at 37 ˚C, 200 RPM until* OD_{600nm} *~0.9, induced with 0.3 M IPTG, and moved to 22 ˚C, 200 RPM. 1 mL samples normalised to* OD_{600nm} *0.7 were taken at 0hrs, 2hrs, 5hrs and 21hrs post induction, washed with PBS and resuspended in 100 µL 1x SDS loading buffer. 7 µL was added to each well of a 12% SDS-PAGE gel. Protein at ~46 kDa (indicated by red arrows) roughly corresponds to the expected size of the AaaAt3 mutant protein.* ***Panel B*** *shows SDS-PAGE gel of whole cell pellets from 0-0.5 mM IPTG induction of AaaAt3. 50 mL cultures of* E. coli *BL21 (DE3)* pET21a::aaaAt3 *were grown until* OD_{600} *~1.0 and induced with 0-0.5 mM IPTG. At both 8 hrs and 22 hrs post induction, 1 mL samples were taken for analysis as described for Panel A.*

4.3.2 Lysis method selection

Spheroplast lysis is a method used to isolate only periplasmic proteins, by breaking the cell OM, but not the IM. Sonication and French press lysis lyse the whole cell by the application of ultrasonic frequencies and mechanical pressure, respectively. The three methods (described in 2.4.3) were compared for lysing cells over-producing AaaAt3, as shown in **Figure 4.2a**. Although it gave a good yield and theoretically a cleaner product, spheroplast lysis is difficult to scale up for a cell pellet from a large (> 1L) initial culture volume and also did not yield enzymatically active AaaAt3 (**Figure 4.2b**). French press lysis was initially taken forward as this gave a good yield of active, soluble AaaAt3, although sonication may be considered equally useful in future.

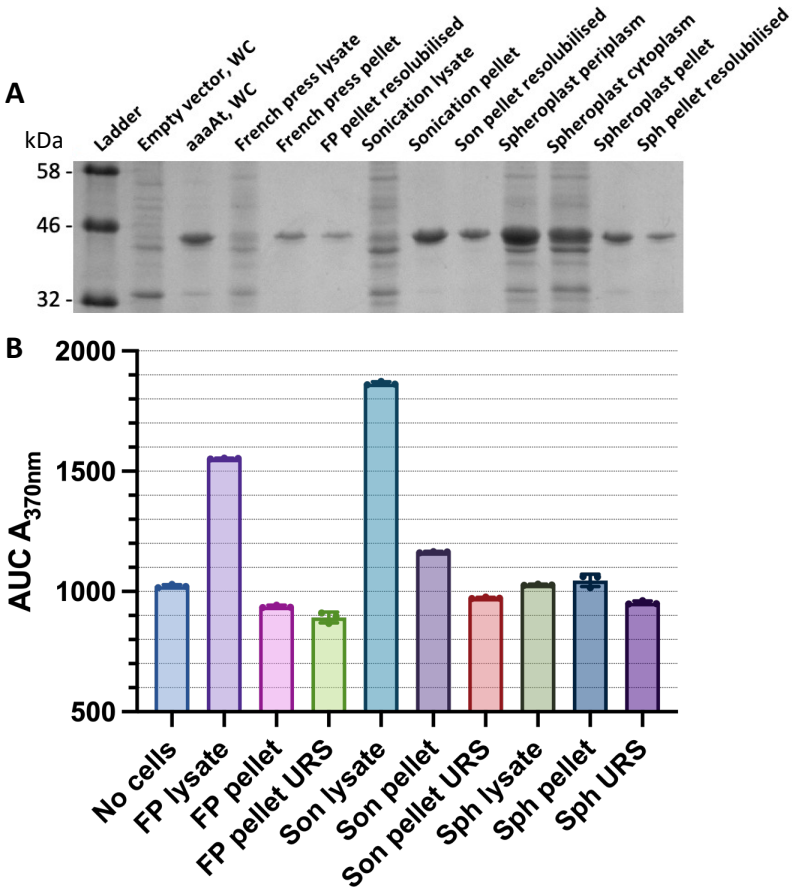


Figure 4.2: Yield of active AaaAt3 in soluble and insoluble fractions of over-produced AaaAt3, from three lysis methods. Panel A *shows an SDS-PAGE gel of either insoluble or 8 M urea-resolubilised (URS) pellets (as in Spencer, 2018), or trichloroacetic acid-precipitated lysates. Each resuspended in 300 µL 1 x SDS-PAGE sample buffer and 7 µL loaded onto a 12% SDS-PAGE gel.* ***Panel B*** *shows the median (± range) area under the curve (AUC) of L-arginine-p-nitroanilide cleavage (AaaAt3 activity) in each lysis fraction (n=3). No cells control contains L-arginine p-nitroanilide only.*

4.3.3 A portion of AaaAt3 is soluble, but it cannot be efficiently purified by metal ion affinity chromatography

4.3.3.1 Purification by Ni-HisTrap column or Co-TALON resin

AaaAt3 did not bind well to either a nickel-charged His-Trap column (**Figure 4.3**) or cobalt-charged TALON resin (**Figure 4.4**), where it was detected primarily in the column flow-through and wash. Anti-His Immunoblotting, (**Figure 4.5**) as described in 2.3.2, confirmed that the His-tag is still present on the soluble AaaAt3, meaning the lack of efficient metal affinity binding was due to the His-tag being inaccessible, rather than removed. There was also AaaAt3 in both French press pellets, indicating incomplete lysis. Where there are large amounts of non-specific, background absorbance in the 370-410nm range, the cleavage of L-arginine-p-nitroanilide is shown over time, so that AaaA activity can be inferred from the shape and gradient of each curve, rather than the area under the curve (AUC), with a steeper curve indicating higher activity levels.

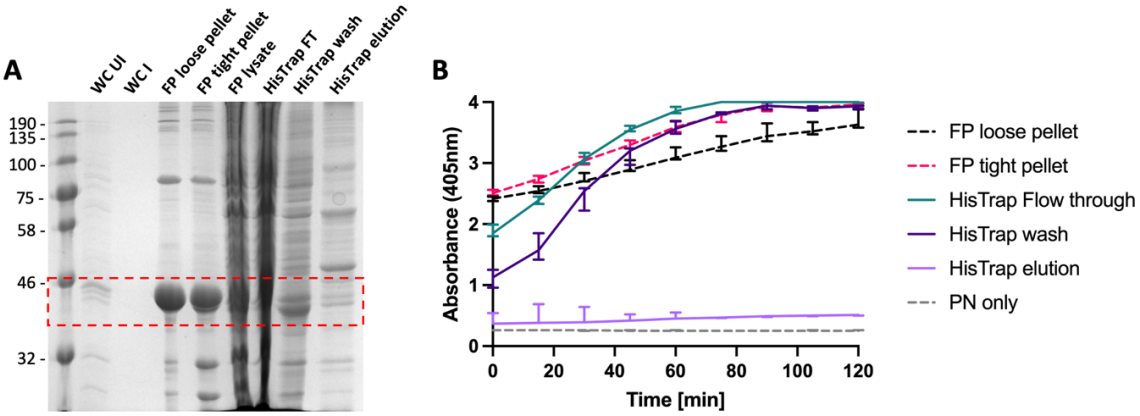


Figure 4.3: AaaAt3 does not appear to bind to the Ni-affinity chromatography column. Panel A *shows SDS-PAGE gel of AaaAt3 at each stage of over-production and nickel affinity (Ni-HisTrap) chromatography purification. Fractions shown are whole cell (WC) induced (I) and uninduced (UI), French press (FP) lysis fractions and Ni-HisTrap flow through (FT), wash and elution fractions. Proteins of the predicted size of AaaAt3 (~43 kDa) are highlighted (red box, dashed lines). For whole cell fractions, 1 mL of culture, normalised to* OD_{600} *0.5 was resuspended in 300 µL 1 x SDS-PAGE sample buffer. Each of the loose and tight pellets and filtered lysate were volume-normalised by resuspending in 25 mL and 160 µL of each, and HisTrap FT, wash and elution were each added to 40 µL 5 x SDS-PAGE sample buffer. 5 µL of the whole cell samples and pellets, and 10 µL of the lysate and HisTrap fractions were added to a 12% SDS-PAGE gel.* ***Panel B*** *shows a L-arginine-p-nitroanilide cleavage (AaaAt3 activity) as read by absorbance at 405 nm, over 2 hrs in lysis and Ni-HisTrap fractions (median ± 95% CI, n=3). PN only- L-arginine p-nitroanilide (no cells) control.*

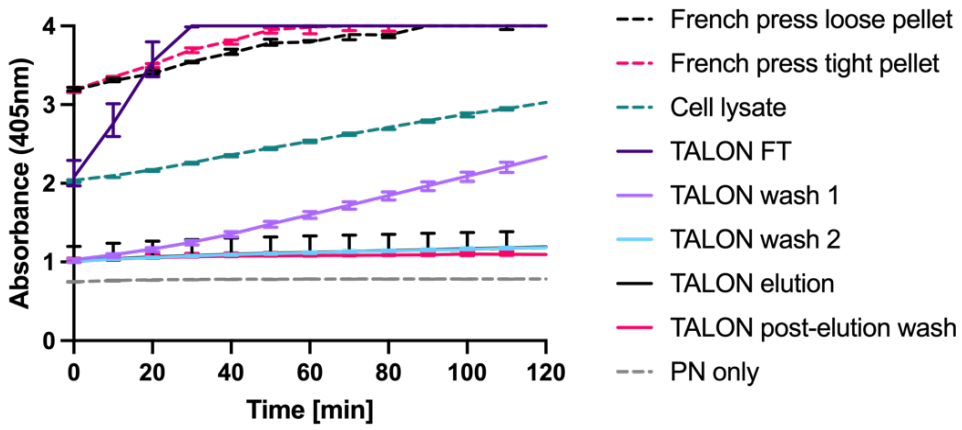


Figure 4.4: L-arginine-p-nitroanilide cleavage (AaaA activity) assay of AaaAt3 purification by Co-TALON resin chromatography shows AaaAt3 does not bind well to Co-TALON resin*. Samples were prepared as in* ***Figure 4.3****. Graph shows median (95% CI) absorbance (405nm), n=3. FT- flow through, PN only- L-arginine p-nitroanilide (no cells) control.*

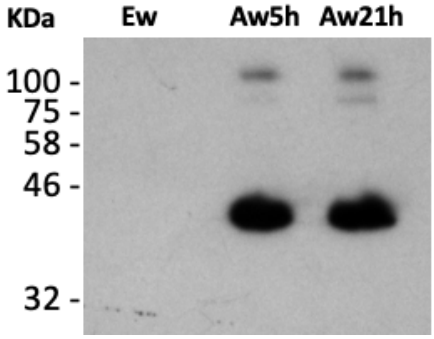


Figure 4.5: Immunoblot to detect AaaA using α-his, HRP-conjugated antibody against whole cell inductions. *100 mL cultures of* E. coli *BL21 (DE3)* pET21a::aaaAt3 *(A) and empty* pET21a *vector were grown until* OD_{600nm} *~0.9 and induced with 0.3 mM IPTG. Aw5h and Aw21h refer to 1 mL samples taken at 5hrs and 21hrs post induction respectively, which were normalised to* OD_{600nm} *0.7 and resuspended in 100 µL SDS loading buffer and diluted further 1:100 in sample buffer, and 5 µL were added to a 12% SDS-PAGE gel before immunoblotting. Ew refers to the pre-induced empty vector.*

4.3.3.2 Extension of the 6-His-tag to 8-His and 10-His does not drastically improve AaaAt3 metal ion affinity

To improve the accessibility of the hexa-histidine-tag, two extended versions were constructed, AaaAt3H8 (8x His-tag) and AaaAt3H10 (10x His-tag) according to the methods in 2.5.2, with a T_m for the SDM PCR of 67°C. Neither construct bound much better to a HisTrap column than the hexa-histidine-tag-carrying construct. Any protein which did bind eluted either at very low imidazole concentrations, or alongside many contaminants, as can be seen in **Figure 4.6** and **Figure 4.7**.

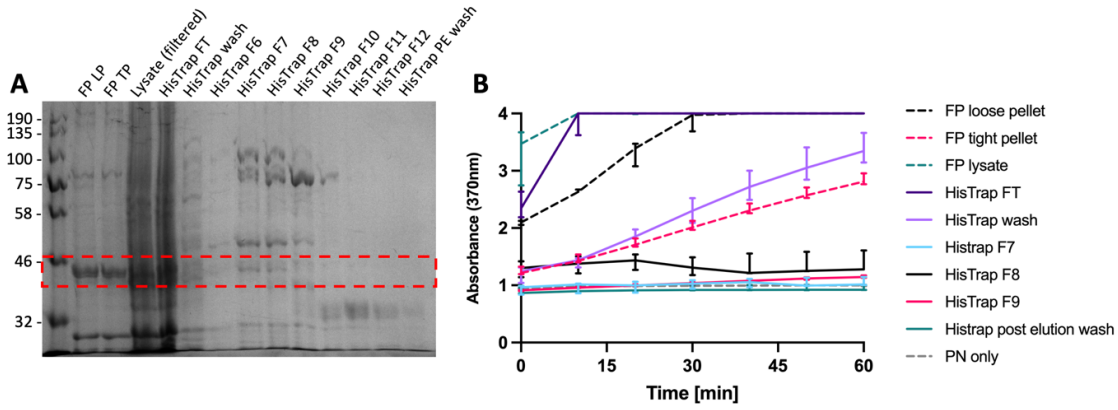


Figure 4.6: AaaAt3 with 8xHis only slightly improves binding to the HisTrap column. Panel A *shows a 12% SDS-PAGE gel of 7 µL each of French press (FP) lysis fractions- loose (LP), tight pellet (TP) and lysate- followed by HisTrap flow through (FT), wash and elution fractions (F6-12). Proteins of the predicted size of AaaAt3 (~43 kDa) are highlighted (red box, dashed lines).* ***Panel B*** *shows the L-arginine-p-nitroanilide cleavage (AaaA activity) assay of the same fractions. PN only- L-arginine p-nitroanilide (no cells).*

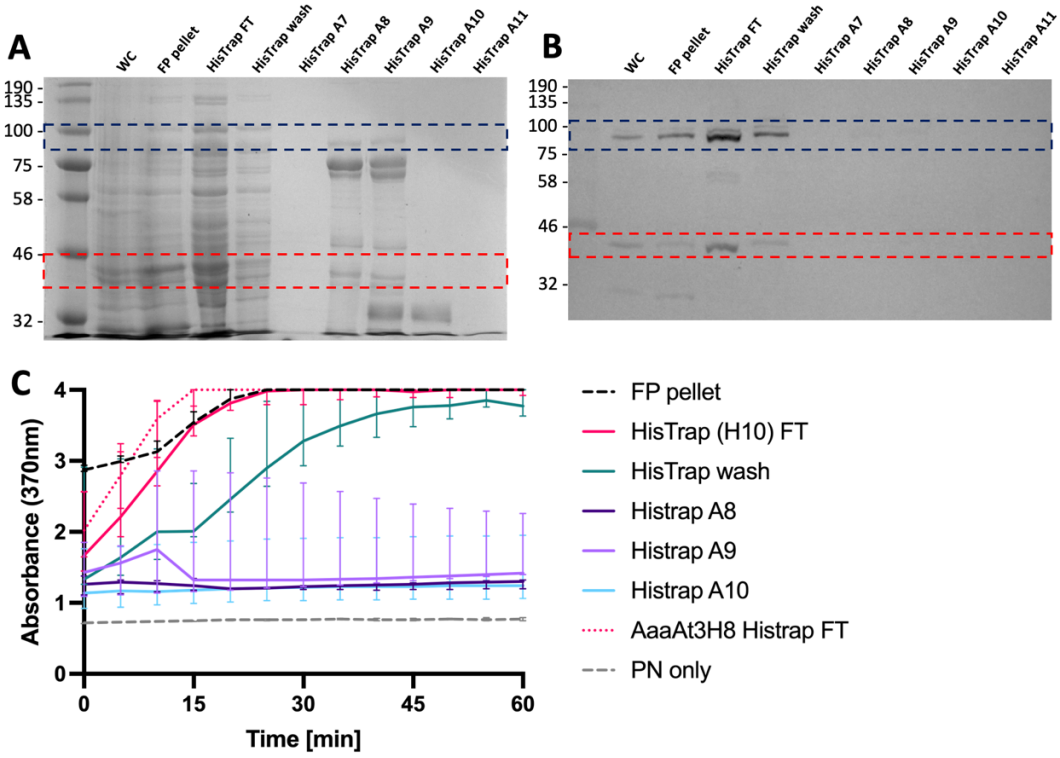


Figure 4.7: AaaAt3 with an 10x His tag also does not bind better to a HisTrap Ni column. Panel A *and* ***B*** *show a 12% SDS-PAGE gel and α-AaaA immunoblot, respectively, with whole cell (WC, OD_{600nm} 0.1 cells in 100 µL SDS-PAGE sample buffer) and French press (FP) pellet, HisTrap flow through (FT), wash and elution fractions. Proteins of the predicted size of AaaAt3 (~45 kDa) and proteins ~90kDa, a possible AaaAt3 dimer, are highlighted by red and blue dashed boxes, respectively. All other samples besides WC were in 160 µL plus 40 µL 5x SDS-PAGE sample buffer and 7 µL was loaded per well.* ***Panel C*** *shows an L-arginine-p-nitroanilide cleavage (AaaAt3 activity) assay alongside a sample of AaaAt3H8 HisTrap flow through (FT) for comparison. PN- L-arginine p-nitroanilide (no cells) control.*

To eliminate the possibility that lack of fraction activity was due to loss of Zn^{2+} ions during HisTrap binding, samples from **Figure 4.7** were re-pooled and split in two. One half was incubated for 10 mins with 10 mM EDTA to strip any nickel ions, then dialysed overnight with 25 mM Tris, 300 mM NaCl, pH 7.5, 1 mM $ZnCl_2$, and the activity assay was repeated alongside the untreated samples, which had been kept at 4 ˚C. However, treatment with $ZnCl_2$ did not restore enzymatic activity (**S-Figure 0.1**).

4.3.3.3 AaaAt3 may form an SDS-resistant dimer

Curiously, a protein was seen consistently at ~80-100 kDa in both anti-His (**Figure 4.5**) and anti-AaaA immunoblots (**Figure 4.7b**), as well as prominently in many SDS-PAGE gels. It is unclear if this is an SDS-resistant aggregate of AaaAt3 as it is the approximate size of a possible dimer.

It may alternatively be a contaminating protein with some level of metal-ion affinity which also cross-reacts with both anti-His (monoclonal) and anti-AaaA (polyclonal) antibodies. In anticipation of this possibility, both constructs were transformed into the *E. coli* NiCo21 (DE3) strain for further overexpression experiments. This strain is designed to eliminate co-contaminants such as GlmS (~66.9 kDa), and make removal of other potential co-contaminates (SlyD, ArnA and Can) possible by chitin affinity chromatography (New England Biolabs, 2022). However, as AaaAt3 was not isolated in elution fractions, but in the flow through, it is possible that if it is a co-contaminant rather than a dimer, it has no metal ion affinity and mass spectroscopy would likely be needed to identify it.

4.3.4 Anionic exchange can be used to partially purify AaaAt3 from crude lysate

As the isoelectric point (pI) of AaaAt3 is predicted to be 5.07, anionic exchange of crude lysate was next tested, as described in 2.4.4.2. Due to the localisation of AaaAt3 to the periplasm, a spheroplast lysis method was again used to maximise lysis efficiency and reduce contaminants. As shown in **Figure 4.2**, spheroplast lysis did not initially yield active AaaAt3, however modification

of the protocol to 10 mg/mL lysozyme with 0.5 mM EDTA rather than 1mM EDTA improved activity, as can be seen in **Figure 4.8**.

As mentioned previously, spheroplast lysis efficiency is lowered when scaled up to the large culture volumes need for protein purifications. However, this is offset by the reduced contamination from the cytoplasmic fraction, meaning that this crude lysate is suitable for partial purification by crude anionic exchange. As shown in **Figure 4.8a**, AaaAt3 can clearly be seen following anionic exchange in fractions 16-24. These fractions (fraction 18 is representative) also have high levels of L-arginine-p-nitroanilide cleavage activity, indicative of enzymatically active AaaAt3 (**Figure 4.8b**). The remaining contaminants following this step could then be removed by size exclusion chromatography (SEC).

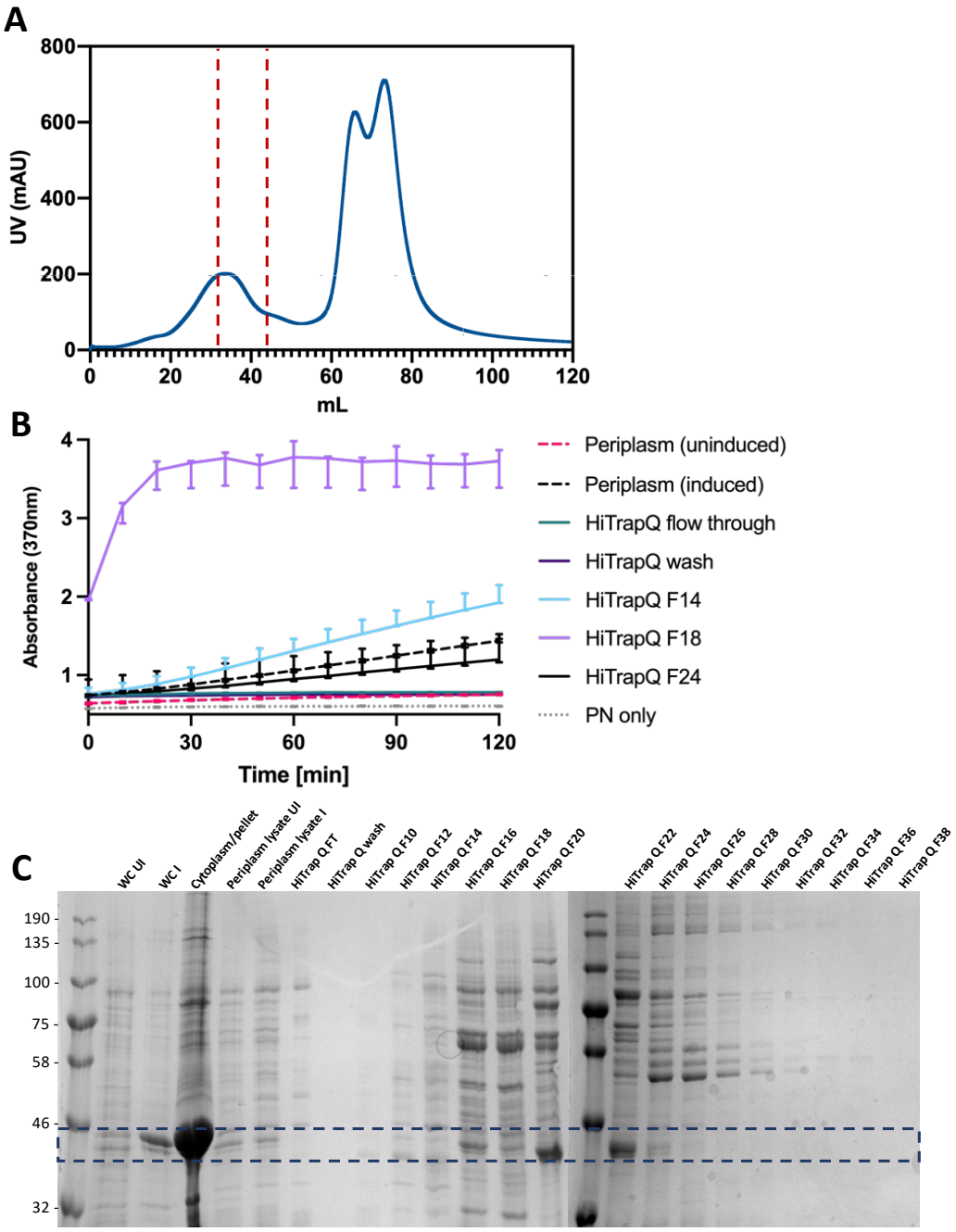


Figure 4.8: AaaAt3 can be partially purified from crude lysate by anionic exchange. Panel A *shows a UV absorbance chromatogram of the outflow from anionic exchange chromatography. Fractions 18-24, containing AaaAt3, are highlighted by dotted red lines; each fraction is 2 mL.* ***Panel B*** *show a L-arginine-p-nitroanilide cleavage assay on spheroplast lysis products (periplasmic fraction) and anionic exchange chromatography fractions over 2 hrs (median ± 95% CI, n=3). PN only- L-arginine p-nitroanilide (no cells) control.* ***Panel C*** *shows an SDS-PAGE gels of whole cell, spheroplast lysis products (periplasmic and cytoplasm/insoluble fractions) and anionic exchange fractions. The protein at ~43 kDa, the predictive size of AaaAt3, is highlighted by a dotted blue outline. Whole cell (WC) sample of induced (I) and uninduced (UI) culture were normalised to* OD_{600nm} *0.1 and resuspended in 300 µL 1 x SDS sample buffer. 160 µL of the resuspended pellet, the periplasmic lysate, the HiTrap Q flow through (FT), wash and elution fractions (F10-F38) were added to 40 µL 5 x SDS sample buffer, with 5 µL loaded onto a 12% SDS-PAGE gel.*

4.3.5 Size exclusion chromatography can further purify active, soluble AaaAt3

Semi-pure, active AaaAt3, can be obtained using size exclusion chromatography as described in 2.4.4.3. Curiously, though a protein ~43 kDa, the expected size of AaaAt3, is can be seen by SDS-PAGE in fractions A6-A11 (**Figure 4.9a**), only fraction A8, and to a lesser extent, A7 and A9 exhibited AaaA enzymatic activity (**Figure 4.9c**), and it is unclear why. It is possible that the 'activity' attributed to AaaAt3 is contributed to by additional non-specific activity from one of the co-contaminants in the fractions. In all cases in this chapter, there have been other proteins present, alongside AaaAt3, in fractions which exhibited enzymatic activity. Moreover, there were moderate levels of activity observed in empty pET21a lysates, which cannot be attributed to AaaAt3 (**S-Figure 0.1**).

More curiously, as can be seen by the chromatogram in **Figure 4.9b**, protein with AaaA activity eluted in two peaks, A8 and A12-A13, which roughly correspond to the expected elution fractions of a monomer and a dimer, respectively. Though no protein is visible for fractions A12 and A13 on the SDS-PAGE gel in **Figure 4.9a**, AaaA enzymatic activity is higher than that observed in fractions A7 and A9 (**Figure 4.9c**). It is possible that, due to their separation by size, fractions A12-13 contain the potential AaaAt3 dimer postulated in 4.3.3.3. An immunoblot on these fractions would potentially amplify the signal and help to identify what, if any, proteins are contributing to this activity.

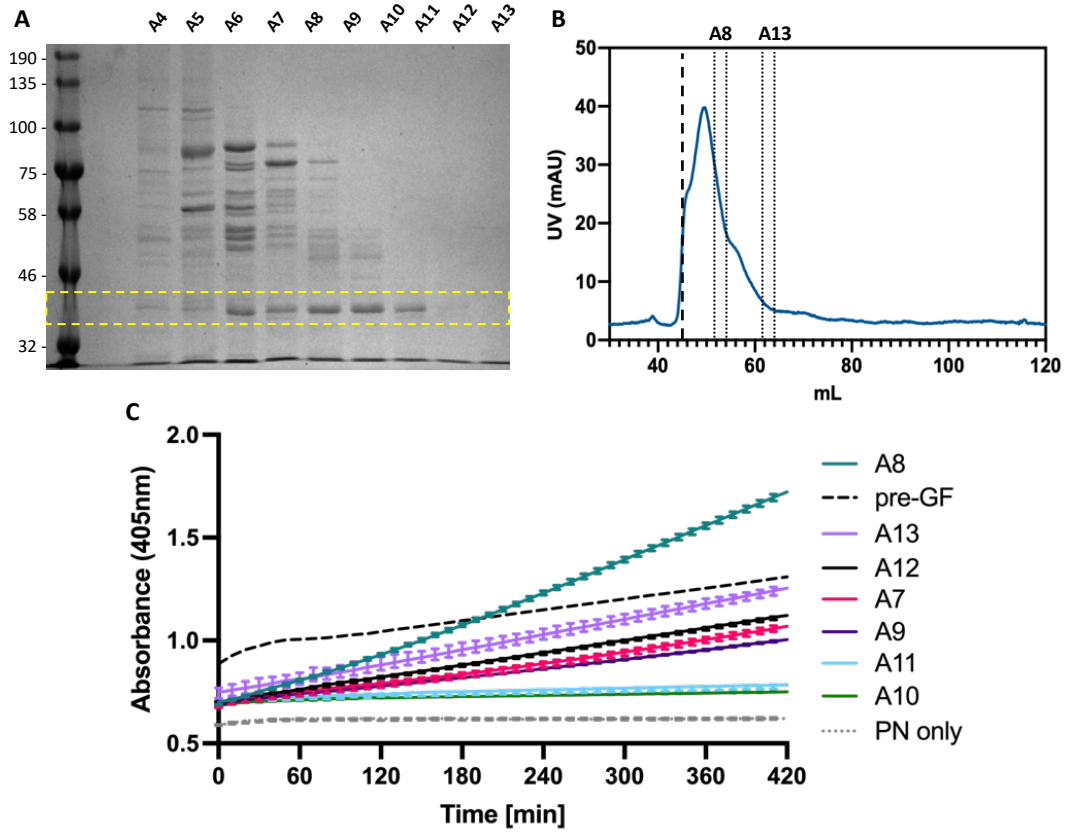


Figure 4.9: Anionic exchange fractions can be further purified by size exclusion chromatography. Panel A *shows SDS-PAGE gels of size exclusion chromatography fractions. HiTrap Q fractions 18-24 from* ***Figure 4.8*** *were pooled, concentrated to 2 mL using a spin concentrator and loaded onto an S75 size exclusion column. AaaAt3 (predicted size ~43 kDa, highlighted by a dashed yellow box) comes off the column in fractions A6-A11, with the greatest purity in A8-11.* ***Panel B*** *is a UV absorbance chromatogram of outflow from the size exclusion column. End of the void is marked by a dashed line (~45 mL). There is a large peak at ~50 mL and smaller peak at ~55 mL, which are indicative of the proteins equivalent to an AaaAt3 dimer and monomer, in size, respectively. Fractions A8 and A13 are marked with dotted lines.* ***Panel C*** *shows L-arginine-p-nitroanilide (PN) cleavage assay of SEC fractions, as well as a sample pre-SEC (pre-GF), over 7 hrs (median ± 95%, n= 3).*

The resolution by size exclusion here is not ideal, and could be improved, for example by decreasing the loading volume. Additionally, a Superdex™ 75 column, which resolves 3-70 kDa proteins, was used here due to the expected size of AaaAt3 being ~43 kDa. However, given the potential that AaaAt3 exists in a complex with itself or another protein, it may be useful to perform future size exclusion instead using a Superdex™ 200 column, which can resolve proteins of 10-600 kDa.

4.4 Discussion

This chapter has detailed work to purify a soluble version of AaaA, AaaAt3, with some success using crude anionic exchange followed by size exclusion chromatography. However, AaaA activity was reduced following purification. Interestingly, this work has provided, for the first time, some evidence that AaaA, in its AaaAt3 form at least, could exist as a dimer. Though the AaaAt3 construct shows some promise, these data suggest that activity and solubility may still be compromised by the truncation, with large portions of AaaAt3 remaining in insoluble fractions. The technology available for purification of membrane proteins is ever evolving, and it is therefore worth exploring purification of the full-length membrane protein alongside continued work on truncated versions.

4.4.1 Alternative protein tags for affinity chromatography of AaaA

Given that the histidine tag does not work well with the AaaAt3 construct, one option would be to clone in an alternative tag, such as a maltose-binding-protein (MBP) tag which is larger and more soluble (Lebendiker & Danieli, 2011) or a streptavidin (Strep II) tag which has better column binding specificity (Schmidt & Skerra, 2007). Placement of the tag on the N-terminus, between the signal peptide and passenger domain, may also improve accessibility of the tag.

4.4.2 Optimisation of strain background

Further optimisation of strain background may also be useful. This study used the NiCo construct to reduce common metal-affinity co-contaminants, but further strain backgrounds, such as *E. coli* C41 or LEMO could be explored to improve on AaaAt3 production and folding. C41 and LEMO *E. coli* are both specialised for difficult-to-express proteins, particularly toxic or membrane-bound proteins which are translocated via Sec (Schlegel *et al.*, 2012). Over-expression may also be more favourable when expressed in a *P. cold* vector, specialised for overproduction of toxic proteins in cold conditions (12-16˚C) to minimise toxic aggregates (Qing *et al.*, 2004).

4.4.3 Possibilities for over-production of a near-native AaaA construct

Alternatively, a new construct could be designed containing a TEV (or other) protease cleavage site in the α-linker between the β-barrel and passenger domain where it is likely most exposed, between residues T315-L319. A hexa-histidine or potentially other tag could also be cloned onto the N-terminus, between Q24-Y25, which should be retained following signal peptide cleavage, yet still accessible to bind to resin, as illustrated in **Figure 4.10**.

The full-length protein would therefore be produced and processed as 'near-native', and inclusion of the β-barrel will allow for any folding that occurs during translocation to the extracellular face of the OM, as per the 'hairpin' model of autotransporter secretion. Following growth and lysis, the membrane can then be collected by ultracentrifugation, homogenised, and incubated with TEV protease and NTA-nickel beads overnight as illustrated in **Figure 4.10**. The passenger domain of the protein, having been released from the membrane by cleavage of the linker anchoring it to the β-barrel, can then be eluted from the beads and taken forward for further purification and crystallography. One potential issue with the approach is that over-expression of a membrane protein could be toxic to the cell, and thus induction conditions would need to be carefully optimised, and a fermenter may be required to grow high volumes of cells.

The pJM220 vector developed by Meisner et al. (2016) has been shown in this study (6.2.6) to over-produce AaaA in *P. aeruginosa* in a tightly inducible manner, and has proven amenable to HiFi cloning in this study, as discussed in detail in 0. Thus, it may be possible to clone this construct directly into *P. aeruginosa,* for maximally robust AaaA folding, although again, with lower expression levels than for example pET21a in BL21 *E. coli*.

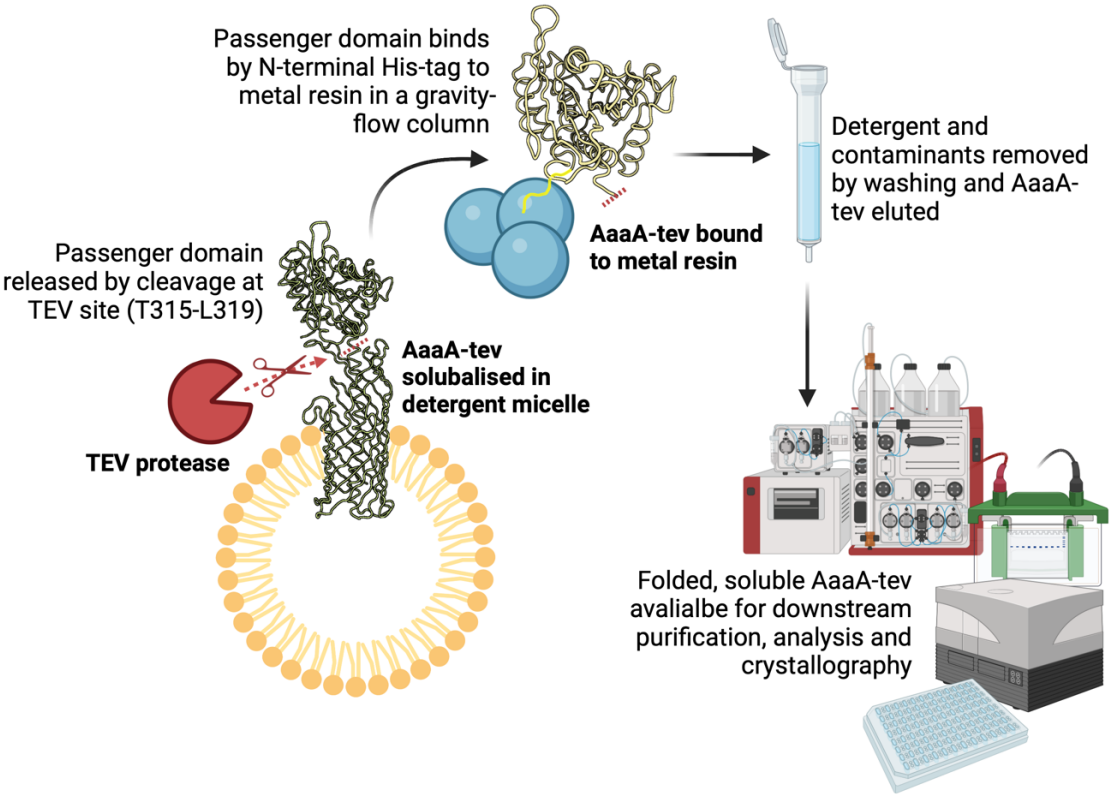


Figure 4.10: Illustration of proposed AaaA-tev construct and purification methodology. *AaaA-tev would contain a TEV protease site between the beta barrel and passenger domain of AaaA. The full-length construct would be over-expressed, the membrane isolated and solubilised in detergent micelles where TEV would cleave at its site to release the soluble passenger domain for purification and downstream analysis. Created using BioRender.com, model for AaaA from* AlphaFold Protein Structure Database (2022) *and modified using PyMOL.*

4.4.3.1 Use of next-generation techniques for full-length membrane-bound crystallography

Another possibility is to attempt crystallography of the full length, membrane bound AaaA protein, using either SMA (styrene maleic acid) or lipidic cubic phase crystallography techniques. The SMA copolymer can be used instead of detergent to extract small discs of lipid bilayer encapsulated by the polymer, known as SMA lipid particles (SMALP). This allows the target protein of interest to be maintained within a lipid bilayer, yet with greater solubility and crystallisation properties than when maintained in traditional detergent micelles (Broadbent *et al.*, 2022). Lipidic cubic phase crystallography instead maintains the protein in a quasi-solid matrix containing the required proportions of water, protein and lipid to create a three-dimensional lipid array which supports the growth of more ordered crystals than those formed in micelles (Caffrey, 2015; Landau & Rosenbusch, 1996). If possible, the benefits of crystallising the full

length, native protein over a truncated mutant are clear, and this is something that could be explored further with collaborators.

4.4.4 Conclusions

This chapter has detailed work to over-produce and purify AaaA for further downstream study. Although IPTG-induced over-production of the protein seems strong, simple metal ion affinity-based purification does not seem possible. Other options, including crude anionic exchange followed by further purifications steps may be effective. However, the solubility and activity of the AaaAt3 construct is questionable, and the construct may need re-engineering to improve this. For example, the construct could be re-engineered with a different tag, such as an-MBP tag, which may improve solubility. Over-production in a different strain background should also be considered, as this may improve folding and translocation, reducing the likelihood of aggregate formation. Finally, an AaaA-tev construct, a version of the full-length membrane-bound protein, modified to contain a cleavage site between the β-barrel and passenger domain, could be designed. Following induction and isolation of the membrane fraction, the passenger domain could be cleaved and purified. This, as well as attempts to purify the native membrane-bound protein could be considered, along with the many possibilities these approaches would offer, and challenges they entail.

Chapter 5. Probing *aaaA* expression and regulation in a synthetic chronic wound model

5.1 The state of current knowledge regarding *aaaA* regulation

Previous researchers have approached the question of *aaaA* regulation using three methods, which will be discussed, along with their findings, in the following section:

1) Bioinformatic interrogation of the *aaaA* promoter to identify putative regulatory binding sites.
2) Monitoring *aaaA* promoter activation, using the luciferase-based transcriptional reporter miniCTX-P_{aaaA}-*lux* in different growth media, or backgrounds lacking putative *aaaA* regulators.
3) Measuring AaaA activity, using the L-arginine-p-nitroanilide cleavage assay also in backgrounds lacking putative regulators.

5.1.1 Previous interrogation of the *aaaA* promoter found potential *aaaA* regulators

Previous work by Spencer (2018) to interrogate the *aaaA* promoter has revealed a number of putative binding sites, which are summarised in **Figure 5.1**. The Shine-Dalgarno (ShD) sequence and transcriptional start site predicted by Spencer (2018) are also annotated along with identification of a putative RpoN-binding site, suggesting RpoN is the sigma factor controlling *aaaA* transcription. These locations are supported by transcriptomic data which found the *aaaA* 5' untranslated region (UTR) was 45 nucleotides long (Wurtzel *et al.*, 2012). Further details on each putative regulator binding site are provided in **Table 5.1**, where those experimentally shown to impact P_{aaaA} activation (discussed further in 5.1.2) are also identified.

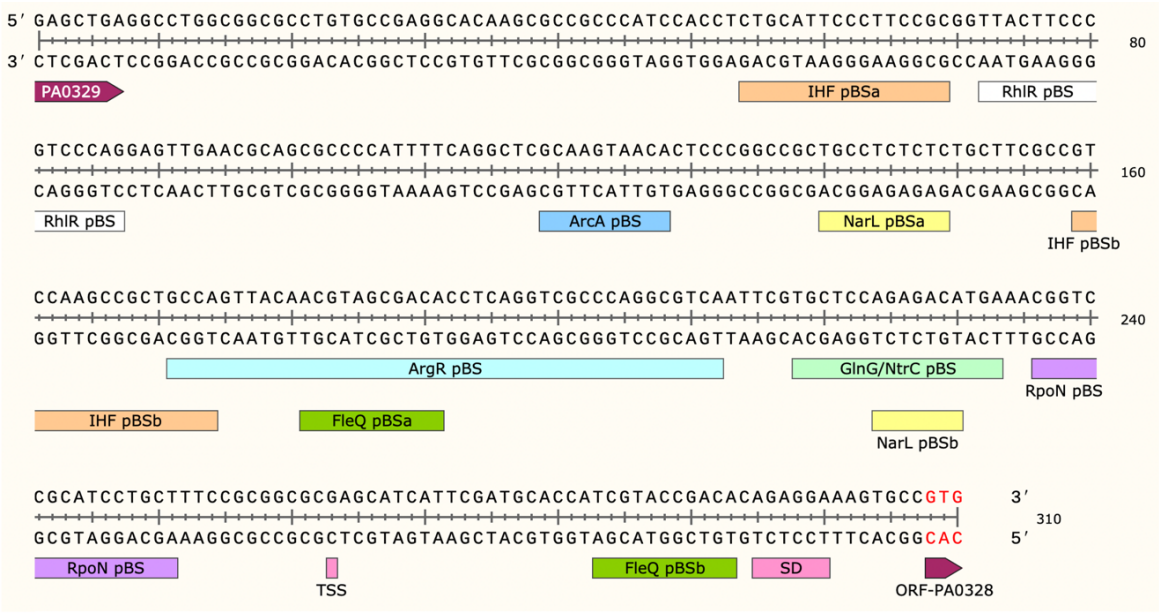


Figure 5.1: Schematic of the aaaA promoter and its putative transcriptional regulator binding sites identified by Spencer (2018)***.*** *Putative binding sites (pBS) are annotated at their predicted location for ArcA, NarL, IHF, ArgR, FleQ, GlnG/NtrC and RpoN. Where there are two putative binding sites, they are labelled a and b, arbitrarily. The Shine-Dalgarno (SD) sequence and transcriptional start site (TSS) are also labelled. The end and beginning of the open reading frames (ORF) for PA0329 and PA0328 (aaaA), respectively are shown in mauve, with the start codon (GTG) for aaaA in red.*

Table 5.1: List of transcriptional factors with putative binding sites in the aaaA promoter. *Selections were based on Virtual Footprint (PRODORIC) in silico screening with manual interrogation of the aaaA promoter (P_{aaaA}) by* Spencer (2018)*. Effect on P_{aaaA} are from previous observations using* miniCTX-P_{aaaA}*-lux transcriptional fusions where available* (Alhagh, 2020; Paredes-Osses, 2014)*. An upward, or downward-pointing arrow indicates this regulator was shown to upregulate or downregulate P_{aaaA}, respectively.*

Sigma factor	***Motif in P_{aaaA}***			***Effect on P_{aaaA}-lux***
	Sequence	**From**	**To**	
ArgR	GCCAGTTACAACGTAGCGACACCTCAGGTCGCCC AGGCGTCA	-137	-96	↑
ArcA	TGTTACTTGC	-189	-179	**No data**
RhlR	TTACTTCCCGTCCCAG	-236	-220	↓
RpoN	CGGTCCGCATCCTGCT	- 72	-57	**Variable**
FleQ	A: ACGTAGCGACA B: TCGTACCGACA	-127	-116	**Not significant**
GlnG (NtrC)	TGCTCC	-90	-74	↑
IHF	A: CTGCATTCCCTTCCGC B: GTCCAAGCCGCTGCCA	-254 -149	-238 -133	**No data**
NarL	A: TGCCTCTCTC B: AGAGACA	-168 -84	-158 -77	**No data**

5.1.2 Overview of experimentally determined *aaaA* regulators

5.1.2.1 Introduction to *lux* transcriptional reporters

Transcriptional reporters are a means of measuring transcription *in situ*, by monitoring the activation of a given promoter, linked to a reporter such as *lacZ* or luciferase genes which are transcribed upon promoter activation. The miniCTX-*lux* vector, developed in the Schweizer lab in 2000, fuses the target promoter to the *lux* gene cluster (*luxABCDE*) from *Xenorhabdus luminescens*, which is stably inserted as a single copy into the chromosome, via the CTX chromosomal phage attachment site (Becher & Schweizer, 2000; Hoang *et al.*, 1998). The *luxAB genes* encode a thermostable luciferase, and *luxCDE* encodes the fatty acid reductase complex which synthesises the natural luciferase substrate tetra-decanal, meaning no exogenous substrate is needed (Voisey & Marincs, 1998). Thus, whenever the native promoter of the target gene is activated, the cloned promoter fused to the *lux* operon similarly activates transcription, producing a bioluminescent signal which is proportional to promoter activation in the population. This can be measured over time using a plate reader, as illustrated in **Figure 5.2**. Previously, a reporter containing the promoter region for *aaaA* (P_{aaaA}), which contains the first four codons of the PA0328 open reading frame (ORF), plus 312 bp upstream, was created by Paredes-Osses (2014), with the resulting vector named miniCTX::P_{aaaA}-*lux*.

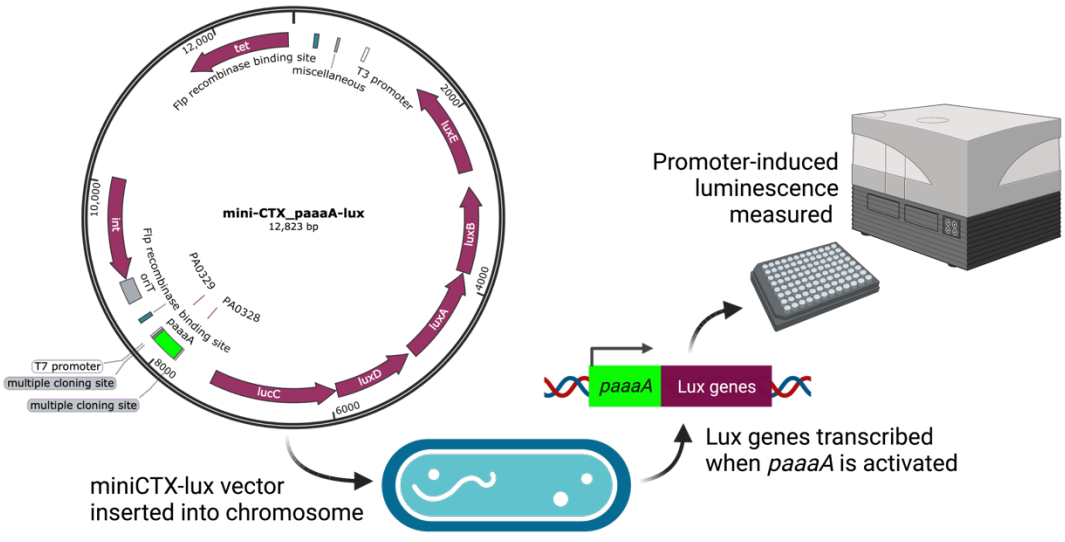


Figure 5.2: Illustration of miniCTX::P_{aaaA}-lux transcriptional reporter assay.

As a positive control, a *mini-CTX-lux* cassette fused to the constitutively active promoter derived from the kanamycin resistance gene cassette (P_{kan}), constructed by Chen (2014) was also used. Bioluminescence created by constitutive activation of P_{kan} can generally be used as an indication of metabolic activity of the cells, given the requirement of oxygen and ATP for the production of bioluminescence, but is also an indicator of any secondary effects on luciferase production or activity, which could bias results.

5.1.2.1.1 A note about the use of different lux data analysis methods

A caveat when comparing findings from *lux* transcriptional reporters, is that there are variations in the chosen data analysis methodology. Paredes-Osses (2014) chose to report the maximum RLU value for a given strain across the whole experiment (18-24 hrs). Alhagh (2020) instead appears to have chosen to report the mean *lux* expression over 16 hrs in triplicate, which integrates data from multiple points and therefore includes both breadth and magnitude of all activation peaks but omits information on error arising from variation across replicates. The methodology used in this study, and by (Spencer, 2018), was to report the AUC, to integrate as much data as possible. The impact of choosing different analysis methodologies on the overall conclusions drawn is unclear, but it is worth bearing in mind when comparing results across different studies.

5.1.2.2 AaaA autoregulation

Transcriptional reporter data has consistently shown that *aaaA* negatively regulates its own expression, with ~3-5 fold higher *aaaA* promoter activation in the *aaaA* mutant (Alhagh, 2020; Kiney-Whitmore, 2017; Paredes-Osses, 2014; Spencer, 2018). This was unaffected by addition of exogenous arginine and persisted across a range of conditions, with the single exception that when grown in minimal media supplemented with 20 mM succinate and 10 mM KNO_3, Spencer (2018) did not observe this difference (Figure 6.2.5 in Spencer, 2018). The mechanism for autoregulation is unclear. Spencer (2018) suggested that it may be mediated by RpoN (see: 5.1.2.3.3), however beside the link to nitrate supplementation, there is not a lot of evidence to support this.

5.1.2.3 Existing evidence points to ArgR and FleR/RpoN as the major regulators of *aaaA*

The regulators and related genes shown to effect P_{aaaA} activation and AaaA activity are summarised in **Table 5.2** and **Table 5.3**, respectively. There are some conflicting findings across the existing body of work, particularly between findings from PAO1 *vs* PA14. However, two regulators were consistently associated with increased *aaaA* expression, ArgR and FleR (in concert with FleQ and RpoN), with ArgR having the larger and stronger evidence base. Data has also been generated with respect to arginine metabolism and acquisition genes, and some other RpoN EBPs, as well as MvaT/U. However, a lack of consistency in the data generated has made it difficult to draw conclusions about their respective roles in *aaaA* regulation, particularly in terms of their directionality (i.e., up or downregulation).

Table 5.2: Genes previously shown to alter P_{aaaA}-lux activation. *Based on previous observations using* miniCTX-P_{aaaA}-lux *transcriptional fusions* (Alhagh, 2020; Kiney-Whitmore, 2017; Paredes-Osses, 2014; Spencer, 2018)*. [a] Indicates 'effects' with a p-value of 0.05-0.1. [b] Indicates genes found to significantly affect* P_{aaaA} *activity (p<0.05) but with significant 'non-specific' differences in expression driven from* P_{kan}*. [c] Indicates genes where there are conflicting data, which in the case of speA means it is listed on both sides. Genes highlighted in green have consistent findings which agree with the activity assay data in* ***Table 5.3****.*

⬆ P_{aaaA}	Role	⬇ P_{aaaA}	Role
argR	Global arginine regulator	[c]***aaaA***	Undetermined. Predicted role in arginine acquisition.
[a]***anr***	Anaerobic regulator of arginine deiminase and nitrate reductase	[c]***rpoN***	Alternative sigma factor 54 (nitrogen assimilation etc.)
argJ, argG, argF	Arginine biosynthesis	***aruB***	Arginine catabolism: AST pathway
aotJ, arcD	Arginine importer across the IM	[b]***rsaL***	Downregulates *lasI*
[c]***sfa2, sfa3***	T6SS regulation	***rhlR***	QS regulator
[bc]***speA***	Arginine catabolism: ADC pathway	[c]***speA***	Arginine catabolism: ADC pathway
ntrC	General nitrogen response regulator	***mvaT, mvaU***	small DNA-binding proteins, control *aotJQMOP-argR*
fleR	flagella biogenesis regulator		
acoR	2,3-butanediol utilisation		

Table 5.3: Genes previously shown to effect AaaA enzymatic activity. *Summary of findings by* Alhagh (2020)*. AaaA activity was measured by whole-cell L-arginine-p-nitroanilide cleavage assay of PA14 transposon library mutants (^) or PAO1 mutants (*). FD: approx. fold-difference in endpoint L-arginine-p-nitroanilide cleavage by mutant compared with wild-type. Those highlighted in green have no conflicting data and findings consistent with the* P_{aaaA} *transcriptional reporter data shown in* ***Table 5.2****.*

⇧ AaaA activity	FD	⇩ AaaA activity	FD	No/negligible effect
sfa3*	-4.3	***rhlR****	+2.0	***sfa3^***
fleR^	-4.0	***dctD^***	+1.4	***acoR^***
sfa2*	-3.25	***pilA^***	+1.4	***ntrC^***
mvaU*	-1.8	***lasR****	+1.2	***fleQ^***
ddaR^	-1.6	***pqsE****	+1.2	***39360^***
sfa2^	-1.6	***cbrB^***	+1.1	***rtcR^***
nirQ*	-1.6			***gscR^***
argR^	-1.5			***mifR^***
phhR^	-1.4			
norR^	-1.4			
rpoN*	-1.3			
mvaT*	-1.3			
13000^	-1.3			
eatR^	-1.2			
hbcR^	-1.1			
algB^	-1.1			

5.1.2.3.1 ArgR and arginine metabolism genes

As discussed in 1.2.1, ArgR is the global response regulator to arginine and there is evidence to suggest it plays a key role in *aaaA* regulation. Firstly, it was shown, using transcriptomics, that deletion of *argR* leads to a 3.7-fold decrease in *aaaA* mRNA levels (Lu *et al.*, 2004). Secondly, Spencer (2018) identified a putative ArgR binding site in the *aaaA* promoter region **Figure 5.1**. However, no work has yet been carried out to interrogate this site experimentally. Thirdly, Paredes-Osses (2014) showed that deletion of *argR* in PA14 led to a ~2-fold decrease in P_{aaaA} activation, suggesting that ArgR positively regulates *aaaA* expression. These findings were replicated by Alhagh (2020), who additionally showed that deletion of *argR* also led to reduced AaaA activity when tested using the L-arginine-p-nitroanilide cleavage assay, with endpoint cleavage ~1.5-fold lower in the Δ*argR* mutant than wild-type PA14.

Taken together, these findings provide a strong experimental and theoretical argument that ArgR is a major regulator of *aaaA*.

The mechanistic basis for ArgR regulation is a little less clear. ArgR responds to increases in arginine by up and downregulating arginine catabolic and anabolic pathways, respectively. As AaaA is thought to involved in arginine acquisition, it is unclear why it would be upregulated in response to arginine, given that the arginine is already present, thus acquisition is perhaps less required. On the other hand, ArgR also upregulates arginine transporters *arcD* and *aotJ*, suggesting there is a positive feedback loop for arginine acquisition, where the presence of arginine signals that further arginine may be available, driving AaaA and other acquisition genes, as part of a nutrient scavenging mechanism.

5.1.2.3.2 Arginine uptake and metabolism genes

Paredes-Osses (2014) found that deletion of genes involved in arginine uptake (*aotJ* and *arcD*) and arginine biosynthesis lead to decreases in P_{aaaA} activation, suggesting they positively regulate *aaaA*. As mentioned above, these genes are involved in arginine acquisition, so it would make sense that they would be functionally linked to AaaA. It is most likely that this regulation is occurring via ArgR, as functions which lead to increased intracellular arginine would induce ArgR, which would in turn increase *aaaA* expression (Lu *et al.*, 2004). Interestingly, Paredes-Osses (2014) also found that deletion of *speA* and *aruB*, enzymes involved in the ADC and AST pathways or arginine catabolism, respectively, caused small but significant increases in P_{aaaA} activation, suggesting they negatively regulate *aaaA,* possibly also via ArgR, by reducing intracellular arginine concentrations.

The finding regarding *speA*, however, was not replicated by Kiney-Whitmore (2017), who, using the same construct and experimental conditions, found the opposite, suggesting SpeA was in fact a positive regulator of *aaaA*. Although, unlike Paredes-Ossess, they also noted differences in P_{kan} activation between the wild-type and *speA* mutant, suggesting possible non-specific confounding due to disruptions in luminescence. Kiney-Whitmore (2017) also found that addition of exogenous spermidine, which was predicted to

compliment for loss of *speA*, reduced both P_{aaaA} and P_{kan} activation, which could support the findings by Paredes-Ossess. However, this may have been a result of non-specific effects of spermidine, and other polyamines, on transcription due to its affinity for forming complexes with RNA/DNA (Sun *et al.*, 2014).

5.1.2.3.3 RpoN, the nitrogen limitation sigma factor 54

As discussed in 1.1.2.3.3, RpoN is the nitrogen-limitation alternative sigma factor (σ^{54}) involved in nitrogen assimilation, amino acid uptake and degradation and expression of flagella and pili. As summarised in **Figure 5.1**, interrogation of the *aaaA* promoter identified an RpoN binding site, suggesting it has a key role in *aaaA* regulation (Spencer, 2018). It was also discussed in 1.1.2.3.3 that RpoN requires the assistance of EBPs to transition from the closed to the open confirmation and initiate transcription. The regulators, or EBPs, most likely working with RpoN to regulate *aaaA* are FleQ and NtrC, based on the presence of putative binding sites in the *aaaA* promoter region. However, there is some evidence that other proteins with Sigma-54 interaction domains could acts as EBPs to affect *aaaA* expression in an RpoN-dependant manner.

5.1.2.3.3.1 FleR and FleQ

FleR is part of a TCS with FleS, which together with RpoN and FleQ upregulate flagella biogenesis genes in a c-di-GMP-dependant manner (Dasgupta *et al.*, 2002; Matsuyama *et al.*, 2016). Thus, FleR positively regulates biofilm formation, as well as being essential for swimming motility (Zhou *et al.*, 2021). There is also recent evidence that FleR/S represses the H1 cluster of the type 6 secretion system via AmrZ, a global regulator which controls the production of several diguanylate cyclases and PDEs (Zhou *et al.*, 2022). FleR is also essential for both swimming and swarming motility, and loss of FleRS decreases biofilm formation, as the flagella is essential for both phenotypes (Kollaran *et al.*, 2019). It is unclear how FleR is functionally related to AaaA, but there is some evidence that it affects *aaaA* expression.

Alhagh (2020) found that deletion of *fleR* in PA14 lead to reduced P_{aaaA} activation, suggesting it positively regulates *aaaA.* They also showed that AaaA activity was almost totally lost in mutants lacking *fleR,* suggesting that it essential for AaaA production and activity. However, deletion of *fleQ* in PA14

did not significantly affect P_{aaaA} activation, suggesting regulation does not depend on FleQ. It is worth noting that changes in *aaaA* transcription were not identified by transcriptional profiling using genome microarrays of *P. aeruginosa* PAK strains with deletions in either *fleQ*, *fleR* or *rpoN* (Dasgupta *et al.*, 2003). It may be that under different conditions, or using a more sensitive transcriptomic technique, differences may be observed, and it would be useful to further explore this, given the existing and consistent evidence for a regulatory relationship between FleR/S, RpoN and AaaA.

5.1.2.3.3.2 NtrC

NtrC is part of a TCS, NtrBC, and is important for utilisation of nitrogen-only sources, such as nitrate and nitrite (Li & Lu, 2007). In PA14 *ntrC* and *ntrB* deletion mutants, *aaaA* was downregulated by 3.14 and 3.15-fold, respectively, compared with wild-type, suggesting NtrC positively regulates *aaaA* (Alford *et al.*, 2020). This is supported by Alhagh (2020), who found that deletion of *ntrC* in PA14 led to reduced P_{aaaA} activation, although it did not affect AaaA enzymatic activity, suggesting it is not essential for AaaA production, at least under those conditions.

5.1.2.3.3.3 AcoR

Alhagh (2020) found that deletion of *acoR*, which also has a Sigma-54 binding domain, in PA14 led to reduced P_{aaaA} activation, suggesting it positively regulates *aaaA*, however, like NtrC, there was no difference in AaaA enzymatic activity between the PA14 wild-type and Δ*acoR* mutant. This suggests again that it may be a non-essential activator of *aaaA* transcription.

AcoR is the regulator responsible for expression of the *bdu* operon, encompassing genes necessary for utilisation of the metabolite 2,3-butanediol via conversion into acetyl-CoA and acetaldehyde (Liu *et al.*, 2018). 2,3-butanediol is produced under anaerobic conditions by fermenting bacteria commonly found co-habiting with *P. aeruginosa* in the CF airway, such as *S. aureus*, *Klebsiella pneumoniae* and *Enterobacter aerogenes* (Liu *et al.*, 2018). 2,3-butanediol has been shown to enhance virulence and biofilm formation of *P. aeruginosa in vitro* as well as improving its persistence in a murine lung infection model and inducing enhanced TNF-α and IL-6 responses (Nguyen *et*

al., 2016; Venkataraman *et al.*, 2014). AcoR is therefore likely to be important in nutrient acquisition under competitive, anaerobic conditions and it makes sense that it could act to simultaneously increase *aaaA* expression under such conditions.

5.1.2.3.3.4Sfa2 and Sfa3

Sfa2 and Sfa3 are EBPs which are encoded on the H2-T6SS and H3-T6SS gene clusters, respectively, and while Sfa2 associates with RpoN to control the expression of H2-T6ss cluster, the role of Sfa3 is less clear (Sana *et al.*, 2013).

Paredes-Osses (2014) showed that deletion of either *sfa2* or *sfa3* in PA14 led to a decrease in P_{aaaA} activation, suggesting that they act as positive regulators of AaaA. However, when repeated by Alhagh (2020), no significant difference in P_{aaaA} activation was observed in either strain. Deletion of *sfa2* did lead to a near-total loss of AaaA activity in PAO1, and partial loss in PA14, in L-arginine-p-nitroanilide cleavage assays, supporting their role as important positive regulators of AaaA. However, while deletion of sfa3 lead to loss of AaaA activity in PAO1, it had a negligible effect in PA14. This highlights the heterogeneity of regulator action, which appears to depend on the background strain.

5.1.2.3.3.5 Deletion of rpoN alone curiously leads to increased P_{aaaA} activation

Curiously, mutation of *rpoN* itself lead to varied, strain-specific changes in *aaaA* expression. For example, Paredes-Osses (2014) found that in a PA14 Δ*rpoN* mutant, P_{aaaA} activation was increased, suggesting that the loss of RpoN somehow had a positive regulatory effect on *aaaA* expression. However, when repeated by Spencer (2018), using a PAO1 Δ*rpoN* mutant and PAO1 Δ*rpoN T7::rpoN*$^{+}$ overexpressing complemented strain, there was no difference in P_{aaaA} activation between the wild-type and mutant. Overexpression of *rpoN* also increased P_{aaaA} activation 2-3-fold (although P_{kan} activation was also greatly increased ~1.5-fold in the Δ*rpoN T7::rpoN*$^{+}$ suggestive of non-specific increases in *lux* expression in this strain). Additionally, assessment of the effect of RpoN on AaaA activity by L-arginine-p-nitroanilide cleavage assay found that

AaaA activity was slightly reduced (~1.3-fold) in both the PAO1 Δ*rpoN* mutant and Δ*rpoN T7::rpoN*$^{+}$ complementation strain, (Alhagh, 2020).

Taken together, the picture that emerges is that when RpoN is lost, P_{aaaA} activation is sometimes, but not always, increased, but actual AaaA production is likely decreased. This suggests that in the absence of RpoN, which typically acts in concert with EPBs (FleR/FleQ, NtrC, AcoR and Sfa2/3) to upregulate *aaaA*, the effects of other regulators acting on the *aaaA* promoter may become more pronounced. Given the density of putative binding sites in the *aaaA* promoter region, especially around -50 to -150 bp from the TSS, this would not be surprising. For example, the ArgR binding site coincides with one of the putative FleQ sites, as well as MvaT/U (5.1.2.3.4) sites overlapping with the RpoN binding site, suggesting there may be competition for binding that could impact upon regulation. Further work to construct double mutants, for example Δ*argR*Δ*rpoN*, would be helpful to establish what interactions predominate in this region.

5.1.2.3.3.6 Mutation of the RpoN binding site in the aaaA promoter ablated aaaA autoregulation

To further examine the role of RpoN on *aaaA* regulation, Alhagh (2020) mutated the putative RpoN binding site predicted by Spencer (2018), within the miniCTX-P_{aaaA}-*lux* reporter to replace the conserved GG and GC residues to adenosine as follows: CAATCCGCATCCTAAT, and named this miniCTX-$P_{aaaA\,\Delta rpoNmotif}$–*lux.* Using this construct, they found that when the putative RpoN binding site in the *aaaA* promoter is mutated, $P_{aaaA\,\Delta rpoNmotif}$ activation was almost totally lost in wild-type PAO1N and PAO1 Δ*rpoN T7::rpoN*$^{+}$ complementation, yet $P_{aaaA\,\Delta rpoNmotif}$ activation was still high in the PAO1N Δ*rpoN* mutant. This suggests that when RpoN is present, it is acting at its binding site to positively regulate *aaaA* in concert with its EBPs. However, when RpoN is absent, *aaaA* expression is increased, suggesting that an alternative activator, possibly ArgR, is more able to upregulate *aaaA* in the absence of RpoN, potentially due to reduced steric interference.

Interestingly, Alhagh (2020) also found that there was no difference in $P_{aaaA\,\Delta rpoNmotif}$ activation between PAO1N and PAO1N Δ*aaaA*, suggesting that the

consistently observed *aaaA* auto-repression is mediated by RpoN binding at this site, as suggested by Spencer (2018).

It is also worth noting there were some conflicting data reported by Alhagh (2020), including some evidence that was anomalous, such as Figure 4.18a in their thesis, where luminescence values are either much higher or lower than typical. There were other issues similar to these within the data presented by Alhagh (2020), and their findings, in general, therefore require further validation and careful appraisal.

5.1.2.3.4 MvaT and MvaU

MvaT and MvaU are homologous to the *E. coli* histone-like nucleoid-structuring proteins (H-NS). Unlike sigma factors, which bind to specific promoter binding sites, MvaT/U alter gene expression by binding to AT-rich regions in curved sections of DNA (Castang *et al.*, 2008; Winardhi *et al.*, 2014). They've been shown to directly repress the *aotJQMOP-argR* operon, and which could indirectly repress *aaaA* (Li *et al.*, 2009a).

Both Paredes-Osses (2014) and Alhagh (2020) separately showed that deletion of either *mvaT* or *mvaU* leads to greater *aaaA* promoter activation, suggesting they negatively regulate *aaaA*. Curiously, this effect was reversed in experiments using the $P_{aaaA\,\Delta rpoNmotif}$ reporter. While deletion of either *mvaT* or *mvaU* leads to increased P_{aaaA} activation, it conversely led to a total loss of $P_{aaaA\,\Delta rpoNmotif}$ activation. These data require further validation, but if replicable, this suggests that if this putative RpoN binding site is disrupted, MvaT and MvaU switch from being negative to positive regulators of *aaaA*. The reason for this is unclear, but could be explained by findings from Spencer (2018), who identified 2-3 potential AT rich regions which MvaT/U could bind to in the *aaaA* promoter, one of which correlated to the RpoN binding site, further supporting the suggestion that the two regulators could modulate each other's regulatory actions on *aaaA*. More simply, it may be that exchanging the GG and GC in the RpoN binding site for AA and AA inadvertently altered the AT-richness of this site, which then indirectly impacted upon MvaT/U binding within this region.

The MvaT/U story is further confounded by evidence from L-arginine-p-nitroanilide cleavage (AaaA activity) assays by Alhagh (2020), showing that

AaaA activity was slightly lowered in Δ*mvaT* and Δ*mvaU* mutants with endpoint cleavage ~1.3-fold and ~1.75-fold reduced, respectively. If replicable, this suggests the possibility that MvaT and MvaU can have positive effects on AaaA production. As these experiments only assess *aaaA* promoter activation, and then active AaaA protein, respectively, it would be necessary to assess some of the intermediate stages between upregulation and surface expression, to really understand what is happening.

5.1.2.3.5 RhlR and other QS components

As discussed in 1.1.2.2, RhlR is a transcriptional regulator which controls QS genes, along with, and under the control of, LasR. Both Paredes-Osses (2014) and Alhagh (2020) showed with varying degrees of certainty, that deletion of *rhlR* led to an increase in P_{aaaA} activation (although Paredes-Osses controversially designated $p<0.1$ as the significance cut-off), suggesting RhlR negatively regulates *aaaA*. This was supported by the finding that AaaA enzymatic activity was increased in all three QS mutants, with a ~1.2-fold increase in endpoint cleavage in Δ*pqsE* and Δ*lasR* mutants, and a ~2-fold increase in the Δ*rhlR* mutant (Alhagh, 2020). Spencer (2018) also identified a putative RhlR binding site in the *aaaA* promoter, suggesting this regulation may be via direct binding. However, the binding site is rather far from the transcriptional start site, and there is no evidence from transcriptomics that *aaaA* is directly controlled by RhlR or other QS mechanisms (Chugani *et al.*, 2012; Schuster *et al.*, 2003; Wagner *et al.*, 2004).

In summary, evidence from various studies has implicated a number of regulators in *aaaA* regulation, in what is likely to be a complex, and even competitive web of interactions. These have been synthesised into a working model for *aaaA* regulation, which is shown in **Figure 5.3**.

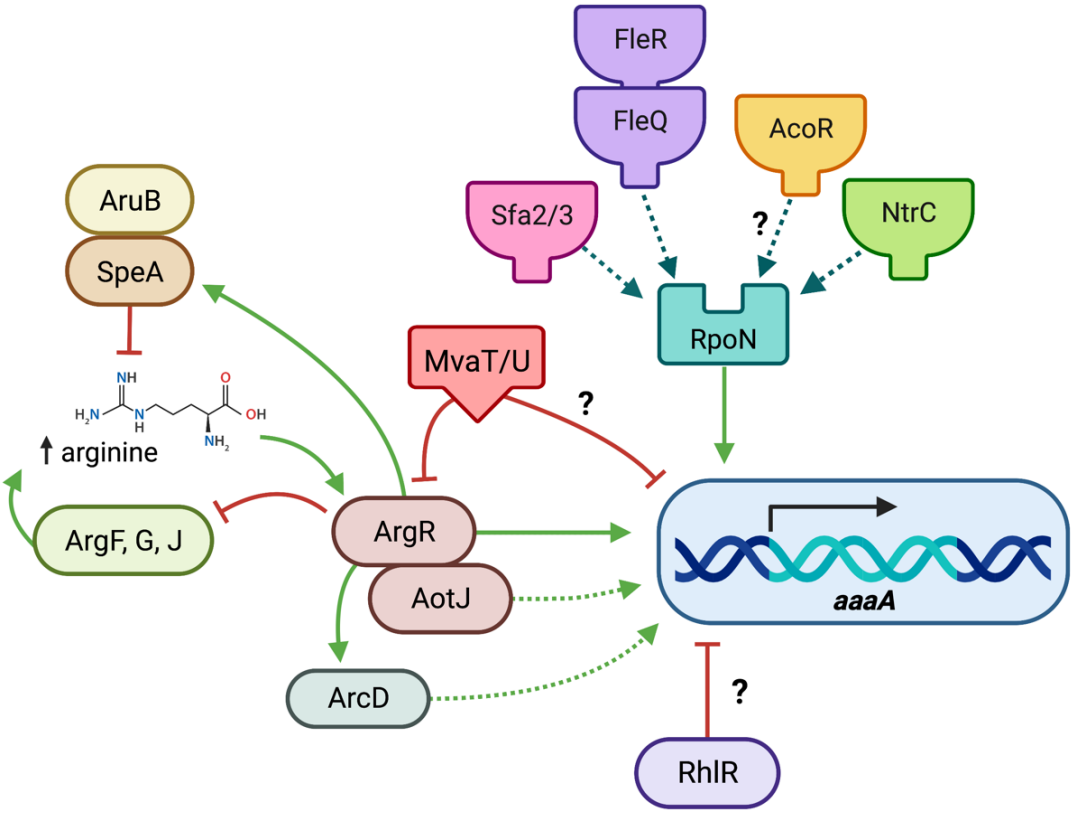


Figure 5.3: Summary of known and putative regulators of aaaA. *Positive regulation is indicated by a green arrow, negative regulation by a flat headed red arrow and enhancer binding protein regulators by dashed black arrows. Potential indirect positive regulation is shown by a dotted green arrow. Question marks denote regulators which are theoretical or have limited data to support the proposed action/mechanism. Created using BioRender.com.*

5.1.2.4 Effect of environmental conditions and oxygen on AaaA

5.1.2.4.1 Salt and *pH*

Previous work by Spencer (2018) found that salt concentration did not affect P_{aaaA} activation or AaaA activity in LB. However, in the Δ*aaaA* mutant, high (~300 mM) salt concentrations reduced P_{aaaA} activation. Salt concentration did not however, affect AaaA enzymatic activity using the L-arginine-p-nitroanilide cleavage assay.

Altering pH also had no effect on *aaaA* promoter activation. However, AaaA enzymatic activity was higher at a neutral pH (6.8-7.2) than a more acidic pH (<6.2) in the wild type, when grown in LB, although activity at pH >7.2 was not tested.

5.1.2.4.2 Arginine and nitrate

Interestingly, addition of exogenous arginine (5 mM) did not increase P_{aaaA} activation (Paredes-Osses, 2014). However, given the established link between ArgR and *aaaA*, it would be worthwhile to repeat this experiment with different concentrations of arginine.

The link between RpoN/NtrC and *aaaA* suggested that nitrate may impact *aaaA* expression. Indeed, addition of exogenous nitrate to LB reduced *aaaA* promoter activation in both PAO1N and PAO1N Δ*aaaA but* did not affect wild-type enzymatic activity of AaaA. However, in CF *P. aeruginosa* isolates, addition of 10mM exogenous nitrate did generally lead to lower *aaaA* promoter activation, though again, this did not cleanly translate into lower AaaA activity (Spencer, 2018). As mentioned above, is not entirely clear why the transcriptional reporter data and activity assay data often contradict each other, and this highlights the need for experiments to study intermediate stages of AaaA production to determine this.

Additionally, as mentioned in 5.1.2.2, to date, the only condition in which *aaaA* auto-repression was not observed was when PAO1N was grown in minimal media supplemented with 20 mM succinate and 10 mM potassium nitrate, though it is unclear why (Spencer, 2018).

5.1.2.4.3 Oxygen

Spencer (2018) showed that AaaA activity increased ~4-fold in microaerobic conditions, providing strong evidence that oxygen availability is a cue for *aaaA* regulation. If the primary purpose of AaaA is in nutrient acquisition in low oxygen or nutrient environments, then it would follow that its expression is tightly controlled. As such, it is imperative that *aaaA* regulation is studied in a physiologically relevant context, particularly if the nutritional environment is a driver of *aaaA* expression and function, and this is discussed further in 5.4.1.

5.2 Chapter aims

Previous work has provided insight into *aaaA* regulation, however, there have often been inconsistencies between *aaaA* expression and AaaA activity. This chapter therefore aims to build on previous work examining factors

upstream of aaaA transcription, by focusing on events downstream of *aaaA* transcription, to better understand how *aaaA* gene expression and regulation translates into to AaaA function, and its interplay with other genes. As AaaA has been shown to be essential for virulence in a chronic wound model, and also to have a marginal phenotype in planktonic conditions, this was examined in a chronic wound environment, to maximise disease relevance. Thus, here, *aaaA* transcription, as well as the transcriptomic impact of *aaaA* deletion, were investigated in a synthetic chronic wound model, to identify co-regulated genes which could give greater insight into AaaA function. The aims of this chapter were two-fold:

1. Construct new in-frame *aaaA* deletion mutants in PAO1L and chronic wound isolate backgrounds, to better improve the disease relevance of the strains used to study AaaA.
2. Examine *aaaA* expression and identify co-regulated genes in a synthetic chronic wound model, using transcriptional reporters, RT-qPCR and RNA-Seq.

5.3 Construction of Δ*aaaA* mutants in PAO1L and a chronic wound isolate

5.3.1 The PAO1N subline and its drawbacks

PAO1 was originally isolated in Melbourne, Australia in 1954, from a wound (Holloway *et al.*, 1979). A derivative of this, PAO1-UW, was later sequenced at the University of Washington, making it the largest sequenced bacterial genome at the time (Stover *et al.*, 2000). Two near-saturation libraries of transposon insertion mutants were also constructed in PAO1, which combined with the ubiquitous use of the PAO1 genome as a reference sequence, has made PAO1 the major strain for the study of *P. aeruginosa* (Klockgether *et al.*, 2010). It has since become widely accepted that PAO1 sublines around the world have deviated from the original isolate, including sublines with sequence variations in MexT, which regulates the MexEF-OprN multidrug efflux system (Maseda *et al.*, 2000).

Recent interrogation of the PAO1 Nottingham collection wild type (PAO1N) found a large (~59 kb) deletion in the PAO1N chromosome and reduced swarming motility compared to the PAO1 Lausanne (PAO1L) subline. Swarming in PAO1N is also ablated entirely upon knock out of *rsmA*, whereas in PAO1L, deletion of *rsmA* only modestly reduces swarming motility (Dubern *et al.*, 2022). Genes which restore swarming motility when complemented back into PAO1N included *proE*, *toxR* and five genes encoding hypothetical proteins: *PA2072, PA0285, PA3825, PA2567, PA0707.* These all have either been shown or predicted to contain conserved domains involved in c-di-GMP metabolism, which could explain their role in swarming motility. Notably, ToxR, which is known for regulating the production of endotoxin A, was shown to bind, but not directly degrade, c-di-GMP, and negatively affect c-di-GMP signalling. Thus, loss of *toxR* is likely to have wide ranging impacts on processes controlled by c-di-GMP signalling. As well as the large deletion, PAO1N also had a missense mutation in *bifA*, the gene encoding the c-di-GMP degrading PDE BifA. This resulted in the substitution, Y442D, in close proximity to the glutamate (Q443) involved in c-di-GMP binding.

Previous study of AaaA have largely been carried out in PAO1N, from which the PAO1 Δ*aaaA* strain was also derived, with some use of PA14 backgrounds (Luckett *et al.*, 2012). Arginine effects substantial changes in cell behaviour, largely via c-di-GMP signalling, as discussed in 1.2.5. Therefore, it was important to use a background strain which did not have compromised c-di-GMP signalling pathways, in order to better understand the regulatory network surrounding AaaA. Moreover, other genes lost in the ~59 kb deletion in PAO1N include those encoding the alkaline phosphatases LapA and LapB, as well as the *hxc* genes involved in alkaline phosphatase secretion. LapA has been shown to be highly expressed in *ex-vivo* chronic wound biofilms and may play important roles in adhesion and colonisation (Tan *et al.*, 2021). Therefore, a strain background lacking LapA is unlikely to be suitable for studying *P. aeruginosa* in a chronic wound biofilm context.

Considering all this, a new Δ*aaaA* mutant was constructed, with PAO1L as an alternative background due to the usefulness of PAO1 as a reference. To

better reflect the populations of bacteria in which AaaA is likely to be most relevant, a Δ*aaaA* mutant was also constructed in a clinical *P. aeruginosa* chronic wound isolate background, CW2-BI1 (discussed in 5.3.2.5).

5.3.2 Successful construction of an in-frame *aaaA* deletion in PAO1L and CW2-BI1

5.3.2.1 Construction of Δ*aaaA* suicide vector

A pDM4 suicide vector was previously used to construct PAO1N Δ*aaaA*, however this relied on chloramphenicol sensitivity, which was only possible because of the unique susceptibility of PAO1N to chloramphenicol (Luckett *et al.*, 2012). This is likely due to loss of *cat* in PAO1N, which encodes a chloramphenicol acetyltransferase conferring the typically intrinsic chloramphenicol resistance of *P. aeruginosa*. Alongside this, PAO1N also contains mutations in multidrug efflux genes *mexF* which encode MexEF-OprN, an efflux pump for which chloramphenicol is a substrate, and *mexT*, which encodes the regulator MexT which upregulates MexEF-OprN expression (Dubern *et al.*, 2022; Fetar *et al.*, 2011). Therefore, to ensure versatility, a new suicide plasmid for allelic exchange was constructed which did not rely on chloramphenicol sensitivity. The *pEX18Gm* vector was chosen, which contains a gentamicin resistance cassette, and the *sacB* gene, for counter-selection based on sucrose toxicity sensitivity after the second cross over (Hoang *et al.*, 1998).

Primers were designed to amplify the regions ~600 bp up and downstream of *aaaA*, including the first and last codons of *aaaA* to create an in-frame, non-functional 4 amino acid-long peptide, as well as add the restriction sites *Xba*I, *Bam*HI and *Eco*RI (Δ*aaaA*-Rx-Up/DnFrag-F/R) to each fragment as shown in the schematic in **Figure 5.4**.

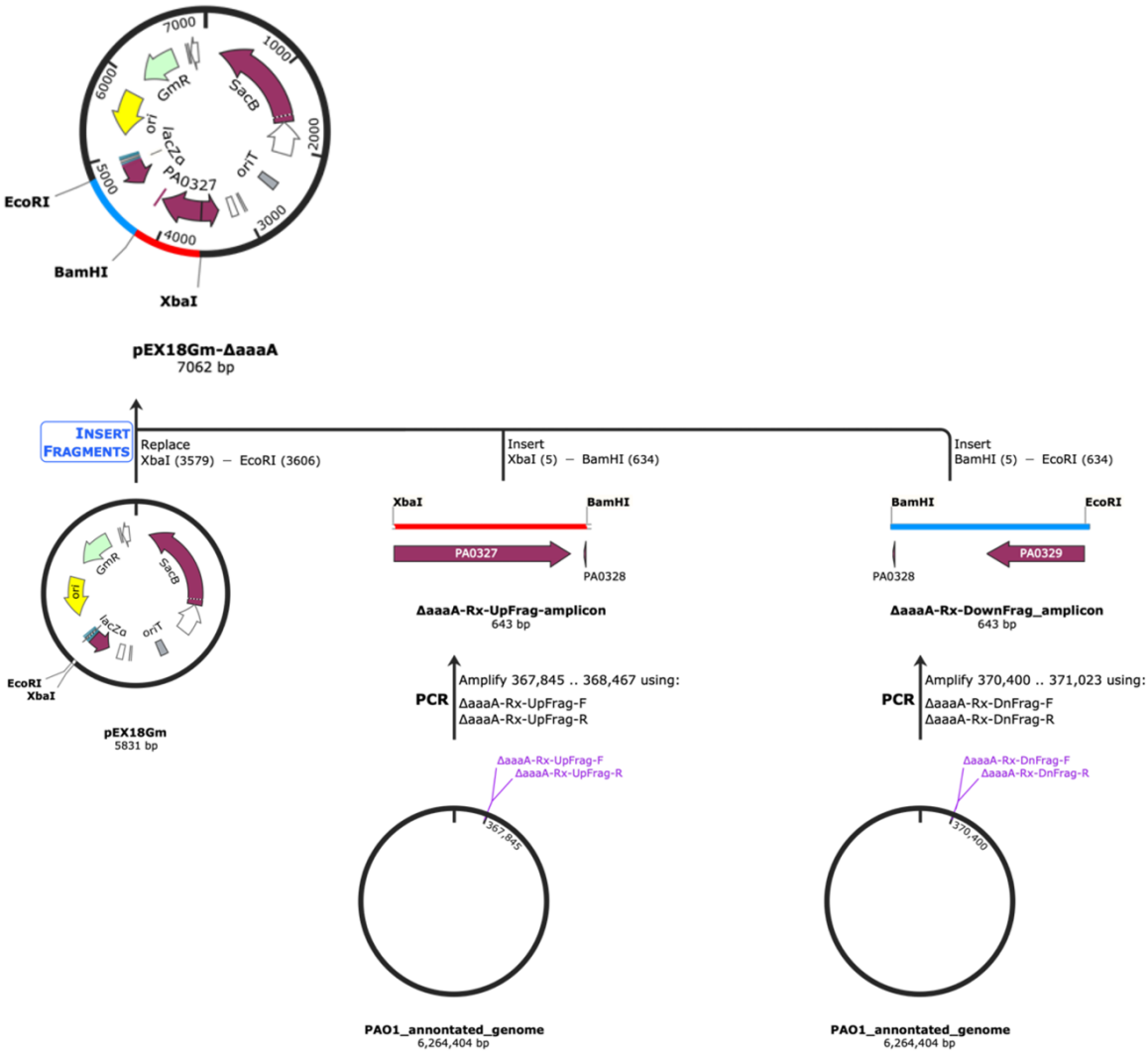


Figure 5.4: Schematic of cloning strategy for creation of suicide vector for making an in-frame aaaA deletion. *Generated using Snapgene.*

Detailed descriptions of cloning methodologies are described in 2.5. Briefly, extracted PAO1L gDNA was used as a template, and the amplicons (**Figure 5.5a**) and *pEX18Gm* suicide vector were cut using double restriction enzyme digests. Digest products (**Figure 5.5b**) were cleaned by agarose gel separation, and the three products ligated together and transformed into *E. coli* DH5α, which were plated onto selective blue/white screening plates (2.5.4.4). Colony PCR of selected white colonies confirmed insertion (**Figure 5.5c**). Sanger sequencing of plasmids extracted from colony 3 and 10 (**S-Figure 0.2**) showed good alignment to the target sequence with only 1 and 2 ambiguous bases in C3 and C10, respectively, thus C3 was selected for allelic replacement.

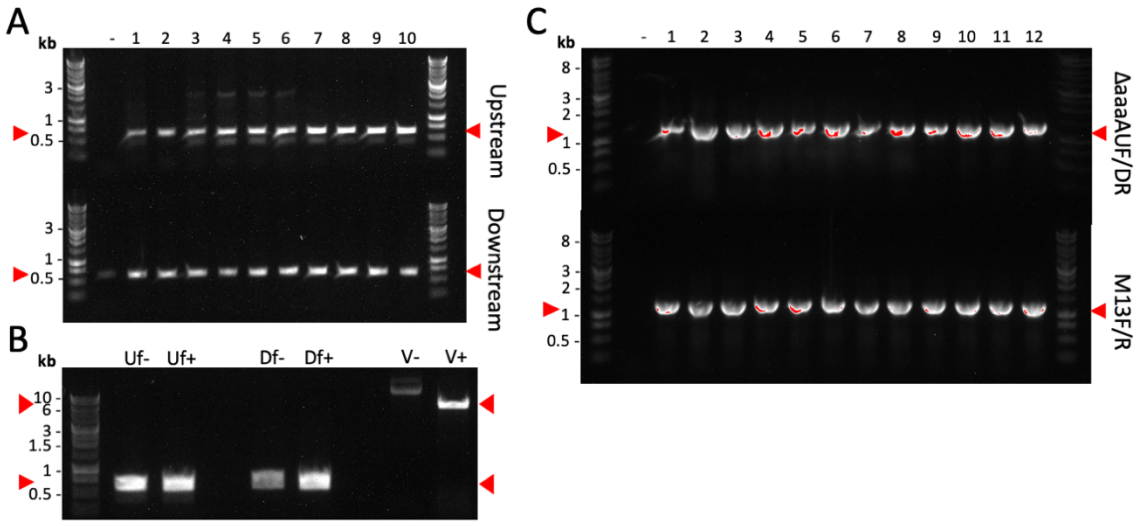


Figure 5.5: Cloning for construction of* pEX18Gm::ΔaaaA *suicide vector, using restriction-based cloning. Panel A *shows an agarose gel of gradient PCR amplicons of upstream and downstream regions either side of* aaaA*. Lanes are different T_m temperatures in ℃: 1. 48.6, 2. 56.3, 3. 52.8, 4. 55.9, 5. 58.9, 6. 62.1, 7. 65.1, 8. 68.1, 9. 70.7, 10. 72.4. Fragment 8 (T_m 68.1) for both were taken forward.* ***Panel B*** *shows an agarose gel of products from restriction digests with XbaI, BamHI and EcoRI as indicated in* ***Figure 5.4****.* ***Panel C*** *shows an agarose gel of colony PCR products from 12 transformed colonies (numbered lanes). Expected product: 1341bp with M13F/R and 1272 bp with ΔaaaA-Rx-UpFrag-F/ΔaaaA-Rx-DnFrag-R. Abbreviations: NTC- no template control, Uf – upstream fragment, Df- downstream fragment, V- Vector, +/- cut/uncut. Lane labelled – in panel A and C indicates a no-template PCR control.*

5.3.2.2 Insertion *of pEX18Gm-ΔaaaA* into PAO1L

The *pEX18Gm-ΔaaaA* extracted from C3 was transformed into electrocompetent *E. coli* S17-1. This strain acted as a donor for conjugation into PAO1L and merodiploids were selected for using PIA with gentamicin. The culture went through 3 passages in no-salt LB broth (recipe the same as LB, minus NaCl) with 10% sucrose, to screen for double-crossover clones, before plating onto LB no salt agar with 10% sucrose as in Paiva (2019). Colonies were re-streaked in parallel onto LB and LB with gentamicin. Colonies that were gentamicin sensitive were screened by colony PCR as above, using primers spanning outside the cloned up- and downstream regions (**Figure 5.6**).

The PCR product for selected colonies 7 and 13 were cleaned and sequenced by Sanger sequencing using these same primers, plus a primer mid-way through the region for full coverage. Sequencing (**S-Figure 0.3**) showed for both colonies, an in-frame deletion, with no mismatches, indicating the previous ambiguity in the reads for the suicide plasmid (**S-Figure 0.2**) were

sequencing-related. Due to a slightly clearer and longer sequencing read, colony 13 was selected and renamed PAO1L Δ*aaaA*.

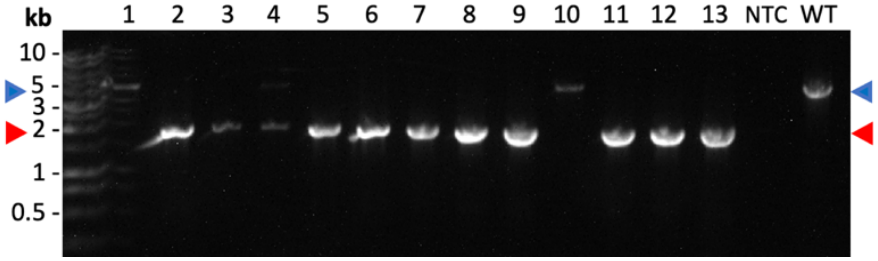


Figure 5.6: Colony PCR of PAOL ΔaaaA candidate post-sucrose selection. *Amplicon size in* ΔaaaA *is 1864bp, indicated by red arrow, and wild type (WT) is 3790bp, blue arrow. NTC = no template control.*

5.3.2.3 Verification of PAO1L Δ*aaaA* by WGS

The knock-out was further examined using whole genome sequencing (WGS) to ensure no secondary mutations were introduced. Preparation and WGS analysis are described in 2.5.6. The resulting raw reads for both PAO1L Δ*aaaA* and the original background PAO1L (both strain and raw sequencing reads kindly provided by Danni Scales) were uploaded to PATRIC for assembly, annotation, alignment, and SNP calling.

For PAO1L Δ*aaaA* there were 77 contigs, an estimated genome length of 6,276,498 bp, and an average G+C content of 66.50%. For PAO1L, there were 118 contigs, an estimated genome length of 6,225,891 bp, and an average G+C content of 66.55%. Both complete genomes were designated 'good' quality by PATRIC. The nucleotide sequence for PA0328 was searched for using PATRIC-BLAST in both genomes (search in contigs, e-value threshold, 10) (Boratyn *et al.*, 2013). PA0328 was found in the wild type only, (sequence identity 100%, score 3591), further confirming successful AaaA knock-out.

The PAO1L Δ*aaaA* genome was aligned to the PAO1L as a reference using PATRIC using four different methodologies (described in 2.5.6), to identify any possible secondary mutations and predict their potential effect (Wattam *et al.*, 2017). There were 8 secondary mutations which resulted in a non-synonymous variation; these are listed in **Table 5.4**. Of these, one was in the intergenic region, 6 were predicted to be of moderate or undefined severity by SnpEff, including mutations in PA5425 and PA5292 and one, in PA0412, was

predicted to be of high severity. There were also 7 synonymous variations in PA0720, PA4022, PA2402, PA1474 and PA5292.

A search of Pseudomonas.com found that PA0412 encodes PilK, a CheR-type methyltransferase and member of the Chp gene cluster (Darzins, 1995; Winsor *et al.*, 2016). Chp has roles in regulating surface attachment, twitching motility, type IV pilli-mediated swarming motility, biofilm formation, virulence and regulation of intracellular cAMP (McGowan, 2017). PilK, along with ChpB, are 'adaptation' proteins, involved in return chemosensory signalling back to 'resting' state, thereby assisting, for example, in pili retraction. The frame shift at proline 219 was identified in PAO1L Δ*aaaA* by one of the four methodologies (BWA-MEM-strict plus SAMTools) and occurs due a 67 bp deletion at position 656. This is toward the C-terminus of the protein, and interrogation of the amino acid sequence with IntroProScan suggests that this would disrupt the CheR methyltransferase domain, and lead to loss of function (Jones *et al.*, 2014). However, the twitching phenotype of Δ*pilK* was previously found to be indistinguishable from wild type in twitching motility assays, both microscopically and macroscopically (Darzins, 1995). This may be due to an unknown redundancy, or compensatory changes in ChpB (McGowan, 2017). Furthermore, many of the Chp-mediated process listed above were shown to be reliant on ChpB, suggesting the role for PilK is potentially less essential (Caiazza *et al.*, 2007; Fulcher *et al.*, 2010; Whitchurch *et al.*, 2004).

PA5425 encodes PurK, the *E. coli* homolog of which is involved in purine biosynthesis, and catalyses the ATP dependant conversion of 5-amino-1-(5-phospho-beta-D-ribosyl)imidazole into 5-carboxyamino-1-(5-phospho-D-ribosyl)imidazole (Mueller *et al.*, 1994). The in-frame deletion identified in PAO1L Δ*aaaA* spans arginine 282 to valine 314. This is toward the C-terminus of the protein, and according to annotations by InterPro, has a three amino acid overlap with the ATP-grasp domain (102-285) (Blum *et al.*, 2021). The impact of the loss of these 32 amino acids was predicted by SNP_Eff to be moderate rather than high, possibly because it occurs away from the functional domain of the protein.

PA5292 encodes PchP, a phosphorylcholine phosphatase which works alongside haemolytic phospholipase C (PlcH) to break down choline-containing phospholipids in the host cell membrane as an energy source, which also contributes to pathogenesis (Beassoni *et al.*, 2007; Massimelli *et al.*, 2005). The crystal structure of PchP has been resolved (PDB: 4AS2), and based on this, the region from valine 321 to the C-terminus is designated by the PDB database as disordered (Infantes *et al.*, 2012). This suggests that a frame shift deletion at leucine 332 may not have a large impact on the folding and function of the protein and may also be why SNP_Eff was not able to designate an effect type.

The remaining mutations occurred in as yet uncharacterised proteins, and it is therefore not clear what the impact of non-synonymous variations would be. Overall, though it is not likely that the mutations found in PAO1L Δ*aaaA* listed in **Table 5.4** and discussed above will have strong phenotypes, the possibility cannot be excluded and should be borne in mind when using this strain.

Table 5.4: List of non-synonymous variations identified in PAO1L ΔaaaA compared with the PAO1L reference genome. *Alignments were made using PATRIC according to 5.3.2.3. Variations designated by SnpEff as high severity, moderate/undefined severity and modifiers are coloured yellow, white, and blue respectively. Abbreviations: nt- nucleotide, AA- amino acid, del- deletion, ins- insertion, fs- frame shift.*

Pos	**Locus_tag**	**Function**	**nt change**	**AA change**	**Score**	**Var_ cov**	**Var_ frac**	**Snp_Eff type**
684	**PA2421**	Hypothetical protein, co-expressed with pyoverdine biosynthesis regulon	623_709delGCCTGGGCGCAAACGACGTACAGCCGGGCCAGGCGCTGCACGCGAACTTCTCGATCTCCTACGCGCTGGACCCGCACTGGCGGATCG	R208_G237del	214.458	20	1	in-frame deletion
4416	**PA2625**	Nudix-like NDP and NTP phosphohydrolase NudJ	418ACA>CCC	T140P	1094.66	35	0.97	missense variant
4435			437T>A	V146N	1533.71	48	1	missense variant
10294	**PA5425 (*purK*)**	N5-carboxyaminoimidazole ribonucleotide synthase (EC 6.3.4.18)	846_941delCGCCGTCGCCGGCCTGCCGCTGGGCTCGACCGCCAAGGTCGGCGAGAGCGCGATGCTCAATTTCATCGGCGCGGTCCCCCCGGTGGCTCAGGTGGT	R282_V314del	214.458	22	1	in-frame deletion
18148	**PA4686**	Chromosome segregation ATPases	1452_1532delCCTGCGCGACCAGAAGGAGCGCCTGGAGAAGGAGCTGAAGCAGCTCAAGACCCAGCAGAGCGTCGCTGCCGACCGCGCCGC	A484_A511del	100.973	4	1	disruptive in-frame deletion
25067	**PA0412 (*pilK*)**	Chemotaxis protein methyltransferase CheR (EC 2.1.1.80)	656_723delCCTGGTCCGGCATGGACGTGATTTTTTGTCAGAACCTGCTGATCTACTTCCGTCGCTGGCGACGGCGC	P219fs	114.973	4	1	frameshift variant
31456	**PA3906**	co-chaperone, co-TecT	275_276insGCCGGC	P92_L93insRRP	34.2251	2	1	disruptive in-frame insertion
77234					136.973	4	1	intergenic region
82243	**PA5292 (*pchP*)**	Phosphoserine phosphatase (EC 3.1.3.3)	995_1082delTGCCGGTGACGGCGGACAGGAACTGGGTGATCGTCACTCCGGAGCAGATCCAGTAGGCGAAACGGCGAGGGGCCGGCATTGGCCGGCC	L332fs	214.459	40	0.95	

5.3.2.4 Complementation and phenotypic verification of PAO1L Δ*aaaA*

A complementation strain was created using the chromosomally stably inserted miniCTX-*aaaA* created by Paredes-Osses (2014), according to methods in 2.5.4.5. Phenotypic confirmation of AaaA knock-out and complementation was then completed using a whole cell AaaA activity assay (**Figure 5.8**).

5.3.2.5 Construction of wound isolate CW2-BI1 Δ*aaaA*

To create a more clinically relevant background for further study, a Δ*aaaA* strain of *P. aeruginosa* CW2-BI1 (kindly provided by Dr Ana da Silva) was also created, as described above using the *pEX18Gm-ΔaaaA* suicide plasmid. This work was done with the kind assistance of Bethan Roberts and Danni Scales, as described in 5.3.2, although sucrose selection occurred by passage on LB no-salt agar plates rather than in broth.

P. aeruginosa CW2-BI1 was originally isolated by Dr Tim Sloan (Clinical Microbiology Department, Nottingham University Hospitals) from a diabetic foot (big toe) ulcer wound in 78-year-old Patient V, who also had osteomyelitis. The bacteria had been present in the wound for 3 months, and was isolated by bone biopsy (BI, bone isolate), alongside a methicillin-sensitive *S. aureus.* The following work to characterise the isolate was carried out by Dr Ana da Silva (2018). Antimicrobial susceptibility assays found CW2-BI1 was sensitive to ciprofloxacin, gentamicin, piperacillin-tazobactam, ceftazidime and meropenem, and intermediately resistant to aztreonam. CW2-BI1 was exoprotease positive and found to produce similar levels of QS signals NHQ, HQNO, NQNO, PQS, and C9- PQS to PAO1, with the exception of reduced HHQ, as well as exhibiting low pyoverdine, and moderate pyochelin and pyocyanin production. The isolate was also previously sequenced using Illumina whole genome sequencing (WGS) by da Silva (2018).

The sequence for *aaaA* was extracted from the CW2-BI1 WGS, translated and aligned with PA0328 using PATRIC (Davis *et al.*, 2020). **Figure 5.7** shows an alignment between the amino acid sequence for AaaA in the PAO1 reference and CW2-BI1. Variations include: K55E, A167T, Q314L N368D

and H516R. Of these, only A167T is not one of the common mutations listed and investigated in 3.5, although it is included in **Figure 3.7** and **Figure 3.8** for illustrative purposes. The start codon, M1 in PA0328, is GTG in both strains, thus could be designated either M or V. The earlier start codon in the alignment for CW2-BI1 reflects the alternative start codon for AaaA present in both genomes, dubbed the 'AaaA long version' by Darch (2009). These positions were assigned automatically by the respective databases, although the evidence suggests the start codon in PA0328 is the true one (Davis *et al.*, 2020; Winsor *et al.*, 2016; Wurtzel *et al.*, 2012).

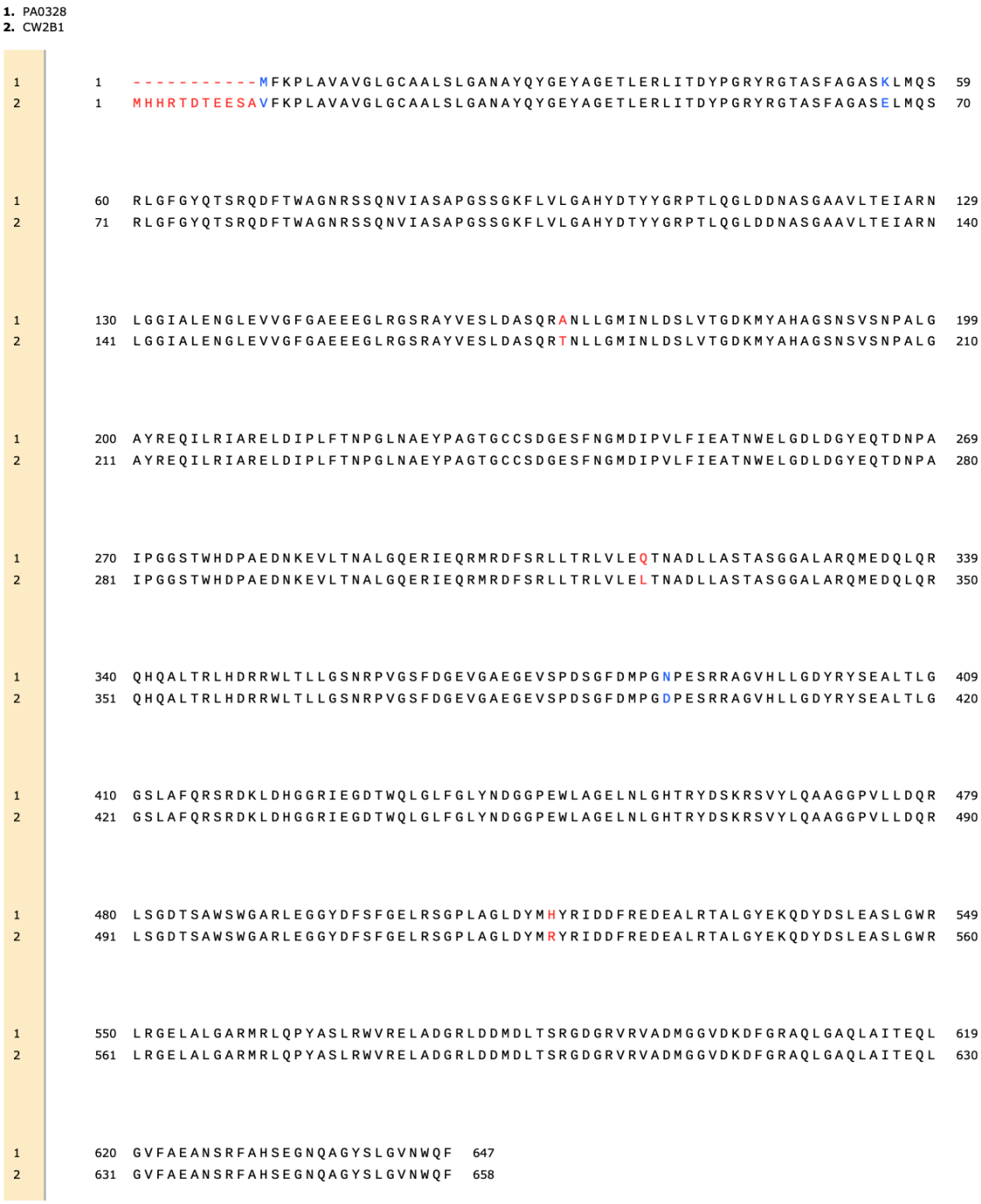


Figure 5.7 Amino acid alignment between PA0328 (1) and* aaaA *from CW2-BI1 (2). *Variant amino acids are highlighted as either blue (similar residue) or red (not similar residue). Aligned using Snapgene.*

5.3.2.6 AaaA activity is intrinsically low in CW2-BI1

To assess the successful deletion of AaaA from CW2-BI1, a whole cell activity assay was carried out, as in 5.3.2.3. CW2-BI1 lacked intrinsic AaaA activity (**Figure 5.8**) and curiously, this was not rescued by complementation with miniCTX*::aaaA*. This suggests that AaaA is being directly suppressed in this strain, as this complementation vector retains the *aaaA* promoter.

Moreover, CW2-BI1 AaaA contains the as-yet unexplored amino acid variation, A167T (**Figure 5.7**). PROVEAN analysis was undertaken as in 3.4 and A167T was predicted to be neutral. However, cloning AaaA from CW2-BI1 into an inducible vector such as pET21a should be used to confirm whether the presence of this, or other variations, impact upon AaaA activity.

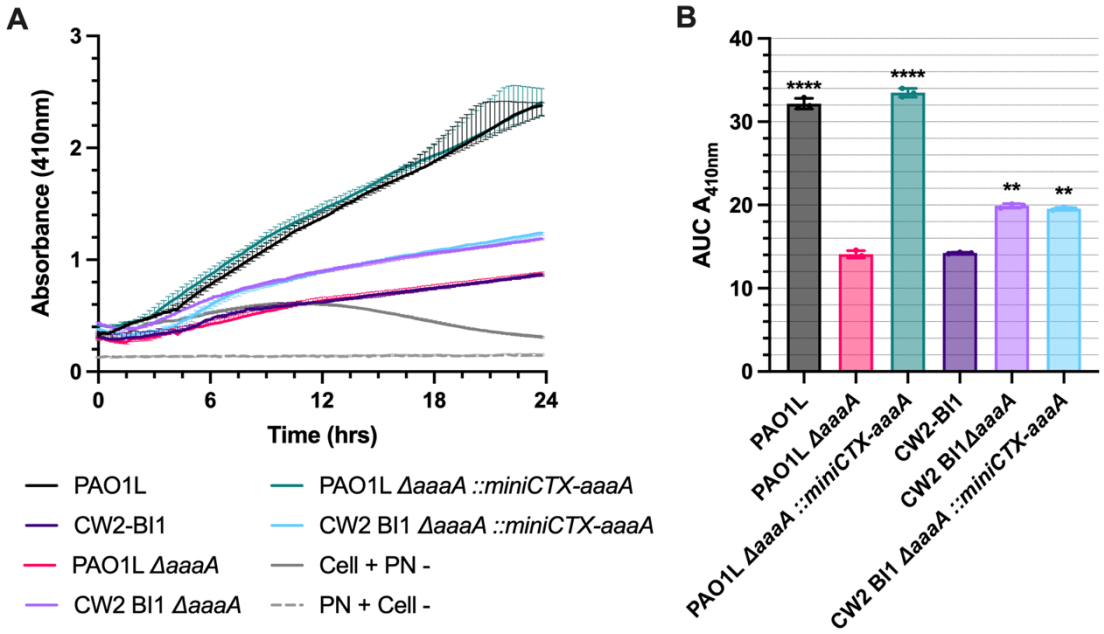


Figure 5.8: AaaA activity is lost in PAO1L ΔaaaA, but intrinsically low in CW2-BI1, with greater non-specific activity in CW2-BI1 ΔaaaA, which is unchanged by complementation*. Following washing, a final OD_{600nm} of 1.0 whole cell suspension was added to the activity assay and L-arginine-p-nitroanilide (PN) cleavage was measured over 24 hrs.* ***Panel A*** *shows values plotted over time (line at median ±95% CI).* ***Panel B*** *shows area under the curve (AUC), with line at mean ±SD (n=3). Differences were determined by Brown-Forsythe and Welch's ANOVA, with Dunnett's T3 multiple comparisons comparing each group to PAO1L ΔaaaA. **p<0.005, ****p<0.0001. cell+ PN- control contains cells from the PAO1L without L-arginine p-nitroanilide and PN+ cell- control contains L-arginine p-nitroanilide only (no cells). Data generated by Bethan Roberts.*

5.3.2.7 Search for an alternative clinical wound isolate with higher intrinsic AaaA activity

An upside of this finding is that it provided the opportunity to construct a further knock out in another clinical isolate from the same strain collection, providing both a 'low' and 'high' AaaA activity variant for comparison. In order to select one with high intrinsic AaaA activity, a whole cell AaaA activity assay was carried out on other isolates. Preliminary results, shown in **Figure 5.9**, identified a number of potential candidates. One of those that showed high levels of AaaA activity was CW4-T1, which as well as CW2-BI1, was also co-

isolated with a *S. aureus*, which could be used in future studies in a polymicrobial context.

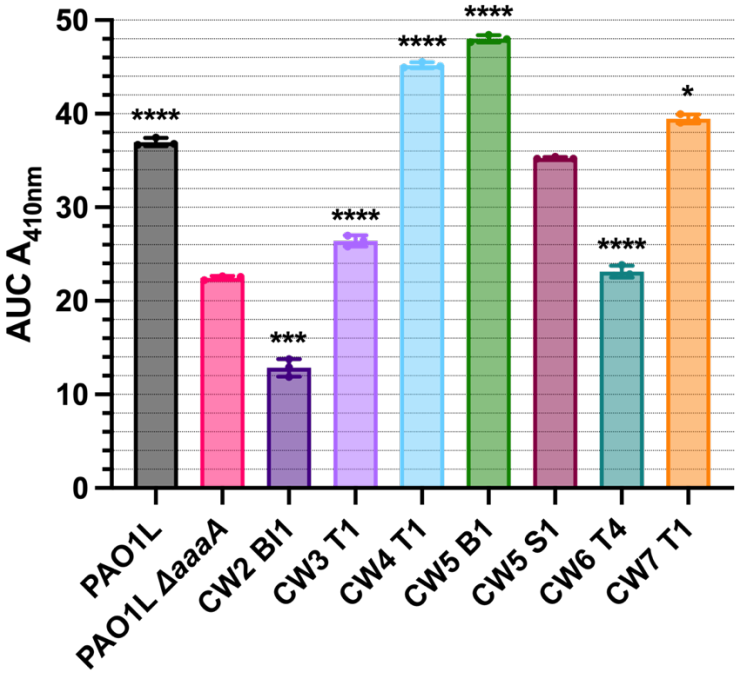


Figure 5.9: AaaA activity varies between P. aeruginosa chronic wound isolates. *L-arginine-p-nitroanilide cleavage by AaaA in PAO1L, PAO1L ΔaaaA, and seven clinical strains isolated from chronic wounds (CW 2-7) was measured over 24 hrs. Planktonic cells in exponential phase in LB were washed in MMP, a final OD_{600nm} of 1.0 was added to the activity assay. Differences were determined by Brown-Forsythe and Welch's ANOVA, with Dunnett's T3 multiple comparisons comparing each group to PAO1L wild type. *p<0.05, ***p<0.001, ****p<0.0001. Data generated by Bethan Roberts.*

The sequences for AaaA in each clinical isolate, the WGS of which were provided by Dr Ana da Silva, were interrogated by Bethan Roberts using PATRIC, and variations from PAO1 are shown in **Table 5.5**. Along with CW2-BI1, isolates CW3-T1 and CW6-T4 also showed near-knock-out levels of AaaA activity. CW6, like CW2, had the H516R variation. As shown in **Figure 3.8**, H516 sits at the top of the membrane-spanning β-barrel. Both histidine and arginine are positively charged, although the positive charge in histidine comes from its imidazole side chain, whereas in arginine it comes from a guanidino group. This difference in structure could conceivably have a marginal impact for example on the interaction of the β-barrel with other membrane constitutions such as LPS. However, a more complete in-membrane 3D model would be required to know for sure. In contrast though, other clinical isolates known to carry the H516R substitution, such as CL10 (**Figure 3.13**) have near-

PAO1 wild-type activity levels. CW6 did also carry the D580E substitution shown in 3.6.1 to have a moderately detrimental effect on AaaA activity when over-expressed in *E. coli*. However, this mutation is not present in CW2 or CW3, so cannot explain their low levels of AaaA activity.

Table 5.5: Amino acid variations in AaaA in clinical wound isolates. *Data generated and collated by Bethan Roberts.*

Strain	***Amino acid variation***								
	T38P	K55E	A167T	N195S	Q314L	N386D	H516R	D580E	S586N
CW2 BI1		✓	✓		✓	✓	✓		
CW3 T1		✓		✓	✓				
CW4 T1		✓		✓	✓				
CW5 B1	✓	✓		✓	✓				
CW5 S1		✓		✓	✓				
CW6 T4		✓			✓		✓	✓	
CW7 T1		✓		✓	✓				✓

Taken together, these data support the conclusions made in 3.7, that no single acid substitution is likely to account for the differences in AaaA activity among clinical isolates, but rather differences are driven by either the combination of substitutions, or more likely, differing levels of *aaaA* expression. It is also possible that the levels of AaaA secretion differs between strains, and interrogation of the Sec, chaperone and Bam-encoding genes could be carried out to further investigate this. An effect from transcriptional regulators, rather than *aaaA* sequence mutations, is supported by the lack of AaaA activity in the complemented CW2-BI1 *aaaA* mutant, as the complementation vector contains PA0328 and its promoter, and there were still intrinsically low AaaA activity levels, likely due to transcriptional suppression. Therefore, to examine AaaA in a context where it was likely to be more highly expressed, a synthetic chronic wound model was employed.

5.4 Development of the synthetic chronic wound model to study AaaA

5.4.1 Introduction to the synthetic chronic wound model

Besides showing AaaA to be essential for survival where tripeptide is the sole carbon/nitrogen source, previous work has not detected a strong

phenotype for Δ*aaaA* under laboratory conditions (i.e., planktonic aerobic growth in LB media). This may be related to the use of the PAO1N subline as a background, or possibly more likely be due to the growth conditions. AaaA was found to be highly abundant in the extracellular biofilm matrix (Toyofuku *et al.*, 2012), which suggests that AaaA is predominantly produced in biofilms rather than planktonic growth.

As the aims of Chapter 5 and 0 were to assess the expression and function of AaaA, and identify co-regulated genes, it was important to do this in an environment where AaaA was likely to be relevantly produced. It was shown by Luckett (2012) that AaaA is most relevant in the chronic, but not acute, stages of a murine wound model. Thus, a synthetic chronic wound model (SCW) model was employed in this study (**Figure 5.10**), described previously by Harrison and colleagues (2017) and da Silva (2018), which were developments upon the model from Werthén and colleagues (2010).

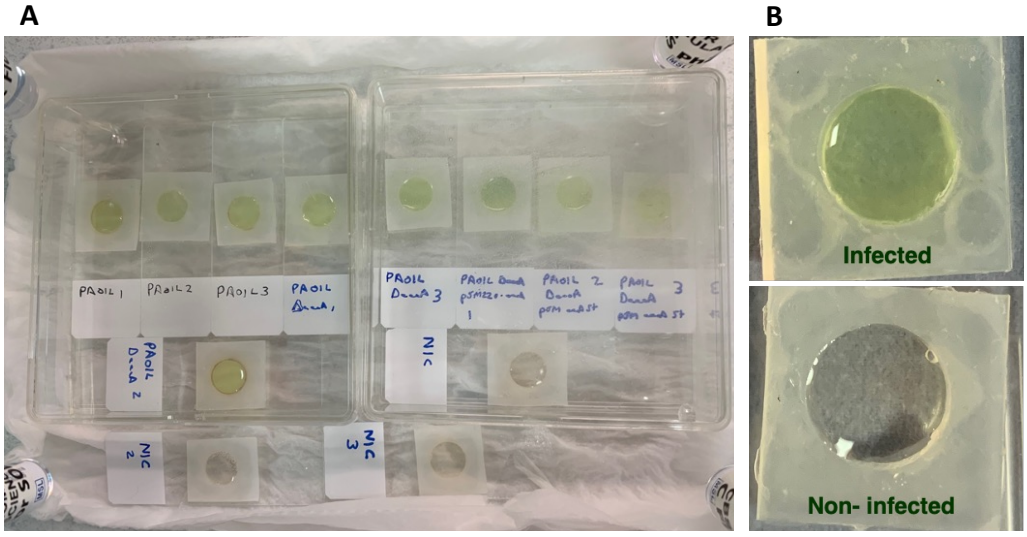


Figure 5.10: The SCW model used in this study, infected with PAO1L wild type and Δ*aaaA*. *The SCW (slide format) shown here was incubated statically for 16 hrs at 37 ℃, in covered square dishes inside a sterilised, humidity-tight box (**Panel A**). **Panel B** shows a close-up, highlighting the visible difference in infected vs uninfected SCW plugs.*

The SCW model was originally designed to recreate, with simplicity, the major matrix, and chemical components of a chronic wound bed. The matrix is made from type I rat tail collagen, and the wound exudate is made from synthetic wound fluid (SWF), which consists of foetal bovine serum, peptone and sometimes horse or sheep blood. Chronic wound models have been

subject to a great deal of development and advancement over recent years, and have been reviewed extensively elsewhere (Thaarup & Bjarnsholt, 2020). For this study simplicity was prioritised, with the potential for further iterations of increasing complexity in future work. For example, though relevant to a chronic wound, the blood component was not added in this study as red blood cells can complicate imaging due to their autofluorescence, and consistency across experiments was preferred. However, inclusion of blood and other components such as cholesterol would be useful to explore in future work.

5.4.2 AaaA confers a growth advantage in 24 hr SCW biofilms

To assess the impact of loss of *aaaA* in the SCW, growth of the PAO1L wild-type and Δ*aaaA* mutant were compared by counting CFUs in the SCW following incubation. After 24 hrs, significantly higher (~1.3-log) viable cells were recovered from samples infected with wild-type than Δ*aaaA* (**Figure 5.11**).

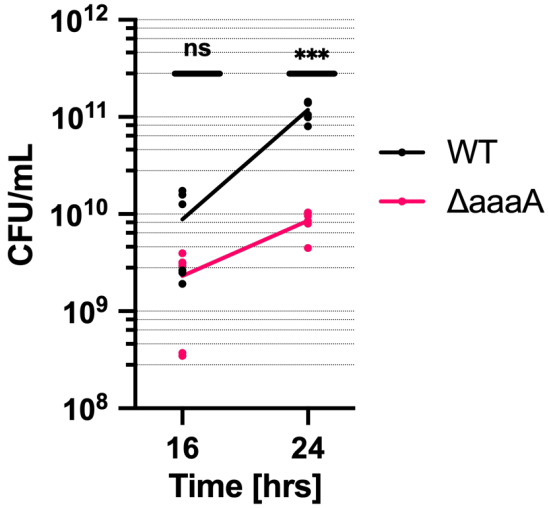


Figure 5.11: Colony forming units provide evidence that aaaA confers a survival advantage at 24 hrs in the SCW. *Data are mean ±SD (n=6, 3 biological, 2 technical) CFU/mL recovered from the SCW (slide mounted) after 16 and 24 hrs, comparing PAO1L wild type (black) and* ΔaaaA *mutant (pink). Differences were determined using Brown-Forsythe and Welch's ANOVA test with Dunnett T3 multiple comparisons; only differences between strains are displayed, though all were compared (*p≤0.05, ***p≤0.0005).*

This was further investigated using simple live/dead staining with SYTO-9 and propidium iodide (PI) in 8-well chambers as described in 2.10.1. There was a significantly higher proportion of live cells in the wild-type infection than

Δ*aaaA* after 24 hrs (**Figure 5.12**). Non-infected controls were also examined, and no significant autofluorescence was observed (not shown).

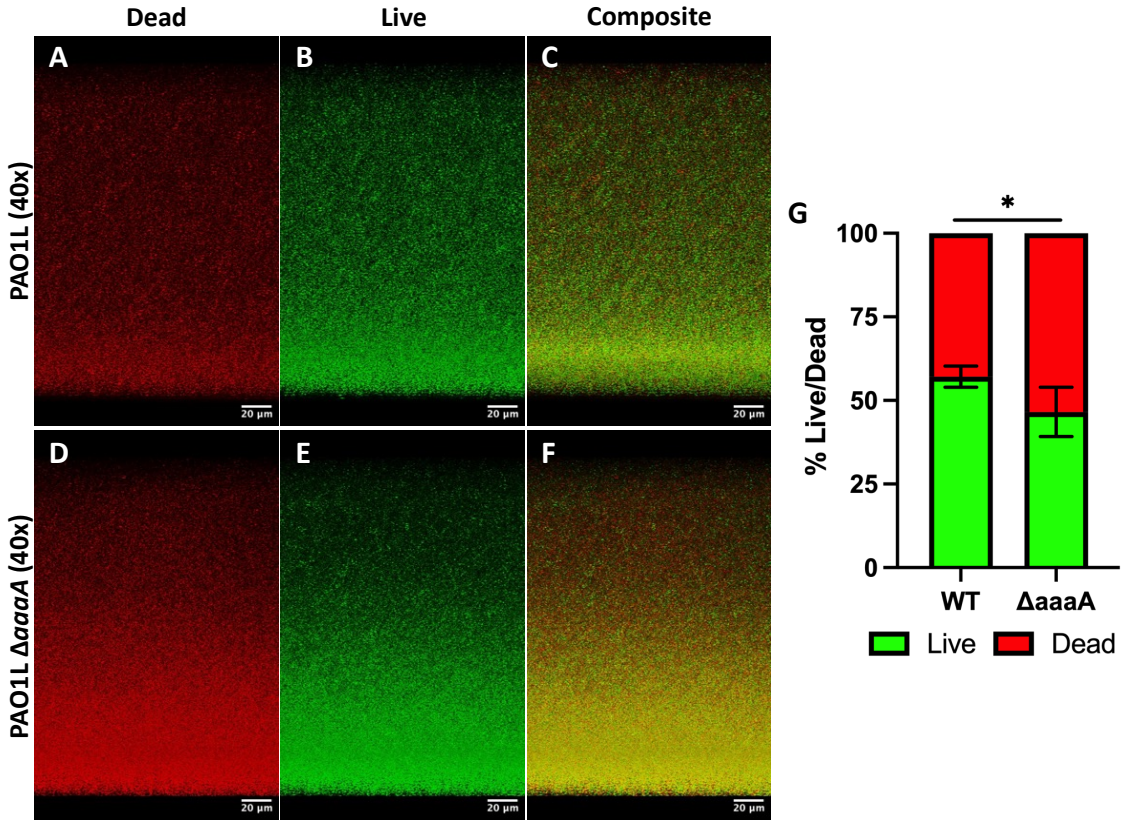


Figure 5.12: Live/Dead staining shows* aaaA *confers a survival advantage at 24 hrs in the SCW. Panel A-D *show 3D projections of representative respectively PI (dead) and SYTO9 (live) stained SCW, infected with either wild-type (WT) PAO1L (**A-C**) or PAO1L* ΔaaaA *(**D-F**) at 40x magnification. Scale bars are all 20 µm. **Panel G** shows the mean (±SD) ratio of live: dead stained cells of as a percentage (live above, dead below) calculated from 3 biological repeats of each condition, imaging 2 points per well and compared using a Welch's t-test (n= 6, *p= 0.0159).*

There were also some indications of a more diffuse pattern of cell distribution in the Δ*aaaA* mutant, suggesting poorer biofilm formation, however it was not possible to draw conclusions on biofilm structure from these data. Further experiments specifically staining for biofilm components such as exopolysaccharides or eDNA, or by crystal violet quantification, would be required to determine this. A role for AaaA in biofilm formation is supported by preliminary findings by Spencer (2018), who found that when grown in artificial sputum medium (ASM), both in falcon tubes, and in a BioFlux flow system, wild-type PAO1 formed larger aggregates, whereas the *aaaA* mutant was more diffuse with sporadic smaller aggregates. Therefore, given the potential

importance of AaaA in assisting with survival within a biofilm, this should be explored further.

5.5 Examining *aaaA* transcription and translation in the SCW

5.5.1 *aaaA* promoter activation is higher in the SCW than in LB, but previously observed autorepression was not replicated

As discussed in 5.1.2.3, it has been consistently observed previously that *aaaA* promoter activation is higher in *aaaA* mutants, indicative of direct or indirect auto-repression by AaaA. To examine *aaaA* expression and auto-repression in the SCW, activation of P_{aaaA} and P_{kan} were compared between PAO1N, PAO1L and CW2-BI1 in both the wild type and Δ*aaaA* mutant, and in the SCW and LB, in a 96-well plate, as described in 2.6.

Though there is a high degree of variability, in general, *lux* expression from both promoters has a broader peak in the SCW than LB, as well as a lag in maximum activation (**S-Figure 0.4**), suggesting *lux* expression is non-specifically more sustained in the SCW. **Figure 5.13** shows the area under the curve for each strain and interestingly, the predicted auto-repression of P_{aaaA} is lost in both PAO1L and PAO1N, but not in CW2-BI1. CW2-BI1 also appears to have higher levels of luminescence from both P_{kan} and P_{aaaA} activation.

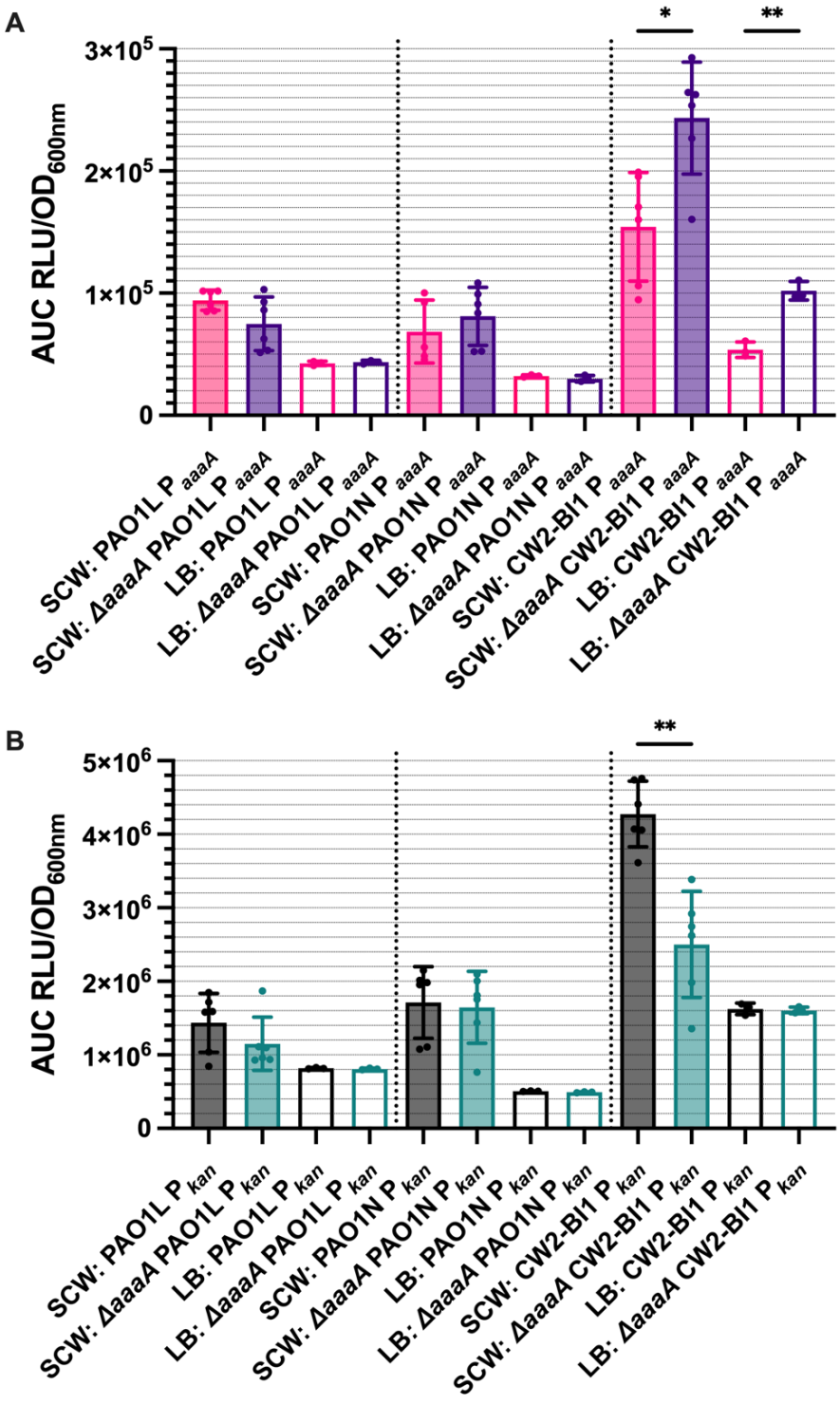


Figure 5.13: Activation of both P_{aaaA} and P_{kan} is higher in the SCW and P_{aaaA} auto-repression is only present in CW2-BI1. Panel A *and* ***B*** *show the area under the curve (AUC) for P_{aaaA} and P_{kan}, respectively, from data shown in* ***S-Figure 0.4****. AUC for each replicate group is plotted, with bars at mean ± SD. (For SCW, n=6, 3 technical replicates from 2 biological replicates on different days, for LB n=3). Differences were determined using Brown-Forsythe and Welch's ANOVA test with Dunnett T3 multiple comparisons comparing between the wild-type and ΔaaaA in each condition (*p≤0.05, **p≤0.005). Data collected with assistance from Bethan Roberts.*

Activation of P_{kan} is significantly higher in the CW2-BI1 wild type compared to the Δ*aaaA* mutant. This could be interpreted as a metabolic advantage in the wild-type over the *aaaA* mutant in the SCW, as being constitutively expressed, P_{kan}-*lux* expression is indicative of ongoing cell metabolism. However, this could also be due to non-specific interference with *lux* expression, either at the transcriptional level or downstream. In the case of CW2-BI1 though, the directionality of the difference in P_{kan} and P_{aaaA} activation is opposite; in the Δ*aaaA* mutant, P_{aaaA} activation is higher, yet P_{kan} activation is lower. This lends support to the conclusion that the variations in P_{aaaA}-*lux* expression are more likely to be related to increased *aaaA* expression, rather than being non-specific effects, as otherwise one would expect the differences to be consistent between promoters.

It is unclear why the previously observed (See **Figure 3.14**) difference in P_{aaaA} activation in the *aaaA* mutant was not observed here in either PAO1 strain. Experiments shown in **Figure 5.13** were collected in a 96-well plate, incubating statically, covered by a gas-permeable seal to retain humidity. Previous experiments have typically been incubated with a plastic lid, rather than adhesive seal. Though difficult to determine to exact oxygen concentration in each well, there is some evidence that use of these particular gas-permeable seals reduce aeration to a greater extent that use of a lid in this type of experiment (Dimitra Panagiotopoulou, University of Nottingham, unpublished data, 2021). This may have impacted either P_{aaaA} regulation, or *lux* expression in general. To test this, the experiment was repeated with a lid instead of a seal. As shown in **Figure 5.14**, the general patterns and magnitudes of *lux* expression seen in **Figure 5.13** are replicated, suggesting that the seal cannot explain the lack of auto-repression.

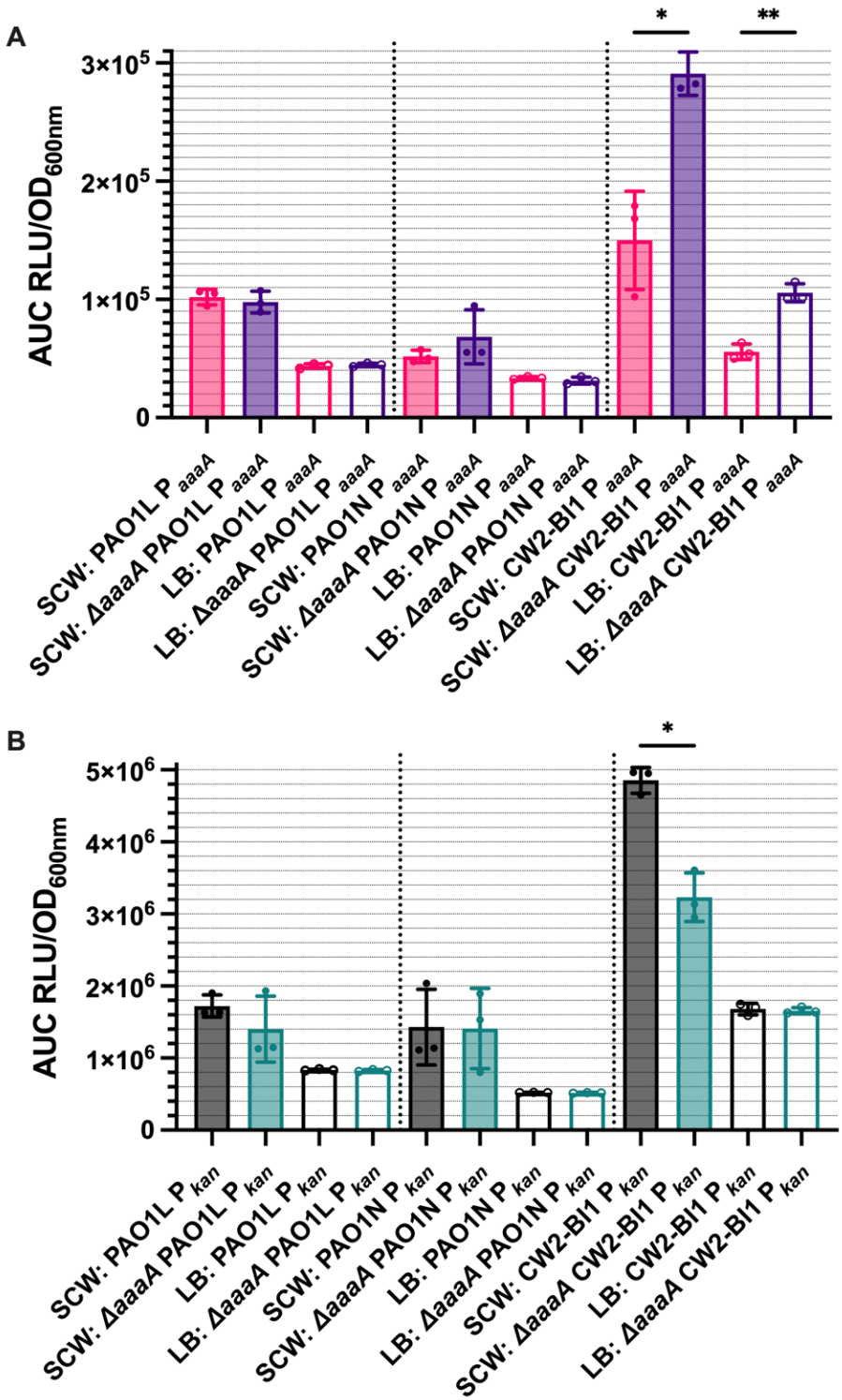


Figure 5.14: Use of a lid instead of a gas-permeable seal makes a negligible difference to P_{aaaA} or P_{kan} lux expression. Panel A *and* ***B*** *show the area under the curve (AUC) for P_{aaaA} and P_{kan}, respectively, using the same method as in* ***S-Figure 0.4****, but with a plastic plate lid instead of a gas-permeable seal. AUC for each replicate group is plotted, with bars at mean ± SD. (n=3). Differences were determined using Brown-Forsythe and Welch's ANOVA test with Dunnett T3 multiple comparisons comparing between the wild-type and ΔaaaA in each condition (*p≤0.05, **p≤0.005). Data collected by Bethan Roberts.*

It may be that a disruption in the CTX site in the PAO1N strain has abrogated *lux* expression over time. Re-sequencing and/or re-conjugation of the reporter plasmid into the background strains could assess this. Moreover, this study lacks an empty miniCTX-*lux* control, where there is no promoter controlling the *lux* genes. Inclusion of this control in future work would provide a background level of 'leaky' *lux* expression in each strain, to give a better indication of the levels of 'true' expression above a baseline.

5.5.2 AaaA activity is higher when grown in SCW than planktonically

Despite the intrinsically low to negligible AaaA activity levels in CW2-BI1 when grown planktonically (**Figure 5.8**), the large level of activation of P_{aaaA} suggests that in the SCW, CW2-BI1 may express *aaaA*. To test this, AaaA activity in both PAO1L and CW2-BI1 was compared when grown planktonically (as in **Figure 5.8**) using a modified L-arginine p-nitroanilide cleavage assay as described in 2.7. Due to the more complex nature of the samples, and the need to compare with planktonic conditions from a different plate, the optical density of samples could not be exactly normalised as they were added to the assay as is typically done. Therefore, data are instead normalised by dividing the absorbance at 410 nm by the absorbance at 600 nm at each timepoint, similarly to the RLU/OD_{600nm} normalisation used in the *lux*-transcriptional reporter assays.

As shown in **Figure 5.15**, PAO1L showed over 3-fold higher AaaA activity (L-arginine-p-nitroanilide cleavage) levels in both the 24 hr and 48 hr SCW compared with planktonic conditions. L-arginine-p-nitroanilide cleavage in CW2-BI1 was similarly significantly higher in the SCW than planktonic conditions, with mean cleavage 5.4-fold and 3.4-fold higher than planktonic in the 24 and 48 hr SCW respectively. There are also higher levels of L-arginine-p-nitroanilide cleavage by the CW2-BI1 Δ*aaaA* mutant in the 24 hr SCW, indicating that non-specific proteolytic cleavage is also occurring at this time. L-arginine-p-nitroanilide cleavage was lower in all strains at 48 hrs compared with 24 hrs, potentially indicative of lower proteolytic activity in general at this time point. These findings, taken together with the evidence from the P_{aaaA}

transcriptional reporters (**Figure 5.13**), provide strong evidence that *aaaA* is preferentially, or in the case of CW2-BI1, exclusively expressed in the SCW compared with planktonic conditions, as indicated by the production of active AaaA. These data further support the use of the SCW as an ideal model to study AaaA in a disease-relevant context.

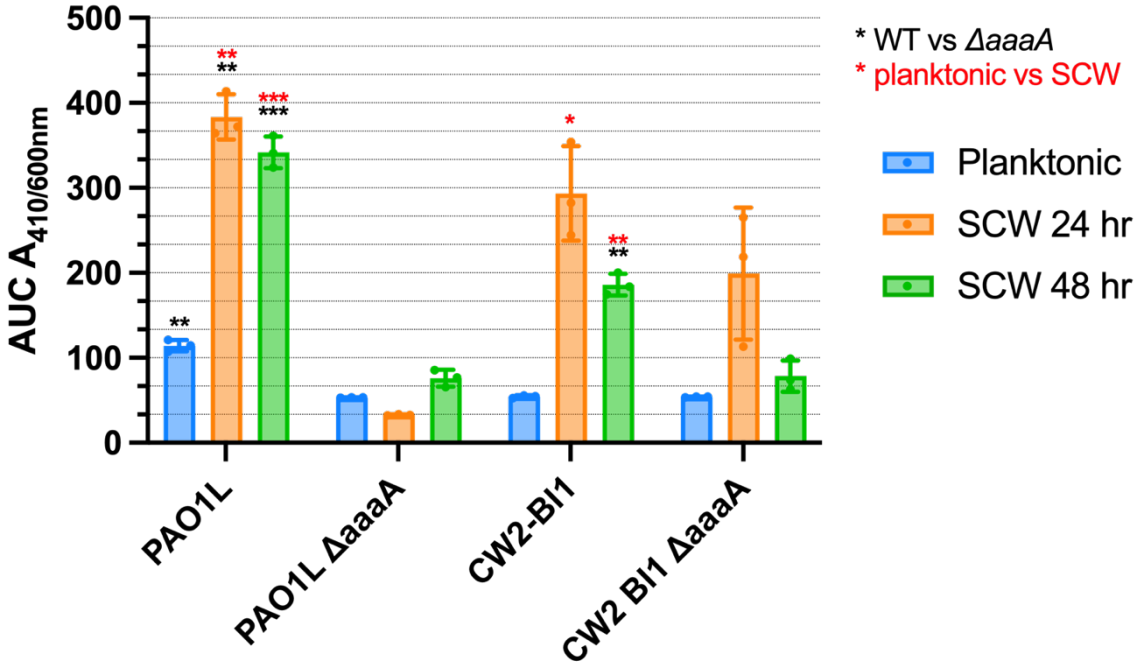


Figure 5.15: Comparison PAO1L and CW2-BI1 AaaA activity in planktonic and synthetic chronic wound contexts. *Cells were grown either planktonically in LB, as in* ***Figure 5.8****, or in the SCW for 24 or 48 hrs as described in 2.7, and 100 µL added to the activity assay (n=3 biological replicates). Differences were determined by Brown-Forsythe and Welch's ANOVA test with Dunnett T3 multiple comparisons, comparing both wild-type with* aaaA *mutant strains. Significant difference between wild type and* aaaA *mutant are shown in black and between growth conditions (wild-type only) shown in red (*p≤0.05 **p≤0.005, ***p≤0.001). Data collected by Bethan Roberts.*

While *lux* transcriptional reporters are indicative of *aaaA* expression over time, they are limited in their temporal range to a maximum of 18 hrs, after which their activities decrease, and a combination of cell aggregation and lysis makes optical density readings erratic. Meanwhile, the L-arginine-p-nitroanilide AaaA activity assay has a more flexible temporal range but is susceptible to detection of non-specific peptidolytic activity, which can make distinguishing AaaA activity from other proteolytic enzymes difficult. It was therefore necessary to determine *aaaA* expression with a high degree of specificity, over a 16-48 hrs period and beyond. To that end, RT-qPCR and

transcriptomics were employed, firstly to reliability identify when in the SCW *aaaA* is most highly transcribed, and secondly to identify co-regulated genes which could give clues to AaaA function.

5.6 Development of RT-qPCR for monitoring *aaaA* expression in the SCW

RT-qPCR (reverse transcriptase, followed by real-time, quantitative PCR) is a molecular biology technique used widely to quantify, in both absolute and relative terms, the amount of a given target RNA in a sample. It can be one-step, where RNA is added directly to the reaction plate, and is converted into cDNA *in situ* by a reverse transcriptase step in the reaction protocol. The advantage is that it is more rapid thus more amenable to high throughput screening, such as in diagnostic lab settings, including in the mass-PCR testing seen worldwide, during the COVID-19 pandemic (Yelin *et al.*, 2020). However, as RNA is liable to degradation, conversion into cDNA on the bench (two-step) allows for longer-term storage of the sample, making it more flexible for research applications. It also allows for standardisation of DNA quantity between samples, as the quantity of RNA added to the reverse transcription can be easily normalised. For these reasons, two-step RT-qPCR was used in this study.

5.6.1 Primer design and validation

5.6.1.1 Selection of genes and primer design

Primers for *aaaA* were previously designed by Spencer (2018) and the set *aaaA*7 (**Table 7.3**) were selected as the preferred pair. Due to the difficulty of identifying and verifying one constant endogenous control, both *rpoS* and *rpoD* were chosen as they have been previously shown to be relatively constantly expressed in biofilms and provide representatives of both exponential and stationary phase (Savli *et al.*, 2003). Additionally, a panel of seven genes was chosen for investigation, due to their known involvement in arginine sensing, uptake, and catabolism, thus likely link to AaaA. These included *rmcA*, *argR*, *arcD*, *aruF*, *speA*, *aruH* and *arcA*, each of which are discussed in 1.2.

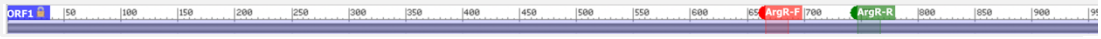
Primers, shown in **Table 5.6** were designed using the Integrated DNA Technologies qPCR design tool and checked using NCBI ORF finder, to minimise potential alternative ORFs. Primer sequences were also checked against the *P. aeruginosa* WGS using Snapgene, to ensure there were no off-target binding sites.

Table 5.6: Genes selected for RT-qPCR analysis and their primer binding sites.
Location of forward and reverse primers are highlighted in red and green, respectively.

rmcA: arginine sensing (PA0575)



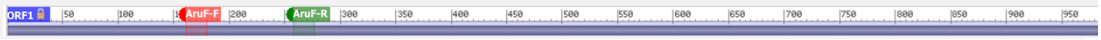
argR: arginine response regulator (PA0893)



arcD: arginine uptake (PA5170)



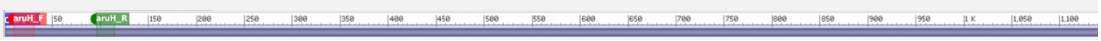
aruF: AST pathway (PA0896)



speA: ADC pathway (PA4839)



aruH: ATA pathway (PA4976)



arcA: ADI pathway (PA5171)



5.6.1.2 Primer efficiency testing

Primer efficiencies were tested using an initial primer concentration of 250 nM, with PAO1L gDNA as the template, with a 1:10 standard curve starting at 400 pg/µL (~6.223 x10^4 copies) down to 0.04 pg/µL (~6.2 copies), as described in 2.9.1.3.1. **Figure 5.16** shows the standard curves for the targets that were successfully amplified. Primers for *aruH* and *arcA* did not produce amplicons and efficiencies for the others are generally lower than the ideal >90%. However, this was accounted for in the analysis by adjusting the expected efficiencies in the ABI7500 software to the values displayed (rounded to 2 decimal places) in **Figure 5.16**.

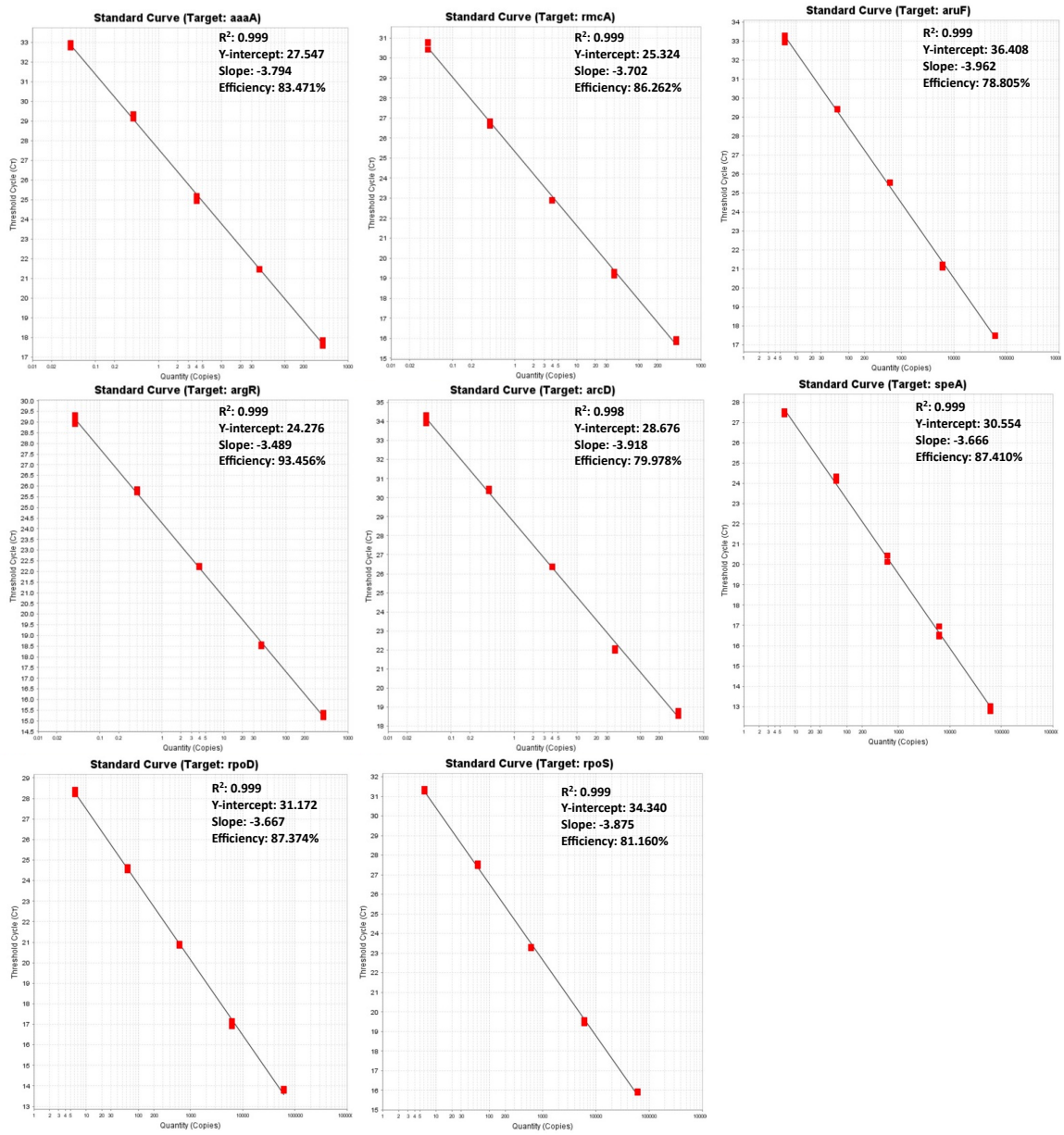


Figure 5.16: Standard curves for selected primer sets on PAO1L gDNA. *R^2 for each standard curve was above 0.998, slopes ranged from -3.489 to -3.961 and efficiencies ranged from 78.8% to 93.5%.*

5.6.2 *aaaA* is transcribed in the SCW at a low level, and is greatest at 16 hrs

Using both *rpoD* and *rpoS* as endogenous controls, the relative mRNA level of *aaaA* was quantified in RNA extracted from PAO1L grown in the SCW (in silicon moulds on glass slides, described in 2.8.1). RNA extraction, cDNA preparation and RT-qPCR are described in 2.9.1. As shown in **Figure 5.17**, transcripts for *aaaA* could be detected in the SCW at 16, 24 and 40 hrs, with the highest relative level at 16 hrs.

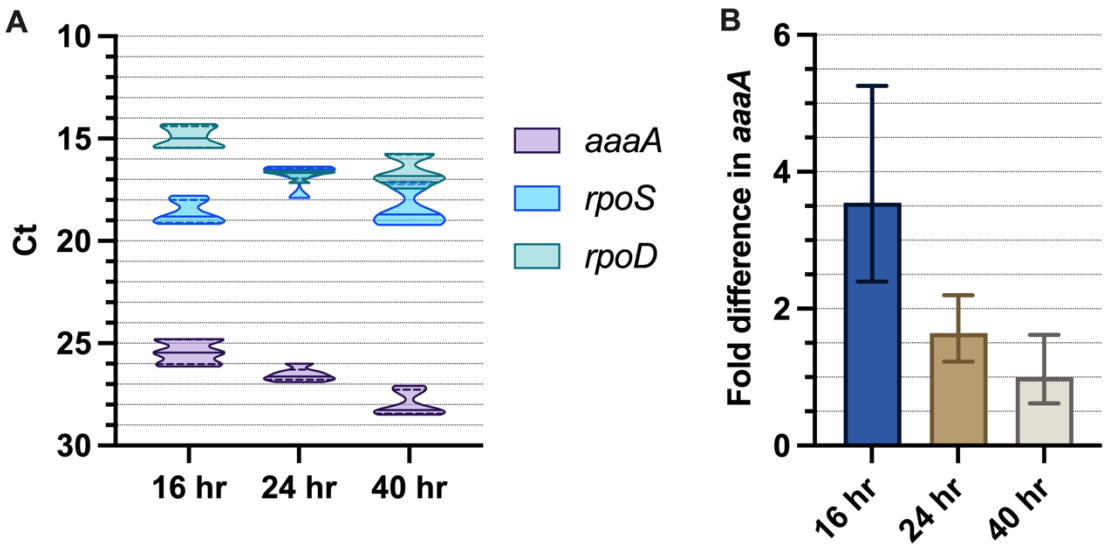


Figure 5.17: RT-qPCR of SCW for relative quantification of* aaaA *over time shows* aaaA *is most highly transcribed at 16 hrs. Panel A *is a truncated violin plot of the Cts for each sample (median as line, quartiles as dashed lines, n=6, 2 technical, 3 biological).* rpoS *and* rpoD *were used as endogenous controls to calculate the relative quantity of* aaaA *at each time point, shown in* ***Panel B*** *as fold difference (40 hr used as an arbitrary reference). ΔΔCt for each time point was -1.8277, -0.7163 and 0 for 16 hrs, 24 hrs and 40 hrs, respectively.*

5.6.3 Arginine-related genes are not differentially expressed in the *aaaA* mutant-infected SCW, but *rpoS* is upregulated

As *aaaA* mRNA levels were highest at 16 hrs, this time point was chosen to compare the mRNA levels of the panel of arginine-related genes between PAO1L and PAO1L Δ*aaaA,* as shown in **Figure 5.18***.* Surprisingly, *rpoS*, which was initially chosen as an additional endogenous control, was found to be significantly upregulated in the Δ*aaaA* mutant (mean fold-change of 3.08). In fact, it was the only gene in the panel that was significantly differentially expressed between the two strains. There were also trends of downregulation of *argR* and upregulation of *aruF* and *arcD* in the *aaaA* mutant, but the differences were not significant, and the high level of variability necessitates further replicates.

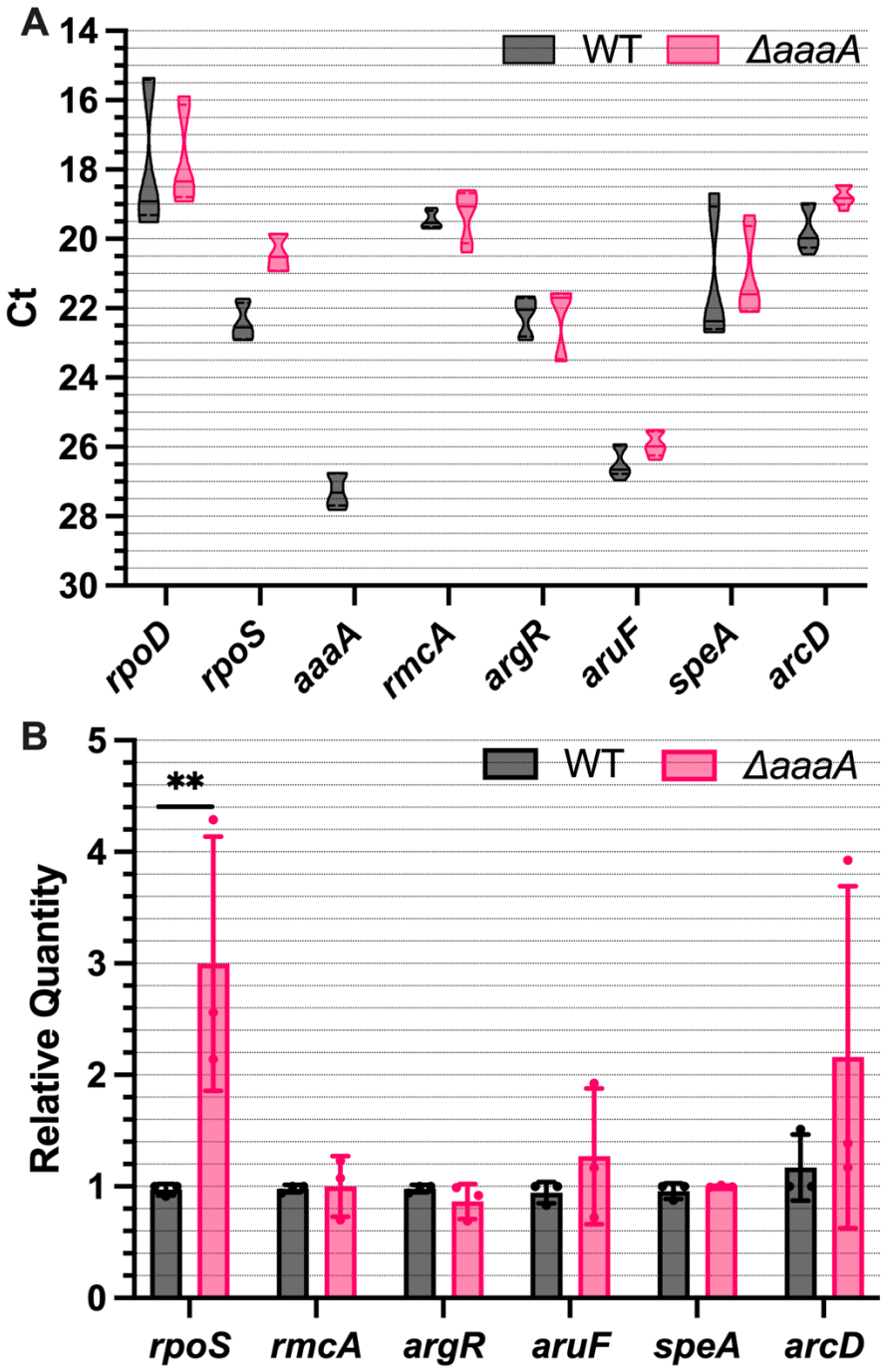


Figure 5.18: RT-qPCR results comparing PAO1L and PAO1L ΔaaaA in the SCW after 16 hrs show* rpoS *is upregulated in the mutant. Panel A *shows a truncated violin plot of Ct values for each sample and target (n = 6, 3 biological x 2 technical).* ***Panel B*** *shows the Relative Quantity of each gene (with 3 biological replicates, each an average of 2 technical replicates) in the wild-type vs mutant, using* rpoD *as the endogenous control (mean ±SD). Differences were determined by One-Way ANOVA test with Dunn's multiple comparisons comparing between wild-type and mutant groups for each target (**p<0.005).*

A BLAST-N search was carried out comparing the sequences for *rpoS* between the wild-type and mutant, to check for any mutations in the region. The nucleotide sequence for *rpoS* ± 500 bp up and downstream was copied

from the Pseudomonas.com database and searched for in the WGS assemblies of both PAO1L and PAO1L Δ*aaaA* using PATRIC BLAST as in 5.3.2.3 (Winsor *et al.*, 2016; Davis *et al.*, 2020). Both sequences had 100% sequence coverage and 100% identity to the query sequence, meaning there were no mutations at all in either strain in this region which could cause changes in transcription, mRNA stability, or RT-qPCR primer binding.

5.7 Using RNA-Seq to further study SCW infection transcriptomics in relation to *aaaA*

5.7.1 Introduction and methodology

To further investigate the impact of deletion of *aaaA* in the SCW, RNA-Sequencing (RNA-Seq) was utilised. This broader, whole transcriptomic approach allowed for the identification of previously unknown or unexplored genes which could be connected to *aaaA*. As with the RT-qPCR, due to the higher levels of *aaaA* expression at 16 hrs, this was the time point chosen for further investigation by RNA-Seq.

5.7.1.1 Sampling methodology

The hypothesised role of AaaA is that it releases arginine from peptides as a local resource for cells. To examine the transcriptional response of cells to small increases in extracellular arginine concentrations, like those hypothesised to be created by AaaA, an arginine supplementation condition was included. The physiological concentration of L-arginine in human sera/plasma ranges from ~40 μM to ~113 μM (2.5^{th} and 97.5^{th} percentile) (Lüneburg *et al.*, 2011), but can reach up to 250-1500 μM for patients with argininemia (Abcam, 2019a). Quantification of the extracellular L-arginine in this SCW, described later in 6.3.1, also found the concentration of arginine in a 200 μL plug of PAO1L-infected SCW at 16 hrs was ~100-200 μM (**Figure 6.9**).

Therefore, supplementation by the addition of a 50 μL aliquot of 100 μM L-arginine was chosen to simulate a modest increase in arginine concentration which remains well within this physiologically relevant range. As the plugs solidify and evaporation occurs over the incubation, it is difficult to give an

exact estimate of the final concentration of L-arginine supplementation, but it is expected to be 20-50 μM, which would constitute a 10-50% increase in extracellular L-arginine. For simplicity, the starting concentration of 100 μM is used throughout this section to refer to the supplemented group.

Samples for the four groups, PAO1L *vs* PAO1L Δ*aaaA*, with and without additional arginine supplementation, were prepared in triplicate, according to 2.9.2.1. RNA extraction, library preparation, sequencing and analysis was completed by a commercial partner, Novogene UK. For data analysis, and to ease randomisation, samples were assigned sample names as follows: PAO1L + SWF (WT1-3), PAO1L + SWF with 100 μM arginine (WT4-6), PAO1L Δ*aaaA* + SWF (Da1-3), PAO1L Δ*aaaA* + SWF with 100 μM arginine (Da4-6)

5.7.1.2 RNA-Seq pipeline

Figure 5.19 outlines the pipelines used by Novogene for library construction, sequencing, QC, and analysis, and these are all described in detail in 2.9.2.2. The data analysis workflow, as illustrated in **Figure 5.19c**, consists of a number of steps which produce different data sets. Not all datasets produced during analysis were interrogated in this work, but the ones which are discussed in this study are briefly described below:

1. **Data QC:** This filters out low quality or adapter-contaminated reads
2. **Mapping clean reads** to the reference genome (PAO1)
3. **Prediction of novel genes**, operons and transcription start sites
4. **Gene expression quantification.** This is calculated as the expected number of Fragments Per Kilobase of transcript sequence per Million base pairs sequenced of each gene (FPKM). This is based on the gene length and read count mapped to this gene, giving an expression quantity for each gene relative to the sequencing depth and coverage.
5. **Differential expression analysis.** This calculates the Log2 fold change (Log2FC) between biological groups, based on FPKM value for each gene. The *p*-values are also calculated and adjusted (padj) for the false discovery rate (FDR).
6. **Functional analysis:** KEGG enrichment analysis assigns predicted functions to uncharacterised genes, as well as determining the biological

pathways which are enriched (i.e., a cluster of genes with similar functions which are all differentially expressed) between conditions.

7. **UTR prediction:** Predicts ShD sequences and other 3' and 5' UTRs.

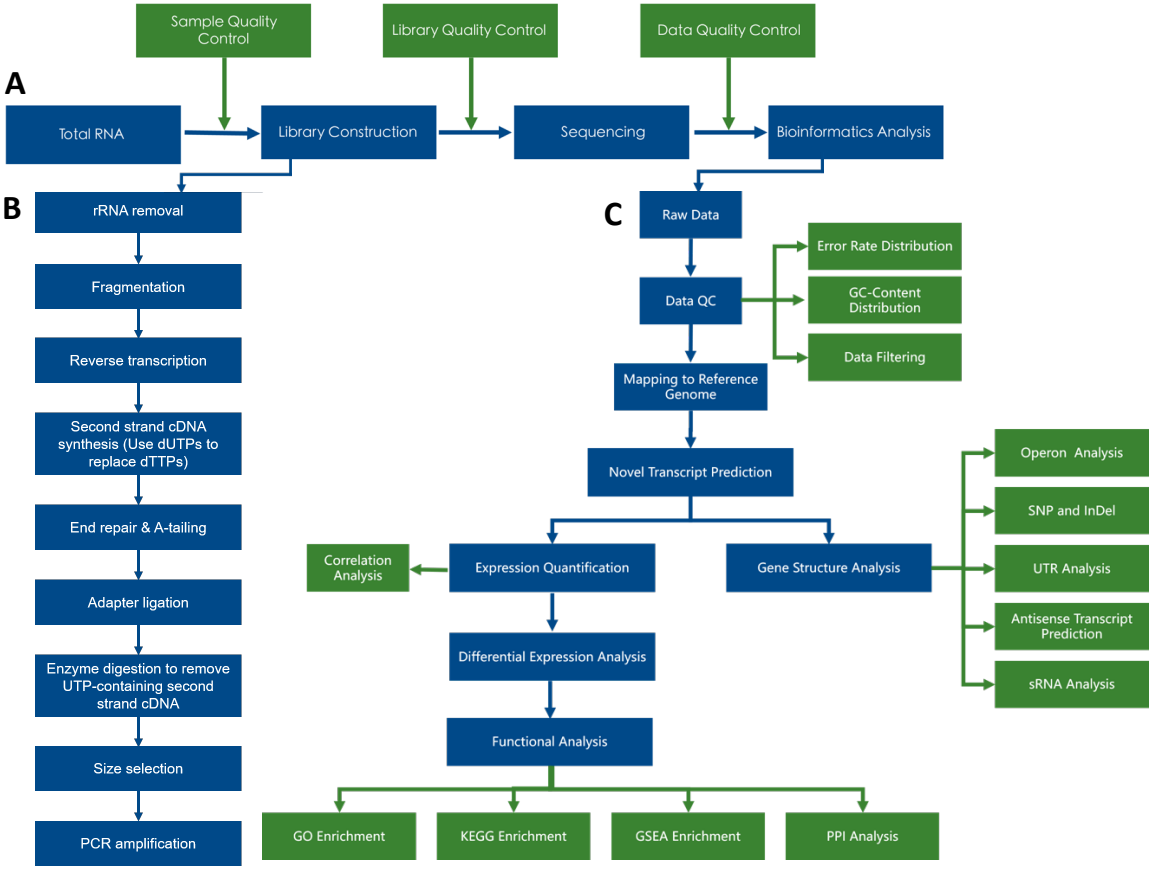


***Figure 5.19: Pipeline for RNA-Seq analysis. Section A** denotes the whole pipeline, from extracted RNA to data analysis. **Section B** provides further detail on the library construction pipeline within this. **Section C** outlines the data analysis pipeline and in green, the data sets generated. Adapted from figure provided by Novogene Co. Ltd.*

5.7.2 RNA-Seq QC and mapping results

5.7.2.1 Sequencing produced high quality reads

Unfortunately, there were high levels of rRNA contamination in one replicate each of the PAO1L group (WT2) and PAO1L Δ*aaaA* + 100 μM L-arginine group (Da5). Those replicates were excluded from analysis, meaning these two groups contain duplicates rather than triplicates.

In all remaining samples, low quality reads constituted <0.01% of all reads, with 1-5% reads across samples removed due to adaptor contamination. For the remaining 10 samples, average single base error rate

was ≤0.03%, indicating high quality sequencing was attained. Full QC analysis is shown in **Table 5.7**, which also shows that 12-23 million clean reads were generated per sample.

Table 5.7: QC of RNA-Seq samples. *'Raw/clean reads' refers to read count from raw and filtered data, respectively. 'Raw/clean bases' = (number of raw/clean reads)* (sequence length), converting the unit to G. Error rate refers to the base error rate of the whole sequencing. Q20 and Q30 refer to the % reads scored (by Phred score) as having a correct rate of 99% and 99.9%, respectively. WT- wild-type PAO1L, Da- PAO1L ΔaaaA, 1-3 were supplemented with SWF, 4-6 supplemented with SWF + 100 μM L-arginine. Data generated by Novogene Co. Ltd.*

Sample	Raw reads	Clean reads	Raw bases	Clean bases	Error rate (%)	Q20 (%)	Q30 (%)	GC content (%)
WT1	16411412	16205388	2.5G	2.4G	0.03	97.78	94.04	61.32
WT3	18865552	18074414	2.8G	2.7G	0.03	97.38	93.51	60.66
WT4	23943560	23181058	3.6G	3.5G	0.03	97.66	93.92	60.4
WT5	23558868	23217312	3.5G	3.5G	0.03	97.84	94.15	61.31
WT6	21141594	20532770	3.2G	3.1G	0.03	97.61	93.91	60.44
Da1	12530194	12297826	1.9G	1.8G	0.03	97.65	93.79	60.71
Da2	15176486	14718324	2.3G	2.2G	0.02	98.1	94.9	59.89
Da3	20612126	20341946	3.1G	3.1G	0.02	98.04	94.56	61.18
Da4	15548952	15087032	2.3G	2.3G	0.02	98.18	94.96	60.63
Da6	15792696	14948534	2.4G	2.2G	0.03	97.72	94.06	61.31

5.7.2.2 Reads generally mapped well to the PAO1 reference genome

As shown in **Table 5.8**, in all 10 samples for analysis, the % total reads mapped to the reference PAO1 genome exceeds 97%, indicating there was no contamination with other species. The % reads mapped to multiple sites was <3.5% in all samples except Da2, where it was 14.34%. This may be due to a higher proportion of shorter reads in this sample, but it is worth noting as multiple mapping reads can make quantification of gene expression more difficult (Deschamps-Francoeur *et al.*, 2020). Given Da2 is part of a triplicate group, any potential bias due to multiple mapping should be outweighed by the other two replicates.

Table 5.8: Mapping of RNA-Seq samples to the reference PAO1 genome. *Generated by Novogene Co. Ltd.*

Sample	Total reads	Total mapped reads	Uniquely mapped reads	Multiple mapped reads	Total mapping rate	Uniquely mapping rate	Multiple mapping rate
WT1	16205388	16100926	15717645	383281	99.36%	96.99%	2.37%
WT3	18074414	17899470	17316018	583452	99.03%	95.80%	3.23%
WT4	23181058	22934204	22326311	607893	98.94%	96.31%	2.62%
WT5	23217312	23073686	22527581	546105	99.38%	97.03%	2.35%
WT6	20532770	20372915	19901284	471631	99.22%	96.92%	2.30%
Da1	12297826	12164307	11811627	352680	98.91%	96.05%	2.87%
Da2	14718324	14560386	12449799	2110587	98.93%	84.59%	14.34%
Da3	20341946	20178453	19769505	408948	99.20%	97.19%	2.01%
Da4	15087032	14693210	14340623	352587	97.39%	95.05%	2.34%
Da6	14948534	14799301	14415652	383649	99%	96.44%	2.57%

5.7.2.3 There was a good degree of replicability between biological replicates

High degrees of correlation between biological replicates are essential for reliability in RNA-Seq experiments. The closer to 1 the Pearson co-efficient (R^2) between samples of the same biological group, the more similar the expression profile within each replicate. **Figure 5.20** shows the R^2 for samples in this experiment. Within the same experimental group (i.e., wild-type *vs aaaA* mutant and arginine supplemented *vs* unsupplemented), the R^2 ranges from 0.905-0.938, indicating a fair degree of replicability between biological replicates.

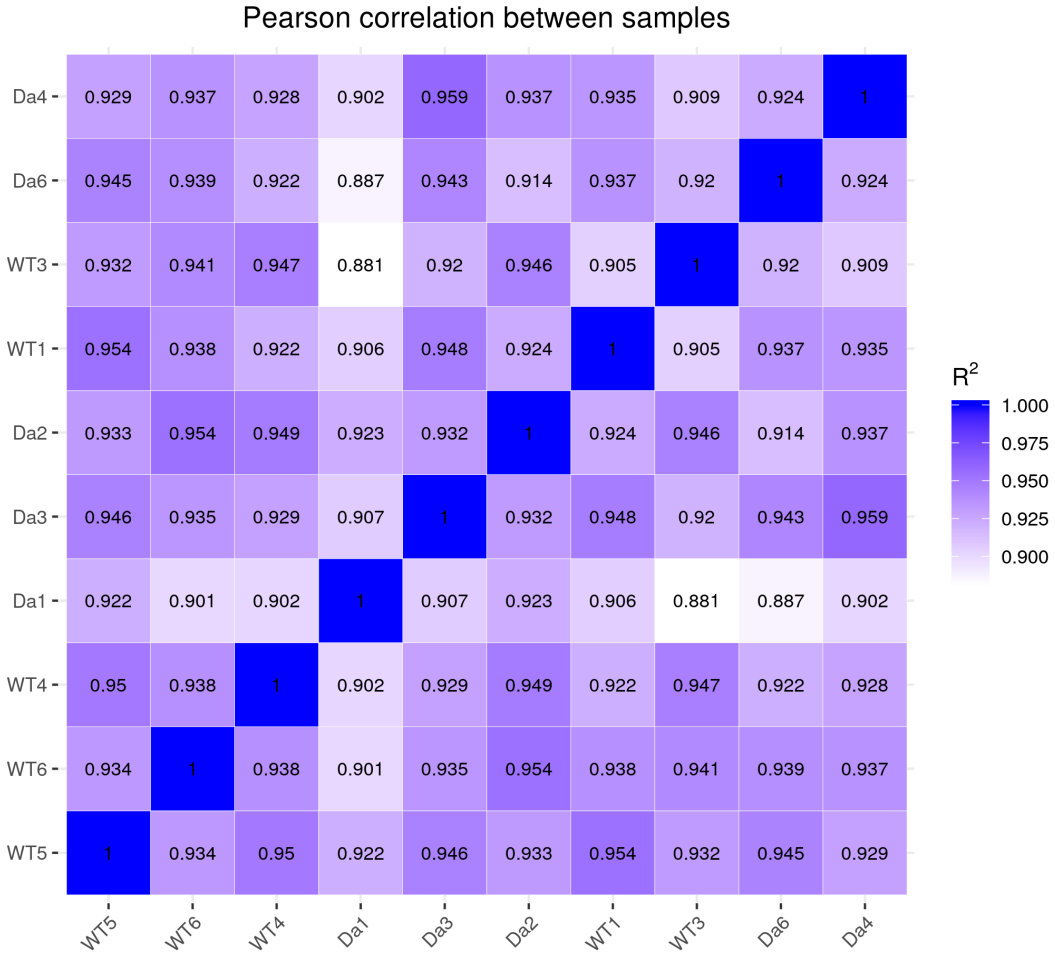


Figure 5.20: Pearson correlation between RNA-Seq samples. *Correlation coefficient (R^2) indicates similarity between biological replicates, with 1 being maximum similarity. Generated by Novogene Co. Ltd.*

5.7.3 Differential expression analysis supports RT-qPCR findings, except differential expression of *rpoS* was not apparent

As expected, and in agreement with the RT-qPCR data, none of the arginine-related genes examined in **Figure 5.18** were significantly differentially expressed in the Δ*aaaA* mutant. Moreover, *arcA* and *aruH*, genes for which the RT-qPCR primers failed to produce amplicons, were also not differentially expressed in the *aaaA* mutant.

Curiously, *rpoS*, which was found by RT-qPCR to be upregulated ~3-fold in the Δ*aaaA* mutant (see: 5.6.3) was not significantly upregulated in the RNA-Seq experiment. The most obvious explanation for this is the slight differences in sampling methodology. For RT-qPCR, samples were grown for 16 hrs then harvested for RNA extraction, whereas for RNA-Seq, in order to

compare between groups with exogenous arginine supplementation, both groups were removed from the incubator at 16 hrs and either 50 µL SWF or SWF + 100 µM L-arginine was added before a further hour incubation. This will no doubt have replenished the nutritional environment, though to different extents. RpoS is a stress or starvation response regulator in *E. coli,* and though it has been shown to have more roles than this in *P. aeruginosa,* this is likely still one of its major functions (Jørgensen *et al.*, 1999; Potvin *et al.*, 2008). Therefore, it is possible that addition of exogenous nutrients caused *rpoS*, which was likely upregulated at 16 hrs, to be subsequently down-regulated prior to sampling. Moreover, even if supplementation did not affect *rpoS* expression, as sampling was at 17 rather than 16 hrs, the window for detecting *rpoS* transcription may have passed. Repetition of the RT-qPCR exactly mirroring the RNA-Seq sample preparation, including the supplementation step, could easily determine which is the case.

5.7.3.1 RpoD is not differentially expressed in the SCW, supporting its use as an RT-qPCR endogenous control

There was no significant differential expression of *rpoD* between the wild-type and Δ*aaaA* mutant in the SCW (Log2FC -0.0515, p= 0.8216), providing useful validation of its applicability as an endogenous control for future RT-qPCR experiments in the SCW. Moreover, other genes were identified that also had equal expression levels across conditions tested including *pilT*, *mexT* and *glnD.* These may serve as useful additional endogenous controls in future experiments, given that *rpoS* is not an appropriate control when examining transcriptomics in the *aaaA* mutant.

5.7.4 Genes related to phenazine and PQS biosynthesis, biofilm formation and oxidative stress response are downregulated in the Δ*aaaA* mutant

There were no large differences in gene transcription between the PAO1L wild-type and Δ*aaaA* mutant in the SCW. This is unsurprising, given the low level of *aaaA* expression and its localisation to the OM. However, there were a few changes which are of interest, and these are shown as a volcano plot identifying up and downregulated genes (**Figure 5.21**) and dot plots

showing enriched KEGG pathways which are downregulated (**Figure 5.22**) and upregulated (**Figure 5.23**). All significantly differentially expressed genes ($p<0.05$) are listed in **S-Table 0.1**.

5.7.4.1.1 A note about statistical significance calculations

An important caveat to the following findings is that when adjusted by testing for the FDR using DESeq2, the padj values were in most cases higher than 0.05. It is not clear if this is due to combined analysis of duplicate and triplicate groups, or reflective of a true lack of significantly differentially expressed genes. While the attrition of one of the three replicates in the wild-type condition increases the risk of false positives, it may also have skewed the padj calculation to be too conservative. To address this, only genes with a more stringent $p<0.005$ are discussed here, with the added caveat that these findings cannot be assessed alone, and should be further validated, as described later. Nevertheless, these findings, where consistent with other existing data, provide some interesting clues to *aaaA* function and regulation as a starting point for further investigation.

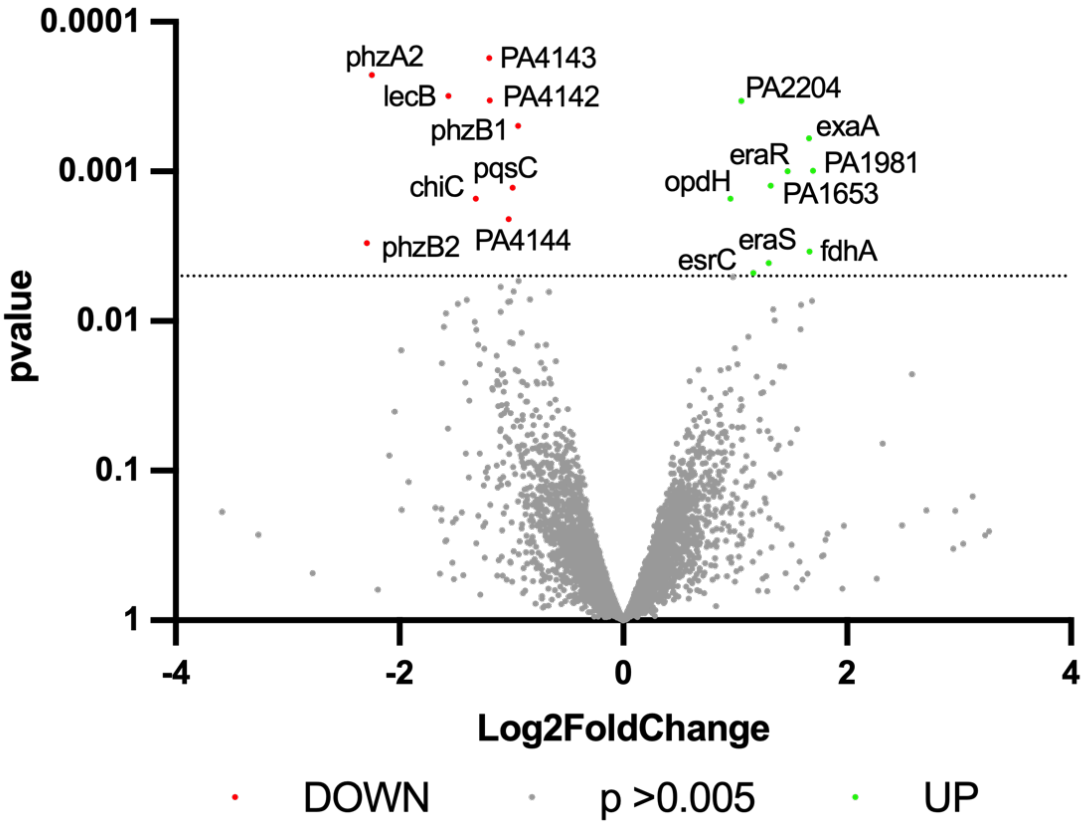


Figure 5.21: Differential expression of genes in PAO1L ΔaaaA compared with PAO1L. *Significantly (p <0.005) up and down regulated genes are shown in red and green, respectively, with the p-value cut-off at 0.005 marked with a dotted line.*

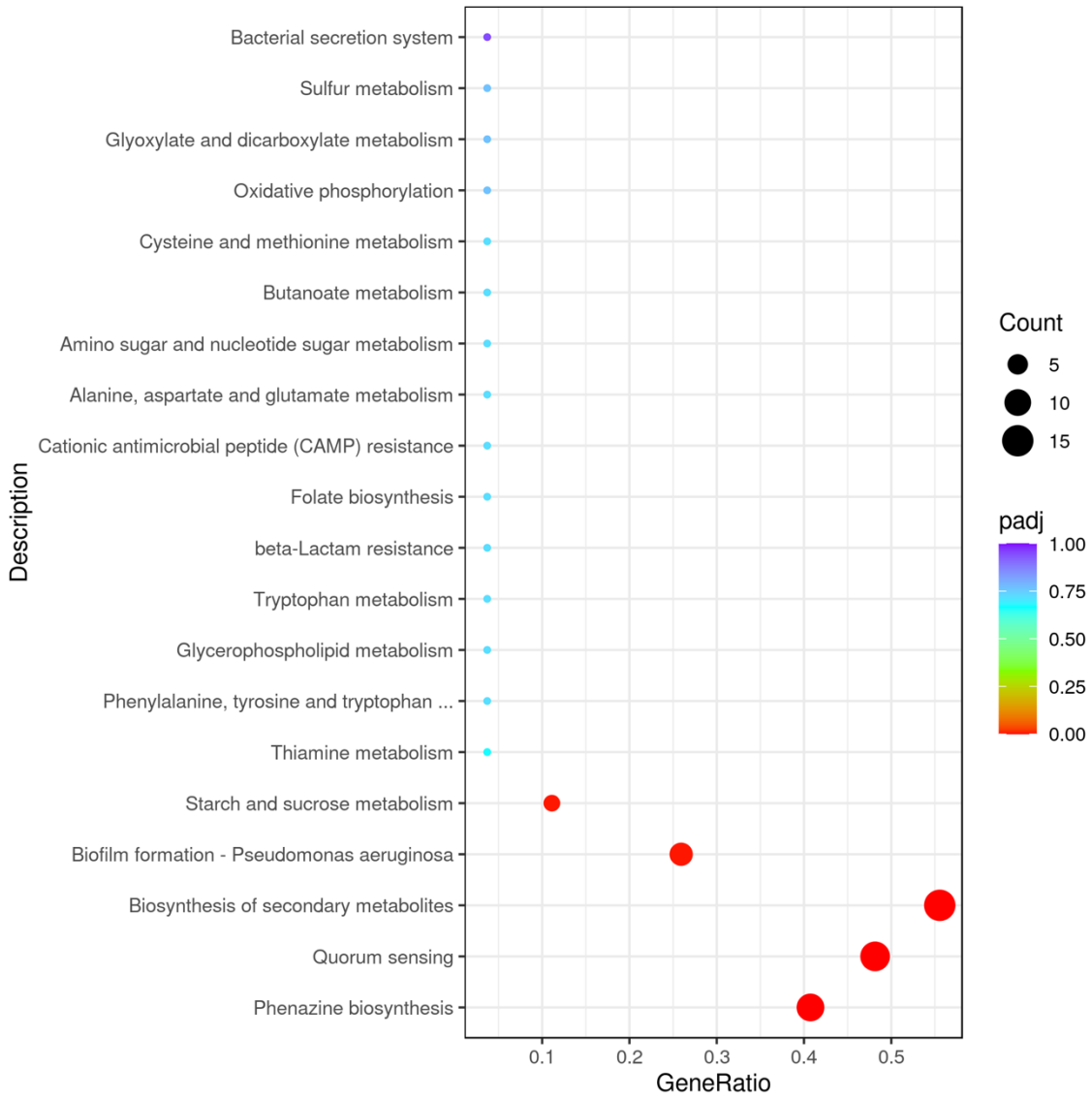


Figure 5.22: Dot plot showing the most significantly enriched KEGG pathways for genes downregulated in the ΔaaaA mutant compared to wild-type. *GeneRatio is the calculated Rich factor, and the size of each node corresponds to the number of significantly downregulated genes identified in that pathway. The colour of the node represents the adjusted p-value of significance. Generated by* Novogene Co. Ltd. (2022)

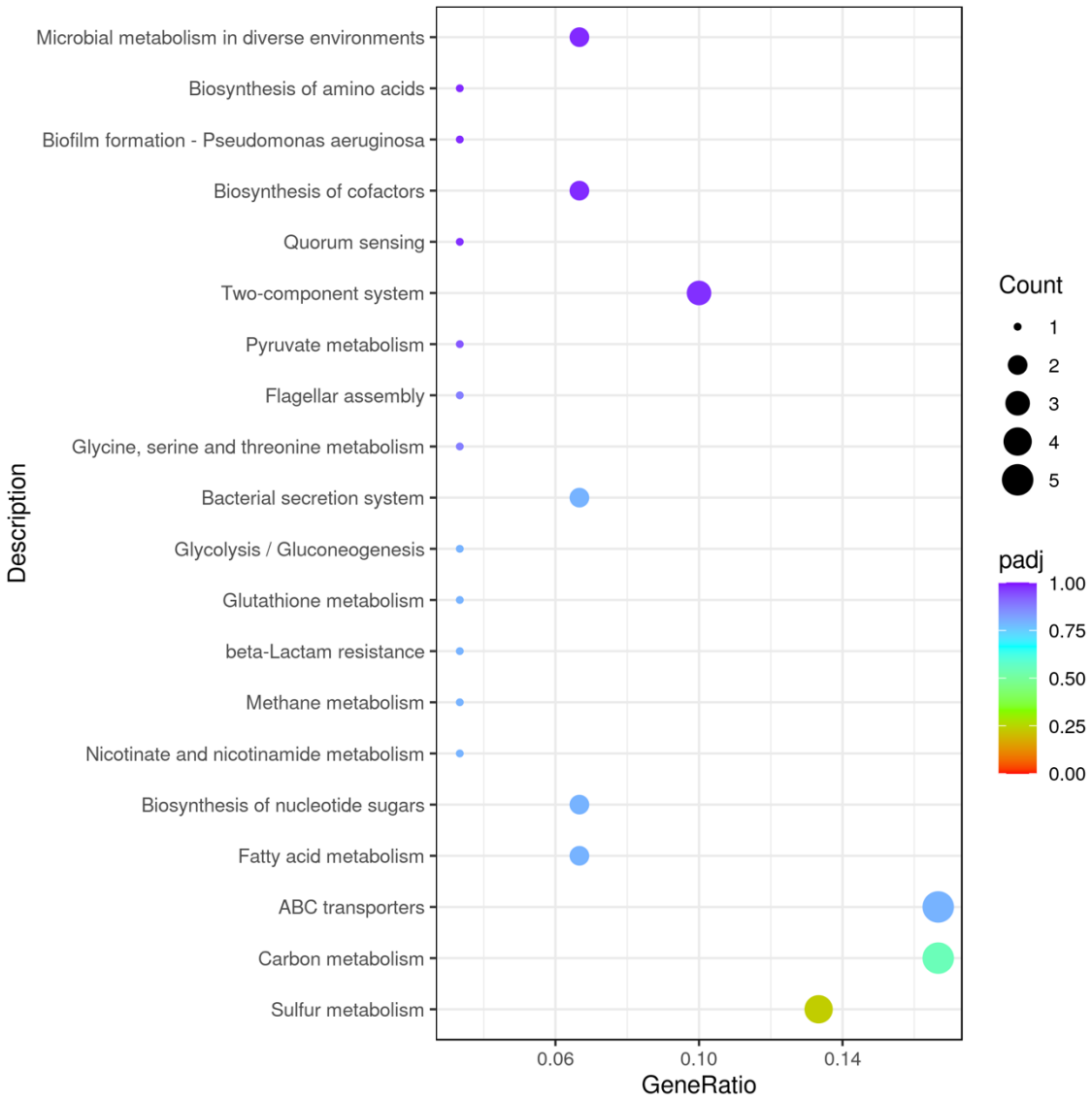


Figure 5.23: Dot plot showing the most significantly enriched KEGG pathways for genes upregulated in the ΔaaaA mutant compared to wild-type. *GeneRatio is the calculated Rich factor, and the size of each node corresponds to the number of significantly upregulated genes identified in that pathway. The colour of the node represents the adjusted p-value of significance. Generated by* Novogene Co. Ltd. (2022)

5.7.4.2 Phenazine and AQ biosynthesis

PhzA2 (PA1899), PhzB2 (PA1900) and PhzB1 (PA4211) are all involved in phenazine biosynthesis and pyocyanin production, while *pqsC* is part of the *pqsABCDE* operon involved in AQ biosynthesis. All of these genes were significantly downregulated in the *aaaA* mutant. Further genes found to be downregulated in the *aaaA* mutant with p=0.05-0.005 included *phzA1* and *phzM*, as well as the full *pqsABCDE* operon (**S-Table 0.1**).

Phenazines act as alternative electron acceptors in the absence of oxygen and therefore oxidize glucose and pyruvate into acetate for ATP production under low oxygen conditions, while keeping redox balanced (Glasser *et al.*, 2014; Price-Whelan *et al.*, 2007). This need for phenazine-dependant redox balancing is not required, however for arginine fermentation via the ADI pathway, which does not involve redox reactions (Glasser *et al.*, 2014). As discussed in 1.1.2.2, the *pqs* operon encodes genes for AQ biosynthesis, including HHQ and PQS, which, in conjunction with PqsR, control the PqsR regulon of QS virulence genes, including phenazine biosynthesis.

Curiously, though the FPKM of *phzA1/2* and *phzB1/2* was 100-350 and 200-600, respectively, there were almost no reads counted for *phzC1-G1* and *phzC2-G2* in any condition. The genes *phzA-G* form an operon, and *phzAB* alone is not sufficient for phenazine biosynthesis in *P. fluoresens* (Mavrodi *et al.*, 1998). A BLAST search confirmed the presence of *phzC1/C2* in both wild-type and *aaaA* mutant with 100% identity to the PAO1 sequence, confirming the genes after *phzAB* have not been mutated. The most likely explanation for this anomaly is that the very high sequence similarity of the *phzC1-G1* and *phzC2-G2* genes could have prevented proper mapping. This phenomenon has occurred previously, even when the presence of phenazines was confirmed by proteomics (Papangeli, 2022). Moreover, the PAO1L-infected SCW samples are visibly green by 16 hrs, which is indicative of pyocyanin biosynthesis (**Figure 5.10**).

5.7.4.3 Quorum sensing/RpoS regulated genes: LecB and ChiC

Both *lecB* and *chiC* are downregulated in the *aaaA* mutant, both with and without arginine, respectively (*lecB* Log2FC -1.56 and -1.40, *chiC* Log2FC -3.12 and -1.88). As briefly mentioned in 1.1.1, *lecB* (PA3361) encodes LecB, a QS-controlled fucose-binding protein which binds to Psl to help stabilise the extracellular matrix and promote biofilm formation (Passos da Silva *et al.*, 2019; Schuster *et al.*, 2003). *chiC* encodes a chitinase, ChiC, which breaks down the ubiquitous environmental polysaccharide chitin. ChiC first accumulates intracellularly before being slowly secreted over several days (Folders *et al.*,

2001). Another chitin binding protein, *cbpD,* which is also controlled by QS, is also downregulated in the *aaaA* mutant, though with $p>0.005$ (Log2FC -0.86, $p=0.0268$), and to a greater and more significant degree in the arginine supplemented condition (Log2FC -1.32, $p=0.0008$).

Both *lecB* and *chiC* are induced by QS, but repressed by RpoS during stationary phase (Schuster *et al.*, 2003, 2004). This suggests that despite *rpoS* upregulation not being detected by RNA-Seq, downstream effects of a previous upregulation in *rpoS* are apparent.

5.7.4.4 The Tes toxin secretion system

The putative operon containing PA4142-PA4144 was downregulated in the Δ*aaaA* mutant by an approximate average Log2FC -1.2, and -1.6 in the arginine-supplemented *aaaA* mutant. PA4142-PA4144 was recently identified as *tesABC*, encoding a type I secretion system cluster formed of TesA (PA4142) TesB (PA4143) and TesC (PA4144) (Zhao *et al.*, 2019). TesA, B and C are homologous to the *E. coli* T1SS CvaA, CvaB and TolC, respectively, which are involved in secretion of the toxin colicin V (Filloux, 2011; Gilson *et al.*, 1990; Zhang *et al.*, 1995; Zhao *et al.*, 2019; Zhong *et al.*, 1996). TesABC was found to likewise be responsible for secretion of the toxin TesG (PA4141), which is specifically expressed during chronic infection, and is thought to assist in immune cell evasion. The reduction in expression of genes exclusive to chronic infection in the *aaaA* mutant suggests that loss of *aaaA* is associated with abrogation of chronic infection phenotypes.

5.7.5 Genes related to alternative nutrient metabolism and uptake are upregulated in the Δ*aaaA* mutant

5.7.5.1 The Exa ethanol oxidation pathway

P. aeruginosa can catabolise ethanol and other secondary alcohols into acetaldehyde by oxidation. Ethanol oxidation in *P. aeruginosa* is carried out by a pyrroloquinoline quinone (PQQ)-dependent quinoprotein ethanol dehydrogenase encoded by *exaA* (PA1982), which forms a homodimer and is activated by ammonia or amines (Diehl *et al.*, 1998; Görisch & Rupp, 1989; Rupp & Görisch, 1988). The electron acceptor for the reaction is cytochrome

c_{550}, encoded by *exaB* (PA1983), which is controlled by a separate promoter (Reichmann & Görisch, 1993; Schobert & Görisch, 1999, 2001). ExaC, encoded by *exaC* (PA1984), then converts acetaldehyde into acetate, which can be subsequently converted to acetyl co-enzyme A and pyruvate, to serve as an energy source (Görisch, 2003; Schobert & Görisch, 1999).

exaA, but not *exaB* is upregulated by the soluble TCS EraSR (Schobert & Görisch, 2001). Both the *eraSR* operon (PA1979-80) and *exaA* are upregulated in the Δ*aaaA* mutant, along with the currently uncharacterised PA1981, which is predicted to be part of a two gene operon with *exaA*. PA1981 contains a pentapeptide repeat, the function of which is still unknown, but has been shown to bind to DNA gyrase in other species (Mérens *et al.*, 2009). It is worth noting that *exaA*-mediated ethanol oxidation is aerobic, and another enzyme, AdhA carries out ethanol oxidation under oxygen-limited conditions (Crocker *et al.*, 2019). However, *adhA* is not differentially expressed in the Δ*aaaA* mutant, suggesting there is still a preference for aerobic ethanol metabolism in the 16 hr SCW, where possible.

These findings suggest that in the absence of *aaaA,* ethanol oxidation pathways are upregulated, possibly to compensate for the lack of arginine as a nutrient source which is theoretically being supplied by AaaA. This is further supported by the finding that these genes are no longer significantly upregulated in the arginine supplemented Δ*aaaA* mutant (**Figure 5.24**).

5.7.5.2 Formaldehyde detoxification: FdhA

fdhA (PA5421), which encodes a glutathione-independent formaldehyde dehydrogenase, FdhA, that catalyses the conversion of formaldehyde to formamide, was upregulated in the *aaaA* mutant.. There a four other orthologs also encoded in *P. aeruginosa* (PA2119, PA2158, PA2188, PA5421) however none of these are differentially expressed in the *aaaA* mutant. FdhA is involved in glycine betaine, choline and potentially sarcosine metabolism (Diab *et al.*, 2006; Wargo *et al.*, 2008). Glycine betaine is abundant in the environment and can be derived from choline stripped from phosphatidylcholine component of host cell membranes. It can be used as sole carbon and nitrogen source for *P. aeruginosa*, and was shown to induce *fdhA*

(Diab *et al.*, 2006). The glycine betaine and related metabolic pathways are repressed by succinate, perhaps suggesting that levels of succinate were reduced in the *aaaA* mutant, which would be consistent with reduced arginine metabolism (Diab *et al.*, 2006). This could be tested by measuring succinate production in the *aaaA* mutant, to determine if this is the basis for the upregulation of *fdhA*.

However, it is worth bearing in mind that a secondary mutation was identified in PchP in PAO1L Δ*aaaA*, which is involved in choline savaging (**Table 5.4**). Though the frameshift mutation was toward the C-terminus and in a disordered region which is less likely to be deleterious to function, it is important to consider the possibility that the secondary mutation in PchP could impact upon downstream choline metabolism genes such as *fdhA*.

5.7.5.3 OpdH (OccK5)

OpdH, also referred to as OccK5, is a member of the OccK, or outer-membrane carboxylate channels, transporting small water-soluble molecules into the periplasm, such as benzoate and pyroglutamate (Eren *et al.*, 2012). OccK5 has a strong selectively for anions, and thus does not allow passage of arginine (Lee *et al.*, 2018). It is not entirely clear why OccK5 specifically is upregulated in the *aaaA* mutant, but it could be to allow entry of other alternative metabolites in a compensatory fashion.

5.7.5.4 PA1653, a putative oxidative stress response regulator

PA1653 was upregulated in the *aaaA* mutant. It is currently uncharacterised, but contains a MarR-type HTH domain, suggesting it is a putative transcriptional regulator of the MarR/SlyA family (Mistry *et al.*, 2021). MarR in *E. coli* controls regulons related to multiple antibiotic resistance and oxidative stress, suggesting that in the absence of AaaA, the cell may need to adapt to higher levels of oxidative stress (Alekshun *et al.*, 2001; Alekshun & Levy, 1999). This could be related to the downregulation of phenazines in the *aaaA* mutant, which act as a redox balancer (Glasser *et al.*, 2014).

5.7.6 Downregulation of *phz*, *chiC* and *tes* genes also occurs in arginine-supplemented Δ*aaaA* mutant, along with upregulation of nitric oxide reductase pathways

Supplementation with arginine was intended to simulate the hypothetical arginine increase provided by AaaA. Under these conditions, there is still differential expression of QS-controlled *phz* and *chiC*, as well as *tes* genes in the *aaaA* mutant, as shown in **Figure 5.24**. However, differential expression of ethanol oxidation and glycine betaine metabolism genes is not present in the arginine-supplemented *aaaA* mutant, suggesting they are nutritional compensation mechanisms. This suggests that if AaaA is acting to increase local arginine, its regulation may tie into both arginine-dependant and independent circuits, particularly considering the lack of differential expression of any arginine metabolism genes. Differential expression of the *pqs* operon is also not apparent in the arginine supplemented condition, suggesting *pqs* regulation here is also more directly related to arginine availability.

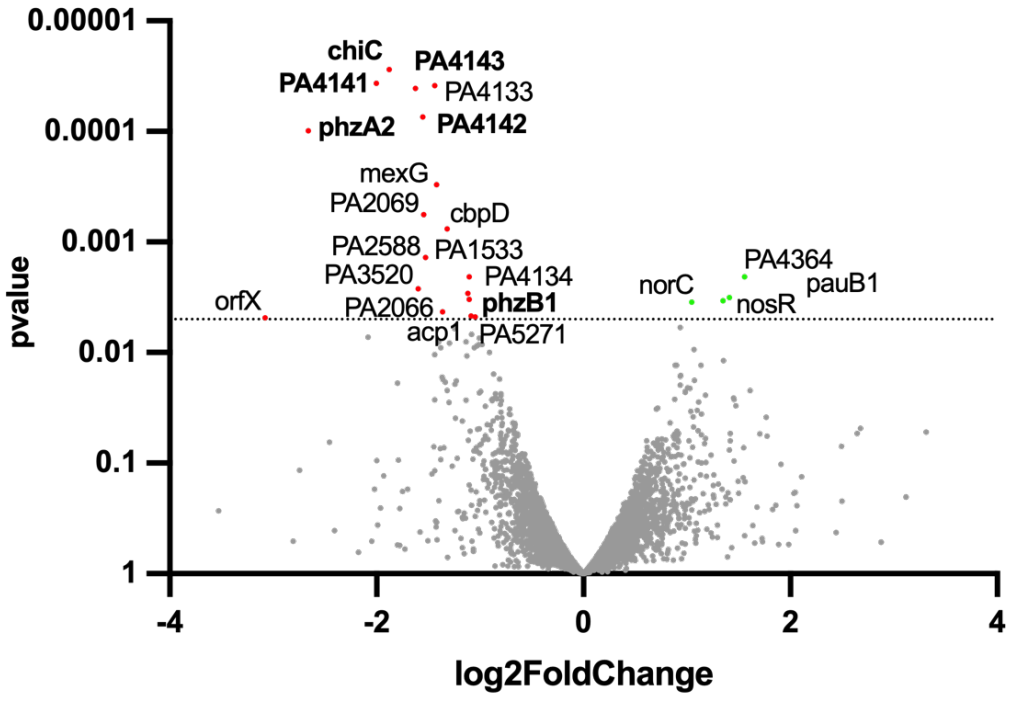


Figure 5.24: Differential expression of genes in PAO1L ΔaaaA supplemented with 100 µM L-arginine, compared with unsupplemented PAO1L. *Significantly (p <0.005) up and down regulated genes are shown in red and green, respectively, with the p-value cut-off of 0.005 marked with a dotted line. Genes shown in bold are those which are also differentially regulated in the unsupplemented PAO1L ΔaaaA compared with PAO1L (p<0.005) in* ***Figure 5.21****, suggesting the regulation is independent of arginine.*

5.7.6.1 Upregulation of integral membrane denitrification proteins NosR and NorC

Interestingly, *norC* (PA0523) and *nosR* (PA3391) were upregulated (Log2FC 2.06 and 2.55, respectively) in the supplemented, but not unsupplemented *aaaA* mutant. Both genes encode integral membrane proteins, NorC and NosR respectively, which are under the control of the ANR/DNR regulatory cascade. They are upregulated in response to nitric- but not nitrous- oxide, and are essential for nitrate/nitrite reduction (Arai *et al.*, 1999, 2003; Borrero-de Acuña *et al.*, 2016).

norC is part of an operon with *norB*; together they form the complex, NorBC, which serves as a major assembly platform for the denitrification protein super-complex (Borrero-de Acuña *et al.*, 2016). NosR associates with NorBC, and facilities regulation and recruitment of the nitrous oxide reductase, NosZ, to the denitrification complex (Borrero-de Acuña *et al.*, 2016; Cuypers *et al.*, 1992). This may suggest that in the absence of AaaA, arginine supplementation could further induce denitrification pathways. However, none of the other denitrification genes are upregulated in these conditions, with the exception of *nosD* (Log2FC 2.05, p =0.034) and *norB* (Log2FC 0.94, p =0.0196) although with p values >0.005. This makes conclusions difficult to draw, as one would expect more differentially expressed dentification genes if the denitrification pathways were truly upregulated. These genes were also not differentially expressed in the wild-type arginine-supplemented vs unsupplemented conditions (**S-Table 0.3**).

5.7.7 Other notable findings from RNA-Seq

5.7.7.1 Arginine addition was not sufficient to cause ArgR upregulation

When comparing the differential expression of genes in the arginine supplemented *vs* non supplemented wild-type-infected SCW, the most notable finding was that the expression of *argR* does not vary significantly due to arginine supplementation (**S-Table 0.3**). This also applied to the *aaaA* mutant and suggests that this concentration of arginine was not sufficient to induce *argR* in the SCW. However, it is clear from this work that the presence of AaaA

does not have a notable impact on expression of *argR*, or other arginine-induced genes, suggesting that if AaaA is releasing arginine, it is at a level below the threshold which would trigger transcriptional changes in arginine response/metabolism genes.

5.7.7.2 No UTR or alternative transcripts were detected for *aaaA*

The FPKM for *aaaA* was 1-3 in the Δ*aaaA* mutant compared with 21-28 in the wild-type, and *aaaA* expression in the wild-type did not significantly vary with arginine addition. This is further evidence that *aaaA* is relevantly expressed in the SCW, as *aaaA* expression is expected to be lower than other genes (for context, the average FPKM for all mapped genes was 352). The negligible FPKM value in the *aaaA* mutant likely reflects reads from the 4 codons left behind in the in-frame deletion, or other non-specific reads, being mapped to the locus.

Upon initial analysis, no ShD sequence nor any 3' or 5' UTR in *aaaA* could be identified. This is not entirely surprising as the ShD identified by Spencer (2018) was non-typical: AGAGGA. Moreover, no novel transcripts were detected on the negative strand immediately up or downstream of *aaaA*, indicating that there is not an alternative, or 'long'-*aaaA* transcript as postulated previously, at least under these conditions (Darch, 2009; Spencer, 2018).

5.7.7.3 Transcripts were detected for genes with potentially deleterious secondary mutations arising from Δ*aaaA* construction, which warrant further interrogation

As discussed in 5.3.2.3, when constructing the PAO1L Δ*aaaA* mutant, WGS identified genes with possible secondary loss of function mutations. Though RNA-Seq cannot be used to glean information about the upstream impacts of missense or nonsense mutations, it can indicate if the gene is transcribed in this context, and thus the potential magnitude of any loss of function of that gene.

Here, the FPKM for PA2421, PA2625 were ~0-3 and ~4-18, respectively, across all strains, suggesting that any potentially deleterious mutation would not be hugely impactful in these experiments. The FPKM across all strains for

the other affected genes were also relatively low: ~12-26 for *pchP,* ~40-85 for PA3906, ~30-75 for *purK,* ~23-43 PA4686 and ~20-45 for *pilK*. None of these genes were significantly differentially expressed in any group. The RNA-Seq analysis completed for this study also generated a list of potential SNPs and indels detected in each transcriptome, and though time constraints prevented their interrogation during this study, it would be useful to do so as part of future work, to verify and better assess the impacts of secondary mutations in PAO1L Δ*aaaA*. This may be particularly worthwhile for *pchP*, as *fdhA* which is also involved in choline metabolism, was differentially expressed in the *aaaA* mutant.

5.8 Discussion

5.8.1 AaaA is produced and relevant in the synthetic chronic wound

5.8.1.1 AaaA production is generally similar or greater in disease relevant strains and growth medium, particularly in the SCW

Work by Spencer (2018) found that, though there was variation, AaaA activity is generally higher in CF isolates than in PAO1N, when grown in LB. Spencer (2018) also showed that AaaA activity in CF isolates was greater still when grown in ASM. However, it is worth noting that AaaA activity in PAO1N was lower in ASM than LB following aerobic growth, but there was no difference following microaerobic growth, which is likely more reflective of the CF lung aero-environment.

Here, AaaA activity and P_{aaaA} activation in PAO1L and chronic wound isolate CW2-BI1 were examined in LB and the SCW and generated similar results. Both P_{aaaA} activation and AaaA activity were higher in the SCW than LB for both strains. Moreover, in LB, AaaA activity of several other chronic wound isolates was higher than that of PAO1L. Interestingly, CW2-BI1 exhibited intrinsically low AaaA activity in planktonic culture, but activity similar to PAO1L in the SCW, suggesting AaaA is preferentially or even exclusively expressed in the SCW for some clinical isolates.

Taken together, these findings suggest that though *aaaA* expression varies between strains, it is generally increased or at least unchanged in disease relevant contexts and strains compared with lab strains grown planktonically in LB. This seems to be more pronounced and consistent in the SCW than in ASM examined by Spencer (2018), suggesting that the SCW may be the more appropriate model for studying AaaA. Work to further characterise both PAO1L and other chronic wound isolates, to determine the potential regulatory causes for differential AaaA production, is ongoing. The construction of the *pEX18Gm-aaaA* suicide vector in this study will allow for the creation of further Δ*aaaA* mutants in a range of strain backgrounds and will serve as an invaluable tool for future work uncovering the complex regulation of AaaA.

This study has also validated the SCW as a useful model for studying AaaA, and there is much that could be expanded upon in terms of increasing the complexity of the SCW to reflect a chronic wound even further. Moreover, while all experiments in the SCW here were incubated aerobically, it would be interesting to explore the role of AaaA in the SCW under anaerobic or microaerobic conditions, where AaaA is even more likely to be relevant (Spencer, 2018).

5.8.2 There is no evidence that AaaA directly or indirectly modulates arginine metabolism gene expression

Data from both RT-qPCR and RNA-Seq indicates that the presence of AaaA itself does not have any regulatory impact on other arginine metabolism genes, nor does it impact expression of ArgR. Due to poor primer efficiency, it was not possible to measure the expression of *arcA*, of the ADI pathway using RT-qPCR. However, expression of *arcD*, which is adjacent to *arcA* on the *arcDABC* operon was measured and did not appear to be differentially expressed in the Δ*aaaA* mutant. This suggests that like the other arginine catabolic pathways, expression of the components of the ADI pathway are similarly unaffected by *aaaA*.

5.8.2.1 Potential future work to probe the ArgR-AaaA relationship

It was posited that the presence of AaaA would increase local arginine, and this increase would activate ArgR, which in turn could have a feedback effect of positively regulating *aaaA*, as well as altering transcription of arginine metabolism genes (Paredes-Osses, 2014; Spencer, 2018). There is solid evidence from *lux*-transcriptional reporters, AaaA activity assays and transcriptomics in *argR* mutants that ArgR does positively regulate AaaA, as well as a putative ArgR binding site in the *aaaA* promoter, suggesting that ArgR plays a key role in AaaA regulation (Paredes-Osses, 2014; Alhagh, 2020; Lu *et al.*, 2004; Spencer, 2018). However, this is no evidence that AaaA, or its action upon arginine, causes ArgR induction. This may be because any arginine released by AaaA is not sufficient to reach the threshold for triggering ArgR induction. This is supported by the finding that ArgR was also not upregulated in the arginine supplemented condition examined in this study by RNA-Seq, which was designed to mimic the hypothesised small increase in L-arginine created by AaaA. Taken together, this suggests that the regulation of AaaA by ArgR is unidirectional, rather than part of a feedback loop.

Further RT-qPCR experiments with higher levels of arginine supplementation would be useful to 1) determine the minimum concentration of exogenous arginine required to upregulate ArgR and 2) examine wild-type *aaaA* expression in relation to ArgR induction. Moreover, interrogation of the putative ArgR binding site by mutating conserved bases, as was done with the $P_{aaaA\,\Delta rpoNmotif}$ construct, would be useful to conclude definitively whether the positive regulation of AaaA by ArgR is via direct binding to the promoter (Alhagh, 2020).

5.8.3 A newly uncovered relationship between RpoS and AaaA

Data from RT-qPCR showed that *rpoS* was significantly upregulated in the *aaaA* mutant in the 16 hr SCW. The potential reasons for a lack of replication in the RNA-Seq are discussed in 5.7.3, but may be related to the different sampling methodologies, in particular the inclusion of a supplementation step followed by an additional hour incubation in the RNA-

Seq method. Interestingly, *lecB, chiC* and the *phz* operons are all downregulated by RpoS (Schuster *et al.*, 2004). This suggests that even though differential *rpoS* transcription itself was not observed in the RNA-Seq, the potential downstream effects of RpoS were apparent.

Given that AaaA is an outer-membrane protein with defined function as an aminopeptidase, it makes more sense that the impact of AaaA on RpoS expression is indirect via its function, which is predicted to be in releasing arginine. RpoS is essential for growth on N-acetyl-L-glutamic acid (NAG) as a sole carbon/nitrogen source, and confers a growth advantage when grown on casein, indicating that RpoS is important for proteolytic-dependant growth (Robinson *et al.*, 2020). The RpoS regulon includes ~800 genes, many of which encode proteolytic secreted enzymes (Schuster *et al.*, 2004). This suggests that in the absence of AaaA proteolysis, *rpoS* may be upregulated to induce other proteolytic enzymes to compensate. One RpoS-controlled gene of particular interest with respect to AaaA is *pepB*, which encodes PaAP.

5.8.3.1 PaAP is an RpoS and QS-controlled secreted M28 aminopeptidase with a homologous active site to AaaA

PaAP, historically referred to as PepB and encoded by *pepB* (PA2939), is a secreted leucine-specific aminopeptidase which is also predicted to be a member of the M28 family (Cahan *et al.*, 2001). Alignment with AaaA in **Figure 5.25** shows conservation of the amino acids predicted to be involved in aminopeptidase activity (Darch, 2009; Luckett *et al.*, 2012). As well as being controlled by *rpoS*, it is one of 43 genes which form a core, conserved QS-controlled regulon, along with the chitin binding protein CbpD (Chugani *et al.*, 2012), which was interestingly downregulated in the *aaaA* mutant. As Cahan et al. (2001) did not include L-arginine-p-nitroanilide cleavage in their activity screen, it may be useful to test this, to determine if PaAP also has arginine-specific aminopeptidase activity. Alternatively, growth on minimal media with the RGD tripeptide could be assessed to determine if, like AaaA, PaAP is able to liberate arginine for growth. This could establish if PaAP shares redundancy with AaaA, which, if so, would explain the relatively marginal phenotype of Δ*aaaA*.

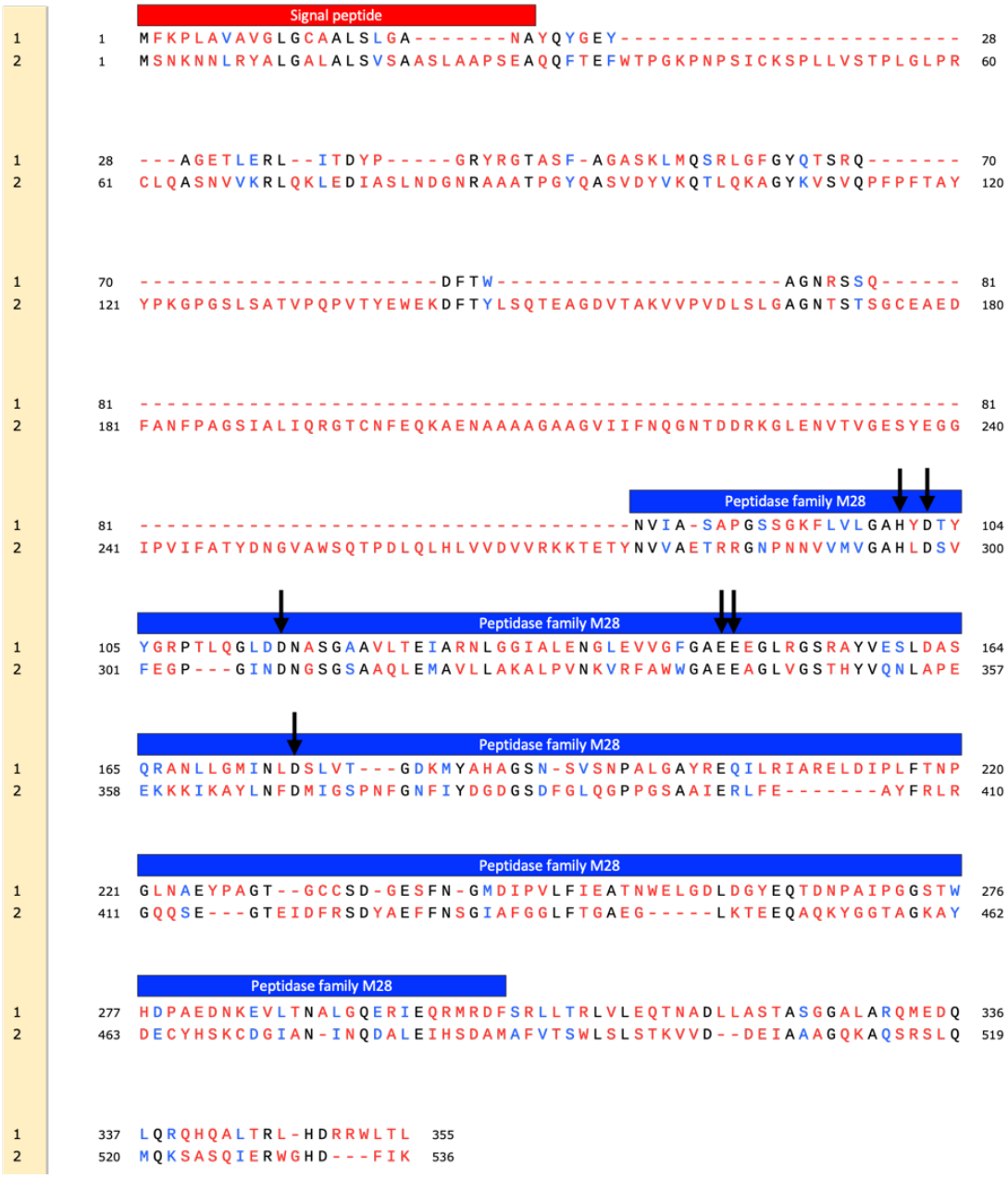


Figure 5.25: Amino acid alignment of AaaA (1, top) with PaAP (2, bottom). *Arrows indicate residues in AaaA known to be essential for secretion and/or aminopeptidase activity, which are conserved in PaAP. Variant amino acids are highlighted as either blue (similar residue) or red (not similar residue). Generated using Snapgene Semi-global alignment version 2.4.2 with Substitution Matrix BLOSUM62, consensus threshold >50%.*

If RpoS were involved in co-ordinating redundancy between PaAP and AaaA, it would be reasonable to expect that PaAP expression would be higher when AaaA is lost. However, there was no differential expression of *pepB* in the *aaaA* mutant. Although, the FPKM for *pepB* across all samples were ~10-fold higher than those for (wild-type) *aaaA*, suggesting that intrinsically higher

expression of *pepB* could provide redundancy without the need for *aaaA*-dependant regulation.

5.8.3.2 RpoS represses genes directly up and downstream of *aaaA*

While transcriptomic analysis of a *rpoS* mutant did not identify *aaaA* as a target of *rpoS*, both genes up and downstream of *aaaA,* albeit on the opposite strand, are repressed by *rpoS*. PA0327 and PA0329 were both upregulated in the *rpoS* mutant by 2.1-fold at late exponential phase, and respectively, by 6.5-fold and 18.4-fold in stationary phase (Schuster *et al.*, 2004). Given their proximity to PA0328, it is worth further exploring the possibility that *rpoS* has either direct or indirect regulatory control of *aaaA*.

5.8.3.3 A search for a *rpoS* binding site in the *aaaA* promoter yields two low-consensus putative sites

The putative -10 promoter sequence for *aaaA* is predicted to be a target for alternative sigma factor RpoN, rather than RpoS (**Figure 5.1**) (Schuster *et al.*, 2004). However, while the conserved putative -10 promoter sequence for RpoS*,* CTAxACT, is not present in the -10, -35 region of the *aaaA* promoter, some conserved elements (CTNNNCT) were present further upstream, which could encode an elongated *aaaA* transcript, illustrated in **Figure 5.26***.* However, based on the findings here, as well as the bioinformatic analysis by Spencer (2018), which was supported by data from Wurtzel et al. (2012), there is currently no evidence of an alternative transcriptional unit encoding a 'long' *aaaA*.

Using mRNA profiling, as well as chromatin immunoprecipitation coupled to high-throughput sequencing (ChIP-Seq), Schulz et al. (2015) identified the sigma factor binding sites for all PA14 genes, as well as genes where there is cross-talk between multiple sigma factors. They identified RpoN as the sigma factor controlling *aaaA*, and furthermore, identified no crosstalk with other sigma factors, including RpoS. This provides further evidence that RpoN is the only sigma factor controlling *aaaA*, suggesting that the relationship between *aaaA* and RpoS is indirect, downstream of *aaaA* transcription, and independent of RpoS binding to the *aaaA* promoter. Importantly, though the data here shows that loss of *aaaA* leads to *rpoS* upregulation, there is currently

no evidence that loss of *rpoS* affects *aaaA* expression. It would therefore be useful to employ the P_{aaaA}-lux transcriptional reporter or RT-qPCR in a Δ*rpoS* background, to determine this.

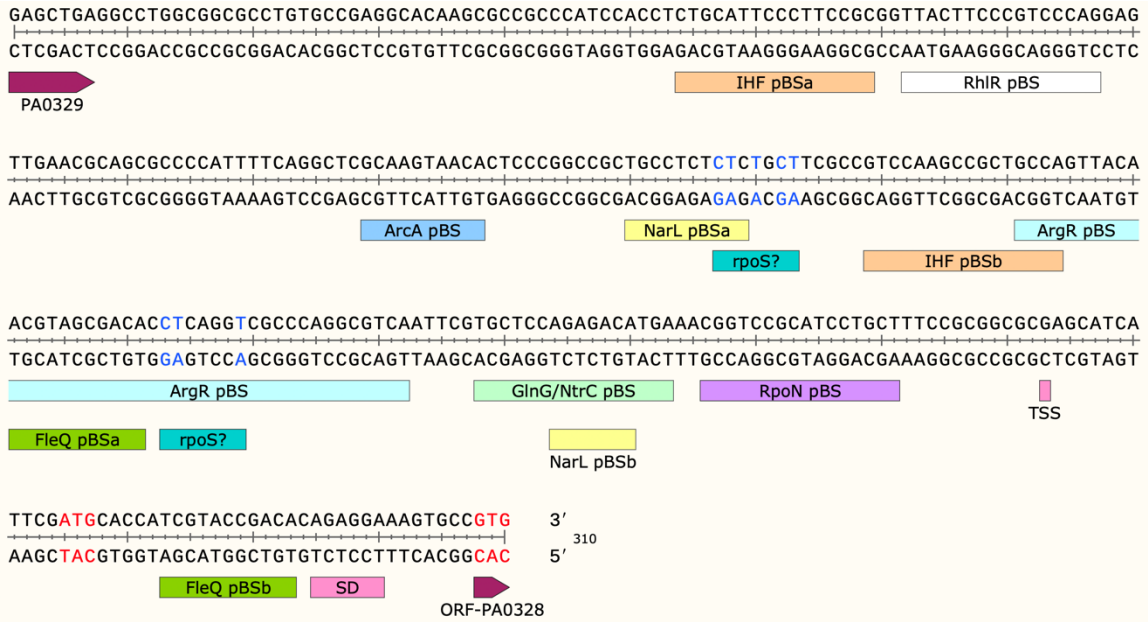


Figure 5.26: Potential poorly conserved binding site for RpoS exist upstream of aaaA. *Start codons for the reference aaaA transcript, as well as the previously posited 'long' AaaA as shown in red. Two poorly conserved rpoS binding sites are highlighted in turquoise, with conserved bases in blue.*

5.8.4 Potential clues from RNA-Seq link AaaA to phenazine biosynthesis and quorum sensing

Using RNA-Seq, this study identified QS-controlled genes, including *chiC*, *lecB* and *pqs*, *phz* and *tes* gene clusters which were all downregulated in the *aaaA* mutant, and this difference was sustained for these genes, with the exception of the *pqs* operon, when supplemented with 100 μM arginine, suggesting they are independent of arginine concentration.

Interestingly, *pqsH* and *pqsL*, the genes encoding the enzymes responsible for converting HHQ into PQS and HQNO were not differentially regulated. This may suggest that the levels of HHQ, but not PQS, are primarily altered in the *aaaA* mutant. As HHQ is the precursor to PQS and HQNO, lower levels of HHQ would likely result in reduced PQS. However, on the transcriptional level, the impact of the *aaaA* mutation appears to be specifically upon HHQ production. Therefore, production of all AQs, particularly HHQ, in the *aaaA* mutant should be determined, as described in 5.8.4.3.

HHQ alone has been shown to be sufficient to induce a range of QS-controlled genes, including ChiC, LecB and the *pqs* operon, although PQS is a more potent QS inducer (Xiao *et al.*, 2006). Additionally, HHQ is known to mediate arginine-induced repression of swarming in *P. aeruginosa* and this is independent of PQS, RhlR/LasR and c-di-GMP (Ha *et al.*, 2011). This arginine-dependant repression of swarming was also related to the phenazine product and pyocyanin precursor, phenazine-1-carboxylate, providing a further link between arginine-mediated swarming repression and phenazines, in addition to that with HHQ (Ha *et al.*, 2011). It is therefore possible that the arginine released by AaaA is sufficient to induce HHQ and phenazine biosynthesis, which in turn contributes to sessility/motility switching. The fact that *pqs* was not differentially expressed when the *aaaA* mutant was supplemented with arginine further supports this.

Moreover, further experimental evidence is needed to confirm this hypothesised link between AaaA and swarming. Swimming and swarming assays conducted by Spencer (2018) did not show any difference in motility between PAO1N and PAO1N Δ*aaaA,* however as PAO1N has known deficiencies in swarming regulation (discussed in 5.3.1), these experiments should be repeated in the PAO1L and CW2-BI1 backgrounds, and ideally with a Δ*aaaA*Δ*pqs* or Δ*aaaA*Δ*phz* double knock-out, to determine the role, if any, for AaaA in arginine-induced, phenazine/HHQ-mediated swarming repression. It is also worth noting that HHQ production in CW2-BI1 was intrinsically lower than other strains, and this should be borne in mind when examining the link between AaaA and HHQ in this and other clinical strains (da Silva, 2018).

5.8.4.1 Previous links between phenazines and *aaaA* provide further evidence that they are co-expressed

The relationship between phenazines and AaaA was first explored by Spencer (2018), who showed that in day-6 aerobic Congo Red colony biofilms, a PA14 Δ*aaaA*Δ*phz* double knock-out showed a stronger ‘wrinkly’ phenotype than the wild-type or either of the single knock-outs. This phenotype is associated with redox-imbalance, and the wrinkling is thought to be an

adaptation which increases surface area and thus oxygen accessibility to counter this (Dietrich *et al.*, 2013).

This stronger phenotype in the double, but not single *phz* mutant suggests that AaaA compensates for phenazine-mediated fermentation in biofilms, presumably by supplying arginine for non-redox active metabolism. The finding in this study, that *phzA2*, *phzB2* and *phzB1* are downregulated in the Δ*aaaA* mutant, regardless of exogenous arginine, supports the theory that these genes are functionally co-expressed.

A possible mechanistic link between AaaA and the phenazine-mutant wrinkly phenotype is through RmcA, which also produces a wrinkly phenotype when deleted. It is thought that, in addition to sensing arginine via its periplasmic, 'Venus-fly trap' domain, RmcA also primarily senses oxidised phenazines, to indicate the redox state of the cytoplasm, via its PAS domain, which triggers c-di-GMP degradation and prevents wrinkling (Okegbe *et al.*, 2017). It is therefore possible that arginine released by AaaA acts in tandem with phenazines to signal, via RmcA, the energy producing capacity of the cell (Scribani Rossi *et al.*, 2022). More work is needed to establish this experimentally, for example by measuring the c-di-GMP levels in cells with combinations of mutations in *aaaA*, *rmcA* and *phz* genes and determine whether c-di-GMP levels are reduced further in double *aaaA*-*phz* knockouts compared to single mutants.

Pyocyanin, a key *P. aeruginosa* virulence factor, is one of the major terminal products of phenazine biosynthesis. Previously, Spencer (2018) compared pyocyanin production between PAO1N and PAO1N Δ*aaaA* and did not find any significant difference when grown aerobically in LB for 18 hrs. However, it is possible that under different conditions more relevant to AaaA expression, differences would be apparent, and this warrants further investigation, as discussed in 5.8.4.3.

5.8.4.2 The TesABC cluster could be a co-expressed with AaaA during chronic infection

The *tesABC* cluster was upregulated ~8-fold during chronic lung infection in mice and was also detected in the sputum of patients with chronic,

but not acute, *P. aeruginosa* lung infections (Zhao *et al.*, 2019). This correlated with pathogenicity; there was no difference in the pathogenicity of PAO1 Δ*tesG* in the acute infection, but a higher mouse survival rate, and lower CFU counts, were observed in the mice chronically infected PAO1 Δ*tesG* compared with wild type. The lungs of mice chronically infected with the Δ*tesG* mutant also showed more severe inflammatory cell infiltration, suggesting TesG has an immunosuppressive effect due to its inhibition of host GTPases (Zhao *et al.*, 2019). This mirrors the patterns seen for AaaA in the mouse wound infection models, where AaaA was shown to be essential for virulence and involved in immune modulation at the chronic, but not acute stage of infections.

The *tesABC* operon and *tesG* have been shown in several studies to be upregulated by the QS system RhlR, LasR, and induced by the QS molecules 3O-C_{12}-HSL and C_4-HSL (Schuster *et al.*, 2003; Wagner *et al.*, 2003; Zhao *et al.*, 2019), providing a further link between QS and *aaaA*. It is possible that signals provided by AaaA, potentially arginine, could induce a QS-mediated switch to a specific chronic infection phenotype, characterised by QS-dependant upregulation of genes including *chiC*, *lecB* and the *tes* cluster, as seen here. It is also possible that this switch is mediated by RpoS, which represses many QS genes, including *chiC* and *lecB.*

5.8.4.2.1 The previously observed autoregulation of AaaA may be mediated by QS systems

Curiously, though it has historically consistently been observed that loss of *aaaA* leads to increased P_{aaaA} activation, this was not observed in this study, either in LB or the SCW. It is unclear why and may be an artifact caused by a faulty transcriptional reporter, thus repetition of the assay with fresh conjugates is needed to be certain. However, the observation that the *pqs* operon, responsible for HHQ production, is downregulated in the *aaaA* mutant may provide a possible mechanism for this autoregulation, as there is some evidence that RhlR, LasR and PqsE, among other QS components reduce P_{aaaA} activation and AaaA activity (Alhagh, 2020; Paredes-Osses, 2014). Thus, if loss of *aaaA* causes downregulation of these components, this reduces their potential repression, direct or indirect, upon the *aaaA* promoter.

5.8.4.3 Potential future experiments to examine the relationship between phenazines, QS, and AaaA

The data generated by RNA-Seq should be validated by quantifying the production of pyocyanin and other phenazines in the SCW and comparing these between the wild-type and *aaaA* mutant. Pyocyanin could be quantified using the same method as Spencer (2018), or more ideally, using LC-MS to quantify the full range of phenazine products, as well as other QS signals such as PQS and HHQ, as in Papangeli (2022). Given the lower statistical reliability of the RNA-Seq data presented here, transcriptomic validation using RT-qPCR may also be useful, particularly to compare phenazine production in the arginine supplemented *vs* unsupplemented infected SCW, and over further time points.

To examine a potential role of QS in *aaaA* regulation and potential auto-regulation, it would be useful to employ the P_{aaaA}-lux transcriptional reporter to examine *aaaA* promoter activation in response to exogenous QS autoinducers $3OC_{12}$-HSL, C_4-HSL, and more pertinently, different AQs, particularly HHQ and PQS. It may be useful to also repeat the experiments in the *rhlR* and *lasR* knock-outs, as well as including a *pqsR* and/or *pqsABCDE* knock-out to add clarification to the existing conflicting data and parse out which genes in the QS network *aaaA* is most affected by, and affective upon (Alhagh, 2020; Paredes-Osses, 2014).

5.8.5 Ethanol oxidation could be a potential compensatory pathway to arginine starvation

The RNA-Seq data presented here showed that ethanol oxidation pathway genes and regulators, as well as the formaldehyde dehydrogenase gene *fdhA*, were upregulated in the *aaaA* mutant, but these differences were not apparent following supplementation with 100 μM L-arginine. This suggests that *P. aeruginosa* may compensate for the lack of AaaA and/or arginine by upregulating alternative metabolic pathways.

As well as being a potential nutrient, at low concentrations ethanol can also serve as a chemotactic signal. A study by Badal et al. (2021) found that ExaA and ExaB, or specifically their action converting ethanol to acetaldehyde

was essential for swarming on peptone growth media agar, which contains 0.1% (vol/vol) ethanol as a solvent for cholesterol. Addition of exogenous acetaldehyde also induced swarming independently of ExaAB and ExaC, suggesting it is the product acetaldehyde, which has signalling roles. The induction of swarming by ethanol was shown to be concentration-dependent and only occurred at low concentrations (0.05-1% v/v), while higher concentrations promote sessility, via intracellular c-di-GMP increases (Lewis *et al.*, 2019). This is a strikingly similar pattern of concentration-dependant chemotaxis seen with arginine, which is also thought to be controlled via c-di-GMP (Everett *et al.*, 2017; Paiardini *et al.*, 2018).

This role for ethanol as a swarming inducer was shown to be regulated not by EraSR, but by ErdR, which though not differentially expressed in the *aaaA* mutant, was consistently expressed in all conditions (FPKM 60-100). Previous findings also show that *erdR* is constitutively expressed (Mern *et al.*, 2010). Badal et al. (2021) also showed that ErdR, along with RhlA, contributed to PA14 swarming on a lawn of *Cryptococcus neoformans,* and ethanol-producing yeast which of co-infects with *P. aeruginosa* in the CF lung.

This was also linked to RpoS by Badal et al. (2021), who showed that a small population of cells at the tip of the moving (swarming) tendril expressed higher levels of RpoS. They suggested that this RpoS upregulation was indicative of a ‘foraging’ phenotype and that under starvation conditions, low concentrations of ethanol acted as a foraging signal. It is possible that the loss of arginine theoretically being produced by AaaA caused upregulation of RpoS to activate a similar foraging phenotype, which subsequently activated parallel ethanol scavenging along with altering the other RpoS-controlled genes identified in this study.

ErdR has been previously shown to downregulate *acoR* 2-fold, and *fdhA* 2-fold and upregulate *sfa2* 1.5-fold. Interestingly, both *acoR* and *sfa2* have been shown to increase P_{aaaA} activation, although with some variability (Alhagh, 2020; Paredes-Osses, 2014). This provides a potential regulatory link between the systems controlling ethanol oxidation and *aaaA.* However, *fdhA*, which is

down-regulated 2-fold by ErdR, was upregulated in the Δ*aaaA* mutant, casting doubt on a role for ErdR on *fdhA* regulation in this study.

Taken together, these findings suggest that in the *aaaA* mutant, the upregulation of ethanol oxidation genes is more likely for utilisation of ethanol as a nutrient source, via EraSR, rather than for chemotaxis, via ErdR, although there is not enough data to conclude either way. The identification, for the first time in this study, of a link between ethanol oxidation and *aaaA,* possibly via RpoS, warrants further study to better understand their potential tandem relationship in nutrient scavenging.

5.8.5.1 Potential future experiments to explore the relationship between ethanol and arginine metabolism, and the involvement of AaaA

Some light could be shed on the strength of ethanol *vs* arginine as a chemotaxis signal by setting up a duel swarming experiment using peptone growth media agar as in Badal et al. (2021), with a gradient of ethanol on one extreme of a plate and RGD tripeptide or L-arginine on the other. The swarming patterns of wild type, Δ*aaaA* and *aaaA* over-expression strains toward each of the two poles could then be compared, to establish which signals are most pertinent to chemotaxis with respect to AaaA. Moreover, a double *ExaAB/aaaA* mutant could be constructed and its growth in the SCW examined, to better determine the strength of ethanol oxidation genes in compensating for *aaaA* with respect to growth and survival in the SCW.

5.8.6 Conclusions

In conclusion, this chapter has shown that AaaA is more highly produced and enzymatically active in the SCW than in planktonic LB culture and confers a growth advantage in the SCW, aligning with previous data from a murine chronic wound model. RT-qPCR identified *aaaA* expression at 16, 24 and 40 hrs in the SCW, with expression highest at 16 hrs. No arginine-related genes were differentially expressed, but surprisingly, *rpoS* was found to be upregulated in *aaaA* mutant in the 16 hr SCW.

RNA-Seq analysis of the *P. aeruginosa* transcriptome in the 16 hr SCW also identified genes involved in ethanol oxidation and alternative metabolic

pathways were upregulated in the *aaaA* mutant. Supplementation with arginine reversed this, and elements of the denitrification pathways were instead upregulated. This suggests that alternative metabolic pathways compensate for the loss of AaaA (and its arginine acquisition capabilities), depending on the nutritional environment. Conversely, genes involved in phenazine/AQ biosynthesis, redox response, chronic-infection specific toxin secretion and QS-controlled biofilm-associated genes were downregulated in the *aaaA* mutant, and this was unaffected or even exaggerated by arginine supplementation. This, combined with the likely reduction in biofilm formation by the *aaaA* mutant suggests that AaaA is involved in maintaining a chronic infection phenotype also typified by these gene clusters.

Taken together, and as summarised in **Figure 5.27**, these data suggest that when *aaaA* is lost, the cell adopts a more 'nutrient foraging' phenotype, characterised by upregulation of *rpoS* which in turn co-ordinates, possibly via QS systems, nutrient acquisition and metabolism genes in relation to alternative available nutrients in the environment. There is an emerging picture suggesting that this in turn restricts biofilm formation and chronic infection phenotypes, however more data is needed to better support and understand this (Luckett *et al.*, 2012; Spencer, 2018). AaaA may also indirectly induce the *pqs* operon to produce HHQ, but not PQS, which could in turn drive *aaaA* autorepression via RhlR, and induce the AaaA-associated QS genes which would otherwise be repressed by RpoS. Additional phenotypic and proteomic analysis is required to further validate the findings of this study, and better understand if and how AaaA, RpoS and QS work together functionally in the context of a chronic wound biofilm.

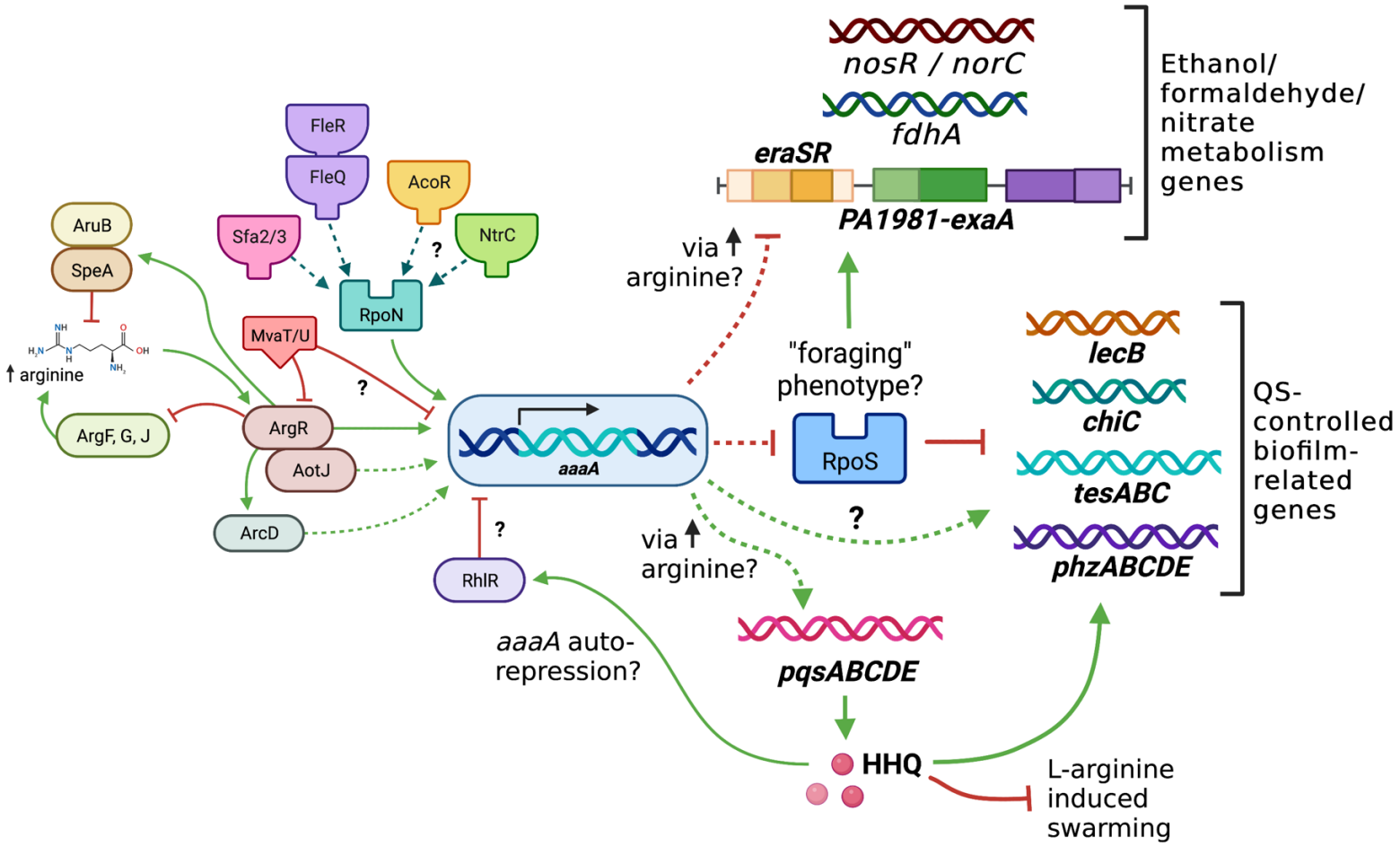


Figure 5.27: Summary of the known and putative* aaaA *regulators as well as potential downstream effects of* aaaA *expression uncovered in this study. *Positive regulation is indicated by a green arrow (dotted arrow if likely to be indirect with unknown intermediaries), negative regulation by a flat headed red arrow and enhancer binding protein regulators by dashed black arrows. Question marks denote relationships which are theoretical or have limited data to support the proposed action/mechanism. Created using BioRender.com.*

Chapter 6. Probing AaaA localisation and function in a synthetic chronic wound model

6.1 Chapter aims

In addition to the insights gained into *aaaA* regulation and expression in Chapter 5, this chapter aimed to further examine the function and localisation of AaaA, both at the single cell level, and in the SCW. This was done by:

1. Designing a Spytagged-AaaA which could be used to localise AaaA *in-situ* first on single cells, and later within a biofilm at the community level.
2. Quantifying extracellular arginine in the SCW, to determine with direct evidence if, and to what extent, AaaA is releasing arginine as theorised.

6.2 Localising AaaA using the Spytag-Spycatcher system

6.2.1 Previous work to localise AaaA using microscopy

Localisation of AaaA recombinantly expressed in *E. coli* was first attempted by Ashawesh (2016), was continued by Alhagh (2020) and produced interesting insights into AaaA cellular localisation. Ashawesh (2016) inserted a tetra-cysteine 'FlAsH' tag (CCPCCC) into a predicted flexible loop region in the M28 domain of AaaA, then following induction of the FlAsH-AaaA, exposed the cells to the fluorescent FlAsH-EDT_2 substrate, which bound to the FlAsH tag, fluorescently labelling AaaA. Using this, they demonstrated that AaaA was localised in a helical or 'spiral' distribution around the cell, similar to the *E. coli* autotransporter EspC. They and Alhagh (2020) also showed that when they incorporated an amber STOP codon to truncate out the β-barrel, the truncated AaaA did not adopt this pattern, instead aggregating in discrete foci. Unlike for

EspC, this apparent spiral localisation seen with AaaA was independent of the putative accessory factor SlyB (Alhagh, 2020).

However, it has been shown that the FlAsH-EDT_2 substrate binds to endogenous cysteine-containing proteins, creating high levels of background fluorescence which can only be overcome by high levels of expression (Stroffekova *et al.*, 2001). As well as being expressed at only low levels, AaaA also contains two consecutive cysteine residues at position 232-233, which likely bound the FlAsH-EDT_2 substrate non-specifically, as signal could be detected in the non-FlAsH-tagged (native) AaaA control. This was confirmed by mutating these cysteines, which abolished this substrate-independent background signal (Ashawesh, 2016).

The substrate, FlAsH-EDT_2, also penetrates the Gram-negative cell membrane, which for Ashawesh (2016), was beneficial in localising AaaA-trafficking to the cell surface. However, localisation of both intra- and extra-cellular populations of AaaA would make it difficult to distinguish the population of mature, surface-expressed AaaA which plays a functional role for the cell. Thus, the aim in this study was to develop a construct of functional AaaA that could be localised on the cell surface with high substrate-binding specificity, with the aim of localising AaaA *in-situ* at both the cellular and biofilm level. To that end, the Spytag-Spycatcher system was employed.

6.2.2 Background on the Spytag-Spycatcher system

The Spytag-Spycatcher system was first developed by the Howarth group at the University of Oxford, and utilises the internal isopeptide bond formation found in the CnaB2 domain of the *Streptococcus pyogenes* virulence factor FbaB, a fibronectin-binding protein (Hatlem *et al.*, 2019; Zakeri *et al.*, 2012). The reaction occurs spontaneously in this domain between the ε-amine of lysine K31 and the side chain carboxyl of aspartate D117, catalysed by the adjacent glutamate E77. They split this domain into two binding partners; Spycatcher, a 138 residue (15kDa) incomplete immunoglobulin-like domain containing the reactive lysine and catalytic glutamate, and Spytag, 13-residues which include the reactive aspartate (Reddington & Howarth, 2015).

The Spytag can be engineered into either termini or within an internal loop of a protein, and when exposed to Spycatcher, they will spontaneously form a covalent, isopeptide bond with each other, as illustrated in **Figure 6.1**. This reaction occurs within minutes, and over a range of different temperatures, pH and buffer conditions making it a versatile and elegant labelling system. The Howarth group later developed improved Spytag/Spycatcher-002 and -003 versions, each with faster rates of reaction by orders of magnitude, and made the technology commercially available (Keeble *et al.*, 2017, 2019).

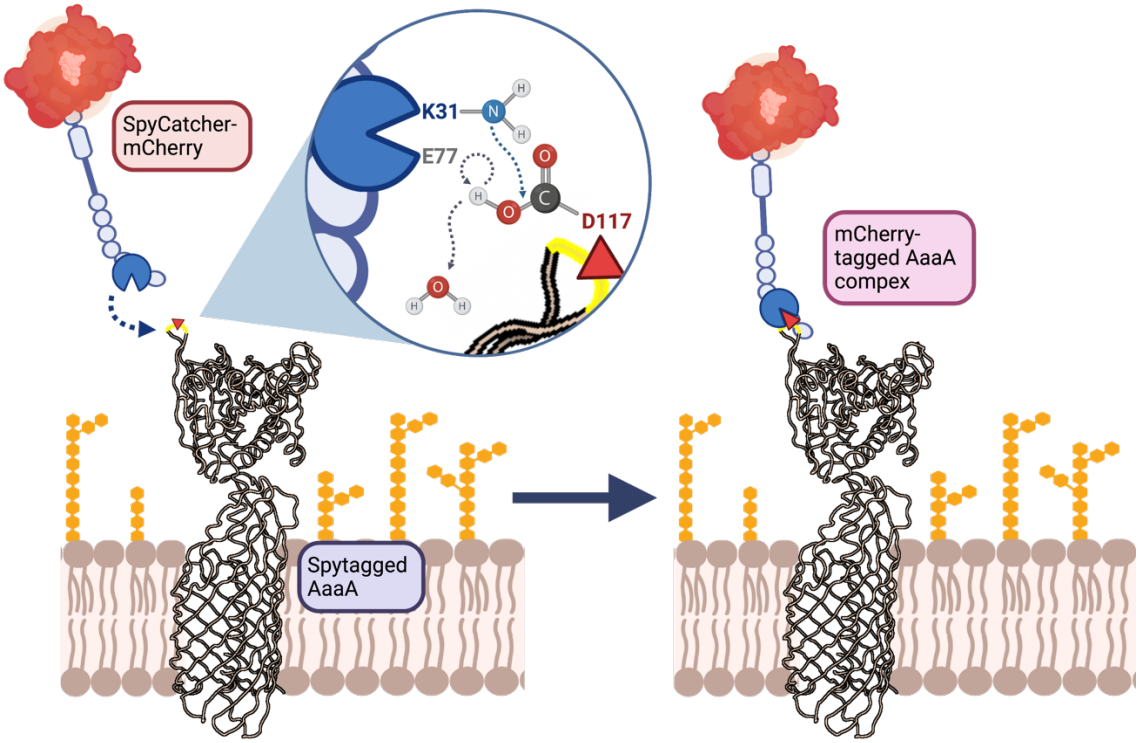


Figure 6.1: Illustration of the Spytag-Spycatcher system for labelling AaaA. *The Spycatcher binding partner, which here is fused to mCherry, forms a covalent bond with Spytag, which has been engineered into a flexible region of AaaA. The reaction occurs by nucleophilic attack of the Asp117 in the Spytag, by Lys31 in the Spycatcher, facilitated by proton transfers (dotted lines) involving Glu77, also in Spycatcher, which forms intermediates. The end product is a stable isopeptide bond and the release of water. N.B. Illustrated amino acid numbers refer to the original positions in the S. pyogenes FbaB CnaB2 domain from which the isopeptide bond originates, rather than their positions in the constructs depicted here. Created using Biorender.*

A version of the Spycatcher protein which is fused to a super-folder GFP protein, as well as containing MBP and poly-histidine tags, was created to quickly and easily label OM proteins bearing the Spytag (Chauhan *et al.*, 2019). Vectors bearing this construct, as well as a yet unpublished version bearing

mCherry instead of GFP, were kindly donated by Dr Jack Leo, for this work. This novel mCherry-Spycatcher construct was used in this chapter to overcome the green-channel (500–560 nm) autofluorescence of *P. aeruginosa* (Yang *et al.*, 2012).

6.2.3 Design of a Spytagged-AaaA construct

Because it is unable to penetrate cell membranes, the Spycatcher/Spytag system is particularly useful for labelling outer-membrane proteins such as AaaA. The previous work by Ashawesh (2016), discussed in 6.2.1, was used as a framework for engineering Spytag002 into AaaA.

The structure of AaaA produced using C-iTasser (**Figure 3.2**) was used to select a range of locations within the N-terminal and flexible loop regions of AaaA to insert the Spytag. Along with the location used by Ashawesh (2016) (L291-G292) for their FlAsH tag, five other locations were selected based on their potential for flexibility and accessibility, as shown in **Table 6.1**.

Table 6.1: Summary of location choices for 6 Spytag construct candidates. *Predictions were taken from the data sheet for the CiTasser model1 of native AaaA. The colour code for each tag, listed in the 'annotations' column, are used from here onward.*

Name	ST location	Predicted SS (Conf.)	Predicted SA	Annotation
STv1	L291:G292	H:H (4:2)	0:1	Blue
STv2	L61:G62	H:H (9:6)	0:7	Red
STv3	A75:G76	C:C (5:7)	5:4	Purple
STv4	G189:S190	C:C (5:7)	2:0	Orange
STv5	G221:L222	C:C (8:8)	4:1	Pink
STv6	A29:G30	H:H (9:9)	0:2	Cyan

ST: spytag, SS: secondary structure (confidence score), SA: solvent accessibility (scored 0, buried to 9, highly exposed)

Due to their slightly different methodologies, iTasser and CiTasser produce variations in the predicted structure of the M28 region (Roy *et al.*, 2010; Yang *et al.*, 2015; Yang & Zhang, 2015; Zheng *et al.*, 2021). Though CiTasser is more sophisticated, it is also more computationally expensive, and given that the active site residues align well in both models, iTasser was used to model the effects of Spytag insertion into the M28 domain. The amino acid sequence of the region Y23-F317 (which includes the extra Spytag residues)

for each construct was fed in succession into iTasser, and the model1 for each was aligned using PyMOL to the Y23-F303 region (M28 passenger domain) of PA0328. As shown in **Figure 6.2**, in every construct, the residues involved in localisation and active site formation align well between the wild-type and Spytagged AaaA, although overall alignment quality, and the predicted accessibility of the Spytag vary.

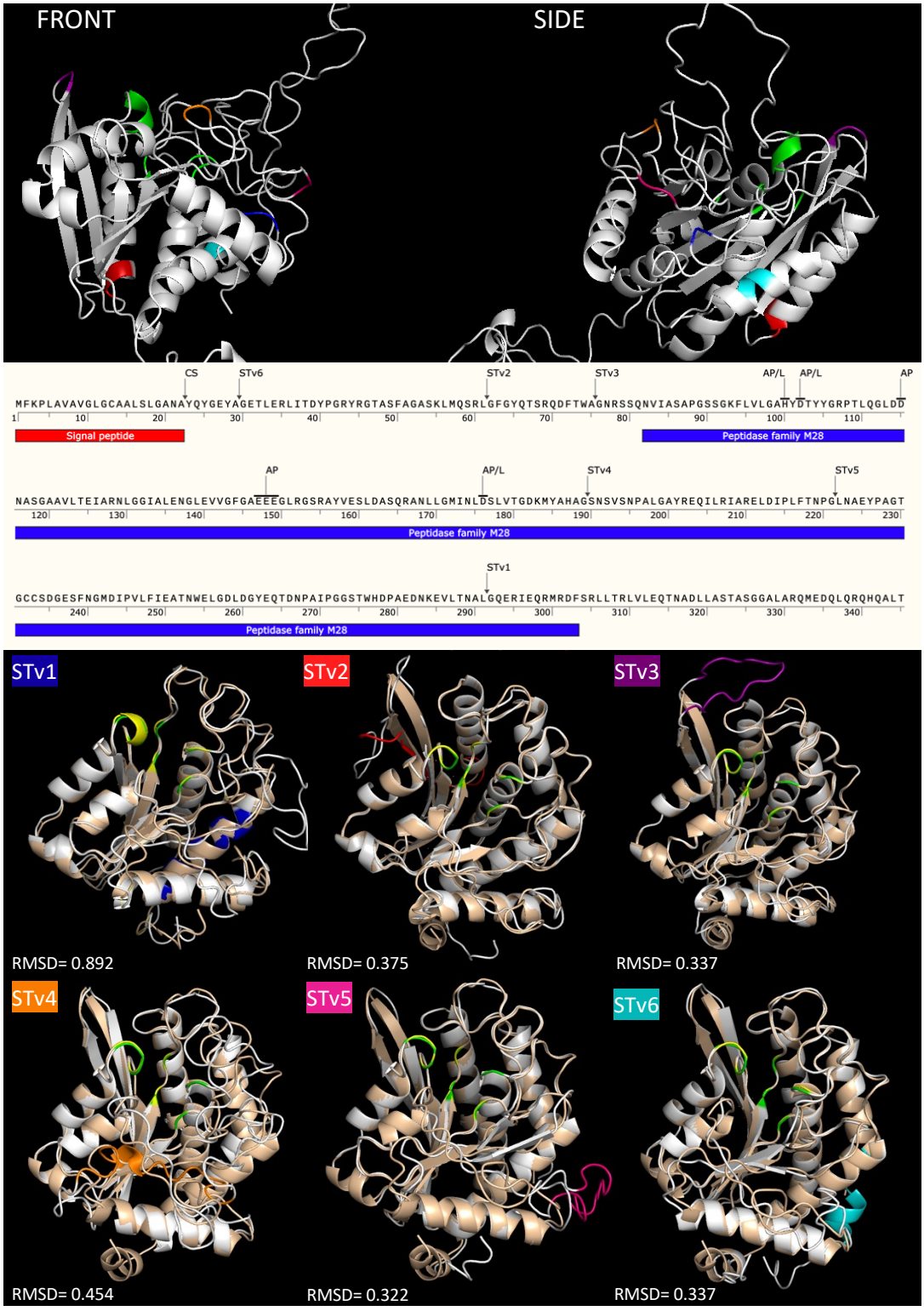


Figure 6.2 Design of Spytag-containing AaaA constructs. Top panel *shows C-iTasser predicted model1 of AaaA, annotated with the location of each Spytag version (STv1-6) coloured according to* ***Table 6.1****. Residues essential for AaaA localisation/ activity (AP/L) are green. Spytag and AP/L locations are shown on the primary AaaA structure in the* ***middle panel****. The* ***Bottom panel*** *shows alignments between iTasser-generated AaaA M28 peptidase domain (Y23-F303, wheat, AP/L residues labelled yellow) and the equivalent models of each Spytagged construct (white, Spytag coloured according to the version, AP/L residues coloured green). The two structures were aligned using PyMOL and the alignment RMSD (Å) is displayed below each.*

6.2.4 Use of the pJM220 vector for rhamnose induction of Spytagged AaaA in *P. aeruginosa*

AaaA has low native levels of expression, as shown in 5.6.2, so to ensure AaaA was easily detectable, an overexpression vector was required. As mentioned in 1.1.1, IPTG can interfere with biofilm formation, therefore a different induction system, compatible with *P. aeruginosa* was chosen. The pJM220 vector (long name: *pUC18T-miniTn7T-gm-rhaSR-PrhaBAD*, accession no. KX777256), designed by Meisner et al. (2016) uses the *rhaSR-PrhaBAD* inducible promoter system, first identified in *E. coli* (Egan & Schleif, 1993). This promoter system allows for tight control of gene expression over a broad range, with very little 'leaky' uninduced expression. This vector also inserts via the *att::Tn7* chromosomal integration site, allowing for stable single-copy insertion. Therefore, it was chosen as the backbone for induction of both native and Spytagged AaaA over-production, to better assess the activity and binding efficiency of each Spytagged construct. The vector was constructed by HiFi assembly and, as discussed in 6.2.8.2, there is also potential in future to construct a Spytagged-AaaA under the control of the *aaaA* promoter using the miniCTX-1 vector, which would allow for native expression levels.

6.2.5 Construction of Spytagged AaaA

Figure 6.3 outlines the cloning strategy taken to construct the Spytagged AaaA and insertion into pJM220. Mutagenic primers for the six constructs, listed in **Table 2.3**, were first used to linearise pBlueScript*::aaaA* and introduce the Spytag by inverse PCR. Successful amplicons for Spytag version (Stv)4, 5 and 6 (**Figure 6.3a**), were taken forward, re-circularised, transformed into ultracompetent *E. coli DH5α* (**Figure 6.3b**) and sequenced (**S-Figure 0.5a**). Colonies containing correct alignments for each of St4-6, as well as non-Spytagged *aaaA* (from pBlueScript*:aaaA,* not shown) were taken forward for insertion into pJM220 by HiFi assembly. Again, both the insert from each respective pBlueScript vector as well as the pJM220 vector were linearised by PCR, and successful amplicons (**Figure 6.3c**) were selected for HiFi assembly and transformed into *E. coli* DH5α (**Figure 6.3d**). Extracted

plasmids were sequenced (**S-Figure 0.5b**), and those which successfully aligned for each of the four constructs were transformed into *PAO1L* and *PAO1L ΔaaaA* for *Tn7* insertion, using the helper vector *pUX-BF13* (Lambertsen *et al.*, 2004).

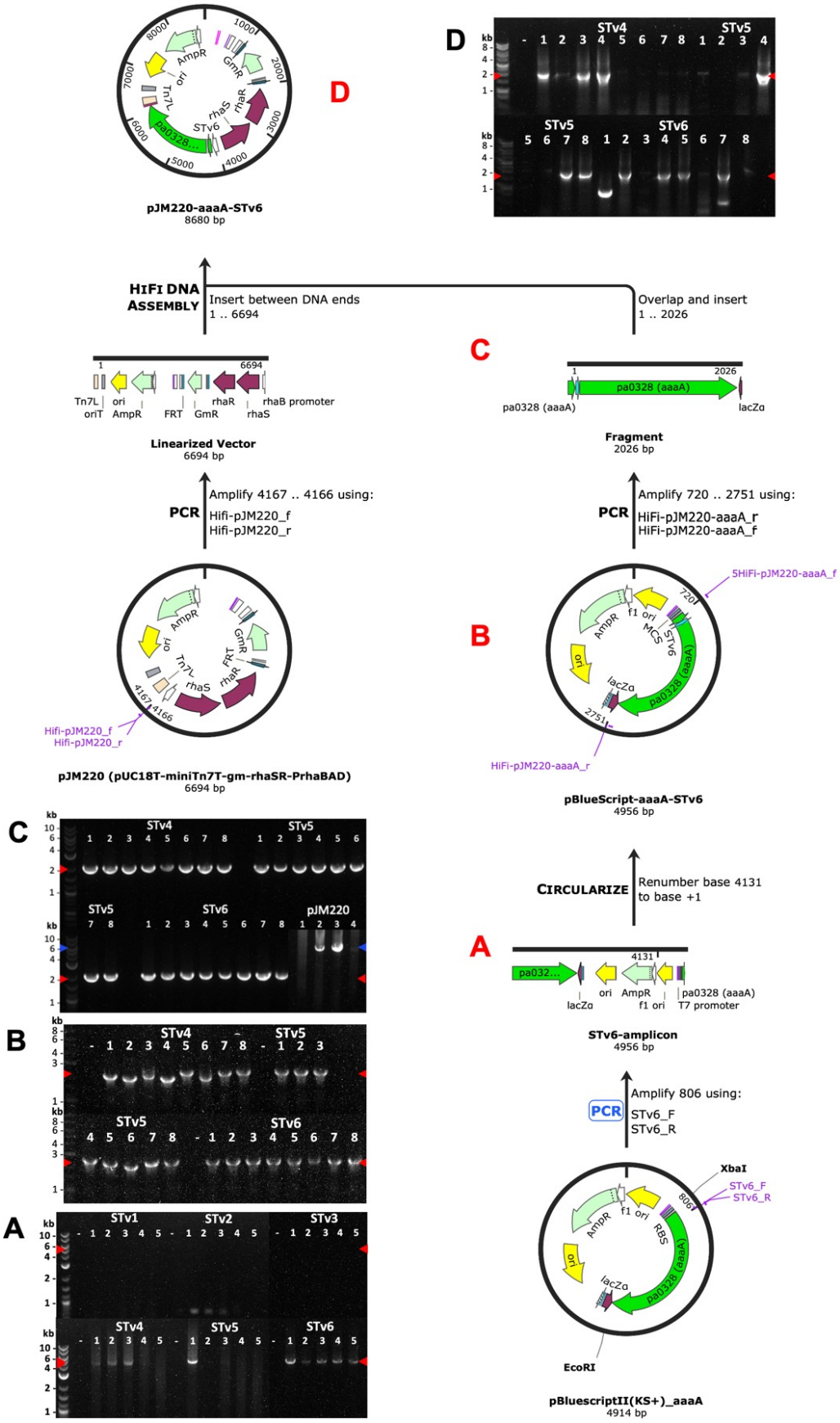


Figure 6.3: Cloning strategy for the construction of Spytagged* aaaA *and insertion into the* pJM220 *rhamnose-inducible mini* Tn*7 vector. *A Snapgene-*

generated schematic for the construction of the representative Spytag version6 (STv6), runs from bottom to top, with the corresponding steps shown in other panels annotated in red letters. ***Panel A*** *shows the gradient inverse PCR to insert Spytag versions 1-6 (STv1-6) in to* pBlueScript::aaaA*. Expected product of 4956 bp (red arrows) was taken forward for STv4 (3, Tm = 64.1 ℃) and STv5 and 6 (1, Tm= 58.4 ℃).* ***Panel B*** *shows the colony PCR of* E. coli *DH5α* pBlueScript::aaaA-STv4-6 *transformants using M13 primers to amplify an expected product of 2023 bp. Colony 7 and 8 were taken forward for STv4, and colonies 2 and 3 for each of STv5 and 6, for Sanger sequencing.* ***Panel C*** *shows the gradient PCR for amplification of HiFi assembly products to insert* aaaA-STv4-6 *into* pJM220*. DNA fragments for the* pJM220 *linearised vector (6694 bp) and* aaaA-ST *linearised fragments (2026 bp) are indicated by blue and red arrows, respectively. Product 8 (Tm= 65 ℃) was selected for each insert, and product 3 for the vector. Products were cleaned and assembled using HiFi assembly master-mix and transformed into ultracompetent* E. coli *DH5α.* ***Panel F*** *shows the colony PCR to screen clones using the primers Hifi-*pJM220*_F and R. Expected products (2026 bp) are indicated by red arrows). In all gels, - indicates the no template control.*

6.2.6 The pJM220 vector can be used to tightly control AaaA expression in a concentration-dependant manner

As shown in **Figure 6.4**, *aaaA* expression can be tightly regulated by rhamnose induction using the *rhaSR-PrhaBAD* promoter system over a broad range, in a concentration-dependant manner. These data also confirm a lack of ‘leaky’ expression, with the uninduced cells having activity levels no different to PAO1L Δ*aaaA*. These data also provide an ‘induced wild-type’ level of activity to compare with Spytagged constructs, to assess whether insertion of each Spytag has a detrimental effect on AaaA translocation or activity.

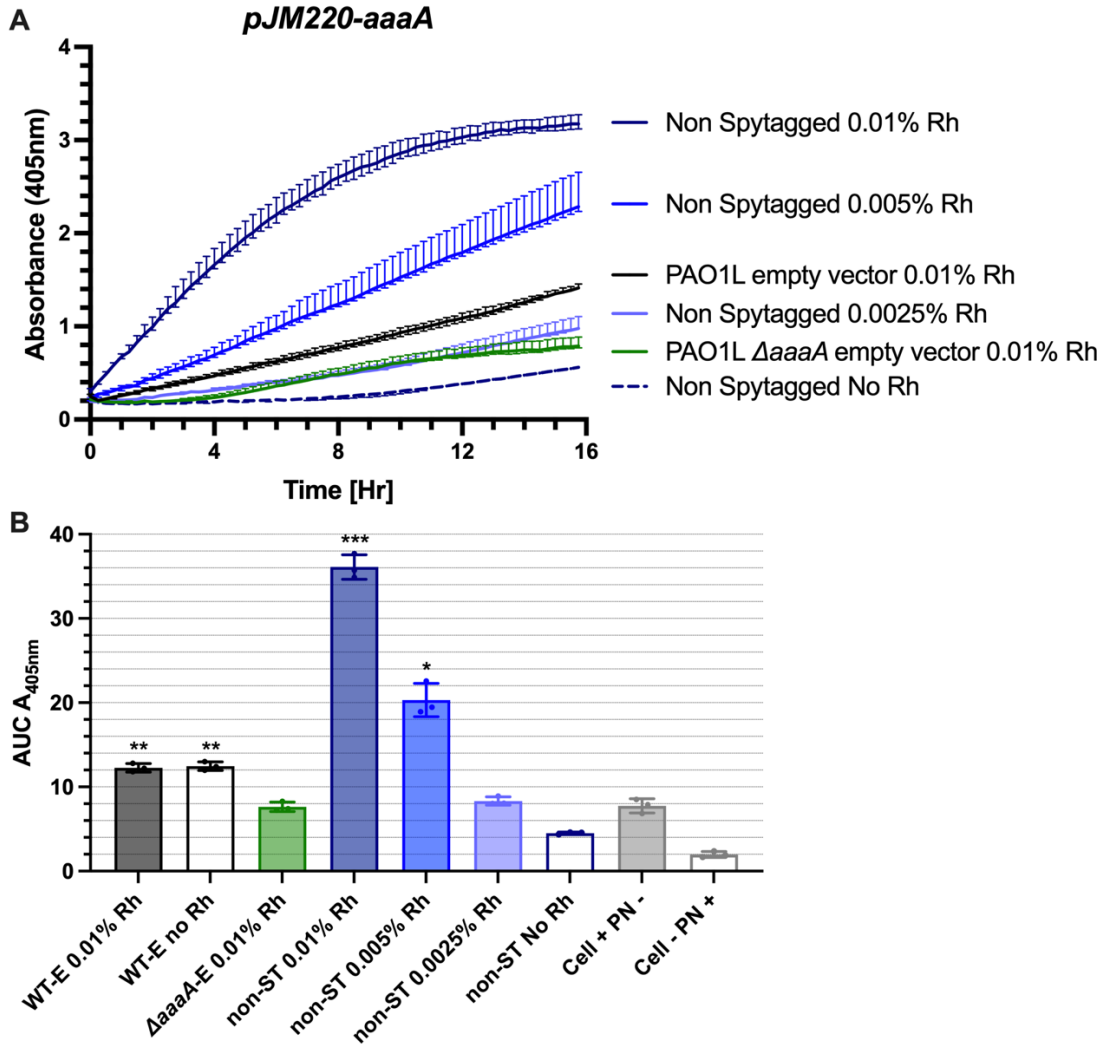


Figure 6.4: AaaA activity can be tightly induced in the* pJM220-aaaA *construct. *Whole cell L-arginine-p-nitroanilide cleavage assay comparing AaaA activity in PAO1L wild type (WT) and* ΔaaaA *carrying empty pJM220 (-E) or* pJM220::aaaA (non-Spytagged/non-ST)*. Overnight cultures were inoculated at* OD_{600nm} *0.05 into 50 mL flasks containing 10 mL LB supplemented with either 0.01%, 0.005% or 0.0025% rhamnose. Following washing, a final* OD_{600nm} *of 0.5 was added to the activity assay. Data is shown over time in* ***Panel A*** *as median Absorbance (405nm) ±95% CI and in* ***Panel B*** *as mean area under the curve (AUC) ±SD (n=3). Differences were determined by Brown-Forsythe and Welch's ANOVA test with Dunnett T3 multiple comparisons, comparing each group to the* ΔaaaA *empty vector control *p≤0.05, **p≤0.005, ***p≤0.001.*

6.2.7 Spytagged AaaA version 4 is active and accessible to Spycatcher binding

Once inserted into *P. aeruginosa*, a whole cell AaaA activity assay was performed on each Spytagged construct, following induction with 0.005% rhamnose. **Figure 6.5** shows that both AaaA-St4 and AaaA-St5, but not AaaA-St6, are active enough to rescue the AaaA phenotype in PAO1L *ΔaaaA*. As L-

arginine-p-nitroanilide should not penetrate the cell membrane, activity in the whole cell assay also indicates that AaaA surface expression has been retained.

To further test whether the Spytag on each construct was accessible to binding with the Spycatcher binding partner, Spycatcher-mCherry (Jack Leo, Nottingham Trent University, unpublished, 2022), a binding assay was undertaken, based on the method used by Chauhan and colleagues (2019), described in 2.12. As shown in **Figure 6.6**, only AaaA-St4 appears to bind Spycatcher-mCherry better than the non-Spytagged control. This makes sense for St6, as its low activity suggests poor expression or surface translocation. The reason for lower accessibility in St5 is less clear, though it is most likely due to some or all of the Spytag residues being buried and inaccessible for binding.

Due to its favourable activity and binding profile, AaaA-St4 was taken forward. It is worth noting however, that the Spycatcher binding assay has proven to be sensitive to cell density, incubation length and other factors which can reduce its replicability (data not shown). Optimisation of the protocol, including increasing the starting number of cells, would be ideal to improve its reliability and gather the most robust data possible on the Spycatcher binding accessibility for each Spytag construct. Moreover, in their study, Chauhan et al. (2019) completed three separate biological repeats to improve the robustness of their data. This was not possible here due to constraints on time and reagent availability, but it would be essential to include going forward.

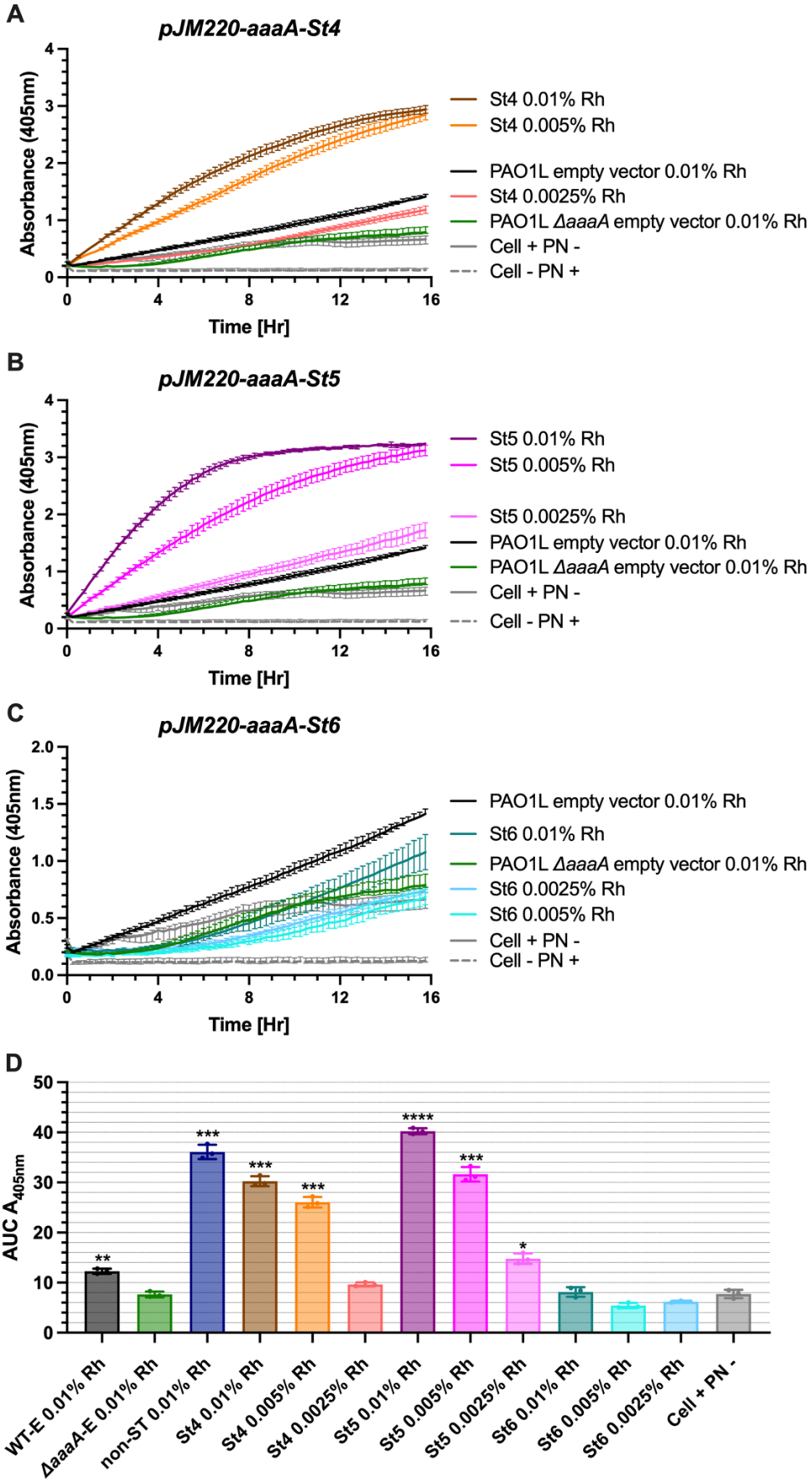


Figure 6.5: AaaA activity assay in rhamnose-induced Spytagged AaaA shows both AaaA-St4 and -St5, but not -St6 have similar activity levels to native AaaA

***(non-Spytagged) induction**. Overnight cultures of PAO1L ΔaaaA* pJM220::*aaaASt4-6 were inoculated at* OD_{600nm} *0.05 into 50 mL flasks containing 10 mL LB supplemented with 0.01%, 0.005%, and 0.0025% rhamnose (Rh). Flasks were incubated at 37℃, 200 RPM for 2.5 hrs, with final* OD_{600nm} *between 0.8-1.2. Following washing, a final* OD_{600nm} *of 0.5 was added to the activity assay. Data is shown in **Panel A-C** as median absorbance (405nm) ±95% CI over time for St4-6, respectively. **Panel D** shows the aggregated area under the curve (AUC) comparing all constructs (mean ±SD, n=3). Activity of data from the non Spytagged* PAO1L Δ*aaaA* pJM220::aaaA*, induced with 0.01% rhamnose, from **Figure 6.4** is shown for comparison. Differences were determined by Brown-Forsythe and Welch's ANOVA test with Dunnett T3 multiple comparisons, comparing each group to the ΔaaaA empty vector control, *p≤0.05, **p≤0.005, ***p≤0.001.*

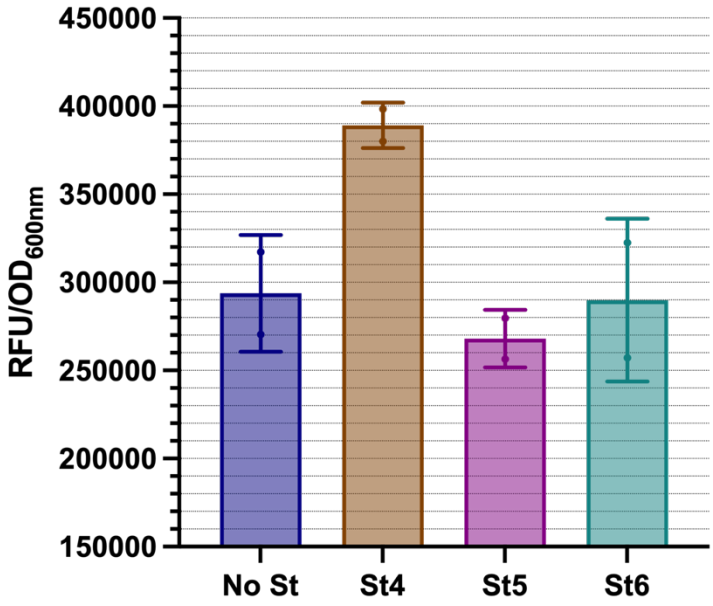


Figure 6.6: Spytag-Spycatcher binding assay shows that only aaaA-St4 binds Spycatcher-mCherry more than the non-Spytagged AaaA.

6.2.8 Spytagged AaaA can be visualised on the surface of single cells

The native (non-Spytagged) and St4 construct were compared using super resolution microscopy with structured illumination as described in 2.10.2. As shown in **Figure 6.7**, clusters of fluorescent signals could be seen associated with the surface of cells expressing Spytagged, but not native, AaaA. This supports previous findings that AaaA is localised to the cell surface (Luckett *et al.*, 2012). Clear conclusions regarding the distribution however cannot be made as resolution was poor due to a large proportion of cells remaining motile.

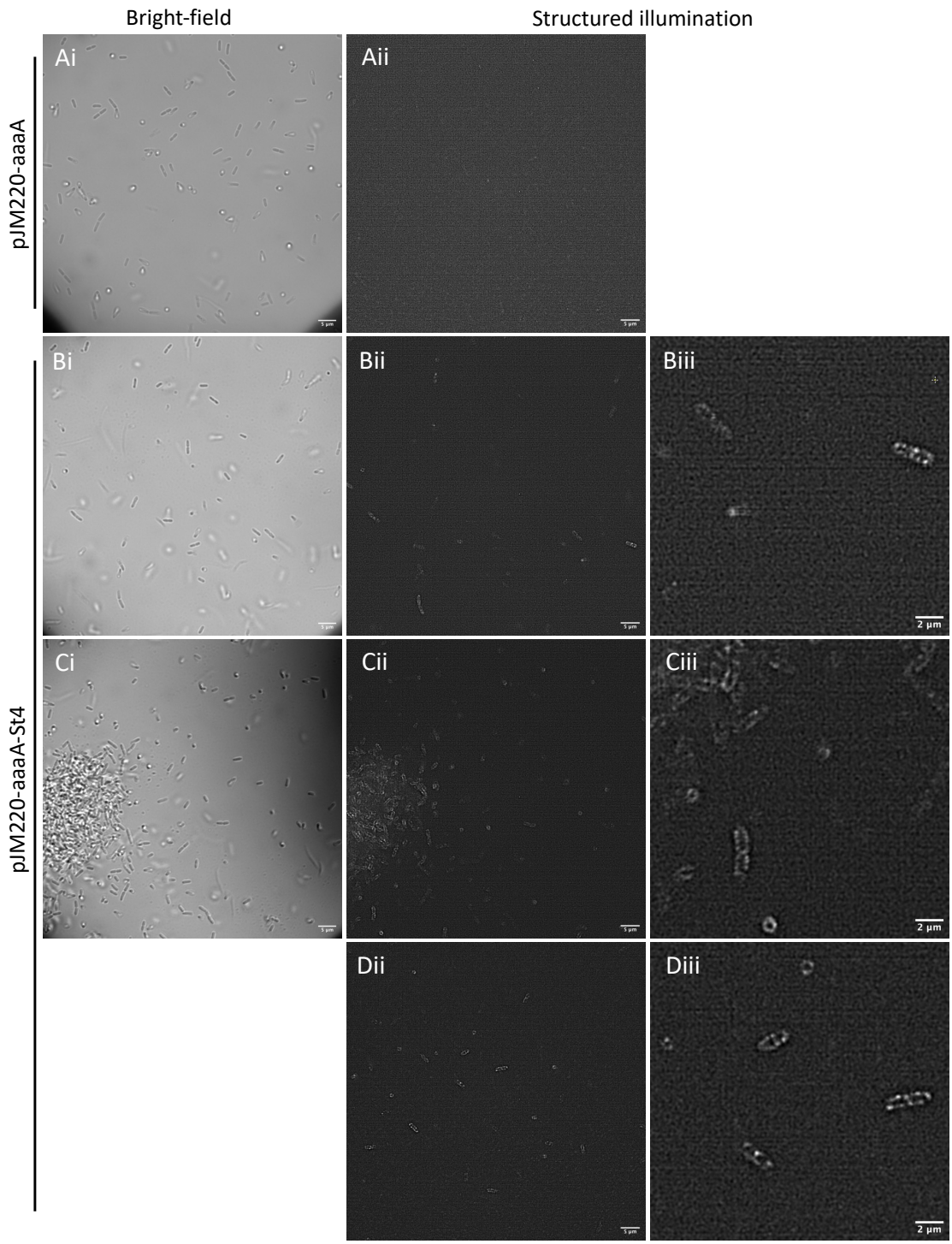


Figure 6.7: Structured illumination microscopy captures Spytagged AaaA on the surface of single cells. *Samples were prepared as described in 2.10.2. Brightfield is shown on the left (i), and fluorescence with structured illumination (λ587/610) on the right (ii, with iii showing an enlarged view of ii).* ***Panel Ai-ii*** *show the non-Spytagged control (PAO1L Δ*aaaA* pJM220::*aaaA*).* **Panel Bi-Diii** show three *separate fields of view from the same sample of PAO1L Δ*aaaA* pJM220::*aaaA-St4. Brightfield snapshots were unavailable for the third field (Di) due to technical issues. Scale bars are 5 µm in Ai-Dii and 2 µm in Biii-Diii.*

Fixation of the cells was attempted to improve upon image resolution. Briefly, on the same day, 30 µL of the samples used in **Figure 6.7** were incubated with 30 µL 0.2% glutaraldehyde for 13 min. Cells were washed in PBS and 20 µL of fixed cells were added to a coverslip, allowed to dry, and pressed onto a glass slide. As shown in **Figure 6.8**, using this method high levels of fluorescent signal could now be detected on the surface of the cells in the Spytagged, but not non-Spytagged control.

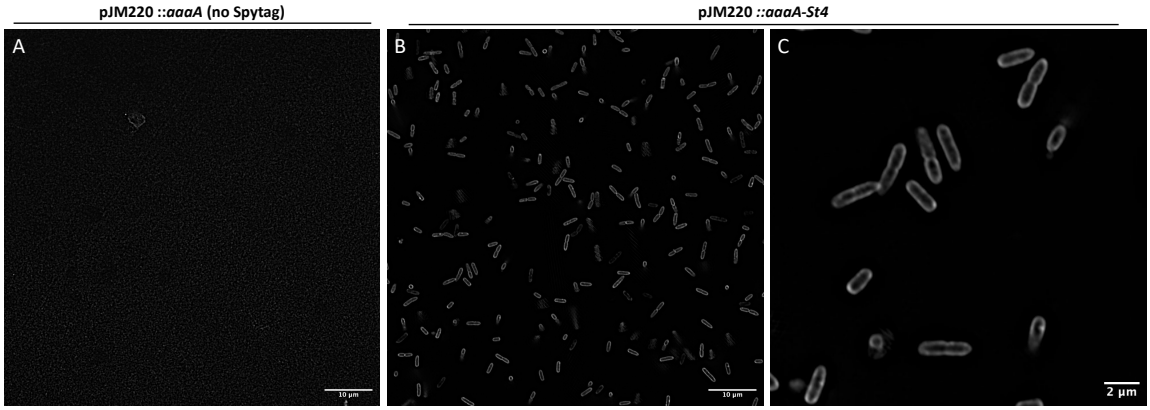


Figure 6.8: Spytagged-AaaA can be detected by structured illumination microscopy on the surface of fixed cells. Panel A *shows fixed non-Spytagged control (PAO1L* ΔaaaA pJM220::aaaA). *Panel* **B** shows fixed *PAO1L* ΔaaaA pJM220::aaaA-St4 (enlarged view in **Panel C**)*. Samples were prepared as described in 2.10.2 and fixed as described in* ***Error! Reference source not found.****. Scale bars are 10 µm in Panels A and B, and 2 µm in Panel C.*

6.2.8.1 Signs of cell toxicity due to over-expression of AaaA-St4

Further attempts on later dates were made to repeat this protocol. However, fluorescent signal could not be detected on the surface of cells, but rather was associated with what looked like cell fragments and only at very high laser power (not shown). Moreover, in later experiments, when examined prior to immobilisation, cell motility as observed under the bright-field lens was very low, indicating the bacteria were not healthy, suggesting that induction of AaaA-St4 was toxic to the cell. This makes sense, as based on the images in **Figure 6.8**, levels of AaaA-St4 expression were sufficient to completely cover the cell surface, which is likely to be placing the cell membrane under great stress. Further work is therefore needed to replicate the experiment with much lower rhamnose induction concentrations more consistent with wild-type AaaA expression levels.

6.2.8.2 Construction of a miniCTX-1::*aaaA*-St for visualisation of native AaaA expression

In addition to the insertion of Spytagged AaaA into pJM220, insertion of *aaaA-St4* into the existing *aaaA* complementation vector, miniCTX-*1::aaaA*, with *aaaA* being swapped for *aaaA-St4* for insertion into *P. aeruginosa* was initiated. The purpose of this was to create a complementation vector containing a Spytagged AaaA, with its native promoter. This could then be used for localisation studies, for example, to monitor native levels of AaaA expression within a biofilm and could potentially circumvent the issue of toxic over-production of AaaA. The same methodology for insertion into pJM220, described by **Figure 6.3**, was used for insertion into miniCTX-*1::aaaA*. However, it was not possible to complete this due to time constraints.

6.3 Testing the arginine liberation hypothesis for AaaA in SCW

It has been proven that AaaA is an arginine-specific aminopeptidase, both by the L-arginine-p-nitroanilide degradation assay, used as a proxy for AaaA activity, and by the observation that *aaaA* is essential for survival when the tripeptide Arg-Gly-Asp was the only carbon and nitrogen source (Luckett *et al.*, 2012). However, the effect of AaaA on the local arginine concentration has not yet been studied, and evidence of the presence of AaaA leading to a rise in local arginine concentrations has not been directly observed. Therefore, commercial arginine quantification kits were used to examine the effect of AaaA on extra-cellular L-arginine concentration in the SCW.

6.3.1 There is no difference in free arginine between wild-type and Δ*aaaA*-infected SCW, but arginine levels are reduced by infection

To determine whether AaaA globally increases arginine, extra-cellular L-arginine was quantified in 16 hr SCW samples (slide mounted) using a commercial arginine quantification kit, as described in **Error! Reference source not found.**. No difference was detected due to AaaA, but arginine levels were significantly higher in uninfected samples (**Figure 6.9**). This

indicates that arginine catabolism is occurring in the SCW, supporting a role for AaaA. Thus, if arginine is being utilised, there may be a need for arginine scavenging, potentially at later time points when oxygen, carbon and nitrogen sources are scarcer. It is also possible that any increases in local arginine produced by AaaA are immediately taken up into the cell, thus would not be detected by an extracellular assay.

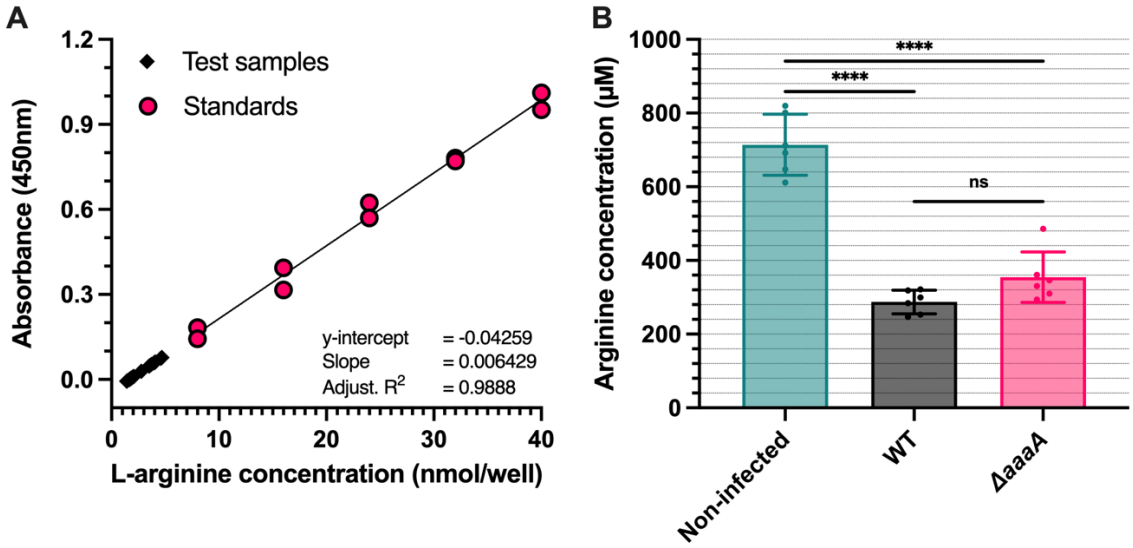


Figure 6.9: Colorimetric arginine quantification assays to compare arginine levels in the SCW between PAO1L wild-type (WT) and ΔaaaA. Panel A*: Shows the standard curve for L-arginine standards (pink circles) and test samples (black diamonds) interpolated from standard curve (Linear best-fit values shown in inset table) from the colorimetric arginine quantification assay.* ***Panel B:*** *shows the dilution factor-adjusted L-arginine concentration values for test samples (bar at mean ± SD, n= 6, 3 biological x 2 technical replicates). Differences were tested using Brown-Forsythe and Welch's ANOVA test with Dunnett T3 multiple comparisons (****p≤0.0001, ns: not significant).*

Absorbance for some of the infected samples was close to the limit of detection for this assay of 1 nmol/well (the lowest value was 1.411 nmol/well). Therefore, a more sensitive fluorometric assay, with a lower detection limit of 100 pmol/well was used to improve arginine quantification precision (Abcam, 2019b). However, the data generated were low quality and so are not included, although further attempts may be fruitful in future. Future experiments could alternatively use a larger starting quantity of sample, or samples spiked with a known concentration of an L-arginine standard, which can be adjusted for in analysis, to lift arginine values further into the detection range. It would also be useful include a AaaA-over-expressing strain such as PAO1L pJM220*::aaaA* to

examine if in an increase in AaaA production correlates with detectable changes in extracellular arginine.

6.4 Discussion

6.4.1 AaaA can be localised at the single cell level using the Spytag-Spycatcher system

A novel Spytagged AaaA was constructed for the purpose of localising AaaA, both on the cell surface, and within a biofilm, including in the SCW. The Spytagged AaaA was shown to be both active and accessible to Spycatcher-mCherry and it was possible to visualise it localised to the cell surface by single cell microscopy. However, as discussed in 6.2.8.1, these findings could not be replicated, as on two separate occasions, once the cells had been induced and prepared for imaging, signal from the bound Spycatcher could not be detected associated with live cells. The low level of signal that was present may have been associated with cell debris, or even OMVs, but it was not possible to determine this. The lack of cell motility, only in these later experiments, was indicative of cell toxicity. Therefore, further work is needed to optimise the induction conditions to attain high, but replicable levels of expression without compromising cell integrity.

6.4.1.1 Over-production of Spytagged AaaA likely caused cell toxicity

Toxicity in this case is not entirely surprising as over-expression of membrane proteins is a known and pervasive issue in membrane protein biology (Miroux & Walker, 1996). Indeed, Ashawesh (2016) also previously found that over-expression of FlAsH-tagged AaaA in *E. coli* impeded cell growth, and lead to cell elongation. A benefit of using the pJM220 vector was that it allows for very tight control of induction, over a broad concentration range (Meisner & Goldberg, 2016). This was shown to be true for AaaA in this study too, and therefore a lower concentration of rhamnose should be tested, to reduce the stress on the cell caused by induction, as mentioned above.

As shown in **Figure 6.5**, at the induction concentration used for microscopy, 0.005%, activity of the AaaA-St4 was far higher than that of the native wild-type AaaA level. This was intentional at this point as it was hoped

a higher expression level would increase the chances of detection, which, once validated, could be tweaked to reflect native expression levels. Given the toxicity that induction with 0.005% rhamnose appears to cause, it would be beneficial to repeat microscopy with the lower 0.0025% induction concentration, which more closely tracked with native AaaA activity levels.

It may also be beneficial to take a step backwards and express the construct in an *E. coli* strain adapted to membrane protein expression, such as the LEMO strain (as discussed in 4.4.2). AaaA has been shown to be properly translocated and retain activity in *E. coli*, thus replications of the findings from **Figure 6.7** in *E. coli* will provide further validation of these data. One of the key benefits of choosing a Tn7 insertion-site vector was its strain versatility, to allow for this eventuality. However, expression in the native *P. aeruginosa* background will always be the ideal.

Another option is to pursue expressing *aaaA-St4* under the native *aaaA* promoter, to both avoid cell toxicity, and also directly observe native levels of AaaA distribution *in-situ*. Attempts were therefore made, but not completed, to clone the AaaASt4-6 constructs into the existing miniCTX-*1::aaaA*, replacing the non-Spytagged *aaaA*. The benefit of unlocking this crucial capability warrants further efforts to complete this construct and allow observation of native-AaaA localisation for the first time.

6.4.1.2 Localisation of Spytagged AaaA in a biofilm

Another major application for the Spytagged-AaaA is to improve specificity of immunoblots, to better quantify AaaA production. Quantification of native levels of AaaA production in this way could support the transcriptional data gathered here, in 5.6.2, to better understand the stages of infection and environment where AaaA is most relevant. Due to time constraints, it was not possible to test this application of the Spytagged AaaA, but initial validation of the system at the single cell level indicates it may be a promising construct for this work. Moreover, alongside immunoblotting, the Spytagged AaaA under native P_{aaaA} control could be used to detect AaaA expression at the community level, to identify for example, pockets within a biofilm where AaaA is more or less present.

6.4.1.3 Potential uses for Spytag-Spycatcher system in purifying AaaA

Attempts to purify AaaA, as discussed in Chapter 4, have been largely unsuccessful to date. It is possible that the Spytagged AaaA construct created here may also be useful for purification purposes, as there are commercially available Spycatcher partners containing histidine, streptavidin and MBP tags, which would allow for very specific capture of AaaA. However, the Spycatcher partner irreversibly forms a covalent bond upon contact with the Spytag, meaning they could not be easily separated following purification. It is unclear whether the addition of the Spycatcher motifs to the protein would help or hinder crystallisation, and this would need to be empirically determined.

6.4.2 AaaA did not observably alter free arginine levels in the SCW, however arginine utilisation was likely occurring

Quantification of extracellular arginine showed that while significantly higher in uninfected SCW samples, there was no detectable difference in extracellular arginine between wild-type and Δ*aaaA* strains. This suggests that in the SCW, arginine is being catabolised, and if so, there may be a need for arginine scavenging, potentially at later time points when nutrient sources are scarcer.

In their thesis, Kiney-Whitmore (2017) used a 'PEG-lid' biofilm assay to determine the effect of exogenous arginine and RGD tripeptide on biofilm formation. They grew 16 hr biofilms in ASM then stained them using crystal violet to quantify biofilm biomass. Their hypothesis was that arginine would increase biofilm formation, and that AaaA would cleave the tripeptide to also liberate free arginine, which would in turn also increase biofilm formation similarly to that by exogenous arginine. They did observe that growth in 5 mM arginine lead to greater biofilm biomass, however there was no difference in biofilm biomass when grown in 5 mM tripeptide. There also did not appear to be a difference in biofilm formation between the wild type and Δ*aaaA* in either experiment, although these experiments were done using the PAO1N background.

It is unclear why AaaA did not appear to liberate arginine to produce a phenotype similar to that when exogenous arginine is added. It was suggested that AaaA was not expressed in ASM, as activation of the *aaaA* promoter has been shown to be reduced when grown in ASM compared with LB (Spencer, 2018). However, AaaA has also been shown to be highly active in ASM using L-arginine-p-nitroanilide cleavage assays (Spencer, 2018). It is also unlikely that arginine liberation by AaaA operates at 100% efficiency, meaning that the equivalent concentration of arginine liberated from 5 mM RGD would not equate to 5 mM pure L-arginine, meaning any phenotype conferred would be less pronounced. This remains a challenging question that may be better answered using the arginine quantification assay, as described in 6.4.3, to directly measure arginine liberation following RGD addition, rather than relying on biofilm biomass as a proxy for the presence of arginine.

6.4.3 Alternative approaches to examining AaaA function

Though extra-cellular arginine could be quantified in this study, it was not possible to detect AaaA-mediated differences in arginine concentration. It is possible that the effect of AaaA is highly localised, which is likely in the context of the microcolony aggregates seen in a chronic wound biofilm. Therefore, a more fine-grain approach using a technique with high spatial resolution may be needed to detect these differences.

During this study, a preliminary attempt at quantifying arginine in the SCW using imaging mass spectrometry (OrbiSIMS) was made in collaboration with Dr Steffi Kern at the University of Nottingham nmRC. However, the data are not presented here as it was difficult to pick out a peak for arginine even in the non-infected controls, and the quality of data acquired was low (not shown). Time constrains prevented the further optimisation necessary to generate meaningful data, especially given the complexity of the SCW media. Nevertheless, this technique has a great deal of potential to examine the effects of AaaA at a highly localised level, and work to optimise this would generate very valuable data.

Previously generated preliminary data using Raman spectroscopy also demonstrated an ability to distinguish PAO1N and PAO1N Δ*aaaA*, with each having significantly distinct chemical profiles when grown in LB (Evita Ning and Kim Hardie, University of Nottingham, unpublished). However, the resolution was not high enough to identify peaks related specifically to arginine metabolism, and a larger data set with longer exposure times would be needed for this. Nevertheless, Raman spectroscopy shows promise as an additional tool for examining the arginine liberation hypothesis for AaaA, potentially in more simple media than the SCW.

To bridge the gap between high complexity and low complexity approaches, liquid chromatography mass spectrometry (LC-MS) could be used to measure the relative quantities of liberated L-arginine by AaaA in minimal media with the RGD tripeptide. In brief, cells could be grown in MMP containing the RGD tripeptide, as in Kiney-Whitmore (2017), then the arginine and RGD content of both the cell lysate and supernatant could be quantified by LC-MS, to compare the quantity of liberated arginine levels between the wild-type, mutant and a pJM220*::aaaA* overexpression strain. This could determine if AaaA causes significantly more arginine to be liberated.

However, it is crucial to determine the extent and speed of arginine uptake in relation to AaaA, as this could also explain the lack of an observed difference in extracellular arginine due to AaaA, if any liberated arginine was immediately taken up into the cell. This would make sense, given that during an infection, *P. aeruginosa* is competing for arginine, it would be disadvantageous to allow accumulation of liberated extra-cellular arginine, that could be scavenged by other microbes, or weaponised by host immune cells into nitric oxide (Das *et al.*, 2010; Grasemann *et al.*, 2014).

To explore this further, it would be necessary to track the uptake and fate of any liberated arginine. This could be tested using LC-MS as described above, although with RGD tripeptide that contains a radio or heavy-labelled arginine. Following exogenous addition of the heavy-RGD, the cells could be fractionated and the proteomics of each fraction of wild type *vs* Δ*aaaA* could be compared. By comparing the amounts of heavy atoms in the arginine still

incorporated in RGD, to heavy atoms which have been incorporated from the liberated arginine into other proteins in each cell fraction, the total level of arginine liberation from RGD, even after cell uptake, could be determined. This may also allow identification of the specific pathways into which AaaA-liberated arginine is preferentially fluxed, as these intermediates would be more greatly enriched by the heavy atoms, and thus more easily detected. Extensive progress has been made in the field of stable isotope labelling in cell culture (SILAC), and though further interrogation would be required to optimise the technique required, it may be possible to also apply these to bacterial culture for this purpose (Beller & Hummon, 2022; Borek *et al.*, 2015; Ong *et al.*, 2003).

Chapter 7. Final remarks

This study has demonstrated that AaaA is highly structurally conserved within the existing *P. aeruginosa* genome database, and expressed in a synthetic chronic wound model, where it confers a survival and growth advantage. Moreover, *aaaA* promoter activation and AaaA activity were found to be significantly higher in the SCW than in planktonic conditions, validating the use of the SCW as ideal for studying AaaA in a disease-relevant context. This study also developed a Spytagged-AaaA, which allowed AaaA to be localised on the surface of *P. aeruginosa*, and this technique has the potential to localise AaaA within SCW biofilms in the future.

While AaaA-mediated liberation of arginine from an exogenous L-arginine-p-nitroanilide substrate was observed in cells grown in the SCW, a quantifiable, AaaA-dependant difference in arginine, which had been liberated from the endogenous SCW components, was not detected- although there was evidence of arginine utilisation by *P. aeruginosa* in the SCW. As discussed in 6.4.3, it may be that it was not possible to observe differences in extracellular arginine because any AaaA-liberated arginine is immediately taken up into the cell.

If so, one might expect differential expression of arginine transporters *arcD* and *aotJ* in response. However as discussed in 5.8.2, genes related to arginine uptake and metabolism, including *arcD* and *argR* (which is in an operon with *aotJ*) were not differentially expressed in the SCW infected with PAO1L Δ*aaaA*, compared with wild-type. Moreover, as discussed in 5.7.7.1, addition of ~20-50 μM exogenous L-arginine was not sufficient to upregulate *argR*, suggesting that if AaaA did release small quantities of arginine, it would not be sufficient to trigger differential gene expression of arginine uptake and metabolism genes. It may be that modulation on a transcriptional level is not required, as if expression of *arcD* and *aotJ* is already sufficient, any additional arginine released by AaaA can modulate metabolic processes simply by altering the pool of available intracellular substrates. This could take place

below the thresholds which required a costly activation of gene expression, but above those required to give *P. aeruginosa* a metabolic advantage.

This is supported by the observation that genes involved in alternative ethanol and formaldehyde metabolism were upregulated in the *aaaA* mutant, but not when the *aaaA* mutant was supplemented with L-arginine, suggesting that AaaA allows for low-level metabolic compensation via arginine acquisition. Taken together, while there is no evidence of AaaA liberating arginine in the SCW in quantities sufficient to alter expression of ArgR, or arginine metabolism genes, any arginine liberated may still contribute to the pool of available substrates to advantage cell survival and the maintenance of chronic infection phenotypes. High-resolution metabolomics, as discussed in 6.4.3, are needed to truly understand the fate, if any, of arginine liberated by AaaA.

An alternative possibility that emerges from this work, that must also be addressed, is that the primary function of AaaA is not for nutritional arginine acquisition, given the lack of altered expression of arginine metabolism genes in the *aaaA* mutant. This seems unlikely, given that AaaA is required for growth on minimal media with the RGD tripeptide, most probably due to its role in releasing arginine for nutrition (Luckett *et al.*, 2012). However, it is possible that AaaA has more than one mode of action, which depend on environmental cues and nutrient availability, and that nutrient acquisition is only one of these.

Analysis of a number of human immune peptides identified a few with arginine residues close to the N-terminus that could potentially be targeted by AaaA, including Human α-Defensin-1, -2 and -4 and Human Neutrophil Peptide-1, -2 and -3 (Lorenza Guadagnini and Kim Hardie, University of Nottingham, Unpublished work). This offers the possibility that AaaA is involved in altering or even inactivating host immune factors, which would explain its previously observed benefit to a chronic mouse wound infection, but not its survival advantage in the SCW, where live immune factors were not present.

Another previously suggested potential role for AaaA in relation to host immunity, could be in altering the balance of arginine available to the host immune system as substrates for iNOS and Arginase, to ultimately create an

immune environment more favourable to chronic infection (Gogoi *et al.*, 2016; Luckett *et al.*, 2012). This is supported by the finding in this study that the *tes* operon, which encodes chronic-infection-specific toxins that suppress the host immune system, was downregulated in the *aaaA* mutant, suggesting they are functionally related. It was not possible to explore host immune factors in this study, but addition of live immune components, such as fibroblasts, to the SCW would allow this question to be probed in the future.

A further possible role for AaaA is in cleaving arginine from *P. aeruginosa* proteins, as part of their maturation or extracellular activation. If so, a deeper bioinformatic analysis of *P. aeruginosa* proteins with N-terminal arginine and structures which fit into the AaaA pocket, and are co-expressed alongside AaaA, is needed to identify these.

Interestingly, this study has uncovered, for the first time, data linking AaaA with the stress response σ factor RpoS, which was found to be upregulated in the *aaaA* mutant. RNA-Seq further uncovered a number of RpoS/QS-controlled genes that were downregulated in the *aaaA* mutant, indicating linked functions. As discussed in 5.8.3, it is possible that the lack of *aaaA* shifts the bacterial phenotype to upregulate compensatory factors, and that this is driven by RpoS. In this model, RpoS would be more associated with the acute, motile, 'foraging' phenotype, and AaaA more associated with a chronic, sessile, phenotype, with each working on opposing axis.

This leads to a novel interpretation, that AaaA is functioning in both a nutritional and signalling capacity by releasing small amounts of arginine, which act as both a nutrient, and a diffuse chemotactic signal to nearby cells, signalling that a rich source of nutrients is nearby. In other words, cells which have located arginine-containing peptides would express AaaA to liberate the arginine, which, then in higher extra-cellular concentrations, would induce sessility, via c-di-GMP, and serve as nutrient to these sessile cells. Some liberated arginine would also diffuse away, laying a trail of 'breadcrumbs' for nearby motile, foraging, RpoS-expressing, cells to follow along a concentration gradient, toward the arginine-rich nutrient source. As these cells approached the arginine source they could then switch to the sessile, AaaA-expressing

phenotype (possibly mediated by HHQ) and further assist with arginine liberation. To investigate this theory, future work should focus on better understanding this relationship between RpoS and AaaA, as well as further exploring the role of arginine in chemotaxis, especially in the context of RpoS, QS and acute *vs* chronic infection stages.

In conclusion, this study has shown that AaaA confers an advantage in the synthetic chronic wound model, replicating previous findings from a murine chronic wound model, though the mechanism behind this is yet to be entirely resolved. The potential mechanism of action, as presented in this study, is that AaaA liberates arginine, either or both as a nutritional source and signalling molecule, which could shift the cell towards a chronic infection phenotype. This switch likely involves QS, in particular HHQ, and works on an opposing axis to RpoS. However, as this has mostly been explored at the transcriptional level, more work is needed to verify and elaborate on these findings at the post-transcriptional level, to gain greater clarity on the role of AaaA in maintaining chronic *P. aeruginosa* wound infections.

References

Abcam. **(2019a).** *ab241028 L-Arginine Assay Kit*. Abcam.

Abcam. **(2019b).** *ab252892 Arginine Assay Kit (Fluorometric)*. Abcam.

Abril, A. G., Rama, J. L. R., Sánchez-Pérez, A. & Villa, T. G. (2020). Prokaryotic sigma factors and their transcriptional counterparts in Archaea and Eukarya. *Appl Microbiol Biotechnol* **104**, 4289–4302.

Agren, M. S., Eaglstein, W. H., Ferguson, M. W., Harding, K. G., Moore, K., Saarialho-Kere, U. K. & Schultz, G. S. (2000). Causes and effects of the chronic inflammation in venous leg ulcers. *Acta Derm Venereol Suppl (Stockh)* **210**, 3–17.

Albus, A. M., Pesci, E. C., Runyen-Janecky, L. J., West, S. E. & Iglewski, B. H. (1997). Vfr controls quorum sensing in *Pseudomonas aeruginosa. J Bacteriol* **179**, 3928–3935.

Alekshun, M. N. & Levy, S. B. (1999). The mar regulon: multiple resistance to antibiotics and other toxic chemicals. *Trends Microbiol* **7**, 410–413.

Alekshun, M. N., Levy, S. B., Mealy, T. R., Seaton, B. A. & Head, J. F. (2001). The crystal structure of MarR, a regulator of multiple antibiotic resistance, at 2.3 A resolution. *Nat Struct Biol* **8**, 710–714.

Alford, M. A., Baghela, A., Yeung, A. T. Y., Pletzer, D. & Hancock, R. E. W. (2020). NtrBC Regulates Invasiveness and Virulence of Pseudomonas aeruginosa During High-Density Infection. *Front Microbiol* **11**, 773.

Alhagh, M. S. (2020, September). *Characterising the role of potential accessory factors that facilitate secretion and gene regulation of the EspC and AaaA autotransporters*. Doctoral dissertation, University of Nottingham.

Almblad, H., Harrison, J. J., Rybtke, M., Groizeleau, J., Givskov, M., Parsek, M. R. & Tolker-Nielsen, T. (2015). The Cyclic AMP-Vfr Signaling Pathway in Pseudomonas aeruginosa Is Inhibited by Cyclic Di-GMP. *J Bacteriol* **197**, 2190–2200.

Almblad, H., Rybtke, M., Hendiani, S., Andersen, J. B., Givskov, M. & Tolker-Nielsen, T. (2019). High levels of cAMP inhibit Pseudomonas aeruginosa biofilm formation through reduction of the c-di-GMP content. *Microbiology (Reading, Engl)* **165**, 324–333.

AlphaFold Protein Structure Database. **(2022, January 27).** Autotransporter domain-containing protein (PA0328): AlphaFold structure prediction.

Alting-Mees, M. A. & Short, J. M. (1989). *pBluescript II:* gene mapping vectors. *Nucleic Acids Res* **17**, 9494.

Arai, H., Kodama, T. & Igarashi, Y. (1999). Effect of nitrogen oxides on expression of the nir and nor genes for denitrification in Pseudomonas aeruginosa. *FEMS Microbiol Lett* **170**, 19–24.

Arai, H., Mizutani, M. & Igarashi, Y. (2003). Transcriptional regulation of the nos genes for nitrous oxide reductase in Pseudomonas aeruginosa. *Microbiology (Reading, Engl)* **149**, 29–36.

Armenteros, J. J. A., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., von Heijne, G. & Nielsen, H. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* **37**, 420–423.

Ashawesh, M. M. (2016, July). *Investigation of the spiral secretion pattern of the serine protease autotransporter, EspC using innovative fluorescent labelling approaches*. Doctoral dissertation, University of Nottingham.

Badal, D., Jayarani, A. V., Kollaran, M. A., Prakash, D., P, M. & Singh, V. (2021). Foraging Signals Promote Swarming in Starving Pseudomonas aeruginosa. *MBio* **12**, e0203321.

Bao, Y., Lies, D. P., Fu, H. & Roberts, G. P. (1991). An improved Tn7-based system for the single-copy insertion of cloned genes into chromosomes of Gram-negative bacteria. *Gene* **109**, 167–168.

Barrientos-Moreno, L., Molina-Henares, M. A., Ramos-González, M. I. & Espinosa-Urgel, M. (2020). Arginine as an environmental and metabolic cue for cyclic diguanylate signalling and biofilm formation in

Pseudomonas putida. *Sci Rep* **10**, 13623.

Beassoni, P., Massimelli, M., Ro, O. & Doménech, C. (2007). A Glance on Pseudomonas aeruginosa Phosphorylcholine Phosphatase, an Enzyme whose Synthesis Depends on the Presence of Choline in its Environment. In *Communicating current research and educational topics and trends in applied microbiology*, pp. 255–262. Edited by A. M. Vilas. Badajoz, Spain: FORMATEX.

Beatson, S. A., Whitchurch, C. B., Sargent, J. L., Levesque, R. C. & Mattick, J. S. (2002). Differential regulation of twitching motility and elastase production by Vfr in Pseudomonas aeruginosa. *J Bacteriol* **184**, 3605–3613.

Becher, A. & Schweizer, H. P. (2000). Integration-proficient *Pseudomonas aeruginosa* vectors for isolation of single-copy chromosomal lacZ and lux gene fusions. *BioTechniques* **29**, 948–50, 952.

Beller, N. C. & Hummon, A. B. (2022). Advances in stable isotope labeling: dynamic labeling for spatial and temporal proteomic analysis. *Mol Omics*.

Benkert, B., Quäck, N., Schreiber, K., Jaensch, L., Jahn, D. & Schobert, M. (2008). Nitrate-responsive NarX-NarL represses arginine-mediated induction of the Pseudomonas aeruginosa arginine fermentation arcDABC operon. *Microbiology (Reading, Engl)* **154**, 3053–3060.

van den Berg, B. (2010). Crystal structure of a full-length autotransporter. *J Mol Biol* **396**, 627–633.

Bernier, S. P., Ha, D.-G., Khan, W., Merritt, J. H. & O'Toole, G. A. (2011). Modulation of *Pseudomonas aeruginosa* surface-associated group behaviors by individual amino acids through c-di-GMP signaling. *Res Microbiol* **162**, 680–688.

Bianconi, I., Alcalá-Franco, B., Scarselli, M., Dalsass, M., Buccato, S., Colaprico, A., Marchi, S., Masignani, V. & Bragonzi, A. (2018). Genome-Based Approach Delivers Vaccine Candidates Against *Pseudomonas aeruginosa*. *Front Immunol* **9**, 3021.

Bleves, S., Viarre, V., Salacha, R., Michel, G. P. F., Filloux, A. & Voulhoux, R. (2010). Protein secretion systems in Pseudomonas aeruginosa: A wealth of pathogenic weapons. *Int J Med Microbiol* **300**, 534–543.

Blum, M., Chang, H.-Y., Chuguransky, S., Grego, T., Kandasaamy, S., Mitchell, A., Nuka, G., Paysan-Lafosse, T., Qureshi, M. & other authors. (2021). The InterPro protein families and domains database: 20 years on. *Nucleic Acids Res* **49**, D344–D354.

Boas, S., Donaldson, S., McBennett, K., Liou, T., Howrylak, J., Johnson, L., Teneback, C., Dozor, A., Sawicki, G. & other authors. (2021). A phase 1b, randomized, double-blind, placebo-controlled, dose-escalation trial of CB-280, an arginase inhibitor, in patients with cystic fibrosis. *J Cyst Fibros* **20**, S250.

Böckelmann, U., Janke, A., Kuhn, R., Neu, T. R., Wecke, J., Lawrence, J. R. & Szewzyk, U. (2006). Bacterial extracellular DNA forming a defined network-like structure. *FEMS Microbiol Lett* **262**, 31–38.

Boratyn, G. M., Camacho, C., Cooper, P. S., Coulouris, G., Fong, A., Ma, N., Madden, T. L., Matten, W. T., McGinnis, S. D. & other authors. (2013). BLAST: a more efficient report with usability improvements. *Nucleic Acids Res* **41**, W29-33.

Borek, W. E., Zou, J., Rappsilber, J. & Sawin, K. E. (2015). Deletion of Genes Encoding Arginase Improves Use of “Heavy” Isotope-Labeled Arginine for Mass Spectrometry in Fission Yeast. *PLoS One* **10**, e0129548.

Borlee, B. R., Goldman, A. D., Murakami, K., Samudrala, R., Wozniak, D. J. & Parsek, M. R. (2010). Pseudomonas aeruginosa uses a cyclic-di-GMP-regulated adhesin to reinforce the biofilm extracellular matrix. *Mol Microbiol* **75**, 827–842.

Borrero-de Acuña, J. M., Rohde, M., Wissing, J., Jänsch, L., Schobert, M., Molinari, G., Timmis, K. N., Jahn, M. & Jahn, D. (2016). Protein Network of the Pseudomonas aeruginosa Denitrification Apparatus. *J Bacteriol* **198**, 1401–1413.

Bourdineaud, J. P., Heierli, D., Gamper, M., Verhoogt, H. J., Driessen, A. J., Konings, W. N., Lazdunski, C. & Haas, D. (1993). Characterization of the arcD arginine:ornithine exchanger of Pseudomonas aeruginosa. Localization in the cytoplasmic membrane and a topological model. *J Biol Chem* **268**, 5417–5424.

Brettin, T., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Olsen, G. J., Olson, R., Overbeek, R., Parrello, B. & other authors. **(2015).** RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep* **5**, 8365.

Broadbent, L., Depping, P., Lodé, A., Vaitsopoulou, A., Hardy, D., Ayub, H., Mitchell-White, J., Kerr, I. D., Goddard, A. D. & other authors. **(2022).** Detergent-Free Membrane Protein Purification Using SMA Polymer. *Methods Mol Biol* **2507**, 389–404.

Busch, A., Richter, A. S. & Backofen, R. (2008). IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. *Bioinformatics* **24**, 2849–2856.

Bush, M. & Dixon, R. (2012). The role of bacterial enhancer binding proteins as specialized activators of σ54-dependent transcription. *Microbiol Mol Biol Rev* **76**, 497–529.

Caffrey, M. (2015). A comprehensive review of the lipid cubic phase or in meso method for crystallizing membrane and soluble proteins and complexes. *Acta Crystallogr Sect F, Struct Biol Commun* **71**, 3–18.

Cahan, R., Axelrad, I., Safrin, M., Ohman, D. E. & Kessler, E. (2001). A secreted aminopeptidase of *Pseudomonas aeruginosa*. Identification, primary structure, and relationship to other aminopeptidases. *J Biol Chem* **276**, 43645–43652.

Caiazza, N. C., Merritt, J. H., Brothers, K. M. & O'Toole, G. A. (2007). Inverse regulation of biofilm formation and swarming motility by *Pseudomonas aeruginosa* PA14. *J Bacteriol* **189**, 3603–3612.

Callaway, E. (2020). “It will change everything”: DeepMind’s AI makes gigantic leap in solving protein structures. *Nature* **588**, 203–204.

Campoccia, D., Montanaro, L. & Arciola, C. R. (2021). Extracellular DNA (edna). A major ubiquitous element of the bacterial biofilm architecture. *Int J Mol Sci* **22**.

Carmona, M. & Magasanik, B. (1996). Activation of transcription at sigma 54-dependent promoters on linear templates requires intrinsic or induced bending of the DNA. *J Mol Biol* **261**, 348–356.

Castang, S., McManus, H. R., Turner, K. H. & Dove, S. L. (2008). H-NS family members function coordinately in an opportunistic pathogen. *Proc Natl Acad Sci USA* **105**, 18947–18952.

Cendra, M. D. M. & Torrents, E. (2021). Pseudomonas aeruginosa biofilms and their partners in crime. *Biotechnol Adv* **49**, 107734.

Chauhan, N., Hatlem, D., Orwick-Rydmark, M., Schneider, K., Floetenmeyer, M., van Rossum, B., Leo, J. C. & Linke, D. (2019). Insights into the autotransport process of a trimeric autotransporter, *Yersinia* Adhesin A (YadA). *Mol Microbiol* **111**, 844–862.

Chen, Y. (2014, March). *Fatty Acid Signalling in* Pseudomonas aeruginosa. Doctoral dissertation, University of Nottingham.

Chevalier, S., Bouffartigues, E., Bodilis, J., Maillot, O., Lesouhaitier, O., Feuilloley, M. G. J., Orange, N., Dufour, A. & Cornelis, P. (2017). Structure, function and regulation of *Pseudomonas aeruginosa* porins. *FEMS Microbiol Rev* **41**, 698–722.

Chevrier, B., D’Orchymont, H., Schalk, C., Tarnus, C. & Moras, D. (1996). The structure of the Aeromonas proteolytica aminopeptidase complexed with a hydroxamate inhibitor. Involvement in catalysis of Glu151 and two zinc ions of the co-catalytic unit. *Eur J Biochem* **237**, 393–398.

Choi, Y., Sims, G. E., Murphy, S., Miller, J. R. & Chan, A. P. (2012). Predicting the functional effect of amino acid substitutions and indels. *PLoS One*

7, e46688.

Chugani, S., Kim, B. S., Phattarasukol, S., Brittnacher, M. J., Choi, S. H., Harwood, C. S. & Greenberg, E. P. (2012). Strain-dependent diversity in the Pseudomonas aeruginosa quorum-sensing regulon. *Proc Natl Acad Sci USA* **109**, E2823-31.

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X. & Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)* **6**, 80–92.

Cornforth, D. M., Dees, J. L., Ibberson, C. B., Huse, H. K., Mathiesen, I. H., Kirketerp-Møller, K., Wolcott, R. D., Rumbaugh, K. P., Bjarnsholt, T. & Whiteley, M. (2018). *Pseudomonas aeruginosa* transcriptome during human infection. *Proc Natl Acad Sci USA* **115**, E5125–E5134.

Couto, N., Schooling, S. R., Dutcher, J. R. & Barber, J. (2015). Proteome Profiles of Outer Membrane Vesicles and Extracellular Matrix of Pseudomonas aeruginosa Biofilms. *J Proteome Res* **14**, 4207–4222.

Crocker, A. W., Harty, C. E., Hammond, J. H., Willger, S. D., Salazar, P., Botelho, N. J., Jacobs, N. J. & Hogan, D. A. (2019). Pseudomonas aeruginosa Ethanol Oxidation by AdhA in Low-Oxygen Environments. *J Bacteriol* **201**.

Cuypers, H., Viebrock-Sambale, A. & Zumft, W. G. (1992). NosR, a membrane-bound regulatory component necessary for expression of nitrous oxide reductase in denitrifying Pseudomonas stutzeri. *J Bacteriol* **174**, 5332–5339.

Dalbey, R. E. & Kuhn, A. (2000). Evolutionarily related insertion pathways of bacterial, mitochondrial, and thylakoid membrane proteins. *Annu Rev Cell Dev Biol* **16**, 51–87.

Danese, P. N., Pratt, L. A., Dove, S. L. & Kolter, R. (2000). The outer membrane protein, antigen 43, mediates cell-to-cell interactions within

Escherichia coli biofilms. *Mol Microbiol* **37**, 424–432.

Darch, O. M. (2009). *A study on the function of the* Pseudomonas aeruainosa *autotransporter PA0328*. Doctoral dissertation, University of Nottingham.

Darzins, A. (1995). The Pseudomonas aeruginosa pilK gene encodes a chemotactic methyltransferase (CheR) homologue that is translationally regulated. *Mol Microbiol* **15**, 703–717.

Das, P., Lahiri, A., Lahiri, A. & Chakravortty, D. (2010). Modulation of the arginase pathway in the context of microbial pathogenesis: a metabolic enzyme moonlighting as an immune modulator. *PLoS Pathog* **6**, e1000899.

Dasgupta, N., Ferrell, E. P., Kanack, K. J., West, S. E. H. & Ramphal, R. (2002). fleQ, the gene encoding the major flagellar regulator of Pseudomonas aeruginosa, is sigma70 dependent and is downregulated by Vfr, a homolog of Escherichia coli cyclic AMP receptor protein. *J Bacteriol* **184**, 5240–5250.

Dasgupta, N., Wolfgang, M. C., Goodman, A. L., Arora, S. K., Jyot, J., Lory, S. & Ramphal, R. (2003). A four-tiered transcriptional regulatory circuit controls flagellar biogenesis in Pseudomonas aeruginosa. *Mol Microbiol* **50**, 809–824.

Davis, J. J., Wattam, A. R., Aziz, R. K., Brettin, T., Butler, R., Butler, R. M., Chlenski, P., Conrad, N., Dickerman, A. & other authors. **(2020).** The PATRIC Bioinformatics Resource Center: expanding data and analysis capabilities. *Nucleic Acids Res* **48**, D606–D612.

Deschamps-Francoeur, G., Simoneau, J. & Scott, M. S. (2020). Handling multi-mapped reads in RNA-seq. *Comput Struct Biotechnol J* **18**, 1569–1576.

Déziel, E., Lépine, F., Milot, S., He, J., Mindrinos, M. N., Tompkins, R. G. & Rahme, L. G. (2004). Analysis of Pseudomonas aeruginosa 4-hydroxy-2-alkylquinolines (HAQs) reveals a role for 4-hydroxy-2-heptylquinoline

in cell-to-cell communication. *Proc Natl Acad Sci USA* **101**, 1339–1344.

Dhodary, B., Sampedro, I., Behroozian, S., Borza, V., Her, S. & Hill, J. E. (2022). The Arginine Catabolism-Derived Amino Acid l-ornithine Is a Chemoattractant for Pseudomonas aeruginosa. *Microorganisms* **10**.

Diab, F., Bernard, T., Bazire, A., Haras, D., Blanco, C. & Jebbar, M. (2006). Succinate-mediated catabolite repression control on the production of glycine betaine catabolic enzymes in Pseudomonas aeruginosa PAO1 under low and elevated salinities. *Microbiology (Reading, Engl)* **152**, 1395–1406.

Diehl, A., von Wintzingerode, F. & Görisch, H. (1998). Quinoprotein ethanol dehydrogenase of Pseudomonas aeruginosa is a homodimer--sequence of the gene and deduced structural properties of the enzyme. *Eur J Biochem* **257**, 409–419.

Dietrich, L. E. P., Okegbe, C., Price-Whelan, A., Sakhtah, H., Hunter, R. C. & Newman, D. K. (2013). Bacterial community morphogenesis is intimately linked to the intracellular redox state. *J Bacteriol* **195**, 1371–1380.

Diggle, S. P., Stacey, R. E., Dodd, C., Cámara, M., Williams, P. & Winzer, K. (2006). The galactophilic lectin, LecA, contributes to biofilm development in Pseudomonas aeruginosa. *Environ Microbiol* **8**, 1095–1104.

Ding, W., Baumdicker, F. & Neher, R. A. (2018). panX: pan-genome analysis and exploration. *Nucleic Acids Res* **46**, e5.

Dötsch, A., Eckweiler, D., Schniederjans, M., Zimmermann, A., Jensen, V., Scharfe, M., Geffers, R. & Häussler, S. (2012). The Pseudomonas aeruginosa transcriptome in planktonic cultures and static biofilms using RNA sequencing. *PLoS One* **7**, e31092.

Dubern, J.-F., Romero, M., Mai-Prochnow, A., Messina, M., Trampari, E., Gijzel, H. N., Chan, K.-G., Carabelli, A. M., Barraud, N. & other authors. **(2022).** ToxR is a c-di-GMP binding protein that modulates

surface-associated behaviour in Pseudomonas aeruginosa. *npj Biofilms and Microbiomes* **8**, 64.

Egan, S. M. & Schleif, R. F. (1993). A regulatory cascade in the induction of rhaBAD. *J Mol Biol* **234**, 87–98.

Eren, E., Vijayaraghavan, J., Liu, J., Cheneke, B. R., Touw, D. S., Lepore, B. W., Indic, M., Movileanu, L. & van den Berg, B. (2012). Substrate specificity within a family of outer membrane carboxylate channels. *PLoS Biol* **10**, e1001242.

Eurofins Genomics. **(2020).** Standard Primer list.

Everett, J., Turner, K., Cai, Q., Gordon, V., Whiteley, M. & Rumbaugh, K. (2017). Arginine Is a Critical Substrate for the Pathogenesis of Pseudomonas aeruginosa in Burn Wound Infections. *MBio* **8**.

Feinbaum, R. L., Urbach, J. M., Liberati, N. T., Djonovic, S., Adonizio, A., Carvunis, A.-R. & Ausubel, F. M. (2012). Genome-wide identification of Pseudomonas aeruginosa virulence-related genes using a Caenorhabditis elegans infection model. *PLoS Pathog* **8**, e1002813.

Fetar, H., Gilmour, C., Klinoski, R., Daigle, D. M., Dean, C. R. & Poole, K. (2011). mexEF-oprN multidrug efflux operon of Pseudomonas aeruginosa: regulation by the MexT activator in response to nitrosative stress and chloramphenicol. *Antimicrob Agents Chemother* **55**, 508–514.

Filloux, A. (2011). Protein Secretion Systems in Pseudomonas aeruginosa: An Essay on Diversity, Evolution, and Function. *Front Microbiol* **2**, 155.

Flemming, H.-C. & Wingender, J. (2010). The biofilm matrix. *Nat Rev Microbiol* **8**, 623–633.

Flemming, H.-C., Neu, T. R. & Wozniak, D. J. (2007). The EPS matrix: the "house of biofilm cells". *J Bacteriol* **189**, 7945–7947.

Flemming, H.-C., Wingender, J., Szewzyk, U., Steinberg, P., Rice, S. A. & Kjelleberg, S. (2016). Biofilms: an emergent form of bacterial life. *Nat Rev Microbiol* **14**, 563–575.

Folders, J., Algra, J., Roelofs, M. S., van Loon, L. C., Tommassen, J. & Bitter, W. (2001). Characterization of Pseudomonas aeruginosa chitinase, a gradually secreted protein. *J Bacteriol* **183**, 7044–7052.

Frank, S., Kämpfer, H., Wetzler, C. & Pfeilschifter, J. (2002). Nitric oxide drives skin repair: novel functions of an established mediator. *Kidney Int* **61**, 882–888.

Fuchs, E. L., Brutinel, E. D., Klem, E. R., Fehr, A. R., Yahr, T. L. & Wolfgang, M. C. (2010). In vitro and in vivo characterization of the Pseudomonas aeruginosa cyclic AMP (cAMP) phosphodiesterase CpdA, required for cAMP homeostasis and virulence factor regulation. *J Bacteriol* **192**, 2779–2790.

Fujita, M., Tanaka, K., Takahashi, H. & Amemura, A. (1994). Transcription of the principal sigma-factor genes, rpoD and rpoS, in Pseudomonas aeruginosa is controlled according to the growth phase. *Mol Microbiol* **13**, 1071–1077.

Fulcher, N. B., Holliday, P. M., Klem, E., Cann, M. J. & Wolfgang, M. C. (2010). The *Pseudomonas aeruginosa* Chp chemosensory system regulates intracellular cAMP levels by modulating adenylate cyclase activity. *Mol Microbiol* **76**, 889–904.

Galimand, M., Gamper, M., Zimmermann, A. & Haas, D. (1991). Positive FNR-like control of anaerobic arginine degradation and nitrate respiration in Pseudomonas aeruginosa. *J Bacteriol* **173**, 1598–1606.

Gallegos, M. T., Michán, C. & Ramos, J. L. (1993). The XylS/AraC family of regulators. *Nucleic Acids Res* **21**, 807–810.

Ghasemi, A. & Zahediasl, S. (2012). Normality tests for statistical analysis: a guide for non-statisticians. *Int J Endocrinol Metab* **10**, 486–489.

Gilboa, R., Spungin-Bialik, A., Wohlfahrt, G., Schomburg, D., Blumberg, S. & Shoham, G. (2001). Interactions of Streptomyces griseus aminopeptidase with amino acid reaction products and their implications toward a catalytic mechanism. *Proteins* **44**, 490–504.

Gilson, L., Mahanty, H. K. & Kolter, R. (1990). Genetic analysis of an MDR-like export system: the secretion of colicin V. *EMBO J* **9**, 3875–3884.

Gjødsbøl, K., Christensen, J. J., Karlsmark, T., Jørgensen, B., Klein, B. M. & Krogfelt, K. A. (2006). Multiple bacterial species reside in chronic wounds: a longitudinal study. *Int Wound J* **3**, 225–231.

Glasser, N. R., Kern, S. E. & Newman, D. K. (2014). Phenazine redox cycling enhances anaerobic survival in *Pseudomonas aeruginosa* by facilitating generation of ATP and a proton-motive force. *Mol Microbiol* **92**, 399–412.

Gogoi, M., Datey, A., Wilson, K. T. & Chakravortty, D. (2016). Dual role of arginine metabolism in establishing pathogenesis. *Curr Opin Microbiol* **29**, 43–48.

Görisch, H. & Rupp, M. (1989). Quinoprotein ethanol dehydrogenase from Pseudomonas. *Antonie Van Leeuwenhoek* **56**, 35–45.

Görisch, H. (2003). The ethanol oxidation system and its regulation in Pseudomonas aeruginosa. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1647**, 98–102.

Grasemann, H., Jaecklin, T., Mehl, A., Huang, H., Rafii, M., Pencharz, P. & Ratjen, F. (2014). Multitracer stable isotope quantification of arginase and nitric oxide synthase activity in a mouse model of pseudomonas lung infection. *Mediators Inflamm* **2014**, 323526.

Green, E. R. & Mecsas, J. (2016). Bacterial secretion systems: an overview. *Microbiol Spectr* **4**.

Grenier, D., Gauthier, P., Plamondon, P., Nakayama, K. & Mayrand, D. (2001). Studies on the aminopeptidase activities of Porphyromonas gingivalis. *Oral Microbiol Immunol* **16**, 212–217.

Gruber, T. M. & Gross, C. A. (2003). Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu Rev Microbiol* **57**, 441–466.

Ha, D.-G., Merritt, J. H., Hampton, T. H., Hodgkinson, J. T., Janecek, M.,

Spring, D. R., Welch, M. & O'Toole, G. A. (2011). 2-Heptyl-4-quinolone, a precursor of the Pseudomonas quinolone signal molecule, modulates swarming motility in Pseudomonas aeruginosa. *J Bacteriol* **193**, 6770–6780.

Haas, D., Matsumoto, H., Moretti, P., Stalon, V. & Mercenier, A. (1984). Arginine degradation in Pseudomonas aeruginosa mutants blocked in two arginine catabolic pathways. *Mol Gen Genet* **193**, 437–444.

Halbert, A. R., Stacey, M. C., Rohr, J. B. & Jopp-McKay, A. (1992). The effect of bacterial colonization on venous ulcer healing. *Australas J Dermatol* **33**, 75–80.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**, 557–580.

Hardie, K. R., Cooksley, C., Green, A. D. & Winzer, K. (2003). Autoinducer 2 activity in Escherichia coli culture supernatants can be actively reduced despite maintenance of an active synthase, LuxS. *Microbiology (Reading, Engl)* **149**, 715–728.

Harrison, F., McNally, A., da Silva, A. C., Heeb, S. & Diggle, S. P. (2017). Optimised chronic infection models demonstrate that siderophore "cheating" in Pseudomonas aeruginosa is context specific. *ISME J* **11**, 2492–2509.

Hashim, S., Kwon, D.-H., Abdelal, A. & Lu, C.-D. (2004). The arginine regulatory protein mediates repression by arginine of the operons encoding glutamate synthase and anabolic glutamate dehydrogenase in Pseudomonas aeruginosa. *J Bacteriol* **186**, 3848–3854.

Hatlem, D., Trunk, T., Linke, D. & Leo, J. C. (2019). Catching a SPY: Using the SpyCatcher-SpyTag and Related Systems for Labeling and Localizing Bacterial Proteins. *Int J Mol Sci* **20**.

Hauser, A. R., Kang, P. J. & Engel, J. N. (1998). PepA, a secreted protein of Pseudomonas aeruginosa, is necessary for cytotoxicity and virulence. *Mol Microbiol* **27**, 807–818.

Heeb, S., Fletcher, M. P., Chhabra, S. R., Diggle, S. P., Williams, P. & Cámara, M. (2011). Quinolones: from antibiotics to autoinducers. *FEMS Microbiol Rev* **35**, 247–274.

Helmann, J. D. & Chamberlin, M. J. (1988). Structure and function of bacterial sigma factors. *Annu Rev Biochem* **57**, 839–872.

Henderson, I. R., Navarro-Garcia, F., Desvaux, M., Fernandez, R. C. & Ala'Aldeen, D. (2004). Type V protein secretion pathway: the autotransporter story. *Microbiol Mol Biol Rev* **68**, 692–744.

Hendrickson, E. L., Plotnikova, J., Mahajan-Miklos, S., Rahme, L. G. & Ausubel, F. M. (2001). Differential roles of the Pseudomonas aeruginosa PA14 rpoN gene in pathogenicity in plants, nematodes, insects, and mice. *J Bacteriol* **183**, 7126–7134.

Hengge, R. (2009). Principles of c-di-GMP signalling in bacteria. *Nat Rev Microbiol* **7**, 263–273.

Henry, R. L., Mellis, C. M. & Petrovic, L. (1992). Mucoid *Pseudomonas aeruginosa* is a marker of poor survival in cystic fibrosis. *Pediatr Pulmonol* **12**, 158–161.

Hoang, T. T., Karkhoff-Schweizer, R. R., Kutchma, A. J. & Schweizer, H. P. (1998). A broad-host-range Flp-FRT recombination system for site-specific excision of chromosomally-located DNA sequences: application for isolation of unmarked *Pseudomonas aeruginosa* mutants. *Gene* **212**, 77–86.

Hofacker, I. L. & Stadler, P. F. (2006). Memory efficient folding algorithms for circular RNA secondary structures. *Bioinformatics* **22**, 1172–1176.

Hogardt, M. & Heesemann, J. (2010). Adaptation of Pseudomonas aeruginosa during persistence in the cystic fibrosis lung. *Int J Med Microbiol* **300**, 557–562.

Holloway, B. W., Krishnapillai, V. & Morgan, A. F. (1979). Chromosomal genetics of Pseudomonas. *Microbiol Rev* **43**, 73–102.

Horcajada, J. P., Montero, M., Oliver, A., Sorlí, L., Luque, S., Gómez-

Zorrilla, S., Benito, N. & Grau, S. (2019). Epidemiology and Treatment of Multidrug-Resistant and Extensively Drug-Resistant Pseudomonas aeruginosa Infections. *Clin Microbiol Rev* **32**.

Huang, H. & Hancock, R. E. (1993). Genetic definition of the substrate selectivity of outer membrane porin protein OprD of Pseudomonas aeruginosa. *J Bacteriol* **175**, 7793–7800.

Infantes, L., Otero, L. H., Beassoni, P. R., Boetsch, C., Lisa, A. T., Domenech, C. E. & Albert, A. (2012). The structural domains of Pseudomonas aeruginosa phosphorylcholine phosphatase cooperate in substrate hydrolysis: 3D structure and enzymatic mechanism. *J Mol Biol* **423**, 503–514.

Insightful Science. **(2022).** *SnapGene | Software for everyday molecular biology*. Computer software, GSL Biotech LLC.

Itoh, Y. (1997). Cloning and characterization of the aru genes encoding enzymes of the catabolic arginine succinyltransferase pathway in Pseudomonas aeruginosa. *J Bacteriol* **179**, 7280–7290.

Jain, S., van Ulsen, P., Benz, I., Schmidt, M. A., Fernandez, R., Tommassen, J. & Goldberg, M. B. (2006). Polar localization of the autotransporter family of large bacterial virulence proteins. *J Bacteriol* **188**, 4841–4850.

Jensen, P. Ø., Bjarnsholt, T., Phipps, R., Rasmussen, T. B., Calum, H., Christoffersen, L., Moser, C., Williams, P., Pressler, T. & other authors. **(2007).** Rapid necrotic killing of polymorphonuclear leukocytes is caused by quorum-sensing-controlled production of rhamnolipid by Pseudomonas aeruginosa. *Microbiology (Reading, Engl)* **153**, 1329–1338.

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A. & other authors. **(2014).** InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240.

Jørgensen, F., Bally, M., Chapon-Herve, V., Michel, G., Lazdunski, A., Williams, P. & Stewart, G. S. A. B. (1999). RpoS-dependent stress tolerance in Pseudomonas aeruginosa. *Microbiology (Reading, Engl)* **145 (Pt 4)**, 835–844.

Jose, J., Jähnig, F. & Meyer, T. F. (1995). Common structural features of IgA1 protease-like outer membrane protein autotransporters. *Mol Microbiol* **18**, 378–380.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A. & other authors. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589.

Junker, M., Schuster, C. C., McDonnell, A. V., Sorg, K. A., Finn, M. C., Berger, B. & Clark, P. L. (2006). Pertactin beta-helix folding mechanism suggests common themes for the secretion and folding of autotransporter proteins. *Proc Natl Acad Sci USA* **103**, 4918–4923.

Kadowaki, T., Nakayama, K., Okamoto, K., Abe, N., Baba, A., Shi, Y., Ratnayake, D. B. & Yamamoto, K. (2000). Porphyromonas gingivalis proteinases as virulence determinants in progression of periodontal diseases. *J Biochem* **128**, 153–159.

Katharios-Lanwermeyer, S., Whitfield, G. B., Howell, P. L. & O'Toole, G. A. (2021). Pseudomonas aeruginosa Uses c-di-GMP Phosphodiesterases RmcA and MorA To Regulate Biofilm Maintenance. *MBio* **12**.

Keeble, A. H., Banerjee, A., Ferla, M. P., Reddington, S. C., Anuar, I. N. A. K. & Howarth, M. (2017). Evolving accelerated amidation by spytag/spycatcher to analyze membrane dynamics. *Angew Chem Int Ed* **56**, 16521–16525.

Keeble, A. H., Turkki, P., Stokes, S., Khairil Anuar, I. N. A., Rahikainen, R., Hytönen, V. P. & Howarth, M. (2019). Approaching infinite affinity through engineering of peptide-protein interaction. *Proc Natl Acad Sci USA*.

Kida, Y., Taira, J., Yamamoto, T., Higashimoto, Y. & Kuwano, K. (2013). EprS, an autotransporter protein of Pseudomonas aeruginosa, possessing serine protease activity induces inflammatory responses through protease-activated receptors. *Cell Microbiol* **15**, 1168–1181.

Kida, Y., Taira, J. & Kuwano, K. (2015). EprS, an autotransporter serine protease, plays an important role in various pathogenic phenotypes of Pseudomonas aeruginosa. *Microbiology (Reading, Engl)* **162**, 318–329.

Kiney-Whitmore, J. (2017, July). *Exploring Regulation and Function of the Arginine-Specific Aminopeptidase Autotransporter of* Pseudomonas aeruginosa *AaaA*. Master thesis, University of Nottingham.

Kingsford, C. L., Ayanbule, K. & Salzberg, S. L. (2007). Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. *Genome Biol* **8**, R22.

Kirketerp-Møller, K., Jensen, P. Ø., Fazli, M., Madsen, K. G., Pedersen, J., Moser, C., Tolker-Nielsen, T., Høiby, N., Givskov, M. & Bjarnsholt, T. (2008). Distribution, organization, and ecology of bacteria in chronic wounds. *J Clin Microbiol* **46**, 2717–2722.

Kjaergaard, K., Schembri, M. A., Hasman, H. & Klemm, P. (2000). Antigen 43 from Escherichia coli induces inter- and intraspecies cell aggregation and changes in colony morphology of Pseudomonas fluorescens. *J Bacteriol* **182**, 4789–4796.

Klein, K., Sonnabend, M. S., Frank, L., Leibiger, K., Franz-Wachtel, M., Macek, B., Trunk, T., Leo, J. C., Autenrieth, I. B. & other authors. (2019). Deprivation of the Periplasmic Chaperone SurA Reduces Virulence and Restores Antibiotic Susceptibility of Multidrug-Resistant Pseudomonas aeruginosa. *Front Microbiol* **10**, 100.

Klockgether, J., Munder, A., Neugebauer, J., Davenport, C. F., Stanke, F., Larbig, K. D., Heeb, S., Schöck, U., Pohl, T. M. & other authors. (2010). Genome diversity of Pseudomonas aeruginosa PAO1 laboratory strains. *J Bacteriol* **192**, 1113–1121.

Kloosterman, T. G. & Kuipers, O. P. (2011). Regulation of arginine acquisition and virulence gene expression in the human pathogen Streptococcus pneumoniae by transcription regulators ArgR1 and AhrC. *J Biol Chem* **286**, 44594–44605.

Knowles, T. J., Scott-Tucker, A., Overduin, M. & Henderson, I. R. (2009). Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. *Nat Rev Microbiol* **7**, 206–214.

Kollaran, A. M., Joge, S., Kotian, H. S., Badal, D., Prakash, D., Mishra, A., Varma, M. & Singh, V. (2019). Context-Specific Requirement of Forty-Four Two-Component Loci in *Pseudomonas aeruginosa* Swarming. *iScience* **13**, 305–317.

Kufareva, I. & Abagyan, R. (2012). Methods of protein structure comparison. *Methods Mol Biol* **857**, 231–257.

Lambertsen, L., Sternberg, C. & Molin, S. (2004). Mini-Tn7 transposons for site-specific tagging of bacteria with fluorescent proteins. *Environ Microbiol* **6**, 726–732.

Landau, E. M. & Rosenbusch, J. P. (1996). Lipidic cubic phases: a novel concept for the crystallization of membrane proteins. *Proc Natl Acad Sci USA* **93**, 14532–14535.

Langmead, B. & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357–359.

Latifi, A., Foglino, M., Tanaka, K., Williams, P. & Lazdunski, A. (1996). A hierarchical quorum-sensing cascade in Pseudomonas aeruginosa links the transcriptional activators LasR and RhlR (VsmR) to expression of the stationary-phase sigma factor RpoS. *Mol Microbiol* **21**, 1137–1146.

Lebendiker, M. & Danieli, T. (2011). Purification of proteins fused to maltose-binding protein. *Methods Mol Biol* **681**, 281–293.

Lee, J. & Zhang, L. (2015). The hierarchy quorum sensing network in *Pseudomonas aeruginosa*. *Protein Cell* **6**, 26–41.

Lee, J., Pothula, K. R., Kleinekathöfer, U. & Im, W. (2018). Simulation Study

of Occk5 Functional Properties in Pseudomonas aeruginosa Outer Membranes. *J Phys Chem B* **122**, 8185–8192.

Lenis, Y. Y., Elmetwally, M. A., Maldonado-Estrada, J. G. & Bazer, F. W. (2017). Physiological importance of polyamines. *Zygote* **25**, 244–255.

Leo, J. C., Grin, I. & Linke, D. (2012). Type V secretion: mechanism(s) of autotransport through the bacterial outer membrane. *Philos Trans R Soc Lond B, Biol Sci* **367**, 1088–1101.

Lewis, K. A., Baker, A. E., Chen, A. I., Harty, C. E., Kuchma, S. L., O'Toole, G. A. & Hogan, D. A. (2019). Ethanol Decreases Pseudomonas aeruginosa Flagellar Motility through the Regulation of Flagellar Stators. *J Bacteriol* **201**.

Leyton, D. L., Rossiter, A. E. & Henderson, I. R. (2012). From self sufficiency to dependence: mechanisms and factors important for autotransporter biogenesis. *Nat Rev Microbiol* **10**, 213–225.

Li, C. & Lu, C.-D. (2009). Arginine racemization by coupled catabolic and anabolic dehydrogenases. *Proc Natl Acad Sci USA* **106**, 906–911.

Li, C., Wally, H., Miller, S. J. & Lu, C.-D. (2009a). The multifaceted proteins MvaT and MvaU, members of the H-NS family, control arginine metabolism, pyocyanin synthesis, and prophage activation in Pseudomonas aeruginosa PAO1. *J Bacteriol* **191**, 6211–6218.

Li, H. (2014). Toward better understanding of artifacts in variant calling from high-coverage samples. *Bioinformatics* **30**, 2843–2851.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. & 1000 Genome Project Data Processing Subgroup. **(2009b).** The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079.

Li, W. & Lu, C.-D. (2007). Regulation of carbon and nitrogen utilization by CbrAB and NtrBC two-component systems in Pseudomonas aeruginosa. *J Bacteriol* **189**, 5413–5420.

Li, Y., Kawakami, N., Ogola, H. J. O., Ashida, H., Ishikawa, T., Shibata, H.

& Sawa, Y. (2011). A novel L-aspartate dehydrogenase from the mesophilic bacterium Pseudomonas aeruginosa PAO1: molecular characterization and application for L-aspartate production. *Appl Microbiol Biotechnol* **90**, 1953–1962.

Lind, P. A., Arvidsson, L., Berg, O. G. & Andersson, D. I. (2017). Variation in Mutational Robustness between Different Proteins and the Predictability of Fitness Effects. *Mol Biol Evol* **34**, 408–418.

Liu, Q., Liu, Y., Kang, Z., Xiao, D., Gao, C., Xu, P. & Ma, C. (2018). 2,3-Butanediol catabolism in Pseudomonas aeruginosa PAO1. *Environ Microbiol* **20**, 3927–3940.

Liu, Z., Hossain, S. S., Morales Moreira, Z. & Haney, C. H. (2022). Putrescine and Its Metabolic Precursor Arginine Promote Biofilm and c-di-GMP Synthesis in Pseudomonas aeruginosa. *J Bacteriol* **204**, e0029721.

Love, M. I., Huber, W. & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550.

Lu, C. D. & Abdelal, A. T. (2001). The gdhB gene of Pseudomonas aeruginosa encodes an arginine-inducible NAD(+)-dependent glutamate dehydrogenase which is subject to allosteric regulation. *J Bacteriol* **183**, 490–499.

Lu, C.-D., Yang, Z. & Li, W. (2004). Transcriptome analysis of the ArgR regulon in Pseudomonas aeruginosa. *J Bacteriol* **186**, 3855–3861.

Luckett, J. C. A., Darch, O., Watters, C., Abuoun, M., Wright, V., Paredes-Osses, E., Ward, J., Goto, H., Heeb, S. & other authors. (2012). A novel virulence strategy for Pseudomonas aeruginosa mediated by an autotransporter with arginine-specific aminopeptidase activity. *PLoS Pathog* **8**, e1002854.

Lüneburg, N., Xanthakis, V., Schwedhelm, E., Sullivan, L. M., Maas, R., Anderssohn, M., Riederer, U., Glazer, N. L., Vasan, R. S. & Böger, R. H. (2011). Reference intervals for plasma L-arginine and the L-

arginine:asymmetric dimethylarginine ratio in the Framingham Offspring Cohort. *J Nutr* **141**, 2186–2190.

Lüthi, E., Baur, H., Gamper, M., Brunner, F., Villeval, D., Mercenier, A. & Haas, D. (1990). The arc operon for anaerobic arginine catabolism in Pseudomonas aeruginosa contains an additional gene, arcD, encoding a membrane protein. *Gene* **87**, 37–43.

Ma, L., Wang, S., Wang, D., Parsek, M. R. & Wozniak, D. J. (2012). The roles of biofilm matrix polysaccharide Psl in mucoid Pseudomonas aeruginosa biofilms. *FEMS Immunol Med Microbiol* **65**, 377–380.

Ma, S., Selvaraj, U., Ohman, D. E., Quarless, R., Hassett, D. J. & Wozniak, D. J. (1998). Phosphorylation-independent activity of the response regulators AlgB and AlgR in promoting alginate biosynthesis in mucoid Pseudomonas aeruginosa. *J Bacteriol* **180**, 956–968.

Madsen, S. M., Westh, H., Danielsen, L. & Rosdahl, V. T. (1996). Bacterial colonization and healing of venous leg ulcers. *APMIS* **104**, 895–899.

Malinverni, J. C. & Silhavy, T. J. (2009). An ABC transport system that maintains lipid asymmetry in the gram-negative outer membrane. *Proc Natl Acad Sci USA* **106**, 8009–8014.

Marth, G. T., Korf, I., Yandell, M. D., Yeh, R. T., Gu, Z., Zakeri, H., Stitziel, N. O., Hillier, L., Kwok, P. Y. & Gish, W. R. (1999). A general approach to single-nucleotide polymorphism discovery. *Nat Genet* **23**, 452–456.

Martin, D. W., Schurr, M. J., Mudd, M. H., Govan, J. R., Holloway, B. W. & Deretic, V. (1993). Mechanism of conversion to mucoidy in Pseudomonas aeruginosa infecting cystic fibrosis patients. *Proc Natl Acad Sci USA* **90**, 8377–8381.

Maseda, H., Saito, K., Nakajima, A. & Nakae, T. (2000). Variation of the mexT gene, a regulator of the MexEF-oprN efflux pump expression in wild-type strains of Pseudomonas aeruginosa. *FEMS Microbiol Lett* **192**, 107–112.

Massimelli, M. J., Beassoni, P. R., Forrellad, M. A., Barra, J. L., Garrido, M.

N., Domenech, C. E. & Lisa, A. T. (2005). Identification, cloning, and expression of Pseudomonas aeruginosa phosphorylcholine phosphatase gene. *Curr Microbiol* **50**, 251–256.

Matsuyama, B. Y., Krasteva, P. V., Baraquet, C., Harwood, C. S., Sondermann, H. & Navarro, M. V. A. S. (2016). Mechanistic insights into c-di-GMP-dependent control of the biofilm regulator FleQ from Pseudomonas aeruginosa. *Proc Natl Acad Sci USA* **113**, E209-18.

Mattick, J. S., Whitchurch, C. B. & Alm, R. A. (1996). The molecular genetics of type-4 fimbriae in Pseudomonas aeruginosa--a review. *Gene* **179**, 147–155.

Mavrodi, D. V., Ksenzenko, V. N., Bonsall, R. F., Cook, R. J., Boronin, A. M. & Thomashow, L. S. (1998). A seven-gene locus for synthesis of phenazine-1-carboxylic acid by Pseudomonas fluorescens 2-79. *J Bacteriol* **180**, 2541–2548.

McGowan, D. B. T. (2017). *Characterization of the Chp system of* Pseudomonas aeruginosa*.* . Doctoral dissertation, Monash University.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S. & other authors. **(2010).** The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20**, 1297–1303.

McKnight, S. L., Iglewski, B. H. & Pesci, E. C. (2000). The Pseudomonas quinolone signal regulates rhl quorum sensing in Pseudomonas aeruginosa. *J Bacteriol* **182**, 2702–2708.

Meisner, J. & Goldberg, J. B. (2016). The Escherichia coli rhaSR-PrhaBAD Inducible Promoter System Allows Tightly Controlled Gene Expression over a Wide Range in Pseudomonas aeruginosa. *Appl Environ Microbiol* **82**, 6715–6727.

Mercenier, A., Simon, J. P., Haas, D. & Stalon, V. (1980a). Catabolism of L-arginine by Pseudomonas aeruginosa. *J Gen Microbiol* **116**, 381–389.

Mercenier, A., Simon, J. P., Vander Wauven, C., Haas, D. & Stalon, V. (1980b). Regulation of enzyme synthesis in the arginine deiminase pathway of Pseudomonas aeruginosa. *J Bacteriol* **144**, 159–163.

Mérens, A., Matrat, S., Aubry, A., Lascols, C., Jarlier, V., Soussy, C.-J., Cavallo, J.-D. & Cambau, E. (2009). The pentapeptide repeat proteins MfpAMt and QnrB4 exhibit opposite effects on DNA gyrase catalytic reactions and on the ternary gyrase-DNA-quinolone complex. *J Bacteriol* **191**, 1587–1594.

Mern, D. S., Ha, S.-W., Khodaverdi, V., Gliese, N. & Görisch, H. (2010). A complex regulatory network controls aerobic ethanol oxidation in Pseudomonas aeruginosa: indication of four levels of sensor kinases and response regulators. *Microbiology (Reading, Engl)* **156**, 1505–1516.

Meuskens, I., Saragliadis, A., Leo, J. C. & Linke, D. (2019). Type V secretion systems: an overview of passenger domain functions. *Front Microbiol* **10**, 1163.

Miroux, B. & Walker, J. E. (1996). Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol* **260**, 289–298.

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., Tosatto, S. C. E., Paladin, L., Raj, S. & other authors. **(2021).** Pfam: The protein families database in 2021. *Nucleic Acids Res* **49**, D412–D419.

Mori, H. & Ito, K. (2001). The Sec protein-translocation pathway. *Trends Microbiol* **9**, 494–500.

Motulsky, H. (1995). GraphPad Prism 9 Statistics Guide. *GraphPad Prism 9 Statistics Guide*.

Mueller, E. J., Meyer, E., Rudolph, J., Davisson, V. J. & Stubbe, J. (1994). N5-carboxyaminoimidazole ribonucleotide: evidence for a new intermediate and two new enzymatic activities in the de novo purine biosynthetic pathway of Escherichia coli. *Biochemistry* **33**, 2269–2278.

Munder, M., Eichmann, K., Morán, J. M., Centeno, F., Soler, G. & Modolell, M. (1999). Th1/Th2-regulated expression of arginase isoforms in murine macrophages and dendritic cells. *J Immunol* **163**, 3771–3777.

Müsken, M., Di Fiore, S., Dötsch, A., Fischer, R. & Häussler, S. (2010). Genetic determinants of Pseudomonas aeruginosa biofilm establishment. *Microbiology (Reading, Engl)* **156**, 431–441.

Nahm, F. S. (2016). Nonparametric statistical tests for the continuous data: the basic concept and the practical use. *Korean Journal Anesthesiol* **69**, 8–14.

Nakada, Y., Jiang, Y., Nishijyo, T., Itoh, Y. & Lu, C. D. (2001). Molecular characterization and regulation of the *aguBA* operon, responsible for agmatine utilization in *Pseudomonas aeruginosa* PAO1. *J Bacteriol* **183**, 6517–6524.

Nakada, Y. & Itoh, Y. (2003). Identification of the putrescine biosynthetic genes in Pseudomonas aeruginosa and characterization of agmatine deiminase and N-carbamoylputrescine amidohydrolase of the arginine decarboxylase pathway. *Microbiology (Reading, Engl)* **149**, 707–714.

Natale, P., Brüser, T. & Driessen, A. J. M. (2008). Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane--distinct translocases and mechanisms. *Biochim Biophys Acta* **1778**, 1735–1756.

Nathwani, D., Raman, G., Sulham, K., Gavaghan, M. & Menon, V. (2014). Clinical and economic consequences of hospital-acquired resistant and multidrug-resistant Pseudomonas aeruginosa infections: a systematic review and meta-analysis. *Antimicrob Resist Infect Control* **3**, 32.

Naziat, H. (2020, August). *Investigating the role of arginine as therapeutics for Pseudomonas aeruginosa related infection*. Master thesis, University of Nottingham.

New England Biolabs. **(2022).** NiCo21(DE3) Competent E. coli. *New England Biolabs*.

Nguyen, M., Sharma, A., Wu, W., Gomi, R., Sung, B., Hospodsky, D., Angenent, L. T. & Worgall, S. (2016). The fermentation product 2,3-butanediol alters P. aeruginosa clearance, cytokine response and the lung microbiome. *ISME J* **10**, 2978–2983.

Nikaido, H. (2003). Molecular Basis of Bacterial Outer Membrane Permeability Revisited. *Microbiol Mol Biol Rev* **67**, 593–656.

Nishijyo, T., Park, S. M., Lu, C. D., Itoh, Y. & Abdelal, A. T. (1998). Molecular characterization and regulation of an operon encoding a system for transport of arginine and ornithine and the ArgR regulatory protein in Pseudomonas aeruginosa. *J Bacteriol* **180**, 5559–5566.

NIST. **(2021).** p-nitroaniline. *National Institute for Standards and Technology Chemistry WebBook, SRD 69*.

Nivaskumar, M. & Francetic, O. (2014). Type II secretion system: a magic beanstalk or a protein escalator. *Biochim Biophys Acta* **1843**, 1568–1577.

Norbury, W., Herndon, D. N., Tanksley, J., Jeschke, M. G. & Finnerty, C. C. (2016). Infection in Burns. *Surg Infect (Larchmt)* **17**, 250–255.

Novagen. **(2021).** pET-21a(+) DNA | 69740.

Novogene Co. Ltd. (2022). *Prokaryotes RNA-seq Report*. Novogene Co. Ltd.

Ochs, M. M., Lu, C. D., Hancock, R. E. & Abdelal, A. T. (1999). Amino acid-mediated induction of the basic amino acid-specific outer membrane porin OprD from Pseudomonas aeruginosa. *J Bacteriol* **181**, 5426–5432.

Okegbe, C., Fields, B. L., Cole, S. J., Beierschmitt, C., Morgan, C. J., Price-Whelan, A., Stewart, R. C., Lee, V. T. & Dietrich, L. E. P. (2017). Electron-shuttling antibiotics structure bacterial communities by modulating cellular levels of c-di-GMP. *Proc Natl Acad Sci USA* **114**, E5236–E5245.

Okuda, S., Sherman, D. J., Silhavy, T. J., Ruiz, N. & Kahne, D. (2016). Lipopolysaccharide transport and assembly at the outer membrane: the PEZ model. *Nat Rev Microbiol* **14**, 337–345.

Ong, S.-E., Kratchmarova, I. & Mann, M. (2003). Properties of 13C-substituted arginine in stable isotope labeling by amino acids in cell culture (SILAC). *J Proteome Res* **2**, 173–181.

Paiardini, A., Mantoni, F., Giardina, G., Paone, A., Janson, G., Leoni, L., Rampioni, G., Cutruzzolà, F. & Rinaldo, S. (2018). A novel bacterial l-arginine sensor controlling c-di-GMP levels in Pseudomonas aeruginosa. *Proteins* **86**, 1088–1096.

Paiva, A. C. (2019, September). *Surface Signalling and Quorum Sensing in Pseudomonas aeruginosa*. Undergraduate thesis, University of Nottingham.

Papangeli, M. (2022). *SadB is a Pseudomonas aeruginosa global regulator*. Doctoral dissertation, .

Paredes-Osses, E. (2014, August). *A novel virulence strategy for Pseudomonas aeruginosa mediated by an autotransporter with arginine- specific aminopeptidase activity*. Doctoral dissertation, University of Nottingham.

Park, A. J., Murphy, K., Surette, M. D., Bandoro, C., Krieger, J. R., Taylor, P. & Khursigara, C. M. (2015). Tracking the Dynamic Relationship between Cellular Systems and Extracellular Subproteomes in Pseudomonas aeruginosa Biofilms. *J Proteome Res* **14**, 4524–4537.

Park, S. M., Lu, C. D. & Abdelal, A. T. (1997). Cloning and characterization of argR, a gene that participates in regulation of arginine biosynthesis and catabolism in Pseudomonas aeruginosa PAO1. *J Bacteriol* **179**, 5300–5308.

Passos da Silva, D., Matwichuk, M. L., Townsend, D. O., Reichhardt, C., Lamba, D., Wozniak, D. J. & Parsek, M. R. (2019). The Pseudomonas aeruginosa lectin LecB binds to the exopolysaccharide Psl and stabilizes the biofilm matrix. *Nat Commun* **10**, 2183.

Paulson, N. B., Gilbertsen, A. J., Dalluge, J. J., Welchlin, C. W., Hughes, J., Han, W., Blackwell, T. S., Laguna, T. A. & Williams, B. J. (2014). The arginine decarboxylase pathways of host and pathogen interact to

impact inflammatory pathways in the lung. *PLoS One* **9**, e111441.

Pena, R. T., Blasco, L., Ambroa, A., González-Pedrajo, B., Fernández-García, L., López, M., Bleriot, I., Bou, G., García-Contreras, R. & other authors. **(2019).** Relationship between quorum sensing and secretion systems. *Front Microbiol* **10**, 1100.

Pesci, E. C., Milbank, J. B., Pearson, J. P., McKnight, S., Kende, A. S., Greenberg, E. P. & Iglewski, B. H. (1999). Quinolone signaling in the cell-to-cell communication system of Pseudomonas aeruginosa. *Proc Natl Acad Sci USA* **96**, 11229–11234.

Peterson, J. H., Tian, P., Ieva, R., Dautin, N. & Bernstein, H. D. (2010). Secretion of a bacterial virulence factor is driven by the folding of a C-terminal segment. *Proc Natl Acad Sci USA* **107**, 17739–17744.

Potvin, E., Sanschagrin, F. & Levesque, R. C. (2008). Sigma factors in Pseudomonas aeruginosa. *FEMS Microbiol Rev* **32**, 38–55.

Price-Whelan, A., Dietrich, L. E. P. & Newman, D. K. (2007). Pyocyanin alters redox homeostasis and carbon flux through central metabolic pathways in *Pseudomonas aeruginosa* PA14. *J Bacteriol* **189**, 6372–6381.

QIAGEN. **(2020).** *RNAprotect® Bacteria Reagent Handbook*, 6H8B-1704-002. QIAGEN.

Qing, G., Ma, L.-C., Khorchid, A., Swapna, G. V. T., Mal, T. K., Takayama, M. M., Xia, B., Phadtare, S., Ke, H. & other authors. **(2004).** Cold-shock induced high-yield protein production in Escherichia coli. *Nat Biotechnol* **22**, 877–882.

Ramos, J.-L. (Ed.). **(2004).** *Pseudomonas: Volume 1 Genomics, Life Style and Molecular Architecture*, illustrated. Springer Science & Business Media.

Reddington, S. C. & Howarth, M. (2015). Secrets of a covalent interaction for biomaterials and biotechnology: SpyTag and SpyCatcher. *Curr Opin Chem Biol* **29**, 94–99.

Reguant, R., Antipin, Y., Sheridan, R., Luna, A. & Sander, C. (2018). Alignmentviewer: sequence analysis of large protein families. *BioRxiv*.

Reichhardt, C., Wong, C., Passos da Silva, D., Wozniak, D. J. & Parsek, M. R. (2018). CdrA Interactions within the Pseudomonas aeruginosa Biofilm Matrix Safeguard It from Proteolysis and Promote Cellular Packing. *MBio* **9**.

Reichmann, P. & Görisch, H. (1993). Cytochrome c550 from Pseudomonas aeruginosa. *Biochem J* **289 (Pt 1)**, 173–178.

Rippe, K., Guthold, M., von Hippel, P. H. & Bustamante, C. (1997). Transcriptional activation via DNA-looping: visualization of intermediates in the activation pathway of E. coli RNA polymerase x sigma 54 holoenzyme by scanning force microscopy. *J Mol Biol* **270**, 125–138.

Robinson, C. & Bolhuis, A. (2001). Protein targeting by the twin-arginine translocation pathway. *Nat Rev Mol Cell Biol* **2**, 350–356.

Robinson, T., Smith, P., Alberts, E. R., Colussi-Pelaez, M. & Schuster, M. (2020). Cooperation and Cheating through a Secreted Aminopeptidase in the Pseudomonas aeruginosa RpoS Response. *MBio* **11**.

Roussel-Jazédé, V., Jongerius, I., Bos, M. P., Tommassen, J. & van Ulsen, P. (2010). NalP-mediated proteolytic release of lactoferrin-binding protein B from the meningococcal cell surface. *Infect Immun* **78**, 3083–3089.

Roy, A., Kucukural, A. & Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* **5**, 725–738.

Ruksakiet, K., Stercz, B., Tóth, G., Jaikumpun, P., Gróf, I., Tengölics, R., Lohinai, Z. M., Horváth, P., Deli, M. A. & other authors. (2021). Bicarbonate Evokes Reciprocal Changes in Intracellular Cyclic di-GMP and Cyclic AMP Levels in Pseudomonas aeruginosa. *Biology (Basel)* **10**.

Rupp, M. & Görisch, H. (1988). Purification, crystallisation and characterization of quinoprotein ethanol dehydrogenase from Pseudomonas aeruginosa. *Biol Chem Hoppe Seyler* **369**, 431–439.

Sambrook, J. & Russell, D. W. (2001). *Molecular Cloning: A Laboratory Manual, Volume 2*. CSHL Press.

Samsudin, F. & Khalid, S. (2019). Movement of Arginine through OprD: The Energetics of Permeation and the Role of Lipopolysaccharide in Directing Arginine to the Protein. *J Phys Chem B* **123**, 2824–2832.

Sana, T. G., Soscia, C., Tonglet, C. M., Garvis, S. & Bleves, S. (2013). Divergent control of two type VI secretion systems by RpoN in Pseudomonas aeruginosa. *PLoS One* **8**, e76030.

Sana, T. G., Laubier, A. & Bleves, S. (2014). Gene transfer: conjugation. *Methods Mol Biol* **1149**, 17–22.

Sandell, L. & Sharp, N. P. (2022). Fitness effects of mutations: an assessment of PROVEAN predictions using mutation accumulation data. *Genome Biol Evol* **14**.

Santajit, S. & Indrawattana, N. (2016). Mechanisms of antimicrobial resistance in ESKAPE pathogens. *Biomed Res Int* **2016**, 2475067.

Savli, H., Karadenizli, A., Kolayli, F., Gundes, S., Ozbek, U. & Vahaboglu, H. (2003). Expression stability of six housekeeping genes: A proposal for resistance gene quantification studies of Pseudomonas aeruginosa by real-time quantitative RT-PCR. *J Med Microbiol* **52**, 403–408.

Schlegel, S., Löfblom, J., Lee, C., Hjelm, A., Klepsch, M., Strous, M., Drew, D., Slotboom, D. J. & de Gier, J.-W. (2012). Optimizing membrane protein overexpression in the Escherichia coli strain Lemo21(DE3). *J Mol Biol* **423**, 648–659.

Schmidt, T. G. M. & Skerra, A. (2007). The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nat Protoc* **2**, 1528–1535.

Schobert, M. & Görisch, H. (1999). Cytochrome c550 is an essential component of the quinoprotein ethanol oxidation system in Pseudomonas aeruginosa: cloning and sequencing of the genes encoding cytochrome c550 and an adjacent acetaldehyde

dehydrogenase. *Microbiology (Reading, Engl)* **145 (Pt 2)**, 471–481.

Schobert, M. & Görisch, H. (2001). A soluble two-component regulatory system controls expression of quinoprotein ethanol dehydrogenase (QEDH) but not expression of cytochrome c(550) of the ethanol-oxidation system in Pseudomonas aeruginosa. *Microbiology (Reading, Engl)* **147**, 363–372.

Schulz, S., Eckweiler, D., Bielecka, A., Nicolai, T., Franke, R., Dötsch, A., Hornischer, K., Bruchmann, S., Düvel, J. & Häussler, S. (2015). Elucidation of sigma factor-associated networks in *Pseudomonas aeruginosa* reveals a modular architecture with limited and function-specific crosstalk. *PLoS Pathog* **11**, e1004744.

Schumacher, J., Zhang, X., Jones, S., Bordes, P. & Buck, M. (2004). ATP-dependent transcriptional activation by bacterial PspF AAA+protein. *J Mol Biol* **338**, 863–875.

Schuster, M. & Greenberg, E. P. (2006). A network of networks: quorum-sensing gene regulation in Pseudomonas aeruginosa. *Int J Med Microbiol* **296**, 73–81.

Schuster, M., Lostroh, C. P., Ogi, T. & Greenberg, E. P. (2003). Identification, timing, and signal specificity of Pseudomonas aeruginosa quorum-controlled genes: a transcriptome analysis. *J Bacteriol* **185**, 2066–2079.

Schuster, M., Hawkins, A. C., Harwood, C. S. & Greenberg, E. P. (2004). The Pseudomonas aeruginosa RpoS regulon and its relationship to quorum sensing. *Mol Microbiol* **51**, 973–985.

Scribani Rossi, C., Barrientos-Moreno, L., Paone, A., Cutruzzolà, F., Paiardini, A., Espinosa-Urgel, M. & Rinaldo, S. (2022). Nutrient Sensing and Biofilm Modulation: The Example of L-arginine in Pseudomonas. *Int J Mol Sci* **23**.

Serra, R., Grande, R., Butrico, L., Rossi, A., Settimio, U. F., Caroleo, B., Amato, B., Gallelli, L. & de Franciscis, S. (2015). Chronic wound infections: the role of Pseudomonas aeruginosa and Staphylococcus

aureus. *Expert Rev Anti Infect Ther* **13**, 605–613.

Sigma-Aldrich. **(2022).** Sigma-Aldrich catalogue.

da Silva, A. C. (2018, November). *Understanding the diversity and evolution of complex wound infections*. Doctoral dissertation, University of Nottingham.

Simon, R., Priefer, U. & Pühler, A. (1983). A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in gram negative bacteria. *Nat Biotechnol* **1**, 784–791.

Skillman, K. M., Barnard, T. J., Peterson, J. H., Ghirlando, R. & Bernstein, H. D. (2005). Efficient secretion of a folded protein domain by a monomeric bacterial autotransporter. *Mol Microbiol* **58**, 945–958.

Smith, R. S., Wolfgang, M. C. & Lory, S. (2004). An adenylate cyclase-controlled signaling network regulates Pseudomonas aeruginosa virulence in a mouse model of acute pneumonia. *Infect Immun* **72**, 1677–1684.

Spencer, D. (2018). *Surviving and thriving in the cystic fibrosis lung: contribution of an arginine specific aminopeptidase to sustained infection by* Pseudomonas aeruginosa. Doctoral dissertation, University of Nottingham.

Stalder, T. & Top, E. (2016). Plasmid transfer in biofilms: a perspective on limitations and opportunities. *npj Biofilms and Microbiomes* **2**.

Stock, A. M., Robinson, V. L. & Goudreau, P. N. (2000). Two-component signal transduction. *Annu Rev Biochem* **69**, 183–215.

Stover, C. K., Pham, X. Q., Erwin, A. L., Mizoguchi, S. D., Warrener, P., Hickey, M. J., Brinkman, F. S., Hufnagle, W. O., Kowalik, D. J. & other authors. **(2000).** Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. *Nature* **406**, 959–964.

Stroffekova, K., Proenza, C. & Beam, K. G. (2001). The protein-labeling reagent FLASH-EDT2 binds not only to CCXXCC motifs but also non-specifically to endogenous cysteine-rich proteins. *Pflugers Arch* **442**,

859–866.

Suh, S. J., Silo-Suh, L., Woods, D. E., Hassett, D. J., West, S. E. & Ohman, D. E. (1999). Effect of rpoS mutation on the stress response and expression of virulence factors in Pseudomonas aeruginosa. *J Bacteriol* **181**, 3890–3897.

Sultan, M., Arya, R. & Kim, K. K. (2021). Roles of Two-Component Systems in Pseudomonas aeruginosa Virulence. *Int J Mol Sci* **22**.

Sun, L., Frykholm, K., Fornander, L. H., Svedhem, S., Westerlund, F. & Akerman, B. (2014). Sensing conformational changes in DNA upon ligand binding using QCM-D. Polyamine condensation and Rad51 extension of DNA layers. *J Phys Chem B* **118**, 11895–11904.

Suzek, B. E., Ermolaeva, M. D., Schreiber, M. & Salzberg, S. L. (2001). A probabilistic method for identifying start codons in bacterial genomes. *Bioinformatics* **17**, 1123–1130.

Tamber, S. & Hancock, R. E. W. (2006). Involvement of two related porins, OprD and OpdP, in the uptake of arginine by Pseudomonas aeruginosa. *FEMS Microbiol Lett* **260**, 23–29.

Tamber, S., Ochs, M. M. & Hancock, R. E. W. (2006). Role of the novel OprD family of porins in nutrient uptake in Pseudomonas aeruginosa. *J Bacteriol* **188**, 45–54.

Tan, X., Cheng, X., Hu, M., Zhang, Y., Jia, A., Zhou, J. & Zhu, G. (2021). Transcriptional analysis and target genes discovery of Pseudomonas aeruginosa biofilm developed ex vivo chronic wound model. *AMB Express* **11**, 157.

Tanaka, K. & Takahashi, H. (1991). Cloning and analysis of the gene (rpoDA) for the principal σ factor of Pseudomonas aeruginosa. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1089**, 113–119.

Thaarup, I. C. & Bjarnsholt, T. (2020). Current in vitro biofilm-infected chronic wound models for developing new treatment possibilities. *Adv Wound Care (New Rochelle)*.

Thanassi, D. G. & Hultgren, S. J. (2000). Multiple pathways allow protein secretion across the bacterial outer membrane. *Curr Opin Cell Biol* **12**, 420–430.

Thanassi, D. G., Stathopoulos, C., Karkal, A. & Li, H. (2005). Protein secretion in the absence of ATP: the autotransporter, two-partner secretion and chaperone/usher pathways of Gram-negative bacteria (Review). *Mol Membr Biol* **22**, 63–72.

Thompson, L. S., Webb, J. S., Rice, S. A. & Kjelleberg, S. (2003). The alternative sigma factor RpoN regulates the quorum sensing gene*rhlI* in*Pseudomonas aeruginosa*. *FEMS Microbiol Lett* **220**, 187–195.

Thornley, M. J. (1960). The Differentiation of *Pseudomonas* From Other Gram-Negative Bacteria on the Basis of Arginine Metabolism. *Journal of Applied Bacteriology* **23**, 37–52.

Toyofuku, M., Roschitzki, B., Riedel, K. & Eberl, L. (2012). Identification of proteins associated with the Pseudomonas aeruginosa biofilm extracellular matrix. *J Proteome Res* **11**, 4906–4915.

Trias, J. & Nikaido, H. (1990). Protein D2 channel of the Pseudomonas aeruginosa outer membrane has a binding site for basic amino acids and peptides. *J Biol Chem* **265**, 15680–15684.

Ude, J., Tripathi, V., Buyck, J. M., Söderholm, S., Cunrath, O., Fanous, J., Claudi, B., Egli, A., Schleberger, C. & other authors. (2021). Outer membrane permeability: Antimicrobials and diverse nutrients bypass porins in Pseudomonas aeruginosa. *Proc Natl Acad Sci USA* **118**.

Vander Wauven, C., Piérard, A., Kley-Raymann, M. & Haas, D. (1984). Pseudomonas aeruginosa mutants affected in anaerobic growth on arginine: evidence for a four-gene cluster encoding the arginine deiminase pathway. *J Bacteriol* **160**, 928–934.

Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G. & other authors. (2022). AlphaFold Protein Structure Database: massively expanding the structural

coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res* **50**, D439–D444.

Venkataraman, A., Rosenbaum, M. A., Werner, J. J., Winans, S. C. & Angenent, L. T. (2014). Metabolite transfer with the fermentation product 2,3-butanediol enhances virulence by Pseudomonas aeruginosa. *ISME J* **8**, 1210–1220.

Verhoogt, H. J., Smit, H., Abee, T., Gamper, M., Driessen, A. J., Haas, D. & Konings, W. N. (1992). arcD, the first gene of the arc operon for anaerobic arginine catabolism in Pseudomonas aeruginosa, encodes an arginine-ornithine exchanger. *J Bacteriol* **174**, 1568–1573.

Voisey, C. R. & Marincs, F. (1998). Elimination of Internal Restriction Enzyme Sites from a Bacterial Luminescence (luxCDABE) Operon. *BioTechniques* **24**, 56–58.

Wagner, V. E., Bushnell, D., Passador, L., Brooks, A. I. & Iglewski, B. H. (2003). Microarray analysis of Pseudomonas aeruginosa quorum-sensing regulons: effects of growth phase and environment. *J Bacteriol* **185**, 2080–2095.

Wagner, V. E., Gillis, R. J. & Iglewski, B. H. (2004). Transcriptome analysis of quorum-sensing regulation and virulence factor expression in Pseudomonas aeruginosa. *Vaccine* **22 Suppl 1**, S15-20.

Waibel, A., Hanazawa, T., Hinton, G., Shikano, K. & Lang, K. J. (1989). Phoneme recognition using time-delay neural networks. *IEEE Trans Acoust* **37**, 328–339.

Wargo, M. J., Szwergold, B. S. & Hogan, D. A. (2008). Identification of two gene clusters and a transcriptional regulator required for Pseudomonas aeruginosa glycine betaine catabolism. *J Bacteriol* **190**, 2690–2699.

Wattam, A. R., Davis, J. J., Assaf, R., Boisvert, S., Brettin, T., Bun, C., Conrad, N., Dietrich, E. M., Disz, T. & other authors. (2017). Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. *Nucleic Acids Res* **45**, D535–D542.

Werthén, M., Henriksson, L., Jensen, P. Ø., Sternberg, C., Givskov, M. & Bjarnsholt, T. (2010). An in vitro model of bacterial infections in wounds and other soft tissues. *APMIS* **118**, 156–164.

West, S. E., Sample, A. K. & Runyen-Janecky, L. J. (1994). The vfr gene product, required for Pseudomonas aeruginosa exotoxin A and protease production, belongs to the cyclic AMP receptor protein family. *J Bacteriol* **176**, 7532–7542.

Whistler, C. A., Corbell, N. A., Sarniguet, A., Ream, W. & Loper, J. E. (1998). The two-component regulators GacS and GacA influence accumulation of the stationary-phase sigma factor sigmaS and the stress response in Pseudomonas fluorescens Pf-5. *J Bacteriol* **180**, 6635–6641.

Whitchurch, C. B., Leech, A. J., Young, M. D., Kennedy, D., Sargent, J. L., Bertrand, J. J., Semmler, A. B., Mellick, A. S., Martin, P. R. & other authors. **(2004).** Characterization of a complex chemosensory signal transduction system which controls twitching motility in *Pseudomonas aeruginosa*. *Mol Microbiol* **52**, 873–893.

Whiteley, M., Parsek, M. R. & Greenberg, E. P. (2000). Regulation of quorum sensing by RpoS in Pseudomonas aeruginosa. *J Bacteriol* **182**, 4356–4360.

Wilhelm, S., Gdynia, A., Tielen, P., Rosenau, F. & Jaeger, K.-E. (2007). The autotransporter esterase EstA of Pseudomonas aeruginosa is required for rhamnolipid production, cell motility, and biofilm formation. *J Bacteriol* **189**, 6695–6703.

Winardhi, R. S., Castang, S., Dove, S. L. & Yan, J. (2014). Single-molecule study on histone-like nucleoid-structuring protein (H-NS) paralogue in Pseudomonas aeruginosa: MvaU bears DNA organization mode similarities to MvaT. *PLoS One* **9**, e112246.

Winsor, G. L., Griffiths, E. J., Lo, R., Dhillon, B. K., Shay, J. A. & Brinkman, F. S. L. (2016). Enhanced annotations and features for comparing thousands of *Pseudomonas* genomes in the *Pseudomonas* genome database. *Nucleic Acids Res* **44**, D646-53.

Wolfgang, M. C., Lee, V. T., Gilmore, M. E. & Lory, S. (2003). Coordinate regulation of bacterial virulence genes by a novel adenylate cyclase-dependent signaling pathway. *Dev Cell* **4**, 253–263.

Woolwine, S. C. & Wozniak, D. J. (1999). Identification of an Escherichia coli pepA homolog and its involvement in suppression of the algB phenotype in mucoid Pseudomonas aeruginosa. *J Bacteriol* **181**, 107–116.

van der Woude, M. W. & Henderson, I. R. (2008). Regulation and function of Ag43 (flu). *Annu Rev Microbiol* **62**, 153–169.

Wu, G. & Morris, S. M. (1998). Arginine metabolism: nitric oxide and beyond. *Biochem J* **336 (Pt 1)**, 1–17.

Wurtzel, O., Yoder-Himes, D. R., Han, K., Dandekar, A. A., Edelheit, S., Greenberg, E. P., Sorek, R. & Lory, S. (2012). The single-nucleotide resolution transcriptome of Pseudomonas aeruginosa grown in body temperature. *PLoS Pathog* **8**, e1002945.

Xiao, G., Déziel, E., He, J., Lépine, F., Lesic, B., Castonguay, M.-H., Milot, S., Tampakaki, A. P., Stachel, S. E. & Rahme, L. G. (2006). MvfR, a key Pseudomonas aeruginosa pathogenicity LTTR-class regulatory protein, has dual ligands. *Mol Microbiol* **62**, 1689–1699.

Yang, J. & Zhang, Y. (2015). I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Res* **43**, W174–W181.

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J. & Zhang, Y. (2015). The I-TASSER Suite: protein structure and function prediction. *Nat Methods* **12**, 7–8.

Yang, L., Zhou, Y., Zhu, S., Huang, T., Wu, L. & Yan, X. (2012). Detection and quantification of bacterial autofluorescence at the single-cell level by a laboratory-built high-sensitivity flow cytometer. *Anal Chem* **84**, 1526–1532.

Yang, Z. & Lu, C.-D. (2007). Functional genomics enables identification of genes of the arginine transaminase pathway in Pseudomonas

aeruginosa. *J Bacteriol* **189**, 3945–3953.

Yelin, I., Aharony, N., Tamar, E. S., Argoetti, A., Messer, E., Berenbaum, D., Shafran, E., Kuzli, A., Gandali, N. & other authors. **(2020).** Evaluation of COVID-19 RT-qPCR Test in Multi sample Pools. *Clin Infect Dis* **71**, 2073–2078.

Yoshimura, F. & Nikaido, H. (1982). Permeability of Pseudomonas aeruginosa outer membrane to hydrophilic solutes. *J Bacteriol* **152**, 636–642.

Zakeri, B., Fierer, J. O., Celik, E., Chittock, E. C., Schwarz-Linek, U., Moy, V. T. & Howarth, M. (2012). Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proc Natl Acad Sci USA* **109**, E690-7.

Zhang, L. H., Fath, M. J., Mahanty, H. K., Tai, P. C. & Kolter, R. (1995). Genetic analysis of the colicin V secretion pathway. *Genetics* **141**, 25–32.

Zhao, K., Li, W., Li, J., Ma, T., Wang, K., Yuan, Y., Li, J. S., Xie, R., Huang, T. & other authors. **(2019).** TesG is a type I secretion effector of Pseudomonas aeruginosa that suppresses the host immune response during chronic infection. *Nat Microbiol* **4**, 459–469.

Zheng, W., Zhang, C., Li, Y., Pearce, R., Bell, E. W. & Zhang, Y. (2021). Folding non-homologous proteins by coupling deep-learning contact maps with I-TASSER assembly simulations. *Cell Reports Methods* **1**.

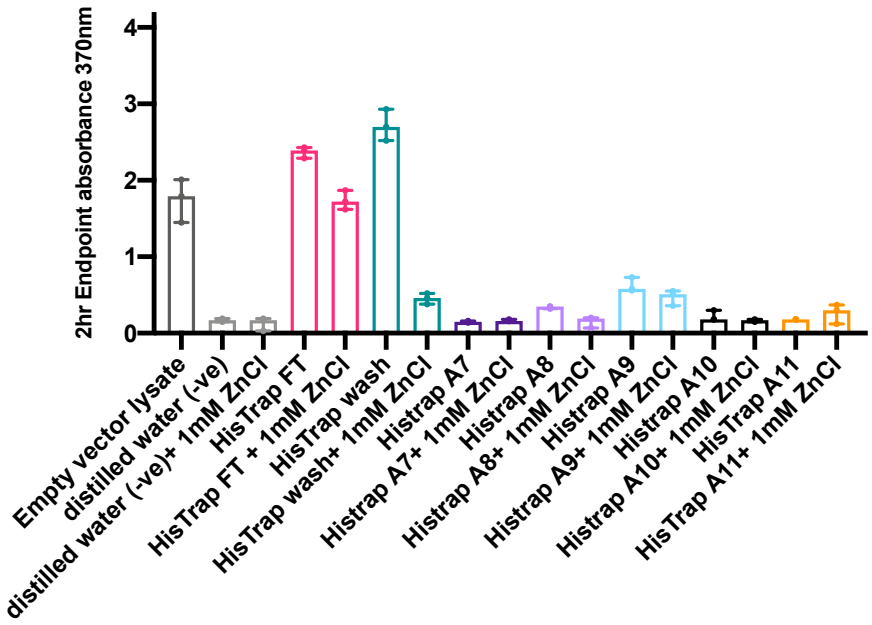
Zhong, X., Kolter, R. & Tai, P. C. (1996). Processing of colicin V-1, a secretable marker protein of a bacterial ATP binding cassette export system, requires membrane integrity, energy, and cytosolic factors. *J Biol Chem* **271**, 28057–28063.

Zhou, T., Huang, J., Liu, Z., Xu, Z. & Zhang, L.-H. (2021). Molecular Mechanisms Underlying the Regulation of Biofilm Formation and Swimming Motility by FleS/FleR in Pseudomonas aeruginosa. *Front Microbiol* **12**, 707711.

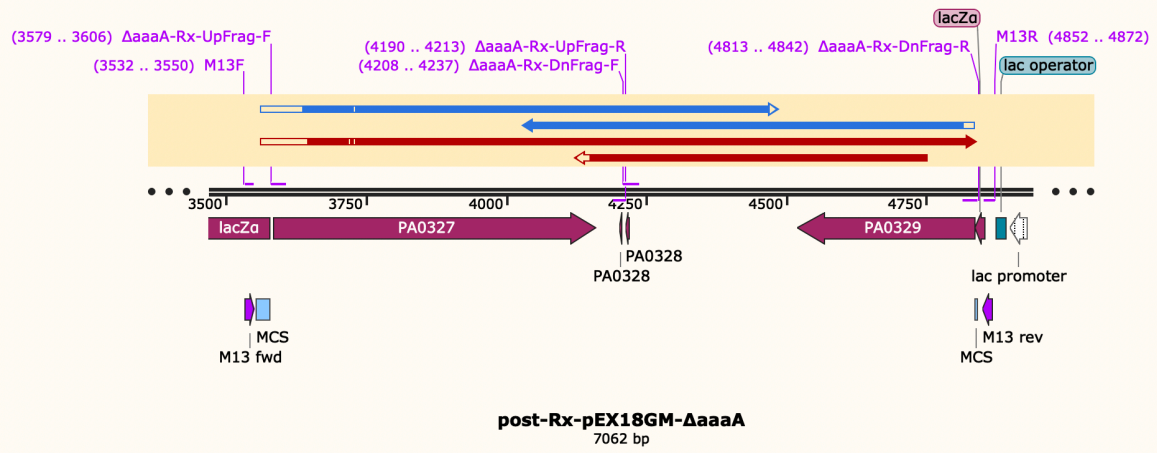
Zhou, T., Huang, J., Liu, Z., Lin, Q., Xu, Z. & Zhang, L.-H. (2022). The Two-

Component System FleS/FleR Represses H1-T6SS via Cyclic di-GMP Signaling in Pseudomonas aeruginosa. *Appl Environ Microbiol* **88**, e0165521.

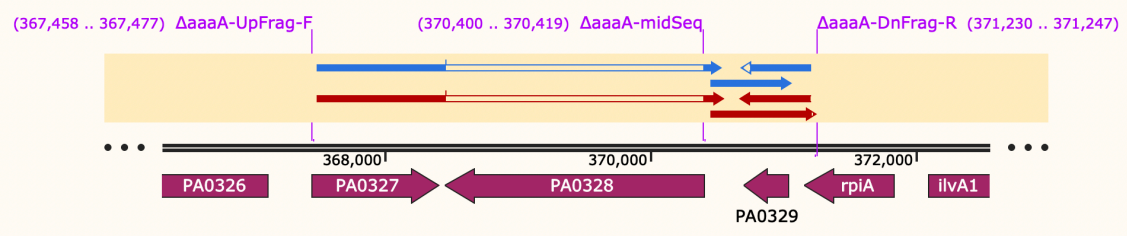
Supplementary Data



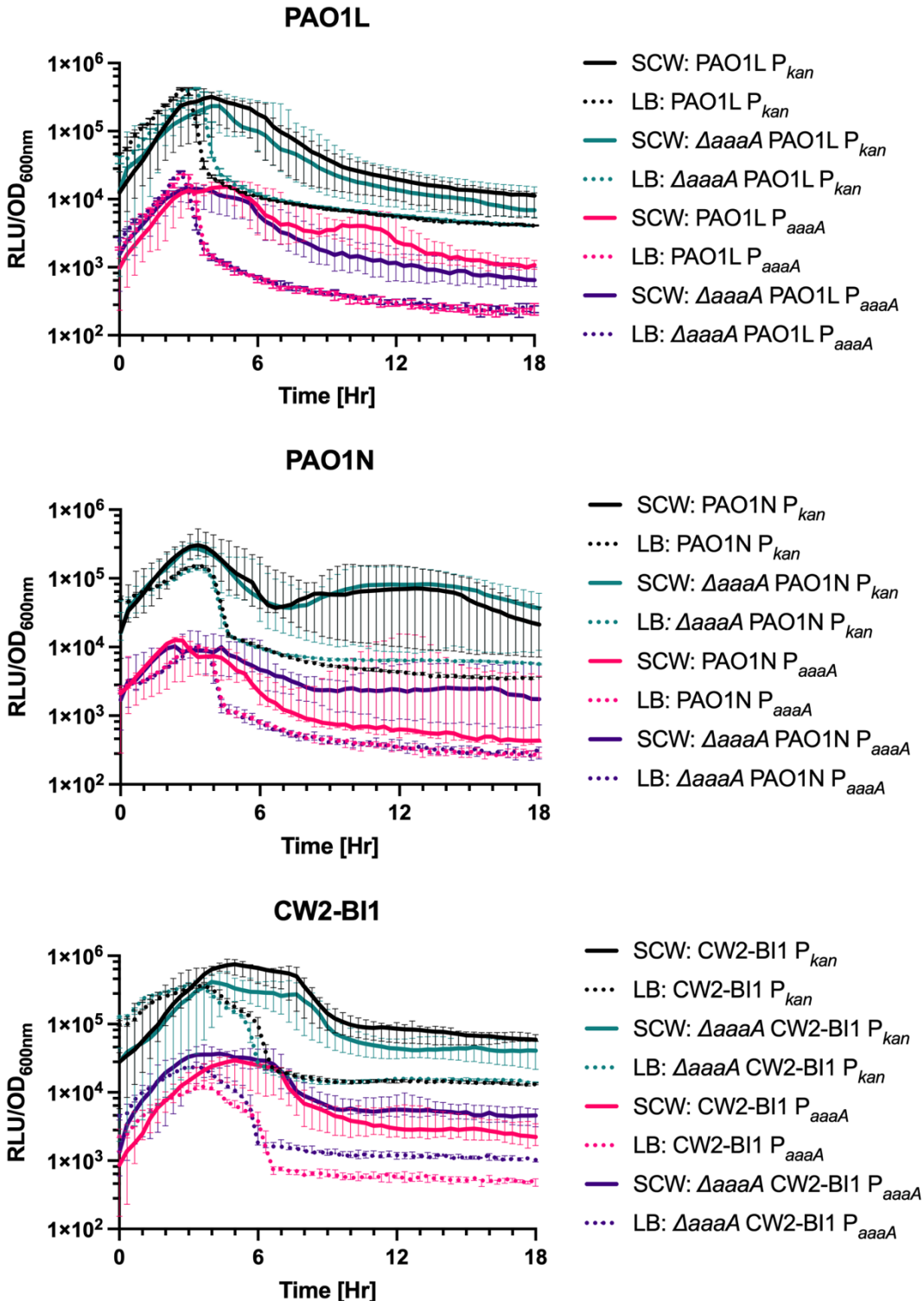
S-Figure 0.1: Addition of $ZnCl_2$ to fractions of AaaAt3H10 following HisTrap chromatography yielded no difference in activity. *Graph shows L-arginine-p-nitroanilide cleavage (AaaA activity) assay endpoint absorbance at 370 nm after 2 hrs (median ± 95% CI, n= 3). Moderate activity is seen in the crude lysate fraction from an empty* pET21a *vector control and the HisTrap flow through and washes while low activity is seen in all elution fraction. Addition of $ZnCl_2$ seems to have a negligible impact on activity.*



S-Figure 0.2: Alignment of sequenced inserts mapped onto pEX18Gm-ΔaaaA vector. *Alignments from sequencing reactions using the M13F/R primers are shown in arrows above the map, C3 (blue) and C10 (red). There are one and two ambiguous bases (indicated by gap in arrow around 3700 bp) in C3 and C10, respectively. Both potential mismatches were assigned as S (G or C) where the base in the reference was C. Genetic features on the vector are labelled below the map, and binding sites for primers used in cloning are indicated in purple above. Generated using Snapgene.*



S-Figure 0.3: Alignments of sequenced PCR products from PAO1L ΔaaaA candidates C7 and C13, spanning the region around PA0328, mapped onto the PAO1 genome. *Alignments from multiple sequencing reactions using the primers: ΔaaaA-UpFrag-F, ΔaaaA-midSeq and ΔaaaA-DnFrag-R are shown in arrows above the map: C7 (blue) and C13 (red). Across the whole length sequenced, there are no mismatches for either colony. Genetic features are labelled below the map, and binding sites for primers used in cloning are indicated in purple above. Generated using Snapgene.*



S-Figure 0.4: Activation of promotors of Kan and aaaA genes in the SCW vs LB over 19 hrs. *PAO1L wild-type and ΔaaaA mutant containing* miniCTX-*lux fused to either the kan or aaaA promoter were grown to mid-exponential phase (OD_{600nm} 0.5-0.9) before inoculation into a black 96-well plate. On different halves of the same plate, 100 µL SCW media was inoculated with OD_{600nm} 0.005 bacteria, and LB was inoculated with OD_{600nm} 0.01. The plate was covered with an optically clear gas-permeable seal and incubated statically at 37°C for 18 hrs. Individual replicates of RLU/OD_{600nm} are plotted (median ±95% CI, for SCW, n=6, 3 technical replicates on two different days, for LB n=3). Data collected with assistance from Bethan Roberts.*

S-Table 0.1: List of all differentially expressed genes in PAO1L ΔaaaA compared with PAO1L in the 16-hr SCW (p<0.05). *Genes are ordered by log2 fold change (FC) in ascending order, data generation, analysis and compilation by Novogene.*

Gene	Log2FC	pvalue	padj	gene_description	gene_start	strand
aaaA	-3.70	<0.0001	<0.0001	arginine-specific autotransporter of *Pseudomonas aeruginosa*, AaaA	368462	-
phzB2	-2.29	0.0030	0.7449	probable phenazine biosynthesis protein	2071209	+
phzA2	-2.25	0.0002	0.2321	probable phenazine biosynthesis protein	2070685	+
PA2146	-2.04	0.0403	0.9998	conserved hypothetical protein	2361706	+
cupE3	-1.99	0.0157	0.9998	Pilin subunit CupE3	5216203	+
cupE1	-1.62	0.0191	0.9998	Pilin subunit CupE1	5215091	+
PA2171	-1.60	0.0109	0.9998	hypothetical protein	2393708	+
PA1784	-1.58	0.0089	0.9998	hypothetical protein	1932129	-
lecB	-1.56	0.0003	0.2321	fucose-binding lectin PA-IIL	3773029	+
PA3734	-1.48	0.0077	0.9998	hypothetical protein	4185522	+
PA2134	-1.41	0.0258	0.9998	hypothetical protein	2349488	+
PA0624	-1.40	0.0072	0.9998	hypothetical protein	683085	+
PA2170	-1.38	0.0341	0.9998	hypothetical protein	2393424	+
PA2190	-1.33	0.0101	0.9998	conserved hypothetical protein	2409837	+
chiC	-1.32	0.0015	0.5229	chitinase	2530389	-
PA3371	-1.31	0.0114	0.9998	hypothetical protein	3780128	+
PA4078	-1.30	0.0144	0.9998	probable nonribosomal peptide synthetase	4555254	+
oprO	-1.28	0.0194	0.9998	Pyrophosphate-specific OM porin OprO precursor	3673008	-
PA2169	-1.24	0.0153	0.9998	hypothetical protein	2392945	+
PA4143	-1.20	0.0002	0.2152	probable toxin transporter	4634139	+
PA4142	-1.19	0.0003	0.2321	probable secretion protein	4632873	+
PA2151	-1.17	0.0281	0.9998	conserved hypothetical protein	2366106	+
glgP	-1.17	0.0288	0.9998	glycogen phosphorylase	2358364	+
PA0636	-1.13	0.0171	0.9998	hypothetical protein	690721	+
lapA	-1.13	0.0266	0.9998	low-molecular-weight alkaline phosphatase A, LapA	747470	+
PA0633	-1.13	0.0253	0.9998	hypothetical protein	689537	+
PA4384	-1.12	0.0213	0.9998	hypothetical protein	4914739	+
PA4141	-1.12	0.0370	0.9998	hypothetical protein	4632477	+
phaC2	-1.11	0.0266	0.9998	poly(3-hydroxyalkanoic acid) synthase 2	5698359	+
PA2565	-1.10	0.0087	0.9998	hypothetical protein	2899924	-
eddB	-1.10	0.0059	0.9998	Extracelullar DNA degradation protein, EddB	4377311	-
tadA	-1.09	0.0425	0.9998	TadA ATPase	4825923	-
PA2588	-1.09	0.0229	0.9998	probable transcriptional regulator	2928539	+
PA3368	-1.09	0.0364	0.9998	probable acetyltransferase	3778704	+
PA2566	-1.07	0.0226	0.9998	conserved hypothetical protein	2900372	-

Gene	Log2FC	pvalue	padj	gene_description	gene_start	strand
PA2148	-1.07	0.0296	0.9998	conserved hypothetical protein	2364311	+
phoA	-1.07	0.0412	0.9998	alkaline phosphatase	3689521	+
PA3369	-1.05	0.0255	0.9998	hypothetical protein	3779562	+
PA2635	-1.04	0.0358	0.9998	hypothetical protein	2979530	+
PA2031	-1.04	0.0390	0.9998	hypothetical protein	2220903	-
pqsB	-1.03	0.0075	0.9998	PqsB	1080009	+
PA4144	-1.03	0.0021	0.6137	probable OM protein precursor	4636298	+
mexG	-1.01	0.0074	0.9998	hypothetical protein	4705956	+
PA1870	-1.01	0.0139	0.9998	hypothetical protein	2031988	+
PA2881	-1.01	0.0412	0.9998	probable two-component response regulator	3233343	-
pqsC	-0.99	0.0013	0.4955	PqsC	1080853	+
pqsD	-0.99	0.0141	0.9998	3-oxoacyl-[acyl-carrier-protein] synthase III	1081942	+
pqsA	-0.98	0.0063	0.9998	PqsA	1078462	+
PA2160	-0.96	0.0212	0.9998	probable glycosyl hydrolase	2379310	-
PA0635	-0.95	0.0462	0.9998	hypothetical protein	690420	+
phzB1	-0.94	0.0005	0.3066	probable phenazine biosynthesis protein	4714314	+
PA2069	-0.94	0.0216	0.9998	probable carbamoyl transferase	2267638	-
phzA1	-0.93	0.0054	0.9779	probable phenazine biosynthesis protein	4713796	+
pprB	-0.92	0.0316	0.9998	two-component response regulator, PprB	4820532	+
phnA	-0.91	0.0120	0.9998	anthranilate synthase component I	1083972	+
pqsE	-0.91	0.0253	0.9998	Quinolone signal response protein	1082949	+
rcpA	-0.90	0.0455	0.9998	RcpA	4828379	-
PA2564	-0.90	0.0426	0.9998	hypothetical protein	2899107	-
PA0643	-0.90	0.0256	0.9998	hypothetical protein	699744	+
PA0622	-0.87	0.0345	0.9998	probable bacteriophage protein	681394	+
cbpD	-0.86	0.0268	0.9998	chitin-binding protein CbpD precursor	930653	-
plcH	-0.86	0.0346	0.9998	hemolytic phospholipase C precursor	919258	-
rhlA	-0.83	0.0072	0.9998	rhamnosyltransferase chain A	3892121	-
PA0673	-0.81	0.0429	0.9998	hypothetical protein	733673	-
rcpC	-0.80	0.0495	0.9998	RcpC	4829643	-
PA0698	-0.78	0.0394	0.9998	hypothetical protein	770847	+
PA3924	-0.77	0.0254	0.9998	probable medium-chain acyl-CoA ligase	4398259	-
PA2068	-0.77	0.0262	0.9998	probable major facilitator superfamily (MFS) transporter	2266431	-
PA3586	-0.77	0.0146	0.9998	probable hydrolase	4018458	-
PA2375	-0.75	0.0440	0.9998	hypothetical protein	2626780	-
PA1202	-0.74	0.0190	0.9998	probable hydrolase	1302696	+
PA4377	-0.71	0.0260	0.9998	hypothetical protein	4909162	-
katA	-0.70	0.0209	0.9998	catalase	4752260	-

Gene	Log2FC	pvalue	padj	gene_description	gene_start	strand
sahH	-0.69	0.0144	0.9998	S-adenosyl-L-homocysteine hydrolase	482706	-
PA4573	-0.69	0.0497	0.9998	hypothetical protein	5121337	-
nuoD	-0.67	0.0064	0.9998	NADH dehydrogenase I chain C,D	2983963	+
PA3123	-0.66	0.0243	0.9998	RidA subfamily protein	3505865	-
PA2609	-0.66	0.0356	0.9998	hypothetical protein	2949967	+
PA1545	-0.65	0.0261	0.9998	hypothetical protein	1682051	+
muiA	-0.63	0.0441	0.9998	mucoidy inhibitor gene A	1621472	-
phzM	-0.61	0.0336	0.9998	probable phenazine-specific methyltransferase	4712095	-
PA1789	-0.60	0.0185	0.9998	hypothetical protein	1938549	-
aspA	-0.59	0.0396	0.9998	aspartate ammonia-lyase	6109260	+
PA2953	-0.50	0.0388	0.9998	electron transfer flavoprotein-ubiquinone oxidoreductase	3312791	+
Novel00118	0.57	0.0223	0.9998	-	2012279	+
Novel00049	0.58	0.0356	0.9998	PF00171:Aldehyde dehydrogenase family	854595	-
Novel00350	0.59	0.0398	0.9998	PF00206:Lyase	4998858	+
pqqH	0.59	0.0253	0.9998	PqqH	2175062	+
nalD	0.59	0.0368	0.9998	NalD	4006510	+
Novel00163	0.60	0.0263	0.9998	PF00975:Thioesterase domain|PF03621:MbtH-like protein	2693736	+
PA3574a	0.63	0.0460	0.9998	Uncharacterized protein	4007130	-
Novel00166	0.64	0.0413	0.9998	PF00668:Condensation domain	2711655	+
Novel00114	0.66	0.0443	0.9998	PF00330:Aconitase family (aconitate hydratase)	1935075	+
cysA	0.67	0.0211	0.9998	sulfate transport protein CysA	313938	-
fecA	0.68	0.0438	0.9998	Fe(III) dicitrate transport protein FecA	4368837	+
Novel00415	0.69	0.0347	0.9998	PF00294:pfkB family carbohydrate kinase	5613356	+
Novel00169	0.71	0.0207	0.9998	PF00668:Condensation domain|PF13193:AMP-binding enzyme C-terminal domain|PF00501:AMP-binding enzyme|PF00550:Phosphopantetheine attachment site	2715348	+
Novel00345	0.72	0.0284	0.9998	PF13466:STAS domain	4985836	+
PA2500	0.72	0.0287	0.9998	probable major facilitator superfamily (MFS) transporter	2817449	+
hasE	0.73	0.0412	0.9998	metalloprotease secretion protein	3810610	-
PA4804	0.75	0.0466	0.9998	POTRA-like domain-containing usher, CupB3	5388702	-
PA4658	0.78	0.0326	0.9998	hypothetical protein	5226026	+
Novel00024	0.82	0.0497	0.9998	-	507849	-
PA1328	0.84	0.0392	0.9998	probable transcriptional regulator	1440639	-
Novel00305	0.86	0.0245	0.9998	PF07549:SecD/SecF GG Motif	4276686	+
PA1232	0.86	0.0492	0.9998	hypothetical protein	1332635	+
sphD	0.86	0.0421	0.9998	SphD	5997057	-
PA0817	0.87	0.0214	0.9998	probable ring-cleaving dioxygenase	894952	+

Gene	Log2FC	pvalue	padj	gene_description	gene_start	strand
PA0201	0.87	0.0367	0.9998	hypothetical protein	229954	+
PA1316	0.87	0.0491	0.9998	probable major facilitator superfamily (MFS) transporter	1425912	+
Novel00267	0.90	0.0121	0.9998	PF00171:Aldehyde dehydrogenase family	4002381	+
Novel00320	0.91	0.0104	0.9998	PF00885:6,7-dimethyl-8-ribityllumazine synthase	4534230	+
PA0752	0.91	0.0303	0.9998	conserved hypothetical protein	819855	-
Novel00348	0.92	0.0233	0.9998	PF00081:Iron/manganese superoxide dismutases, alpha-hairpin domain	4997701	+
PA3939	0.94	0.0207	0.9998	hypothetical protein	4417014	-
fepD	0.96	0.0286	0.9998	ferric enterobactin transport protein FepD	4655368	+
opdH	0.96	0.0015	0.5229	cis-aconitate porin OpdH	822915	-
sRNA00052	0.97	0.0206	0.9998	-	3585120	+
PA1282	0.98	0.0051	0.9523	probable major facilitator superfamily (MFS) transporter	1392564	-
lsfA	1.00	0.0152	0.9998	1-Cys peroxiredoxin LsfA	3855493	-
carO	1.02	0.0195	0.9998	calcium-regulated OB-fold protein CarO	359982	+
PA4889	1.04	0.0393	0.9998	probable oxidoreductase	5485097	-
Novel00436	1.05	0.0035	0.7449	PF06945:Protein of unknown function (DUF1289)|PF01491:Frataxin-like domain	5940598	+
ercS'	1.05	0.0333	0.9998	ErcS'	2160363	+
PA2204	1.06	0.0003	0.2321	putative binding protein component of ABC transporter	2422819	-
soxA	1.06	0.0371	0.9998	sarcosine oxidase alpha subunit	6097024	+
PA0751	1.12	0.0127	0.9998	conserved hypothetical protein	818825	-
Novel00078	1.12	0.0032	0.7449	PF00158:Sigma-54 interaction domain	1188092	-
esrC	1.16	0.0048	0.9501	EsrC	5149064	+
PA2112	1.19	0.0235	0.9998	conserved hypothetical protein	2322778	-
Novel00234	1.20	0.0246	0.9998	PF00479:Glucose-6-phosphate dehydrogenase, NAD binding domain	3573355	+
desB	1.22	0.0456	0.9998	acyl-CoA delta-9-desaturase, DesB	5483988	-
Novel00256	1.22	0.0011	0.4802	-	3817023	+
PA3505	1.23	0.0304	0.9998	hypothetical protein	3921270	-
Novel00280	1.25	0.0026	0.7377	PF00814:Glycoprotease family	4126576	+
PA1977	1.25	0.0298	0.9998	hypothetical protein	2163010	-
eraS	1.30	0.0041	0.8395	sensor kinase, EraS	2165213	+
PA1653	1.32	0.0012	0.4955	probable transcriptional regulator	1800071	-
PA3445	1.34	0.0259	0.9998	conserved hypothetical protein	3850830	-
PA3441	1.34	0.0084	0.9998	probable molybdopterin-binding protein	3847718	-
Novel00057	1.34	0.0001	0.1107	PF00501:AMP-binding enzyme|PF13193:AMP-binding enzyme C-terminal domain	971077	-
xphA	1.35	0.0099	0.9998	XphA	2028454	+
PA4894	1.40	0.0200	0.9998	hypothetical protein	5489041	+

Gene	Log2FC	pvalue	padj	gene_description	gene_start	strand
tauD	1.44	0.0202	0.9998	taurine dioxygenase	4413192	-
eraR	1.47	0.0010	0.4727	response regulator EraR	2165876	+
PA3936	1.59	0.0113	0.9998	probable permease of ABC taurine transporter	4414132	-
PA1975	1.59	0.0078	0.9998	hypothetical protein	2159270	+
exaA	1.66	0.0006	0.3373	quinoprotein ethanol dehydrogenase	2167281	-
fdhA	1.67	0.0034	0.7449	glutathione-independent formaldehyde dehydrogenase	6101884	+
exaB	1.69	0.0074	0.9998	cytochrome c550	2169464	+
PA1981	1.70	0.0010	0.4727	hypothetical protein	2166580	-
Novel00177	1.71	0.0059	0.9998	-	2869852	+
PA2311	2.58	0.0227	0.9998	hypothetical protein	2549748	+

S-Table 0.2: List of all differentially expressed genes in PAO1L + 100 µM L-arginine compared with PAO1L in the 16-hr SCW. *Genes are ordered by log2 fold change (FC) in ascending order. Data generation, analysis and compilation by Novogene.*

Gene	Log2FC	pvalue	padj	gene_description	gene_start	strand
PA3758	-1.61	0.0142	0.9997	probable N-acetylglucosamine-6-phosphate deacetylase	4208168	+
phzA2	-1.51	0.0127	0.9997	probable phenazine biosynthesis protein	2070685	+
PA3734	-1.35	0.0049	0.9997	hypothetical protein	4185522	+
PA3371	-1.14	0.0051	0.9997	hypothetical protein	3780128	+
PA4078	-1.10	0.0429	0.9997	probable nonribosomal peptide synthetase	4555254	+
PA3293	-1.10	0.0387	0.9997	hypothetical protein	3685759	-
PA0269	-1.01	0.0124	0.9997	conserved hypothetical protein	305725	+
Novel00119	-0.99	0.0133	0.9997	PF01717:Cobalamin-independent synthase, Catalytic domain	2106983	-
metE	-0.98	0.0095	0.9997	5-methyltetrahydropteroyltriglutamate-homocysteine S-methyltransferase	2106580	+
PA0270	-0.92	0.0384	0.9997	hypothetical protein	306174	+
PA2096	-0.89	0.0472	0.9997	probable transcriptional regulator	2306776	-
PA3369	-0.88	0.0269	0.9997	hypothetical protein	3779562	+
PA1888	-0.82	0.0391	0.9997	hypothetical protein	2057930	-
Novel00118	-0.79	0.0432	0.9997	PF14467:Domain of unknown function (DUF4426)	429458	+
PA3041	-0.70	0.0228	0.9997	hypothetical protein	3404145	+
ppgL	-0.69	0.0441	0.9997	periplasmic gluconolactonase, PpgL	4704139	+
fpvH	-0.66	0.0156	0.9997	FpvH	2688705	+
PA2915	-0.65	0.0367	0.9997	hypothetical protein	3270827	-
alpA	-0.64	0.0498	0.9997	lysis phenotype activator, AlpA	992714	+
phzA1	-0.62	0.0154	0.9997	probable phenazine biosynthesis protein	4713796	+
gltB	0.49	0.0443	0.9997	glutamate synthase large chain precursor	5667696	-
PA5515	0.51	0.0378	0.9997	hypothetical protein	6208042	-
rsmN	0.53	0.0359	0.9997	RsmN	5836470	-
PA5438	0.54	0.0310	0.9997	probable transcriptional regulator	6121172	-
slyD	0.54	0.0197	0.9997	peptidyl-prolyl cis-trans isomerase SlyD	913086	+
lpxC	0.54	0.0456	0.9997	UDP-3-O-acyl-N-acetylglucosamine deacetylase	4938276	-
PA3904	0.56	0.0263	0.9997	PAAR4	4373938	+
nuoA	0.57	0.0471	0.9997	NADH dehydrogenase I chain A	2982781	+
PA4816	0.58	0.0447	0.9997	hypothetical protein	5406165	+
PA2199	0.59	0.0296	0.9997	probable dehydrogenase	2417760	+
PA0310	0.60	0.0243	0.9997	hypothetical protein	350890	+
PA0358	0.62	0.0363	0.9997	hypothetical protein	402020	+
PA0142	0.63	0.0289	0.9997	hypothetical protein	161906	+

Gene	Log2FC	pvalue	padj	gene_description	gene_start	strand
PA3033	0.63	0.0421	0.9997	hypothetical protein	3397096	-
betT3	0.63	0.0312	0.9997	BetT3	4408924	+
PA4940	0.65	0.0432	0.9997	conserved hypothetical protein	5544636	-
rho	0.67	0.0148	0.9997	transcription termination factor Rho	5898865	-
PA0257	0.69	0.0392	0.9997	hypothetical protein	288384	-
secE	0.74	0.0049	0.9997	secretion protein SecE	4783771	-
PA4753	0.74	0.0121	0.9997	conserved hypothetical protein	5338618	+
betI	0.75	0.0211	0.9997	transcriptional regulator BetI	6050718	-
fabA	0.76	0.0016	0.9997	beta-hydroxydecanoyl-ACP dehydrase	1752903	-
PA3403	0.78	0.0325	0.9997	hypothetical protein	3808318	-
PA4889	0.82	0.0283	0.9997	probable oxidoreductase	5485097	-
PA4746	0.84	0.0029	0.9997	conserved hypothetical protein	5331503	-
PA2358	0.89	0.0446	0.9997	hypothetical protein	2605435	-
-	0.89	0.0442	0.9997	sRNA	3526607	+
capB	0.92	0.0221	0.9997	cold acclimation protein B	3653667	+
PA2506	0.93	0.0269	0.9997	hypothetical protein	2824283	+
PA3519	0.93	0.0272	0.9997	hypothetical protein	3935800	-
PA0839	0.94	0.0261	0.9997	probable transcriptional regulator	914412	+
Novel00045	0.94	0.0398	0.9997	-	791459	-
PA0645	0.94	0.0227	0.9997	hypothetical protein	701151	+
PA2204	0.96	0.0031	0.9997	putative binding protein component of ABC transporter	2422819	-
opdH	0.96	0.0055	0.9997	cis-aconitate porin OpdH	822915	-
PA2910	0.96	0.0221	0.9997	conserved hypothetical protein	3264642	-
PA1112.1	1.00	0.0013	0.9997	sRNA	1205031	-
pscR	1.00	0.0031	0.9997	translocation protein in type III secretion	1842568	-
Novel00331	1.01	0.0172	0.9997	-	4785855	+
PA3009	1.02	0.0004	0.5739	hypothetical protein	3369036	-
PA4805	1.03	0.0374	0.9997	probable class III aminotransferase	5390148	-
PA1425	1.09	0.0066	0.9997	probable ATP-binding component of ABC transporter	1550984	+
PA2790	1.10	0.0124	0.9997	hypothetical protein	3147765	+
xphA	1.11	0.0307	0.9997	XphA	2028454	+
esrC	1.17	0.0001	0.1678	EsrC	5149064	+
PA5471.1	1.23	0.0072	0.9997	PA5471 leader peptide	6160912	-
PA4188	1.23	0.0486	0.9997	conserved hypothetical protein	4685090	+
PA1282	1.35	0.0164	0.9997	probable major facilitator superfamily (MFS) transporter	1392564	-
PA0252	1.35	0.0011	0.9997	hypothetical protein	283818	-
PA0805.1	1.37	0.0000	0.1678	sRNA	883307	-
Novel00177	1.41	0.0389	0.9997	-	2869852	+
PA4148	1.51	0.0473	0.9997	probable short-chain dehydrogenase	4641711	+

Gene	Log2FC	pvalue	padj	gene_description	gene_start	strand
P15	1.52	0.0298	0.9997	sRNA P15	3299022	+
eftM	1.91	0.0051	0.9997	SAM-dependent methyltransferase , EftM	4673963	+
PA2311	2.33	0.0248	0.9997	hypothetical protein	2549748	+
PA1541	2.48	0.0001	0.1678	probable drug efflux transporter	1678584	-
PA2836	2.84	0.0352	0.9997	probable secretion protein	3189150	+

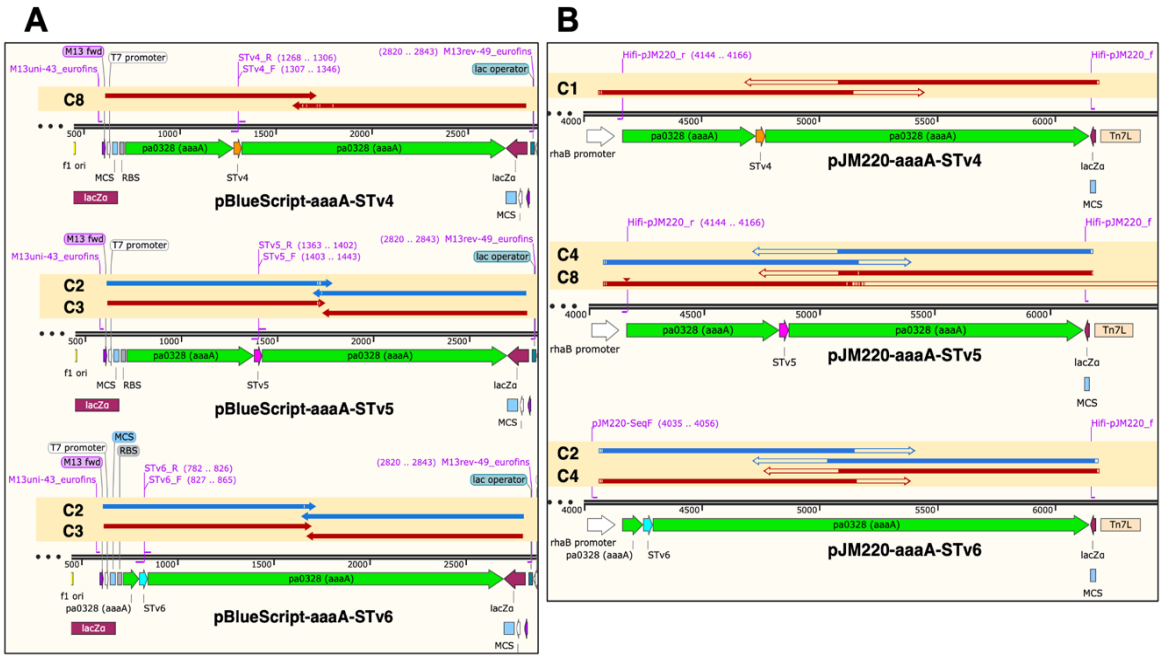
S-Table 0.3: List of all differentially expressed genes in PAO1L ΔaaaA + 100 µM L-arginine compared with PAO1L. *Genes are ordered by log2 fold change (FC) in ascending order. Data generation, analysis and compilation by Novogene.*

Gene	Log2FC	pvalue	padj	gene_description	gene_start	strand
aaaA	-3.81	0.0000	<0.0001	arginine-specific autotransporter of *Pseudomonas aeruginosa*, AaaA	368462	-
orfX	-3.08	0.0049	0.9992	OrfX	1814995	+
phzA2	-2.66	0.0001	0.0677	probable phenazine biosynthesis protein	2070685	+
phzB2	-2.08	0.0072	0.9992	probable phenazine biosynthesis protein	2071209	+
PA4141	-2.00	0.0000	0.0359	hypothetical protein	4632477	+
chiC	-1.88	0.0000	0.0359	chitinase	2530389	-
PA3237	-1.80	0.0190	0.9992	hypothetical protein	3624837	+
PA4143	-1.63	0.0000	0.0359	probable toxin transporter	4634139	+
PA3520	-1.60	0.0026	0.9992	putative periplasmic substrate binding protein	3937238	-
PA4142	-1.56	0.0001	0.0567	probable secretion protein	4632873	+
PA2069	-1.54	0.0006	0.3173	probable carbamoyl transferase	2267638	-
PA2588	-1.53	0.0014	0.6559	probable transcriptional regulator	2928539	+
PA2171	-1.44	0.0266	0.9992	hypothetical protein	2393708	+
metE	-1.44	0.0105	0.9992	5-methyltetrahydropteroyltriglutamate-homocysteine S-methyltransferase	2106580	+
PA4133	-1.44	0.0000	0.0359	cytochrome c oxidase subunit (cbb3-type)	4622812	+
mexG	-1.42	0.0003	0.1872	hypothetical protein	4705956	+
Novel00125	-1.41	0.0118	0.9992	PF02543:Carbamoyltransferase N-terminus	2268900	+
lecB	-1.40	0.0080	0.9992	fucose-binding lectin PA-IIL	3773029	+
rmf	-1.38	0.0091	0.9992	ribosome modulation factor	3414401	+
PA4078	-1.36	0.0167	0.9992	probable nonribosomal peptide synthetase	4555254	+
PA2066	-1.36	0.0043	0.9992	hypothetical protein	2265126	-
hsiF2	-1.36	0.0175	0.9992	HsiF2	1807241	+
cupE4	-1.34	0.0186	0.9992	Pilin assembly chaperone CupE4	5216764	+
cbpD	-1.32	0.0008	0.3926	chitin-binding protein CbpD precursor	930653	-
PA4739	-1.32	0.0221	0.9992	conserved hypothetical protein	5322757	-
pra	-1.30	0.0284	0.9992	protein activator	5140701	-
phzH	-1.30	0.0082	0.9992	potential phenazine-modifying enzyme	66303	+
PA4144	-1.25	0.0061	0.9992	probable OM protein precursor	4636298	+
PA3734	-1.24	0.0193	0.9992	hypothetical protein	4185522	+
PA2792	-1.24	0.0394	0.9992	hypothetical protein	3148723	-
PA2067	-1.23	0.0181	0.9992	probable hydrolase	2265761	-
Novel00285	-1.19	0.0310	0.9992	PF02868:Thermolysin metallopeptidase, alpha-helical domain|PF01447:Thermolysin metallopeptidase, catalytic domain|PF07504:Fungalysin/Thermolysin Propeptide Motif|PF03413:Peptidase propeptide and YPEB domain	4168966	+

Gene	Log2FC	pvalue	padj	gene_description	gene_start	strand
PA5507	-1.17	0.0414	0.9992	hypothetical protein	6198917	+
Novel00123	-1.16	0.0473	0.9992	PF16861:Carbamoyltransferase C-terminus	2267429	+
Novel00119	-1.16	0.0341	0.9992	PF01717:Cobalamin-independent synthase, Catalytic domain|PF08267:Cobalamin-independent synthase, N-terminal domain	2106983	-
hsiG2	-1.13	0.0080	0.9992	HsiG2	1808193	+
PA3496	-1.13	0.0270	0.9992	hypothetical protein	3913876	+
ahpF	-1.12	0.0108	0.9992	alkyl hydroperoxide reductase subunit F	158907	+
PA4134	-1.12	0.0029	0.9992	hypothetical protein	4624323	+
phzB1	-1.11	0.0033	0.9992	probable phenazine biosynthesis protein	4714314	+
PA1533	-1.10	0.0021	0.8493	conserved hypothetical protein	1672080	+
Novel00421	-1.09	0.0150	0.9992	PF01197:Ribosomal protein L31	5687053	+
Novel00054	-1.09	0.0277	0.9992	PF18416:N-acetylglucosamine binding protein domain 2|PF03067:Lytic polysaccharide mono-oxygenase, cellulose-degrading	930647	+
acp1	-1.09	0.0046	0.9992	Acp1	2031466	+
hsiB2	-1.08	0.0068	0.9992	HsiB2	1805218	+
hsiC2	-1.06	0.0091	0.9992	HsiC2	1805753	+
PA5271	-1.05	0.0048	0.9992	hypothetical protein	5935988	+
rpmG	-1.03	0.0087	0.9992	50S ribosomal protein L33	5985716	-
PA0754	-1.02	0.0494	0.9992	hypothetical protein	821902	-
hcnA	-1.01	0.0085	0.9992	hydrogen cyanide synthase HcnA	2412546	+
PA5148	-1.01	0.0231	0.9992	conserved hypothetical protein	5797065	+
PA1135	-1.01	0.0489	0.9992	conserved hypothetical protein	1226427	-
rpmE	-1.01	0.0074	0.9992	50S ribosomal protein L31	5687105	-
lasB	-0.99	0.0224	0.9992	elastase LasB	4168987	-
phzA1	-0.98	0.0085	0.9992	probable phenazine biosynthesis protein	4713796	+
Novel00022	-0.96	0.0179	0.9992	PF00670:S-adenosyl-L-homocysteine hydrolase, NAD binding domain|PF05221:S-adenosyl-L-homocysteine hydrolase	482700	+
liuC	-0.95	0.0306	0.9992	putative 3-methylglutaconyl-CoA hydratase	2202649	-
PA4645	-0.93	0.0449	0.9992	probable purine/pyrimidine phosphoribosyl transferase	5212107	-
glyA3	-0.91	0.0100	0.9992	serine hydroxymethyltransferase	5160738	+
hsiH2	-0.90	0.0453	0.9992	HsiH2	1809737	+
metK	-0.89	0.0268	0.9992	methionine adenosyltransferase	603706	-
rsaL	-0.88	0.0271	0.9992	regulatory protein RsaL	1558880	-
hcnB	-0.88	0.0261	0.9992	hydrogen cyanide synthase HcnB	2412857	+
hisH2	-0.88	0.0415	0.9992	glutamine amidotransferase	3535971	-
clpV2	-0.87	0.0156	0.9992	clpV2	1810751	+
speC	-0.85	0.0237	0.9992	ornithine decarboxylase	5060115	-

Gene	Log2FC	pvalue	padj	gene_description	gene_start	strand
liuA	-0.85	0.0255	0.9992	putative isovaleryl-CoA dehydrogenase	2205190	-
wbpG	-0.84	0.0261	0.9992	LPS biosynthesis protein WbpG	3533948	-
wbpI	-0.83	0.0388	0.9992	UDP-N-acetylglucosamine 2-epimerase WbpI	3531751	-
rpsM	-0.81	0.0173	0.9992	30S ribosomal protein S13	4756492	-
rplS	-0.80	0.0295	0.9992	50S ribosomal protein L19	4195008	-
sahH	-0.80	0.0236	0.9992	S-adenosyl-L-homocysteine hydrolase	482706	-
liuD	-0.80	0.0346	0.9992	methylcrotonyl-CoA carboxylase, alpha-subunit (biotin-containing)	2200685	-
aspA	-0.80	0.0246	0.9992	aspartate ammonia-lyase	6109260	+
dadA	-0.79	0.0273	0.9992	D-amino acid dehydrogenase, small subunit	5972178	-
PA5220	-0.79	0.0458	0.9992	hypothetical protein	5877896	+
liuB	-0.79	0.0389	0.9992	methylcrotonyl-CoA carboxylase, beta-subunit	2203460	-
tseT	-0.78	0.0438	0.9992	TOX-REase-5 domain-containing effector, TseT	4375230	+
rplW	-0.72	0.0361	0.9992	50S ribosomal protein L23	4765713	-
rplV	-0.71	0.0324	0.9992	50S ribosomal protein L22	4764243	-
rpsS	-0.71	0.0365	0.9992	30S ribosomal protein S19	4764588	-
atpB	-0.68	0.0432	0.9992	ATP synthase A chain	6253542	-
rpsL	-0.67	0.0443	0.9992	30S ribosomal protein S12	4771756	-
rplB	-0.67	0.0436	0.9992	50S ribosomal protein L2	4764880	-
rplK	-0.66	0.0470	0.9992	50S ribosomal protein L11	4782680	-
rpsF	-0.66	0.0450	0.9992	30S ribosomal protein S6	5537319	-
secE	0.70	0.0325	0.9992	secretion protein SecE	4783771	-
PA5275	0.72	0.0318	0.9992	conserved hypothetical protein	5940747	-
czcA	0.84	0.0433	0.9992	Resistance-Nodulation-Cell Division (RND) divalent metal cation efflux transporter CzcA	2837334	-
PA0805.1	0.85	0.0182	0.9992	sRNA PA0805.1	883307	-
PA3669	0.88	0.0266	0.9992	hypothetical protein	4107505	-
PA4746	0.89	0.0130	0.9992	conserved hypothetical protein	5331503	-
PA3574a	0.90	0.0130	0.9992	Uncharacterized protein	4007130	-
PA3009	0.93	0.0059	0.9992	hypothetical protein	3369036	-
fabA	0.94	0.0164	0.9992	beta-hydroxydecanoyl-ACP dehydrase	1752903	-
norB	0.94	0.0196	0.9992	nitric-oxide reductase subunit B	582455	+
PA2691	0.94	0.0160	0.9992	conserved hypothetical protein	3044947	-
PA3601	0.98	0.0230	0.9992	conserved hypothetical protein	4035758	-
PA2204	0.99	0.0225	0.9992	putative binding protein component of ABC transporter	2422819	-
PA0104	1.01	0.0208	0.9992	hypothetical protein	126518	+
nosD	1.04	0.0340	0.9992	NosD protein	3798916	+
Novel00436	1.04	0.0328	0.9992	PF06945:Protein of unknown function (DUF1289)|PF01491:Frataxin-like domain	5940598	+
norC	1.05	0.0035	0.9992	nitric-oxide reductase subunit C	582015	+
PA0714	1.06	0.0383	0.9992	hypothetical protein	784698	+

Gene	Log2FC	pvalue	padj	gene_description	gene_start	strand
PA1864	1.07	0.0469	0.9992	probable transcriptional regulator	2023281	-
esrC	1.07	0.0094	0.9992	EsrC	5149064	+
PA0525	1.07	0.0179	0.9992	probable dinitrification protein NorD	583857	+
PA3039	1.10	0.0334	0.9992	probable transporter	3402134	-
PA3600	1.11	0.0281	0.9992	conserved hypothetical protein	4035606	-
PA1425	1.14	0.0130	0.9992	probable ATP-binding component of ABC transporter	1550984	+
nrdG	1.14	0.0358	0.9992	class III (anaerobic) ribonucleoside-triphosphate reductase activating protein, 'activase', NrdG	2093444	-
ntrB	1.18	0.0242	0.9992	two-component sensor NtrB	5772298	+
PA5471.1	1.25	0.0224	0.9992	PA5471 leader peptide	6160912	-
nosR	1.35	0.0034	0.9992	regulatory protein NosR	3794819	+
hxcV	1.36	0.0118	0.9992	HxcV	738111	-
pauB1	1.41	0.0032	0.9992	FAD-dependent oxidoreductase	593273	+
pelG	1.45	0.0257	0.9992	PelG	3421696	-
PA3444	1.45	0.0266	0.9992	conserved hypothetical protein	3849604	-
xphA	1.47	0.0302	0.9992	XphA	2028454	+
PA4364	1.56	0.0021	0.8493	hypothetical protein	4892345	+
Novel00177	1.61	0.0385	0.9992	-	2869852	+
PA0021	1.61	0.0220	0.9992	conserved hypothetical protein	22872	+
nasS	1.77	0.0385	0.9992	NasS	1933649	-
PA2784	2.68	0.0483	0.9992	hypothetical protein	3141719	+



S-Figure 0.5: Alignments of sequenced products show successful construction of Spytagged AaaA. Panel A *shows the Snapgene alignment of sequencing for* pBlueScript*:aaaA-STv4-6 with good alignment around the site of Spytag insertion in all. For STv4, on the reverse read, there are 4 possible mutations: C p.1643 and p.1663 both G/C (R309G and R315R, respectively) and C p.1732 and C p. 1799 both to C/T (A338A and R360R, respectively). On the forward read only, there are 2 possible mutations: C p.1714 to A/C (D332E) and C p.1722 to C/G (A335G). On both F/R reads there is 1 possible mutation: C p.1718 to C/G (L334V). For Stv5, on the C2 reverse read, there is 1 possible mutation: C p.1735 to C/T (S339F). On the C2 forward read, there are 3 possible mutations: C p.1714 to A/C (D332E), C p.1722 to C/G (A335G) and A p.1727 to A/C (T337P), with one possible mutation on reverse read: C p. 1734 to C/T (S339F). On the C3 forward read, there are 2 possible mutations: C p.1714 to A/C (D332E) and C p.1722 to C/G (A335G), no mutations on the C3 reverse read. For STv6, on the C2 forward read, there is 1 possible mutation: G p. 1662 to G/C (R315P).* ***Panel B*** *shows the Snapgene alignment of sequencing for* pJM220*-aaaA-STv4-6 with good alignment and no errors across the entire read for STv4 C1, STv5 C4 (blue arrows) and both STv6 colonies (C2, blue arrows, was taken forward).*