Uncovering the role of Macrophage Infectivity Potentiator (Mip) proteins in *Pseudomonas aeruginosa* (PA01) virulence

Benjamin David Scott

John Innes Centre

Department of Molecular Microbiology

Norwich Research Park, Colney Lane, Norwich, NR4 7UH, UK

A thesis submitted to the University of East Anglia, Medical School for the degree of Doctor of Philosophy

December 2022

Uncovering the role of Macrophage Infectivity Potentiator (Mip) proteins in *Pseudomonas aeruginosa* (PA01) virulence

Benjamin David Scott, John Innes Centre, December 2022

Abstract

Currently, in a clinical setting, Microbiologists face an ever-growing challenge in treating multi-drug resistant bacterial infections using the conventional antibiotics currently at our disposal. One gram-negative ESKAPE pathogen, *Pseudomonas aeruginosa*, is currently listed at the top of the priority one pathogens by the World Health Organisation (WHO) that require urgent research and development of new antimicrobials. With a limited number of antibiotics currently in clinical trials, new drug targets are desperately needed, with recent investigations targeting bacterial virulence proteins. One promising virulence target are a group of proteins termed Macrophage Infectivity Potentiators (Mips). These proteins belong to the FK506 binding protein family of peptidyl-prolyl *cis/trans* isomerases (PPIases). Mip proteins have been discovered in several gram-negative pathogens including *Legionella pneumophila* and *Burkholderia pseudomallei,* where they have been shown to be essential in the invasion of macrophages, in murine-infection models, and in several other virulence determinants. Inhibition of these Mip-proteins by FK506 or rapamycin reduces virulence of the pathogens drastically.

Here, we investigated whether Mip-like protein candidates exist in the *P. aeruginosa* genome. This study describes the identification and characterisation of three Mip-proteins, PaMip1, PaMip2 and PaMip3, that are required for the full virulence of *P. aeruginosa* in two *in vivo* models. To further unpick the pleiotropic functions of these PaMips, and understand how they may exert their virulence effects, we performed *in vitro* plate based virulence assays with Mip mutants which revealed attenuation in several important virulence processes including biofilm formation and twitching motility. To elucidate these phenotypes, RNA sequencing analysis was performed which revealed the significant downregulation in an unstudied cluster of genes associated with the Type Six Secretion System hcp-island III (T6SS HSI-III). Strains housing deletions in the T6SS HSI-III exhibited attenuation in *in vivo* infection models similar to the PaMip knockout strains.

Acknowledgements

This thesis is dedicated to my Grandad, Dennis Arthur Sykes, who sadly passed away during its curation.

My first and most immediate thanks goes to my supervisor Prof. Barrie Wilkinson for his outstanding guidance and supervision throughout this entire process. He has helped inordinate amounts both professionally and personally. He has ensured through difficult times, that my mental health was paramount and has been beyond supportive. I have enjoyed countless discussions with him about this project along with personal matters on innumerable occasions. I am greatly looking forward to continuing this project with him in the future. I would also like to give thanks to my supervisory team Prof. Mark Webber and Dr. Eleftheria Trampari who have provided excellent discussions and thoughts on the project over the past four years. I would further like to thank Dr. Silke Alt who began this project and created many strains which were handed over to me while helping me hit the ground running when I first started my PhD and taught me all the molecular tools I needed to complete this thesis.

Further thanks goes to everyone in the JIC Molecular Microbiology department who have helped me throughout this work. It has been an honour to be a part of such a world-renowned department that has provided me with so many invaluable skills. The environment has been nurturing, providing a professional place of work while simultaneously a relaxed environment where any member of the Chatt building is approachable and keen for a scientific chat or to provide any advice where needed. Special thanks goes to Dr Govind Chandra for his extensive bioinformatics support on this project. I would like to thank Dr Carlo Martins and Dr Gerhard Saalbach for their proteomics expertise and analysis. Great thanks goes to my friends who began their PhD journeys with me, (future Drs) Jonathon Liston and Hannah McDonald. Jonathon provided invaluable protein expertise and general scientific discussion. A big thank you to Prof. Sam Fountain and Sonia Paz who have been excellent collaborators on our macrophage work.

Finally, I would like to thank my family, without whom I would not have been able to complete this doctorate. My Mum, Karen, who has always supported me and been by my side, and Ian, who has also been incredibly supportive and always provided words of encouragement. My grandparents, Dennis and Sheila, who have always supported and believed in me even if they didn't understand what I was talking about most of the time! My partner Jemma, who has been there throughout, always providing whatever it is I may need I as 'my emotional support crew' through both the good times and the rough times, in and out of this project.

Author's declaration

The research described in this thesis was conducted entirely at the John Innes Centre between September 2018 and December 2022. All the data described are original and were obtained by the author, except where specific acknowledgement has been made. No part of this thesis has previously been submitted as for a degree or any other academic institution.

Abbreviations

3-oxo-C_{12}-HSL	*N*-3-oxo-dodecanoyl homoserine lactone
C_4-HSL	*N*-butanoyl lactone
'A'	Adenine
AHL	*N*-Acyl homoserine lactones
Ala	Alanine
ATP	Adenosine triphosphate
AU	Absorbance
bFKBP	bovine FKBP
Bp	Base pair
BpML1	*Burkholderia pseudomallei* Mip-like protein 1
BPSL1823	*Burkholderia pseudomallei* 1823 (Mip)
'C'	Cytosine
Carb	Carbenicillin
cDNA	Complementary DNA
CF	Cystic Fibrosis
CsA	Cyclosporin A
Cbmip	*Coxiella burnetiid* Macrophage Infectivity Potentiator
ci-di-GMP	Cyclic diguanylate guanosine monophosphate
CFU	Colony forming unit
CtMip	*Chlamydia trachomatis* Macrophage Infectivity Potentiator
'D'	Aspartate
DMEM	Dulbecco's Modified Eagle Media
Dot	Defect in organelle trafficking
SEM	Scanning Electron Microscopy
EMBL	European Molecular Biology Laboratory
FDR	False Discovery Rate
Fha	Forkhead-associated protein
FKBP	FK506-binding protein
Gac	Global Activator of antibiotic and Cyanide synthesis
'G'	Guanine
Hcp	Haemolysin coregulated protein

HSI	Hcp Secretion Island
hFKBP	Human FK506-binding protein
HPLC	High-performance liquid chromatography
Hrp	Hypersensitive response and pathogenicity
Icm	Intracellular multiplication
IgG	Immunoglobulin G
IPTG	Isopropyl ß-D-1-thiogalactopyranoside
kDa	Kilo-Dalton
LadS	Lost adherence sensor
LgMip	*Legionella pneumophilia* Macrophage Infectivity Potentiator
Lip	Lipoprotein
LPS	Lipopolysaccharide
LogFC	Log Fold Change
MCS	Multiple cloning site
Mip	Macrophage Infectivity Potentiator
ml	Millilitres
mg	Milligrams
mm	Millimetres
MOI	Multiplicity of Infection
NBRF	National Biomedical Research Foundation
NcFKBP	*Neurospora crassa* FK506-binding protein
NGS	Next Generation Sequencing
NgMip	*Neisseria gonorrhoeae* Macrophage Infectivity Potentiator
NmMip	*Neisseria meningitidis* Macrophage Infectivity Potentiator
NMR	Nucleic magnetic resonance
OD	Optical Density
O/N	Overnight
ORF	Open Reading Frame
OMV	Outer membrane vesicles
p-NPPC	*p*-nitrophenol phosphorylcholine
PA01	*Pseudomonas aeruginosa*
PaMip	*Pseudomonas aeruginosa* Macrophage Infectivity Potentiator

PAMP	Pathogen-Associated Molecular Pattern
PCH	Pyochelin
PCR	Polymerase Chain Reaction
PGPR	Plant Growth Promoting Rhizobacteria
PlcA	Phospholipase C
Primer 'F'	Forward
Primer 'R'	Reverse
Pro	Proline
PPIases	Peptidlyl-prolyl *cis-trans* isomerases
PpmA	Parvulin proteinase maturation protein A
ppk	Pseudomonas protein kinase
ppp	Pseudomonas protein phosphatase
Pyr	Pyridine
PVD	Pyoverdine
QS	Quorum Sensing
RetS	Regulator of Exopolysaccharide and Type Three Secretion
Rpm	Revolutions per minute
SEM	Scanning electron microscopy
Sfa	Sigma activator factor
SlrA	Streptococcal rotamase A
Stk	Serine/threonine kinase
Stp	Serine/threonine phosphatase
StMip	*Salmonella typhimurium* Macrophage Infectivity Potentiator
'T'	Thymine
Tet	Tetracycline
TF	Trigger Factor
T6SS	Type Six secretion System
T3SS	Type Three Secretion System
TcMip	*Trypanosoma cruzi* Macrophage Infectivity Potentiator
T3E	Type Three Effector protein
UTI	Urinary Tract Infection
UV	Ultraviolet

VgrG	Valine-glycine repeats protein
WHO	World Health Organisation
WT	Wild Type
‘Y’	Tyrosine
µl	Microlitre
µg	Microgram

Mutant table key

Genes deleted in each mutant strain that was created in assessing Mip-like gene candidates which are discussed in Chapter 1. All have been assigned a letter for reference. Some of the possible combinations were not possible to make and are thought to be deleterious to the organism suggesting synthetic lethality, these combinations can be seen without a letter assignment. This table can be referred to throughout reading this thesis as a reminder to which genes are deleted in a given strain. This same table can be found in Chapter 3, Table 3.1.

Strain alphabetical assignment	Previous genotype annotation	New genotype annotation
D	PA01Δ*pa4558*	PA01ΔPaMip1
E	PA01Δ*pa4572*	PA01ΔPaMip2
F	PA01Δ*pa5254*	PA01ΔPaMip3
I	PA01Δ*pa4558* / Δ*pa4572*	PA01Δ*pamip1* / Δ*pamip2*
J	PA01Δ*pa4572* / Δ*pa5254*	Pa01Δ*pamip2* / Δ*pamip3*

Following phenotypic assays, Mip-like gene candidates assessed throughout Chapter 3 and listed in the above table were reassigned a new nomenclature displayed in this table which can be found in Chapter 3, Table 3.4. Table displaying PA01 mutant strains with their assigned letter for reference depending on what gene(s) are knocked out. Initially gene(s) knocked out in each strain were referred to as their chromosomal gene number (central column), however have now been replaced with an annotation in the far-right hand column. This table can be referred to throughout reading this thesis as a reminder to which genes are deleted in a given strain.

Table of contents

Index of Figures and Tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 1:

Introduction

Chapter 1. Introduction

1.1 Introducing bacterial diversity

Prokaryotes are a group of organisms that are distinct from eukaryotes that live in, on and all around us. We interact with them daily, and for the most part these interactions are harmless and often beneficial, but, there are some instances where these interactions have unfavourable consequences. There are two distinct domains within the prokaryotes, bacteria, and archaea. Organisms in these domains lack a nucleus and several other cell components that are present in eukaryotes. Although bacteria and archaea are similar in size and shape, archaea are grouped separately from bacteria as they possess genes and metabolic pathways more similar to those of eukaryotes, most notably for transcription and translation [2]. Furthermore, unlike bacteria, no pathogenic archaea have yet been identified. Within the bacterial domain there is enormous diversity, bacterial taxonomy provides a consistent means to classify organisms, which in turn helps us understand how known or novel bacteria may be of value to us agriculturally, in the production of new molecules or whether they offer health benefits [3]. By contrast, their identification helps elucidate potential clinical or agricultural threats. How bacteria are identified and classified has changed through the years since their initial discovery as 'animalcules' in 1676.

1.1.1 Taxonomically identifying bacteria using physical and biochemical methods

Bacteria today are taxonomically classified using a combination of physical characteristics and their genetics. Physical characteristics have been used for many years and are useful for the initial classification of a bacterium. These characteristics include their morphological appearance and biochemical tests [4]. An important classification method that was developed back in 1884 is still used today in assessing whether a bacterium is gram-positive or gram-negative [5]. Gram-positive bacteria possess a thick cell wall comprised of peptidoglycan layers (sugars and amino acids meshed), whereas gram-negative bacteria are slightly more complex, consisting of a thin peptidoglycan layer and an outer membrane external to this (Figure 1.1.a) [6]. The importance of the gram-negative structural complexities and its role in antibiotic resistance will be discussed later. To elucidate whether a bacterium is gram-positive or gram-negative, a gram staining test has been developed, involving the addition of crystal violet, iodine, alcohol and safranin. Gram-positive bacteria will retain crystal violet in their thick peptidoglycan cell wall, giving cells a purple appearance. Gram-negative bacteria will not retain crystal violet after the iodine and alcohol wash retaining only safranin and this gives cells a red appearance (Figure 1.1.b) [6]. Gram-positive and gram-negative bacteria can be further divided using their distinct morphologies including spherical (cocci), rod (bacilli), spiral (spirilla), comma (vibrios) or

corkscrew (spirochaetes). Gram-positive bacteria morphologies are typically made up of cocci and bacilli, whereas gram-negative bacteria possess all morphology types (Figure 1.1.c).

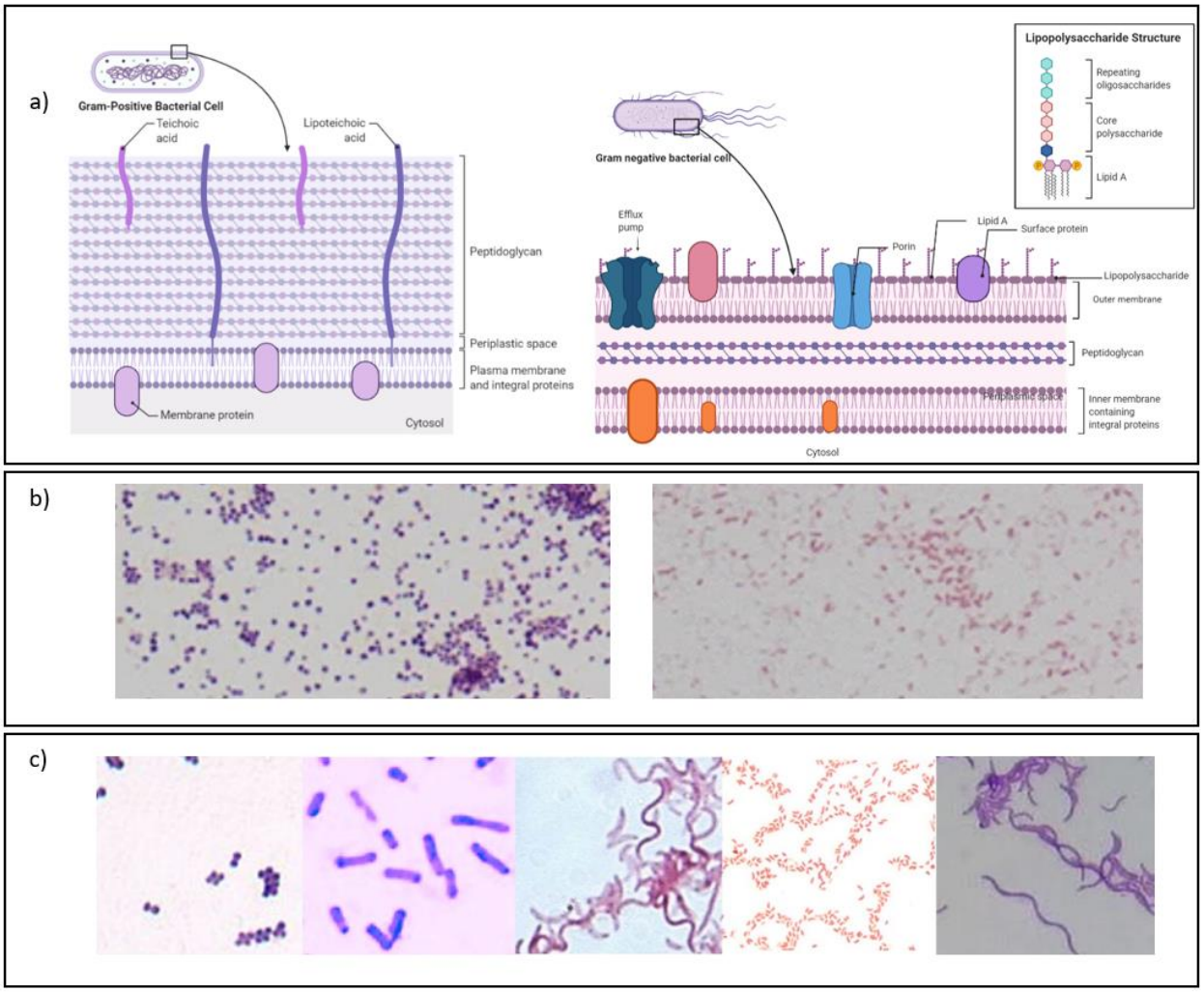


Figure 1.1. Bacterial cell structures. a) Diagram showing the cell wall structures of gram-positive (left) and gram-negative (right) bacteria. b) Images of gram-positive and gram-negative bacteria following the typical gram-staining procedure. On the left is a typical gram-positive cocci, on the right is a typical gram-negative bacillus [7]. c) Images of the differing morphologies typically possessed by bacteria. From image left to right shows; Cocci, Baccili, Spirilla, Vibrios and Spirochaetes

Physical differences can only tell us so much about a bacterium, and to confirm a bacterium down to the Genus level, further biochemical tests can be used. This includes catalase tests (assess whether the organisms can produce catalases), oxidase tests (identifies whether a bacterium produces cytochrome c oxidase) and blood agar tests (assess whether a bacterium is able to produce haemolysins) (Figure 1.2). However, these biochemical tests have only been shown to account for 5-20% of the genetic potential of a bacterium, meaning major differences may go undetected [8]. The ideal means to truly identify bacteria today is by applying a phylogenetic approach, comparing the total DNA of a bacterium to those of other bacteria. Using this approach can tell us whether they are part of the same Domain (e.g. Animalia / Bacteria), Phylum (e.g. Chordata / Pseudomonadota), Class (e.g. Mammalia / Gammaproteobacteria), Order (e.g. Primates / Pseudomonodales), Family (e.g. Hominidae / Pseudomonadaceae), Genus (e.g. *Homo* / *Pseudomonas*), Species (e.g. *H. sapiens* / *P. aeruginosa*). Often morphological and biochemical tests can narrow down a bacterium to a Family or Genus level, but identifying a bacterium down to a Species level is difficult using these tools alone.

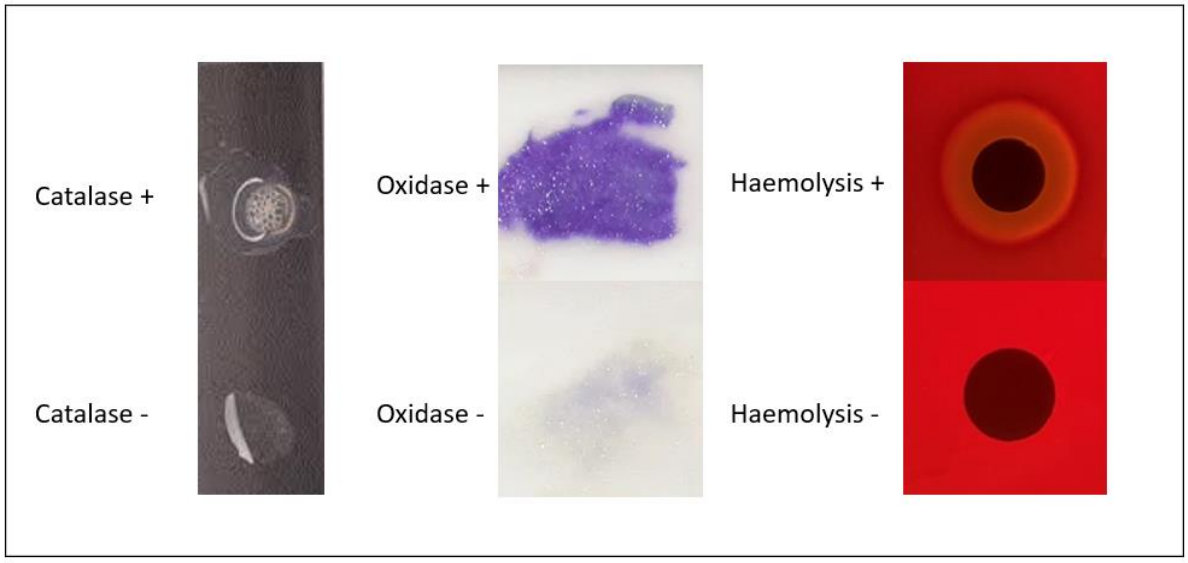


Figure 1.2. Biochemical tests used to help identify a bacterium. The left image shows the catalase test which looks for the presence of the catalase enzyme produced by a bacterium. If the bacterium in question is able to produce the enzyme, it breaks down hydrogen peroxide giving off the bubbles. If it does not produce the enzyme, no bubbles are formed. The middle image shows an oxidase test which assess whether a bacterium is able to produce the cytochrome c oxidase enzyme. The final image shows whether a bacterium produces haemolysins that are able to break open red blood cells, then subsequently use the haemoglobin inside. There are three types of haemolysis, alpha, beta and gamma.

1.1.2 Importance of identifying and categorising bacterial species

So why is identifying and categorising bacteria so important? As mentioned previously, they are everywhere. Taking a clinical setting as an example, identifying a bacterium from a given sample is important to elucidate whether the bacterium is a commensal organism, non-pathogenic or pathogenic. Not only that, but deciphering the Genus and Species tells us more about its pathogenicity behaviours such as whether it is an intracellular (an organism that infects and replicates within host tissues) or extracellular (an organism that infects a host existing as a free-living pathogen in an environment) pathogen. Identifying the Genus tells us whether it is an intracellular or extracellular pathogen, giving us an idea of what broad-spectrum treatment pathway to follow. However, further identifying the pathogen down to a Species level will have an even greater impact on the treatment pathway and antibiotic regime prescribed, potentially differing to the treatment pathway for only identifying to a Genus level [9].

Difficulties in identifying and distinguishing a Species responsible for an infection by applying only morphological and biochemical tests can be seen for the *Pseudomonas* Genus. Typically, *Pseudomonas aeruginosa* accounts for more than 80% of *Pseudomonas* infections, however there are cases where a *Pseudomonas putida* or *Pseudomonas fluorescens* strain is the cause of clinical infection [9]. The typical treatment of an infection caused by the Genus Pseudomonas is ceftazidime, a cephalosporin antibiotic, with the assumption the Species will be *P. aeruginosa*. However, the preferential and most effective treatment for a *P. putida* infection is arbekacin (an aminoglycoside) [10]. Here, using morphological and biochemical tests alone are not always accurate in elucidating which *Pseudomonas* Species is the cause of infection. Using DNA sequencing, coupled with a proteomic approach using an instrument with a MALDI-TOF detector, would provide the correct identification down to the *Pseudomonas* Species level and aid in the correct choice of treatment for the infection (Figure 1.3) [11]. To further complicate matters, even within a Species there can be genetic variation, this is termed a strain. In the single Species *Pseudomonas aeruginosa,* there are several clinically relevant strains that possess antibiotic resistance genes that other strains do not. In these instances, strains are often sent to clinical reference laboratories where their genomes are fully sequenced. As previously mentioned, this is the ideal means to identify a bacterium of interest.

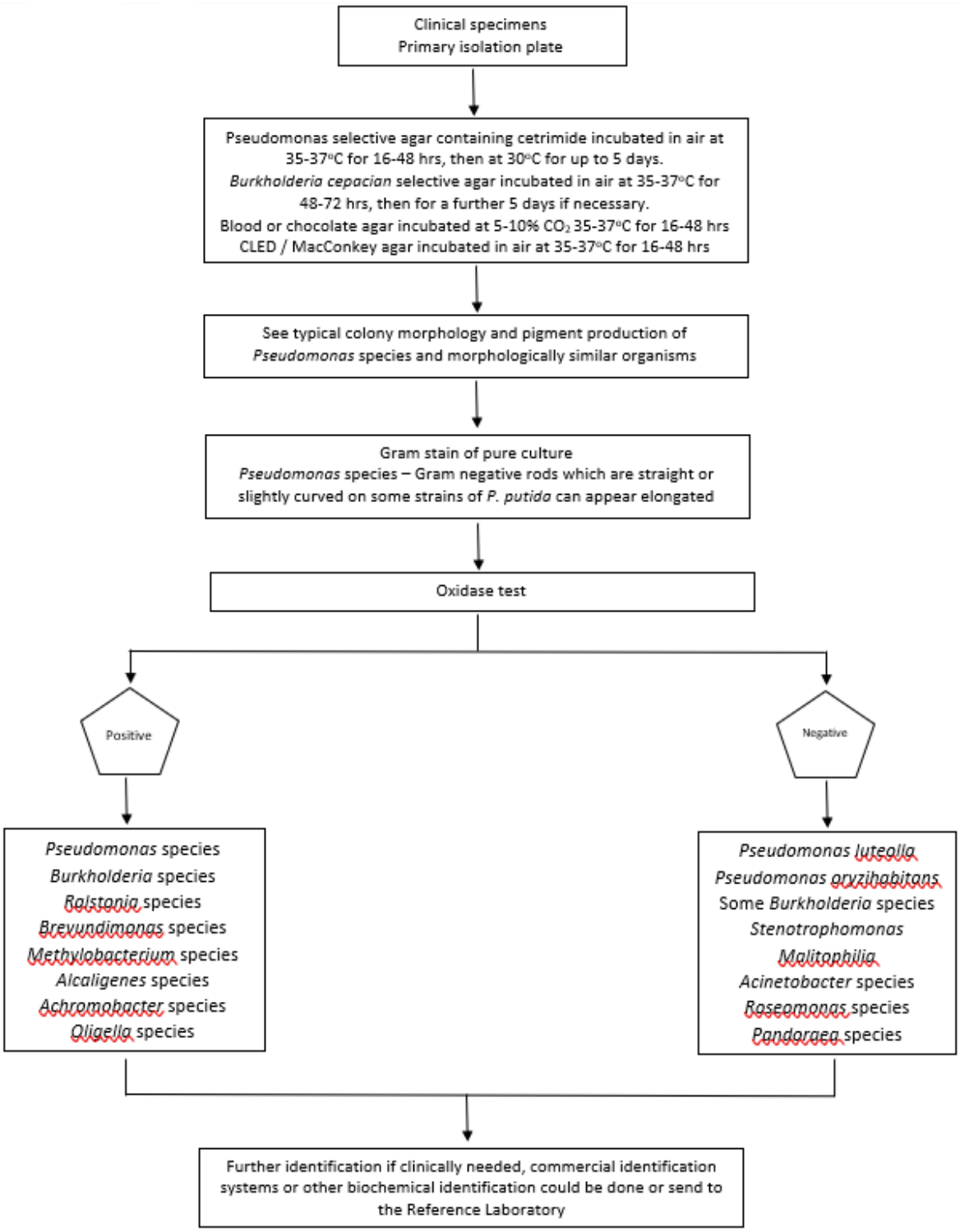


Figure 1.3. Flow chart for identifying microbial infections. Flow chart taken from the UK Government Public Health England UK Standards for Microbial Investigations. The flow chart describes the process of identifying a bacterium and cause of infection using morphological and biochemical tests, with these tests narrowing down to a Genus level. Once at a Genus level, the guidance suggests to further identify the cause of infection down to a Species level using other means. These other methods include DNA sequencing by a reference laboratory, or other inhouse testing such as MADLI-TOF identification.

Given their extraordinarily ubiquitous nature, identifying bacteria helps us understand how they may be of value to us or, whether they pose a potential threat. Considering the diversity that lies within even a single Species, it is no surprise that within a Genus, there lies even greater diversity. As briefly touched upon, there are several Species within the *Pseudomonas* Genus that pose significant health threats to humans, one of which, *Pseudomonas aeruginosa* (PA01), will be discussed in more detail later. However, within the *Pseudomonas* Genus there are also several Species that are incredibly useful to humans.

1.2 The *Pseudomonas* Genus and its relationship to humans

The *Pseudomonas* Genus belongs to the Class Gammaproteobacteria that was first described in 1894 and is currently known as one of the most complex in the bacterial domain, housing the largest number of annotated Species [12, 13]. This Genus, and the Species it comprises, has constantly been revised over the past few decades due to the continual improvement in taxonomic approaches, with Species frequently being distinguished from one another. Initially, the gold standard at distinguishing a Species was to look at a single phylogenetic marker, the 16S rRNA gene, and assess how conserved the sequence was between Species. However, through the years it has been evidenced this method is inadequate when the Species are closely related. This initial approach, coupled with using similarities in morphologies, led to categorising several bacterial isolates as the same Species, in some instances as the same Genus, when in fact they were vastly different. The most notable reclassification was in 1990 for *Pseudomonas cepacia*, now known as *Burkholderia cepacia*. Taxonomists proved using DNA-DNA homology values that *P. cepacia* was so vastly different from the *Pseudomonas* Genus, it should not only be reclassified as its own Genus, but also it belonged to an entirely different Class known as a Betaproteobacteria [14].

To establish different Species within the *Pseudomonas* Genus, it is now standard practice to use several genes as phylogenetic markers in DNA-DNA taxonomic studies; this is also known as multilocus sequence analysis (MLSA) [15, 16]. These marker genes are housekeeping genes conserved across the *Pseudomonas* Genus such as 16S rRNA gene, *gyrB*, *rpoB*, *rpoD, recA* and *atpD*. Thanks to Next-generation sequencing (NGS) there has been an exponential increase in the number of available whole genome sequences, allowing a more comprehensive comparative analysis of these genes and entire genomes. This is only a brief snapshot into the complexities of taxonomy and how ambiguous the previous classification system was at identifying new Species.

The *Pseudomonas* Genus is an excellent example of how selective pressures result in allelic differences in common marker genes, predominantly involved in regulation and metabolic pathways. These differences have aided in their adaption and colonisation of different environments, while also allowing each Species within a Genus to be taxonomically separated [12]. The wide variety of metabolic pathways shown by Species within the *Pseudomonas* Genus has allowed them to adapt and survive in a variety of niches [17]. This Genus can be found anywhere from water and soil to our own gastrointestinal tract. Depending on the environment, different Species of the Genus can be found, each with their own biological importance [18]. The most well studied Species include the plant pathogen *Pseudomonas syringae*, the soil dwelling *Pseudomonas putida*, the plant-growth promoting *Pseudomonas fluorescens* and finally the human pathogen, *Pseudomonas aeruginosa*, [19]. Although all are members of the same Genus, each Species has their own biological significance to humans which will now be briefly discussed.

1.2.1 *Pseudomonas syringae*

First isolated from the lilac tree *Syringa vulgaris*, *P. syringae* is a gram-negative plant pathogen of the Genus *Pseudomonas*. It has been reported to infect a wide range of plant species with over 50 pathovars (plant pathogen of the same Species, however labelled a different strain depending on the plant species they infect) [20]. They are of agricultural significance as they cause disease in several annual crops such as bacterial speck on tomato plants and halo blight on beans, and this results in significant crop yield losses. Most recently this plant pathogen has caused an endemic of bacterial cankers in kiwifruits resulting in a global shortage of kiwis [21]. This pathogen has two stages of growth, the epiphytic phase, growth of the bacteria on the outside and surface of plant tissues, and the endophytic phase that begins once the bacteria has colonised the plant tissues via the apoplast intracellular space. The disease only begins to set in once *P. syringae* has entered the plant tissue and begins to multiply in the endophytic phase.

Once inside plant tissues, it has been shown that this Species possess a gene cluster termed the hypersensitive response and pathogenicity (hrp) cluster, encoding a secretion system known as the **T**ype **III S**ecretion **S**ystem (T3SS) responsible for delivering a variety of T3SS bacterial effector proteins (T3Es) to the hosts tissues [22]. These T3Es have shaped *P. syringae* broad host ranges with some T3Es remaining conserved across all pathovars acting broadly. However, others are exclusive to different pathovars, allowing specialisation to a specific host. These effectors have various biological roles, and most importantly they have been shown to suppress pathogen-associated molecular pattern (PAMP) triggered immunity, the first line of defence in plants [23]. Without this PAMP immunity, the plant does not recognise an infection, therefore host defences such as stomatal closure are not activated to achieve disease resistance by, preventing further infection and damage. This is only one of many mechanisms by which *P. syringae* can interfere

with plant defensive pathways, with a plethora of other tools within its genome at its disposable making it a difficult pathogen to control. Internationally, as a result of *P. syringae* infections, crop yield losses have been reported from 5-50% through the years, having a significant economic impact on several countries [24].

1.2.2 *Pseudomonas putida*

P. putida is an interesting Species within the Pseudomonas Genus. It possesses many traits for use in diverse areas of industry, however in some circumstances it also poses a clinical threat. As with all members of the *Pseudomonas* Genus it is gram-negative, it is predominately found in soil and is saprotrophic (able to extracellularly digest decaying organic matter). It has been demonstrated as a potential biocontrol candidate acting as an antagonist against common plant pathogens from the Genera *Pythium* and *Fusarium. Pythium* Species are the most common cause of root-rot disease, which has been estimated to result in loses of roughly $715 per hectare per year [25]. Equally, *Fusarium* Species commonly infects barely crops, the fourth most grown crop globally, resulting in almost $3 billion in losses annually, so new approaches to control these diseases is essential. With the metabolic diversity of Pseudomonads, in particular *P. putida*, it has been shown they produce a variety of antimicrobial compounds that inhibit the growth of plant pathogens [26, 27]. Furthermore, this metabolic diversity has allowed a commensal relationship to develop between *P. putida* and plants, making it a Plant Growth Promoting Rhizobacteria (PGPR). Plants feed *P. putida* in the rhizosphere via root exudates which results in *P. putida* producing many plant hormone precursors required for plant growth [28]. With this rhizosphere colonisation, *P. putida* acts as a further antagonist to plant pathogens by outcompeting them for essential resources such as iron, indirectly acting as a biocontrol agent. This iron acquisition trait is widespread in the genomes of the *Pseudomonas* Genus. The genomes of nearly all *Pseudomonas* Species house genes capable of producing a wide range of siderophores, a group of small molecules allowing an organism to solubilise and sequester iron from an environment (Figure4) [29]. Given these properties, research is ongoing as to the potential use of *P. putida* in controlling several crop destroying pathogens.

Not only does the metabolic diversity of *P. putida* aid in the protection of plants against plant pathogens and the growth of plants, but it has also been shown to have potential in bioremediation and biodegradation [30]. These are processes that break down and detoxify soil contaminants as well as complex materials such as plastics. Currently there is a growing issue globally of plastic pollution. Plastics are synthetic or semi-synthetic polymerized products currently used in all industries with a variety of uses. However, most uses of plastics revolve around packaging and single use, resulting in plastics ending up in landfill (79%) if they are not recycled (9%). Plastics do not decompose and when left in landfill or generally dumped, they

release harmful chemicals into environments such as soil and water. Recently, it has been shown that several microorganisms are capable of degrading these synthetic plastic polymers on a small scale, potentially offering a solution to the plastic pollution problem. *P. putida* has been characterised with this plastic degrading ability. In a study it was shown a garden soil sample housed a strain of *P. putida* that was able to metabolise and convert complex plastic material; this was determined by weight loss of the plastic. Over the course of one month, *P. putida* was able to reduce the weight of plastic items including polythene bags, plastic bags, plastic cups and milk bottles by 75.3% [31]. How *P. putida* does this remains largely unknown, although recent studies have begun to elucidate key enzymes involved in the process, paving the way for future technologies for plastic degradation [32].

Not only does *P. putida* have useful applications agriculturally and environmentally, but it is also valuable in a lab and industrial setting. As with all *Pseudomonas* species, they are highly promiscuous. This promiscuity allows the frequent uptake of plasmids from other bacteria in their environment through horizontal gene transfer (movement of genetic material from one cell to another), this trait has contributed towards their ability in colonising a variety of niche environments [33]. Furthermore, due to their excellent genetic tractability (easy to genetically modify), robust metabolic pathways, and broad tolerance to toxic compounds, they are useful tools in research as 'cell factories'[34]. Using *P. putida* as a model is a popular choice in research, it can be transformed and genetically engineered to overexpress and produce a wide range of natural products that it ordinarily would not produce. This has allowed the overproduction of proteins and products of interest to further purify and use in studies [35].

Finally, *P. putida* has its role as a human pathogen. To a healthy individual infection caused by *P. putida* is rarely seen. Although not as virulent as *P. aeruginosa*, *P. putida* can still present as an uncommon opportunistic pathogen predominantly presenting as urinary tract and catheter-related infections. However in neonatal, neutropenic or immunocompromised patients it can cause bacteraemia (bacteria presence in the bloodstream) and sepsis (severely life-threatening infection where bacterial toxins are produced) [36]. Mortality rates due to *P. putida* infections can range anywhere between 8-40%, but appropriate antibiotic therapy can result in favourable clinical outcomes. Unlike *P. aeruginosa*, *P. putida* strains do not possess quite as many intrinsic antibiotic resistance mechanisms and are usually susceptible to antibiotic regimes using carbapenems, fluoroquinolones and aminoglycosides [37]. However, in recent years there has been an increase of antibiotic resistant *P. putida* isolates, particularly towards carbapenems with a 20% increase from 2005 to 2011 [38].

1.2.3 *Pseudomonas fluorescens*

Like *P. putida*, *P. fluorescens* has been proven as a useful biocontrol agent and acts as a PGPR contributing towards plant health and growth by suppressing and reducing the severity of several plant fungal diseases. In the case of the fungal pathogen *Fusarium oxysporum, P. fluorescens* is able to induce plant systemic resistance against the fungus [39, 40]. In these studies it was demonstrated that *P. fluorescens* essentially primed the plant for an infection, so when the plant was challenged by an actual pathogen, the plant defensive systems are preconditioned to resist the *F. oxysporum* fungus [19, 39]. Furthermore, *P. fluorescens* can outcompete potential plant pathogens in the rhizosphere by sequestering iron by producing siderophores similar to those of *P. putida*. It was also shown that mutant strains that do not produce siderophores were not able to grow as well as producer strains [41].

Similar to how *P. putida* acts as an antagonist, *P. fluorescens* is able to produce antimicrobial molecules such as phenazine-type antibiotics as well as hydrogen cyanide [42]. It was shown a mutant strain of *P. fluorescens* unable to produce phenazine-1-carboxylate was significantly attenuated in the control of take-all infection of wheat caused by *Gaeumannomyces graminis var. tritici.* compared to the WT phenazine-1-carboxylate producing strain [43, 44]. A similar study was undertaken to look at the effects of hydrogen cyanide production by *P. fluorescens* in the control of black root rot caused by the fungal pathogen *Thielaviopsis basicola. P. fluorescens* mutants unable to produce hydrogen cyanide were less effective at protecting tobacco plants from the pathogen compared to strains that could [45]. Although *P. fluorescens* and *P. putida* are similar in how they help protect plants, *P. putida* is not able to produce either phenazine-1-carboxylate or hydrogen cyanide, with both species producing different antimicrobial molecules and secondary metabolites highlighting just how genetically diverse the *Pseudomonas* Genus is.

As previously mentioned, *P. putida* is able to infect humans, although the occurrence is rare. *P. fluorescens* like *P. putida*, is also able to cause infections but to an even rarer occurrence and through the years has been of scarce clinical significance. Ordinarily, *P. fluorescens* is a commensal organism occurring at low levels in the gastrointestinal tract. There have been instances whereby *P. putida* was reported as the causative pathogen of urinary tract infections (UTIs) and several other non-life threatening acute opportunistic infections. [46]. However, over the past few decades there has been an increase in nosocomial infections (an infection that takes place within 48 h of hospital admission) [47]. An incident in the USA led to four cases of *P. fluorescens* bacteraemia, the first recorded cases of bacteraemia caused by *P. fluorescens*. These incidents arose due to contamination of heparinized saline intravenous flush syringes with *P. fluorescens*. These syringes were used in immunocompromised cancer patients, leading to bacteraemia. This highlights just how opportunistic the *Pseudomonas* Genus can be [48]. This

suggests the *Pseudomonas* Genus and the Species it comprises is continually evolving and adapting to colonise new niches, something that is of great clinical concern.

Thus far, each Species within the *Pseudomonas* Genus described above has been demonstrated to have a variety of biological uses while still posing a threat to humans. Even within the same Species all biological uses and threats have been evidenced; from acting as cell factories in industry for the production of useful materials, to helping protect crops in agriculture, while also, given the right opportunity, posing a threat to human health. This highlights just how adaptable this Genus and the Species it encompasses really is. They adapt to survive in an environment they are thrown into, be it soil, water, or human fluids. Understanding how different mechanisms and metabolic pathways allow *Pseudomonas* Species to adapt to such diverse environments, as well as their regulation, will help further elucidate how a Species may be of use to us today in solving the serious global problems we face such as pollution and sustainable food production for an ever-growing global population. However, the biggest global problem we are currently facing in the world of Microbiology is antibiotic resistance, something which the next *Pseudomonas* species to be discussed plays a huge role in.

1.3. *Pseudomonas aeruginosa* – a global threat

Pseudomonas aeruginosa, a common gram-negative, aerobic bacilli pathogen of both plants and humans, is currently listed as a priority one ESKAPE pathogen (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* Species) by the World Health Organisation (WHO) that requires further research and the development of new antibiotics. It possesses a large, Guanine and Cytosine (G+C) rich genome (67% GC content), that is around 6.3 million base pairs in size. Its genome has been fully sequenced and annotated as housing incredible versatility that is responsible for many antibiotic resistance mechanisms including; intrinsic resistance (a bacteria naturally resistant to an antibiotic without the need for mutation or gain of further genes), acquired resistance (the result of an evolutionary process and selection pressures resulting in the mutation of drug targets) and adaptive resistance (resistance occurring due to environmental circumstances that results in transcriptional changes of genes involved in complex lifestyle states and is often reversible) [49-51].

Agriculturally, *P. aeruginosa* is a plant pathogen of low importance as it does not result in quite as severe crop yield losses as the previously described *Pseudomonas* species. It primarily infects higher plants such as *Arabidopsis thaliana* and *Lactuca sativa* (lettuce), while also playing an important role in bioremediation through similar processes described above for other *Pseudomonas* Species [52].The primary focus of research into *P. aeruginosa* has been looking in

depth at its pathogenicity towards humans and the plethora of systems it utilises to cause life threatening infections. *P. aeruginosa* is what is known as an opportunistic pathogen. Pathogens of this description would ordinarily not be able to mount a serious infection in a healthy host, however in patients who are immunocompromised such as cancer patients, or patients that suffer from cystic fibrosis, these infections pose a serious health risk.

As previously mentioned, *P. aeruginosa* has two lifestyle stages, sessile and free roaming, each is associated with the expression of different structures, metabolic pathways and virulence mechanisms, all of which are tightly controlled by several intertwined, complex regulation systems. Given over 20% of *P. aeruginosa* genome is dedicated to regulation systems, it is no surprise this Species has evolved to colonise a variety of niches orchestrated by the expression of different genes involved in lifestyle switches, metabolic pathways and virulence mechanisms optimal to its current environment. Gaining a better understanding on how these regulation systems are orchestrating the lifestyle transitions will improve our knowledge of the pathogenesis of *P. aeruginosa*. With these lifestyle transitions comes the expression of different cellular structures, activated metabolic pathways and the expression of different virulence mechanisms. Elucidating essential genes and their products involved in the expression of any virulence determinants will help provide potential drug targets. If, for example, a single protein that is essential in the regulatory system cascade that orchestrates the expression of a given virulence mechanism, can be inhibited, the overall virulence of *P. aeruginosa* can be reduced resulting in improved patient outcomes.

In the coming sections I will briefly discuss the following: important cellular structures involved in the intrinsic, acquired and adaptive antibiotic resistance mechanisms of *P. aeruginosa,* as well as structures that contribute towards the virulence processes of *P. aeruginosa.* I will further discuss how different regulation systems promote lifestyle transitions, and with this, changes in the expression of genes involved in several virulence mechanisms along with metabolic pathways involved in antibiotic resistance. Finally I will consider how these different lifestyles, sessile and free roaming, promote different infectious processes, chronic and acute infections, respectively.

1.4 Structural Components of *P. aeruginosa* involved in antibiotic resistance and virulence

1.4.1 Cell envelope, outer membrane and lipopolysaccharide (LPS)

As briefly described before, as a gram-negative pathogen, the cell envelope of *P. aeruginosa* plays an important role in its intrinsic antibiotic resistance and consists of an inner membrane, an outer membrane and a peptidoglycan layer sat in the periplasmic space between the inner and outer membranes. The outer membrane is composed of several proteins, phospholipids, and lipopolysaccharide (LPS). The structure of the outer membrane is what contributes towards *P. aeruginosa* intrinsic resistance to many antibiotics as it provides low membrane permeability. The LPS structure is comprised of three components, lipid A, core oligo and an O antigen polysaccharide chain extending out into the extracellular environment (Figure 1.4). This structure has been implicated in a variety of essential virulence determinants such as; interacting with host cell receptors, involved in inhibiting host defence systems, influencing biofilm formation and mediating resistance to some antimicrobials [53].

1.4.2 Porins

Another important structural component of the outer membrane are structures called porins, protein β-barrels made of antiparallel β-pleated sheets. There are many different families of porins with *P. aeruginosa* possessing at least 26. The general purpose of most porin families is fairly non-specific (e.g. OprD) and allows the passive movement and diffusion of simple sugars, amino acids and small peptides into the cell [54]. This non-specific movement has been a means for β-lactam antibiotics (antibiotics that possess β-rings in their structures and usually inhibit cell wall synthesis), such as meropenem, to enter the cell. However, mutations in these genes has been linked to antibiotic resistance, with mutations in porins also resulting in carbapenem resistance [55, 56]. Some mutations result in a conformational change in the porin OprD which prevents entry of either type of antibiotic into the cell (Figure 1.4). Other mutations have led some *P. aeruginosa* strains to either replace the non-specific porin families with porins that have affinity to specific useful molecules, or in some cases the loss in expression of a porin such as OprD altogether [57].

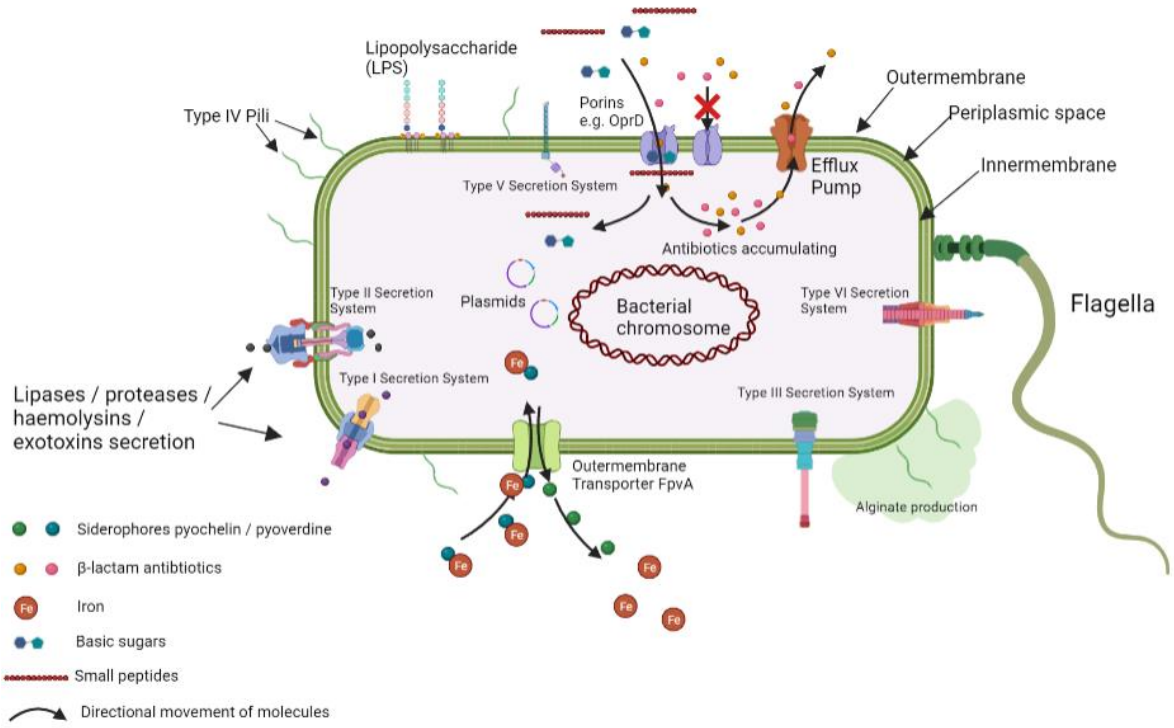


Figure 1.4. Detailed diagram of *P. aeruginosa* cell showing different structural components. The characteristic gram-negative cell envelope consisting of the outer membrane, periplasmic space containing a thin peptidoglycan layer and inner membrane can be seen. There are several virulence systems spanning the cell envelope all of which are responsible for delivering effectors and toxins into the environment or directly into a host cell. Structures responsible for movement includes the tail-like appendage termed the flagella and shorter hair like structures termed the type IV pili. Outer membrane transporters are shown expelling molecules such as siderophores into the environment to sequester iron, an essential nutrient. Structures termed porins (e.g. OprD) responsible for the passage of small molecules are shown spanning the cell envelope. Small molecules such as simple sugars and short peptides diffuse through these porin channels into the cell. Some antibiotics use these porin channels to enter the cell. Mutations in these porins (denoted by the red cross) result in the loss of porins that allow antibiotics into the cell, and conformational changes of the non-specific porins resulting in selection of molecules allowed into the cell both of which lead to antibiotic resistance. An intrinsic resistance mechanism of *P. aeruginosa* is the expression of efflux pumps. These pumps selectively pump out harmful molecules from inside the cell into the environment. Antibiotics enter via a porin and begin to accumulate in the cytoplasm. The efflux pumps, using an energy source, pump these antibiotics back out. The overexpression of genes responsible for efflux pumps leads to several acquired and adaptive resistance mechanisms to several antibiotics.

1.4.3 Efflux Pumps

Efflux pumps are structures part of the innate antibiotic resistance of *P. aeruginosa* that span the entire cell envelope and are a great contributor to antibiotic resistance. They are essentially large protein macromolecules that are active transporters (require an energy source to work) able to move unwanted toxic molecules from the cytoplasm of a cell back out into the environment (Figure 1.4). There are six main families of efflux pumps, categorised by their structural component similarities, the energy source used by the pump and the types of molecules that the pump exports. These efflux pumps are multi-protein complexes that require energy to pump solutes out of a cell, allowing regulation of their internal environment. *P. aeruginosa* possess efflux pumps conferring intrinsic resistance to several antibiotics including penicillins and carbapenems [58]. The most important and constitutively expressed efflux pump system is MexAB-OprM, which confers resistance to a broad range of β-Lactam antibiotics, with nearly all clinical isolates possessing genes for this pump [59]. Several pumps have been shown to become constitutively overexpressed as a result of mutations in genes that negatively regulate their expression. This overexpression has resulted in the acquired resistance of fourth generation cephalosporins and some aminoglycosides. A further three efflux pump systems have been studied, MexXY/OprM(OprA), MexCD-OprJ, and MexEF-OprN, each able to pump out different antibiotics, further increasing the antibiotic resistance abilities of *P. aeruginosa* [57].

1.4.4 Flagella

An important structure key to the movement of *P. aeruginosa* within a host is the flagellum. The flagellum is a polar hair-like structure (Figure 1.4) that can occur singularly, in pairs (termed twinning) or in short chains that rotate in either a clockwise or anticlockwise propeller-like manor, requiring energy in the form of adenosine triphosphate (ATP) to do so [60]. The building of flagella is complex, often involving around 50 genes, and is essential to the survival of a bacterium to disseminate within a host from areas of low nutrients, to high. Phenotypically, the presence of flagella is required for different types of movement of *P. aeruginosa*, swarming and swimming. Swarming is defined as multicellular, flagella-mediated surface migration of bacteria that also requires surface surfactants and an increase in the number of flagella expressed, often associated with solid surfaces such as epithelial cells [61]. Swimming is the movement of *P. aeruginosa* cells, often as individual cells rather than multi-cellular, through a liquid such as mucus or blood. In studies, mutants lacking the genes required for the assembly of a flagellum were less invasive in a mouse infection model [62]. Flagella are expressed in the lifestyle known as the free-roaming, which is associated with acute infections, which will be shortly be discussed in more detail.

1.4.5 Type IV Pili

With flagella responsible for the free-roaming lifestyle of *P. aeruginosa*, the expression of structures termed type IV pili are associated with the sessile lifestyle switch *P. aeruginosa* can transition to. This lifestyle switch is associated with chronic infections. Type IV pili are assembled using thousands of single pilin subunits, and once formed they are involved in several processes including a form of motility termed twitching, host cell adhesion, microcolony and biofilm formation [63]. The twitching movement is required by *P. aeruginosa* to navigate a semi-solid surface such as mucosal epithelial cells. This movement involves the extension and retraction of the pili, dragging the cells along a surface. The initial attachment of *P. aeruginosa* to a surface is the beginning of the lifestyle transition from free-living to sessile. It has been demonstrated that *P. aeruginosa* mutant strains lacking type IV pili were unable to successfully attach to a surface, which subsequently affected microcolony and biofilm formation [64].

1.5 *P. aeruginosa* lifestyle changes and how they are regulated

1.5.1 GacS / RetS two component system

As mentioned above, *P. aeruginosa* has two very distinct forms of living, the free-swimming lifestyle associated with acute infections, and the more sessile, stationary lifestyle associated with chronic infections (Figure6), these lifestyles and their importance will be discussed in more detail shortly. This transition between lifestyles, is mediated by complicated, intertwined regulatory systems that despite extensive research, are still not fully understood. These lifestyle changes are a consequence of environmental queues that spark a regulatory phosphorylation cascade known as the **G**lobal **A**ctivator of antibiotic and **C**yanide synthesis (Gac) two-component system, most commonly known as the GacS/RetS two-component system which results in the expression and repression of genes involved in contrasting lifestyles (Figure 1.5).

P. aeruginosa has over 60 two-component signal transduction systems, with the GacS/RetS being the most well studied following its initial discovery in *P. syringae* in 1992 [65]. This system gained special interest due to its role in mediating life-style transitions, which consequently affects pathogen-host interactions. The sessile lifestyle is of particular concern when Cystic Fibrosis (CF) patients are involved. This lifestyle is involved with biofilm formation, an adaptive resistance mechanism, whose population is highly resistant to antibiotics. The importance and relevance of biofilm formation to CF patients will be discussed in more detail in due course.

When in the free-living lifestyle associated with acute infections, a small regulatory RNA-binding protein, RsmA, is bound to the promoters of genes involved in flagella-dependant motility such as swimming and swarming, lipase production, rhamnolipid production and the effector

delivering **T**ype **III** **S**ecretion **S**ystem (T3SS), activating their transcription. At the same time, RsmA is also bound to the DNA coding region of a gene or binds to prevent translation initiation in the mRNA near to the RBS for genes assocaited with the sessile lifestyle and involved in chronic infections, preventing translation initiation (Figure 1.5). Autophosphorylation of the transmembrane sensor kinase GacS ignites the cascade system. The now phosphorylated GacS transfers a phosphate group to the global activator GacA. This phosphorylation activates GacA leading to the upregulation of genes encoding further small regulatory RNAs, RsmY and RsmZ. RsmY and RsmZ go on to bind RsmA, releasing RsmA bound to the promoters of genes it was previously activating transcription of that were involved in the planktonic, high motility lifestyle (leading to a reduction of transcription), while simultaneously releasing RsmA from the promoters of genes it was previously repressing involved in biofilm formation, pyocanin production, alginate production and the expression of a second effector delivery system, the **T**ype **S**ix **S**ecretion **S**ystem (T6SS). This system was shown to have a profound affect in an acute pneumonia infection model of mice. It was shown using that an RsmA mutant strain that the sessile lifestyle was favoured, with a reduction in initial the colonisation and infection of mice [66].

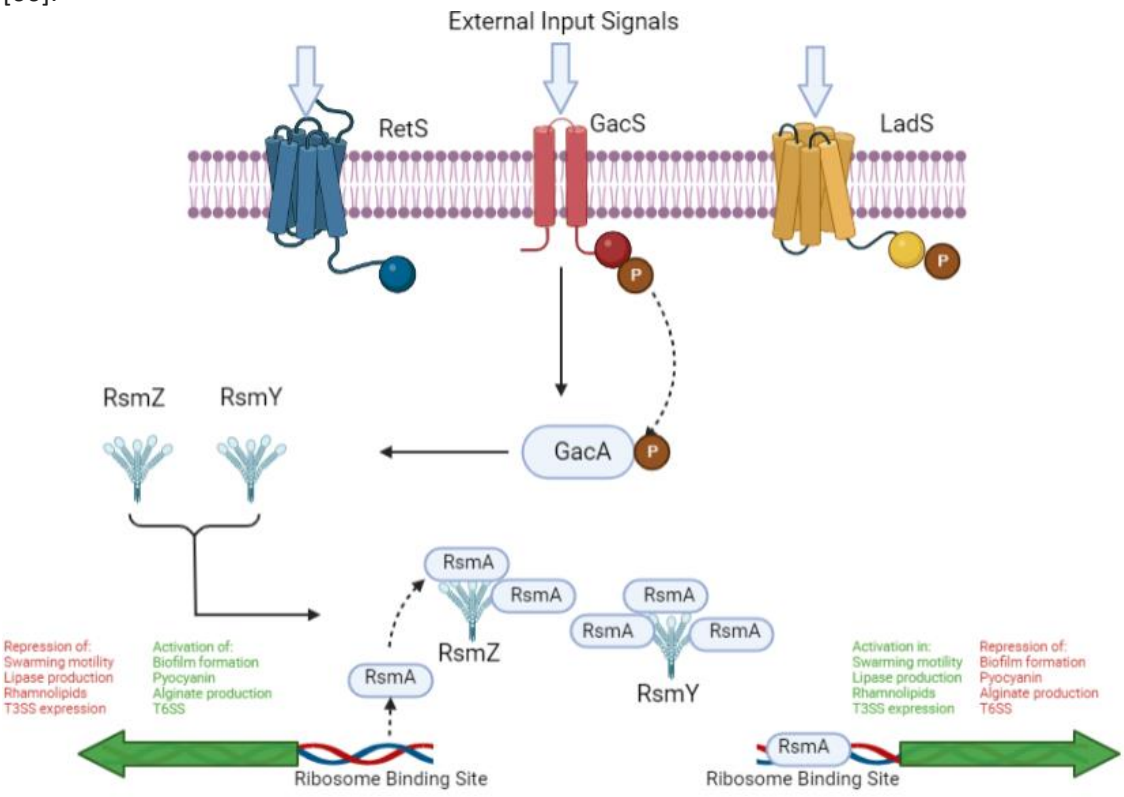


Figure 1.5. Diagram explaining the GacS / RetS two component system cascade. Upon phosphorylation GacS phosphorylates GacA which then causes the upregulation of RsmZ and RsmY. These two smalls regulatory RNAs then bind to RsmA which is bound to promoters of genes involved with the chronic infection lifestyle, repressing them. Upon the binding of RsmZ and RsmY to RsmA, RsmA is released from the promoters of these genes which are the allowed to be transcribed and upregulated.

This system is further complicated with the involvement of two additional sensor kinases, LadS **(L**ost **a**dherence **s**ensor) and RetS (**R**egulator of **e**xopolysaccharide and **T**ype III Secretion **S**ystem). LadS works in conjunction with GacS, acting as a positive regulator of GacA and genes in the pel operon that is heavily involved in biofilm formation, while repressing the expression of genes involved in the T3SS [67, 68]. RetS acts in an opposite manor to GacS and LadS by repressing genes involved in biofilm formation and promoting genes involved in acute infection processes. It was shown that RetS exerts its biofilm repressive affects by forming a heterodimer with GacS, preventing its autophosphorylation, consequently preventing the phosphorylation of GacA. Without phosphorylated GacA, RsmY and RsmZ are not upregulated, resulting in RsmA remaining bound to and repressing the promoters of genes involved in biofilm formation and other chronic infectious processes while also continuing to promote expression of genes involved in infectious processes [68].

1.5.2 Quorum sensing (QS) systems, *Las* and *Rhl*

Although responsible for the transition from planktonic to sessile lifestyles, and with-it different virulence mechanisms, there are many more complex systems at play that are interlinked with the GacS/ RetsS system. Two systems that coexist in a hierarchical relationship to one another, and which are partly controlled by the Gac/RetS system, are the Las and Rhl Quorum Sensing systems [69]. These two systems rely on cell signalling molecules known as *N*-**A**cyl **h**omoserine **l**actones (AHLs) [69]. These molecules have been evidenced to control and regulate gene expression in a cell-density dependant manor, known as **Q**uorum **S**ensing (QS) [70]. AHLs are generally comprised of fatty acids and are commonly synthesised by the LuxI family of proteins, which are subsequently sensed by LuxR proteins, a transcriptional regulator family. Once the threshold concentration of AHLs has been reached, a concentration that is correlated to a certain bacterial density, a complex is formed with LuxR transcriptional regulators. This complex is able to bind DNA, and subsequently allows binding to several promoters of virulence genes, altering their expression.

The Las and Rhl systems respond to two different AHLs, *N*-3-oxo-dodecanoyl homoserine lactone (3-oxo-C_{12}-HSL) and *N*-butanoyl lactone (C_4-HSL), respectively. The Las system responds to 3-oxo-C_{12}-HSL which is produced by the LasI synthase and recognised by the transcriptional regulator LasR. This system is involved in controlling the expression of multiple virulence factors involved in acute infection and promotes the planktonic lifestyle. Conversely, the Rhl system responds to the AHL C_4-HSL which is produced by the RhlI synthase and sensed by the transcriptional regulator RhlR. This transcriptional regulator responds by promoting the expression of genes associated with the transition towards the sessile lifestyle, particularly repressing genes involved in the T3SS assembly. These two systems exist as previously

mentioned in a hierarchical manor. The Las system controls the Rhl system and is preferentially active, much like how the planktonic lifestyle is more readily observed, with an environmental queue needed to transition lifestyles mediated by the GacS/RetS cascade. The 3-oxo-C_{12}-HSL-LasR complex is responsible for the upregulation of *rhlR* transcription, so the activation of the Las system allows the later activation of the Rhl system.

Regulatory mechanisms such as the above are vital to the overall virulence of *P. aeruginosa*, coordinating the expression of virulence factors at the correct time and place to enable optimal pathogenicity. It is no coincidence that the Las and Rhl systems act in an opposite manor, regulating genes depending on the lifestyle stage. It has been demonstrated that the GacS / RetS partially controls these QS systems in several ways [71]. Ordinarily the GacS / RetS cascade acts as a positive regulator for the expression of C_4-HSL ultimately promoting the sessile lifestyle and expression of the T6SS [72]. The cascade causes the overexpression of RsmY and RsmZ, which in turn act as antagonists towards RsmA, which when unchallenged, negatively controls the expression of QS pathways. In mutant strains lacking the small RNAs RsmY and RsmZ or GacA, the synthesis of the AHL C_4-HSL was severely impaired as RsmA was no longer antagonised, resulting in high motility and other processes associated with the planktonic lifestyle [73]. These studies suggested that for lifestyle transitions, both regulatory systems are important and interlinked. However, the GacS / RetS system is top of the hierarchy having significant control over the Las and Rhl QS systems, which have been suggested as an additional regulatory layer [74].

1.5.3 Cyclic-di-GMP

A final molecule that plays a role amongst the chaos of these systems is cyclic diguanylate also known as cyclic-di-GMP (c-di-GMP), an important secondary messenger in signal transduction pathways. This secondary messenger has been implicated in the transition of lifestyles from motile to stationary. As previously mentioned, in a fully functional GacS/RetS cascade, RetS exerts its biofilm repressive affects by forming a heterodimer with GacS, keeping P. aeruginosa in the planktonic, acute infection stage. In a study looking at the effects of a RetS mutation, it was noted that biofilm formation was induced, the T3SS repressed, and the chronically associated T6SS, was activated [75]. A further observation noted was that high levels of c-di-GMP induced similar phenotypes as the RetS mutant strain. To assess whether these levels of high c-di-GMP were able to directly affect biofilm formation and repression of the T3SS, c-di-GMP levels were manipulated in the WT strain. Unexpectedly, high levels of c-di-GMP initiated the switch from the T3SS to the T6SS (Figure 1.6) [75]. Furthermore, elevated concentrations of intracellular c-di-GMP promoted the increased production of the exopolysaccharide psl, an important component in biofilm formation, further promoting, along with the above regulatory

systems, the transition to the sessile lifestyle [76]. These results prove that these systems are not independent of one another and are all instrumental in coordinating the lifestyle transitions of *P. aeruginosa* [77] .

1.6 Lifestyles of *P. aeruginosa*

1.6.1 Free-swimming, planktonic lifestyles causing acute infections

Having described the complexities involved in regulating the lifestyle transitions of *P. aeruginosa*, I will now discuss the importance of each lifestyle and the challenges faced when trying to treat these infections. The acute / free swimming lifestyle is associated with the bacteria adopting a planktonic lifestyle (ability to travel and colonise new niches). Acute *P. aeruginosa* infections are often associated with a rapid dissemination rate, leading to tissue damage and in the worst cases, pneumonia and sepsis, which comes with a high mortality rate ranging from 18-61% [78]. Infections that are more treatable and do not present as life threating, unless there are severe complications, include, burn sites, wound infections and UTIs. During an acute infection swarming and swimming motility is up, lipase and protease production is increased, rhamnolipids (surfactants that aid in swarming motility) are produced and the T3SS virulence system responsible for injecting effector proteins into the host is upregulated. As a consequence, type IV pili expression is downregulated, alginate production and biofilm formation are repressed along with downregulation of the T6SS virulence system [67].

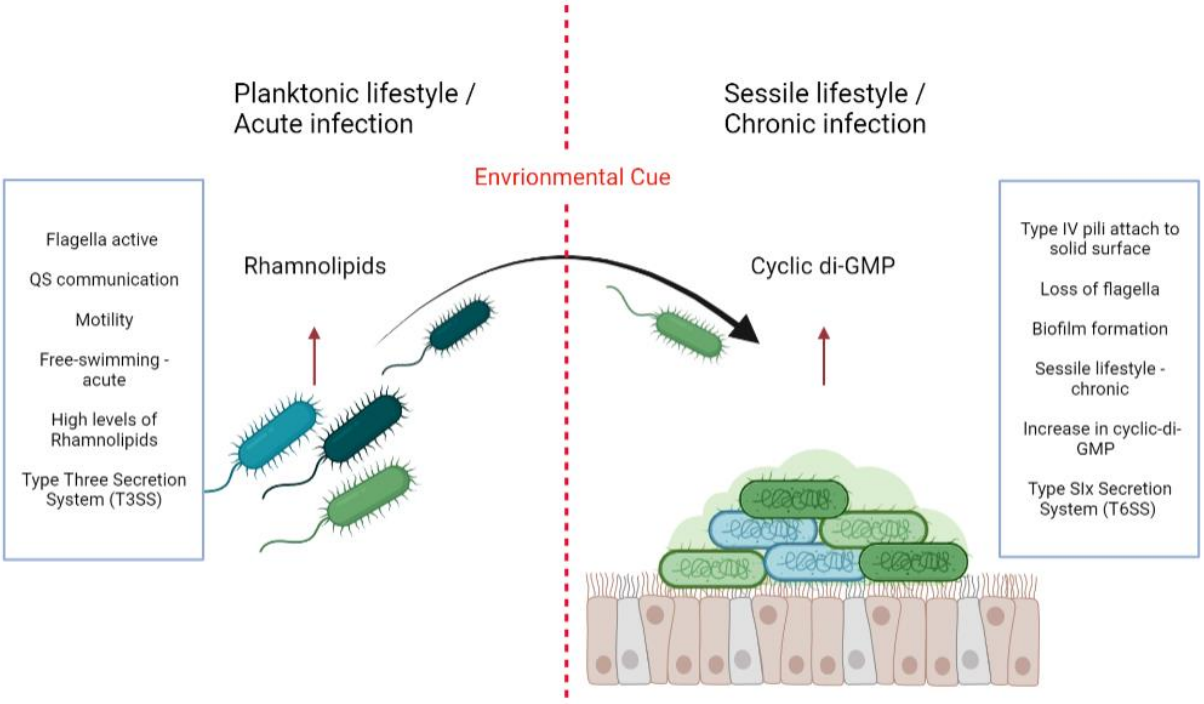


Figure 1.6. Diagram representing the two lifestyles of *P. aeruginosa*. On the left is the planktonic lifestyle associated with acute infections, high motility that is predominantly flagella-dependant, the T3SS system is expressed injecting toxins into host cells and high levels of rhamnolipids are produced. Cell-to-cell communication (QS) is on-going at all times, allowing bacteria to assess population densities and neighbour communication. The red line down the centre is an environmental queue telling *P. aeruginosa* to begin the transition to a sessile lifestyle associated with chronic infections. Here, motility is down, flagella are lost and type IV pili are expressed and begin to attach to epithelial cells, this initiation of cell surface attachment begins biofilm formation and the expression of the T6SS.

1.6.2 Stationary, sessile lifestyles responsible for chronic infections

The sessile lifestyle deployed by *P. aeruginosa* is commonly associated with chronic infections. These are associated with progressive infections getting worse over time, leading to repeat respiratory infections and resulting in constant inflammation of tissues. These chronic infections can occur on burn wounds; however, they predominantly affect patients suffering from cystic fibrosis (CF). CF is a genetic condition caused by a defective CFTR gene that affects mucus producing cells. Normally this mucus is thin and slippery, however in CF patients this mucus is thick and sticky. Ordinarily this mucus is designed to aid in the lung's immune response, trapping any airway irritants or pathogens, which can then be expelled. Due to the sticky nature of the mucus produced by CF patients, it is difficult to expel. This not only clogs up airways making it difficult to breathe, but it also traps in pathogens that it was designed to get rid of, essentially making it a host defensive defect. Not only does the mucus trap in the bacteria, due to its composition, it makes it an ideal environment for colonisation, providing the nutrients such as carbon and nitrogen sources that opportunistic pathogens such as *P. aeruginosa* thrive upon [79].

The transition from acute to chronic infection is often described at the point a biofilm begins to form which is initiated by the loss of flagella and the attachment to a solid surface by the type IV pili (Figure 1.7) [60]. Biofilms are an architectural structure that bacteria build which is initiated by the attachment to a surface and coupled with other environmental inputs leading to the GacS/RetS cascade. Biofilms are primarily comprised of polysaccharides, extracellular DNA, proteins and lipids. There are five main stages to biofilm formation; i) Initial surface attachment mediated by the flagella and the type IV pili, and the production of three exopolysaccharides; *psl* (required for the initial adhesion of cells), *pel* (biofilm matrix component linked with promoting aminoglycoside antibiotic tolerance) and alginate (involved in the maturation of the biofilm and protection from phagocytosis) [80],; ii) cells begin to replicate and transition from reversible to irreversible attachment with loss of flagella; iii) increased cell-cell interaction co-ordinating progression of a more structured architecture termed a microcolony; iv) development into a mature biofilm that has a three dimensional structure with increased production of extracellular polymeric substances and, v) breaking away of cells transitioning back into a planktonic lifestyle allowing dispersion to uncolonized spaces (Figure 1.7) [81].

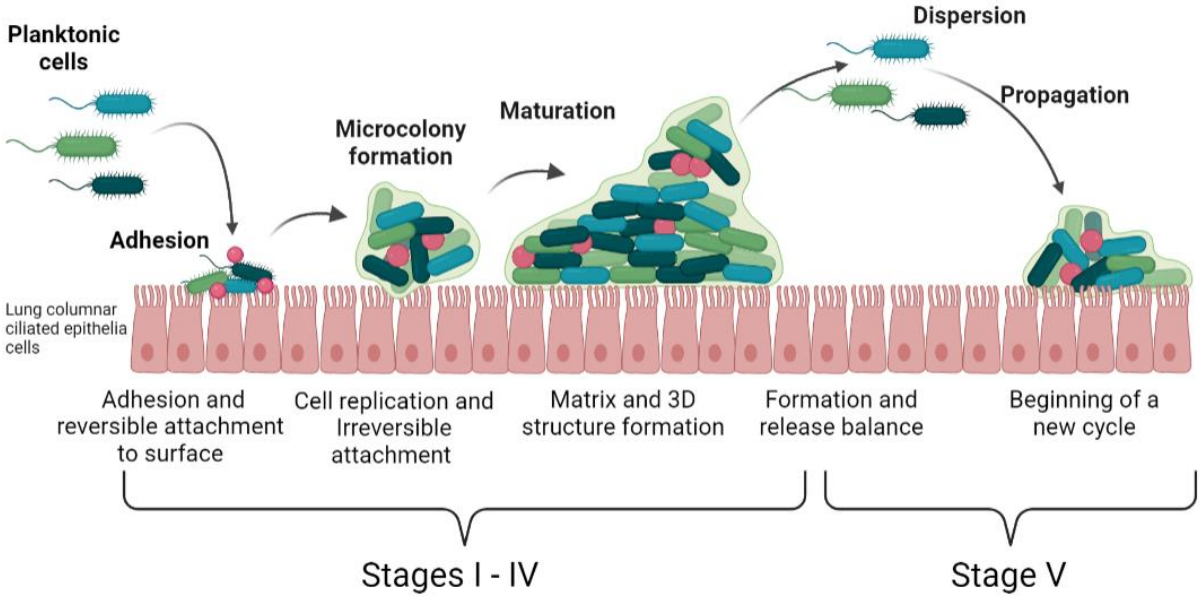


Figure 1.7. Diagrammatic representation explaining how biofilm formation begins. It starts with planktonic cells receiving an external environmental queue to begin adhering to cell surfaces, in this case lung columnar ciliated epithelium cells.

1.7 Biofilms

Biofilm formation is particularly problematic for CF patients. Biofilm formation, and the chronic recurrent infections that comes with it, leads to constant irritation of lung tissues ending in fatal lung failure. Biofilms provide several mechanisms that makes the treatment of *P. aeruginosa* chronic infections more difficult to treat than planktonic cells [82]. The first difficulty lies in the inability of an antibiotic to penetrate the thick, polysaccharide matrix encasing the biofilm which

acts as a diffusion barrier, with few able to do so. Even if antibiotics can penetrate and disseminate within a biofilm, they face further challenges. Multidrug-tolerant persister cells can remain due to a sub-optimal therapeutic dose reaching cells within the biofilm matrix; this selects for antibiotic resistant strains, and ultimately leads to persistent and recurring infections. Furthermore, within a biofilm there are differences in the growth rates of cells. Some cells within a biofilm are slow growing and metabolically inactive. This has been shown to be problematic for the effectiveness of antibiotics that require a cell to be metabolically active in order to have an effect, e.g. ciprofloxacin, tetracycline and tobramycin [83]. The final hurdle antibiotics encounter is to enter and remain in the cell long enough to take effect. It has been shown that within a biofilm population there are range of phenotypically different cells expressing a variety of different efflux pumps, that are often more upregulated in biofilm environments than when in the planktonic lifestyle, further conferring resistance to many antibiotics [84].

In some instances, *P. aeruginosa* infection will not clear from the lungs. In these patients, constant long term oral and inhaled antibiotic treatments are required to keep the bacteria at a low enough level to ease symptoms. This has led to the debate of whether chronic infections are recurrent or a relapse and reactivation of a previous strain that has remained and persisted [85]. The issue with long term use of antibiotics in staving off these infections is, their recurrent and overuse leads to antibiotic resistance. This additional antibiotic resistance in bacteria such as *P. aeruginosa,* that are already intrinsically resistant to many already, has contributed towards the dilemma we face today.

With the antibiotics currently at our disposable proving inadequate, along with the slow rate of the development and approval of new antibiotics looking bleak, new approaches are desperately needed [86]. One such approach being undertaken is assessing new druggable targets integral to cell function and virulence. Given that virulence processes are essential for the overall pathogenicity of bacteria, targeting proteins involved in virulence is an attractive prospect, and has the advantage that is a low selection pressure for resistance as it is not necessarily detrimental to a pathogens growth and survival [87]. Furthermore, targeting virulence will reduce pathogenesis, allowing valuable time to establish the exact cause of an infection and its antibiotic sensitivity, resulting in a more targeted narrow spectrum therapy helping to prevent antibiotic resistance selection. One issue that has been highlighted for targeting virulence is that, given the incredible bacterial diversity as discussed above, virulence mechanisms vary massively from bacteria-to-bacteria. This means that identifying a broad-spectrum virulence target present a major challenge [87]. This important point has brought researchers to a fascinating group of enzymes that are ubiquitous in nature, which over the years

have been shown to possess a plethora of biological roles, including as will be discussed, in bacterial virulence [88].

1.8 Peptidyl-prolyl *cis-trans* isomerases, FKBPs and Macrophage Infectivity potentiators discovery, categorisation and functions

1.8.1 Peptidyl-prolyl *cis-trans* isomerases

Amongst the great diversity between eukaryotes and prokaryotes lies a ubiquitous family of proteins termed the peptidyl-prolyl *cis-trans* isomerases (PPIases). These are isomerase enzymes that interconvert the *cis*/*trans* structure of peptide bonds involving the amino acid L-proline (Xaa-Pro, Xaa being any amino acid and Pro being short for proline). Within this large family of proteins there are three subgroups: i) parvulins, ii) cyclophilins and iii) **FK**506-**b**inding **p**roteins; (FKBP). These subgroups are based on their structural homology as well as their sensitivity to naturally occurring inhibitor molecules [89]. Cyclophilins and FKBPs are the most closely related and together make up the immunophilin branch of PPIases, with parvulins being placed in a group on their own due to their unrelated protein sequences and drug binding properties (Figure 1.8) [90]. The immunophilin branch are termed as such due to their function as receptors for immunosuppressive (natural product derived) drugs, cyclosporin, FK506 (tacrolimus), ascomycin and rapamycin. The presence of these enzymes was first documented in 1984 in pig kidney extracts as well as several other tissues [91]. The Fischer group termed this new enzyme peptidyl-prolyl *cis-trans*-isomerase and went on to study it further.

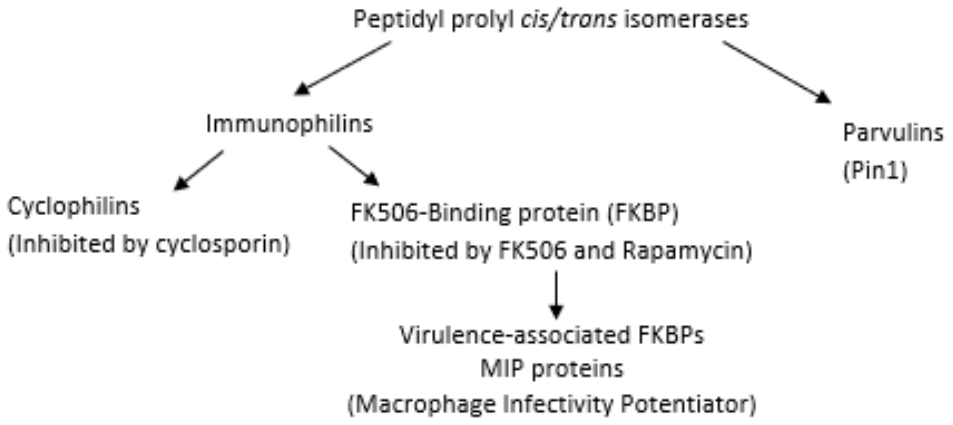


Figure 1.8. A schematic diagram showing how the large PPIase family is divided based on their sequence similarities and drug binding properties. The immunophilin branch are related based on their inhibition resulting in immunosuppression in humans where they were first discovered. The parvulins are placed in a group on their own as they possess no sequence similarities to the immunophilin branch, along with a different cellular function involved in transition from G2 phase to M phase in the cell cycle [92]. Within the immunophilin family, the FKBP subgroup possess a subgroup of its own. Today, the majority of FKBPs across eukaryotes and prokaryotes belong to the FKBP subgroup and are annotated with housekeeping / chaperone functions. However recently FKBPs have been shown to play a role in virulence processes of several pathogens. In these instances that have been annotated as termed Macrophage Infectivity Potentiators (Mips) and sit in their own subgroup within the FKBP subgroup.

PPIase's are often involved in protein folding as they catalyse the slow *cis-trans* isomerisation of proline-peptide bonds which is often a key rate-limiting reaction in protein folding [93, 94]. In nature, the *trans* configuration of an amino acid is observed in the majority (99.7%) of peptide bonds due to steric hinderance between the functional groups attached to α-carbon atoms when in the *cis* isomer [95]. However, the *cis* isomer is more readily observed (occurrence increases to 6.5%) upon involvement of the amino acid proline in peptide bonds. This is due to the cyclic nature of proline's side chain (Figure 1.9) giving the *cis* and *trans* configurations almost equivalent energies [96]. Proline residues that are preceded by glycine and other aromatic residues yield increasing fractions of the *cis* isomer with up to 40% identified for aromatic-proline peptide bonds. However, the interconversion between the *cis* and *trans* isomers of peptide bonds between Xaa-Pro does not occur readily due to the high activation energy which must be overcome for isomerisation.

Nature has thus evolved enzymes to fulfil this role, namely the PPIases which are described as protein folding chaperones. The highest sequence conservation among the FKBP subgroup, is observed at the C-terminus region of these proteins, for both eukaryotic and prokaryotic organisms[97]. In fact, the PPIase activity of the FKBPs is reliant on the conservation of specific amino acid residues in the C-terminus and this has been evidenced by substituting out individual amino acid residues thought to be responsible for PPIase activity. For instance, substitution of the single amino acid, aspartate 37 in human FKBP (hFKBP) and the corresponding aspartate 142 in the *Legionella pneumophila* FKBP, termed **M**acrophage **I**nfectivity **P**otentiator (LgMip), resulted in the loss of PPIase activity [97]. This LgMip will be discussed in more detail in the coming sections as its initial discovery and elucidation as an FKBP with PPIase activity took several years of unexpectedly intertwined studies. Furthermore, it was shown that the substitution of tyrosine 82 in human FKBP and the corresponding tyrosine 185 in *L. pneumophila* are essential for the binding of the inhibitors FK506 and rapamycin.

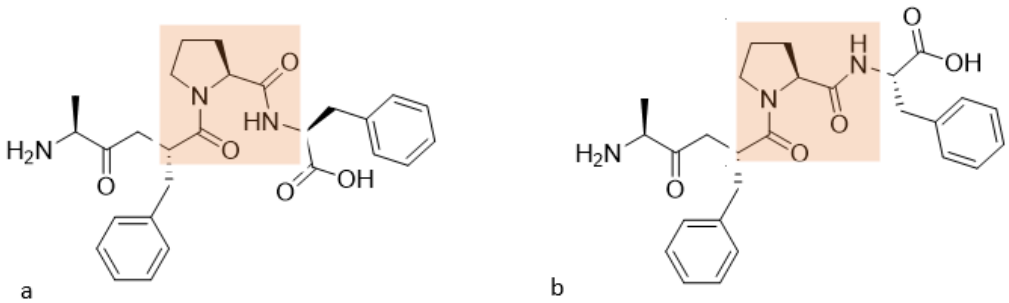


Figure 1.9. Chemical structures of the two different isomers of the peptide bond involved in amino acid proline. Two identical peptides containing the different isomers of the peptide bond (*trans* and *cis*) involving the amino acid proline. a) highlighted shows the *trans* isomer of proline and b) highlights the *cis* isomer of proline. Both these bonds display high steric hinderance and interconversion requires a high activation energy. PPIase enzymes are required to speed this reaction up.

1.8.2 The discovery and differentiation of PPIases

Today we know that FKBPs, cyclophilins and parvulins, belong to the large superfamily that are the PPIases, however this has not always been the case [94]. Each PPIase subgroup was established following the discovery of their own naturally occurring inhibitor, the first of which was the immunosuppressant Cyclosporin A (CsA) [98]. The target of CsA was shown to be a cytoplasmic binding protein, cyclophilin. This target was only determined eight years following CsA's initial discovery [99]. Cyclophilins were shown to have an important biological function in assisting protein folding by catalysing the *cis/trans* isomerisation of peptide bonds, a process which is inhibited by the addition of CsA [94, 100]. In mammalian systems, inhibition of cyclophilin and it's PPIase activity causes an immunosuppressive effect by interfering with the calcineurin–phosphatase pathway cascading down to the suppression of T cell activation [101, 102]. The FKBP family of proteins was identified and their functions elucidated in a similar fashion to that of cyclophilin in that the natural inhibitor FKBP inhibitor, FK506, was found first using a phenotypic screen and its binding target only identified subsequently.

FK506 (commercial name: Tacrolimus), was first isolated from the fermentation broth of *Streptomyces tsukubaensis* by Fujisawa Pharmaceutical Company in 1988 [102]. FK506 was originally used as an immunosuppressive drug following allogenic organ transplantations, to lower the risk of organ rejection [102]. Before the commercial sale of this new FK506 immunosuppressive molecule it's mechanism of action and binding target remained unknown. It was reported that the immunopharmacological properties of FK506 were similar to that of CsA, with both resulting in the suppression of T cell activation. Therefore, it was proposed that they would bind the same target, cyclophilin [103, 104]. However, both Harding *et al.* and Siekierka *et al.* disproved this theory by noting that CsA and FK506 were structurally unrelated so would unlikely bind the same target [104, 105]. Both groups tested their theories by using experiments that involved the use of affinity matrices. They successfully proved that both immunosuppressants CsA and FK506 bound to different cytoplasmic binding proteins; CsA to cyclophilin and FK506 to a newly elucidated target **FK**506 **B**inding **P**rotein (FKBP). In these two independent studies, each group used purified cytosol extracts from both bovine thymus and human spleen which were absorbed on to an FK506 affinity matrix column and then eluted after washing away non-binding proteins. Each sample eluted a single 14kDa protein that was different to cyclophilin, and they termed these proteins bovine FKBP (bFKBP) and human FKBP (hFKBP). The two groups looked for any differences between the known cyclophilin protein and the new FKBP proteins via immunoblotting using affinity-purified rabbit anti-cyclophilin IgG. As expected, the purified cyclophilin reacted strongly, whereas the newly purified FKBPs did not,

confirming FKBP was antigenically unrelated to cyclophilin[104, 105]. These two new FKBP proteins, along with cyclophilin, were subsequently named immunophilins.

The first 40 amino acids and first 16 amino acids of the N-terminus protein sequences were deduced by using Mass Spectrometry in both the bFKBP and hFKBP, respectively. Using these partial protein sequences, Harding *et al.,* looked to establish whether there were any published FKBP homologues known at that time using the protein sequence database NBRF [104]. However, this search failed to find any significant similarities to other known protein sequences for several reasons; firstly, they had only obtained a partial sequence for their proteins, with the acquired sequence belonging to only the N-terminus. As mentioned previously, the C-terminus is responsible for FKBP sequence conservation, thus, having only the N-terminus protein sequence was unavailing. Secondly, they only searched using the protein sequence as they had not yet established the nucleotide sequence. This searching error was later accounted for in work carried out by Standaert *et al.* [106].

Subsequently, Harrison *et al*. independently confirmed the high degree of specificity each immunophilin had for its cognate immunosuppressant drug [104, 105, 107]. This was done by conducting a PPIase assay using purified bFKBP, hFKBP, cyclophilin and a tetrapeptide Suc-Ala-Ala-Pro-Phe substrate, then adding in FK506 to inhibit the reaction [104]. This assay was performed by the addition of the tetrapeptide substrate to each purified protein in the absence of an immunosuppressant. The isomerization of the substrate by the purified proteins was measurable by isomerisation kinetics of the substrate using UV/Vis spectrophotometry which returned different absorbance readings for either the *cis* or *trans* isomer. This showed both FKBP and cyclophilin act as catalysts in the *cis-trans* isomerization of the Ala-Pro amide bond in the tetrapeptide substrate. To test for substrate specificity, FK506 was added into each separate reaction, as expected only FKBP activity was inhibited by the addition of FK506. Therefore it was concluded and widely accepted that although cyclophilin and FKBP were both members of the new immunophilin family of PPIase enzymes, they were distinct from one another and bound different inhibitors [105].

1.8.3 The FKBP homologue search conundrum

Shortly following these works which established that FKBP and cyclophilins were distinct enzymes activity, the first complete hFKBP sequence was published independently by the groups Standeart and Maki, who both reported identical sequences [106, 108]. They independently managed to obtain the full cDNA sequence of hFKBP through a four stage PCR approach. Using the already known N-terminal amino acid sequence, primers were designed to obtain a 64 base pair explicit sequence, which was subsequently used to find the

remainder of the unknown C-terminal sequence by a one-sided PCR procedure [109, 110]. Finally, a full-length cDNA clone was generated using primers corresponding to each end of the cDNA. A 108 amino acid full length protein sequence was deduced along with a 538bp nucleotide sequence [106]. It was also shown that there was only a three-amino acid difference between the hFKBP and that of the bFKBP. The groups deposited both the full protein and nucleotide sequences of the hFKBP to relevant databases. This sequencing work proved to be essential in paving the way to uncovering future FKBP homologues.

In the same study, Standaert *et al.* noted that previous searches for FKBP homologues relied on using only the partially sequenced N-terminus of the hFKBP protein published by Harding *et al*. Having established the full-length protein sequence, they performed their own hFKBP homologue search using their newly elucidated protein sequence on NBRF. This search failed to uncover any compelling similarities to other published protein sequences. However, having elucidated the nucleotide sequence of hFKBP, they performed a second search on EMBL, a database of published primary nucleotide sequences for both eukaryotic and prokaryotic genomes. This search revealed a high degree of nucleotide sequence conservation (61%) between the now establish hFBKP open reading frame (ORF) and that of an unstudied cryptic ORF in the partially sequenced genome of the gram-negative bacterium *Neisseria meningitidis* (Figure 1.10) [106, 111]. This was the first instance an FKBP nucleotide sequence had been used to search for homologues, instead of solely protein sequences. Searching in this way would account for proteins that were not yet annotated, but whose nucleotide sequence was available as part of any partly published genome sequences. Once translated, the protein sequence of the cryptic ORF in *N. meningitidis* was shown to be 56% identical to the hFKBP over the C-terminal stretch, with sequence conservation dropping off at the N-terminal end. It was subsequently found that the ORF of *N. meningitidis* did in fact encode an FKBP (termed NmMip) with C -terminus PPIase activity that was inhibitable by the addition of FK506 [112]. How this was ultimately achieved will be discussed shortly.

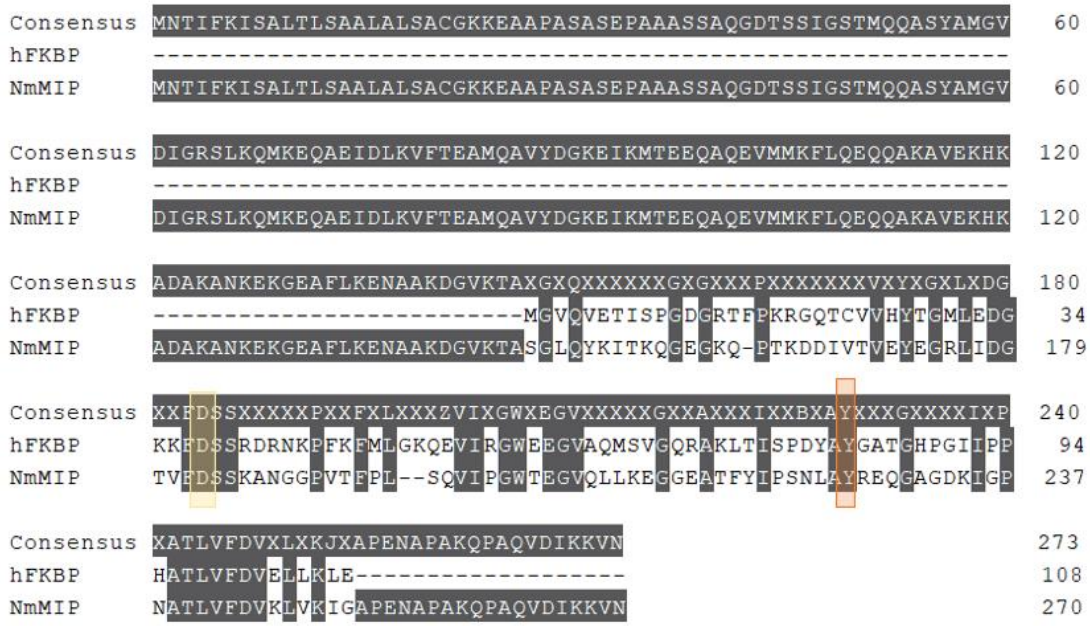


Figure 1.10. Sequence alignment of hFKBP and *N. meningitides* FKBP (NmMip). Initial protein alignment performed by Standaert *et al.* of both hFKBP (FKBP12) and *N. meningitides* FKBP (now termed NmMip) obtained by translating the ORF in each genome, then amino acid sequences were aligned. The coloured boxes highlight essential amino acids required for PPIase activity. The yellow box highlights a key residue in the active site (D – Aspartate), where-as the orange box highlights an essential amino acid in the binding of FK506/rapamycin (Y - Tyrosine). The black highlighted regions display the highly conserved amino acids in the C-domain region.

These studies represent huge contributions to improving how homology searches were conducted. They highlighted the importance of incorporating both nucleotide and protein sequences in searching for homologues, as doing so would increase the number of viable hits on various databases. They also highlighted difficulties encountered in these types of searches in the field at that time. One such difficulty in the late 1980s may have been attributed to the finite number of annotated proteins in the NRBF database, as well as the scarce number of completely sequenced genomes that were publicly available. Although Sanger sequencing was available in 1977, only small viral genomes of around 55kb in size were able to be sequenced, while fully sequencing larger, more complicated, bacterial genomes was thought to be unfeasible. This is evidenced by the delay of nearly 18 years from the introduction of Sanger sequencing to the first fully sequenced bacterial genome in 1995 via shotgun sequencing. Sequencing at this time was a rather inaccessible, timely and expensive process compared with today's readily available, fast, inexpensive NGS services.

1.9 The first Macrophage infectivity potentiators (Mip)

While the above studies were ongoing into eukaryotic FKBPs, Engleberg *et al.,* began to separate and characterise antigens in *Legionella pneumophila*, the causative agent of Legionnaires' disease [113]. *L. pneumophila* is a gram-negative, intracellular pathogen able to infect alveolar macrophages via interacting directly with a target cell, or by releasing outer membrane vesicles (OMVs) that can travel long distances to deliver virulence factors [114]. The Engleberg group stated their long-term research goals were to better understand the molecular immunobiology of *L. pneumophila*, something that was poorly understood in 1989, and their initial hypothesis suggested that a key factor in *L. pneumophila's* capacity to cause disease was its ability to enter and colonise human alveolar macrophages [113]. Their first works described, in detail, proteins of *L. pneumophila* whose antigens were localised to the cell surface, particularly those that invoked immune responses in a host. Through these cloning experiments, they identified surface protein antigens that were potential virulence factors involved in pathogenicity [115-117]. Having annotated these proteins, they began to focus on surface protein antigens specific to *L. pneumophila* that shared no known similarities in other organisms. Subsequent studies focussed on investigating a gene encoding an unannotated *L. pneumophila* specific surface protein, 24 kilo-Dalton (kDa) in size, as a target for site-directed mutagenesis [115].

To test whether this protein was involved in intracellular infection, the group designed an *in-vitro* model using monolayers of transformed U937 cells and human alveolar macrophages. The U937 cell line was developed from a human histiocytic lymphoma cell line, to behave like a human macrophage cell line that supported intracellular growth of *L. pneumophila* [118]. It was found that *L. pneumophila* strains with mutations in the 24-kDa surface protein were severely attenuated at infecting both the U937 cell line and human alveolar macrophages. When the wild type (WT) gene was reintroduced, the strains regained full infectivity. They appropriately named the *L. pneumophila* gene and its product **M**acrophage **I**nfectivity **P**otentiator (Mip), with it being shortened to LgMip for the *L. pneumophila* protein [115]. Although in this study it was established as a protein essential for full infectivity, it's mechanism of action remained unknown. To date, the mechanism of action of Mips, and how they are involved in the virulence of several gram-negative pathogens remains unknown. This will be discussed in more detail in the coming sections.

Concomitantly, the same group published a separate paper with the full nucleotide and protein sequence of their LgMip. Much like how Harding *et al.* had searched for FKBP homologues, they searched using only the LgMip protein sequence in the hope of uncovering any potential activity based on homology to other annotated prokaryotic proteins with a known function. [119]. This search initially uncovered no homologues of LgMip, suggesting it may be a unique virulence

protein belonging to *L. pneumophila* with several potential biological functions based on its chemical structure [119].

Two years following the sequencing of LgMip and its categorisation as a virulence factor, a Mip-like homologue was found in the gram-negative pathogen *Chlamydia trachomatis*. In similar fashion to Engleberg *et al.,* Lundemose *et al.* began work on proteins localised to the cell surface of *C. trachomatis*. Initial work uncovered two surface proteins responsible for the ability of *C. trachomatis* to bind to the surface of HeLa cells, a human cell line used in infection assays [120]. In a subsequent publication, they uncovered and began to categorise a third, potentially important, 27kDa surface membrane protein of *C. trachomatis* [121]. In this initial work they described cloning and sequencing part of the gene encoding the 27kDa protein. Using this partial gene sequenced, they performed a homology search which returned a single hit with a strong sequence identity match to the C-terminus of LgMip. Due to the high sequence conservation with the LgMip, they annotated this gene and its product as a *C. trachomatis* Mip-like protein (CtMip). The group deemed CtMip to be homologous to LgMip, the activity of which had still not been established other than it being required for full virulence processes in *L. pneumophila*.

Although they noted how similar the two proteins were, as well as the similarities in how both *C. trachomatis* and *L. pnuemophila* infections became established, they did not test in this present study whether the CtMip played a role in *C. trachomatis* virulence, like that of LgMip. They did surmise, however, that given an LgMip knockout drastically reduced virulence of *L. pneumophila*, it would be prudent to test whether this homologous CtMip-like protein had similar affects in an infection model. This hypothesis was tested several years later, following several other entwined publications linking these Mips not only to their roles in virulence, but to their biological activity as FKBPs. The complexities in how this was finally deduced will be discussed below.

1.10 The exponential rise in published FKBPs

During this early period of the 1990s, sequencing services became more accessible [2]. This allowed groups to procure and publish sequences of their studied proteins, depositing both nucleotide and protein sequences in the relevant databases. With this, there was an exponential increase of deposited sequences leading to a greater chance of sequence searches uncovering potential homologues. Following the above works and, publishing of the fully sequenced hFKBP, Tropschug *et al.* became interested in studying slow-protein reactions accelerated by PPIase activity in lower eukaryotes. The group reported a new FKBP with full PPIase activity in the fungus *Neurospora crassa*, at the time this was the second known FKBP uncovered following

the initial characterisation of hFKBP. Although Mip proteins above are FKBPs, at the time of the reporting of *N. crassa*'s FKBP, Mips had not yet been annotated as FKBPs with PPIase activity.

The Tropschug group initially found a protein in *N. crassa* displaying PPIase activity that was inhibited by the addition of FK506 but not CsA. Following this result, they decided to use a nucleotide and protein database search of their protein to further confirm its relation to hFKBP. This search revealed their protein to have a conserved, homologous region (44%) with that of the hFKBP C-terminus, as well as, more importantly, homology to the newly published LgMip sequence (Figure 1.11). This was the first instance an FKBP homology search returned LgMip as a potential FKBP relative, revealing some homology (39% sequence identity) to the C-terminal sequence of the LgMip. They suggested that given that LgMip had no known function, coupled with its sequence likeness to both *N. crassa* FKBP and hFKBP C-terminus, an investigation as to whether LgMip demonstrated PPIase activity would be interesting [122]. They concluded their work by annotating their protein as an FKBP (NcFKBP) due to its inhibitable PPIase activity by the addition of FK506, coupled with its sequence identity to hFKBP (44%). They also proposed that the FKBP sub-group of PPIases was not solely reserved to higher organisms involved in the immune response, but rather, they were likely to have a variety of functions across both higher and lower eukaryotes, as well as in prokaryotes.

With the reporting of the first FKBP in a lower eukaryote, many research groups between 1990-1992 uncovered additional FKBPs across a range of higher eukaryotes, lower eukaryotes and prokaryotes. These included; an FKBP isolated from calf thymus, a second human FKBP (hFKBP13), FKBP1 of the yeast *Saccharomyces cerevisiae* and finally an FKBP in the gram-positive bacteria *Streptomyces chrysomallus* [123-126]. Once groups had established PPIase activity in their protein, they elucidated the nucleotide and protein sequences. Using these sequences, each group performed their own independent homology search to further confirm their protein belonged to the FKBP subgroup using nucleotide and protein sequence databases (EMBL, NBRF and Genbank).

As expected, each group independently revealed their protein contained remarkable conservation in the C-terminus part of their protein sequences to the well described hFKBP and NcFKBP (Figure 1.11). More excitingly and unexpectedly, each of their searches revealed remarkable similarities with the virulence proteins LgMip and CtMip. This finding, coupled with the previous observation and hypothesis made by Tropschug *et al.* that that LgMip had no known function, along with its sequence likeness to both *N. crassa* FKBP and hFKBP C-terminus, an investigation as to whether LgMip demonstrated PPIase activity would be interesting, began to reveal there may be a potential link between FKBPs and Mips, with the LgMip potentially possessing PPIase activity. Furthermore, the nucleotide searches revealed similarities to an

unannotated open reading frame in the gram-negative pathogen *Pseudomonas aeruginosa* (Pa5254) and, for the second time, the cryptic ORF of *N. meningitidis*. Each group suggested that it would be interesting to investigate whether these unannotated ORFs of these pathogenic gram-negative bacteria housed FKBPs with functional PPIase activity based on the high sequence conservation highlighted in their homology searches.

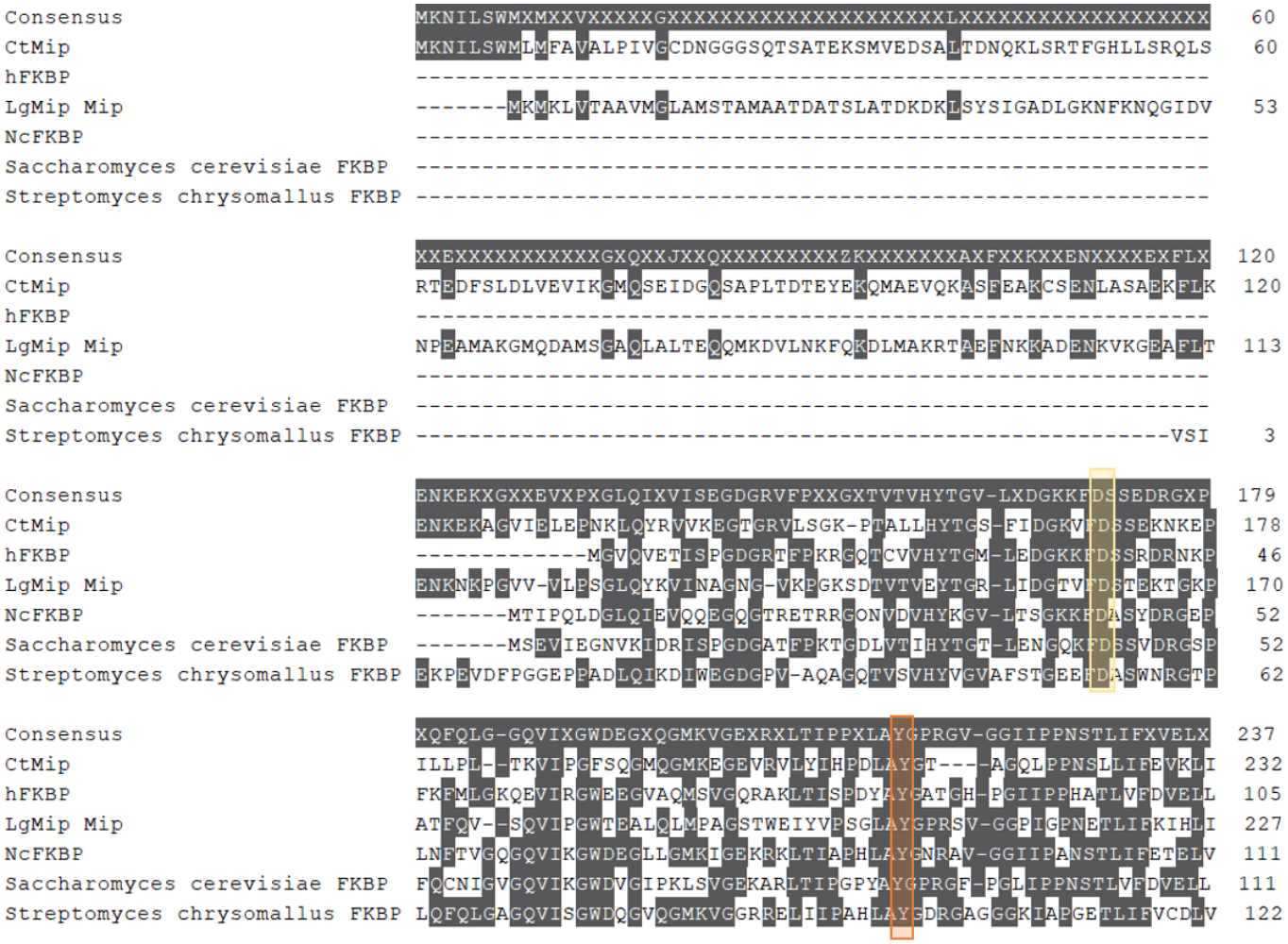


Figure 1.11. Protein sequence alignments of the *C. trachomatis* Mip (CtMip), human FKBP (hFKBP*), L. pneumophila* Mip (LgMip), *N. crassa* FKBP (NcFKBP), *N. meningitidis* Mip (NmMip), *Saccharomyces cerevisiae* FKBP, *Streptomyces chrysomallus* FKBP. The coloured boxes highlight essential amino acids required for PPIase activity. The yellow box highlights a key residue in the active site (D – Aspartate), where-as the orange box highlights an essential amino acid in the binding of FK506/rapamycin (Y - Tyrosine). The black highlighted regions display the highly conserved amino acids in the C-domain region.

Given that two separate works several years apart had now highlighted the unannotated ORF of *N. meningitidis* as potentially housing a gene encoding an FKBP, Sampson and Gotschlich decided to investigate this hypothesis [106, 112]. They expressed the gene and isolated its protein product and showed that it possessed PPIase activity that was inhibitable in the presence of FK506 or rapamycin. Further evidence that it belonged in the FKBP subgroup of PPIases came from their sequence alignment which showed 61% sequence identity conservation at the C-terminus with that of hFKBP (Figure 1.11). However, they failed to assess whether this Mip-like FKBP was involved in *N. meningitidis* virulence. Here, the group were agonisingly close to making the final link that Mips are a virulence associated subgroup of FKBPs. Eventually studies found this protein to be a Mip, termed NmMip, important for the survival of *N. meningitidis* in the bloodstream [127]. A second Mip-like FKBP from the *Neisseria spp.* was later investigated in 2005 in works by Leuzzi *et al.* who studied a homologous protein in the closely related *Neisseria gonorrhoeae* and concluded it was a Mip (NgMip) required for full virulence, homologous to LgMip and CtMip [128].

1.11 The final link – Mips are a subgroup of virulence associated FKBPs with PPIase activity

Finally, drawing together all of the above observations, the Fischer group sought to finally uncover a biological activity and function of LgMip beyond its only annotation as a virulence associated protein [129, 130]. They successfully achieved this by showing that recombinant LgMip exhibited PPIase activity localised to the C-terminus which had activity comparable to that of eukaryotic FKBPs, which was inhibited by FK506 but not CsA. They concluded that hFKBP and LgMip were members of the same superfamily of immunophilin proteins, with LgMip being placed in a new FKBP subgroup as the first virulence associated FKBP (Figure 1.8.). Thus, two years after the initial discovery of the *L. pneumophila* as a virulence factor, LgMip's activity had finally been elucidated thanks to overlapping works, the continual improvements in how scientists performed homology searches and, most importantly, due to the rapid advancements in sequencing platforms and the exponential increase of deposited sequences to databases allowing improved homology searches. They did not however, establish whether this new-found PPIase activity of LgMip was linked to its role in virulence.

This revelation that LgMip was part of a virulence-associated subgroup of FKBPs with PPIase activity, led the Lundemose group to expand on their previous investigation into the CtMip-like protein. They sought to uncover whether the CtMip was also involved in virulence while also establishing if, like its LgMip homologue, it belonged to the FKBP subgroup of PPIases with inhibitable activity [131]. Their experimental design for establishing PPIase activity was similar to that conducted in the LgMip study by Fischer *et al.*. They showed recombinant CtMip-like protein possessed PPIase activity which was inhibitable by the addition of FK506 and rapamycin,

but not CsA. Here, they concluded that the CtMip-like protein was an FKBP and moved on to test whether it was required for full virulence.

Using McCoy cells, which are useful for *C. trachomatis* cultivation deriving from human synovial fluid, they performed an infection assay. Unlike other studies that had used Mip knockouts to study effects on virulence, this group chose to use WT *C. trachomatis* that had been pre-treated with either FK506 or rapamycin before inoculation of McCoy cells, or McCoy cells dosed with FK506 or rapamycin at regular intervals during the infection assay. They showed pre-treating *C. trachomatis* resulted in a 30% reduction in infectivity of the McCoy cell line, while the dosing regimen with FK506 at intervals disrupted optimal intracellular infection by almost 80%. They also noted that during the dosing treatment, abnormal *C. trachomatis* cells were present throughout the infection assay suggesting using the inhibitors affected the overall fitness of *C. trachomatis* [131]. This experimental design was useful as it demonstrated that functional PPIase activity of CtMip was important for full virulence, with its inhibition leading to a reduction in virulence. This was a refreshing approach compared to previous studies, as not only did they show CtMip is implicated in virulence, but its PPIase activity played a role in doing so.

In summary these studies told us that; 1) Mips possessed PPIase activity and belonged to a subgroup of FKBPs (Figure 1.8) that had evolved as virulence factors and may be ubiquitous in prokaryotes, not were not unique to *L. pneumophila* and *C. trachomatis*; 2) the PPIase activity of Mips was likely to play a role in virulence. However, quite how the PPIase activity was involved in virulence remained unclear, and to this day remains unsolved. These studies made the final the link that Mips are FKBPs with PPIase activity and belong to their own subgroup of virulence associated FKBPs. Given this new unique subgroup of FKBPs, further interest around FKBPs in prokaryotes was raised pertaining to what other biological functions prokaryotic FKBPs may possess, and how they may fit in the large superfamily of PPIases.

1.12 Categorising further FKBPs in prokaryotes

Given the biological significance of FKBPs in eukaryotes, as well as their newly established importance for the virulence of clinically relevant pathogens, much interest had been generated in these proteins. Many groups now sought to uncover further FKBPs across a range of organisms, with the goal of characterising them and elucidating further interesting biological functions. Research into the various roles FKBPs play in eukaryotic cellular processes identified their involvement in protein folding, signal transduction pathways, receptor signalling and their secondary role in modulating the immune response [92, 132, 133]. With only the above prokaryotic FKBPs discussed above discovered, a further FKBP was found in the bacteria *S. chrysomallus.* Thus far this FKBP had only been annotated as

possessing PPIase activity for which no other biological function had been uncovered. Despite the high sequence conservation between the *S. chrysomallus* FKBP and the other bacterial Mips, it was noted that it would be unlikely to play a part in virulence like its LgMip and CtMip homologues, as *S. chrysomallus* was non-pathogenic. The *S. chrysomallus* FKBP was later found to have an important biological function in facilitating the correct folding of proteins that were aberrantly produced at high levels [134].

With only the above prokaryotic FKBPs established, many groups wished to reveal how widely distributed FKBPs were within the prokaryotes, along with their cellular functions [135, 136]. This included categorising proteins with PPIase activity as either cyclophilins, 'housekeeping' FKBPs, or virulence associated Mip-like FKBPs. It was not assumed; however, an organism would only possess one particular group of PPIases whereby if an organism was shown to house parvulins they would unlikely possess FKBPs also. It was further noted that an organism would only encode on particular type of FKBP, only housing either housekeeping FKBPs or virulence associated Mip FKBPs, and would likely possess a combination of the two with both housekeeping FKBPs and virulence associated Mip FKBPs. These searches, now routine courtesy of the studies described above, often began with rigorous homology searches using both protein and nucleotide sequences. One group studying the model organism *Escherichia coli* uncovered several ORFs that were predicted to encode FKBPs that were related to both NcFKBP and Mips [137]. It was believed that studying FKBPs in *E. coli* would give an excellent insight into the variety of functions FKBPs may possess given *E. coli* was a model organism known to have both pathogenic and non-pathogenic strains.

The first *E. coli* protein to be studied, SlyD*,* was shown to be an FKBP required for the lysis of cells infected by the bacteriophage phi X174 [137]. They noted that the PPIase activity of SlyD was required for the folding and stabilising of an envelope protein, E protein, allowing it to accumulate to levels required for the lysis of infected cells. They described SlyD as a housekeeping chaperone protein whose PPIase activity promoted correct protein folding [138]. They also noted it was unlikely to play a role in virulence as there was little homology between this FKBP, and the virulence associated LgMip (18%) and CtMip (15%) FKBPs.

A second *E. coli* gene of interest, *fkpA*, was shown to encode a protein with a similar sequence identity (42%, 35%) to *L. pneumophila* and *C. trachomatis* Mips respectively. Despite acknowledging the high sequence identity of FkpA to Mips, the group did not assess whether it may be involved in virulence, rather, they initially described it is non-essential gene as its deletion resulting in no detrimental effect on growth. It was further described as an FKBP protein with PPIase activity involved in protein folding and chaperoning of envelope proteins as well as being integral to the heat shock stress response [139]. They further noted that despite its high

sequence homology to Mips, it was unlikely to be involved in virulence as although this protein was studied in both pathogenic and non-pathogenic strains of *E. coli* and, unlike Mips, it was not located to the cell surface.

These studies shed light on the varying functions FKBPs could play in prokaryotes, highlighting their versatility while demonstrating how much remained unknown about these proteins. Furthermore, it highlighted that categorising FKBPs as either housekeeping or virulence associated was more complicated than simply comparing sequence homologies to establish function as demonstrated by the *fkpA* studies. One might have assumed based on the protein and nucleotide alignments of *fkpA* to LgMip, that *fkpA* would be a Mip-like FKBP involved in *E. coli* pathogenesis, but this was not the case [136]. What could be concluded however, was that by using the sequences of known FKBPs in homology searches would give an indication as to whether an organism of interest would contain something resembling an FKBP, but this would not help determine what its function in the organism may be. These types of searches were further used by groups interested in searching for novel FKBPs in organisms that had not yet been annotated with FKBPs, and what biological functions these novel FKBPs may possess ranging from important cellular process such as protein folding or, what we believe to be most interesting, FKBPs involved in bacterial virulence such as LgMip.

In summary, by the mid-1990s FKBPs and Mips were well established as; 1) a subgroup belonging to the large PPIase superfamily of proteins; 2) The FKBP subgroup had its own internal subdivision of virulence associated FKBPs (termed Mips) [136]. 3) Not all bacterial FKBPs were Mips, however all Mips currently described were FKBPs with PPIase activity; 4) FKBPs were ubiquitous across eukaryotes and prokaryotes with a variety of important organism dependent biological functions. Furthermore, sequence searches had become a reliable tool for uncovering potential FKBPs within bacteria whose genomes were partly or fully sequenced. This proved useful for many groups as attentions now turned towards finding and studying FKBPs in pathogenic bacteria in the hope of elucidating more about their functions and uncovering more Mips.

1.13 The Mip explosion – uncovering of Mips

The prospect of uncovering more Mips was now of great interest to the world of infectious microbiology. These proteins were essential for the full virulence of certain intracellular pathogens that were difficult to treat in a clinical setting, and they achieved this by enhancing their ability to survive within macrophages and epithelial cells. Furthermore, the cellular role that Mips played during infection could be inhibited by known compounds (that were well characterised drugs), and with the PPIase activity of Mips had been shown to play a role in the

virulence process. When combined this knowledge suggested that Mips might have potential as drug target candidates. These observations led to an explosion of groups identifying new FKBPs in human and plant pathogens including, *Coxiella burnetii*, *Trypanosoma cruzi, Salmonella typhimurium* and *Burkholderia pseudomallei* [135, 140, 141]. These will be discussed below

1.13.1 *Coxiella burnetii* Mip (CbMip)

Coxiella burnetii is a gram-negative obligate intracellular pathogen that is the cause of Q fever, a potentially life-threatening infection. *C. burnetii* was of interest to *Cianciotto et al*., who were responsible for the original work on LgMip, due to its phylogenetical closeness to *L. pneumophila*. They hunted for Mip-homologues using Mip specific probes and low-stringency Southern hybridisations, coupled with *in-silico* searches and identified a Mip homologue (CbMip) with sequence identities of 46% and 30% to LgMip and CtMip respectively [142]. Low-stringency Southern hybridisation was used to enable the hybridisation of nucleotide sequences with a particular size that have low-moderate homology [143]. Their subsequent works showed recombinant CbMip possessed PPIase activity that was inhibited by FK506, but failed to establish its role in virulence [140]. They did suggest, however, that as it was present in the genomes of all known pathogenic *Coxiella* species coupled with its phylogenetical closeness to *Legionella* species, the CbMip would most likely be involved in pathogenesis but that further studies were needed [144].

1.13.2 *Trypanosoma cruzi* Mip (TcMip)

Trypanosoma cruzi is a protozoan and the causative parasite of Chagas disease, a tropical disease affecting 16-18 million people worldwide and was shown unexpectedly to possess a Mip homologue [141, 145]. This parasite enters the human bloodstream via a bite from an infected insect who then defecates near the bite site. The faeces of the insect house the mobile *T. cruzi* parasite which uses its singular flagella to navigate into the bloodstream. The group studying this parasite found a protein released into the culture medium during the infective stage of this protozoan. Once purified they found the protein to possess PPIase activity inhibitable by FK506 and the corresponding gene was shown to have similar sequence identity to that of LgMip. The group demonstrated that in an infection assay, the addition of FK506 reduced *T. cruzi* invasion of HeLa cells and concluded that this protein was a Mip (TcMip) involved in intracellular invasion that was potentially mediated by its PPIase activity [141].

1.13.3 *Salmonella typhimurium* (StMip / fkpA)

Salmonella typhimurium is an obligate intracellular gram-negative human pathogen, most commonly responsible for cases of gastroenteritis that possesses homologues of the genes encoding *fkpA* and LgMip. With *fkpA* in *E. coli* being shown as non-essential for pathogenesis, there was an interest in deciphering whether this protein was involved in virulence or if its activity was more aligned with that of its homonym [135]. Given the Mips described to this point belonged to obligate intracellular pathogens, it was suggested Mips may belong predominately to intracellular Enterobacteriaceae pathogens such as *L. pneumophila*, *C. trichomonas* and now, potentially, *S. typhimurium*. It was further suggested that Mip-like homologues found in bacteria that were extracellular pathogens were less likely to be involved in virulence and were more likely to serve housekeeping roles much like the *E. coli* FkpA.

To test whether *fkpA* of *S. typhimurium* was essential for full virulence two different cell types were used for infection assays, Caco-2 cells (an intestinal epithelial cell line) and the macrophage-like cell line J774.A1, which had been established to extensively test the susceptibility of intracellular pathogens to phagocytosis [146]. Both cell types were used because during previous infection works, virulence related loci in *S. typhimurium* had been shown to be required for survival in either Caco-2 cells or J774.A1, but never both [147]. It was found that an *fkpA* mutant of *S. typhimurium* was able to enter both Caco-2 cells and J774.A1 macrophage cell lines as equally well as the WT. However, 4 h post incubation the number of viable *fkpA* mutant cells was reduced by nearly 80%, with the number of bacteria not able to recover after 6 h. In contrast the number of WT cells remained consistent across the 6 h period. The results were similar in the J774.A1 cell line, with both the *fkpA* mutant strain and the WT able to enter the J774.A1 cells at comparable levels. However, in this cell line, the mutant strain began to show a decrease in the number of viable cells earlier than the Caco-2 cells at 2 h post incubation. The number of WT cells was able to persist and increase, while the number of *fkpA* mutant cells were able to persist following the initial decrease after 2 h. These results showed that functional *fkpA* contributed towards the intracellular survival of *S. typhimurium* during the initial stages of infecting macrophages. This activity was able to be restored when complemented with the WT *fkpA* in both cell lines [135]. These findings suggested that *fkpA* was a Mip, however they did not annotate it as such. Although *fkpA* is not annotated as a Mip in the literature, I will refer to it from now on as StMip, to avoid confusion with the *fkpA* of *E. coli.*

The group carrying out this study proposed several potential mechanisms by which StMip and Mips of other intracellular pathogens contributed towards virulence. Firstly, they suggested StMip and other Mips might play an indirect role in pathogenesis by using their PPIase activity to correctly fold several bacterial proteins that are themselves involved in virulence. This idea was supported by the works by Lundemose *et al.* on the CtMip, where it was shown that inhibition of PPIase activity led to a decrease in virulence and intracellular survival [121, 131]. Alternatively, they suggested that StMip itself, or a protein it interacts with may enhance the ability of *S. typhimurium* to transition from growth outside of a host cell to intracellular invasion and growth of the bacteria once inside the host cell. Their final proposal was that StMip itself may promote virulence directly should its substrate be a defensive protein in the host cell, which when bound rendered the defensive protein ineffective [135]. Given that StMip and Mips at this time were shown to belong to intracellular pathogens, it was suggested that Mips belonged mainly to intracellular gram-negative pathogens, and that Mip-like FKBPs in extracellular pathogens likely played non-virulence roles.

More recent studies have identified gram-positive pathogens possessing PPIases important in virulence, and which exert their effects by modulating processes such as virulence protein secretion in *Bacillus subtilis* and evasion of phagocytes in *Streptomyces pneumoniae* [148-150]. However, these PPIases do not fit within the FKBP subgroup, aligning more with the cyclophilin subgroup and parvulin branch respectively, and therefore they do not qualify as Mips.

1.13.4 *Neisseria gonorrhoeae* and *Neisseria meningitidis* Mips (NgMip and NmMip)

Following the initial work in the discovery of an FKBP and potential Mip in *N. meningitidis*, Leuzzi *et al.* looked to study a similar protein in the closely related intracellular, gram-negative pathogen *N. gonorrhoeae* [128]. Although the two bacteria have very different clinical presentations, their genomes share 90% sequence identity [151]. Here, the group identified and characterised a novel surface exposed FKBP in *N. gonorrhoeae,* with inhibitable PPIase activity, and named it NgMip. To investigate the function of this protein and its potential role as a Mip, a series of infection assays using an NgMip knockout and examining its ability to adhere to, and to internalise and survive intracellularly in mouse macrophages.

Much like StMip, the NgMip knockout had comparable internalisation of macrophage cells to the WT. Unlike the StMip however, during the initial infection stages both the WT and NgMip strains rapidly decreased in number, not just the NgMip. However, drastic changes in their numbers were apparent during the later stages of intracellular infection. Over the course of the experiment around 50% of WT strain were able to survive. The NgMip strain proved to be

considerably more sensitive to macrophage-mediated killing in a time-dependent manor. After 3 and 6 h survival dropped to 21% and 11% respectively. After a 24 h period only 3% of the initially inoculated *N. gonorrhoeae* NgMip mutants were recovered [128]. When complemented with the WT NgMip gene, the strain became less sensitive to macrophage killing.

The NmMip of *N. meningitidis* has been of further interest in recent times since its initial study and categorisation as an FKBP in 1992 by Sampson and Gotschlich [112]. Recent studies have shown that NmMip is highly upregulated in *N. meningitidis* when exposed to blood in an *ex vivo* whole human blood model. The NmMip mutant strain showed a significant sensitivity to killing by human blood after 30 minutes. This showed a similar sensitivity to that seen in the related NgMip mutant which had already been shown to be a lipoprotein involved in the intracellular survival of *N. gonorrhoeae* in macrophages [128]. This led to the suggestion NmMip may be involved in facilitating *N. meningitidis* interaction and survival within macrophages [127].

The NmMip was also a focus of interest in research looking into potential antigens to be included in meningococcal vaccines**,** where it was shown recombinant NmMip expressed in *E. coli* induces serum bactericidal activity – killing of bacterial cells as a result of antibody production [127, 152]. This was a direct method developed to measure the functional activity of antibodies in a patient's blood serum against a specific microorganism causing an infection. This assay measured the ability of antibodies produced in the blood serum to kill off bacteria by assessing their capacity to bind to the target bacteria and activation of complement leading to a proteolytic cascade resulting in cell lysis and killing of the target bacteria [153, 154]. The group assessed the sequence similarities of the NmMip across 13 different, well categorised meningococcal strains that had been isolated from patients, whereby each strain belonged to a different serogroup (bacteria possessing a common antigen), serotype (closely related bacteria with a common antigen set) and subtype. Their amino acid alignments using NmMip from each strain showed 98-100% sequence conservation across all strains, sequences with distinct sequence types were designated I, II, and III. All strains expressed their NmMip at similar levels to one another.

This study found that in the serum bacterial activity assay, mice inoculated with type I recombinant NmMip induced high levels of surface-reactive antibodies with serum bactericidal activity. Furthermore, antisera (blood serum that contains specific antibodies against an infective organism) raised against type I, were also able to kill strains with NmMip sequence types II and III, showing that recombinant NmMip can provide cross-strain protection against various meningococcal strains with different serotypes. Thus far, the issue with developing a vaccine towards *N. meningitidis* is that it possesses a huge level of antigenic diversity spanning

many *N. meningitidis* serogroups. This study highlighted that as NmMip is ubiquitous across meningococcal strains and serogroups, with nearly identical sequence conservation across all, NmMip would be an excellent candidate to be included as an antigen in future meningococcal vaccines [155].

1.13.5 *Legionella pnuemophila* further studies

Since its initial characterisation as a virulence factor and subsequent elucidation as an FKBP, the LgMip has been studied in much more detail. These studies have focussed on deciphering the LgMip structure, potential binding targets in a host and attempting to elucidate further functions including, whether its PPIase activity was essential for its virulence role [156-158]. The crystal structure of LgMip was determined with a view to developing future drugs targeting Mip proteins of intracellular pathogens (Figure 1.12) [156]. This revealed that LgMip exists as a homodimer, with each monomer being split into three distinct regions, an N-terminal domain comprised of two α-helices, a long connecting α-helix in the centre, and a C-terminal FKBP domain responsible for PPIase activity. Prior to this study the function of the N-terminal domain was unknown, but the structural data showed that it facilitated dimerization of the protein. Dimerization is likely critical for the *in vivo* action of Mips, as a monomeric mutant exerted far less PPIase activity [159]. The function of the long α-helix centre was not elucidated but nucleic magnetic resonance (NMR) data assigned the helix structure as flexible, presumably allowing the Mips dimer to bind a range of substrates as a dimer, a property for protein interactions and subsequently, virulence.

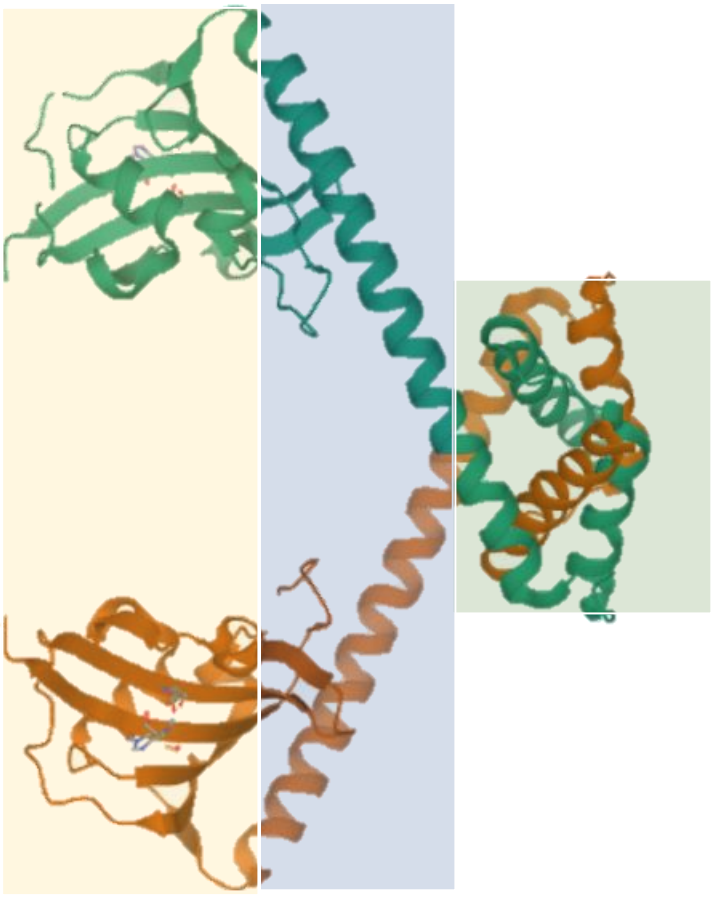


Figure 1.12. Crystal structure of LgMip. The 2.4 angstrom crystal structure of LgMip as a homodimer reported by Riboldi-Tunnicliffe *et al.* [156]. Each monomer of LgMip is shown in green or orange. The green box highlights the N-terminals of each monomer forming a dimerization module. The blue box highlights the long α-helix connecting the C-terminal to the N-terminal. The Yellow box 'highlights the FKBP domain responsible for PPIase activity.

In further the LgMip was truncated at the centre of the linking α-helix, creating monomeric LgMip's consisting of either the N-terminal dimerization domain or the FKBP PPIase C-terminal domain. Although the resulting C-terminal monomer retained PPIase activity comparable to full length LgMip in PPIase assays, however when introduced into a protein folding assay with a denatured substrate, the C-terminal monomer showed a significant loss in efficiency when refolding the substrate compared to the intact LgMip dimer. This suggested that for LgMip to perform efficient protein folding activity it must exist as a dimer, requiring the N-terminal dimerization domain. Furthermore, it was shown that strains expressing dimerization deficient Mip, or low PPIase activity, were severely attenuated in a guinea pig infection model [160].

Several extensive studies were conducted to identify the key residues required for full PPIase activity in both LgMip and hFKBP. It was found for hFKBP that nine residues were involved in PPIase activity, with two key residues shown to be essential, aspartate 37 part of the active site, and tyrosine 82 part of the FK506 binding domain [97]. When these two residues in

hFKBP were deleted, PPIase activity was lost. These studies led Ludwig *et al.* to surmise that, given the homology of LgMip to hFKBP, and coupled with sequence alignments highlighting that LgMip also possessed these key corresponding residues, that deletion of these residues would also likely result in a loss of PPIase activity of LgMip (Figure 1.13) [161]. Single amino acid changes were thus made using site directed mutagenesis replacing the aspartate 142 residue (corresponds to hFKBP aspartate 37) with leucine, and tyrosine 185 (corresponds to hFKBP tyrosine 82) with alanine. These LgMip mutants displayed strongly reduced PPIase activity of LgMip *in vitro* from 5.3% to 0.6%, and this activity was restored when complemented with the WT sequence.

Furthermore, they assessed whether *L. pneumophilia* mutants expressing LgMip lacking PPIase activity were attenuated an *in vivo* model at invading and surviving intracellularly in U937 cells and human blood monocytes. Mutants lacking PPIase activity in the U937 cell-line showed no significant difference of intracellular survival compared to the WT. However, in the monocytes, the survival of PPIase lacking mutants was severely compromised by a factor of 100. The group noted however, that the result obtained in the U937 cells may not be entirely representative as the 0.6% PPIase activity that remained in the *in vitro* experiment for the mutant may still have been sufficient to fulfil its intracellular role. They further noted that in hFKBP although mutations in these residues resulted in no PPIase activity, it did not reduce their ability to bind a substrate. Therefore, they surmised it was plausible that the PPIase activity of LgMip may not be entirely associated with Mip-dependant invasion and survival of eukaryotic cells by *L. pneumophila,* as long as the mutant enzyme was still capable of biding its substrate [161].

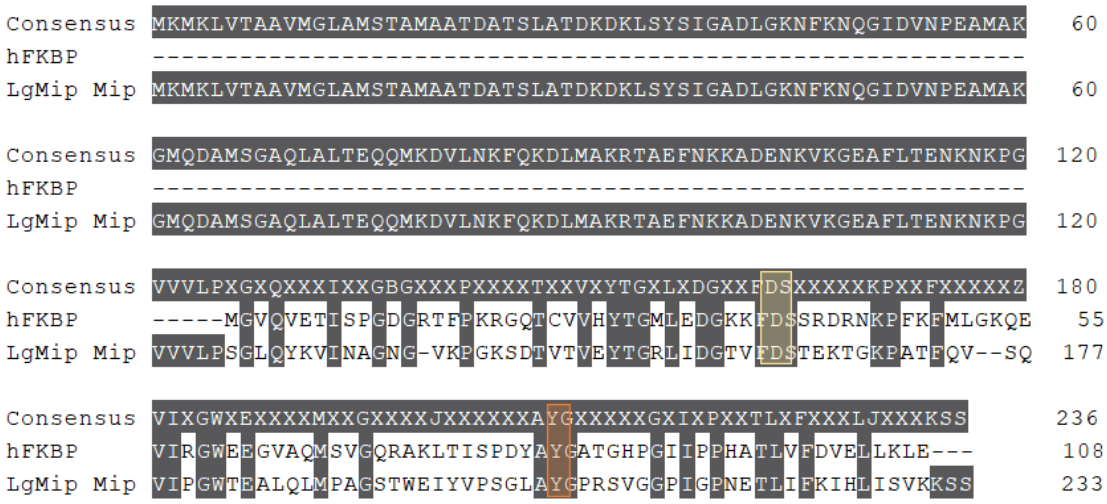


Figure 1.13. Sequence alignment highlighting key residues for PPIase activity. Protein sequence alignment of human FKBP (hFKBP) and *L. pneumophila* Mip (LgMip) The coloured boxes highlight essential amino acids required for PPIase activity. The yellow box highlights a key residue in the active site (D – Aspartate), where-as the orange box highlights an essential amino acid in the binding of FK506/rapamycin (Y - Tyrosine). The black highlighted regions display the highly conserved amino acids in the C-domain region.

Additional studies as to the importance of PPIase activity of LgMip in virulence were conducted by Fischer *et al.* following the landmark paper in which they had established LgMip as an FKBP with PPIase activity [129]. Given that in the CtMip it had been shown its PPIase activity was required for full virulence, they investigated if this may be the case for the LgMip [162]. They raised monoclonal antibodies that bound the LgMip C-terminus that were able to significantly inhibit LgMip PPIase activity *in vitro*; no loss of activity was observed for antibodies that bound the N-terminus. Using these antibodies in an infection study, they successfully demonstrated that blocking the PPIase activity site significantly reduced the ability of *L. pneumophila* in the establishment and initiation of intracellular infection of U937 cells. This demonstrated for the first time that an uninhibited, intact active site of LgMip was required for its full virulence enhancing abilities [162].

Further structural studies of LgMip demonstrated that the LgMip homodimer is situated on the outer membrane of *L. pneumophila* and binds specifically to collagens I-VI, with the highest affinity for collagen IV which is the most prevalent type of collagen in the lung. Binding of LgMip to collagen IV in guinea pig lung tissue allowed *L. pneumophila* to transmigrate the lung epithelial cell barrier and disseminate throughout lung tissues [163]. The addition of FK506 or rapamycin inhibited this transmigration across the barrier further suggesting LgMip PPIase activity, or at least the active site, was required for full virulence and dissemination in host tissues. This result led to the discovery of the surface exposed LgMip binding sequence of the NC1 domain of collagen IV. This binding sequence was shown to be completely localised to the surface of cells, making it ideal for LgMip to interact with. The binding sequence was elucidated and given the

name P290. Using this sequence, the group synthesised a P290 peptide to use as a competitor to bind LgMip, preventing it from binding the P290 sequence of collagen IV [158]. Using a competitive binding assay, LgMip was shown to have significantly reduced binding to collagen IV in the presence of the P290 peptide.

Given LgMip's involvement in the transmigration across lung epithelial cells via collagen IV binding, the group then tested whether the addition of P290 reduced this transmigration. The group compared *L. pneumophila* WT, *L. pneumophila* with an LgMip lacking PPIase activity, and *L. pneumophila* WT pre-treated with P290. Their results showed that only *L. pneumophila* WT was able to transmigrate across the lung epithelial cell barrier efficiently, while the WT pre-treated with P290 showed a 75% reduction in transmigration. Interestingly, this reduction was equivalent to the poor transmigration of the *L. pneumophila* strain with LgMip lacking PPIase activity, as well as comparable to the control group of *L. pneumophila* WT pre-treated with the known LgMip PPIase activity inhibitor, rapamycin [158]. These results suggested that interaction of *L. pneumophila* with collagen IV is mediated by the LgMip FKBP C-domain allowing adhesion of the bacteria to epithelial cells promoting transmigration, with the interaction inhibitable by the addition of FK506 or rapamycin or P920 [158]. Together, these studies showed that the LgMip C-terminus and its PPIase activity were not only essential in the invasion and intracellular survival of *L. pneumophila* in macrophages and lung epithelial cells, but LgMip also acted as a moonlighting protein (a protein with multiple relevant biophysical functions) by binding to collagen IV [164].

In unrelated work, research investigating effectors of the type II secretion system in *L. pneumophila*, a secretion system involved in promoting pathogenesis via the excretion of effectors such as exoenzymes (an enzyme that acts outside of a cell), uncovered an unexpected function of LgMip. *L. pneumophila* phospholipase C (PlcA) is a virulence determinant and enzyme that is capable of hydrolysing eukaryotic cell membrane structures leading to cell lysis and in other pathogens, such as *P. aeruginosa*, it has been shown to haemolyse red blood cells [165]. The group studied phospholipase C activity of LgMip mutants by fractioning the supernatants of LgMip mutants and comparing them to equivalent fractions of supernatant obtained from the WT. Interestingly, the phospholipase C activity of the LgMip mutant was reduced (40-70%) compared to the WT. When the LgMip was reintroduced by complementation, the secretion levels and activity of phospholipase C returned to that of the WT. This data provided the first evidence for a pleotropic role of LgMip in virulence. It was suggested that the PPIase activity of LgMip might be involved as a chaperone in the secretion or activation / correct folding of this hydrolase [155, 166, 167]. Although these studies have shed light on the role of LgMip in *L. pneumophila* pathogenesis, intracellular functions and targets of LgMip remain unknown. It is

likely LgMip executes a variety of intra- and extracellular functions during *L. pneumophila* pathogenesis.

1.13.6 *Xanthomonas campestris* Mip (XcMip)

For the first time, a Mip was discovered in a plant pathogen, *Xanthomonas campestris* pv. *Campestris,* a gram-negative plant pathogen and the cause of black rot disease in cruciferous plants, and was shown to possess a protein with 49% similarity at the amino acid level to that of LgMip [168]. The black rot disease causes a V shaped brown and yellow dead area on affected leaves (Figure13). Eventually the stem tissue will turn black where the plants will eventually show signs of wilting. Creating a Mip mutant (XcMip) strain using homologous suicide plasmid integration, they performed an infection assay on a host plant, the Chinese radish, by a method known as leaf clipping (introduces the bacteria directly into the vascular system) [169]. Ten days following inoculation of the clipped leaves with either the WT strain or XcMip-mutant it was found that leaves inoculated with the WT strain produced a V-shape lesion of 15.1mm (mean length), compared to 2.1mm (mean length) for the XcMip mutant (Figure 1.14) [168]. The mean lesion length was restored when complemented with the WT Mip gene.

To assess whether pathogen proliferation inside the affected tissues was affected for the XcMip mutant bacterial cells were recovered from the inoculated leaves. The number of XcMip mutant cells recovered was significantly less than that of the WT strain. On this basis it was concluded that the XcMip plays an important role in the pathogenesis of the plant pathogen *X campestris* pv. *Campestris*. Interestingly, it was also shown that the production of exopolysaccharides as well as the activity of extracellular proteases were significantly reduced for the XcMip mutant. This is important as it had previously been proven that exopolysaccharide production and extracellular protease activity of *X campestris* pv. *Campestris* were important virulence factors [170]. Like the StMip, they proposed several mechanisms by which the XcMip affected exopolysaccharide production and protease activity by acting as a chaperone in the correct folding of factors involved in their secretion or that the XcMip may be responsible for the correct folding of the exopolysaccharides and proteases themselves. This study was not only interesting as it uncovered a Mip gene involved in virulence of a plant pathogen, but also demonstrated for the first-time what virulence processes they may contribute towards [168].

Figure 1.14. Chinese radish lesions caused by *X campestris* pv. *Campestris* infection. Image from work performed by Zang *et al.* comparing the lesion lengths in Chinese radish after 10 days caused by infection of the WT, Mip mutant strain, and Mip mutant complement strain of *X campestris* pv. *Campestris* [168]. On the far left is the WT displaying the characteristic V-shaped yellowing of an infected leaf. The middle image is with the Mip mutant, showing only a small V-shaped yellowing, and indicating a significant reduction in virulence. The final image on the right shows the Mip mutant complemented with the WT gene restoring full infectivity.

1.13.7 *Burkholderia pseudomallei* BPSL1823 (BpML1) & BPSL0918

In recent years, the first organism containing two Mips fully involved in virulence has been discovered in the pathogen *Burkholderia pseudomallei. B. pseudomallei* is an intracellular, gram-negative pathogen known to cause melioidosis, or Whitmore's disease, with a mortality rate of 10-40%, even following proper treatment [171]. Until the mid-1990s, *B. pseudomallei* was known as *Pseudomonas pseudomallei* as part of the Pseudomodale order, within the Gammaproteobacteria class (these are paraphyletic where only some members of the group share a recent common ancestor). Taxonomists realised it was sufficiently different to merit being renamed, along with several other annotated members of the Pseudomodales and were placed into a different class, from Gammaproteobacteria to Betaproteobacteria (monophyletic groups sharing a most common recent ancestor) [172]. Burkholderia as a genus are one of the most, adaptable bacteria and are able to survive in hostile conditions, in some cases surviving in areas that have been disinfected. Much like *Pseudomonas aeruginosa*, they are intrinsically resistant to numerous antibiotics due to a host of efflux pumps and have recently been touted a potential bioterror agent [173]. The treatment options for melioidosis (infection caused by *B. pseudomallei*) have always been limited. For example, although an *in vitro* agar plate antibiotic disc diffusion test may return the organism as sensitive to an antibiotic, the antibiotic's use *in vivo* treatment can be suboptimal, often requiring much larger therapeutic doses than what is recommended and safe. The current first line treatment of *B. pseudomallei* infection is ceftazidime and other beta-lactam antibiotics, however resistant strains have begun to be isolated from clinical settings [174, 175].

Given the issues and threats posed by *B. pseudomallei*, further research into its pathogenicity has been of interest to many groups. To determine novel virulence determinants and genes required for the intracellular lifestyle of *B. pseudomallei*, Norville *et al.* used transposon mutagenesis in the hope of finding novel virulence determinants [176, 177]. The resulting mutants were assessed by examining their ability to perform intracellular spreading via the formation of plaques on PtK2 cell monolayers, as well their ability to replicate within mammalian cells. This led to the identification of a mutant carrying a disruption in the gene *BSPL0918*, a cytoplasmic protein, which displayed a reduced ability to replicate within mammalian cells and was also shown to be severely attenuated *in vivo* using a mouse infection model. They found that *BSPL0918 e*ncoded a putative FKBP with a sequence identity of 37% to the StMip. As *B. pseudomallei* is an intracellular pathogen, and coupled with the previous studies demonstrating the involvement of Mip PPIase activity in the pathogenicity of intracellular pathogens, the group performed a PPIase assay [177].

Until this point, and based on research described above, it was thought that PPIase activity was an inherent and ubiquitous trait of all Mips, and that it played a role in how Mips exerted their virulence affects [131, 162]. The PPIase assay performed using purified BPSL0918 resulted in no measurable PPIase activity, and sequence alignments further revealed that BSPL0198 lacked six out of the nine residues required for full PPIase activity, including the essential aspartate and tyrosine residues previously highlighted as essential for LgMip and hFKBP (Figure 1.18.). The absence of these residues would explain the lack of measurable PPIase activity. Thus, based on the previous observations with other Mips that PPIase activity was necessary for full virulence, had a bioinformatic analysis been undertaken prior to the biochemical assay, it could have been assumed that BSPL0918 would not be required for full virulence as it lacked this PPIase activity. However, as mentioned above BSPL0918 mutants were severely attenuated in their ability to resist intracellular killing and replicate within macrophages and were severely attenuated in a murine infection model [177]. Thus, BSPL0918 was the first Mip described to have no PPIase activity, however, is important for full virulence, suggesting that the virulence effect of Mips is not always mediated by PPIase activity.

A concurrent publication by the same group reported the identification of a second Mip from *B. pseudomallei* encoded by the gene BPSL1823, a periplasmic protein, which was later annotated as BpML1 (*B. pseudomallei* Mip-like protein 1, [178]. Unlike BPSL0918, this Mip was shown to exhibit full PPIase activity which was inhibitable by FK506 and rapamycin with sequence identities of 40%, 45%, and 42% amino acid identity to LgMip, NgMip, and CtMip, respectively. Further analysis showed that BpML1 possesses the key residues necessary for PPIase activity, unlike BPSL0918, with the crystal structure revealing an almost identical C-

terminus active site to that of LgMip (Figure 1.15). However, unlike the LgMip, BpML1 lacks N-terminal dimerization and exists as a monomer [179]. Although confirmed to exhibit PPIase activity, the involvement and importance of BpML1 PPIase activity in virulence was not studied here.

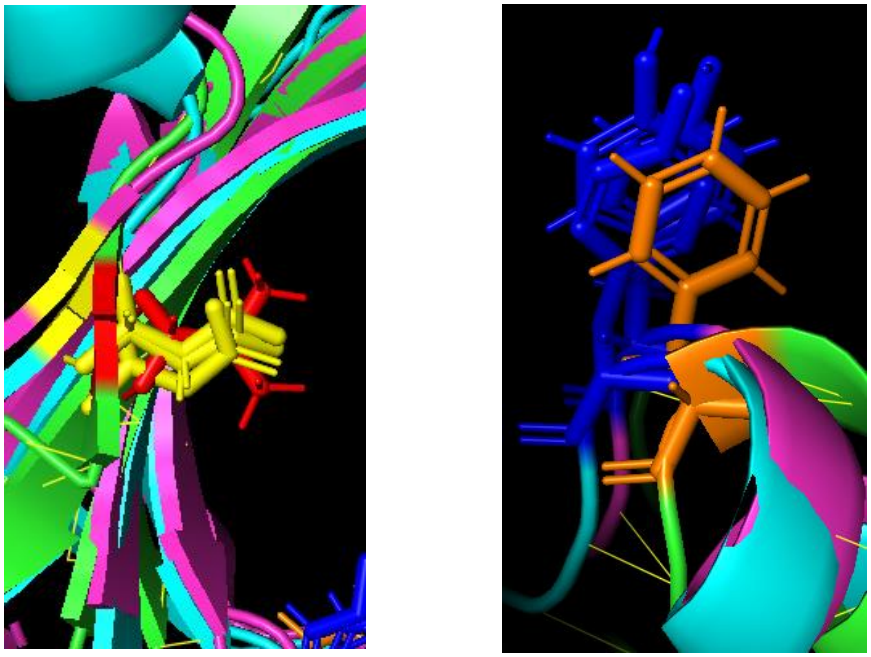


Figure 1.15. Crystal structure alignments of LgMip, BPSL0918 and BpML1. Aligned crystal structures of LgMip, BPSL0918 and BpML1 in both the left and right images. In the left image is a crystal structure alignment highlighting the position of the key aspartate residue (in yellow) required for PPIase activity. Both LgMip and BpML1 display this residue in a nearly identical alignment. This is replaced with the red valine residue in BSPL0918. The image on the right shows a crystal structure alignment of LgMip, BPSL0918 and BpML1, the blue residues highlight the essential tyrosine in both LgMip and BpML1 required for substrate binding. This is replaced by an orange phenylalanine residue in BSPL0918 where the tyrosine residue should be if binding of a substrate were to happen and PPIase activity to be achieved.

To assess the role of BpML1 in virulence, a mutant strain lacking BpML1 was created which exhibited a reduced ability to survive and replicate within J774A.1 murine macrophages or A549 human epithelial cells. The growth of this mutant was comparable to that of the WT in LB media suggesting no growth defects of the mutant *in vitro.* Furthermore, the virulence of this BpML1 mutant was significantly attenuated in a mouse infection model with only 50% of mice deceased 40days post infection, compared to 100% of mice succumbing to infection one day post inoculation with the WT (Figure 1.16) [178].

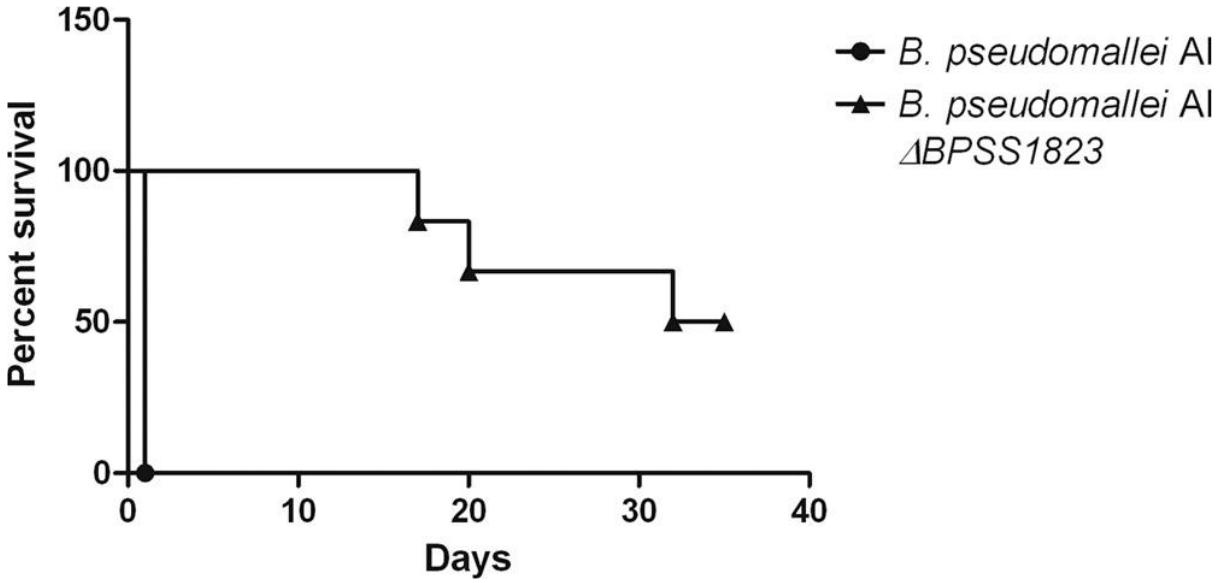


Figure 1.16. Line graph plotting the percentage survival of mice inoculated with *B. pseudomallei*. Graph from *Norville et al.* showing the percentage survival of mice inoculated with either WT *B. pseudomallei* or the *B. pseudomallei* BpML1 mutant [178]. Over the course of one day 100% of mice inoculated with the WT strain died. IN contrast the BpML1 mutant shows a significantly reduced ability to kill mice over a 40 day period, with the first death recorded 17days post inoculation; overall only 50% of inoculated mice were deceased by the end of the study.

To uncover how BpML1 may play a role in the virulence of *B. pseudomallei*, a variety of membrane-associated virulence mechanisms were investigated including swarming motility and protease secretion. Swarming motility is a flagella-dependant movement across a solid surface, often aided by the production of surfactants. To study this type of movement *in vitro* assays on 0.3% agar plates were used, inoculating the surface with a drop of culture, and assessing bacterial spread (swarming) (Figure 1.17). Here it was shown that the BpML1 mutant lacked any bacterial spread on the agar plate in contrast to the WT, and scanning electron microscopy (SEM) was used to show that the BpML1 mutant did not possess an intact flagellum (Figure 1.17). Upon *in trans* complementation of the WT *BpML1* gene, bacterial spread as well as the presence of a flagellum was restored.

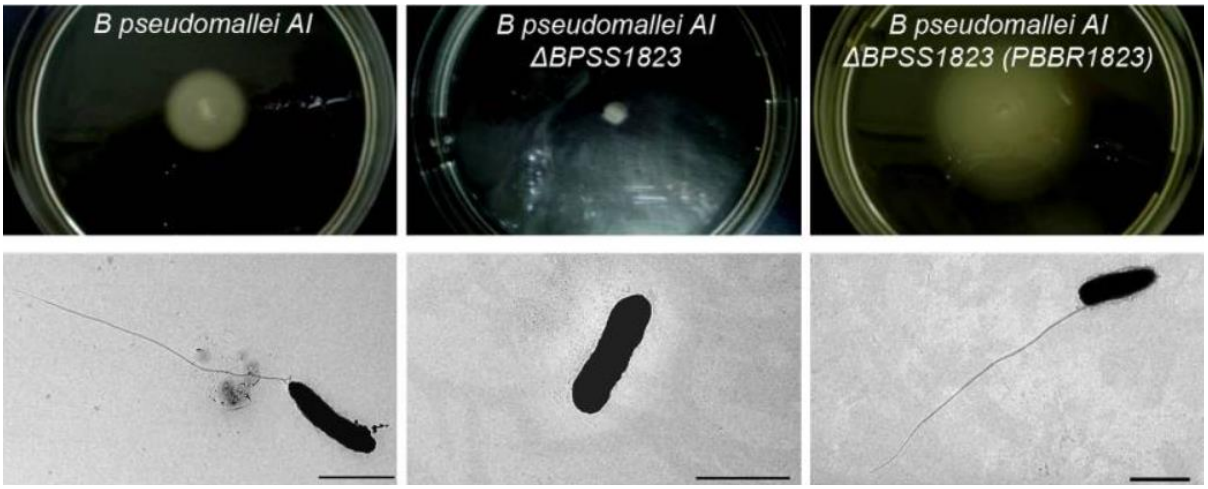


Figure 1.17. Swarming phenotypes and SEM images of *B. pseudomallei* and a BpMlL1 mutant strain. Images taken from Norville *et al*. showing swarming phenotypes on agar plates and EM images of *B. pseudomallei* cells [178]. The two panels on far the left show WT *B. pseudomallei* performing swarming on an agar plate (top left image), with the image below showing the WT and its intact flagella using SEM. The middle panel shows the BpML1 mutant lacks any swarming motility (top middle panel) which is caused by the lack of flagellum shown in the SEM image beneath. The far-right panels show that the swarming phenotype can be complemented when the BpML1 gene is reintroduced *in trans* and that this leads to the return of the flagellum.

A second membrane-associated virulence associated phenotype examined in this study was secreted protease activity. The secretions of proteases by pathogenic bacteria has many important biological is implicated in survival in the host by aiding in; evasion of the host immune system, enduring phagocytosis, and escaping entanglement in host fibrin meshes [180]. Using azocasein (a nonspecific protease substrate) as a substrate it was shown that although both the WT and BpML1 mutant strains were able to hydrolyse azocasein, the BpML1 activity was four-fold lower than the WT [178]. This study was the first to identify that a Mip gene is involved in modulating a broad range of virulence-associated phenotypes, further evidencing how diverse and pleotropic the role of Mip genes are at mediating pathogenicity. The means by which BpML1 Mips exerts these effects were not studied in this paper, and to date remains unclear for all Mips.

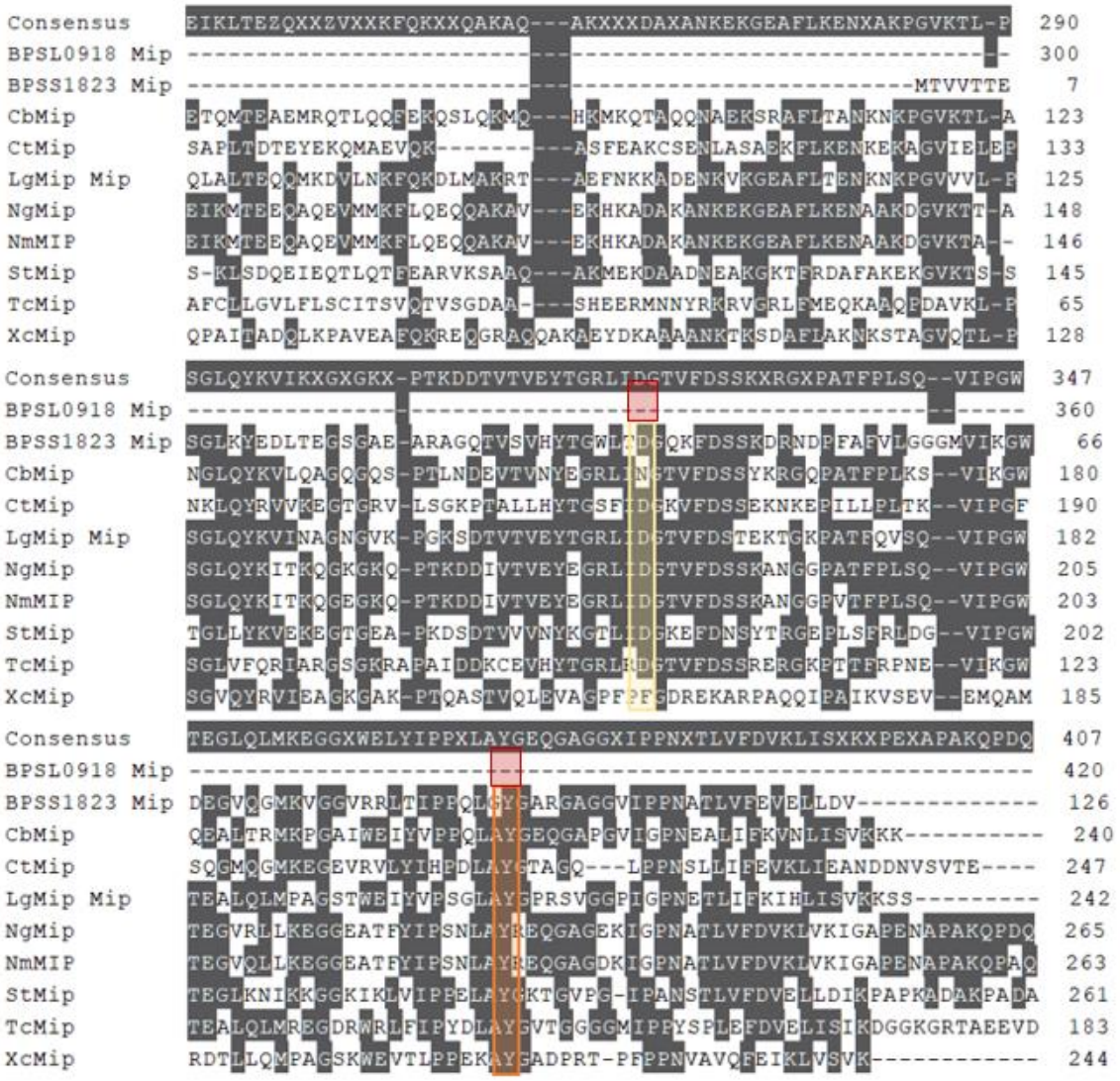


Figure 1.18. Protein alignment of the Mips mentioned in the introduction; *B. pseudomallei* 0918 Mip, *B. pseudomallei* 1823 (BpML1), *C. burnetii* Mip (CbMip), *C. trachomatis* (CtMip), *L. pneumophila* Mip (LgMip), *N. gonorrhoeae* Mip (NgMip), *N. meningitidis* Mip (NmMip), S. typhimurium fkpA (StMip), *T. cruzi* Mip (TcMip), *X. campestris* Mip (XcMip). The coloured boxes highlight essential amino acids required for PPIase activity. The yellow box highlights a key residue in the active site (D – Aspartate), where-as the orange box highlights an essential amino acid in the binding of FK506/rapamycin (Y - Tyrosine). The black highlighted regions display the highly conserved amino acids in the C-domain region. The red box highlights the missing aspartate and tyrosine of the BPSL0918 Mip, essential for PPIase activity and FK506 binding in other Mips, partially explaining the lack of PPIase activity observed for this FKBP.

1.14 Summary: the current understanding of Mip biology

To summarise, the virulence associated subgroup of FKBPs termed macrophage infectivity potentiators, or Mips, have predominantly been identified in intracellular, gram-negative pathogens of humans, animals and plants (Table 1.1). This is not to say that gram-positive bacteria do not possess PPIases involved in virulence. Several gram-positive, extracellular pathogens have been shown to possess PPIases, that act in the chaperoning and protein folding of other virulence determinants including the parvulin EF0685, and the cyclophilin EF1534 of *Enterococcus faecalis which play a role in the folding of secreted virulence proteins [181]. The gram-positive pathogen Streptococcus pneumonia was found to possess PPIases directly involved in virulence,* the cyclophilin SlrA (streptococcal rotamase A) and parvulin PpmA (proteinase maturation protein A)*. These two PPIases were shown to modulate the immune response in a mouse infection model, with mutants lacking PpmA showing reduced virulence [182]. Although interesting,* they will not be discussed further as they are not phylogenetically or structurally related to Mips which belong to the FKBP subgroup of PPIases that are inhibited by FK506 or rapamycin, which are the focus of this thesis.

The evidence discussed here, be for gram-negative or gram-positive pathogens, highlights the important role that PPIases, and their enzymatic activity, play in bacterial virulence. How they do this, for the most part, is poorly understood. It is clear, especially for Mips, that their involvement in virulence is exerted in a multitude of ways. It has been evidenced that majority of Mips studied require an intact active site for full virulence. As discussed, the active site of LgMip plays a direct role in virulence by binding collagen IV and allowing *L. pneumophila* to transmigrate across epithelial cells and disseminate within host tissues [158, 163]. The active site and PPIase activity of CtMip is essential for full virulence in *C. trachomatis,* with its inhibition by rapamycin resulting in a reduction in infectivity towards McCoy cells [131]. Most recently, the BpML1 mutant has been shown to have secondary affects in modulating certain virulence mechanisms in *B. pseudomallei* such as swarming motility [178]. Most of the more recent studies into Mips has suggested the active site amongst annotated Mips is relatively conserved and is needed to play a role in virulence process. Again, for the most part, Mips as a protein are conformationally flexible which allows them to; bind several substrates, mediate protein-protein interactions and bind directly to host cells, all of which alludes to the pleotropic nature of Mips [179]. Thus, it is evident Mips play a role in virulence, involved in a plethora of biological processes crucial to a pathogen's virulence processes, although the exact mechanisms responsible for these phenotypes remains elusive. The most important conclusion to be drawn from the studies discussed above for each Mip is that Mips

are somehow pivotal in virulence, with their inhibition leading to a reduction in the virulence of an organism. This makes them an attractive drug target, more specifically an anti-virulence drug target[173, 183].

Organism	Protein	PPIase activity	Association
Burkholderia pseudomallei -	BpML1	Yes	Mutant has severe attenuation in murine infection model, reduced ability to survive in adverse environments (e.g. pH<5), reduced ability to swam, reduced expression of virulence associated proteases
Burkholderia pseudomallei -	BSPL0918	No	Mutants unable to resist intracellular killing and to replicate in human macrophages
Chlamydia trachomatis -	CtMip	Yes	Exposure to FK506/rapamycin significantly reduced (80%) initial infection of McCoy cells, reduced intracellular survival and transformation to the reproductive state
Coxiella burnetii -	CbMip	Yes	Immunogenic Cross-reactive with polyclonal anti-Mip for LMip
Legionella pneumophila	LgMip	Yes	Mutants were impaired in infection of macrophages and amoebae, significant attenuation in the guinea pig infection model
Neisseria gonorrhoeae	NgMip	Yes	Mutant showed decreased persistence in macrophages
Neisseria meningitidis-	NmMIp	Yes	Induces high levels of surface-reactive antibodies, potential serogroup B vaccine candidate
Salmonella Typhimurium-	StMip / fkpA	Yes	Mutant showed a reduced ability to survive intracellularly
Trypanosoma cruzi	TcMip	Yes	Shown to enhance invasion of HeLa cells
Xanthomonas campestris	XcMip	Yes	Mutant showed significant reductions in virulence and ability to replicate in host plant Chinese radishes.

Table 1.1. A table listing Mips in a variety of species, whether they have PPIase activity and their involvement in virulence. Gram negative pathogens are indicated with *-*.

Given the biological importance of Mips discussed, their candidacy for the development of new anti-virulence drugs is attractive [184]. Anti-virulence drugs are an attractive new approach for the treatment of multi-drug resistant infections as inhibiting virulence does not necessarily affect the ability of a pathogen to survive, thus they should exhibit a low propensity for the selection of resistance. Potentially, an anti-virulence drug will help slow down infections and limit pathogenicity, which in turn will help limit the prescribing of broad-spectrum antibiotics until a proper pathology report can be generated allowing the optimal narrow spectrum antibiotic to be determined and prescribed [185].

As discussed in detail, *P. aeruginosa* is a pathogen that currently presents as a serious global health concern. Fully antibiotic resistant strains of *P. aeruginosa* responsible for serious infections have been isolated in clinical settings and are associated with high mortality rates [186]. Furthermore, *P. aeruginosa* infections are of even greater concern in immunocompromised patient groups, predominantly those affected by CF who are more likely to develop multi-drug resistant *P. aeruginosa* infections. With the plethora of virulence tools at *P. aeruginosa's* disposal, assessing whether any of these tools are potential anti-virulence drug targets to improve patient outcomes is a promising area of research. Given that several gram-negative pathogens have been shown to possess Mip proteins playing a role in virulence, which when inhibited reduce the pathogenicity of a pathogen, in this work we have investigated whether *P. aeruginosa* has Mip-like proteins and whether these are involved in virulence processes.

Chapter 2:

Materials and Methods

2. Materials and Methods

2.1 General materials

Reagents and chemicals were purchased from Sigma-Aldrich, Formedium, Oxoid Thermo Fisher, Calbiochem and E&O Laboratories. All were used without further purification.

Commercially available enzymes, reagents, DNA purification kits and DNA ladders used in this study are listed in Table (2.1)

Supplier	**Enzymes and kits**
New England Biolabs®	1Kb DNA ladder Gibson Assembly Master Mix® T4 DNA Ligase Phusion® High-Fidelity DNA Polymerase NEB Restriction Endonucleases
Promega	GoTaq® Green Master Mix Deoxynucleotide Triphosphate (dNTPs) Set PCR Grade
Qiagen	QIAprep Spin Miniprep Kit QIAquick gel exraction kit RNeasy Mini kit (MPbio – Lysing Matrix B 100 ml tube RNase-DNase free) RNAprotect Tissue Reagent (50 ml)

Table 2.1. Lab consumables and reagents used in this study.

2.2 Bacterial strains, plasmids and oligonucleotides

2.2.1 Bacterial strains

A list of bacterial strains used in this study can be found in Table 2.2. Glycerol stocks were generated from a single colony streaked out on LB agar and picked into liquid LB then left overnight in the 37 ^{o}C shaking incubator at 250rpm (revolutions per minute). Glycerol stocks were then generated from a 1:1 mixture of overnight liquid culture and 50% glycerol and stored at -80^{o}C.

Strain	Genotype / description	References
Pseudomonas aeruginosa PA01	Wild type strain from which all mutant strains were generated. This strain will be abbreviated to PA01 below.	This study
Escherichia coli DH10B: pUC19	Stock strain used for pUC19 plasmid extraction. Plasmid was used for cloning purposes to align and ligate homologous arm fragments up and downstream of a gene to be deleted. Once fragments were ligated to one another they were digested out of pUC19 and ligated into pTS1. Tetracycline resistance and *SacB* counter selection.	This study.
Escherichia coli DH10B: pTS1	Stock strain to use for plasmid extraction of the pTS1 plasmid used for gene knockouts in *P. aeruginosa*. This plasmid is a suicide vector designed to integrate into the chromosome via homologous recombination with two overlapping arms to each side of a target gene. Tetracycline resistance.	This study.
Escherichia coli DH10B: pUC18-mini-TN7T-Gm	Stock strain to use for plasmid extraction of pUC18-mini-TN7T-Gm the plasmid used for complementation of gene knockouts in *P. aeruginosa*. Carbenicillin and Gentamycin resistance.	This study.
Escherichia coli DH10B: pJH10TS	Stock strain to use for plasmid extraction of the pJH10TS plasmid used for complementation of gene knockouts in *P. aeruginosa.* Tetracycline resistance.	This study.
Escherichia coli DH10B: pME6032	Stock strain to use for plasmid extraction of pME6032 the plasmid used for complementation of gene knockouts in *P. aeruginosa*. Tetracycline resistance.	This study.
Escherichia coli DH10B: pUC19 - *pa3262* Frag A	pUC19 containing 800-1200bp overlapping region upstream of *pa3262* and the first 45bp of *pa3262* gene termed Frag A. Carbenicillin resistance.	This study.
Escherichia coli DH10B: pUC19 - *pa3262* Frag A+B	pUC19 containing 800-1200bp overlapping region upstream of *pa3262* and the first 45bp of *pa3262* gene termed Frag A. Ligated next to Frag A is Frag B coding for an overlapping region to the last 45bp of *pa3262* and 800-1200bp downstream. Carbenicillin resistance.	This study.
Escherichia coli DH10B: pTS1- Frag A + B *pa3262* knockout construct	Suicide vector designed to integrate into the chromosome via homologous recombination and truncate *pa3262* with two overlapping arms to each side of a target gene. Tetracycline resistance and *SacB* counter selection.	This study.
Escherichia coli DH10B: pUC19 - *pa3717* Frag A	pUC19 containing 800-1200bp overlapping region upstream of *pa3717* and the first 45bp of *pa3717* gene termed Frag A. Carbenicillin resistance.	This study.
Escherichia coli DH10B: pUC19 - *pa3717* Frag A+B	pUC19 containing 800-1200bp overlapping region upstream of *pa3717* and the first 45bp of *pa3717* gene termed Frag A. Ligated next to Frag A is Frag B coding for an overlapping region to the last 45bp of *pa3717* and 800-1200bp downstream. Carbenicillin resistance.	This study.

Escherichia coli DH10B: pTS1- Frag A + B *pa3717* knockout construct	Suicide vector designed to integrate into the chromosome via homologous recombination and truncate *pa3717* with two overlapping arms to each side of a target gene. Tetracycline resistance and *SacB* counter selection.	This study.
Escherichia coli DH10B: pUC19 - *pa4558* Frag A	pUC19 containing 800-1200bp overlapping region upstream of *pa4558* and the first 45bp of *pa4558* gene termed Frag A. Carbenicillin resistance.	This study.
Escherichia coli DH10B: pUC19 - *pa4558* Frag A+B	pUC19 containing 800-1200bp overlapping region upstream of *pa4558* and the first 45bp of *pa4558* gene termed Frag A. Ligated next to Frag A is Frag B coding for an overlapping region to the last 45bp of *pa4558* and 800-1200bp downstream. Carbenicillin resistance.	This study.
Escherichia coli DH10B: pTS1- Frag A + B *pa4558* knockout construct	Suicide vector designed to integrate into the chromosome via homologous recombination and truncate *pa4558* with two overlapping arms to each side of a target gene. Tetracycline resistance and *SacB* counter selection	This study.
Escherichia coli DH10B: pUC19 - *pa4572* Frag A	pUC19 containing 800-1200bp overlapping region upstream of *pa4572* and the first 45bp of *pa4572* gene termed Frag A. Carbenicillin resistance.	This study.
Escherichia coli DH10B: pUC19 - *pa4572* Frag A+B	pUC19 containing 800-1200bp overlapping region upstream of *pa4572* and the first 45bp of *pa4572* gene termed Frag A. Ligated next to Frag A is Frag B coding for an overlapping region to the last 45bp of *pa4572* and 800-1200bp downstream. Carbenicillin resistance.	This study.
Escherichia coli DH10B: pTS1- Frag A + B *pa4572* knockout construct	Suicide vector designed to integrate into the chromosome via homologous recombination and truncate *pa4572* with two overlapping arms to each side of a target gene. Tetracycline resistance and *SacB* counter selection,	This study.
Escherichia coli DH10B : pUC19 - *pa5254* Frag A	pUC19 containing 800-1200bp overlapping region upstream of *pa5254* and the first 45bp of *pa5254* gene termed Frag A. Carbenicillin resistance.	This study.
Escherichia coli DH10B: pUC19 - Frag A + B *pa5254* knockout construct	pUC19 containing 800-1200bp overlapping region upstream of *pa5254* and the first 45bp of *pa5254* gene termed Frag A. Ligated next to Frag A is Frag B coding for an overlapping region to the last 45bp of *pa5254* and 800-1200bp downstream. Carbenicillin resistance.	This study.
Escherichia coli DH10B: pTS1- Frag A + B *pa5254* knockout construct	Suicide vector designed to integrate into the chromosome via homologous recombination and truncate *pa5254* with two overlapping arms to each side of a target gene. Tetracycline resistance and *SacB* counter selection	This study.
PAO1/ Δ*pa3262*	*pa3262* Mip gene knockout. Denoted strain B.	This study
PAO1/ Δ*pa3717*	*pa3717* Mip gene knockout. Denoted strain C.	This study
PAO1/ Δ*pa4558*	*pa4558* Mip gene knockout. Denoted strain D.	This study
PAO1/ Δ*pa4572*	*pa4572* Mip gene knockout. Denoted strain E.	This study

PAO1/ Δ*pa5254*	*pa5254* Mip gene knockout. Denoted strain F.	This study
PAO1/ Δ*pa3262* / Δ*pa3717*	*pa3262* / *pa3717* double Mip gene knockout. Denoted strain K.	This study
PAO1/ Δ*pa3262* / Δ*pa4558*	*pa3262* / *pa4558* double Mip gene knockout. Denoted strain L.	This study
PAO1/ Δ*pa3262* / Δ*pa4572*	*pa3262* / *pa4572* double Mip gene knockout. Denoted strain G.	This study
PAO1/ Δ*pa3262* / Δ*pa5254*	*pa3262* / *pa5254* double Mip gene knockout. Denoted strain M.	This study
PAO1/ Δ*pa3717* / Δ*pa4558*	*pa3717* / *pa4558* double Mip gene knockout. Denoted strain N.	This study
PAO1/ Δ*pa3717* / Δ*pa4572*	*pa3717* / *pa4572* double Mip gene knockout. Denoted strain H.	This study
PAO1/ Δ*pa3717* / Δ*pa5254*	*pa3717* / *pa5254* double Mip gene knockout. Denoted strain O.	This study
PAO1/ Δ*pa4558* / Δ*pa4572*	*pa4558* / *pa4572* double Mip gene knockout. Denoted strain I.	This study
PAO1/ Δ*pa4572* / Δ*pa5254*	*pa4572* / *pa5254* double Mip gene knockout. Denoted strain J.	This study
PAO1/ Δ*pa3262* / Δ*pa3717* / Δ*pa4572*	*pa3262* / *pa3717* / *pa4572* triple Mip gene knockout. Denoted strain Q.	This study
PAO1/ Δ*pa3262* / Δ*pa4572* / Δ*pa5254*	*pa3262* / *pa4572* / *pa5254* triple Mip gene knockout. Denoted strain P.	This study
Escherichia coli : pUC57-*pa3262* + 200bp upstream	Strain housing a construct to complement *pa3262* ordered from GenScript. Vector contains 200bp upstream of *pa3262* where the native promoter sits, and the full *pa3262* gene to complement. Carbenicillin resistance.	This Study
Escherichia coli : pUC57-*pa3717* + 200bp upstream	Strain housing a construct to complement *pa3717* ordered from GenScript. Vector contains 200bp upstream of *pa3717* where the native promoter sits, and a copy of the WT *pa3717* gene to complement. Carbenicillin resistance.	This Study
Escherichia coli : pUC57-*pa4558* + 280bp upstream	Strain housing a construct to complement *pa4558* ordered from GenScript. Vector contains 200bp upstream of *pa4558* where the native promoter sits, and a copy of the WT *pa4558* gene to complement. Carbenicillin resistance.	This Study
Escherichia coli: pUC57-*pa4572* + 84bp upstream	Strain housing a construct to complement *pa4572* ordered from GenScript. Vector contains 200bp upstream of *pa4572* where the native promoter sits, and a copy of the WT *pa4572* gene to complement. Carbenicillin resistance.	This Study

Escherichia coli: pUC57-*pa5254* + 125bp upstream	Strain housing a construct to complement *pa5254* ordered from GenScript. Vector contains 200bp upstream of *pa5254* where the native promoter sits, and a copy of the WT *pa5254* gene to complement. Carbenicillin resistance.	This Study
Escherichia coli: pUC18-Mini-Gm-TN7T- *pa3262*	Strain housing complementation vector designed to integrate a single copy of the WT *pa3262* gene under its native promoter into the chromosome at the neutral *att*Tn7 site. Carbenicillin and Gentamycin resistance.	This Study
Escherichia coli: pUC18-mini-TN7T-Gm-*pa3717*	Strain housing complementation vector designed to integrate a single copy of the WT *pa3717* gene under its native promoter into the chromosome at the neutral *att*Tn7 site. Carbenicillin and Gentamycin resistance.	This Study
Escherichia coli: pUC18-mini-TN7T-Gm *pa4558*	Strain housing complementation vector designed to integrate a single copy of the WT *pa4558* gene under its native promoter into the chromosome at the neutral *att*Tn7 site. Carbenicillin and Gentamycin resistance.	This Study
Escherichia coli: pUC18-mini-TN7T-Gm - *pa4572*	Strain housing complementation vector designed to integrate a single copy of the WT *pa4572* gene under its native promoter into the chromosome at the neutral *att*Tn7 site. Carbenicillin and Gentamycin resistance.	This Study
Escherichia coli: pUC18-mini-TN7T-Gm - *pa5254*	Strain housing complementation vector designed to integrate a single copy of the WT *pa5254* gene under its native promoter into the chromosome at the neutral *att*Tn7 site. Carbenicillin and Gentamycin resistance.	This Study
PA01 : pUC18-mini-TN7T-Gm	Wild-type *P. aeruginosa* containing pUC18-Mini-Gm-TN7T inserted at the neutral *att*Tn7 site. Carbenicillin and Gentamycin resistance.	This Study
PAO1 Δ*pa3262* : pUC18-mini-TN7T-Gm - *pa3262*	*P. aeruginosa* Δ*pa3262* containing pUC18-Mini-Gm-TN7T *pa3262* complementation construct inserted at the neutral *att*Tn7 site. Carbenicillin and Gentamycin resistance.	This Study
PA01 Δ*pa3717*: pUC18-mini-TN7T-Gm -*pa3717*	*P. aeruginosa* Δ*pa3717* containing pUC18-Mini-Gm-TN7T *pa3717* complementation construct inserted at the neutral *att*Tn7 site. Carbenicillin and Gentamycin resistance.	This Study
PA01 Δ*pa4558* : pUC18-mini-TN7T-Gm - *pa4558*	*P. aeruginosa* Δ*pa4558* containing pUC18-Mini-Gm-TN7T *pa4558* complementation construct inserted at the neutral *att*Tn7 site. Carbenicillin and Gentamycin resistance.	This Study
PA01 Δ*pa4572* : pUC18-mini-TN7T-Gm - *pa4572*	*P. aeruginosa* Δ*pa4572* containing pUC18-Mini-Gm-TN7T *pa4572* complementation construct inserted at the neutral *att*Tn7 site. Carbenicillin and Gentamycin resistance.	This Study
PA01 Δ*pa5254* : pUC18-mini-TN7T-Gm - *pa5254*	*P. aeruginosa* Δ*pa5254* containing pUC18-Mini-Gm-TN7T *pa5254* complementation construct inserted at the neutral *att*Tn7 site. Carbenicillin and Gentamycin resistance.	This Study

Escherichia coli: pJH10TS - *pa3262*	Strain housing complementation vector carrying a WT copy of the *pa3262* gene on the pJH10TS plasmid to complement by expression *in trans***.** Tetracycline resistance.	This study.
Escherichia coli: pJH10TS - *pa3717*	Strain housing complementation vector carrying a WT copy of the *pa3717* gene on the pJH10TS plasmid to complement by expression *in trans***.** Tetracycline resistance.	This study.
Escherichia coli: pJH10TS - *pa4558*	Strain housing complementation vector carrying a WT copy of the *pa4558* gene on the pJH10TS plasmid to complement by expression *in trans***.** Tetracycline resistance.	This study.
Escherichia coli: pJH10TS - *pa4572*	Strain housing complementation vector carrying a WT copy of the *pa4572* gene on the pJH10TS plasmid to complement by expression *in trans.* Tetracycline resistance.	This study.
Escherichia coli: pJH10TS - *pa5254*	Strain housing complementation vector carrying a WT copy of the *pa5254* gene on the pJH10TS plasmid to complement by expression *in trans.* Tetracycline resistance.	This study.
PA01 : pJH10TS	Wild-type *P. aeruginosa* housing empty pJH10TS plasmid as a control ensuring the plasmid itself has no effects on phenotypes**.** Tetracycline resistance.	This study.
PAO1 Δ*pa3262* : pJH10TS - *pa3262*	*P. aeruginosa* Δ*pa3262* complemented with pJH10TS – carrying a copy of the WT *pa3262* plasmid by expression *in trans***.** Tetracycline resistance.	This study.
PA01 Δ*pa3717*: pJH10TS - *pa3717*	*P. aeruginosa* Δ*pa3717* complemented with pJH10TS – carrying a copy of the WT *pa3717* plasmid by expression *in trans***.** Tetracycline resistance.	This study.
PA01 Δ*pa4558* : pJH10TS - *pa4558*	*P. aeruginosa* Δ*pa4558* complemented with pJH10TS – carrying a copy of the WT *pa4558* plasmid by expression *in trans***.** Tetracycline resistance.	This study.
PA01 Δ*pa4572* : pJH10TS - *pa4572*	*P. aeruginosa* Δ*pa4572* complemented with pJH10TS – carrying a copy of the WT *pa4572* plasmid by expression *in trans***.** Tetracycline resistance.	This study.
PA01 Δ*pa5254* : pJH10TS - *pa5254*	*P. aeruginosa* Δ*pa5254* complemented with pJH10TS – carrying a copy of the WT *pa5254* plasmid by expression *in trans***.** Tetracycline resistance.	This study.
Escherichia coli: pME6032 - *pa3262*	Strain housing vector carrying a copy of the WT *pa3262* gene for complementation by induced expression under the *lac* promoter**.** Tetracycline resistance.	This study.
Escherichia coli: pME6032 - *pa3717*	Strain housing vector carrying a copy of the WT *pa3717* gene for complementation by induced expression under the *lac* promoter**.** Tetracycline resistance.	This study.
Escherichia coli: pME6032 - *pa4558*	Strain housing vector carrying a copy of the WT *pa4558* gene for complementation by induced expression under the *lac* promoter**.** Tetracycline resistance.	This study.
Escherichia coli: pME6032 - *pa4572*	Strain housing vector carrying a copy of the WT *pa4572* gene for complementation by induced	This study.

	expression under the *lac* promoter**.** Tetracycline resistance.	
Escherichia coli: pME6032 - *pa5254*	Strain housing vector carrying a copy of the WT *pa5254* gene for complementation by induced expression under the *lac* promoter**.** Tetracycline resistance.	This study.
Escherichia coli: pME6032 - *pa4558* / *pa4572*	Strain housing vector carrying a copy of the WT *pa4558* and WT *pa4572* genes for complementation by induced expression under the *lac* promoter**.** Tetracycline resistance.	This study.
Escherichia coli: pME6032 - *pa4572* / *pa4558*	Strain housing vector carrying a copy of the WT *pa4572* and WT *pa4558* genes for complementation by induced expression under the *lac* promoter**.** Tetracycline resistance.	This study.
Escherichia coli: pME6032 - *pa4572* / *pa5254*	Strain housing vector carrying a copy of the WT *pa4572* and WT *pa5254* genes for complementation by induced expression under the *lac* promoter**.** Tetracycline resistance.	This study.
Escherichia coli: pME6032 – *pa5254* / *pa4572*	Strain housing vector carrying a copy of the WT *pa5254* and WT *pa4572* genes for complementation by induced expression under the *lac* promoter**.** Tetracycline resistance.	This study.
PA01 : pME6032	Wild-type *P. aeruginosa* housing empty pME6032 plasmid as a control ensuring the plasmid itself has no effects on phenotypes**.** Tetracycline resistance.	This study.
PAO1 Δ*pa3262* : pME6032 - *pa3262*	*P. aeruginosa* Δ*pa3262* complemented with the pME6032 plasmid carrying a WT copy of *pa3262* for induced expression under the control of the *lac* promoter**.** Tetracycline resistance.	This study.
PA01 Δ*pa3717*: pME6032 - *pa3717*	*P. aeruginosa* Δ*pa3717* complemented with the pME6032 plasmid carrying a WT copy of *pa3717* for induced expression under the control of the *lac* promoter**.** Tetracycline resistance.	This study.
PA01 Δ*pa4558* : pME6032	*P. aeruginosa* Δ*pa4558* housing empty pME6032 plasmid as a control ensuring the plasmid itself has no effects on phenotypes**.** Tetracycline resistance.	This study.
PA01 Δ*pa4558* : pME6032 - *pa4558*	*P. aeruginosa* Δ*pa4558* complemented with the pME6032 plasmid carrying a WT copy of *pa4558* for induced expression under the control of the *lac* promoter**.** Tetracycline resistance.	This study.
PA01 Δ*pa4572* : pME6032	*P. aeruginosa* Δ*pa4572* housing empty pME6032 plasmid as a control ensuring the plasmid itself has no effects on phenotypes**.** Tetracycline resistance.	This study.
PA01 Δ*pa4572* : pME6032 - *pa4572*	*P. aeruginosa* Δ*pa4572* complemented with the pME6032 plasmid carrying a WT copy of *pa4572* for induced expression under the control of the *lac* promoter**.** Tetracycline resistance.	This study.
PA01 Δ*pa5254* : pME6032	*P. aeruginosa* Δ*pa5254* housing empty pME6032 plasmid as a control ensuring the plasmid itself has no effects on phenotypes**.** Tetracycline resistance.	This study.
PA01 Δ*pa5254* : pME6032 - *pa5254*	*P. aeruginosa* Δ*pa5254* complemented with the pME6032 plasmid carrying a WT copy of *pa5254* for	This study.

	induced expression under the control of the *lac* promoter**.** Tetracycline resistance.	
PA01 Δ*pa4558* / Δ*pa4572* : pME6032	*P. aeruginosa* Δ*pa4558* / Δ*pa4572* housing empty pME6032 plasmid as a control ensuring the plasmid itself has no effects on phenotypes**.** Tetracycline resistance.	This study.
PA01 Δ*pa4558* / Δ*pa4572* : pME6032 - *pa4558* / *pa4572*	*P. aeruginosa* Δ*pa4558* / Δ*pa4572* complemented with the pME6032 plasmid carrying a WT copy of both *pa4558* and *pa4572* for induced expression under the control of the *lac* promoter**.** Tetracycline resistance.	This study.
PA01 Δ*pa4558* / Δ*pa4572* : pME6032 - *pa4572* / *pa4558*	*P. aeruginosa* Δ*pa4572* / Δ*pa4558* complemented with the pME6032 plasmid carrying a WT copy of both *pa4572* and *pa4558* for induced expression under the control of the *lac* promoter**.** Tetracycline resistance.	This study.
PA01 Δ*pa4572* / Δ*pa5254* : pME6032	*P. aeruginosa* Δ*pa4572* / Δ*pa5254* housing empty pME6032 plasmid as a control ensuring the plasmid itself has no effects on phenotypes**.** Tetracycline resistance.	This study.
PA01 Δ*pa4572* / Δ*pa5254* : pME6032 - *pa4572* / *pa5254*	*P. aeruginosa* Δ*pa4572* / Δ*pa5254* complemented with the pME6032 plasmid carrying a WT copy of both *pa4572* and *pa5254* for induced expression under the control of the *lac* promoter**.** Tetracycline resistance.	This study.
PA01 Δ*pa4572* / Δ*pa5254* : pME6032 - *pa5254* / *pa4572*	*P. aeruginosa* Δ*pa4572* / Δ*pa5254* complemented with the pME6032 plasmid carrying a WT copy of both *pa5254* and *pa4572* for induced expression under the control of the *lac* promoter**.** Tetracycline resistance.	This study.
Escherichia coli DH10B: pUC19- *pa2371* (ClpV3) Frag A	pUC19 containing 800-1200bp overlapping region upstream of *pa2371* (ClpV3) and the first 45bp of *pa2371* (ClpV3) gene termed Frag A. Carbenicillin resistance.	This study.
Escherichia coli DH10B: pUC19- Frag A + B *pa2371* (ClpV3) knockout construct	pUC19 containing 800-1200bp overlapping region upstream of *pa2371* (ClpV3) and the first 45bp of *pa3271* (ClpV3) gene termed Frag A. Ligated next to Frag A is Frag B coding for an overlapping region to the last 45bp of Pa3262 and 800-1200bp downstream. Carbenicillin resistance.	This study.
Escherichia coli DH10B: pTS1- Frag A + B *pa2371* (ClpV3) knockout construct	Suicide vector designed to integrate into the chromosome via homologous recombination and truncate *pa2371* (ClpV3) with two overlapping arms to each side of a target gene**.** Tetracycline resistance and *SacB* counter selection.	This study.
Escherichia coli DH10B: pUC19- *pa2373* (VgrG3) Frag A	pUC19 containing 800-1200bp overlapping region upstream of Pa2373 (VgrG3) and the first 45bp of *pa2372* (VgrG3) gene termed Frag A. Carbenicillin resistance.	This study.
Escherichia coli DH10B: pUC19- Frag A + B *pa2373* (VgrG3) knockout construct	pUC19 containing 800-1200bp overlapping region upstream of *pa2373* (VgrG3) and the first 45bp of *pa2373* (VgrG3) gene termed Frag A. Ligated next to Frag A is Frag B coding for an overlapping region to the last 45bp of *pa2373* (VgrG3) and 800-1200bp downstream. Carbenicillin resistance.	This study.

Escherichia coli DH10B: pTS1- Frag A + B *pa2373* (VgrG3) knockout construct	Suicide vector designed to integrate into the chromosome via homologous recombination and truncate *pa2373* (VgrG3) with two overlapping arms to each side of a target gene. Tetracycline resistance and *SacB* counter selection.	This study.
Escherichia coli DH10B: pTS1- Frag A + B *pa1662* (ClpV2) knockout construct	Suicide vector designed to integrate into the chromosome via homologous recombination and truncate *pa1662* (ClpV2) with two overlapping arms to each side of a target gene. Construct ordered from TWIST Bioscience. Tetracycline resistance and *SacB* counter selection.	This study.
Escherichia coli DH10B: pTS1- Frag A + B *pa1511* (VgrG2) knockout construct	Suicide vector designed to integrate into the chromosome via homologous recombination and truncate *pa1511* (VgrG2) with two overlapping arms to each side of a target gene. Construct ordered from TWIST Bioscience. Tetracycline resistance and *SacB* counter selection.	This study.
Escherichia coli DH10B: pTS1- Frag A + B *pa0090* (ClpV1) knockout construct	Suicide vector designed to integrate into the chromosome via homologous recombination and truncate *pa0090* (ClpV1) with two overlapping arms to each side of a target gene. Construct ordered from TWIST Bioscience. Tetracycline resistance and *SacB* counter selection.	This study.
Escherichia coli DH10B: pTS1- Frag A + B *pa0091* (VgrG1) knockout construct	Suicide vector designed to integrate into the chromosome via homologous recombination and truncate *pa0091* (VgrG1) with two overlapping arms to each side of a target gene. Construct ordered from TWIST Bioscience. Tetracycline resistance and *SacB* counter selection.	This study.
PA01 Δ*pa2371* (ClpV3)	*P. aeruginosa* strain with *pa2371* (ClpV3) knocked out	This study.
PA01 Δ*pa2373* (VgrG3)	*P. aeruginosa* strain with *pa2373* (VgrG3) knocked out	This study.
PA01 Δ*pa1662* (ClpV2)	*P. aeruginosa* strain with *pa1662* (ClpV2) knocked out	This study.
PA01 Δ*pa1511* (VgrG2)	*P. aeruginosa* strain with *pa1511* (VgrG2) knocked out	This study.
PA01 Δ*pa0090* (ClpV1)	*P. aeruginosa* strain with *pa0090* (ClpV1) knocked out	This study.
PA01 Δ*pa0091* (VgrG1)	*P. aeruginosa* strain with *pa0091* (VgrG1) knocked out	This study.
PA01 Δ*pa2371* (ClpV3) : pME6032- *pa2371* (ClpV3)	*P. aeruginosa* Δ*pa2371* (ClpV3) complemented with the pME6032 plasmid carrying a WT copy of *pa2371* (ClpV3) for induced expression under the control of the *lac* promoter. Tetracycline resistance.	This study.
PA01 Δ*pa2373* (VgrG3) : pME6032 *pa2373* (VgrG3)	*P. aeruginosa* Δ*pa2373* (VgrG3) complemented with the pME6032 plasmid carrying a WT copy of *pa2373* (VgrG3) for induced expression under the control of the *lac* promoter. Tetracycline resistance	This study.

Escherichia coli: pME6032 – *pa2371* (ClpV3)	Strain housing vector carrying a copy of the WT *pa2371* (Clpv3) gene for complementation by induced expression under the *lac* promoter**.** Tetracycline resistance.	This study.
PA01 Δ*pa2371* (ClpV3) : pME6032 *pa2371* (ClpV3)	*P. aeruginosa* Δ*pa2371* (ClpV3) complemented with the pME6032 plasmid carrying a WT copy of *pa2371* (ClpV3) for induced expression under the control of the *lac* promoter**.** Tetracycline resistance.	This study.

Table 2.2 List of all bacterial strains used in this study

2.2.2 Plasmids

Plasmids used for genetic manipulation in this experiment.

Plasmid	Genotype / Description	Reference
pUC19	Plasmid used to align two separate gene fragments to then cut and ligate into pTS1. Carbenicillin resistance.	New England Biolabs
pTS1	Suicide vector designed to integrate into the chromosome via homologous recombination with two overlapping arms to each side of a target gene. Tetracycline resistance.	[187]
pUC18-mini-TN7T-Gm	Vector designed to reintroduce a single copy of a WT gene under its native promoter into the chromosome at the neutral *att*Tn7 site. Carbenicillin and Gentamicin resistance	[188]
pTNS2	Helper plasmid to aid the integration of the pUC18-mini-TN7T-Gm into the *P. aeruginosa* chromosome	[188]
pJH10TS	Plasmid used for complementation of gene knockouts in *P. aeruginosa* by expression *in trans*. Tetracycline resistance.	[187]
pME6032	Vector used for complementation by induced expression under the *lac* promoter. Tetracycline resistance.	[189]
pUC19 - *pa3262* Frag A	pUC19 containing 800-1200bp overlapping region upstream of *pa3262* and the first 45bp of *pa3262* gene termed Frag A. Restriction / ligation XbaI – SacI. Carbenicillin resistance.	This study.
pUC19 - *pa3262* Frag A + Frag B	pUC19 containing 800-1200bp overlapping region upstream of *pa3262* and the first 45bp of *pa3262* gene termed Frag A. Ligated next to Frag A is Frag B coding for an overlapping region to the last 45bp of *pa3262* and 800-1200bp downstream. Restriction / ligation AvrII – EcoRI. Carbenicillin resistance.	This study.
pTS1- *pa3262* Frag A + Frag B	Suicide vector designed to integrate into the chromosome via homologous recombination and truncate *pa3262* with two overlapping arms to each side of a target gene. Restriction / ligation XbaI – BmtI. Tetracycline resistance.	This study.
pUC19 - *pa3717* Frag A	pUC19 containing 800-1200bp overlapping region upstream of *pa3717* and the first 45bp of *pa3717* gene termed Frag A. Restriction / ligation XbaI – SacI. Carbenicillin resistance.	This study.
pUC19 - *pa3717* Frag A + Frag B	pUC19 containing 800-1200bp overlapping region upstream of *pa3717* and the first 45bp of *pa3717* gene termed Frag A. Ligated next to Frag A is Frag B coding for an overlapping region to the last 45bp of *pa3717* and 800-1200bp downstream. Restriction / ligation AvrII – EcoRI. Carbenicillin resistance.	This study.

pTS1- *pa3717* Frag A + Frag B	Suicide vector designed to integrate into the chromosome via homologous recombination and truncate *pa3717* with two overlapping arms to each side of a target gene. Restriction / ligation XbaI – BmtI. Tetracycline resistance.	This study.
pUC19 - *pa4558* Frag A	pUC19 containing 800-1200bp overlapping region upstream of *pa4558* and the first 45bp of *pa4558* gene termed Frag A. Restriction / ligation XbaI – SacI. Carbenicillin resistance.	This study.
pUC19 - *pa4558* Frag A + Frag B	pUC19 containing 800-1200bp overlapping region upstream of *pa4558* and the first 45bp of *pa4558* gene termed Frag A. Ligated next to Frag A is Frag B coding for an overlapping region to the last 45bp of *pa4558* and 800-1200bp downstream. Restriction / ligation AvrII – EcoRI. Carbenicillin resistance.	This study.
pTS1- *pa4558* Frag A + Frag B	Suicide vector designed to integrate into the chromosome via homologous recombination and truncate *pa4558* with two overlapping arms to each side of a target gene. Restriction / ligation XbaI – BmtI. Tetracycline resistance.	This study.
pUC19 - *pa4572* Frag A	pUC19 containing 800-1200bp overlapping region upstream of *pa4572* and the first 45bp of Pa4572 gene termed Frag A. Restriction / ligation XbaI – SacI. Carbenicillin resistance.	This study.
pUC19 - *pa4572* Frag A + Frag B	pUC19 containing 800-1200bp overlapping region upstream of *pa4572* and the first 45bp of *pa4572* gene termed Frag A. Ligated next to Frag A is Frag B coding for an overlapping region to the last 45bp of *pa4572* and 800-1200bp downstream. Restriction / ligation AvrII – EcoRI. Carbenicillin resistance.	This study.
pTS1 - *pa4572* Frag A + Frag B	Suicide vector designed to integrate into the chromosome via homologous recombination and truncate *pa4572* with two overlapping arms to each side of a target gene. Restriction / ligation XbaI – BmtI. Tetracycline resistance.	This study.
pUC19 - *pa5254* Frag A	pUC19 containing 800-1200bp overlapping region upstream of *pa5254* and the first 45bp of *pa5254* gene termed Frag A. Restriction / ligation XbaI – SacI. Carbenicillin resistance.	This study.
pUC19 - *pa5254* Frag A + Frag B	pUC19 containing 800-1200bp overlapping region upstream of *pa5254* and the first 45bp of *pa5254* gene termed Frag A. Ligated next to Frag A is Frag B coding for an overlapping region to the last 45bp of *pa5254* and 800-1200bp downstream. Restriction / ligation AvrII – EcoRI. Carbenicillin resistance.	This study.
pTS1 - *pa5254* Frag A + Frag B	Suicide vector designed to integrate into the chromosome via homologous recombination and truncate *pa5254* with two overlapping arms to each side of a target gene. Restriction / ligation XbaI – BmtI. Tetracycline resistance.	This study.

pUC57 - WT *pa3262* + 200 bp upstream	Construct to complement *pa3262* ordered from GenScript. Vector contains 200bp upstream of *pa3262* where the native promoter sits, and the full *pa3262* gene to complement. Restriction / ligation KpnI – SacI. Carbenicillin resistance.	This study.
pUC57 - WT *pa3717* + 200 bp upstream	Construct to complement *pa3717* ordered from GenScript. Vector contains 200bp upstream of *pa3717* where the native promoter sits, and a copy of the WT *pa3717* gene to complement. Restriction / ligation BamHI – HindIII. Carbenicillin resistance	This study.
pUC57 - WT *pa4558* + 280 bp upstream	Strain housing a construct to complement *pa4558* ordered from GenScript. Vector contains 280bp upstream of *pa4558* where the native promoter sits, and a copy of the WT *pa4558* gene to complement. Restriction / ligation BamHI – HindIII. Carbenicillin resistance.	This study.
pUC57 - WT *pa4572* + 84bp upstream	Construct to complement *pa4572* ordered from GenScript. Vector contains 84bp upstream of *pa4572* where the native promoter sits, and a copy of the WT *pa4572* gene to complement. Restriction / ligation BamHI – HindIII. Carbenicillin resistance.	This study.
pUC57 - WT *pa5254* + 125 bp upstream	Construct to complement Pa5254 ordered from GenScript. Vector contains 125bp upstream of Pa5254 where the native promoter sits, and a copy of the WT Pa5254 gene to complement. Restriction / ligation BamHI – HindIII. Carbenicillin resistance.	This study.
pUC18-mini-TN7T-Gm - WT *pa3262* + 200 bp upstream	Complementation vector designed to integrate a single copy of the WT *pa3262* gene under its native promoter into the chromosome at the neutral *att*Tn7 site. Restriction / ligation KpnI – SacI. Carbenicillin and Gentamycin resistance.	This study.
pUC18-mini-TN7T-Gm - WT *pa3717* + 200 bp upstream	Complementation vector designed to integrate a single copy of the WT *pa3717* gene under its native promoter into the chromosome at the neutral *att*Tn7 site. Restriction / ligation BamHI – HindIII. Carbenicillin and Gentamycin resistance.	This study.
pUC18-mini-TN7T-Gm - WT *pa4558* + 280 bp upstream of beginning of operon	Complementation vector designed to integrate a single copy of the WT *pa4558* gene under its native promoter into the chromosome at the neutral *att*Tn7 site. Restriction / ligation BamHI – HindIII. Carbenicillin and Gentamycin resistance.	This study.
pUC18-mini-TN7T-Gm - WT *pa4572* + 84 bp upstream	Complementation vector designed to integrate a single copy of the WT *pa4572* gene under its native promoter into the chromosome at the neutral *att*Tn7 site. Restriction / ligation BamHI	This study.

	– HindIII. Carbenicillin and Gentamycin resistance.	
pUC18-mini-TN7T-Gm - WT *pa5254* + 125 bp upstream	Complementation vector designed to integrate a single copy of the WT *pa5254* gene under its native promoter into the chromosome at the neutral *att*Tn7 site. Restriction / ligation BamHI – HindIII. Carbenicillin and Gentamycin resistance.	This study.
pJH10TS - WT *pa3262*	Complementation vector carrying a WT copy of the *pa3262* gene on the pJH10TS plasmid to complement by expression *in trans***.** Gibson Assembly. Tetracycline resistance.	This study.
pJH10TS - WT *pa3717*	Complementation vector carrying a WT copy of the *pa3717* gene on the pJH10TS plasmid to complement by expression *in trans.* Gibson Assembly. Tetracycline resistance.	This study.
pJH10TS - WT *pa4558*	Complementation vector carrying a WT copy of the *pa4558* gene on the pJH10TS plasmid to complement by expression *in trans.* Gibson Assembly. Tetracycline resistance.	This study.
pJH10TS - WT *pa4572*	Complementation vector carrying a WT copy of the *pa4572* gene on the pJH10TS plasmid to complement by expression *in trans.* Gibson Assembly. Tetracycline resistance.	This study.
pJH10TS - WT *pa5254*	Complementation vector carrying a WT copy of the *pa5254* gene on the pJH10TS plasmid to complement by expression *in trans.* Gibson Assembly. Tetracycline resistance.	This study.
pME6032 - WT *pa3262*	Complementation vector carrying a copy of the WT *pa3262* gene for complementation by induced expression under the *lac* promoter**.** Restriction / ligation BamHI – KpnI. Tetracycline resistance.	This study.
pME6032 - WT *pa3717*	Complementation vector carrying a copy of the WT *pa3717* gene for complementation by induced expression under the *lac* promoter**.** Restriction / ligation BamHI – KpnI. Tetracycline resistance.	This study.
pME6032 - WT *pa4558*	Complementation vector carrying a copy of the WT *pa4558* gene for complementation by induced expression under the *lac* promoter**.** Restriction / ligation BamHI – KpnI. Tetracycline resistance.	This study.
pME6032 - WT *pa4572*	Complementation vector carrying a copy of the WT *pa4572* gene for complementation by induced expression under the *lac* promoter**.** Restriction / ligation BamHI – KpnI. Tetracycline resistance	This study.
pME6032 - WT *pa5254*	Complementation vector carrying a copy of the WT *pa5254* gene for complementation by induced expression under the *lac* promoter**.**	This study.

	Restriction / ligation BamHI – KpnI. Tetracycline resistance.	
pME6032 - WT *pa4558* / *pa4572*	Complementation vector carrying a copy of the WT *pa4558* and WT *pa4572* genes for complementation by induced expression under the *lac* promoter**.** Restriction / ligation BamHI – KpnI / KpnI – SacI. Tetracycline resistance.	This study.
pME6032 - WT *pa4572* / *pa4558*	Complementation vector carrying a copy of the WT *pa4572* and WT *pa4558* genes for complementation by induced expression under the *lac* promoter**.** Restriction / ligation BamHI – KpnI / KpnI – SacI. Tetracycline resistance.	This study.
pME6032 - WT *pa4572* / *pa5254*	Complementation vector carrying a copy of the WT *pa4572* and WT *pa5254* genes for complementation by induced expression under the *lac* promoter**.** Restriction / ligation BamHI – KpnI / KpnI – SacI. Tetracycline resistance.	This study.
pME6032 - WT *pa5254* / *pa4572*	Complementation vector carrying a copy of the WT *pa5254* and WT *pa4572* genes for complementation by induced expression under the *lac* promoter**.** Restriction / ligation BamHI – KpnI / KpnI – SacI. Tetracycline resistance.	This study.
pUC19 - *pa2371* (ClpV3) Frag A	pUC19 containing 800-1200 bp region upstream of *pa2371* (ClpV3) and the first 45bp of *pa2371* gene termed Frag A. Restriction / ligation XbaI – SacI. Carbenicillin resistance.	This study.
pUC19 - Frag A + B *pa2371* (ClpV3)	pUC19 containing 800-1200bp region upstream of *pa2371* and the first 45bp of *pa2371* gene termed Frag A. Ligated next to Frag A is Frag B coding for the last 45 bp of *pa2371* and 800-1200 bp downstream. Restriction / ligation AvrII – EcoRI. Carbenicillin resistance.	This study.
pTS1 - Frag A + B *pa2371* (ClpV3)	Suicide vector designed to integrate into the chromosome via homologous recombination and truncate *pa2371* with two overlapping arms to each side of a target gene. Construct ordered from TWIST Bioscience. Restriction / ligation XbaI – BmtI. Tetracycline resistance and *SacB* counter selection.	This study.
pUC19 - *pa2373* (VgrG3) Frag A	Plasmid containing homologues upstream region and first 45 base pairs of the *pa2373* (VgrG3) gene, termed Fragment (Frag) A. Restriction / ligation XbaI – SacI. Carbenicillin resistance.	This study.
pUC19 - Frag A + B *pa2373* (VgrG3)	pUC19 containing 800-1200 bp region upstream of *pa2373* and the first 45bp of *pa2373* gene termed Frag A. Ligated next to Frag A is Frag B coding for the last 45 bp of *pa2373* and 800-1200 bp downstream Restriction / ligation AvrII – EcoRI. Carbenicillin resistance.	This study.
pTS1 - Frag A + B *pa2373* (VgrG3)	Suicide vector designed to integrate into the chromosome via homologous recombination and truncate *pa2373* with two overlapping arms	This study.

	to each side of a target gene. Construct ordered from TWIST Bioscience. Restriction / ligation XbaI – BmtI. Tetracycline resistance and *SacB* counter selection.	
pTS1 - Frag A + B *pa1662* (ClpV2)	Suicide vector designed to integrate into the chromosome via homologous recombination and truncate *pa1662* with two overlapping arms to each side of a target gene. Construct ordered from TWIST Bioscience. Restriction / ligation XbaI – BmtI. Tetracycline resistance and *SacB* counter selection	This study.
pTS1 - Frag A + B *pa1511* (VgrG2)	Suicide vector designed to integrate into the chromosome via homologous recombination and truncate *pa1511* with two overlapping arms to each side of a target gene. Construct ordered from TWIST Bioscience. Restriction / ligation XbaI – BmtI. Tetracycline resistance and *SacB* counter selection	This study.
pTS1 - Frag A + B *pa0090* (ClpV1)	Suicide vector designed to integrate into the chromosome via homologous recombination and truncate *pa0090* with two overlapping arms to each side of a target gene. Construct ordered from TWIST Bioscience. Restriction / ligation XbaI – BmtI. Tetracycline resistance and *SacB* counter selection	This study.
pTS1 - Frag A + B *pa0091* (VgrG1)	Suicide vector designed to integrate into the chromosome via homologous recombination and truncate *pa0091* with two overlapping arms to each side of a target gene. Construct ordered from TWIST Bioscience. Restriction / ligation XbaI – BmtI. Tetracycline resistance and *SacB* counter selection	This study.
pME6032 - WT *pa2371* (ClpV3)	Complementation vector carrying a copy of the WT *pa2371* (ClpV3) gene for complementation by induced expression under the *lac* promoter**.** Restriction / ligation BamHI – KpnI. Tetracycline resistance	This study.

Table 2.3 List of all the plasmids created and used in this study

2.2.3 Oligonucleotides

All Oligonucleotides were ordered from Integrated DNA Technologies (IDT) and designed on Geneious Prime software.

Primer Name	Primer sequence	Tm (°C)
*pa4558*_A_FW	GCTCTAGACACACCGCCGAGGAAGTCTG	60.5
*pa4558*_A_RV	CGGAGCTCCCTAGGGCGGCTCTCCTGGCCAAT	60.5
*pa4558*_B_FW	CCCCTAGGGTCGAGATCATCGACGTCCA	59.3
*pa4558*_B_RV	CGGAATTCGCTAGCCAGTTCCTGCTCCTCCGAC	61.0
*pa4572*_A_FW	TGCAAGCTTCAGCAAGGCTATATCGAAGCCG	62.1
*pa4572*_A_RV	TGCTCTAGAGGCTTCGTCAGTGGTCAGGTTG	64.0
*pa4572*_B_FW	CGGTCTAGAGTGTTCGACGTCGAACTGCTGGAAATC	66.5
*pa4572*_B_RV	CGGGGTACCGAGGACACCGTGCAGTTGCTGAAG	66.1
*pa5254*_A_FW	CGTCTAGACTTCGCAGCAGGTTTGGTCG	61.4
*pa5254*_A_RV	GCGAGCTCGGATCCTACGGCATAGGCCAGTTCGTC	61.8
*pa5254*_B_FW	CCAAGCCTCAAGCCTATGGTCACGAAGGC	61.8
*pa5254*_B_RV	CGGCTAGCCTTCTCCATCGTGCAGTTCCTC	62.1
*clpv3*_A_FW	GGCGGCTCTAGAGCCGACAAGGAGGTGAAGACA	61.8
*clpv3*_A_RV	GGCGGCGAGCTCCCTAGGACAGTCAGGGTTGAGGCGG	61.0
*clpv3*_B_FW	AAAACGCCTAGGCAGGTGACGGGGAAGTTGG	61.0
*clpv3*_B_RV	AAAACGGAATTCGCTAGCCGAAGACCTGGGTGACGATGT	61.8
*vgrg3*_A_FW	AAGCCTTCTAGAGAGGAGTCTATATGAACTATTTATTCGTGCC	64.2
*vgrg3*_A_RV	AATGCTGAGCTCCCCGGGCGGAAAGCGTCTGGGGGTAAAG	64.0
*vgrg3*_B_FW	AAAACGCCCGGGGGTGGACGGCGGCGGCAT	65.1
*vgrg3*_B_RV	AAAACGGAATTCGCTAGCCGCGCTCGGGAGGACGG	64.8
*clpv2*_A_FW	AAGCCTTCTAGAGCCAGCATCAACTCCTACCACG	64.0
*clpv2*_A_RV	AATGCTGGTACCCCTAGGCAGCGCCTGAATCAGTTGTTGCA	62.4
*clpv2*_B_FW	AATGCTCCTAGGCTGGGCGAGCACGGGC	62.0

*clpv2*_B_RV	AAAACGGAGCTCGCTAGCCTGGTTGGAAGGAAATCTCGATCATG	63.2
*vrgg2*_A_FW	AAGCCTAAGCTTGCTCAGGCACCAATCCATACCA	62.1
*vgrg2*_A_RV	AATGCTTCTAGACCTAGGCCTGGGCAAGGGCAGTCACAA	63.7
*vgrg2*_B_FW	AATGCTCCTAGGCGAGTTTCCCTTCGCCGACG	63.4
*vgrg2*_B_RV	AAGCCTGGATCCGCTAGCCCGCACCTCCGGTTCCC	62.4
*clpv1*_A_FW	AAGCCTTCTAGAGTGTTCCTTCTCGGCGCTGTGTT	64.2
*clpv1*_A_RV	AATGCTGAGCTCCCTAGGCCTTGTAGGCCAGGCTGTTCAG	64.0
*clpv1*_B_FW	AAAACGCCTAGGCCTCACCCGCATGCTCGAAGG	65.7
*clpv1*_B_RV	AAAACGGAATTCGCTAGCCGTACAGCGGGTAGTCGGCC	65.5
*vgrg1*_A_FW	AAGCCTTCTAGAGGCCAGGACCACGCGCTG	65.1
*vgrg1*_A_RV	GGCGGCGGATCCCCTAGGACGTCCGGCCCCAGCGG	64.8
*vgrg1*_B_FW	AAAACGCCTAGGGTCGACGCCAAGGGCAAGGG	65.5
*vgrg1*_B_RV	AAAACGGGTACCGCTAGCCCGGCAGCGGTGACCTCC	65.1

Table 2.4. Gene knockout construct primers. Primers used for the creation of knockout constructs in PA01. Restriction sites are highlighted in red.

Primer Name	Primer sequence	Tm (°C)
pa4558 colony PCR check FW	TAGATGGCACCACCGTTGCTGG	64.0
pa4558 colony PCR check RV	GGTGCCGCCTTCAGCTTCCT	63.4
pa4572 colony PCR check FW	GGCATGGCGCGGACCTTCAC	65.5
pa4572 colony PCR check RV	GACAGGGTTCCGGTGGAGTTGG	65.8
pa5254 colony PCR check FW	CGCTTGTTTGCACGCACCTTCC	64.0
pa5254 colony PCR check RV	CGAGCAACTGACGGCGGAAG	63.4
clpv3 colony PCR check FW	CTCAGTCTCGGTCCTCTGGATG	64.0
clpv3 colony PCR check FV	CGGTGGTTCTCTGCCGCG	62.8
vgrg3 colony PCR check FW	GAAAAGCTGGAGTTCCTCGACAATAC	63.2
vgrg3 colony PCR check RV	GCGGCCTGTCGTGCTCG	62.4
clpv2 colony PCR check FW	CGCATTCATATCCGCCAGCTCGA	64.2
clpv2 colony PCR check RV	AGATACAACTGGCTCAGCTCGCA	62.4
vgrg2 colony PCR check FW	CGGCCTCCAGACCTCGC	62.4
vrgg2 colony PCR check RV	CCGGTGGCAGGGTGAGTTG	63.1
clpv1 colony PCR check FW	CCTGGGTCGCCGAATACCTGG	65.7
clpv1 colony PCR check RV	TACCAGGTGCAGCTCGTAGGCG	65.8
vgrg1 colony PCR check FW	CTGATCGTCTCGCGCTGCAC	63.4
vgrg1 colony PCR check RV	CCGGAACTCTCCAGGGCCC	65.3

Table 2.5. Gene knockout colony PCR primers. Primers used for colony PCR and sequencing to check the correct construction of knockout vectors and deletion of target genes

Primer Name	Primer sequence	Tm (°C)
*pa3262*_FW_ pJH10TS	ATTCTAGTCAATTGGTCATTAATTAACTGCGCTAGCATGAAAC AACATCGGTTGGCAGCAG	63.0
*pa3262*_RV_ pJH10TS	AATGAGGCGGTCACGCTCTCCAGCGAGCTCTCTAGATTACTTC TTCGCGTCCGCCTTGG	64.2
*pa3717*_FW_ pJH10TS	ATTCTAGTCAATTGGTCATTAATTAACTGCGCTAGCATGAACG ACCGACTGCAGATCGAA	62.7
*pa3717*_RV_ pJH10TS	AATGAGGCGGTCACGCTCTCCAGCGAGCTCTCTAGATCAGTCG TCGCGGGTGAGC	63.1
*pa4558*_FW_ pJH10TS	ATTCTAGTCAATTGGTCATTAATTAACTGCGCTAGCATGGCTG AGTCGACCCGCATT	61.8
*pa4558*_RV_ pJH10TS	AATGAGGCGGTCACGCTCTCCAGCGAGCTCTCTAGATCAGGCC GGCTGGACGTC	62.8
*pa4572*_FW_ pJH10TS	ATTCTAGTCAATTGGTCATTAATTAACTGCGCTAGCATGAGCG AACTCAACCTGACCACTGA	64.8
*pa4572*_RV_ pJH10TS	AATGAGGCGGTCACGCTCTCCAGCGAGCTCTCTAGATCAGAG GATTTCCAGCAGTTCGACGTC	66.5
*pa5254*_FW_ pJH10TS	ATTCTAGTCAATTGGTCATTAATTAACTGCGCTAGCATGCCGC GCCGCCTTTTCATCGGTCTTTTT	69.5
*pa5254*_RV_ pJH10TS	AATGAGGCGGTCACGCTCTCCAGCGAGCTCTCTAGATCAGCG GAAGCCCAGCAGGTCGATCT	69.5
*pa4558*_FW_ pME6032	AAGCCTGAATTCATGGCTGAGTCGACCCGCATT	61.8
*pa4558*_ RV_ pME6032	GGCGGCGGTACCTCAGGCCGGCTGGACGTC	62.8
*pa4558*_FW_ pME6032_ Linear (ligate into pME6032 *pa4572*)	AAGCCTGGTACCATGGCTGAGTCGACCCGCATT	61.8
*pa4558*_RV pME6032_ Linear (ligate into pME6032 *pa4572*)	GGCGGCCTCGAGTCAGGCCGGCTGGACGTC	62.8
*pa4572*_FW_ pME6032	AAGCCTGAATTCATGAGCGAACTCAACCTGACCACTGA	64.8
*pa4572*_RV_ pME6032	GGCGGCGGTACCTCAGAGGATTTCCAGCAGTTCGACGTC	66.5
*pa4572*_FW_ pME6032 Linear (ligate into pME6032	AAGCCTGGTACCATGAGCGAACTCAACCTGACCACTGA	64.8

pa4558 or PME6032 *pa5254*)		
*pa4572*_RV_ pME6032_ Linear (ligate into pME6032 *pa4558* or pmE6032 *pa5254*)	GGCGGCCTCGAGTCAGAGGATTTCCAGCAGTTCGACGTC	66.5
*pa5254*_FW_ pME6032	AAGCCTGAATTCATGCCGCGCCGCCTTTTCATCGGTCTTTTT	69.5
*pa5254*_RV_ pME6032	GGCGGCGGTACCTCAGCGGAAGCCCAGCAGGTCGATCT	69.5
*pa5254*_FW_ pME6032 Linear (ligate into pME6032 *pa4572*)	AAGCCTGGTACCATGCCGCGCCGCCTTTTCATCGGTCTTTTT	69.5
*pa5254*_RV_ pME6032 Linear (ligate into pME6032 *pa4572*)	GGCGGCCTCGAGTCAGCGGAAGCCCAGCAGGTCGATCT	69.5

Table 2.6. Gene amplification primers. Primers used to amplify genes and insert into complementation vectors. Restriction sites are highlighted in red.

Primer Name	Primer sequence	Tm (°C)
pJ10TS FW External Sequencing	GTAAATCACTGCATAATTCGTGTCGC	61.6
pJ10TS RV External Sequencing	CAGCCGCCGGTACTCCC	62.4
pME6032 FW External Sequencing	GGTCATCCACCGGATCAATTC	59.8
pME6032 RV External Sequencing	GCACTCCCGTTCTGGATAATG	59.8
pTS1 FW External Sequencing	CATTCTCTGTATTTTTTATAGTTTCTGTTGCAT	60.8
pTS1 ReV External Sequencing Primer	CAAGATCCTGGTATCGGTCTGC	62.1
pUC18-mini-Gm FW Sequencing	TCAATTCGATCCATTGCTGTTGAC	59.3
pUC18-mini-Gm RV Sequencing	GGGTGGAAATGGAGTTTTTAAGGA	59.3
pUC18-mini-Gm FW *att*TN7 insertion site	CTGTGCGACTGCTGGAGCTGA	63.7
pUC18-mini-Gm RV *att*TN7 insertion site	GAGAGCACGTCGCCGATGTGC	65.7
pUC19 FW Sequencing	GCCTCTTCGCTATTACGCCAG	61.8
pUC19 RV Sequencing	GCTTCCGGCTCGTATGTTGTG	61.8

Table 2.7. Vector colony PCR primers. Primers used for colony PCR and sequencing of vectors to check the correct assembly of constructs in various plasmids.

2.3 Culture media, solutions and buffers

2.3.1 Culture media

2.3.1.1 Liquid media

All liquid media was made using the recipes below added into 900 ml distilled water

LB (Miller) Broth

Made using Either

LB Broth Miller powder	25 g/L

Or

Tryptone	10 g/L
Yeast extract	5 g/L
Sodium chloride	10 g/L

SOC

Tryptone	20 g/L
Yeast extract	5 g/L
NaCl	0.58 g/L
$MgCl_2$	2 g/L
$MgSO_4$	2.5 g/L

Succinate media

K_2HPO_4	12 g/L
KH_2PO_4	6 g/L
$(NH_4)_2PO_4$	2 g/L
Succinate acid	8 g/L
NaOH	2.2 g/L

Adjusted to pH 7.0 using whole NaOH pellets, one pellet at a time.

2.3.1.2 Solid media

For all solid media, unless stated otherwise, 10 g/L Formedium agar was added to each recipe below and added to sterile water. All media was autoclaved.

LB (Miller) agar

Made using either

LB Broth Miller powder	25 g

Or

Tryptone	10 g/L
Yeast extract	5 g/L
Sodium chloride	10 g/L

Haemolysis agar

Columbia blood agar	39 g/L
Defibrillated horse blood	50 ml

Protease agar

Mueller-Hinton	38 g/L
Milk solution (**2.3.2)**	100 ml/L

Swarming agar

Formedium agar	5.5 g/L (for final conc. of 0.55%)
5 x M8 solution (**2.3.2)**	250 ml/L
50% glucose	2 ml/L (final conc. of 02%)
20% Casamino acids (**2.3.2)**	12.5 ml/L (final conc. of 0.5%)
1M $MgSO_4$	0.5ml/L

Twitching agar (1% LB-Lennox agar)

Formedium agar	10 g/L (for final conc. of 1%)
Tryptone	10 g/L
Yeast extract	5 g/L
NaCl	5 g/L

2.3.2 Solutions

All solutions were made up in distilled water and autoclaved unless stated otherwise.

Milk solution

Kendamil lactose free powder	10 g/L

5 x M8 solution

Na_2HPO_4	30 g/L
KH_2PO_4	15 g/L
NaCl	2.5 g/L

20% Casamino acid Solution

Omnipur Casamino acids (Calbiochem)	200 g/L

TM developer solution – store at 4°C

Distilled water	400 ml
100% glacial acetic acid	100 ml
MeOH	500 ml

2.3.3 Buffers

SET buffer – to make up 20 ml add;

5M NaCL	133 µl
0.25M EDTA	2 ml
1M Tris	400 µl
RNase	5 mg/ml
Lysozyme	100 µl 50 mg/ml
H_2O	17.47 ml

TE buffer

1M Tris	20 µl
0.25M EDTA	8 µl
H_2O	1972 µl

2.3.4 Antibiotics

The stocks and working concentrations of antibiotics used in this study are listed in Table 2.5.

Antibiotic	Abbr.	Solvent	Stock concentration (mg/ml)	Working concentration (μg/ml)
Carbenicillin	Carb	ddH_2O	100	100
Kanamycin	Kan	ddH_2O	50	50
Gentamicin	Gent	ddH_2O	10	10
Tetracycline	Tet	Ethanol	100	100

Table 2.8. Antibiotic stocks used in this study

2.4 Cloning, screening and transformations

2.4.1 Preparation of *P. aeruginosa* genomic DNA (gDNA)

Wild-type *P. aeruginosa* was taken from -80°C glycerol stock and grown up in 40 ml LB culture O/N at 37°C. The next day this culture was transferred to a 50 ml falcon and centrifuged at 4000 rpm for 15 min. The pellet was then resuspended in 5 ml SET buffer (**2.3.3**) and incubated for 30-60 min at 37°C. Once the incubation period was over, 140 μl of proteinase K (at a concentration of 20 mg/ml) was added along with 600 μl 10% SDS, this was mixed carefully by inverting 25 times. This was then incubated at 55°C for 2 h, occasionally inverting. Following incubation, 2 ml of 5M NaCL was added and mixed thoroughly by inversion and allowed to cool in a water bath set at 37°C. Once cooled, 5 ml chloroform was added using a glass pipette and mixed by inversion for 30 min at room temperature. The sample was then centrifuged for 15 min at 4500 *x g*. The supernatant was then carefully transferred to a fresh falcon tube, avoiding any white precipitate. Once in the falcon 0.6 volume of isopropanol was added, this was mixed by inversion. The DNA was then spooled into a sealed Pasteur pipette and rinsed in 70% EtOH. This was then left to airdry and then dissolved in 1-2 ml TE buffer and stored at 4°C. The quality and concentration was checked on a nanodrop looking for an absorbance ratio of 260/280 between 1.8-2.0. Samples had to be within the 1.8-2.0 parameters to pass quality control.

2.4.2 Polymerase Chain Reaction (PCR)

All PCR reactions were performed in a SimpliAmp Thermal Cycler from Thermo Fischer Scientific. PCR products for cloning and / or sequencing were generated using Phusion polymerase in 25 μl reaction volumes as per the manufacturers protocol (Table 2.6). Either *P. aeruginosa* gDNA or a vector containing DNA of interest was used as a template. PCR cycles used can be found in Table 2.6. varying annealing and elongation temperatures were used based on oligonucleotide

length and overall length of fragment to be amplified. PCR products were cleaned up using QIAquick gel extraction kit (**2.4.5**) and sent for sequencing to confirm correct amplification.

Phusion PCR reaction mix		
Component	**Volume (µl)**	**Final concentration**
DNA template	1	~50 ng
Forward primer	0.5	0.5 µM
Reverse primer	0.5	0.5 µM
dNTPs	1	200 µM
DMSO	1.25	0.03%
5x Phusion HF buffer	5	1x
Phusion polymerase	0.5	0.5 U

Phusion PCR programme			
Cycle	**Temperature ^{o}C**	**Time**	**Repeat (x)**
Initial denaturation	95	3 min	1
Denaturation	98	20 sec	30-40
Annealing	55-69	30 sec	
Elongation	72	45 sec / 1Kb	
Final elongation	72	7 min	1
Cooling	12	1 min	1

Table 2.9. Phusion PCR reaction mix and programme for amplification of DNA fragments. PCR performed in 25 µl volume, made up to this volume using sterile water.

2.4.3 Colony Polymerase Chain Reaction (PCR)

Colony PCR was used to screen colonies for introduced gene deletions or verify whether DNA was successfully introduced to the chromosome or via plasmid uptake. A single colony was picked into 50 µl of sterile water and mixed, leaving to stand for 10 min. Following the 10 min standing period, 10 µl of the colony mixed in the sterile water was taken out and used as DNA template in the reaction. Appropriate volumes of components were added to the PCR tube and reactions were ran (Table 2.7). PCR products were cleaned up using QIAquick gel extraction kit (**2.4.5**) and sent for sequencing to confirm gene deletions or introduced DNA.

Colony PCR reaction mix		
Component	**Volume (µl)**	**Final concentration**
DNA template	10	Cell solution
Forward primer	0.5	1.0 µM
Reverse primer	0.5	1.0 µM
GoTaq G2 Master Mix	10	2x

Colony PCR programme			
Cycle	**Temperature °C**	**Time**	**Repeat (x)**
Initial denaturation	95	3 min	1
Denaturation	98	20 sec	30-40
Annealing	55-69	30 sec	
Elongation	72	1 min / 1Kb	
Final elongation	72	7 min	1
Cooling	12	1 min	1

Table 2.10. Colony PCR reaction mix and programme. PCR performed in 25µl volume, made up to this volume using sterile water.

2.4.4 Agarose gel electrophoresis

Agarose gels were prepared by adding agarose powder to TBE buffer for a final concentration between 0.8 and 1.0% depending on the size of the DNA fragment amplified. Either 50 ml, 100 ml or 150 ml gels were poured depending on number of samples, with 2 µl, 4 µl or 6 µl of ethidium bromide added respectively, prior to pouring the gel for visualisation of DNA under ultraviolet (UV). Before DNA samples were added to the gel, loading dye was added to each sample. This step was not necessary for colony PCR samples. 1kb DNA ladder (NEB) was ran in a separate well for size fragment analysis. Gel electrophoresis was performed in TBE buffer using a PowerPac™ Universal Power Supply (Bio-Rad) ran between 90-120 V until DNA fragments were deemed sufficiently separated. Gels were visualised using a UV Transilluminator GBox Syngene.

2.4.5 Cloning of DNA fragments – restriction digestion & ligation

Cloning was performed using standard molecular biology protocols. Following analysis of agarose gels, amplified PCR fragments or vector DNA were cut out of the gel and purified using QIAquick gel extraction kit. Eluted DNA was then digested using appropriate restriction endonucleases. The restriction digests were performed in a total volume of 50 µl and left for 2 h at 37°C (table 2.8). Digests were purified from an agarose gel to size-separate fragments following electrophoresis. These fragments were cut from the gel and purified using QIAquick gel extraction kit. Once eluted, ligations were performed overnight (O/N) at room temperature in a total volume of 20 µl (Table 2.8). Ligation mixes were directly used for transformation of electro competent *E. coli* DHB10 cells (**2.4.6**).

Restriction digest mix	
Component	**Volume (µl)**
Nuclease free H_2O	11
DNA	30
Restriction enzyme A	2
Restriction enzyme B	2
CutSmart Buffer	5

Table 2.11. Restriction digest reaction mix

Ligation reaction mix	
Component	**Volume (µl)**
Nuclease free H_2O	7
DNA fragment	8
Vector	2
T4 Ligase	1
T4 Ligase buffer	2

Table 2.12. Ligation reaction mix

2.4.6 Cloning of DNA fragments – Gibson Assembly

Gibson assembly was used when ligations were proving unsuccessful, particularly when using the pJH10TS plasmid for gene complementation. Mip genes were amplified from *P. aeruginosa* gDNA using oligonucleotides containing the beginning and end of the Mip genes allowing full WT Mip gene amplification, along with a 20 bp overlapping region up and downstream of the pJH10TS multiple cloning site backbone. The pJH10TS plasmid was linearised with BmtI and cleaned up using QIAquick gel extraction kit. Components of the Gibson reaction were assembled following manufacturer's instructions in containing 10 µl of Gibson Assembly (2X) Master mix (NEB) and a suitable ratio of PCR products and linearised vector, with a two-to-five-fold excess of insert over vector and made up to 20 µl with nuclease free H_2O. The Gibson reaction was performed in a thermocycler at 50°C for 60 min followed by transformation into *E. coli* DH10B cells by electrorotation.

2.4.7 Transformation of electrocompetent *E. coli* DHB10 cells

Following overnight ligation of DNA fragments and vector, 2 µl of the ligation mix was added to a 1.5 ml Eppendorf tube containing 50 µl of electro competent *E. coli* DHB10 cells (Thermo Fisher) cooled on ice. The Eppendorf was flicked to mix the contents and left to stand for 3 min. This was then transferred to an EP-102 2 mm electroporation cuvette (Cell Projects) and electroporated at 2.5 V for 5 sec**.** Cells were immediately transferred to a fresh sterile 1.5 ml Eppendorf and recovered in 150 µl SOC media and incubated at 37^{o}C and 250rpm. After 1 h, 50 µl of cells were plated on LB agar supplemented with the appropriate antibiotic to select for colonies. The plates were incubated O/N at 37^{o}C. On the following day several colonies were screened via colony PCR and DNA sequencing to verify successful cloning of the gene and transformation of the plasmid. Glycerol stocks were made of verified colonies and stored at -80^{o}C.

2.4.8 Plasmid purification and sequencing

A sample of glycerol stocks made from verified colonies were inoculated into in 10 ml LB liquid culture with suitable antibiotic and grown O/N at 37^{o}C, 250 rpm. Plasmids were extracted and purified using QIAprep Spin Miniprep Kits. Once plasmid DNA (or purified PCR products) was purified it was sent for Sanger sequencing at Eurofins Genomics. Following the Eurofins genomic protocol, 15 µl of DNA sample and 2 µl of primer were mixed in a sequencing tube. All plasmid constructs, generated in this project were sent for confirmatory Sanger sequencing, as were PCR products generated to verify gene deletions introduced into strains as well as gene complementation strains.

2.5 Transformation and mutagenesis of *P. aeruginosa*

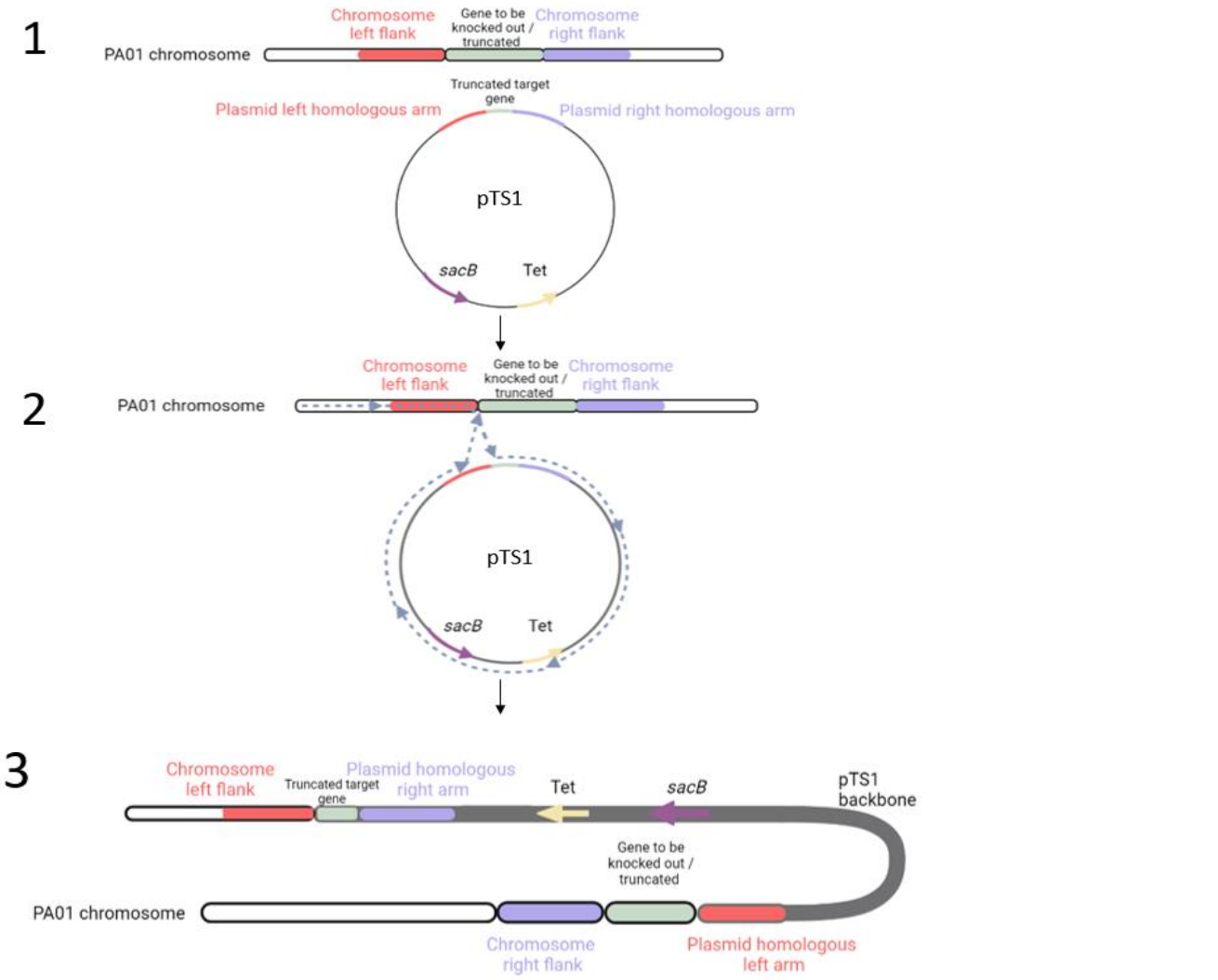
2.5.1 pUC19 & pTS1 construct design

For gene knockouts the pTS1 suicide plasmid was used. The pTS1 mediated knockout constructs, primers were designed to amplify 800-1300 bp flanking regions (up and down stream) of the gene to be knocked out. These two regions were termed ‘Fragment A’ and ‘Fragment B’ respectively. To minimise possible polar effects of a gene deletion, 9-51 bp at the beginning and end of the gene were left, creating a truncated version of the gene to be deleted in multiples of 3 to ensure the reading frame remained intact. Fragment A and Fragment B were aligned in pUC19 due to occasional difficulties in cloning directly into pTS1. Fragment A was ligated into the pUC19 multiple cloning site (MCS) between XbaI and SacI, with AvrII placed just before the SacI restriction site remaining intact for the second ligation involving Frag B. Fragment B was ligated between the introduced AvrII site and the pUC19 EcoRI site, with BmtI place just before the EcoRI restriction site. Once Fragment A and B were assembled it created a truncated copy

of the target gene to be deleted, with homologous flanking regions to facilitate homologous recombination. Once Frag A and Frag B were confirmed and aligned in pUC19, the linear fragment was cut out and ligated into pTS1 between the XbaI and BmtI restriction sites.

2.5.2 Transformation and mutagenesis of *P. aeruginosa* using pTS1

Once the pTS1 gene deletion construct was assembled and confirmed by sequencing, it was used to transform *P. aeruginosa.* A 10 ml culture of *P. aeruginosa* was inoculated and incubated O/N at 37°C, 250 rpm. Following overnight incubation, the culture was centrifuged for 10 min at 4000 *x g*. The *P. aeruginosa* pellet was the resuspended in 2 ml of 300 mM sterile filtered sucrose solution. This was centrifuged at 13,300 rpm, and the supernatant poured away leaving a *P. aeruginosa* cell pellet which was resuspended in 2 ml of 300 mM sterile filtered sucrose solution and centrifuged. The supernatant was again poured away, and the pellet resuspended in the ~100 µl of flowback that remained. The resuspended cells were mixed with 15 µl of a given pTS1 construct and transferred into an EP-102 2 mm electroporation cuvette and electroporated at 2.5 V for 5 sec. The cells were then transferred into a fresh sterile 1.5 ml Eppendorf tube and recovered in 150 µl of SOC media. This was incubated for 2 h in a shaking incubator at 42°C, 250 rpm. Cells were then plated on LB agar plates supplemented with Tet^{100} to select for single crossover colonies. These plates were incubated at 37°C O/N. The following day, single colonies were re-streaked on fresh LB Tet^{100} agar plates and incubated at 37°C O/N to ensure true tetracycline resistance. Single colonies from the re-streak plates were then picked and put into 10 ml of fresh LB liquid media with no antibiotic and incubated O/N in a shaking incubator at 37°C, 250 rpm. This step is to facilitate a double crossover event, resulting in loss of the pTS1 backbone. The following day, 10^{-5} dilutions of the overnight cultures were prepared and 50 µl was plated on LB 10% sucrose plates. The pTS1 plasmid possess a copy of the *SacB* gene which encodes levansucrase, meaning if the double crossover event has not occurred, when plated on sucrose the cells that have not undergone the double crossover event will not grow due to conversion of sucrose to levan, a toxic compound. These plates were incubated at 37°C O/N. The following day colonies growing on the plate should have either the WT genotype or encode a truncated gene. Colony PCR followed by Sanger sequencing of gel purified PCR products was used to confirm gene deletions. Glycerol stocks were prepared for three separate verified clones per gene deletion attempt and stored at -80°C.



First homologous recombination event that incorporates the entire pTS1 vector including the antibiotic resistance cassette, sucrose sensitivity and truncated target gene

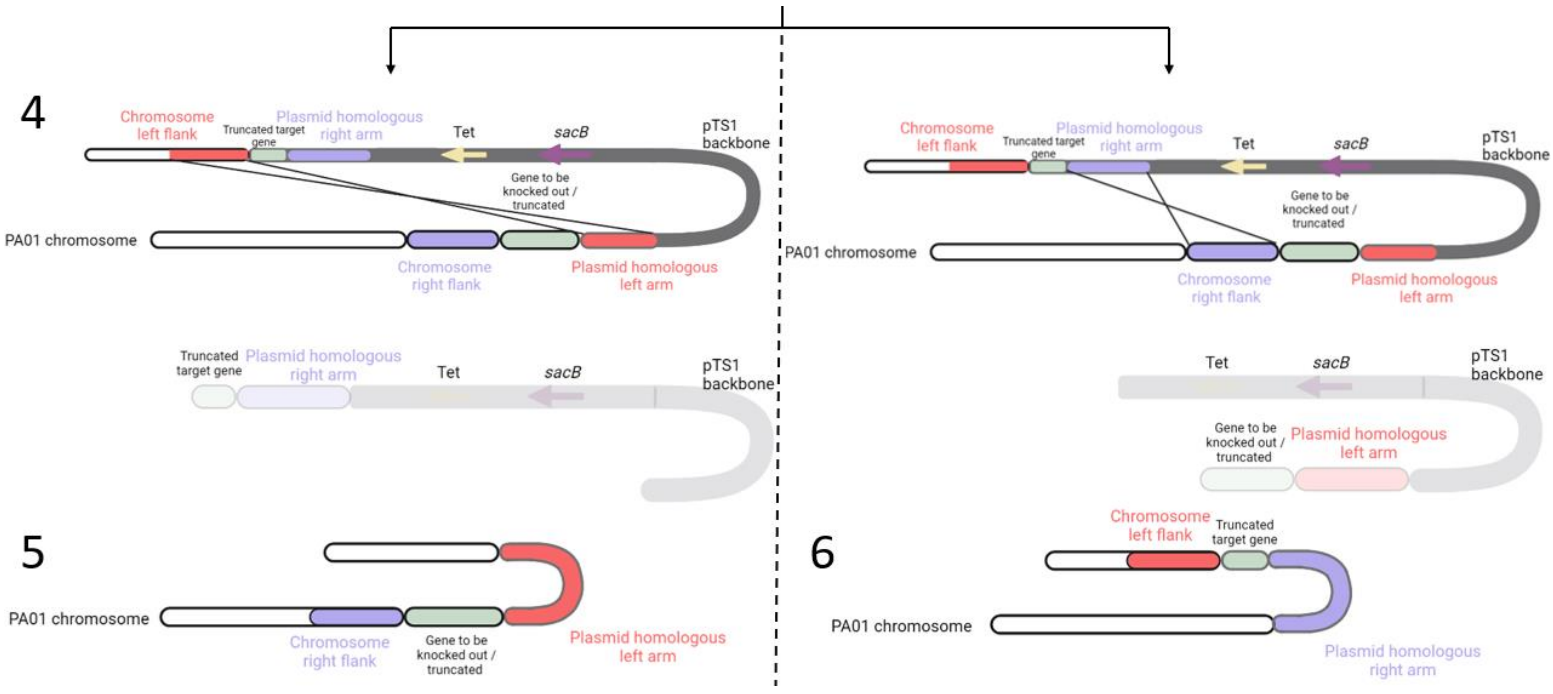


Figure 2.1. Image depicting the gene deletion process using the pTS1 suicide knockout vector in *P. aeruginosa*. 1) PA01 cells are transformed by elecroporation with a pTS1 vector housing homologous arms to either side of a gene to be truncated or deleted. 2) The first homolgous recombination event occurs during DNA replication where there is a single strand break, the pTS1 vector will be incorporated into the chromosome at a region on the chromosome homologous to the stretch of flanking DNA on the pTS1 vector, this can happen on either the left or right handside of the target gene on the PA01 chromosome (depending on where the recombination event occurs). 3) The first recombination event here occurs to the left handside of the target gene. If the first recomibation event is successful and the pTS1 knockout vector is incorporated into the chromosome, colonies can be selected for on tetracyline plates. 4) Positive colonies are taken and put into liquid LB with no selection marker where a second homologous recombination crossover event is allowed to happen looping the vector out which will either occur at; 5) the left flank, resulting in the WT gene remaining, or 6) at the right flank resulting in a mutant gene deletion / truncation. This can be selected for on plates supplemented with sucrose, if the second crossover event failed to happen in liquid LB where the vector backbone housing the *sacB* gene failed to loop out, no colonies should be present. Any colonies on the plate will either remain with the WT copy of the gene or will have the gene deletion / truncation.

2.5.3 Complementation of genes in *P. aeruginosa*

2.5.3.1 Complementation using pUC18-mini-TN7T-Gm

The first approach undertaken to complement Mip genes was using the chromosomal integrative plasmid, pUC18-mini-TN7T-Gm [188]. This plasmid facilitates a single copy insertion at an *att*Tn*7* neutral site. Each construct was designed to express a single Mip gene under the control of its native promoter and were ordered from GenScript which arrived in pUC57 plasmids. These constructs were designed by taking the WT gene sequence plus 200bp upstream of the start codon, with the assumption its native promoter lies within. Where the target gene was located within an operon (but not the first gene), 200 bp upstream from the start of the operon was included as the promoter region. To prepare the expression construct the target gene was digested from pUC57 using BamHI and HindIII then ligated into the MCS of pUC18-mini-TN7T-Gm digested with the same enzymes. After incubation, an aliquot of the ligation mixtures was used to transform *E. coli* DH10B and resulting resistant colonies were screened via colony PCR and Sanger sequencing Once confirmed, the pUC18-mini-TN7T-Gm construct was co-transformed along with helper plasmid pTNS2 into PA01 using the previously described transformation protocol (**2.5.2**). To verify the gene was inserted at the correct site, colony PCR was carried out using primers up and downstream of the *att*Tn*7* neutral site.

2.5.3.2 Complementation using pJH10TS

A secondary approach taken to complement Mip genes was to use the pJH10TS plasmid. This plasmid allows complementation by ectopic expression of the target gene in a relevant mutant background under the control of a constitutive promoter. The desired Mip gene was cloned into this plasmid via Gibson Assembly (**2.4.6**). Following confirmation of construct assembly by colony PCR and sequencing, plasmids were transformed into the relevant PA01 mutant background using the transformation protocol described in section **2.5.2**. To verify that the plasmid was successfully taken up, colony PCR and Sanger sequencing was used (pJH10TS primers).

2.5.3.3 Complementation using pME6032

The final approach used to complement Mip genes utilised the pME6032 plasmid. This plasmid allows complementation by expression of the target gene using an inducible *lac* promoter. This promoter is leaky and does not require a large amount of IPTG for induction. A WT copy of the desired Mip gene was cloned into the MCS of pME6032 via restriction / ligation between KpnI and EcoRI sites. Following confirmation of correct construct assembly by colony PCR and sequencing, plasmids were transformed into the relevant PA01 mutant background using the previously described PA01 transformation protocol **2.5.2**. To verify that the plasmid was successfully taken up, colony PCR and Sanger sequencing was used (pME6032 primers). For complementation with two Mip genes, a second Mip gene was cloned into a pME6032 vector already housing a Mip gene. The second Mip gene was cloned directly next to the pre-existing Mip gene already housed in the construct between BamHI and KpnI.

2.6 Plate based virulence assays

2.6.1 Haemolysis assay

Overnight cultures were set up for each PA01 strain of interest and grown in 10 ml LB universals at 37^{o}C, 250 rpm, the following morning they were sub-cultured into fresh 10 ml LB universals and grown to or adjusted to an optical density (OD) of 0.8. To obtain a target OD for a given culture in total volume of 1000 µl the formula; 1000 µl / OD 600 nm measurement (e.g. current OD of culture at 600nm is 0.8) x wanted final OD (e.g. 0.05) was used. Growing cultures to 0.8 OD allowed the cultures to near the end of the log phase, entering the stationary phase. Two 500 ml bottles of Columbia blood agar were prepared and autoclaved. Following autoclaving 25 ml defibrillated horse blood was added to each bottle of warm liquid agar and mixed. This mixture was the haemolysis agar (**2.3.1.2**). 30 ml of haemolysis agar was poured into square petri dishes (12x12cm). From the subculture / adjusted 0.8 OD culture, 10 µl was pipetted onto

one filter disc per strain and placed on the blood plates using sterile forceps, ensuring the forceps were washed in EtOH and H_20 between each sample. Each plate setup was performed in triplicate. These plates were incubated at 37°C O/N in a static incubator. Following ~14h incubation, haemolysis zones was measured using a standard ruler in millimetres (mm). Haemolysis activity is where a clear zone of beta haemolytic activity had occurred, not just typical colony growth.

2.6.2 Protease assay

Overnight cultures were set up for each PA01 strain of interest and grown in 10 ml LB universals at 37°C, 250 rpm, the following morning they were sub-cultured into fresh 10 ml LB universals and grown to or adjusted to an optical density (OD) of 0.8. To obtain a target OD for a given culture in total volume of 1000 µl the formula; 1000 µl / OD 600 nm measurement (e.g. current OD of culture at 600nm is 0.8) x wanted final OD (e.g. 0.05) was used. Growing cultures to 0.8OD allowed the cultures to near the end of the log phase, entering the stationary phase. Two 500 ml bottles of Muller-Hinton agar were prepared and autoclaved. Lactose-free milk solution was made up adding 10 g to100 ml water and autoclaved. Following autoclaving of both the Muller-Hinton agar and Lactose-free milk solution, 50 ml milk solution was added to each bottle of warm liquid agar and mixed. This mixture was the protease agar (**2.3.1.2**). 30 ml of protease agar was poured into square petri dishes (12x12cm). From the subculture / adjusted 0.8 OD culture, 10 µl was pipetted onto one filter disc per strain and placed on the blood plates using sterile forceps, ensuring the forceps were washed in EtOH and H_20 between each sample. Each plate setup was performed in triplicate. These plates were incubated at 37°C O/N in a static incubator. Following ~14h incubation, proteolysis zones were measured using a standard ruler in millimetres (mm). Protease activity is where a clear zone of proteolytic digest of the milk activity had occurred, not just typical colony growth.

2.6.3 Biofilm formation assay

Overnight cultures were set up for each PA01 strain of interest and grown in 10 ml LB universals at 37°C, 250 rpm, the following morning they were sub-cultured into fresh 10 ml LB universals and grown to or adjusted to an optical density (OD) of 0.8. To obtain a target OD for a given culture in total volume of 1000 µl the formula; 1000 µl / OD 600 nm measurement (e.g. current OD of culture at 600 nm is 0.8) x wanted final OD (e.g. 0.05) was used. Growing cultures to 0.8OD allowed the cultures to near the end of the log phase, entering the stationary phase. In a fresh 1.5 ml Eppendorf tube, the 0.8OD culture was further adjusted down to an OD of 0.05 using the above formula. From each 1.5 ml Eppendorf containing 0.05OD culture, 5 µl of culture was taken and added to 195 µl LB in a 96 well plate giving a final OD of 0.001nm. The plate was incubated

for 24 h at 37°C in a static incubator. To stain, the planktonic cells were removed by emptying the well contents into an empty square container, and the tray rinsed with water gently, this process was repeated twice. The plate was turned over and excess water was allowed to drain. 200 µl 0.1% crystal violet was pipetted into each well and was allowed to stand for 10-15 min. Crystal violet was removed from the wells by gently submerging the plate in water to ensure cells that had adhered to the plate were not disturbed. The plate was then left to dry O/N at room temperature in the fume hood. To quantify biofilm production 200 µl 30% acetic acid solution was pipetted into each well to solubilise biofilm and allowed to sit for 10 min. The solution was pipetted up and down to ensure the biofilm was well solubilised. The resulting 200 µl sample was then transferred into a fresh 96 well plate and assayed for absorbance on the BioTek at 550 nm [190].

2.6.4 Swarming motility assay

Overnight cultures were set up for each PA01 strain of interest and grown in 10 ml LB universals at 37°C, 250 rpm. 5.5 g Formedium agar was added to 800 ml water and autoclaved. 5 x M8 solution was prepared by adding 64 g $Na_2HPO_47H_2O$ to 1 L sterile water and autoclaved. For the final swarming media several components were added to the autoclaved warm, liquid Formedium agar including; 100 ml of 5 x M8 solution, 2 ml of 50% glucose (final conc. 0.2% glucose), 12.5 ml 20% casamino acids (final conc. of 0.5%) and 0.5 ml of 1M $MgSO_4$ (**2.3.1.2**). This Agar solution was mixed and cooled slightly prior to pouring plates. 25 ml of this warming agar was poured into each circular petri dish and allowed to solidify for 1 h. The centre of the petri dish was inoculated with 2.5 µl unadjusted O/N culture. Plates were incubated upright at 37°C in a static incubator for 24-48 h and phenotype was observed [191].

2.6.5 Twitching motility assay

Overnight cultures were set up for each PA01 strain of interest and grown in 10 ml LB universals at 37°C, 250 rpm. Following O/N incubation, each strain was plated and streaked out on LB agar and incubated O/N at 37°C in a static incubator. 1% LB-Lennox agar was prepared (**2.3.1.2**) with 10 ml of the LB-Lennox liquid agar poured in petri dishes, these were the twitching agar plates. From the streaked plates, using a small loop, the outer edge of streaked culture was taken and gently mixed in a sterile area of the same plate. Using a toothpick, a match-stick sized clump of this mixed culture was taken and used to stab the centre of the LB-Lennox agar all the way down to the bottom of the petri dish. Plates were inverted and placed in a humid chamber and incubated for 24 h at 37°C. To observe the cloudy halo representing twitch movement, TM developer solution was added to the surface of the agar and left to stand for 10 min. When the halo zone became obvious, it was measured, and its surface area calculated using the formula;

½a x ½b x π (where 'a' is the longest diameter of the halo zone and 'b' is the shortest diameter of the halo zone) [192].

2.7 Biochemical analysis and assays

2.7.1 Rhamnolipid imaging

Swarming media was prepared as per the swarming protocol, with the addition of a glass slide placed at the bottom of the petri dish to be extracted later. This swarming medium was poured on top of glass slide and allowed to solidify. The bacterial culture was seeded on the agar surface above where the glass slide was situated. The glass slides were cut out from the agar and dried overnight at 45°C. Once dried samples were handed to Dr Carlo Martins in the Proteomic platform who performed the following analysis.

The agar on the slides was covered with 2,5-dihydroxybenzoic acid matrix (DHB) to a density of approx. 3 µg mm^{-2} using a SunCollect MALDI Sprayer (SunChrome, Friedrichsdorf, Germany) and a DHB solution of 10 mg/ml^{-1} in 80% methanol/0.05% TFA. MALDI imaging was performed using a Waters Synapt G2-Si mass spectrometer fitted with a MALDI source (Waters, Wilmslow, UK) and equipped with a 2.5 kHz Nd:YAG laser operated at 355 nm. The slides were fixed in the instrument metal holder and then scanned with a flat-bed scanner (Canon). The images were subsequently used to generate pattern files and acquisition methods in the HDImaging software version 1.4 (Waters) using the following parameters: area of a complete section of on average 3.5x35 mm, laser beam diameter at low setting (60 µm) with 187.5 µm step size, resulting in appr. 3000 to 5500 pixels of an area of average 120 mm^2, MALDI-MS positive sensitivity mode, *m/z* range 50-1200, scan time 1.0 s, laser repetition rate 1000 Hz, laser energy 250. Red phosphorous clusters were used for instrument calibration and lock mass correction. Total scan time for a complete section was 1-2 h and the lock mass was acquired every 450 sec for 2 sec. The MS raw files were processed in HDI1.4 with the following parameters: detection of the 2000 most abundant peaks, *m/z* window 0.05, MS resolution 10,000, lock mass 526.554 (red phosphorous cluster). The processed data were loaded into HDI1.4 and normalised by TIC. Images were generated using the HotMetal2 colour scale and exported as png image files.

2.7.2 Rhamnolipid plate extraction

Plates for rhamnolipid extraction were set up as per swarming assays (**2.6.4**). Five plugs were taken beginning immediately next to the initial central inoculum drop, regardless of how far the zone of swarming extended. The first plug was taken from the centre of the plate immediately next to the initial inoculum (labelled plug 1), with plugs 2-5 taken in a straight line, moving towards the outside of the plate, instantly after the previous plug. Plugs were taken and placed

in sterile 2 ml Eppendorf tubes, followed by addition of 2 ml of 100% MeOH which was then mixed on an Invitrogen HulaMixer™ Sample Mixer for 10 min. Once mixed, 1 ml of the MeOH solution was passed through a 0.45 µm porin filter into a fresh Eppendorf tube and labelled according to the plate, strain and plug it originated from. These samples were handed to Dr Carlo Martins for analysis.

High-resolution LCMS data were acquired on a Synapt G2-Si mass spectrometer equipped with an Acquity UPLC (Waters). Aliquots of the samples were injected onto an Acquity UPLC® BEH C18 column, 1.7 µm, 1x100 mm (Waters) and eluted with the following gradient (mobile phases (A) water/0.1% formic acid; (B) acetonitrile/0.1% formic acid) at a flow rate of 0.08 ml/min at 45 °C. The concentration of B was kept at 1% for 1 min followed by a liner increase to 40% B in 9 min, ramping to 99% B in 1 min, kept at 99% B for 2 min and re-equilibrated at 1%B for 4 min. MS1 data were collected in negative ion mode with the following parameters: resolution mode, scan time 0.5 s, mass range *m/z* 50-1200 calibrated with sodium formate; capillary voltage = 2.0 kV; cone voltage = 40 V; source temperature = 120 °C; desolvation temperature = 250 °C. Leu-enkephalin peptide was used to generate a lock-mass calibration with m/z 554.2720, measured every 30 s during the run. Quantification was performed in Skyline (v.21.2.0.536), using the precursor mass of the rhamnolipids and their isotopic distribution for delimitation of the peaks. Results of 'idotp' equal or higher than 0.9 were assigned for identification of peaks.

2.7.3 Siderophore assay culture set up

Iron limiting succinate medium was prepared (**2.3.1.1**) and adjusted to a pH of 7.0. Inoculated O/N cultures in 10 ml LB at 37^{o}C agitation 220 rpm. Poured 100 ml succinate media into sterile conical flask and added the 10ml O/N culture, incubated at 30^{o}C under agitation 220 rpm for 24 h. 10 ml of the 24 h culture was pipetted into 100 ml fresh succinate medium and incubated in a shaking incubator at 30^{o}C, 220 rpm for 24 h.

2.7.4 Pyoverdine (PVD) UV-visible siderophore assay analysis

To analyse siderophores, Lambda 35 UV spectrophotometer was set to wavelength range 300 to 600 nm and a base level established. Into a quartz cuvette 1 ml of pyridine (Pyr) / acetic acid (AcOH) 50 mM buffer pH 5.0 was pipetted and placed in the control position of spectrophotometer. For samples, added 100 µl culture to quartz cuvette, making it up to 1 ml with 900 µl Pyr/AcOH buffer. Spectrophotometer was ran, if pyoverdine was present a specific peak was visible at around absorption 380 nm preceded by a substantial shoulder at 360 nm. If PVD was in the presence of iron, a complex formed and causing a modification of absorbance, mainly disappearing of 360 nm shoulder and a shift of the 380 nm peak to 400 nm [193, 194].

2.7.5 Pyochelin (PCH) extraction

Cultures were inoculated with the final culture centrifuged at 4000 rpm for 8 min collecting the supernatant. Supernatant was filtered using 0.45 µm porin filters into fresh falcon tubes using 15 ml syringe. Filtrate pH was altered to 3.0 by adding citric acid powder and checked on litmus paper. Filtrate was poured into separating funnel with 20 ml dichloromethane added and gently mixed allowing gas out each time. Once it had separated into three phases the bottom phase was collected in a beaker. $MgSO_4$ was added to the collected phase and allowed to stand for 15 min. This was then filtered using filter paper and collected in a falcon tube. The sample was concentrated under reduced pressure on the *Vacuo*. Sample was then solved in 300 µl ethanol and centrifuged with the supernatant poured into a high-performance liquid chromatography (HPLC) vial [193].

2.8 *In vivo* infection assays

2.8.1 *Galleria mellonella* infection assay

Overnight cultures were set up for each PA01 strain of interest and grown in 10 ml LB universals at 37°C, 250 rpm, the following morning they were sub-cultured into fresh 10 ml LB universals and grown to or adjusted to an optical density (OD) of 0.8. To obtain a target OD for a given culture in total volume of 1000 µl the formula; 1000 µl / OD 600nm measurement (e.g. current OD of culture at 600nm is 0.8) x wanted final OD (e.g. 0.05) was used. Growing cultures to 0.8OD allowed the cultures to near the end of the log phase, entering the stationary phase. In a fresh 1.5 ml Eppendorf tube, the 0.8OD culture was further adjusted down to an OD of 0.05 using the above formula. From each 1.5 ml Eppendorf tube containing 0.05 OD culture, 100 µl of culture was taken and added to 900 µl sterile PBS, making a serial dilution of 10^{-1}. This was repeated five more times giving an overall serial dilution of 10^{-6}. This was done in an attempt to obtain a culture containing between 0-4 bacterial cells, the optimal number of cells to initiate an infection to monitor disease progression. Ten *G. mellonella* larvae were used for each PA01 strain analysed, with each injected in the last left pro-leg using 10 µl of the lowest serial dilution culture (10^{-6}) using a Microliter syringe (Figure 2.). Infected larvae were incubated at 28°C and checked every 6 h over a 48 h period with the status of each larvae at each 6 h interval recorded as dead (if melanised) or alive (if non-melanised) [195].

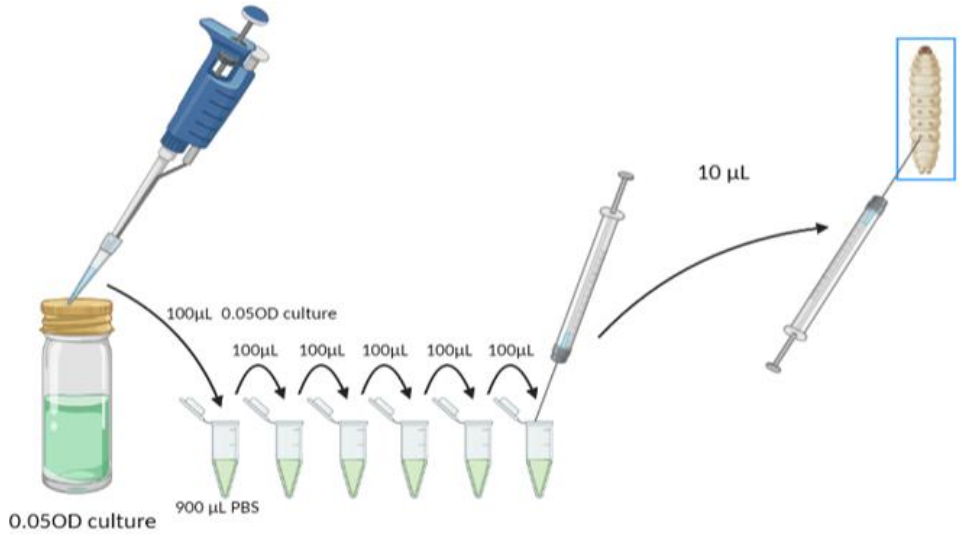


Figure 2.2. Image depicting how the *G. mellonella* infection assay was performed. The image shows how the serial dilution to acquire the initial inoculum of between 0-4 cells was performed. Once acquired, this image displays the precise location the *G. mellonella* larvae were injected with the 0-4 cell inoculum.

2.8.2 Macrophage infection assay

RAW 264.7 is a macrophage cell line that was established from a tumour in a male mouse induced with the Abelson murine leukaemia virus (termed RAW 264.7 cells). These RAW 264.7 macrophages were incubated in Dulbecco's Modified Eagle Media (DMEM) from Lonza containing 0.6 g/L^{-1} L-Glutamine, 10% of foetal bovine serum (FBS) and supplemented with 1% penicillin/streptomycin solution (20 units ml^{-1} penicillin and 50 $\mu g/ml^{-1}$ streptomycin). Cells were grown in an incubator with an atmosphere of 5% CO_2 at 37˚C. The health of cells and their viability was checked using a Calcium mobilisation assay where cells were checked for their responsiveness to 10 µM ATP after 24 h in 96 well plates. When provided with the 10 µM ATP, RAW 264.7 cells to respond by increasing their uptake of $Ca2^+$ ions which checks the cells are able to respond normally to agonists. The shape and basal state of calcium in the cells showed us how healthy cells were at the beginning of experiments.

Cells were plated in 96-well clear-bottomed plates at a concentration of 20,000 cells per 200 µl DMEM in each independent well 24 before experimentation. Bacterial strains to be studied were grown in LB O/N at 37°C at 220 rpm. The following morning bacterial cultures were serial diluted in sterile PBS to the multiplicity of infection (MOI) required (ratio of infectious agent to host cell; MOIs used in this study were 10 and 50). Prior to inoculation of RAW 264.7 cells the DMEM culture medium was removed and replaced with 200 µL fresh DMEM without FBS or pen/strep. 50 µl of bacteria in PBS at the desired MOI was added to each well containing 20,000 RAW 264.7 cells (Figure 2.1). Plates were incubated for 3 h at 37°C in an upright incubator, after which the medium was replaced with 200 µL of fresh DMEM containing 100 µg/ml gentamicin but without FBS or pen/strep to kill off any bacterial cells that had not invaded the RAW 264.7 cells. The plate

was then further incubated for 1 h at 37°C in an upright incubator after which the incubation medium was removed and the RAW 264.7 cells washed once with sterile PBS. After all traces of PBS was removed, and 100 µl of filter sterilised distilled water was added to each well to lyse macrophages. This lysate was added to sterile Eppendorf tubes containing 100 µl PBS creating a 1in2 dilution. 50 µl of the 1in2 diluted lysate was then plated on LB agar and incubated O/N at 37 °C. This was used to determine colony forming units (CFU) on agar the following morning.

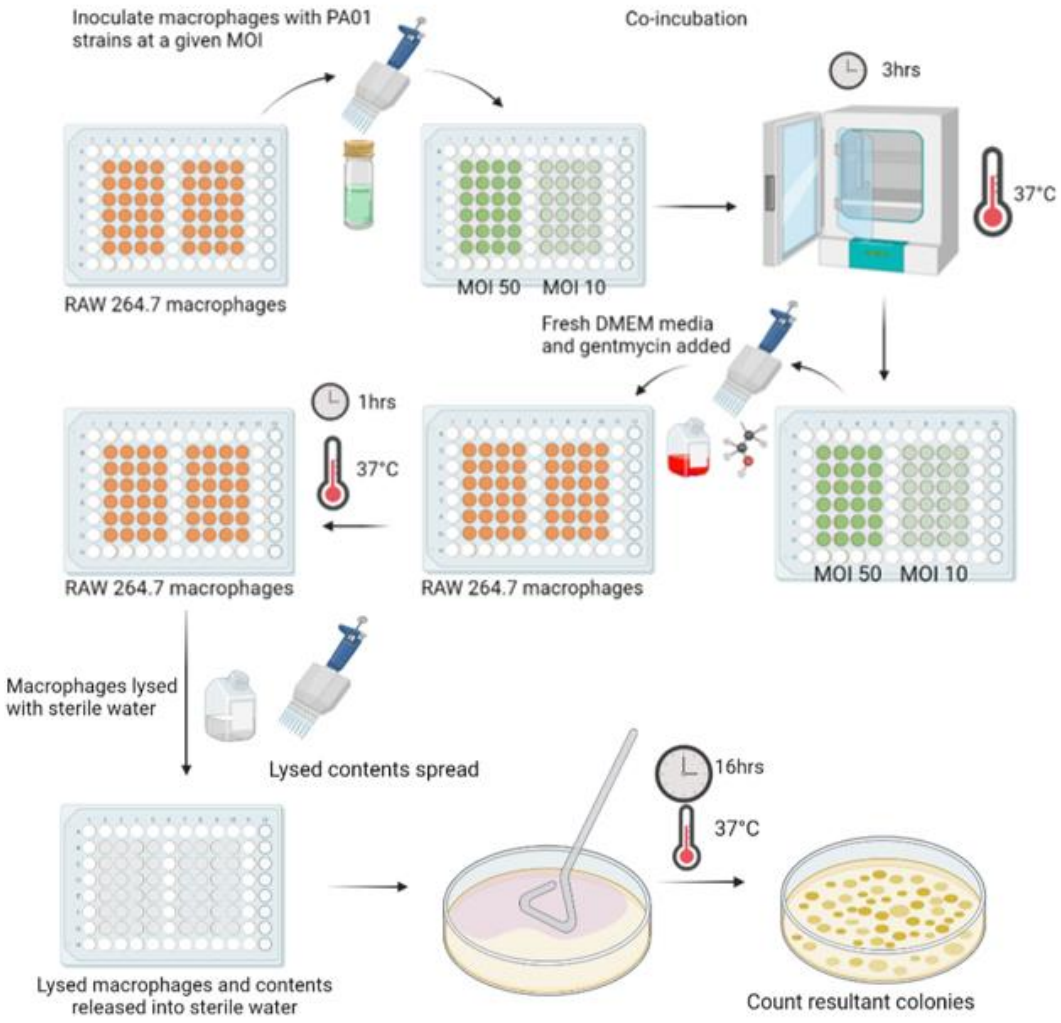


Figure 2.3 Flow diagram of how the macrophage infectivity assay was performed. The flow diagram shows a diagrammatic representation of how the macrophage infectivity assay was performed and what is described above (**2.8.2**).

2.9 RNA sequencing and Whole-Genome Sequencing

2.9.1 RNA isolation for RNA sequencing

Three biological replicates were per strain subjected to RNA sequencing analysis (sent to Novogene UK Company Limited). Prior to starting experiments, all areas were deep cleaned using RNase cleaning spray, and this was frequently repeated between each step throughout the process to ensure of no contamination or breakdown of samples. An O/N culture of the

strain of interest was incubated O/N at 37°C in a shaking incubator at 250 rpm. The following morning this culture was adjusted to an OD of 0.4 with RNase free water and 16x50 µl aliquots were dropped on an LB agar square plate (12 cmx12 cm) which was then incubated at 30°C for 20 h. Following the 20 h incubation, all cell material (~100-200 mg / plate) was scraped off into a sterile 5 ml Eppendorf tube containing 3 ml RNAprotect Tissue Reagent, creating a cell suspension. From this point RNA isolation was performed using the Qiagen RNeasy Minikit. From the 3 ml cell suspension, a 300 µl aliquot was added to a 2 ml Lysing Matrix B glass bead tube from MP Biomedicals. 700 µl buffer RLT from the Qiagen RNeasy Minikit was added to the tube and placed in a cell disruptor for 3x45 sec on cycle 6. The resulting lysate was centrifuged for 10 sec at maximum speed and the supernatant transferred into a fresh sterile 2 ml Eppendorf tube. The volume of the supernatants was determined, and an equal volume of 70% ethanol was added and mixed well by pipetting. 700 µl lysate was added to a RNeasy spin column and placed in a 2 ml collection tube, closing the lid gently before centrifuging for 15 sec at ≥8000 *x g* (≥10,000 rpm). The flow-through was discarded and 700 µl Buffer RW1 was added to the RNeasy spin column and centrifuged for 15 sec at ≥8000 *x g* (≥10,000 rpm). The flow-through was again discarded and 500 µl Buffer RPE was added to the RNeasy column and centrifuged for 15 sec at ≥8000 *x g* (≥10,000 rpm). The flow through was again discarded and 500 µl Buffer RPE was added to the RNeasy column and centrifuged for 2 min at ≥8000 *x g* (≥10,000rpm) to wash the spin column membrane. The RNeasy spin column was then placed into a fresh collection tube and centrifuged at full speed for 1 min to ensure the full drying of the column. The RNeasy spin column was then placed in a 1.5 ml collection tube with 50 µl RNase-free water added directly to the centre of the spin column which was then centrifuged for 1 min at ≥8000 *x g* (≥10,000 rpm). An Invitrogen DNase treatment was then performed by adding 0.1 volume of 10X TURBO DNase Buffer and 1 µl TURBO DNase to the RNA followed by gentle mixing. This was incubated at 37°C for 30 min before 0.1 volume (15 µl) of resuspended DNase Inactivation Reagent was added and the solution mixed well. This was incubated for 5 min at room temperature with occasional mixing. This was pipetted into a Qiagen RNeasy spin column and centrifuged at 10,000 *x g* for 90 sec with the flowthrough discarded. Following DNase treatment the sample in the Qiagen RNasey spin column was further purified by adding 500µl Buffer RPE and centrifuged for 15 sec at ≥8000 *x g* to wash the membrane, discarding the flow-through. 500µl Buffer RPE was added to the spin column and centrifuged for 2 min at ≥8000 *x g* to wash the membrane. The RNeasy spin column was added to a new sterile collection tube (supplied in the Qiagen kit) and centrifuged at full speed for 1 min. The RNeasy spin column was then added to a new 1.5 ml sterile collection tube and 40 µl of RNase-free water was directly pipetted on to the spin column membrane. This was centrifuged at ≥8000 *x g* for 1 min to elute the RNA. RNA quality and quantity was then checked on a Nanodrop looking for an absorbance ratio of 260/280

between 1.9-2.0. Samples had to be within the 1.9-2.0 parameters to pass quality control. Once sent to Novogene, they performed their own QC analysis and RNA library preparation including, rRNA removal, RNA fragmentation, cDNA reverse transcription, stranded RNA library preparation. They then performed Illumina sequencing strategy PE150.

2.9.2 RNA Sequencing analysis

The *subread-align* program of the Subread package in Linux was used to align the sequencing reads to the *Pseudomonas aeruginosa* PAO1 reference genome (NCBI accession NC_002516.2). This resulted in .bam files for each replicate. The *featureCounts* function of the Subread package was then used to count the number of reads mapping to each gene in the reference genome. This resulted in a table of counts for each gene for each replicate which was then read into a data frame in R. This data frame was made into a *DGEList* (edgeR package) object with replicates assigned to their respective groups (all function names mentioned from now on are provided in the edgeR package). Genes with low read counts across all samples were filtered out using the *filterByExpr* function. tRNA and rRNA genes were also filtered out. Normalisation factors for library sizes were calculated using the *calcNormFactors* function. After making a suitable design matrix, common, trended and gene-wise dispersions were estimated using the *estimateDisp* function. The *glmQLFit* and *glmQLFTest* functions were used for fitting genewise negative binomial generalised linear models and statistical testing (empirical Bayes quasi-likelihood F-tests) between the wild type and the other strains (defined in a contrasts object). Finally, the function *topTags* was used to output tables for all comparisons containing (for each gene) logFC, logCPM, F, PValue, FDR and, gene product). R version used was 4.1.2. Performed by Dr Govind Chandra.

2.9.3 Preparing stocks for full genome sequencing

Strains to be sent for full-genome sequencing were inoculated in liquid LB O/N in a shaking incubator at 37^{o}C, 220 rpm. The following day this liquid culture was streaked out onto LB agar and incubated statically O/N at 37^{o}C to obtain single colonies. A single colony was picked and mixed in 100 µl of sterile x1 PBS making a cell suspension. Taking a sterile loop, this cell suspension was streaked out across an LB agar plate ensuring around 1/3 of the plate was spread out to be a thick lawn of bacteria with the remainder of the plate streaked out as a normal spread plate resulting in single colonies to ensure the entire plate was pure and no contaminants present. Plates were then incubated O/N at 37^{o}C in a static incubator. Using a large sterile loop all of the bacterial culture on the plate, including the 1/3 lawn coverage and all single colonies, were scraped from the plate and mixed into barcoded tube supplied by MicrobesNG containing beads and MicrobesNG own liquid media. Each barcoded tube was designated to a specific

sample (e.g. WT1, WT2, WT3, I1, I2, I3 etc.) This tube was then mixed by inversion 10 times. The tubes were sealed and sent back to MicrobesNG at room temperature (didn't require dry ice). MicrobesNG returned the sequencing which provided 30X sequence coverage and was analysed by Dr Govind Chandra.

2.10 Phylogenetic analysis

Sequences of *P. aeruginosa* potential Mip genes were identified using known Mip genes from the literature to perform various searches against the genome sequence of *P. aeruginosa*. Genes that were flagged in these searches were then further analysed using sequence alignment tools in Geneious Prime including MUSCLE and Clustal Omega. Sequences were trimmed at the N-termus to remove the highly unrelated sequences. From these sequence alignments phylogenetic trees were created to assess the relationships using an UPGMA consensus tree on Geneious Prime. and select genes for further investigation as to whether their gene products were likely to be housekeeping or FKBP virulence related factors.

2.11 Software

Software used for *in silico* analysis and generation of figures.

Software	Use
BLAST	Homology searches of DNA and protein sequences
BioRender	The creation of diagrams and figures
ChemDraw (v18.0)	The creation of chemical structures for illustration purposes
Clustal Omega	Amino acid sequence alignment and trimming
Excel	Used for statistical analysis and generation of graphs and charts
Geneious Prime® 2020.0.3	Used to design primers, constructs, assess sequence data, creation of nucleotide alignments and phylogentic trees
GraphPad	Used for statistical analysis and generation of graphs and charts
MUSCLE	Amino acid sequence alignment and trimming
PyMOL 2.5.2	Used to look at crystal structures and create crystal structure overlays to assess FKBP binding C-domains
R Project for Statistical Computing	Used for statistical analysis and generation of graphs and charts

Table 2.13. Software used in this study.

2.12 Statistical analysis

To evaluate the significance of results obtained, T tests were performed to compare the means of two groups to a probability value of ($P<0.05$). All experiments were performed in triplicate (n – number of replicates e.g. n=3), unless stated otherwise. All graphs presented possess standard error (SE) bars. Standard error was calculated by first obtaining a standard deviation (SD) value for each data set (STDEV.S function on Microsoft Excel) then using the formula; SD / (√(n)) =SE. Software used to perform statistical analysis can be found in Table 2.10.

Chapter 3:

Identification of Mip-like genes in *Pa01* and assessing their role in virulence using *in vivo* and *in vitro* assays

Chapter 3. Identification of Mip-like genes in *Pa01* and assessing their role in virulence using *in vivo* and *in vitro* assays

3.1 Introduction to the work and project outlook

As described above, discovering novel drug targets is becoming increasingly important and Mips stand out as attractive candidate targets for the development of new anti-virulence drugs for the use against gram-negative pathogens. Given that multiple studies have demonstrated the importance of Mip genes in the virulence processes of several gram-neg pathogens, work in this thesis was designed to further develop our understanding of the role of Mips in additional gram-negative strains. Specifically, we investigated the clinically relevant strain *Pseudomonas aeruginosa*, PA01, to assess whether any genes potentially encode Mip-like protein candidates. Once Mip-like candidate genes had been established through a sequence homology and phylogenetic approach, a series of gene deletion strains were created and subjected to a variety of *in vivo* and *in vitro* virulence assays to determine whether their gene products represent true Mip proteins required for full virulence. Following assays, to begin to elucidate how these Mip proteins in PA01 may exert their virulence effect, RNA sequence analysis was performed to asses the transcription of virulence related genes in Mip mutant strains that presented with attenuated virulence phenotypes in the *in vivo* and *in vitro* assays.

3.2 Identification and annotation of Mip-like genes in PA01

Initial searches and phylogenetic analysis were performed using homologues of characterised Mip genes from *Burkholderia pseudomallei* and *Legionella pneumophila* as comparator sequences to determine potential Mip-like genes within the PA01 genome. These initial works were performed by previous lab members; Dr Lorena Fernandez-Martinez, Dr Eleftheria Trampari and Dr Silke Alt. Initial searches flagged several potential FKBPs in the PA01 genome. To assess which of these were more likely a Mip-gene candidate, phylogenetic analysis of their sequences was performed. This phylogenetic analysis aided in the elucidation of which were likely to be housekeeping genes and which were likely to encode Mip-like gene candidate(s). Initial phylogenetic analysis by previous group members used only a small number of characterised FKBPs and outliers in their analysis which worked adequately to identify five Mip-like gene candidates (Figure 3.1). Here I have expanded on their initial phylogenetic analysis to include all families within the PPIase superfamily, including cyclophilins and parvulins. By doing this, my analysis shows a clear differentiation between each family of PPIase, as well as the separation between housekeeping FKBPs and virulence associated Mips (Figure 3.1).

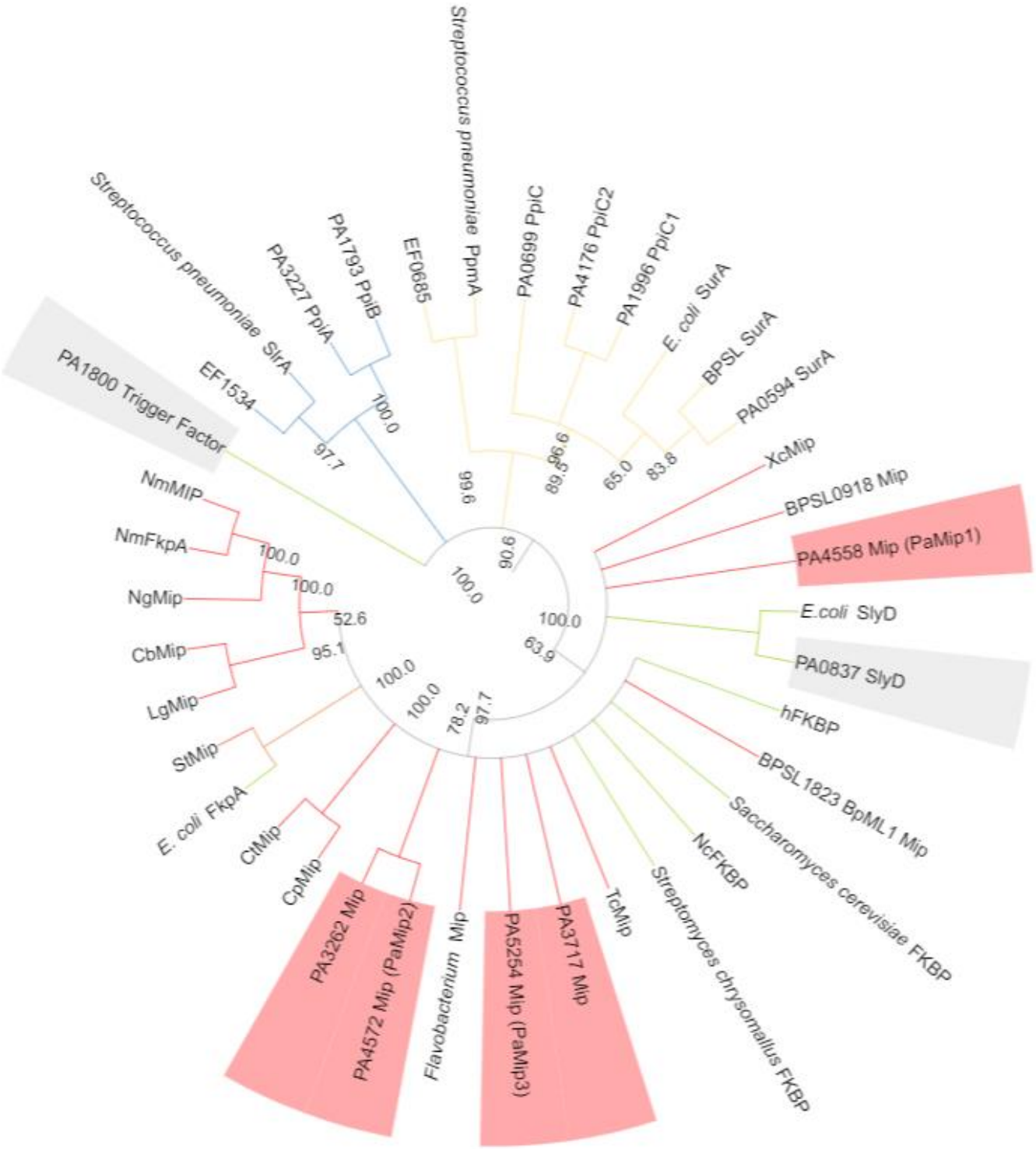


Figure 3.1. Phylogenetic UPGMA protein alignment consensus tree generated using Geneious Prime using protein alignments of a variety of PPIases from different organisms. Sequences were trimmed to remove large gaps, predominantly at the N-terminal domain where conservation is lost given PPIase activity occurs at the C-terminal domain. Each different coloured branch represents a different PPIase group. The blue highlights cyclophilins, yellow highlights parvulins, green highlights known housekeeping FKBPs and red highlights known Mips. Any shaded sections represent PA01 genes predicted to encode an FKBP, red sections highlight which are the more likely to present as virulence associated FKBPs (Mips), the grey sections highlight FKBPs with a function more likely aligned with housekeeping roles.

Based on their grouping with other known Mip genes and distance away from known housekeeping PA01 FKBPs such as SlyD and Trigger Factor (TF) (Figure 3.1). It was clear which PPIase encoding genes in the PA01 genome presented as hypothetical FKBPs, and from these genes, which were likely to encode potential Mip gene candidates. The only exception to this was PA4558 which although grouped closer to known housekeeping FKBPs, it grouped with the well annotated virulence Mip BPSL0918 from *B. pseudomallei* sharing 38% sequence similarity. Interestingly BPSL0918 is the only Mip to have been shown to lack PPIase activity, and like PA4558, is located in the cytoplasm unlike other Mips that are described as surface proteins, this will be discussed in more detail later.

Previous lab members created single gene knockout strains for the genes; *pa3262*, *pa3717*, *pa4558*, *pa4572* and *pa5254* using homologous recombination. From these single knockout strains, double and triple knockout strains were subsequently generated (Table 3.1.). Each knockout strain was annotated with a letter (B-Q) to simplify analysis and annotation of results, with the WT PA01 strain often referred to as either the letter 'A' or WT. Due to complications that arose after whole genome analysis, which will be discussed in detail in Chapter 4, I remade strains D, E and I independently and all assays presented in this chapter were performed with my version of these three knockout strains. My new versions were made in an manner to those of previous lab members, with my strains D, E and I behaving in exactly the manner, with identical phenotypes and results in all assays to those of previous lab members strains. It is worth noting that some Mip-like gene deletion combinations could not be created despite numerous attempts and starting from each of the parent single mutant variants suggesting synthetic lethality. This suggested that some combinations of these genes are essential for the survival of the organism, possible due to one compensating for the activity of the other. Table 3.1 lays out the Mip-like gene knockout combinations that were created and their denoted letter as explained.

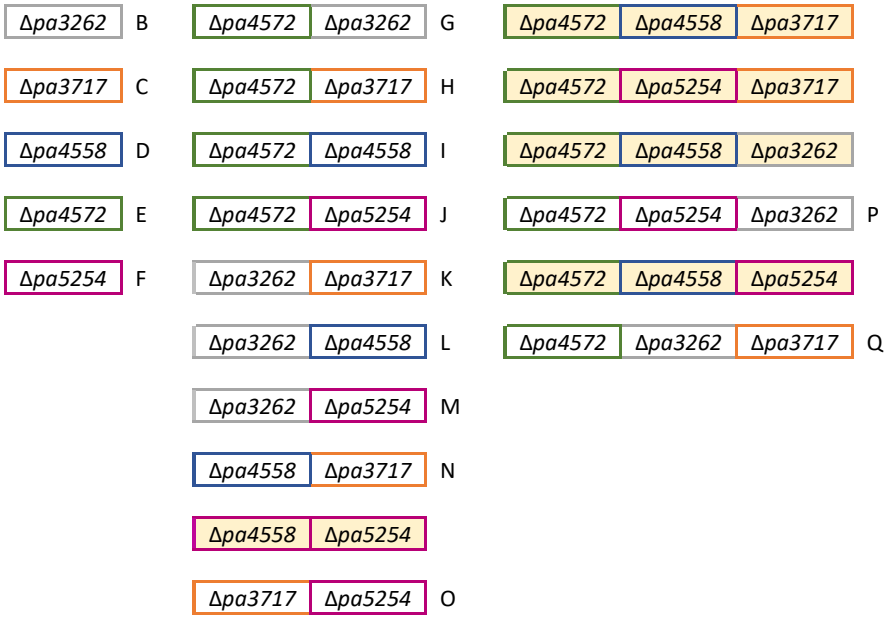


Table 3.1. A table laying out PA01 Mip gene knockout strains. A list of PA01 Mip gene knockout combinations with single mutant strains on the left, double mutant knockout strain in the middle and triple mutant strains on the right. All have been assigned a letter for reference. Some of the possible combinations were not possible to make and are thought to be deleterious to the organism, and hence these combinations are not annotated with a letter and are highlighted in yellow. The wild type is denoted as 'A' or WT not listed here.

With the creation of these Mip-like gene knockout strains phenotyping was undertaken using a variety of *in vitro* and *in vivo* assays. This chapter describes the format and results for each assay, and why it was used to assess virulence of the strains. Furthermore, this chapter will discuss which PA01 Mip-like genes are required for full virulence of PA01, and therefore which were carried forward for further study. All raw data for the assays described below can be found in the relevant Appendices.

3.3 *In vitro* plate based virulence assays

In vitro assays are experimental biology research tools that allow the assessment of cell health after experimental treatment or manipulation. Using *in vitro* assays to assess virulence is useful to help understand what infectious processes may have been affected by gene knockouts and the potential pleotropic affects they may have. Furthermore, they can elucidate what virulence mechanisms may be impaired when in a host, resulting in loss of infectivity, and whether moving to an *in vivo* model is worthwhile, what *in vivo* model is most appropriate and how to design experiments. Prior to running *in vivo* assays, having an idea of what virulence processes may be attenuated *in vitro*, will help us decide whether to undertake *in vivo* studies and if so, how to understand and unpick the *in vivo* data.

3.3.1 Haemolysis assay

Haemolysis assays are used to assess whether a pathogen can produce haemolysins that damage the host cytoplasmic membrane, causing cell lysis and death. Pathogenic bacteria possess three types of haemolysins; α-haemolytic strains oxidize the iron in the haemoglobin (turning agar supplemented with horse blood a dark green), β-haemolytic strains completely rupture the red blood cells (visible as a halo on agar) and γ-haemolytic, or non-haemolytic, strains that do not cause haemolysis and rarely cause illness. PA01 is β -haemolytic, meaning it can completely lyse and rupture red blood cells as part of their infectious cycle [196]. This β-haemolytic haemolysis activity is a result of the *plcH* gene being expressed [197].

For the haemolysis assay each strain was pipetted on to filter discs then placed onto a blood agar plate. If β-haemolytic activity was functional there should be visible halo rings as observed for the WT strain (Figure 3.2). Here, it was found that strains I and J were significantly attenuated in their ability to haemolyse red blood cells compared to the WT (Figure 3.2). Each halo ring was measured across its entire diameter and the mean average zones recorded.

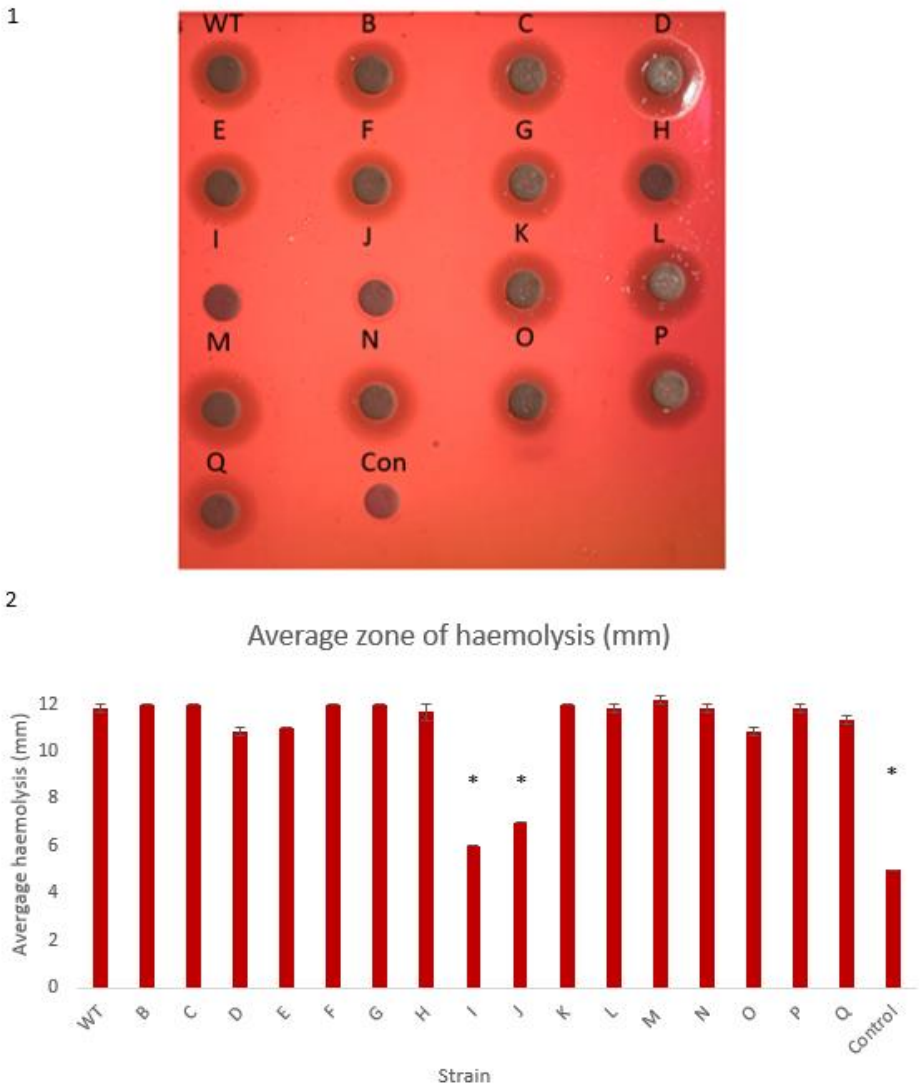


Figure 3.2. Haemolytic activity of all Mip knockout strains B-Q; Panel 1 shows a blood agar plate with strains showing clear zones of haemolytic activity compared to the control of sterile water. The WT has a large activity zone compared to those of strains I and J, both of which have a clear, reduced halo. The bar chart in panel 2 shows the mean average zones of haemolytic activity measured from replica plates (n=3). Along the X axis are the Mip knockout strains as described in Table 3.1. Standard error variance was calculated and plotted as seen by the error bars comparing the difference between the mean with the amount of scatter within the groups, error bars that overlap are not statistically significant. The asterisks above columns indicate ($P<0.05$, n=3) (see methods).

From Figure 3.2, it is evident that strains I and J (comprised of deletions in Mip genes D (*pa4558*), E (*pa4572*) and F (*pa5254*)) are significantly attenuated in haemolysing red blood cells when compared to the WT ($P<0.05$). It is interesting to note that the single knockouts from which I and J are derived (D, E, F) show no difference in activity compared to the WT. In terms of infection, this is highly relevant, as lysing red blood cells is, in part, a key mechanism by which PA01 obtains iron, an essential metal for bacterial growth [198]. Furthermore, PA01 collates iron using siderophores; pyochelin and pyoverdine in iron limiting environments, the importance of which will be discussed later.

3.3.2 Protease assay

Protease activity is the result of excreted protease enzymes, which are enzymes that hydrolyse peptide bonds in protein molecules [199, 200]. Pathogenic protease activity has been proposed as an important virulence factor in disease progression through a variety of mechanisms including, being responsible for the damage of host tissues, involved in evading the host immune system, increase a pathogens ability to endure phagocytosis, and finally a pathogen's ability to escape entanglement in host fibrin meshes [201].

As with the haemolysis assay, each strain was pipetted on to filter discs then placed onto the lacto-free whey milk protein agar plates. Similarly, to the haemolysis assay, proteolytic activity was visualised by visible halo rings where the hydrolysis of casein had occurred. As with the haemolysis assay, we found that strains I and J were significantly attenuated in their ability to break down the whey protein suggesting these strains are unable to excrete optimal protease activity when compared to the WT (Figure 3.3). Each halo ring was measured across its entire diameter and recorded for the mean average zone of activity.

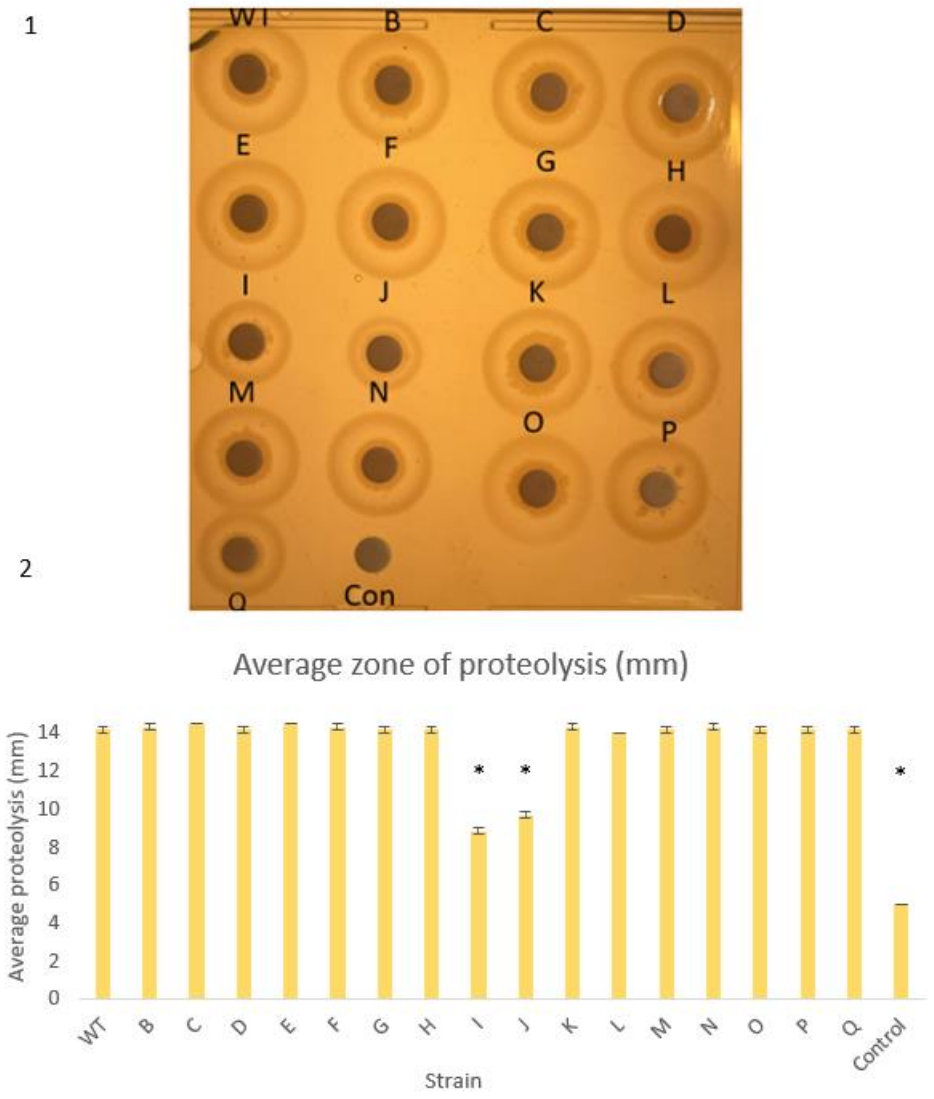


Figure 3.3. Proteolytic activity of assay using all Mip knockout strains B-Q; Panel 1 shows a lactose-free whey protein agar plates with strains showing clear zones of proteolytic activity compared to the control of sterile water. The WT has a large clearance zone compared to those of strains I and J, both of which have a clear, reduced halo. The bar chart in panel 2 shows the average zones of proteolytic activity measured from plates. Along the X axis are the Mip strains as described in Table 3.1. Standard error variance was calculated and plotted as seen by the error bars comparing the difference between the mean with the amount of scatter within the groups, error bars that overlap are not statistically significant. The asterisks above columns indicate ($P<0.05$, $n=3$) (see methods).

As in the haemolytic assay, strains I and J display significantly less protease activity than the WT ($P<0.05$), and once again the single knockouts from which I and J are derived (D, E, F) show no difference in activity compared to the WT. Here, it is observed that strains I and J have a significant reduction in protease activity, implicating PA01 Mip-like genes are involved in processes contributing towards proteolytic activity. It is important to note this experiment was carried out twice, both in triplicate on different days and freshly prepared media. Firstly, using normal powdered milk, however it was noted that the breakdown of the lactose may have resulted in pH changes causing the agar to become more acidic, thus affecting the growth of the

strains and skewering results. We repeated the experiment using lacto-free milk powder ruling out any possible pH changing affects. Both types of powder used presented with the same results and made no difference.

3.3.3 Biofilm formation assay

As described in the introduction chapter, biofilm formation is a key virulence determinant of PA01 and is responsible for its characterisation as a chronic infection following its transition from the planktonic lifestyle. Chronic infection is a severe infection persisting for weeks, months or even years. This chronic phase of infection is predominantly seen in cystic fibrosis (CF) patients where the bacteria thrive in the thick mucus produced in their lungs. The transition from acute to chronic phase is often put down to the formation of a biofilm; a consortium of bacteria adhering to a surface that are embedded within an extracellular matrix. The extracellular matrix that constitutes a biofilm is viscous, predominantly made up of various polysaccharides such as alginate. This vicious matrix often renders antibiotics useless as they cannot penetrate through the matrix to reach the target cells. As biofilms are key to PA01 virulence, it was prudent for us to assess whether PA01 Mip-like genes are involved in their formation by using a biofilm assay. This involved growing PA01 strains in the wells of 96-well plates, removing the media and washing, staining with crystal violent, resuspend the resulting biofilm in acetic acid and measuring UV of the resulting solution; this assay utilises the fact the amount of strain retained is proportional to the number of cells in the biofilm. The more biofilm produced, the more purple the solution, and these correlate with a greater UV absorbance reading (Figure 3.4). If a biofilm is poorly formed, the washing steps (**2.6.3**) will wash away cells not encased within the biofilm matrix, resulting in fewer cells to uptake the crystal violet dye and therefore returning a lower UV measurement.

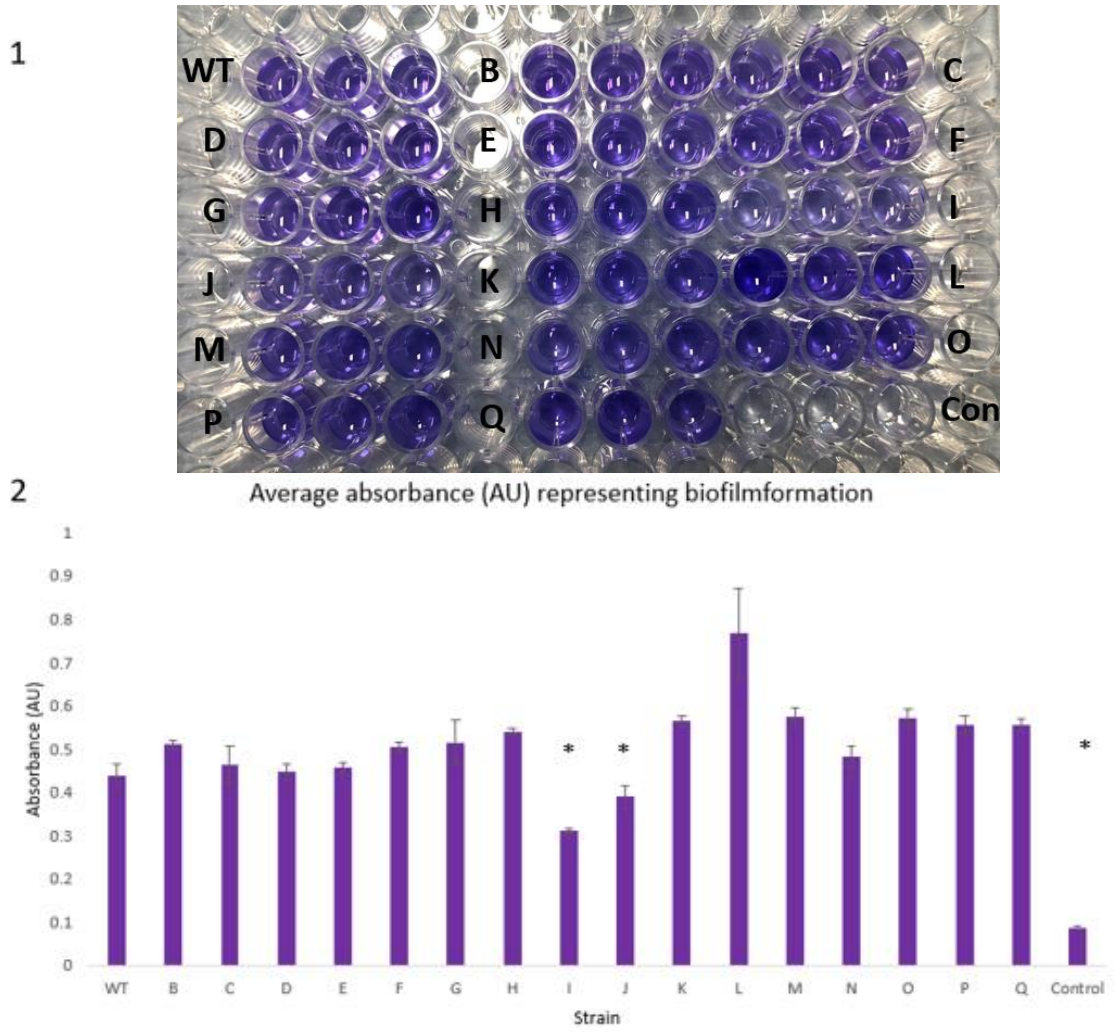


Figure 3.4. Biofilm formation assay using all Mip knockout strains B-Q: Crystal violet strain derived from absorbance by biofilm dissolved in acetic acid. The experiment displays different purple intensities for the different strains (Panel 1). Example plate where strains have been grown in a 96-well plate then stained with crystal violet. Wells housing strains I and J have a much lighter colour when compared to the WT and other strains meaning they have produced less biofilm. Strains are labelled from the WT to strain Q with each strain grown in triplicate. The UV absorbance was measured on a plate reader and plotted in the bar chart in panel 2. Standard error variance was calculated and plotted as seen by the error bars comparing the difference between the mean with the amount of scatter within the groups, error bars that overlap are not statistically significant. The asterisks above columns indicate ($P<0.05$, $n=3$) (see methods).

By comparing the colour intensities in Figure 3.4(1). there is a clear difference between strains I and J, and the WT, suggesting these mutants have an impaired ability to produce a biofilm. This was confirmed by the absorbance readings obtained (Figure 3.4(2)) for I and J which are both significantly lower ($P<0.05$) than that for the WT. Again, the single knockouts from which strains I and J are derived show no significant difference in biofilm formation compared to the WT. These results suggest that certain combinations of PA01 Mip-like genes are required for the formation of biofilms by PA01 (as I and J strains are impaired).

Given that only strains I and J showed any phenotypic deviations from the WT, in the first three assays run, for the majority of subsequent assays I primarily focussed on using these two strains, as well as the three single mutant strains D, E and F, that contain deletions in the three Mip genes that comprise the deletions in strains I and J.

3.3.4 Twitching motility assay

Previously it has been determined that biofilm formation is linked with hair like structures on the surface of PA01 cells termed type IV pili. Initial attachment to the surface prior to biofilm formation is mediated by the type IV pili and a reduction or loss of type IV pili has been shown to affect biofilm formation [64, 202]. These type IV pili are also responsible for a form of movement performed by PA01 known as twitching which is flagellum independent. This movement begins following the initial type IV pili led attachment to a solid surface, whereby they then act as grappling hooks that extend and retract pulling the bacterium along the surface. This form of movement can be visualised by a twitching assay and led to our hypothesis that given these type IV pili are involved in the formation of a biofilm, and strains I and J are impaired at forming biofilms, then these two strains may not twitch or at least should display somewhat impaired twitching when compared to the WT. Furthermore, twitching motility is linked to the virulence of PA01 as an essential movement, without which dissemination in a host is defective [203].

Following the hypothesis that strains I and J would lack, or express a limited number of type IV pili, a twitching movement assay was carried out. Twitching is responsible for chemotaxis toward or away from specific signals when other movement types, such as swarming or swimming which require semi-solid surfaces, are not possible [204]. The strains investigated were inoculated onto twitching plates by stabbing with an inoculated toothpick ensuring it reached the bottom of the plate, making contact with the solid surface where twitching motility occurs then incubated in a humid chamber. In this experiment we only examined the strains WT, D, E, F, I and J. As expected, WT, D, E and F all displayed similar levels of twitching, whereas strains I and J did not (Figure 3.5 and 3.6).

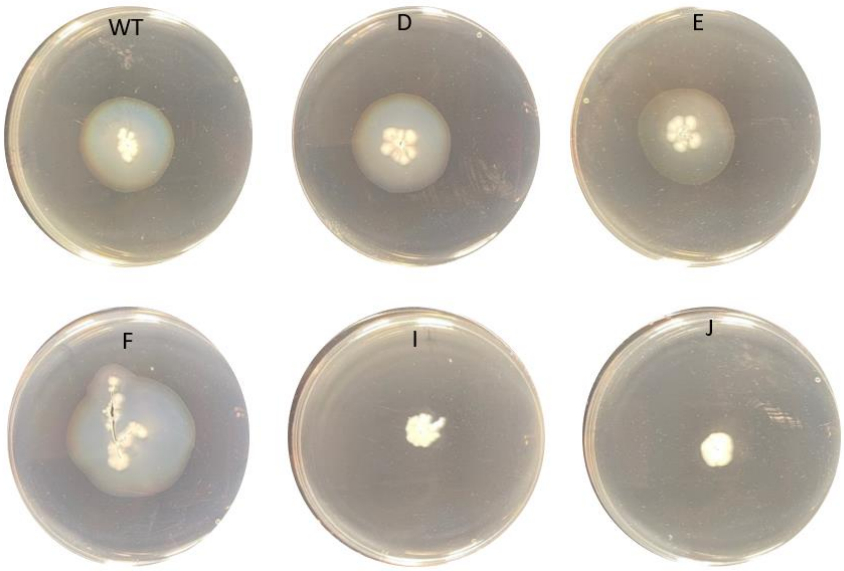


Figure 3.5. Twitching motility assay plates for the WT, D, E, F, I and J strains. The WT has a faint white cloud around the initial inoculum, this is the growth outwards along the solid plastic petri dish as a result of twitching. The growth indicated by the smaller dense white patches is on the surface of the agar due to the initial inoculum. Strains I and J display no visible twitching halo.

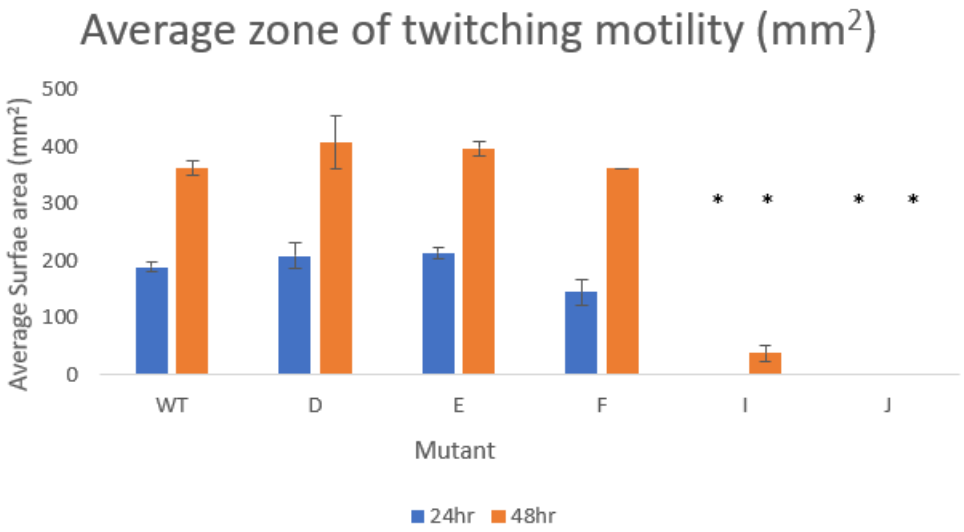


Figure 3.6. Bar chart plotting the average zone of twitching measured after 24 h and 48 h for Mip mutant strains. Twitching motility zones were calculated using the equation; ½a x ½b x π (where 'a' is the longest diameter and 'b' is the shortest diameter). The experiment was performed twice on separate days using freshly prepared media. Standard error variance was calculated and plotted as seen by the error bars comparing the difference between the mean with the amount of scatter within the groups, error bars that overlap are not statistically significant. The asterisks above columns indicate ($P<0.05$, $n=2$) (see methods).

These results suggest that strains I and J have either structurally impaired, or no type IV pili; this would partly explain the results obtained in the biofilm assay. If PA01 is unable to twitch, under nutrient limiting environments it may struggle to survive and exert its full virulence by failing to form biofilms or respond to environmental queues and QS signals. To further elucidate whether these phenotypes were the result of loss or mutated type IV pili in strains I and J, scanning electron microscopy (SEM) images were obtained. Images of the WT strain producing type IV pili

were as a standard to compare with images of strains I and J type IV pili. Obtaining these images was a difficult process as it has been previously noted that *E. coli* only produced their type IV pili in and on certain media types and conditions[205]. Furthermore, type IV pili are fragile structures, that, when handled incorrectly during their preparation and fixation onto the matrix grid for imaging, would break off into the liquid LB in which strains were grown and not be present on the imaging grid resulting in no type IV pili images. After several attempts and trying various different grids, type IV pili were eventually visualsed (Figure 3.7).

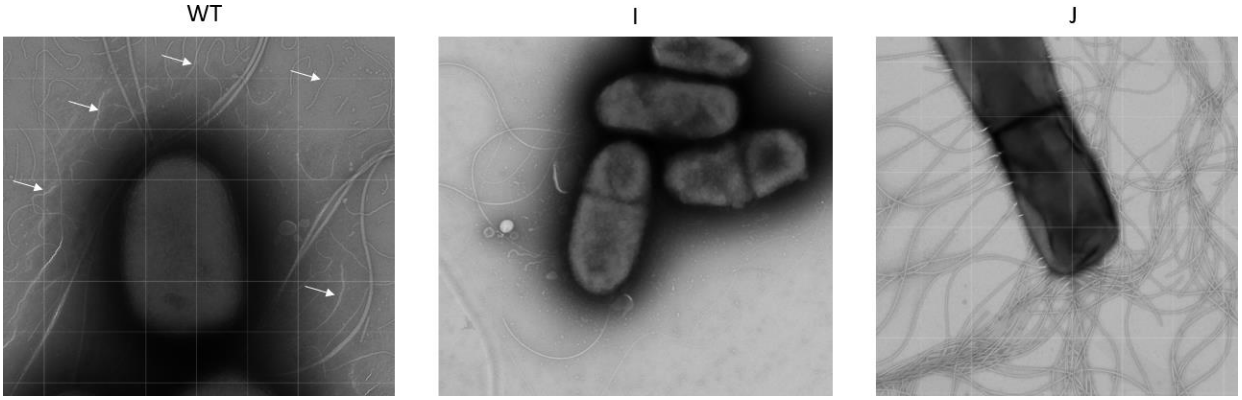


Figure 3.7. Scanning electron microscope images of PA01 cells highlighting type IV pili fixed to a carbon grid. The white arrows in the image labelled WT are pointing to the type IV pili, many of which have separated from the cell. This image is representative of what was seen across the entire grid. Both images for strains I and J show no type IV pili, and this was the case for all areas of the grid. Many further image replicates were captured (Appendix Image 1.1).

The SEM images clearly show the WT possess type IV pili when compared to both strains I and J which are missing the appendage. These images help to elucidate why in the twitching assays, no twitch movement was observed in strains I and J, while also partly explaining why in these strains, biofilm formation was significantly attenuated. What can also be seen from the images is the irregular flagella in the strains leading to the next set of assays performed.

3.3.5 Swarming motility assay

Swarming is a flagella dependent form of movement across a semisolid surface (0.55-0.8% agar) that requires flagellar motility and the production of biosurfactants such as rhamnolipids. This is the rapid coordinated movement across a viscous surface mediated by the flagellum in response to other groups of cells, nitrogen limitation and certain amino acids [206, 207]. Like twitching, this movement type is important on host epithelial cells and plays a role in virulence with recent observations suggesting it plays a part in the adaptive antibiotic resistance against gentamycin and ciprofloxacin. Here, we sought to find out whether our PA01 Mip-like genes play a role in this form of movement. Using swarming plates, we compared the WT to all knockout strains (Figure 3.8). Swarming phenotypes are subjective and based on observation rather than any calculable measurement to perform statistical analysis on.

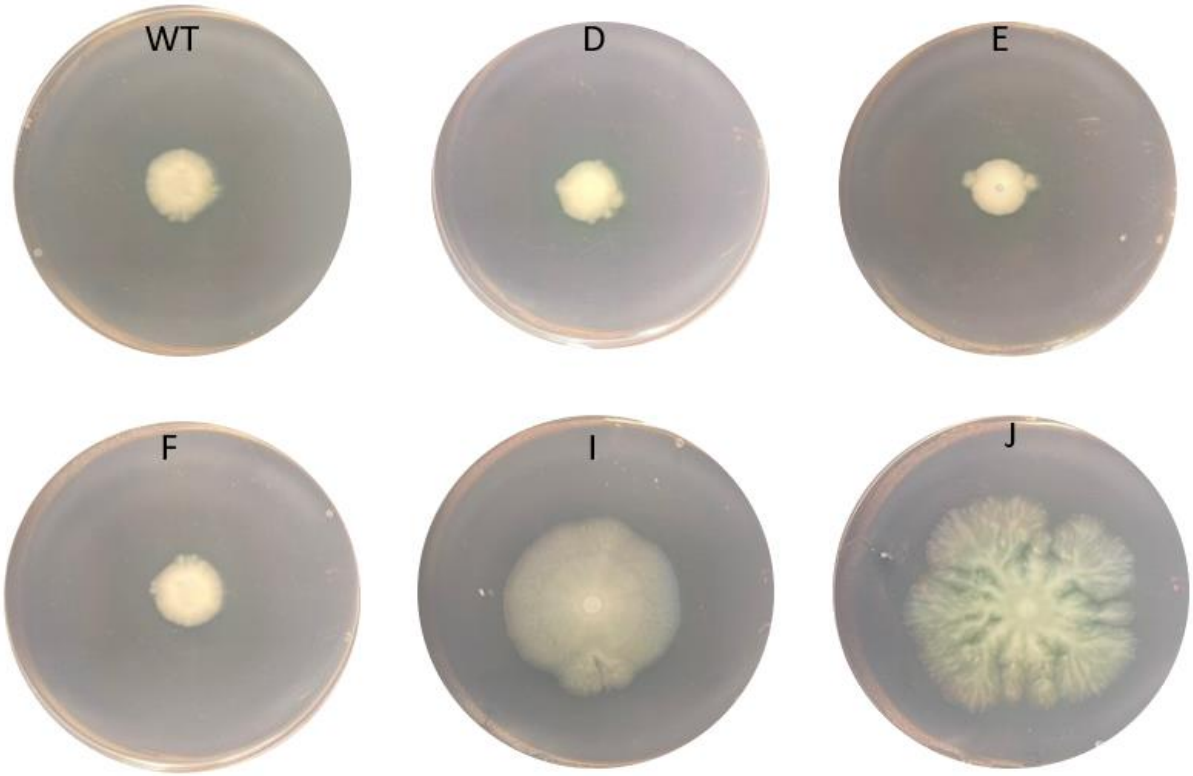


Figure 3.8. Swarming motility plates for the WT and Mip knockout strains D, E, F, I and J strains. These plates compare the WT to single mutant strains D, E and F and the double mutant strains I and J. Single mutant strains display the same phenotype as the WT. Strains I and J show vastly different phenotypes compared to the WT. Between I and J the phenotypes were strikingly different, with J presenting a flower pattern, where as I a much larger, more uniformed circle.

As per other assays performed, strains I and J showed drastic phenotypic deviation from the WT, with single Mip-like gene knockouts displaying no phenotypic difference. The swarming of WT and the single strains is uniform and spreads out over the plate in a structured circular fashion (Figure 3.8), whereas the movement of I and J is appears more erratic and without structure or proper coordination. Even between strains I and J there were phenotypic differences, which proves interesting as they are comprised of different Mip gene knockout combinations, with one shared Mip-like gene, *pa4572*. An interesting observation is that strain J swarming is highly similar to the swarming observed for PA14, a different clinical *P. aeruginosa* strain. This suggests that each PA01 Mip-like gene may possess its own unique function, rather than there being redundancy between them.

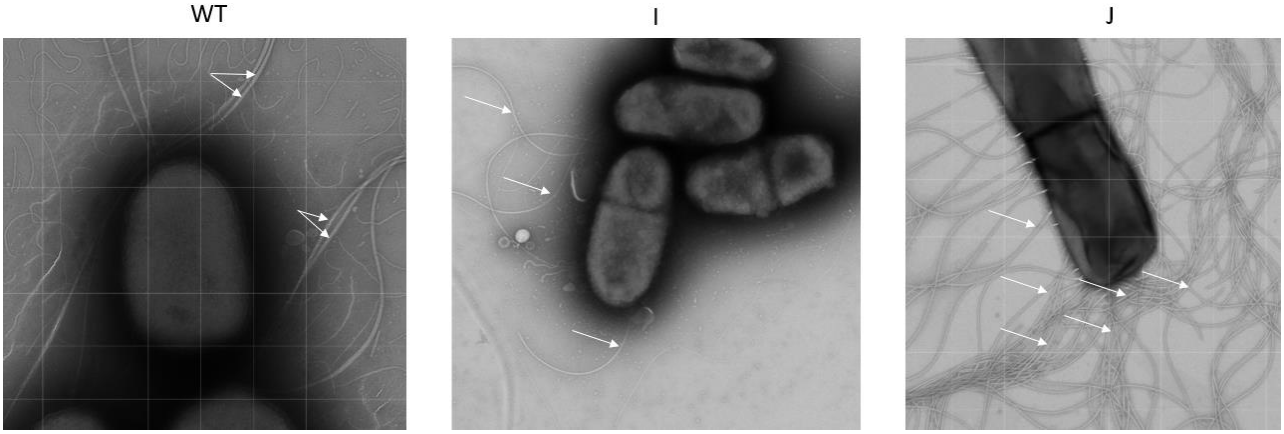


Figure 3.9. Scanning electron microscope images of PA01 cells highlighting flagella fixed to a carbon grid. The white arrows in the images point the flagella. The WT is displaying the flagella phenotype known as twinning (flagella in pairs). Strain I possess a singular flagellum which appears unhealthy and not as pronounced as the WT. Strain J displays hyperflagellation, a bizarre phenotype. This image is representative of what was seen across the entire grid and further image replicates were captured (Appendix Image 1.2)

From the SEM images obtained, the abnormal flagella expression in both I and J have somehow contributed to the phenotypes observed on the swarming plates. Strain I is lacking the twin flagella often associated with optimal swarming motility [208]. Having only one flagellum may partly explain the obscure swarming phenotype observed (Figure 3.9). Strain J presents an incredible hyperflagellation phenotype. Given twinning flagella in the WT have been evidenced to provide a structured, correct swarming phenotype [208], with strain I and its singular flagellum returns a less structured, circular phenotype, it is no surprise the swarming phenotype for strain J is completely abnormal given its hyperflagellation. Ordinarily the twinning flagella allows a coordinated movement pattern on the swarming agar, given strain J presents a large number of randomly located flagella, the completely random movement on agar plates is understandable. For swarming to occur optimally, not only are twinning flagella required, the production of rhamnolipids, biosurfactants that reduce surface and interfacial tension, are also required. Following the observations made here, it was necessary to analyse whether the main rhamnolipids of PA01 were being produced in our strains to further elucidate the swarming phenotypes (**3.4.1**).

3.4 Biochemical analysis and assays

3.4.1 Analysis of rhamnolipid production on swarming plates

PA01 produces a wide range of rhamnolipids often produced in the stationary growth phase where the pathogen may need to explore its environment to scavenge nutrients. Furthermore, rhamnolipids have been shown to disperse through biofilm matrices leaving behind channels through which nutrients and oxygen can flow through into the biofilm. The precursor to rhamnolipids, 3-(3-hydroxyalkanoyloxy)alkanoic acids (HAAs) catalysed by rhamnosyltransferase, encoded by the rhlAB operon, have been shown to be heavily involved in surface movement. The two most abundant rhamnolipid congeners produced involved in swarming motility are MonoC10-C10 and DiC10-C10. These can be analysed using mass spectrometry (MS) and MonoC10-C10 displays a mass to charge ratio (*m/z*) of 503.3226 for the $[M+H]^+$ ion (which shifts to 527.3026 for the $[M+Na]^+$ ion), and an *m/z* of 649 for the DiC10-C10 $[M+H]^+$ ion (which shifts to 673.3608 for the $[M+Na]^+$ ion). It has previously been shown that differences in rhamnolipid production can modulate different swarming patterns, and a *rhlC* mutant unable to produce DiC10-C10 exhibited irregular swarming movement [207].

Given the irregular swarming patterns exhibited by strains I and J, coupled with their irregular flagella, it was hypothesised that their rhamnolipid production would differ to that of the WT. To test this hypothesis two methods were used, the first was imaging mass spectrometry (IMS) with matrix-assisted laser desorption/ionisation time-of-flight (MALDI-TOF imaging), this was performed on a Synapt G2-Si mass spectrometer (Figure 3.10). The second method used extraction of the lipids from swarming agar-plugs which were then analysed by the LCMS looking for specific peaks corresponding to each rhamnolipid that could be quantified (Figure 3.11).

The IMS method enabled us to directly visualise rhamnolipid production and observe how far it had dispersed across the swarming agar relative to the initial inoculum. This was achieved by placing a glass microscope slide at the bottom of a petri dish and pouring a thin layer of agar over the top. The culture was then pipetted onto the agar to one side where the glass microscope slide began, as opposed to the swarming plates described above where I was pipetting the culture in the centre of the plate (Figure 3.10). After the glass slide was carefully cut out, ensuring the agar remained intact on top of the slide with the initial culture inoculum. This was given to Dr Carlo Martins from the JIC proteomics platform who performed all of the IMS analysis.

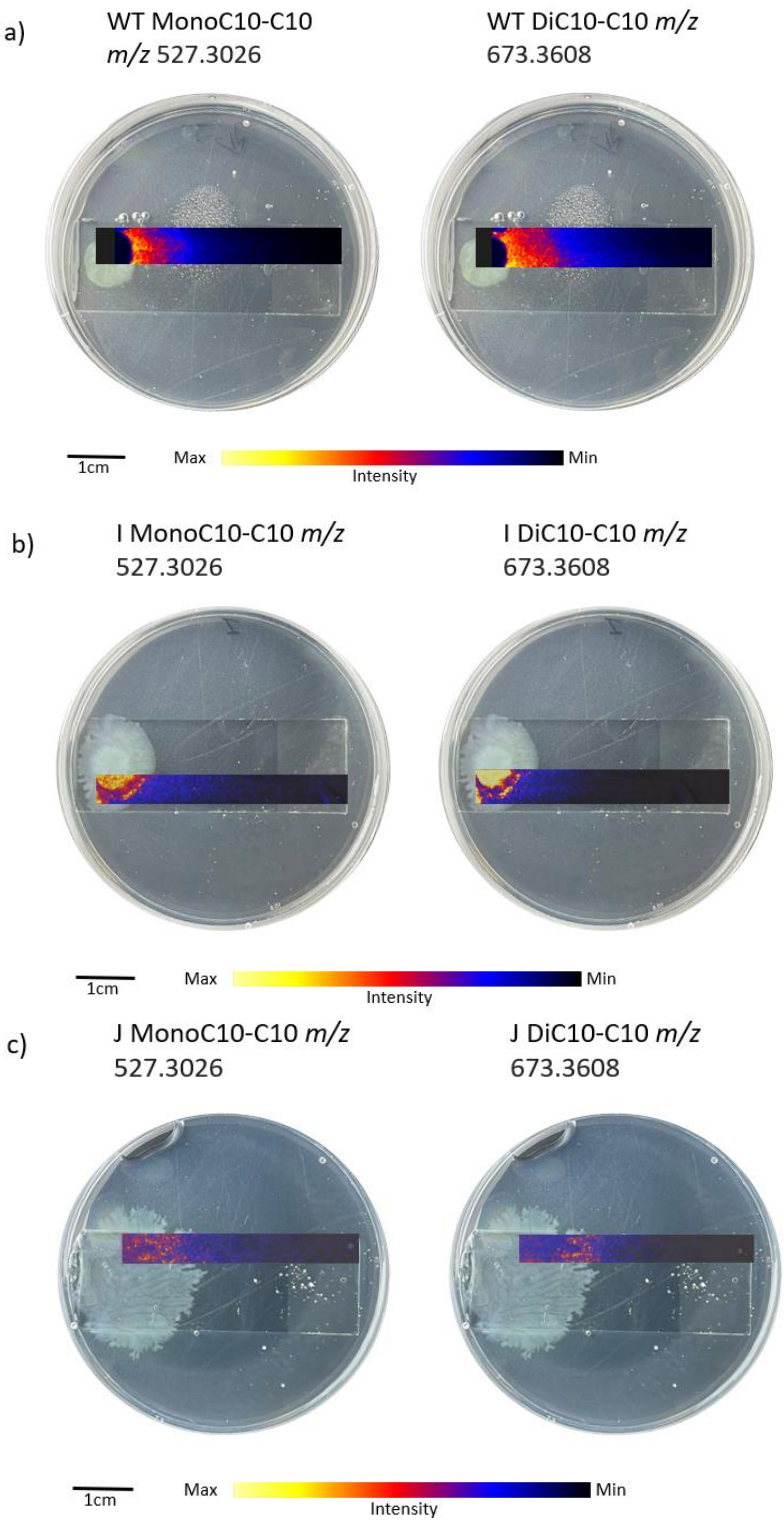


Figure 3.10. Imaging rhamnolipid production of the WT and knockout strains I and J; Images obtained for rhamnolipid production using imaging mass spectrometry on the Synapt G2-Si mass spectrometer. Exact specifications for imaging can be found in methods (**2.7.1)**. Each image shows a photograph of the swarming plates superimposed with images obtained from the MALDI for the rhamnolipids of interest; on the left side of each panel is MonoC10-C10 lipid and on the right DiC10-C10 lipid. Rhamnolipid ion intensity and their dispersal distance from the initial inoculum can be established using the intensity scale bar and distance scale bar beneath each image. Panel a) shows WT plates. Panel b) shows the strain I plates. Panel c) shows the strain J plates.

The images in figure 3.10 suggest that the WT strain has higher levels of rhamnolipids closer to the initial inoculum when compared to both mutant strains (see the greater colour intensity). However, at around 1 cm and beyond from the initial inoculum of all strains, including the WT, there were little to no detectable rhamnolipids, suggesting their dispersal range across the agar is limited regardless of their production intensity near the initial inoculum. To further establish the production intensity and dispersal of rhamnolipids, extractions of rhamnolipids from swarming plates were performed and LCMS methods used to quantify their levels in each sample. Extractions were performed by taking equal sized plugs from swarming plates using sterile 20 µl pipette tips, these were placed into sterile Eppendorf's containing 2 ml MeOH and allowed to mix for 10 min. Following mixing 1ml of the MeOH was removed from the Eppendorf was sterile filtered through 0.45 µm porin filters into a fresh Eppendorf and labelled according to the plate, strain and number plug it originated from, these samples were provided to Dr Carlo Martins for LCMS analysis.

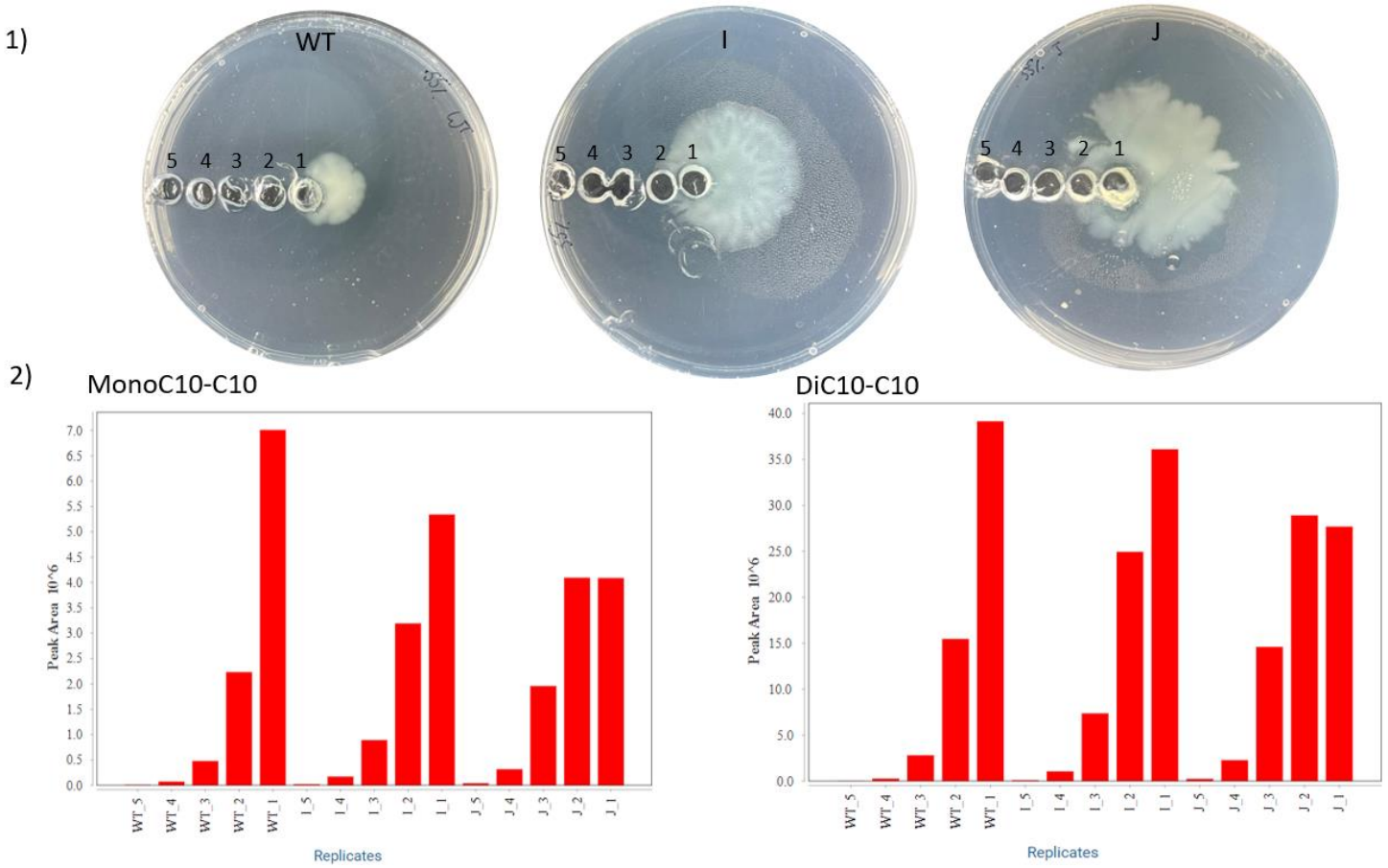


Figure 3.11. Assessing rhamnolipid production of the WT and knockout strains I and J; Panel 1) shows how agar plugs were taken from each agar plate beginning with plug 1 closest to the initial inoculum. Plugs were then taken moving outward to the edge of the plate with plug 1 closest to the centre where the initial inoculum was plated, with plug 5 being the last plug at the edge of the plate. Plugs were removed using the circular end of sterile 20 µl pipette tips and extracted using MeOH. Panel 2) Graphs prepared using data obtained from samples injected into the Synapt G2-Si mass spectrometer equipped with an Acquity UPLC (Waters) showing rhamnolipid intensity for MonoC10-C10 (left) and DiC10-C10 (right) from each plug taken. Each bar corresponds to a plug taken in image 1).

The rhamnolipids extracted from swarming plate plugs were injected onto an Acquity UPLC® BEH C18 column, 1.7 µm, 1x100 mm (Waters) and eluted with the following gradient (mobile phases (A) water/0.1% formic acid; (B) acetonitrile/0.1% formic acid) at a flow rate of 0.08 ml/min at 45 °C. Quantification was performed in Skyline (v.21.2.0.536), using the precursor mass of the rhamnolipids and their isotopic distribution for delimitation of the peaks. The results were consistent with the IMS images obtained in Figure 3.10. The level of rhamnolipids present was higher for plugs taken directly adjacent to the initial inoculum, with a greater intensity seen for the WT at plug 1 (Figure 3.11). However, for plugs 2-5 in each strain the level of rhamnolipids drops off drastically, including for the WT. Furthermore, rhamnolipid DiC10-C10 is produced to a much greater intensity for all strains compared to MonoC10-C10 which agrees with observations in other literature [209]. These results suggest that although the WT strain may produce slightly higher levels of rhamnolipids close to the initial inoculum, when compared to the mutant strains, the levels produced by the mutant strains are only slightly attenuated. However, their spread and dispersal away from the point of inoculation is comparable to the WT.

3.4.2 Analysis of siderophore production

As previously mentioned, the ability of PA01 to collate iron is essential to its growth. In iron limiting environments PA01 produces extracellular compounds called siderophores, that bind free iron which is then actively taken back up through bacterial iron transport mechanisms. There are two main siderophores produced by PA01, pyoverdine and pyochelin, both of which are unstable. Production of pyoverdine can been visualised by the yellowish green fluorescence of cultures [210]. Pyochelin, the more unstable of the two molecules displays a slightly less fluorescent colour with its presence in media undistinguishable unlike the fluorescence and colour given off by pyoverdine (Figure 3.12). As siderophores are a key component of PA01's ability to grow in iron limiting environments, we examined whether our strains were able to produce either or both in iron limiting media. Due to how unstable they both are, they each had different extraction processes and quantification methods. Growing the strains in iron limiting media resulted in varied phenotypes (Figure 3.12) with strain I displaying significantly less fluorescence than the WT, suggesting it produced little to no functional pyoverdine. This was shown to be the case after measuring the pyoverdine content using a Lambda 35 UV spectrophotometer which recorded if pyoverdine was produced as in the presence of iron, a complex is formed causing a modification of absorbance, which sees the shift of 360 nm shoulder to around a 380 nm peak (Figure 3.13).

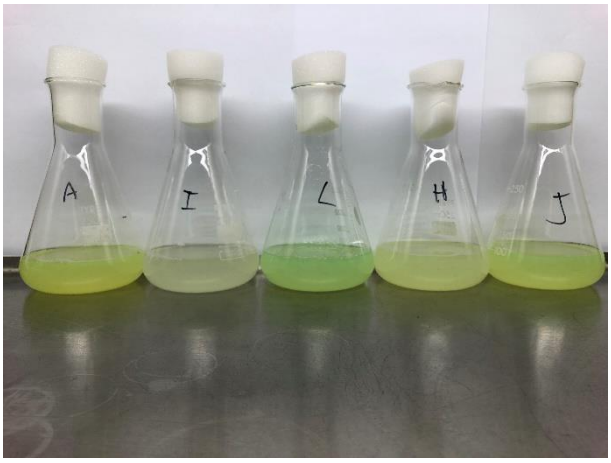


Figure 3.12. Example image of different coloured cultures as a result of siderophore production. Different phenotypes of the WT strain (labelled A) and strains H, I, J and L when grown in succinate media an iron limited media. Strains I and H have a significantly less fluorescent appearance suggesting that little to no pyoverdine is produced when compared to the WT. Strain J appears most similar to the WT suggesting pyoverdine production.

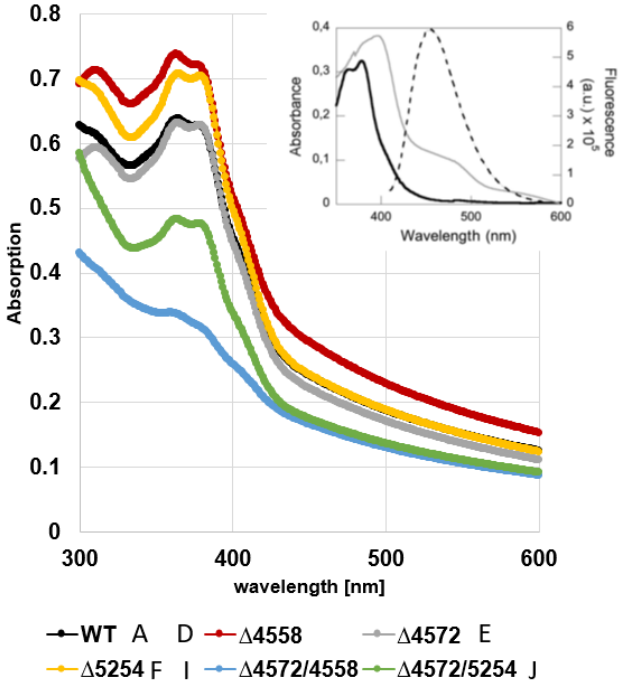


Figure 3.13. Graph presenting Lambda 35 UV spectrophotometer readings for strains D, E, F, I and J searching for the presence of pyoverdine. Spectrophotometer was set to wavelength range 300 to 600 nm and a base level established. The top right image is the expected shape of the absorbance curve in the solid black line for pyoverdine with a shoulder at around 360 nm which shifts to 380 nm in the presence of iron. If pyoverdine was in the presence of iron, a complex would form and cause a change in the absorbance spectra, with the 360 nm shoulder disappearing and a shift to the 380 nm peak.

Figure 3.13 shows UV absorbance spectra for the growth media of the various strains. The spectra of the WT (strain A) and each single Mip deletion strain D, E, F and J have the characteristic change in shoulder peaks associated with the presence of pyoverdine in complex with iron, shifting from 360nm to 380nm. Strain I does not have either of these peaks and looks relatively flat compared when compared to the WT spectra, matching with what was seen phenotypically (Figure 3.12) suggesting no pyoverdine was produced in this strain. These results

were difficult to replicate due to the unstable nature of pyoverdine coupled with other variables that may have affected results. One variable that was investigated was the potential source of extra iron other than that added to the media at specific concentrations. During the time we carried out this experiment, the media kitchen changed how they washed glassware which could have resulted in excess iron initially in flasks in which we made media. Due to this, results were difficult to replicate, especially phenotypes observed in flasks (Figure 3.12). To ensure no iron was present other than that in the media we decided to sterilise and clean all the flasks ourselves to negate any contamination. Upon doing this, results appeared to be more consistent from which we were able to replicate the data shown in Figure 3.13 twice.

Measuring pyochelin was a difficult process due to its unstable nature. After following several different extraction protocols, it was not possible to obtain enough stable pyochelin to consistently run on the HPLC for analysis.

3.5 *In vivo* infection assays

Drawing together the *in vitro* data, it was prudent to assess whether these data translated across into an *in vivo* model. Our *in vitro* data suggests that some strains have lost or are severely attenuated at carrying out several important virulence mechanisms. This theoretically, should correspond to a demonstratable loss of virulence in a eukaryotic host.

3.5.1 *Galleria mellonella* infection model

The *Galleria mellonella* greater wax moth are simple invertebrates and are one of two species of wax moth worm that are well known for their parasitzation of honeybees. Recently they have become an excellent tool and infection models to study host-pathogen interactions. Although insects lack an acquired immune system, they do possess an innate immune system that is structurally and functionally similar to vertebrates. It is rapid and simple to use that requires no ethical approval and provides a cheap alternative to the mouse model which is often time consuming and expensive [195].

In several *in vitro* assays discussed above I have primarily focussed on the single mutant strains D, E and F, and the double mutant strains I and J following the first three *in vitro* experiments suggesting the remaining single (B-C), double (G-H, K-O) and triple mutants (P-Q) were not attenuated in virulence based on *in vitro* assays and presented identical to the WT strain. This was not to suggest that *in vivo* they would not display any attenuation, so now having moved to the first *in vivo* study all mutant strains (B-Q) were included in order to assess all strains virulence and whether previously uninteresting strains presented with an unexpected phenotype. This would allow once and for all to decipher whether previously uninteresting strains exhibited any

phenotypic differences from the WT *in vivo,* a much more important marker for assessing whether genes are involved in virulence than plate-based assays. If these strains remained comparable to the WT they could be completely discontinued from this study. For the infection model, ten wax moth larvae were used per PA01 strain. Each larva was injected with between 0-4 cells into the hind left leg (Figure 3.14.a). To do this, cells were initially grown to an OD of 0.8 then serially diluted in sterile PBS through a log10 series. To do this 100 µl of 0.8 OD culture was diluted into 900 µl PBS, and the resulting dilution was diluted six further times to accomplish a serial dilution of 10^6 (Figure 3.14.a). It was calculated this should result in between 0-8 CFU when plated, however, upon plating to ensure this calculation was a reality, CFU of between 14 and 47 were obtained. Therefore, through trial and error, the optimal starting culture was found to be 0.05 OD, from which the 10^6 serial dilution was performed and when plated gave between 2-8 CFU. Despite all the protocol modification, in reality the number of cells injected could not be exact. However, using the serial dilution (10^{-6}) of an 0.05 OD culture proved the most accurate way of ensuring that a tiny bacterial load (~0-4cells) was injected, preventing the larvae from immediate or quick deaths due to high bacterial load. Once inoculated, the larvae were incubated at 30°C with readings taken every 6 h. These readings were recorded over a 48 h period at which point the infection assay was deemed finished. Categorising whether the larvae were dead or alive was determined by their turning black due to melanisation when dead compared to their normal white appearance (Figure 3.14.d) [211]. This melanisation is an innate part of an insect's response to a pathogen; it produces melanin to encapsulate a pathogen preventing its further spread. Only when a larva was completely black and no longer moving was it recorded as dead.

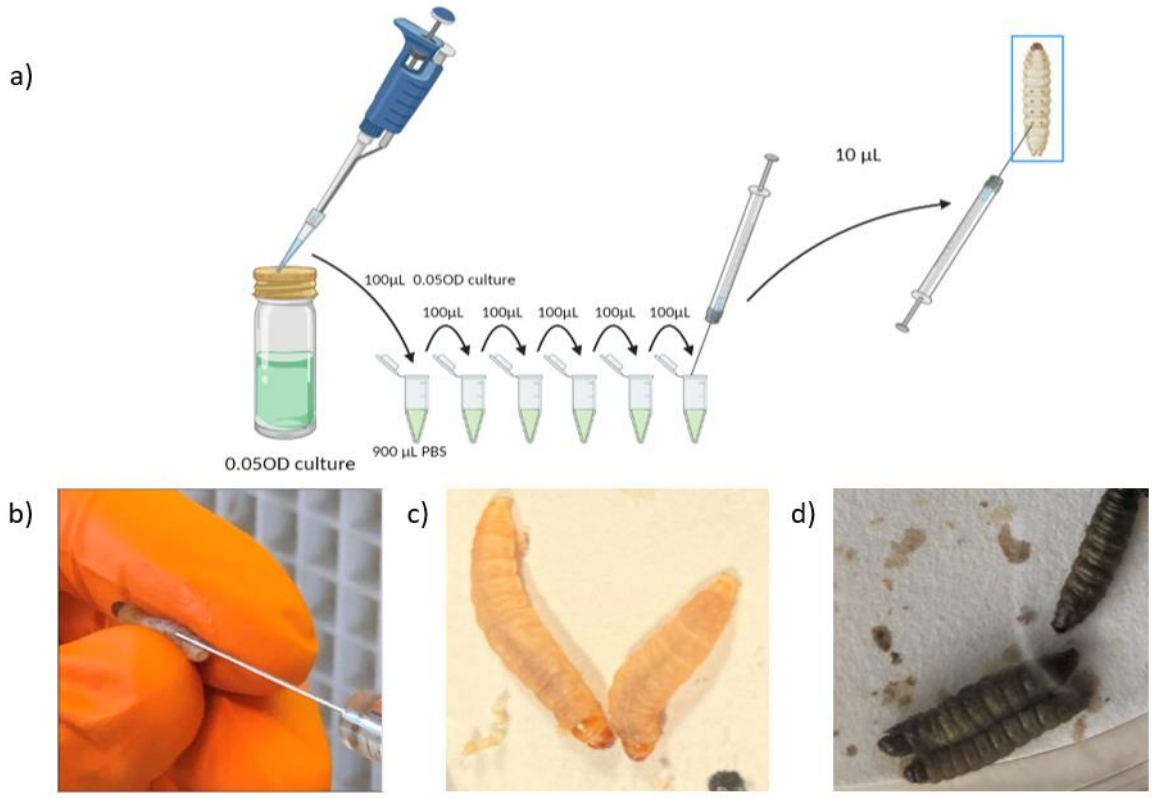


Figure 3.14. *G. mellonella* infection model diagrammatic representation; Image a) is a diagrammatic representation of the serial dilution performed using 0.05 OD culture to inoculate the *G. mellonella* larvae in its hind left leg. Image b) example inoculation of *G. mellonella* using a 10 µl Hamilton glass syringe. Image c) example of a healthy *G. mellonella* larvae pre-inoculation. Image d) example of dead *G. mellonella* larvae post-inoculation and incubation resulting in the characteristic melanisation causing the colour change to black and classification as dead.

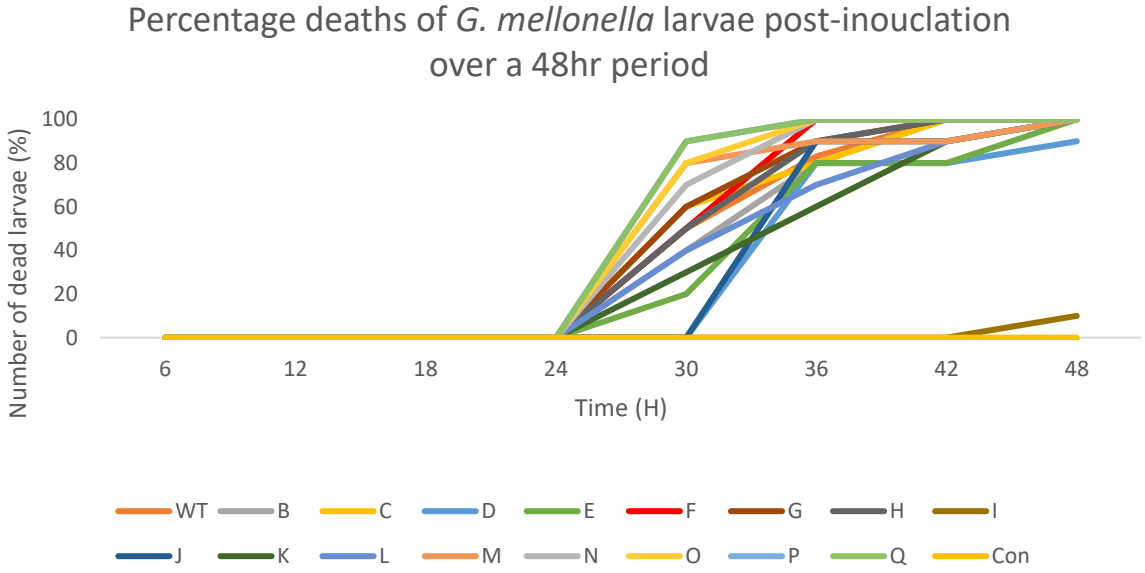


Figure 3.15. Line graph plotting percentage deaths of *G. mellonella* inoculated with Mip knockouts strains B-Q. Percentage of deaths in wax moth larvae for all mutant strains over a 48 h period with recordings taken at 6 h intervals. This experiment was performed twice with results pooled. Each time the experiment was performed 10 larvae were used per strain. Final deaths were recorded as a percentage of deaths.

It is evident from the graph in Figure 3.15 that deaths only begin to occur after 24 h post injection for the WT, with deaths reaching 100% after 42 h. For the first time this assay indicated a slight deviation in phenotype for the single mutant strains D, E and F when compared to the WT. Their killing of the larvae was delayed, starting at around 30 h, 6 h later than the WT, and reaching 100% deaths after 48 h and therefore both indicators were slightly delated when compared to the WT. Strain J followed the same trend as D and E with a delayed death onset and final 100% death rate at 48 h. Finally, strain I was markedly different and only displayed a killing effect in 10% of the larvae inoculated and only doing so after 42 h, remarkably, 90% of larvae still alive at the 48 h mark. This delayed and lack of killing is remarkably reminiscent to the data obtained in a mouse infection model using the *B. pseudomallei* BpML1 Mip mutant described in the introduction.

It is important to note, however, a possible issue causing variation of results using an infection model such as this is the larvae can be fragile and some deaths can be attributed to puncture wounds sustained from the initial inoculum injection. Equally, the delayed-on set of deaths at 30h, along with the few deaths seen for strain I, could be down to the initial inoculum having fewer cells than others following the 10^6-serial dilution. To account for this potential variation, this experiment was carried out twice, each time 10 larvae were used per strain to try and negate these effects. Furthermore, on both occasions the control group, inoculated with sterile PBS resulted in zero deaths, suggesting good technique when injecting the larvae and giving confidence that it was unlikely that deaths were as a result of puncture wounds. For both sets of experiments, similar results were obtained giving confidence as to what was seen was legitimate. These results further suggest that phenotypes observed on plates are significant and that virulence factors of PA01 are severely attenuated as a result of Mip-like gene knockouts. This *in vivo* experiment also suggests the phenotypes observed using *in vitro* assays (i.e. lack of biofilm formation and haemolysing blood cells *in vitro*) are transferable across to an infection model, partly explaining why some strains had a delayed killing time or were overall poor at killing. Given in this assay, coupled with the first three *in vitro* assays performed, only single mutant strains D, E and F, and double mutant strains I and J displayed any phenotypic deviation from the WT, the remaining single, double and triple mutants that presented with no phenotypes associated with attenuated virulence, they were not studied any further.

3.5.2 Macrophage infectivity model

Given that Mip genes were initially termed as such due to their discovery in *L. pneumophila* mutants that were attenuated in invading and colonising macrophages, I determined that it was essential to assess whether the mutants generated were affected in their ability to invade and colonise macrophages. This work was a collaborative effort performed with Sonia Paz (Medical School, University of East Anglia) who has had previous experience at performing assays with RAW 264.7 cells, a macrophage cell line established from a tumour in a male mouse induced with the Abelson murine leukaemia virus [212]. The purpose of this assay is to assess whether pathogenic strains can colonise macrophage cells, and once colonised are able to survive and replicate. This is an essential step in the virulence of an intracellular pathogen such as PA01, without which, it would struggle to mount an attack on the host. Furthermore, upon entry into macrophages, PA01 is able to evade killing and clearance by phagocytes which contributes to persistent infections, something that CF patients routinely suffer [213]. This assay involved inoculation of the RAW 264.7 cells with PA01 strains and incubated for 3 h. After this the cells were washed, incubated again in the presence of gentamicin for 1 h to kill any PA01 cells that had not entered the macrophage cells and remained in the culture. Following a further wash step the cells were lysed open using sterile water. The resulting lysate should contain any PA01 cells recovered from the macrophages that had been able to enter, colonise survive and replicate within the macrophages. This lysate was plated on LB agar and incubated at 37^{o}C O/N. Any PA01 colonies on the plates were then counted the next day to determine the virulence of each strain compared to the WT.

Prior to inoculation of the RAW 264.7 macrophage cells with PA01 strains, the fitness and health of the RAW 254.7 cells was determined using a calcium mobilisation assay [214]. In the RAW 264.7 cell type there is always a certain concentration of calcium ions ($Ca2^{+}$). To check the health of cells, an aminopolycarboxylic acid dye (Fura2) which binds to intracellular free $Ca2^{+}$ ions and causes fluorescence, was used to determine the initial amount of calcium in the cells, with too much or too little $Ca2^{+}$ ions suggesting the cells were not healthy. The fluorescence also allowed us to visualise the size and shape of the cells, an important indicator of cell health with the optimal shape of cells appearing oval rather than circular. Imaging the cells using light microscopy revealed whether cells were healthy and not circular in shape which would have suggested the cells were dead, plus the images obtained revealed cells were not saturated with green fluorescence which would have indicated that calcium levels were too high (Figure 3.16.a). When provided with ATP, these cells are promoted to increase the uptake of $Ca2^{+}$ ions which further checks that cells are able to respond normally to agonists (in this case ATP), which again can be visualised by an increase in green fluorescence (Figure 13.16.b). This increase in

fluorescence was then able to be measured, with a brief peak observed approximately 50 sec post ATP exposure followed by a drastic drop (Figure 3.16.c/d). All cells analysed were confirmed to be healthy as they were able to uptake $Ca2^{+}$ and were therefore ideal to continue forward for use in the PA01 macrophage assay.

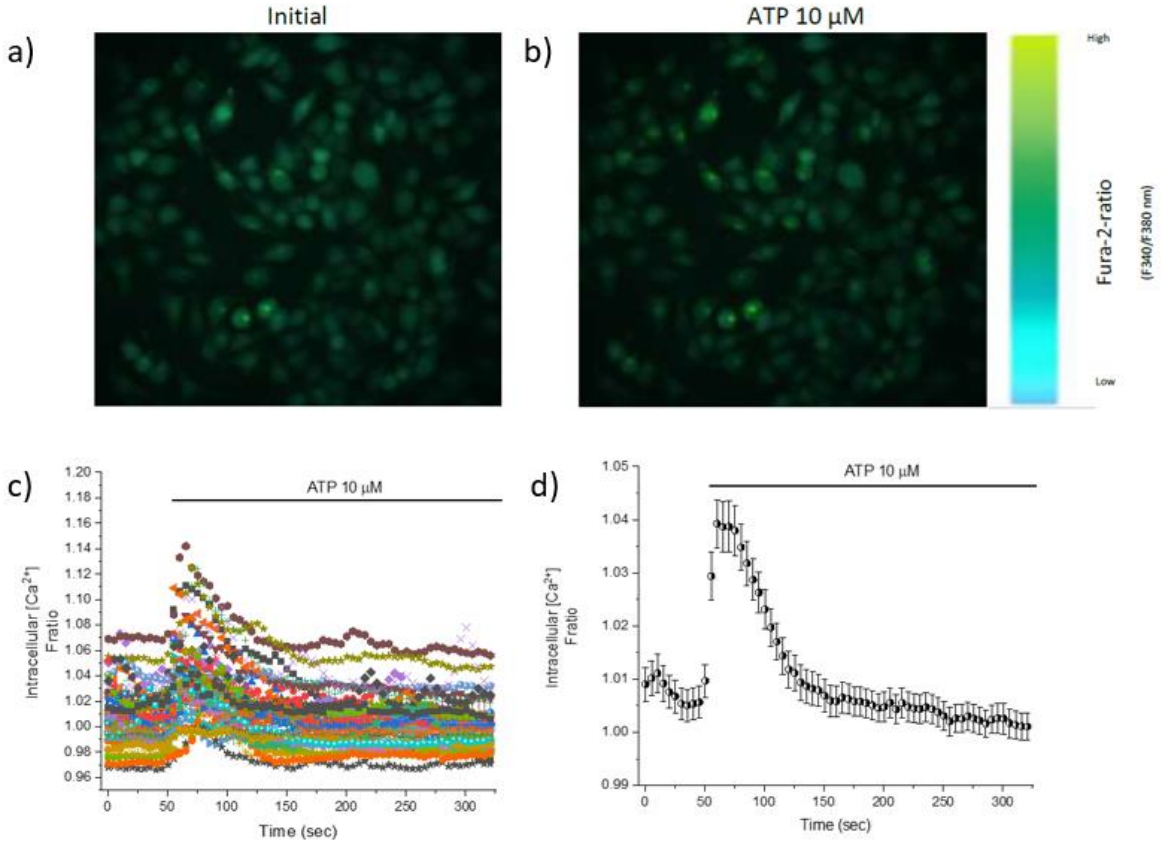


Figure 3.16. Results for the analysis of the fitness and health of RAW 264.7 macrophage cells used in the infection assay. Panel a) shows a light microscope image of RAW 264.7 cells prior to their introduction to ATP with a more subtle turquoise colour coupled with their healthy oval morphology, if the cells were unhealthy, they would appear round. Panel b) shows upon the introduction of ATP cells sequester more $Ca2^{+}$ ions and become greener in appearance. Panel c) is a graph plotted using fluorescent readings obtained from each cell that can be seen in the image in panel a) measured over 300 sec following their introduction to ATP. Each individual cell shows the same trend, with a peak in fluorescence at around 50 sec, demonstrating the uptake of $Ca2^{+}$ in response to ATP, followed by a drastic drop in fluorescence as channels close in response to ATP levels dropping over time. Panel d) shows the data obtained for each individual cell in panel d) a graph of the average readings for each individual cell in panel c) at each timepoint, taking their average fluorescent reading and plotting it overtime. This highlights the trend that RAW 264.7 cells respond to ATP by an increase in uptake of $Ca2^{+}$ which decreases over time. These data suggest cells were perfectly healthy for the infection assay.

Once cells were confirmed to be healthy and fit for the assay, they were pipetted into 96-well plates, with each well containing 20,000 cells per 200 µl. The first few experiments required trial and error to determine the optimal multiplicity of infection (MOI) to use. Multiplicity of infection is the ratio of pathogenic cells to the number of macrophage cells in a single well. The MOI of PA01 cells to use was calculated using serial dilutions starting from a 0.05 OD culture in a similar fashion to the *G. mellonella* infection assay. Initial runs used MOIs of 0, 0.1, 1, 10 and 50 (Table

3.2 and Table 3.3). Upon this analysis, MOIs of 0, 10 and 50 were settled upon for final experimental runs. Here only the WT, D, E, F, I and J strains were studied, with each strain performed with four replicates per 96-well plate with the experiment repeated twice on separate days.

MOI	Number of macrophages	Number of PA01 cells
0	20,000	0
0.1	20,000	2000
1	20,000	20,000
10	20,000	200,000
50	20,000	1,000,000

Table 3.2. Table describing the different multiplicity of infection (MOIs) ratios used in the macrophage infectivity assay. Macrophage infectivity assays using a variety of multiplicity of infection (MOIs) ratios used in trial-and-error runs prior to final experimental trials, after which MOIs of 0, 10 and 50 were settled upon. How the number of PA01 cells were obtained for each MOI can be found in Table 3.3. MOI of 0 is the negative control for each experiment.

Serial Dilution	Number of PA01 cells in culture	Target number of PA01 cells	Serial dilution culture added (µl)	Sterile PBS added (µl)
$4x10^7$ (starting culture 0.05OD)	~40,000,000	~4,000,000	100	900
$4x10^6$	~4,000,000	~1,000,000	250	750
$4x10^5$	~400,000	~200,000	500	500
$4x10^4$	~40,000	~20,000	500	500
$4x10^3$	~4000	~2000	500	500

Table 3.3. Table describing serial dilutions to obtain a given MOI. Dilutions and volumes of culture calculated used to obtain required ~ number of PA01 cells for each MOI ratio. Dilutions for each strain began with cultures at 0.05OD which contained ~40,000,000 ($4x10^7$) cells in 1 ml. From these initial cultures, 100 µl was taken and put into 900 µl sterile PBS giving a serial dilution of $4x10^6$ and ~4,000,000 cells. From this serial dilution the number of cells required for an MOI of 50 was 1,000,000 so 250 µl was taken into 750 µl sterile PBS giving ~1,000,000 Pa01 cells. From the $4x10^6$ serial dilution, further serial dilutions of $4x10^5$ (~400,000 cells), $4x10^4$ (~40,000 cells) and $4x10^3$ (~4000 cells) were made up. Using these serial dilutions, the target number of cells for each MOI were made up by adding 500 µl serial dilution culture and added to 500 µl sterile PBS.

Three 96 well plates were used each time the experiment was performed, with each 96 well plate housing MOIs of either 0, 10 or 50. The MOI of 0 provided the negative control for each experiment. In each plate there were four replicate wells per strain (WT, D, E, F, I and J). Upon reaching the protocol step for macrophage lysis and recovery of PA01, one agar plate was used

per well lysate, (4 agar plates per strain, 24 agar plates per MOI, 72 plates per experiment, experiment performed twice, 144 agar plates in total). Once plated, they were incubated at 37°C O/N. The following morning colonies were counted on each plate (Figure 3.17). In both replicate experiments the negative control plates with MOI 0 had no colonies as expected, this ensured no contamination of wells had occurred during the macrophage inoculation stage. Although the MOI 0 plate in Figure 3.17 appears to have a small number of colonies present, this is a visual artifact and they did not possess any colonies across any MOI 0 replicate plates.

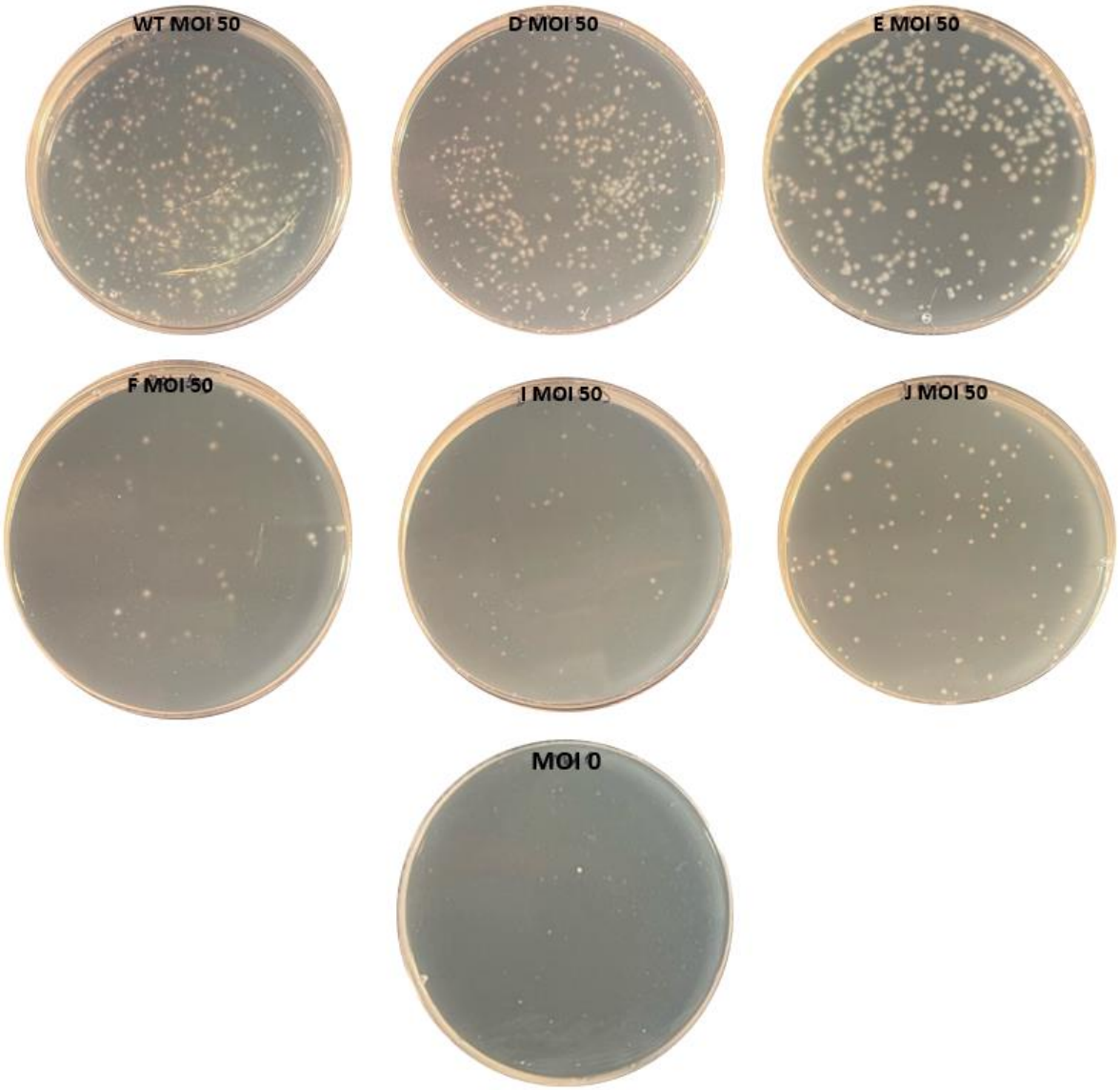


Figure 3.17. Macrophage infectivity assay plates for strains WT, D, E, F, I and J. Plates are spread with macrophage lysate that contained PA01 cells that had colonised, survived, and replicated within host macrophages. The more colonies on a plate, the more successful and virulent a strain was at colonising, surviving, and replicating in host macrophage cells. Each image is replicate of a plate per strain at MOI 50 used for colony counting, with an example of a replicate plate of MOI 0 in the bottom image possessing no colonies, what appear as potential colonies on this plate are bubbles from pouring the LB agar.

The replicate experiments were carried out independently, on different days, and each presented essentially identical across all replicate plates for each strain. Simply by visualising the plates prior to even counting colonies, it was evident that the WT, D and E strains behaved equivalently and as expected i.e. they were able to colonise, survive and replicate within the macrophages. In contrast strains F, I and J returned significantly much lower cell counts and thus appeared to be less virulent (Figure 3.17). This visualisation was more evident in the MOI 50 group which presented more obvious data than the MOI 10 group. Although the MOI 10 group displayed a similar trend to the data obtained from the MOI 50 group and provided statistical significance ($P<0.05$, $n=8$), the number of colonies recovered on the plate were too low to draw a definitive conclusion, most likely due to the small initial inoculum which was five-fold less than the MOI 50 group. However, what is interesting about the data from the MOI 10 group compared to MOI 50 is that when counting the recovered cells, they were always approximately five-fold less than for the MOI 50 plates, the difference in initial inoculum size, thus the number of colonies recovered is dose dependent (Figure 3.18.).

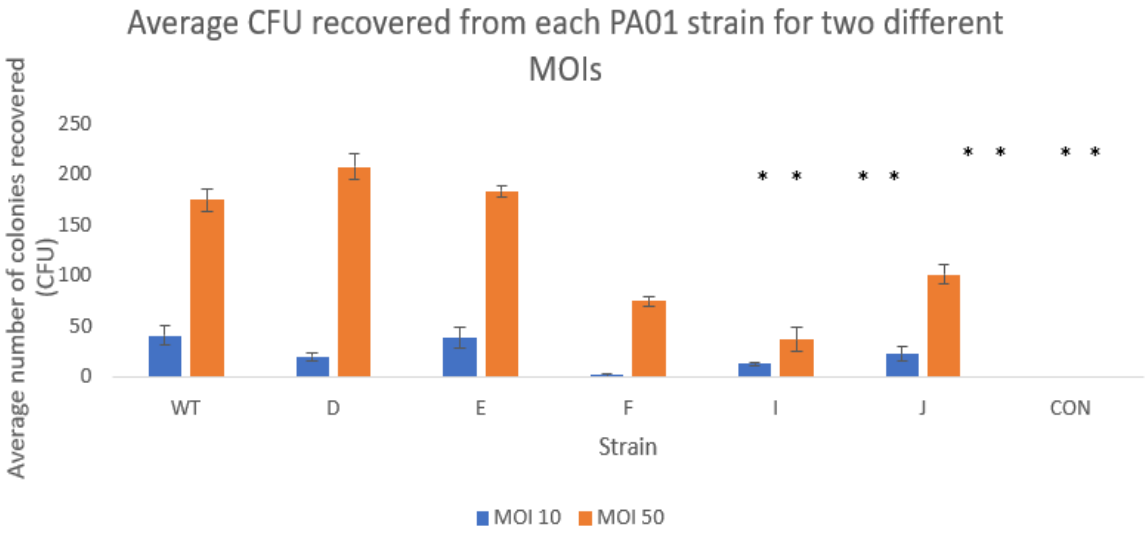


Figure 3.18. Bar chart for macrophage infectivity assay results. The average number of CFU counted on agar plates recovered from the macrophage cell lysate plated up was plotted. In blue presents the MOI 10 group and in orange is the MOI 50 group. Standard error variance was calculated and plotted as seen by the error bars comparing the difference between the mean with the amount of scatter within the groups, error bars that overlap are not statistically significant. The asterisks above columns indicate ($P<0.05$, $n=8$).

This assay was designed to assess the virulence of a strain with a simple readout that the more colonies on the plate that were visualsed and counted, the more virulent the strain was. Upon counting the colonies and performing statistical analysis it was clear that strains F, I and J were significantly attenuated in their ability to enter and colonise macrophage cells when compared to the WT, D or E strains. This was true for both MOIs 10 and 50 tested, although as previously mentioned, the data from the MOI 10 group was not as striking which was most likely due to a lower starting inoculum.

These data tell us that the Mip-like candidates Pa4558, Pa4572 and Pa5254 are important in the virulence processes associated with macrophage infection and colonisation. Of particular interest from this data is for the first time in any assay, strain F has a phenotype that is drastically different than that of the WT. For all of the *in vitro* assays discussed above, its phenotype was the same as the WT, whereas for this assay and for the *G. mellonella* infection model, where it had a slight delayed killing time of 6h. Upon closer analysis, strain F has a *pa5254* gene deletion, which coincidentally is the Mip-like gene in PA01 whose product protein is most closely related to the surface antigen LgMip, with 40% sequence identity. This *pa5254* gene product has been previously flagged as a potential FKBP when its ORF was flagged in several sequence searches performed by groups looking for FKBP homologues which was discussed in the Introduction (**1.10**). Presently, the location of gene products from *pa4558*, *pa4572* and *pa5254* are all hypothetical with both Pa4558 and Pa4572 predicted to reside in the cytoplasm, while Pa5254 is predicted to reside on the outer membrane. These predictions come from the *Pseudomonas* database using the PA01 reference genome which uses amino acid terminal sequences as a best guess as to where the protein would be localised. Given that the LgMip has been studied in depth, and its location on the outer membrane is key to its role as a virulence factor, it is no surprise that at least one potential Mip of PA01 would be located here. The studies of LgMip showed that its expression in the outer membrane was essential for the optimal colonisation and survival of *L. pneumophila* in macrophages [118].

The data presented in this section, coupled with the results obtained from the *in vitro* plate-based assays, further suggests that *pa5254* Mip gene product is located at the outer membrane where it exerts its virulence effects. In the *in vitro* assays, mutant F, which has only a single gene deletion in *pa5254*, has no obvious attenuated virulence phenotypes, this is likely because in these assays there are no host cells to interact with. Furthermore, virulence mechanisms such as haemolysis, protease activity and biofilm formation are mediated by specific transcriptional activators, intracellular signalling, and protein-protein interactions, and are thus unlikely to be affected by a deletion in a gene whose product is most likely located in the outer membrane [215]. However, for assays which require interaction with host cells such as the *G. melonella* model and the macrophage assay, a phenotypic difference is seen. This suggests that the *pa5254* gene product, like its LgMip homologue, may exert its virulence effects by mediating the pathogen-host cell interaction.

What can be concluded from this macrophage assay is that the prospective Mip-like gene candidates *pa4558*, *pa4572* and *pa5254* are all required for optimal colonisation, replication and survival of PA01 in macrophages. Therefore, these genes meet the criteria to be categorised as Macrophage infectivity potentiator encoding genes. On this basis we have re-named them in

line with how other Mips are annotated (e.g. LgMip). Thus, the *pa4558* gene is now annotated with the name *pamip1* (product PaMip1); *pa4572* as *pamip2* (product PaMip2); and *pa5254* as *pamip3* (product PaMip3). When referring to these genes together I will use the collective term; 'PaMips.' In the remainder of this thesis, strains will still be referred to using their alphabetical assignment (D,E,F etc.), but when referring to a particular gene that has been knocked out, its newly annotated gene name will be used (Table 3.4).

Strain alphabetical assignment	**Previous genotype annotation**	**New genotype annotation**
D	PA01Δ*pa4558*	PA01Δ*paMip1*
E	PA01Δp*a4572*	PA01Δ*paMip2*
F	PA01Δp*a5254*	PA01Δ*paMip3*
I	PA01Δp*a4558* / Δp*a4572*	PA01Δ*paMip1* / Δ*paMip2*
J	PA01Δp*a4572* / Δp*a5254*	Pa01Δ*paMip2* / Δ*paMip3*

Table 3.4. Table annotating new nomenclature for PaMip genes described in this study. This Table displays PA01 mutant strains with their assigned letter for reference depending on what gene(s) are knocked out. Initially gene(s) knocked out in each strain were referred to as their chromosomal gene number (central column), however have now been replaced with an annotation in the far-right hand column.

3.6 Growth curves to assess strain fitness

To ensure the phenotypes observed in the experiments described thus far were not due to growth defects, growth curves were performed to assess the fitness of strains D, E, F, I and J in comparison to the WT. Each strain was grown in triplicate in a 96 well plate containing LB. An initial inoculum culture for each strain was grown to an OD of 0.05, and from these cultures 5 µl was added to each well on the 96 well plate containing 195 µl of LB. These plates were then incubated in a Labtech Spectrostar at 37°C and 250rpm for 12 h. Optical density measurements were recorded every 30 min at OD 600 nm.

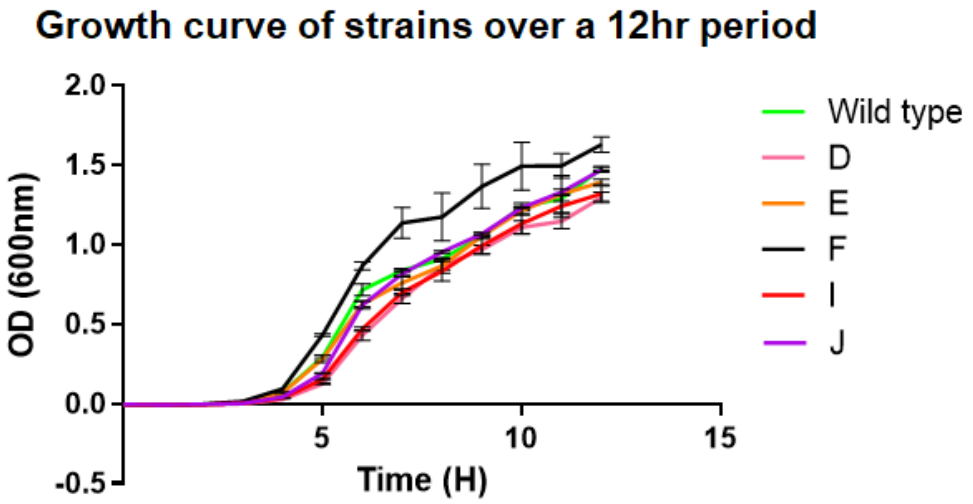


Figure 3.19. Growth curve assessing Mip mutant strains fitness compared to the WT. Growth curves comparing strains D, E ,F, I and J to the WT. Each strain was grown in triplicate with averages plotted. Overlapping error bars signify that there were no statistical differences in growth rates of strains compared to the WT except for strain F which grew to a higher OD and at a slightly faster rate ($P<0.05$, $n=3$).

Based on these growth curve data, it is safe to assume that all strains of interest have no growth fitness defects due to their gene knockouts, with all growing at a similar rate and ending up at similar ODs. The only exception to this is strain F which grows at a slightly faster rate than the WT and reached a higher OD; however, this does not suggest a fitness disadvantage and may serve as a fitness advantage to outgrow competing bacterial cells or improve the virulence of the organism by growing faster and reaching a higher population density sooner. Given all strains grow at a similar rate and are comparable to the WT, it is safe to assume that the phenotypic results obtained *in vitro* and *in vivo* were not because of fitness or growth defects and that they were the result of the Mip-like gene deletions.

3.7 RNA sequencing analysis

Drawing together the plethora of phenotypic data that have been described above as a result of the Mip gene deletions, it suggests that Mips exert their effects in a pleotropic manner and that they may be involved in the regulation of multiple processes at both a transcriptional and a translational level. To further unpick the phenotypes observed, the WT and strains D, E, F, I and J were sent for RNA sequencing (RNA seq) to provide insight to any transcriptome changes that may have occurred as a result of the Mip gene deletions. Transcription is the first step in gene expression where one or more general transcription factors (e.g. RhlR or AlgT) bind to promoter DNA, and the DNA is ‘unzipped’ by RNA polymerase from its normal double stranded structure into single stranded DNA which acts as a template. This single stranded DNA template is ‘read’ by RNA polymerase and copied into a complementary messenger RNA molecule (mRNA) which

is then used to construct a product such as a protein by a process called translation. Each gene within a genome has its own level of transcription, meaning the expression of some genes will lead to more mRNA molecules (transcripts) in the cell than others (Figure 3.20). The more mRNA molecules transcribed from a gene, the more highly expressed a gene is. Not all genes (particularly in bacteria), are expressed at the same time, with different cellular, environmental, physical and chemical signals promoting the expression of genes whose products are more favourable for certain situations [216]. The over-expression or downregulation of a gene is often impacted by the expression or repression of another gene(s). RNA sequencing builds a picture of a particular strain's transcriptome under certain conditions, and this will then be compared to another strain's transcriptome under the same conditions to look for differences. Overall RNA seq analysis provides essential data to allow the interpretation of functional elements within a genome that contribute towards virulence.

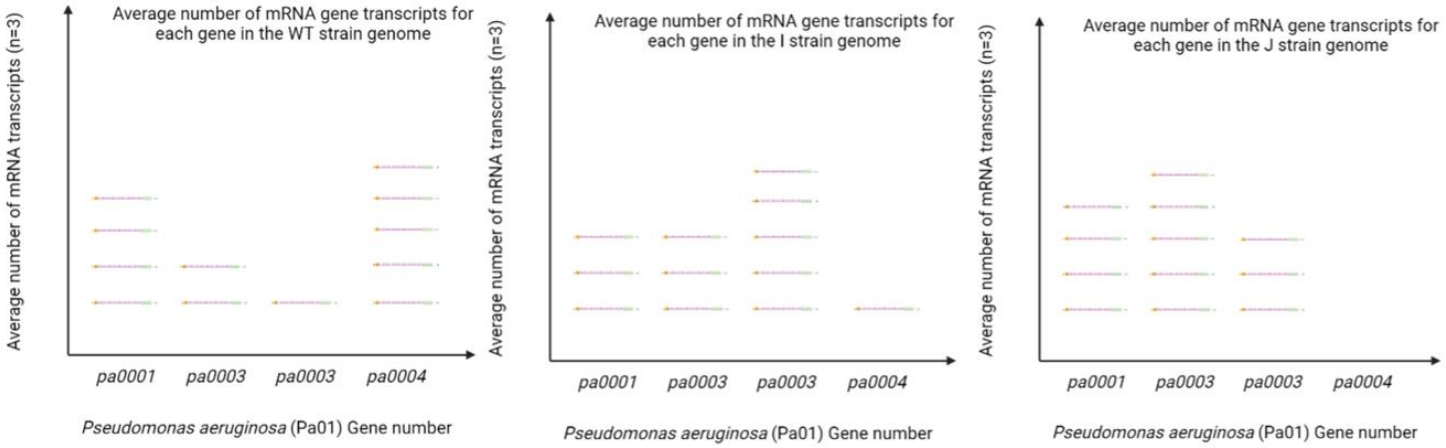


Figure 3.20. Diagrammatic representation of how mRNA transcripts for each gene are compared across different strains. The number of lines represents the average number (n=3) of transcripts for each gene in the genome of a given strain which is then subjected to statistical analysis. The WT readings are used as a baseline to then compare the transcripts of every gene in mutant strains. From this, it is then possible to assess any differences in the transcription of genes between strains and whether these differences are statistically significant. In the diagram, it is clear gene *pa0003* has an increased transcription level in strains I and J when compared to the baseline transcript of the same gene in the WT due to the larger number of transcripts counted. This would suggest this gene is more highly expressed when compared to the WT, subject to statistical analysis ($P<0.05$, FDR<0.05). This hypothetical *pa0003* gene may be involved in the stress response of the organism hence it is more highly expressed in mutant strains I and J which may not be as fit as the WT strain.

In this study three biological replicates were used for each strain to be sent for RNA sequencing. For each strain of interest 16 x 50 µl drops were pipetted on to square agar plates without spreading and allowed to grow O/N at 37°. The next day, the resulting cell material was scraped off the plate into 5 ml Eppendorf tubes and this material used for RNA extraction. Growing the strains on agar plates was used rather liquid culture as I wished to mimic the conditions of the mostly plate-based virulence assays described thus far in order to recreate, as closely as possible, the transcriptome of each strain in a virulence assay. The hypothesis behind this being,

each strain would most likely express some genes significantly different compared to the WT in virulence assays as a consequence of Mip gene deletions in each strain. Here, the WT was used to obtain a 'base' transcriptome level for all genes which could then be used to compare the transcriptomes of all strains, identifying any significant changes as a result of Mip gene deletions. These differences should then point to how Mips exert their virulence affects and allow us to better understand the observed phenotypes. For the RNA seq samples were sent to Novogene who employed high-throughput next generation sequencing which profiles the transcriptome through a stranded RNA library.

The WT PA01 strain and strains D, E, F, I and J generated by previous lab members were all sent for RNA seq along with versions of strains D, E and I that were created by myself. RNA seq for my mutants D, E and I will be discussed below, however RNA seq for the mutant I produced previously by Dr Silke Alt revealed essentially identical results (which can be found in the Appendix Table 1.8). A list of assessed genes and their involvement in phenotypes observed above can be found in Table 3.5. The contents of Table 3.5 displays RNA seq data obtained for my strain I, however this strain I and Silke Alt's strain I, revealed essentially identical results unless stated otherwise. The analysis of RNA seq for Silke Alt's strain I can be found in the Appendix Table 1.8 for comparisons.

Upon receipt of the sequencing data from Novogene, Dr Govind Chandra of the JIC Bioinformatics technology platform aided in the data analysis starting with the WT strain for which had extraction and sequencing had been performed twice independently (with triplicate samples in each run), once by Silke Alt and once by myself. A mean average of all three transcript readings for each gene and subjected to statistical analysis ($P<0.05$, $FDR<0.05$) to establish a base transcript level for every gene were determined in the PA01 genome. A P value of <0.05 would indicate the RNA seq result was of statistical significance, the FDR (false discovery rate) value, which was also set at a significance level of $FDR<0.05$, further ensures statistical significance by correcting for multiple testing by giving the proportion of tests above a threshold will be false positives. From this WT baseline transcriptome, it was then possible to compare the transcriptomes of all strains, gene for gene, against the WT baseline to establish which genes were up- or downregulated. The changes were determined as log fold change (LogFC) values, which is the log-ratio of a gene's or a transcript's expression values in two different conditions. For example, if the data showed a LogFC change of (+/-) 1.0, it meant there was a two-fold change in expression of a given gene. Using the LogFC data to get the fold change of a given gene was done by using its exact LogFC value, so 2 was raised to the power of the LogFC value. A logFC value of (+/-) 2.0 would mean a four-fold (2^2) increase or decrease in transcript levels when compared to the WT. Any changes in transcription of a gene were only deemed significant

if both the P value and FDR values were <0.05 and the fold change was 2, double that was observed in the WT.

Following the statistical analysis of the RNA seq data by Dr Chandra, the first set of genes to be assessed were those surrounding the PaMips. Given most genes encoding PaMips formed part of an operon, their deletion and genetic manipulation may have had unintended consequences on the transcription of surrounding genes. Such unintentional transcriptional changes might indicate that, the phenotypes observed in the various assays were due to polar effects on the expression of other genes and it would not be possible to conclude the results obtained were a direct consequence of PaMip gene deletions per se. To minimise the occurrence of such possible polar effects, 9-51bp at the beginning and end of the gene deleted were left in place, creating a truncated version of the gene rather than a start to stop codon deletion; this was designed in multiples of 3bp to ensure the reading frame remained intact.

For strain D and I, all genes surrounding the *pamip1* (*pa4558*) deletion, including *pa4557* (*lytB*) and *pa4559* (*ispA*), no statistically significant transcriptional changes were observed (P<0.05, FDR<0.05). It was particularly important that there were no changes in the *lytB* gene, whose *E. coli* homolog has been implicated in penicillin tolerance, the stringent response and cell survival, with gene deletions in *lytb* resulting in *E. coli* cell lysis [217-219]. For strains E and I, all genes surrounding the *pamip2* (*pa4572*) deletion there were no transcriptional changes. Again, importantly, *pa4571* houses a probable Cytochrome C, an important oxidase known as a key component of the respiratory chain reducing oxygen to water [220]. Any significant transcriptional changes in this gene in the mutant backgrounds could have suggested a strain's fitness was affected, in turn impacting any phenotypes observed that differed to the WT. Overall all transcript levels for genes surrounding gene deletions in strains D, E and I were unaffected, suggesting no negative, deleterious effects on transcription around the PaMip gene deletions.

In strain J, *pa4573*, a hypothetical gene next to *pamip2* (*pa4572*), but not part of the same operon, was shown to be transcribed at -0.5-fold than the level in the WT, something we would not have deemed a significant change coupled with the transcripts for all other surrounding genes were transcribed to normal WT levels. It is unlikely this transcriptional change is the result of genetic manipulation and knockout of the *pamip2* (*pa4572*) gene in J. For strain F, housing a gene deletion in *pamip3* (*pa5254*), there were no statistically significant changes in genes up or downstream of the truncated gene. For strain J where *pamip3* was truncated, out of all the surrounding genes only one gene, *pa5253* (*algP*) was found to have any transcriptional changes. This gene was found to be transcribed 1.5x more in strain J when compared to the WT. This gene has been previously annotated as a transcriptional activator of genes involved in the production of alginate, a key determinant of mucoidy, which is important for biofilm formation in *P.*

aeruginosa [221]. Given this information, one would have hypothesised an increased transcription of the *algP* gene would in turn result in increased biofilm formation. However, this was found not to be the case in the biofilm assay, with strain J unable to form a dense biofilm (Figure 3.4). Furthermore, a second gene surrounding *pamip3*, *pa5255* encoding AlgQ, a second gene transcriptional activator gene involved in the production of alginate, was shown to be transcribed at the same level as the WT, with no statistical difference. These observations, coupled with the fact that the other genes surrounding *pamip3* remaining transcriptionally unaffected, suggest that any phenotypes described above are not a consequence of the genetic manipulation of strain J and its *pamip3* gene.

Given there were no significant or unexplainable transcriptional changes in genes surrounding truncated *pamip* genes it is reasonable to assume that the phenotypes obtained in the various virulence assays were a direct consequence of the *pamip* gene deletions and not due to any unfavourable consequences resulting from their truncation. Given we could find no obvious deleterious effects on transcription due to the *pamip* gene deletions, the RNA seq data was further analysed in an attempt to elucidate any transcriptional changes within the genome that could explain the observed phenotypes. Below I will predominantly discuss the transcriptomes of strains I and J as strains these strains provided the greatest phenotypic variations in all assays.

3.7.1 Elucidating haemolytic and proteolytic phenotypes using the RNA seq

To begin, both haemolytic and proteolytic activity require the expression of haemolysins, proteases and lipases. For strains I and J, both haemolytic and proteolytic activity are either abolished and significantly reduced, respectively, one would hypothesise genes responsible for these activities would be downregulated, with their transcription levels showing a significant fold decrease when compared to WT.

The primary haemolysin in PA01 responsible for haemolytic activity is Phospholipsae C. PA01 produces two phospholipases, the first PLC-H (*plcH*) is haemolytic, while the second PLC-N (*plcN*) is non haemolytic [222]. Delving into the RNA seq data revealed that *plcH* in strain I and J is expressed to levels comparable to those of the WT suggesting in these strains it remains optimally transcribed (Table3.5) and that deletions of *pamip1*, *pamip2* or *pamip3* genes have no effect on the transcription of *plcH.* What this may suggest however, given these PaMips are FKBPs, is that they may affect the translation of the *plcH* product and / or its correct folding. If improperly folded, the activity of PLC-H would be lacking. A similar observation can be made in assessing the transcription of several proteases and lipases responsible for proteolytic activity in PA01. However, the genes encoding proteases and lipases responsible for proteolytic activity such as LipA (*pa2862*, *lipA*), a lactonizing lipase precursor, MucD (*pa0766*, *mucD*), a serine

protease and CtpA (*pa5134*, *ctpA*), a periplasmic protease that is implicated in the ability of PA01 to mount both acute and chronic infections, showed no transcriptional changes in strain I when compared to the WT [200, 201]. Interestingly, for strain J it displayed much higher transcript levels for these same genes, almost twice that of the WT strain. However, despite these genes being highly upregulated in strain J, its proteolytic activity was still significantly lower than the WT. Although the *plcH*, *lipA, mucD* and *ctpA* genes responsible for haemolytic and proteolytic activity are transcribed to levels comparable of the WT, their products still require correct folding in order to function. Moreover, along with correct folding, their translocation via PA01 secretion systems must also be expressed and assembled for them to reach their target where they can exhibit their activity. This may partly explain the loss of haemolytic and decrease in proteolytic activity in both strains.

The only significant protease gene displaying any downregulation in both strains I and J was *pa1871* (*lasA*), a protease precursor responsible for elastolytic activity, this activity aids in the breakdown of fats and proteins [223, 224]. This gene was downregulated in both strains when compared to the WT with transcript levels half of those in the WT (Table 3.5). Both strains I and J are, to some degree, able to perform proteolytic activity but are impaired when compared to the WT. This suggests this protease precursor protein is essential for complete proteolytic activity of whey protein and its active casein protein. Furthermore, it suggests that deletions in PaMip genes may be involved in its transcription on some level, and ultimately, complete proteolytic activity.

			I			J		
Gene Nº	Gene name	Description	LogFC	p Value	FDR	LogFC	p Value	FDR
Haemolytic and Proteolytic genes								
pa0844	*plcH*	Haemolytic phospholipase C precursor	0.207	0.583	0.746	0.340	0.100	0.170
pa0766	*mucD*	Serine protease MucD precursor	1.299	0.000	0.000	1.169	0.000	0.000
pa1871	*lasA*	Protease precursor	-0.977	0.022	0.079	-1.549	0.001	0.004
pa2862	*lipA*	lipA, lactonizing lipase	-0.018	0.939	0.966	-0.308	0.187	0.243
pa5134	*ctpA*	Carboxyl-terminal protease	-0.021	0.881	0.946	-0.204	0.280	0.342
GacS / RetS Two-Component system genes								
pa0527	*rsmY*	Regulatory RNA	-0.049	0.050	0.100	-0.659	0.004	0.011
pa0905	*rsmA*	Regulatory RNA	0.675	0.100	0.280	-0.677	0.100	0.177
pa0928	*gacS*	Sensor/response regulator hybrid	0.118	0.460	0.650	-0.040	0.800	0.830
pa2586	*gacA*	Response Regulator	-0.690	0.008	0.050	-0.510	0.030	0.080
pa3621.1	*rsmZ*	Regulatory RNA	0.130	0.570	0.770	-0.939	0.002	0.006
pa3974	*ladS*	Lost Adherence Sensor	-0.793	0.000	0.001	-0.422	0.011	0.023
pa4856	*retS*	Regulator of Exopolysaccharide and Type III Secretion	-0.420	0.001	0.010	-0.069	0.723	0.765
Alginate /*psl* / *pel* mucoidy synthesis genes								
pa0762	*algT*	Positive regulator alginate biosynthesis	0.949	0.003	0.029	0.105	0.695	0.772
pa0763	*mucA*	Anti-sigma factor MucA	1.179	0.001	0.013	1.224	0.001	0.003
pa0764	*mucB*	Negative regulator alginate biosynthesis	1.285	0.000	0.007	0.789	0.012	0.030
pa4033	*MucE*	Positive regulator alginate biosynthesis	-0.503	0.135	0.288	-1.221	0.001	0.005
pa3649	*mucP*	Positive regulator alginate biosynthesis	-0.002	0.993	0.997	-0.847	0.006	0.016
pa5253	*algP*	Alginate biosynthesis regulatory protein	-0.177	0.418	0.660	1.481	0.000	0.000
pa3541	*alg8*	Alginate biosynthesis	0.354	0.036	0.146	-0.169	0.286	0.397
pa3542	*alg44*	Alginate biosynthesis	-0.187	0.412	0.654	0.798	0.003	0.010
pa3543	*algK*	Alginate biosynthesis	0.205	0.625	0.809	-0.196	0.640	0.729
pa3544	*algE*	Alginate biosynthesis	-0.030	0.924	0.966	0.591	0.075	0.135
pa3545	*algG*	Alginate biosynthesis	0.218	0.133	0.342	-0.241	0.100	0.172
pa3546	*algX*	Alginate biosynthesis	0.218	0.356	0.609	0.089	0.702	0.778
pa3547	*algL*	Alginate biosynthesis	0.898	0.022	0.108	1.127	0.006	0.018
pa3548	*algI*	Alginate biosynthesis	0.161	0.282	0.531	0.002	0.991	0.994
pa3549	*algJ*	Alginate biosynthesis	-0.497	0.296	0.547	-0.114	0.806	0.857
pa3550	*algF*	Alginate biosynthesis	-0.009	0.986	0.995	0.587	0.260	0.368
pa3551	*algA*	Alginate biosynthesis	0.649	0.002	0.022	0.915	0.010	0.011
pa2232	pslB	pslB, biofilm formation protein PslB	-1.023	0.001	0.008	-0.621	0.029	0.051
pa2233	pslC	pslC, biofilm formation protein PslC	-0.964	0.000	0.004	-0.403	0.091	0.133
pa2234	pslD	pslD, biofilm formation protein PslD	-0.823	0.013	0.056	-0.495	0.116	0.162
pa2235	pslE	pslE, biofilm formation protein PslE	-1.066	0.004	0.021	-0.602	0.076	0.113
pa2236	pslF	pslF, biofilm formation protein PslF	-0.988	0.001	0.008	-0.379	0.149	0.201
pa2237	pslG	pslG, biofilm formation protein PslG	-0.815	0.003	0.018	-0.205	0.400	0.465
pa2238	pslH	pslH, biofilm formation protein PslH	-0.614	0.019	0.071	0.002	0.995	0.996
pa2239	pslI	pslI, biofilm formation protein PslI	-0.756	0.002	0.014	-0.015	0.944	0.956
pa2240	pslJ	pslJ, biofilm formation protein PslJ	-0.750	0.003	0.016	0.048	0.824	0.852
pa2241	pslK	pslK, biofilm formation protein PslK	-0.648	0.001	0.006	0.895	0.000	0.001

pa2242	pslL	pslL, hypothetical protein	-0.485	0.065	0.174	0.208	0.404	0.469
pa2243	pslM	pslM, FAD-binding dehydrogenase	0.173	0.637	0.787	2.428	0.000	0.000
pa2244	pslN	pslN, hypothetical protein	0.233	0.560	0.730	0.773	0.064	0.098
pa3058	pelG	pelG, pellicle/biofilm biosynthesis	0.571	0.004	0.034	0.950	0.000	0.000
pa3059	pelF	pelF, pellicle/biofilm biosynthesis	0.070	0.719	0.864	0.563	0.011	0.028
pa3060	pelE	pelE, pellicle/biofilm biosynthesis	0.523	0.195	0.431	0.636	0.121	0.200
pa3061	pelD	pelD, pellicle/biofilm biosynthesis	0.683	0.148	0.364	1.358	0.009	0.025
pa3062	pelC	pelC, pellicle/biofilm biosynthesis	0.162	0.925	0.966	0.348	0.839	0.882
pa3063	pelB	pelB, pellicle/biofilm biosynthesis	0.749	0.007	0.053	0.423	0.096	0.166
pa3064	pelA	pelA, hypothetical protein	0.021	0.867	0.942	-0.603	0.000	0.001
***las / rhl* Quorum sensing (QS) system genes**								
pa1432	*lasI*	Autoinducer synthesis protein	-0.615	0.073	0.180	-1.850	0.000	0.001
pa1430	*lasR*	Transcriptional regulator	-0.905	0.036	0.115	-1.320	0.004	0.010
pa3476	*rhlI*	Autoinducer synthesis protein RhlI	0.870	0.028	0.094	-1.066	0.009	0.020
pa3477	*rhlR*	Transcriptional regulator	-0.564	0.150	0.310	-0.815	0.043	0.070
Type IV Pili assocaited genes								
pa4462	*rpoN*	Sigma factor	0.070	0.690	0.850	0.480	0.018	0.042
pa4525	*pilA*	Type IV pili synthesis	0.278	0.587	0.750	-0.649	0.215	0.273
pa4526	*pilB*	Type IV pili synthesis	3.229	0.000	0.000	-0.901	0.005	0.013
pa4528	*pilD*	Type IV pili synthesis	0.771	0.000	0.001	0.384	0.016	0.032
pa4546	*pilS*	Type IV pili synthesis	0.915	0.000	0.002	1.070	0.000	0.001
pa4547	*pilR*	Type IV pili synthesis	0.417	0.027	0.092	0.362	0.051	0.081
pa4549	*fimT*	Type IV pili synthesis	1.326	0.002	0.013	1.156	0.005	0.013
pa4550	*fimU*	Type IV pili synthesis	0.059	0.872	0.934	4.912	0.000	0.000
pa4551	*pilV*	Type IV pili synthesis	-0.480	0.189	0.361	4.165	0.000	0.000
pa4552	*pilW*	Type IV pili synthesis	-0.509	0.173	0.340	3.854	0.000	0.000
pa4553	*pilX*	Type IV pili synthesis	-0.335	0.347	0.544	3.357	0.000	0.000
pa4554	*pilY*	Type IV pili synthesis	0.319	0.389	0.584	3.177	0.000	0.000
pa4555	*pilY2*	Type IV pili synthesis	-0.553	0.044	0.132	-0.256	0.325	0.391
pa4556	*pilE*	Type IV pili synthesis	-0.842	0.002	0.011	0.023	0.918	0.933
pa4959	*fimX*	Phosphodiesterase involved in twitching	0.241	0.210	0.388	-0.101	0.593	0.648
Cyclic-di-GMP synthesis genes								
pa1120	*tpbB*	diguanylate cyclase	-0.599	0.004	0.035	-1.009	0.000	0.000
pa4108		Cyclic di-gmp phosphodiesterase - degradataion of cyclic di-gm	-0.212	0.322	0.520	-0.475	0.034	0.058
pa4781		Cyclic di-gmp phosphodiesterase	-0.370	0.048	0.141	2.817	0.000	0.000
pa3702	wspR	directs formation of c-di-GMP	-0.228	0.051	0.186	-0.293	0.016	0.038
pa3703	wspF	directs formation of c-di-GMP	-0.160	0.442	0.636	-0.289	0.170	0.225
pa4332	SadC	Biofilm formation	-0.857	0.001	0.002	0.23	0.216	0.27
Flagella assembly and Rhamnolipid biosynthesis genes								
pa1100	fliE	fliE, flagellar hook-basal body complex	0.259	0.191	0.425	0.627	0.005	0.015
pa1101	fliF	fliF, flagellar MS-ring protein	0.103	0.379	0.626	1.311	0.000	0.000
pa1102	fliG	fliG, flagellar motor switch protein FliG	-0.175	0.304	0.554	2.010	0.000	0.000
pa1103	Hypo	flagellar assembly protein FliH	-0.091	0.451	0.689	-0.301	0.023	0.052
pa1104	fliL	fliI, flagellum-specific ATP synthase	0.038	0.878	0.945	-0.008	0.974	0.984
pa1105	fliJ	fliJ, flagellar biosynthesis chaperone	0.077	0.516	0.731	0.055	0.638	0.728
pa1454	fleN	flagellar synthesis regulator	-0.310	0.020	0.100	-0.770	0.021	0.039

pa3387	*rhlG*	Rhamnolipid biosynthesis	1.514	0.004	0.022	1.070	5.69x10-5	0.000
pa3476	*rhlL*	Rhamnolipid biosynthesis	0.870	0.028	0.094	0.510	0.030	0.060
pa3478	*rhlB*	Rhamnolipid biosynthesis	-0.630	0.050	0.143	-1.097	0.002	0.006
pa3479	*rhlA*	Rhamnolipid biosynthesis	-0.730	0.067	0.178	-1.175	0.006	0.014

Table 3.5. RNA seq analysis table displaying groups of genes involved in a variety of the virulence processes that have been discussed for both strains I and J. The leftmost column provides the annotated gene number as it appears on the Pseudomonas.org website for the PA01 strain, the column labelled gene name is the name given to a particular gene and its product in literature. The Log Fold Change (LogFC) column is the log-ratio of a gene's or a transcript's expression values compared to the WT strain. The LogFC value for each gene was obtained taking the average of three transcript readings for each gene in the Pa01 genome from which statistical analysis could be applied. To obtain the overall fold change in real terms compared to the WT, 2 is raised to the power of the LogFC value. The P value column indicates if the LogFC change of a given gene was of statistical significance when P<0.05. This was further statistically challenged by looking at the FDR value, also set at a significance level of FDR<0.05. Using both values ensured the LogFC value was statistically significant and not down to chance or false positives. Boxes highlighted in green call attention to a gene that has an increased LogFC change when compared to the WT and are statistically significant (P<0.05, FDR<0.05) suggesting the gene is more highly transcribed and upregulated compared to the WT. Boxes highlighted in red call attention to a gene that has a decrease LogFC change when compared to the WT and are statistically significant (P<0.05, FDR<0.05) suggesting the gene is more lowly transcribed and upregulated compared to the WT. Boxes that remain unhighlighted show genes whose LogFC change is not statistically different when compared to the WT,

3.7.2 Elucidating biofilm formation, twitching motility and type IV pili gene expression using RNA seq analysis

Interpreting the RNA seq to uncover whether there are any transcriptional changes in strains I and J that would explain their poor biofilm formation and density is complex. Given the intricacies of biofilm formation I have analysed genes involved in different aspects of the process including genes involved in; regulation of biofilm formation through regulatory systems such as the GacS/RetS two component system and the QS system; genes involved in the production of important biofilm components such as polysaccharides; genes encoding enzymes that synthesise and degrade secondary messenger molecules such as c-di-GMP which promotes the production of these polysaccharides and finally genes encoding cellular structures such as type IV pili. A list of the genes assessed and their involvement in biofilm formation can be found in Table 3.5. As mentioned, the contents of Table 3.5 displays RNA seq data obtained for the strain I made by myself, however both this strain I and the version made by Silke Alt's strain I, revealed similar results unless stated otherwise. The RNA seq data for Silke's strain I can be found in the Appendix Table 1.8. Analysing the transcripts of these genes will begin to elucidate what key biofilm components are not transcribed to levels that would be expected in the WT for strains I and J resulting in their poor biofilm formation.

The first group of genes to be examined were those involved in the GacS/RetS two-component regulatory system. As mentioned in the introduction, this complex regulatory system sits at the top of the hierarchy of systems involved in the transition from planktonic to sessile lifestyles.

Any changes in transcription of genes involved in this system might begin to explain phenotypes observed for Mip mutant strains.

In general, genes encoding components of the GacS/RetS system remained largely unaffected apart from *gacA* (*pa2586*), and *rsmY* (*pa0527*), which are both significantly downregulated in both strains I and J. Both GacA and RsmY are instrumental in the transition of planktonic cells to the sessile. Ordinarily, expression and subsequent phosphorylation of GacA induces the upregulation of genes encoding small regulatory RNAs, RsmY and RsmZ. These small RNAs proceed to bind to RsmA, a transcriptional repressor of genes involved in biofilm formation, polysaccharide production and the expression of the T6SS. Once bound to RsmA, it is released from the promoters of genes involved in biofilm formation, polysaccharide production and the expression of the T6SS and its transition to the sessile, biofilm lifestyle, begins. The biofilm assay results for both mutants I and J suggest they do not form dense biofilms and the RNA seq data sheds some light as to why this may be the case. With the downregulation of both *gacA*, and subsequently *rsmY*, RsmA will remain bound to the promoters of genes involved in biofilm formation, alginate production and other important sessile lifestyle components, thereby repressing them. This, in part, explains the data obtained for I and J which exhibit a compromised ability to form biofilms when compared to the WT.

Further evidence to support this hypothesis lies in the RNA seq for two genes involved in the biosynthesis of alginate, pel and psl. These gene products are required for the synthesis of polysaccharides which are responsible for how mucoidy PA01 is at a given time and are pivotal to biofilm formation. As more of these polysaccharides are produced, biofilm formation increases. As stated above, in mutants I and J, RsmA is likely remaining tightly bound to promoters of genes involved in biofilm formation, which includes polysaccharide biosynthesis. If this is true, one would expect a downregulation of genes involved in alginate, pel and psl biosynthesis, while also an upregulation in genes that negatively regulate their expression. Other than RsmA, there are two main repressors of polysaccharide production, MucA and MucB [225]. Overexpression of these two alginate regulators would result in the decrease of alginate production, which in turn would result in decreased biofilm formation. In both mutants I and J, *mucA (pa0763*) and *mucB* (*pa0764*) are both upregulated more than two-fold. Furthermore, genes encoding positive regulators of alginate biosynthesis such as AlgT (*pa0762*) are downregulated in strain I. In strain J genes encoding positive regulators MucP (*pa3649*) and MucE (*pa4033*) are downregulated by two-fold and five-fold respectively [226, 227]. The only positive regulator gene upregulated in strain J was *algP* as previously mentioned. Alginate biosynthesis genes *(pa3541-pa3551)* are for the most part downregulated, although they do not meet the required statistical significance values ($P<0.05$, $FDR<0.05$) and are therefore not

conclusive readouts. Genes involved in the biosynthesis of the psl polysaccharides (*pa2231-pa2243*) were on the whole down regulated in both strains I and J. Interestingly, genes involved in the biosynthesis of the pel polysaccharide (*pa3058-pa3064*) were highly upregulated in strain J, while no difference was seen in strain I when compared to the WT. It has been noted however, of these two polysaccharides psl plays a more central role in biofilm formation by aiding in the facilitation of surface attachment [76]. Thus far, a pattern is emerging that in the absence of PaMips, there are significant changes in the expression of genes involved in biofilm formation.

A second type of system involved in the transition of planktonic PA01 cells to the sessile state which interlinks with the GacS/RetS two-component system, are the N-Acyl homoserine lactone (AHLs) dependent las / rhl Quorum sensing (QS) systems. One would expect, given the repression of genes involved in the sessile, chronic lifestyle described above that our analysis would likely show a difference in transcription of genes involved in the las / rhl systems. As described in the introduction, these two systems rely on AHL signalling molecules which have been evidenced to control and regulate gene expression in a cell-density dependant manor (QS). Both the las and rhl systems respond to different AHLs, 3-oxo-C_{12}-HSL and C_4-HSL, respectively. 3-oxo-C_{12}-HSL is produced by the LasI synthase (*pa1432*) and is recognised by the transcriptional regulator LasR (*pa1430*) [228]. This system is involved in controlling the expression of multiple virulence factors involved in acute infection and promotes the planktonic lifestyle. The rhl system promotes the expression of genes associated with the transition towards the sessile lifestyle and responds to the AHL C_4-HSL which is produced by the RhlI synthase (*pa3476*) and sensed by the transcriptional regulator RhlR (*pa3477*)[229].

Given the observations thus far, including the poor biofilm formation for strains I and J, one would suspect the las system to be upregulated, promoting the acute infection lifestyle, while the rhl system would be downregulated. In the RNA seq for strain I genes encoding components of both the las and rhl systems were transcriptional unchanged and comparable to the WT. Peculiarly, in strain J the genes involved in both the las and rhl systems are downregulated promoting neither lifestyle. However, overall, RNA seq data for strain J supports the idea it is encouraged to adopt the planktonic lifestyle, not through the las system as mentioned above, but via the GacS/RetS system, a system which sits at the top of the hierarchy of systems involved in lifestyle transitions, including above the las and rhl systems. This is evidenced by the significant reduction in transcription of genes promoting the transition to the sessile lifestyle (GacA and RsmY).

The outputs of the rhl system are additionally involved in the adherence of type IV pili to lung epithelial cells and twitching motility [230]. Twitching motility is orchestrated when cell densities are high and cell-to-cell communication occurs through QS. Our *n vitro* analysis showed that

strain I lack the ability to undergo twitching motility, and it fails to produce on the cells surface (Figures 3.8 and 3.9). For this strain the rhl system and its components are expressed at levels comparable to the WT suggesting this is not the cause of these changes. In contrast, for strain J there is significant downregulation of rhl system components, which may partly explain the absence of type IV pili and twitching motility, although there are other contributing factors to type IV pili expression, such as the sigma factor RpoN, which will be discussed below.

This is only one example of a system thought to be involved in the expression of type IV pili subunits, and their assembly and activity, but on the whole, this is a complex process that is not yet fully understood, which involves over 30 genes distributed across 15 different loci. Beyond the rhl system involvement, it is known that the alternate sigma factor, RpoN (σ54), is involved in the assembly of the type IV pili and their subsequent activities of cell surface attachment and twitching motility. RpoN has been shown to regulate transcription of *pilA* and other type IV pili synthesis genes [231].

Given the rhl system is involved in the expression of type IV pili, however there are no transcriptional changes for genes in this system in strain I that would account for the lack of type IV pili observed, it was thought the RNA seq may instead uncover transcriptional changes in RpoN *(pa4462*) sigma factor, which, as mentioned above, helps promote the expression of type IV pili synthesis genes [232]. In Both strains I and J, RpoN is transcribed to levels comparable at the WT strain, thus it should not effect the transcription of the genes it activates. Furthermore, in strain I, type IV pili biosynthesis genes *pilA-pilD* (*pa4525-pa4526*), *pilS* and *pilR* (*pa4546-pa4547*), *fimT* and *fimU* (*pa4548-pa4549*), *pilV-pilYZ* (*pa4550-pa4555*) and *pilE* (*pa4556*) were all transcribed to levels comparable of the WT. However, for strain J, all these genes were upregulated but from varying degrees e.g. with a one-fold increase in *pilD* all the way to a 29-fold increase for *fimU*. Thus, the RNA seq analysis strains I and J do not provide any explanation as to the lack of type IV pili structures produced. In fact, based on the RNA seq analysis of sets of associated with type IV pili biosynthesis and twitching motility it might be expected that type IV pili expression and twitching motility would be increased. These data do, however, point to a possible issue at the translational or post-translational level for the production and/or maturation of proteins involved in the type IV pili synthesis and assembly, and in the twitching process.

The final contributor involved in the lifestyle transitions of PA01 is cyclic-di-GMP. As described in **1.5.3**, cyclic-di-GMP is an important secondary messenger in bacteria that is involved in signal transduction pathways and controlling many cellular process, such as biofilm formation, at the transcriptional, translational and post-translational level [233]. It is synthesised by enzymes with diguanylate cyclase activity and degraded by enzymes with phosphodiesterase activity. In PA01,

when cyclic-di-GMP levels increase, along with the GacS / RetS and las /rhl systems mentioned above becoming active, the transition from planktonic / acute lifestyles to the more sessile /chronic lifestyle, begins. Conversely, the planktonic / acute lifestyles are more associated with low levels of cyclic-di-GMP. When cyclic-di-GMP levels are high one would expect high levels of diguanylate cyclases synthesising the cyclic-di-GMP, thus encouraging biofilm formation. In contrast, when cyclic-di-GMP levels are low, higher levels of the cyclic-di-GMP degrading enzymes, phosphodiesterases, would be expected along with the planktonic lifestyle [77].

Given that for both strain I and J biofilm formation is low compared to the WT, low levels of diguanylate cyclase genes (*pa1120*, *pa3702* and *pa703*) and high levels of phosphodiesterases (*pa4108* and *pa4781*) would be anticipated. In both strains I and J the genes expressing the diguanylate cyclases TpbB (*pa1120*) and WspR (*pa3702*), both of which help catalyse cyclic-di-GMP formation, are both significantly downregulated in both strains as predicted. In contrast, the gene encoding a phosphodiesterase *pa4781* is upregulated approximately 8-fold in strain J when compared to the WT, suggesting cyclic-di-GMP should be readily degraded therefore present a low level in the cell, leading to suppression of biofilm formation which would be consistent with the phenotype observed in strain J. This phosphodiesterase (*pa4781*) is downregulated in strain I by one-fold however not to any statistical significance. The downregulation of diguanylate cyclases in strains I and J, while an upregulation in phosphodiesterases in strain J, in part, elucidate why biofilm formation in these strains is poor.

The transcript levels of these sets of genes would suggest the synthesis of cyclic-di-GMP to be low, while enzymes responsible for its degradation to be high, preventing both from transitioning into a biofilm, something which cyclic-di-GMP is known to help facilitate. It is unlikely PaMips are transcription factors directly affecting the transcription of these cyclic-di-GMP synthesis / degradation encoding genes, however, are more likely acting as chaperones of transcription factors that promote expression of genes for enzymes such as diguanylate cyclase. In the strains I and J, deletion of the PaMips may potentially result in the loss of correct folding of these transcriptional regulators, which would ordinarily promote the synthesis of cyclic-di-GMP, without which biofilm formation is suppressed.

3.7.3 Examining swarming motility and flagella related gene expression using RNA seq analysis

As described above unusual swarming phenotypes observed for strains I and J and the use of SEM highlight irregularities in flagella expression. To understand what might be driving these effects we assessed the transcription levels of genes involved in flagella biosynthesis and rhamnolipid production for strains to see if they might explain the phenotypes observed. Swarming is the result of flagella dependent movement and is aided by rhamnolipid production

and it is usually associated with the planktonic / acute lifestyle of PA01 [207]. Given the in assays assessing virulence phenotypes associated with the stationary / chronic lifestyle are attenuated, such as twitching motility and biofilm formation, the bizarre highly motile phenotypes observed in the swarming associated with the planktonic / acute lifestyle, is logical.

This was the case for strains I and J who both performed erratic swarming motility reaching further distances towards the edge of the swarming plates than the WT. This phenotype can be partly attributed to the irregularities in flagella expressed in strain I which lacks the twinning flagella, while strain J produces excess flagella (Figures 3.8 and 3.9). To help understand the peculiar flagella expression we looked at genes involved in flagella biosynthesis. There are many genes involved in flagella biosynthesis and expression, however I have focussed on the following; *pa1100-pa1105* with each gene encoding a different flagella subunit; and *pa1454* that encodes FleN, a protein that helps regulate the number of flagella present on the surface of PA01 cells.

In strain I there were no differences in transcription levels for the flagella biosynthesis genes or the gene encoding their regulator, FleN, suggesting the lack of twinning flagella is the result of translational or exportation issues as a consequence of PaMip gene deletions. Furthermore, the larger swarming circumference that is not quite as circular in pattern when compared to the WT may be the result of only a singular flagellum being expressed [234]. Strain J exhibited large transcriptional increases in the flagella biosynthesis genes *pa1100-pa1102*, with no changes in *pa1103-pa1105*. Furthermore, the gene encoding the transcriptional regulator that mediates the expression of these genes, FleN, was significantly downregulated which partly explains the number of flagella expressed and obscure swarming motility observed in strain J. In literature a *fleN* gene deletion resulted in the expression of multiple flagella. They found complementing the *fleN* gene to be difficult in that flagella assembly was highly sensitive to the amount of FleN expressed, and that anything other than WT transcription levels would express no flagella, or multiple [234]. As there is a large two-fold decrease in the transcription of the *fleN* in strain J (almost mimicking the *fleN* gene deletion strain), the SEM images and obscure swarming motility are explainable given the regulator of flagella biosynthesis genes *pa1100-pa1102* is significantly downregulated which resulted in their unchecked overexpression.

As mentioned, rhamnolipid production aids in swarming motility allowing the bacterium to glide across a surface so several genes involved in rhamnolipid biosynthesis were assessed; *pa3387* and; *pa3476-pa3479*. Rhamnolipid biosynthesis genes *pa3387* (RhlG) and *pa3476* (RhlL) were found to be upregulated in strain I by two-fold, while *pa3478* (RhlB) was downregulated by one-fold when compared to the WT. However, these transcriptional differences do not elucidate any further the peculiar swarming phenotype, especially when factoring in results in Figure 3.11

which revealed although key rhamnolipids MonoC10-C10 and DiC10-C10 were produced more highly in the WT adjacent to the initial inoculum compared to strains I and J, their spread and dispersal away from the initial inoculum was no different when compared to the WT therefore unlikely to affect swarming motility. As with strain I, both rhamnolipid genes *pa3387* (RhlG) and *pa3476* (RhlL) were found to be more highly transcribed than the WT, while both *pa3478* and *pa3479* were significantly downregulated. However, like with strain I these observations genes do not elucidate any further the peculiar swarming phenotype, especially when factoring in results in Figure 3.11. These results do however begin to explain the abnormal number of flagella expressed which in itself would explain the swarming phenotype.

3.7.4 Other notable transcriptional changes identified through analysis of the RNA seq data

Having attempted to better understand the phenotypes observed for strains I and J in plate-based assays assessing transcriptional changes in genes thought to be involved in various virulence mechanisms, I next looked in an unbiased manner for genes exhibiting the highest changes in relative transcription levels across the entire genome. The RNA seq data was analysed and arrange in order to list the transcriptional fold change of each gene from highest positive fold changes to highest negative fold changes, with genes exhibiting smaller fold changes in the middle of the list (e.g. highest positive fold change down to lowest negative fold change). This was done for both I strains and J strains which allowed comparisons between the two sets of RNA data for each RNA seq experiment, once performed by myself and once by Silke. Furthermore, I mainly focussed on sets or clusters of genes all exhibiting similar fold changes, rather than individual genes as transcriptional changes in a single gene would unlikely reveal any global context as to the loss of virulence in the PaMip knockout strains.

The most striking observation was for a group of genes involved in the Type Six Secretion System Island III (T6SS HSI-III). These genes displayed between a 4-to-8-fold decrease in transcription when compared to the WT ($P<0.05$, $FDR<0.05$). This cluster was also significantly downregulated in strain J, but not to the extent seen in strain I (between 2-to-4-fold decrease). In PA01 there are three clusters of genes that encode a T6SS. These three T6SS clusters are termed ‘islands’ and annotated as follows; HSI-I (Hcp secretion island I), HSI-II and HSI-III, and each is composed of two operons termed ‘left’ and ‘right’ (Figure 3.21) [235]. To date seven types of secretion systems (I-VII) have been described, with each secretion system type shown to differ in their components and in their molecular mechanisms. Type I, type II, IV and type V usually deliver proteins onto the bacterial cell surface or into the surrounding environment [236], whereas Type III and VI Secretion Systems usually deliver proteins or effectors directly into the periplasm of a target cell.

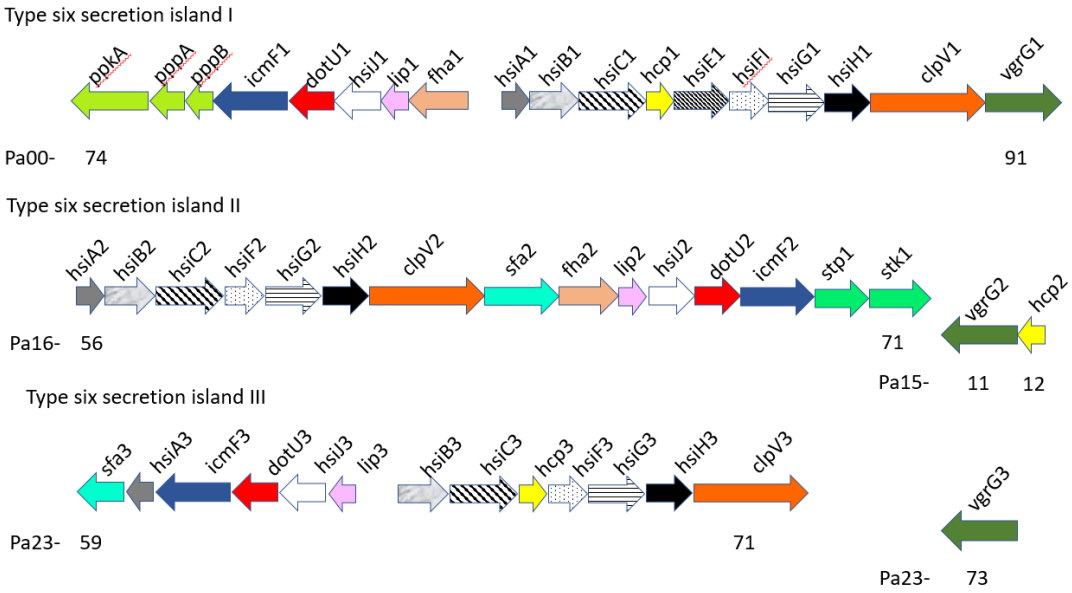


Figure 3.21. Diagram of the three T6SS gene clusters or 'islands' in the *P. aeruginosa* genome. Homologous genes are marked in the same colour. Abbreviations for each gene are as follows; *ppk*, *Pseudomonas* protein kinase; *ppp*, *Pseudomonas* protein phosphatase; *icm*, intracellular multiplication; *dot*, defect in organelle trafficking; *hsi*, hcp secretion island; *lip*, lipoprotein; *fha*, forkhead-associated protein; *hcp*, haemolysin coregulated protein; *sfa*, sigma activator factor; *stk*, serine/threonine kinase; *stp*, serine/threonine phosphatase; *vgrG*, valine-glycine repeats protein.

For strain I (both versions) there was a significant down regulation of the entire cluster of genes encoding the T6SS HSI-III, but no statistical differences in the transcription of the entire gene clusters encoding the T6SS HSI-I and T6SS HSI-II (Table 3.6). In Silke's strain I and J there are transcriptional increases in some genes in the T6SS HSI-I and T6SS HSI-II islands however did not meet the LogFC threshold. Each system has been described with a different role in the literature. The primary and most well studied T6SS in *P. aeruginosa* is HSI-I, this island is responsible for the export of three bacteriolytic toxins Tse1-3, whose genes are known as accessory genes, and are scattered throughout the chromosome [237]. This system is known to be primarily active in outcompeting bacteria in the environment by delivering these bacteriolytic toxins directly into the periplasm of a target bacterial cell [238, 239]. Less is known about the function and control of the two other islands HSI-II and HSI-III. The few studies carried out on HSI-II suggest it is predominantly involved in delivering toxins to competing bacterial cells however has also been shown to be involved in the increased uptake of PA01 into HeLa cells, with mutants displaying a decrease in colonisation of HeLa and lung epithelial cells by 75% [240]. Furthermore, in an *in vivo Caenorhabditis elegans* worm infection model HSI-II was shown to play a role in virulence, with larvae infected with a ClpV2 knockout mutant showing a higher survival rate than for the WT [240]. Subsequent work showed that the phospholipase PldA is an effector specific to the HSI-II island. PldA was shown to degrade phosphatidylethanolamine, a major structural

component of bacterial membranes, thus suggesting the HSI-II, like HSI-I, is associated with virulence against bacteria as well as interacting with eukaryotic cells [241].

The HSI-III island has been poorly studied with most research focussing on the HSI-III in PA014 which appears genetically different to that in PA01. As the other two T6SS islands have been predominantly annotated as delivering toxins to other prokaryotic cells, although recently the HSI-II has shown to potentially interact with eukaryotic cells, it has been assumed the same will be true for the HSI-III. Only two studies have touched upon HSI-III activity and the role it plays in the virulence of PA01. These studies focussed on an acute lung infection model [242] and a *C. elegans* survival experiment [243]. In the study examining a Δ*clpV2* mutant described above, a *C. elegans* survival experiment was carried out and ClpV2 was shown to be required for full virulence, in this same study a Δ*clpV3* mutant also showed a two-day delay in the killing of *C. elegans* compared to the WT suggesting this HSI-III island is important for *in vivo* virulence of eukaryotes [243]. This study also noted that there was no compensation between the two clusters and that the Δ*clpV3* mutant strain displayed a more dominant phenotype compared to the Δ*clpV2* mutant.

			I			J		
Gene N°	Gene name	Description	LogFC	p Value	FDR	LogFC	p Value	FDR
Type Six Secretion System HSI-III genes								
pa2359	*sfa3*	Probable transcriptional regulator	-0.076	0.716	0.863	1.612	0.001	0.005
pa2360	*hsiA3*	T6SS HSI-III hypothetical protein	-1.901	0.000	0.000	-0.481	0.234	0.294
pa2361	*Icmf3*	T6SS HSI-III hypothetical protein	-1.992	0.000	0.000	0.371	0.408	0.473
pa2362	*dotU3*	T6SS HSI-III hypothetical protein	-2.397	0.000	0.000	0.226	0.620	0.673
pa2363	*hsiJ3*	T6SS HSI-III hypothetical protein	-1.678	0.000	0.000	-0.751	0.048	0.077
pa2364	*lip3*	T6SS HSI-III hypothetical protein	-1.488	0.000	0.000	-0.989	0.006	0.015
pa2365	*hsiB3*	T6SS HSI-III hypothetical protein	-2.358	0.000	0.000	-1.726	0.003	0.009
pa2366	*hsiC3*	T6SS HSI-III hypothetical protein	-2.435	0.000	0.000	-1.642	0.003	0.009
pa2367	*hcp3*	Secreted protein Hcp	-2.414	0.000	0.000	-1.629	0.002	0.006
pa2368	*hsiF3*	T6SS HSI-III hypothetical protein	-3.009	0.000	0.001	-1.104	0.038	0.064
pa2369	*hsiG3*	T6SS HSI-III hypothetical protein	-2.570	0.000	0.000	-0.449	0.374	0.439
pa2370	*hsiH3*	T6SS HSI-III hypothetical protein	-3.265	0.000	0.000	-0.332	0.489	0.550
pa2371	*ClpV3*	Probable ClpA/B-type protease	-2.396	0.000	0.000	-1.279	0.005	0.012
pa2373	*VgrG3*	T6SS HSI-III hypothetical protein	-1.561	0.000	0.000	-0.826	0.011	0.024
pa2374	*tseF*	T6SS HSI-III hypothetical protein	-1.405	0.000	0.005	-0.343	0.154	0.208
Type Six Secretion System HSI-II genes								
pa1511	*VgrG2*	T6SS HSI-II hypothetical protein	-0.261	0.343	0.540	-0.173	0.203	0.303
pa1512	*HCP2*	Secreted protein Hcp	-0.185	0.545	0.716	-0.058	0.802	0.854
pa1656	*hsiA2*	T6SS HSI-II hypothetical protein	0.134	0.525	0.701	-0.128	0.512	0.618
pa1657	*hsiB2*	T6SS HSI-II hypothetical protein	-0.305	0.271	0.464	0.975	0.000	0.000
pa1658	*hsiC2*	T6SS HSI-II hypothetical protein	-0.962	0.002	0.016	-0.654	0.000	0.001
pa1659	*hsif3*	T6SS HSI-II hypothetical protein	-0.862	0.056	0.158	-1.499	0.000	0.000
pa1660	*hsiG2*	T6SS HSI-II hypothetical protein	-0.618	0.129	0.280	-0.709	0.001	0.003
pa1661	*hsiH2*	T6SS HSI-II hypothetical protein	-0.559	0.174	0.341	-0.284	0.091	0.159
pa1662	*ClpV2*	Probable ClpA/B-type protease	-0.801	0.029	0.099	-0.401	0.013	0.033
pa1663	*sfa2*	Probable transcriptional regulator	-0.587	0.273	0.466	-0.831	0.002	0.008
pa1664	*fha2*	T6SS HSI-II hypothetical protein	-0.943	0.283	0.477	-2.427	0.011	0.027
pa1665	*lip2*	T6SS HSI-II hypothetical protein	-0.669	0.253	0.444	-0.546	0.102	0.174
pa1666	*hsij2*	T6SS HSI-II hypothetical protein	-0.659	0.240	0.428	-0.030	0.892	0.923
pa1667	*dotu3*	T6SS HSI-II hypothetical protein	-0.549	0.014	0.058	-0.116	0.324	0.435
pa1668	*icmF2*	T6SS HSI-II hypothetical protein	-0.650	0.116	0.261	-0.892	0.001	0.005
pa1669	*stp1*	T6SS HSI-II hypothetical protein	-0.402	0.181	0.350	0.255	0.162	0.253
pa1670	*stk1*	T6SS HSI-II hypothetical protein	-0.210	0.491	0.675	-0.196	0.765	0.828
Type Six Secretion System HSI-I genes								
pa0074	*ppkA*	Serine threonine protein kinase PpkA	-0.225	0.293	0.543	0.317	0.145	0.233
pa0075	*pppA*	serine/threonine phosphatase	-0.645	0.128	0.334	-1.275	0.007	0.019
pa0076	*pppB*	T6SS HSI-I hypothetical protein	-1.056	0.114	0.312	1.579	0.026	0.058
pa0077	*icmF1*	Type VI secretion protein IcmF	-0.845	0.000	0.005	0.703	0.001	0.004
pa0078	*dotU1*	T6SS HSI-I hypothetical protein	-0.016	0.959	0.981	0.244	0.432	0.541
pa0079	*hsiJ1*	T6SS HSI-I hypothetical protein	-0.185	0.389	0.634	-0.873	0.001	0.004
pa0080	*lip1*	T6SS HSI-I hypothetical protein	-0.093	0.696	0.852	-0.935	0.001	0.005
pa0081	*fha1*	Fha domain-containing protein	0.198	0.435	0.675	-1.113	0.001	0.003

pa0082	*hsiA1*	T6SS HSI-I hypothetical protein	0.080	0.647	0.822	-0.645	0.002	0.007
pa0083	*hsiB1*	T6SS HSI-I hypothetical protein	-0.416	0.044	0.169	0.179	0.355	0.468
pa0084	*hsiC1*	T6SS HSI-I hypothetical protein	-0.149	0.388	0.633	-0.457	0.016	0.039
pa0085	*hcp1*	Secreted protein Hcp	-0.242	0.209	0.448	-0.985	0.000	0.001
pa0086	*hsiE1*	T6SS HSI-I hypothetical protein	0.052	0.833	0.924	-0.017	0.945	0.962
pa0087	*hsif1*	T6SS HSI-I hypothetical protein	-0.198	0.665	0.832	-0.117	0.797	0.850
pa0088	*hsiG1*	T6SS HSI-I hypothetical protein	0.291	0.184	0.416	-0.311	0.158	0.247
pa0089	*hsiH1*	T6SS HSI-I hypothetical protein	0.456	0.243	0.488	0.008	0.983	0.989
pa0090	*clpv1*	Probable ClpA/B-type protease	0.018	0.905	0.957	-0.130	0.395	0.506
pa0091	*Vgrg1*	T6SS HSI-I hypothetical protein	0.648	0.001	0.017	0.016	0.923	0.946
Type Three Secretion System genes								
pa1697	*pscN*	Type III SS related	1.892	0.008	0.037	3.710	0.000	0.001
pa1706	*ppcrV*	Type III SS related	2.269	0.000	0.000	1.915	0.000	0.000
pa1708	*popB*	Type III SS related	1.140	0.000	0.000	1.434	0.000	0.000
pa1709	*popD*	Type III SS related	1.221	0.000	0.000	1.480	0.000	0.000
pa1710	*exsC*	Type III SS related	1.049	0.000	0.002	0.298	0.192	0.248
pa1711	*exsE*	Type III SS related	1.004	0.000	0.000	0.327	0.075	0.111
pa1712	*exsB*	Type III SS related	1.217	0.000	0.000	0.663	0.002	0.006
pa1713	*exsA*	Type III SS related	2.167	0.000	0.000	0.471	0.182	0.238
pa1714	*exsD*	Type III SS related	1.706	0.000	0.000	0.954	0.000	0.001
pa1716	*pscC*	Type III SS related	1.082	0.008	0.038	1.612	0.000	0.002
pa1719	*pscF*	Type III SS related	0.943	0.060	0.165	1.849	0.001	0.004
pa1723	*PscJ*	Type III SS related	1.412	0.003	0.018	1.725	0.001	0.003
pa1725	*pscL*	Type III SS related	-0.142	0.812	0.901	2.277	0.001	0.004
pf1 bacteriophage genes								
pa0715		Hypothetical protein	-10.554	0.000	0.000	-0.797	0.035	0.073
pa0716		Hypothetical protein	-10.531	0.000	0.000	0.605	0.104	0.176
pa0718		Hypothetical protein of bacteriophage Pf1	11.090	0.000	0.000	7.839	0.000	0.000
pa0719		Hypothetical protein of bacteriophage Pf1	10.816	0.000	0.000	8.063	0.000	0.000
pa0720		Helix destabilizing protein of bacteriophage Pf1	11.375	0.000	0.000	6.987	0.000	0.000
pa0721		Hypothetical protein of bacteriophage Pf1	10.347	0.000	0.000	6.761	0.000	0.000
pa0722		Hypothetical protein of bacteriophage Pf1	10.953	0.000	0.000	4.595	0.000	0.000
pa0723	*coaB*	Coat protein B of bacteriophage Pf1	11.142	0.000	0.000	4.411	0.000	0.000
pa0724		Phage coat protein A	7.196	0.000	0.000	5.159	0.000	0.000
pa0725		Hypothetical protein of bacteriophage Pf1	6.410	0.000	0.000	5.774	0.000	0.000
pa0726		Hypothetical protein of bacteriophage Pf1	6.571	0.000	0.000	6.328	0.000	0.000
pa0727		Pf replication initiator protein	7.469	0.000	0.000	5.535	0.000	0.000
pa0728		Probable bacteriophage integrase	8.569	0.000	0.000	7.084	0.000	0.000
pa0729		PfiT antitoxin, PfiA	4.140	0.000	0.000	4.599	0.000	0.000

Table 3.6. RNA seq analysis table displaying groups of genes exhibiting the largest fold changes for both strains I and J. The leftmost column provides the annotated gene number as it appears on the Pseudomonas.org website for the PA01 strain, the column labelled gene name is the name given to a

particular gene and its product in literature. The Log Fold Change (LogFC) column is the log-ratio of a gene's or a transcript's expression values compared to the WT strain. The LogFC value for each gene was obtained taking the average of three transcript readings for each gene in the Pa01 genome from which statistical anaylsis could be applied. To get the overall fold change in real terms compared to the WT, 2 is raised to the power of the LogFC value. The P value column indicated if the LogFC change of a given gene was of statistical significance if P<0.05. This was further statistically challenged by looking at the FDR value, also set at a significance level of FDR<0.05. Using both values ensured the LogFC value was statistically significant and not down to chance or false positives. Boxes highlighted in green call attention to a gene that has an increased LogFC change when compared to the WT and are statistically significant (P<0.05, FDR<0.05) suggesting the gene is more highly transcribed and upregulated compared to the WT. Boxes highlighted in red call attention to a gene that has a decrease LogFC change when compared to the WT and are statistically significant (P<0.05, FDR<0.05) suggesting the gene is more lowly transcribed and upregulated compared to the WT. Boxes that remain unhighlighted show genes whose LogFC change is not statistically significant when compared to the WT, therefore one could assume transcript levels of a given gene between the WT and a given strain are no different, neither up or down regulated.

Given the significant down-regulation of the T6SS HSI-III in both strains I and J, and the various phenotypes observed for both strains in the virulence assays described in this thesis, it is possible that the HSI-III island may be significant for interactions with eukaryotic cells. Ordinarily, the phenotypes described here have been associated with the action of the T3SS, however the genes involved in T3SS assembly are all highly upregulated in both strains I and J, this suggests the T3SS is likely active in strains I and J unless the absence of PaMips1-3 has led to their incorrect assembly and protein folding (Table 3.6). This hypothesis will be explored in more detail in Chapter 5. A second region of genes upregulated in both sets of data for strain I are *pa2134-pa2160*, which are most likely associated with the stationary growth phase of PA01 [244].

The most highly upregulated cluster of genes in strain J comprise *pa0718-pa0729* (between 16-to-256-fold increases) belong to genes encoding the filamentous bacteriophage pf1 (Table 3.6). Genes *pa0715* and *pa0716* preceding this region, termed the pre-prophage region, are downregulated in both sets of strain J RNA seq data. There are two types of bacteriophages, lysogenic phages (temperate phage / prophage) a type of phage that infects PA01 and integrates into the circular bacterial chromosome, and, lytic (virulent) bacteriophages that use the host cell machinery to replicate and then destroy and lyse the host cell [245]. The pf1 phage genes that are upregulated in strain J encode a class II filamentous prophage that is incorporated into the PA01 genome and has been shown to promote a more non-invasive lifestyle of PA01, this influences the host inflammatory response whilst promoting resistance to phagocytosis by macrophages [246]. Furthermore, the Pf1 phage has been shown to promote bacterial adhesion to mucin, trapping PA01 in the lung, a phenotype typically associated with chronic infections. Interestingly, although this set of genes are highly upregulated in both sets of strain J RNA seq data, the observed phenotypes suggest the opposite to that expected for an upregulated prophage region.

Curiously, although the results obtained by both Silke and myself for this prophage region in strain J are identical, the results for my strain I to Silke's in this same region were strikingly different. Silke's strain I exhibited a slight upregulation, albeit not statistically different to the WT, in genes *pa0715* and *pa0716* preceding the prophage region while also a significant downregulation in genes *pa0718-pa0729* (Appendix Table 1.9). Conversely, data for my strain I was more similar to that observed in both sets of RNA seq data for strain J, with a significant downregulation in *pa0715* and *pa0716*, as well as a significant regulation in *pa0718-pa0729* genes (Table 3.6). The significance of this is at this stage not known.

This prophage encoding region will be discussed again due to the whole-genome sequencing data obtained for Silke's strain I which is described in Chapter 4. Chapter 4 explains why it was necessary to recreate strains D, E and I as it was found Silke's strain I, over time, had lost a 3275 bp region encompassing the pre-prophage region genes *pa0715* and *pa0716*. Due to the loss of this 3275bp region it was not possible to conclude that the results obtained in assays were a direct consequence of the PaMip gene deletions and not due to the loss of this region. Hence, I recreated strain I and have been presenting RNA seq data obtained for the version of strain I in this chapter (with reference to Silke's strain I where relevant). Furthermore, Silke's strain I was sent for RNA seq immediately upon its creation suggesting the presence of the 3275bp region at one time and why in the RNA seq this region appears and is not statistically different to the WT. Silke's strain I was only sent for whole-genome sequencing three years following its creation suggesting this 3275bp region was lost overtime. Upon the creation of my strain I, it was immediately sent for whole genome sequencing to ensure no genomic changes had occurred (analysis of this data demonstrated that the 3275bp region was present). Once confirmed, assays were performed with the newly created version, which, gratifyingly, exhibited identical phenotypes to those observed with Silke's strain I. However, again, over time, this 3275bp region appeared to be lost in my strain I following continual monitoring via colony PCR. This loss over time is then reflected in the RNA seq for my strain I for genes *pa0715* and *pa0716*, which showed a 1024-fold decrease in transcription, essentially being non-existent. Again, this will be discussed more detail in Chapter 4.

3.8 Discussion

Based on the combined results presented above, I believe it is appropriate to claim that PA01, like other pathogenic gram-negative bacteria previously discussed, does possess virulence associated Mip proteins. This is the also first instance where three Mip proteins have been discovered in a genome and confirmed to have a virulence related role. Furthermore, it is evident that these PaMips are involved in virulence in a pleiotropic manor, somehow affecting several different virulence processes. Moreover, these PaMips appear to be important for processes associated with the planktonic lifestyle and chronic infections, something which would impact CF patients. As previously discovered Mips have been shown to be inhibitable by FK506 and rapamycin, the development of molecules to inhibit PaMips, but which lack the immunosuppressive effect of these molecules, is an attractive future area of research. Such compounds could be utilised as anti-virulence drugs, reducing the pathogenicity of an infection and allowing the appropriate antibiotic to be administered following the identification of the pathogen and its antibiotic sensitivity testing. This would combat the increase of drug resistant bacteria on two fronts. Firstly, by targeting the virulence mechanisms required for the development of the infection rather than the growth and fitness of the pathogen, it would be less likely to select for resistant organisms than current antibiotic strategies [247]. Secondly, it would provide clinicians with time while identifying the pathogen, something that is often limited when deciding the appropriate treatment strategy and would allow correct microbial procedure in identification of organisms, and therefore, a more targeted antibiotic approach instead of the common current treatment strategy of 'best guess' prophylaxis while awaiting the return of microbial sensitivity testing [155].

Although here I have demonstrated PA01 does have Mips involved in virulence, the specific mechanism by which they mediate virulence remains unknown. Unlike previous studies which generally showed that a given organism possess a Mip, and denoting that it is required for macrophage colonisation, here I have begun to explore the molecular nature of how Mips may be exerting their effect by utilising a wide range of phenotypic assays. The accompanying RNA seq analyses indicates the potential involvement of PaMips in regulatory systems responsible for the transition of lifestyles, and with these lifestyle transitions, the switching on and off of different virulence mechanisms. These regulation systems are incredibly complex and still poorly understood which is of little surprise given 20% of the PA01 genome is dedicated to genes predicted to have some involvement in regulation [248]. Based on the phenotypic observations made, coupled with RNA seq data, many genes involved in the sessile biofilm lifestyle in the PaMip mutant strains are negatively expressed when compared to the WT, suggesting that for

the PaMip mutant strains they preferentially adopt the planktonic lifestyle, or are just not able to transition to the sessile lifestyle.

The data described in the RNA seq analysis revealed that several virulence associated gene clusters are significantly up or down regulated in the PaMip knockout strains I and J, and one may have assumed they are directly involved in the transcription of these genes. However, after using several protein prediction tools, it is evident that none of PaMip1, PaMip2 or PaMip3 possess any obvious DNA binding regions and are unlikely to act as transcriptional factors binding directly to promoters of genes causing their transcription (Appendix Image 1.3).

Based on their amino acid sequences and several protein alignments, it can be predicted that PaMip2 and Pamip3 will exhibit PPIase activity as they both possess the key aspartate and tyrosine residues required for PPIase activity that was discussed in the introduction (Appendix Image 1.4). On this basis it is possible they are involved in the correct translation of proteins and in folding and other posttranslational processes. PaMip2 (which is missing in both strains I and J) is predicted to be localised to the cytoplasm and thus might be responsible for such activity. PaMip1 is predicted to localise to the cytoplasm also, however, is also predicted to lack PPIase activity as it is absent in the key aspartate and tyrosine residues required, and as such is unlikely involved in the folding of virulence associated proteins. However, to verify these predictions a PPIase activity assay to measure the ability to perform *cis-trans* isomerisation of proline-peptidyl bonds would need to be conducted. Although PaMip3 is predicted to possess PPIase activity, it is predicted to localise to the cell surface like the LgMip and is unlikely involved in the processes described above. Instead, given the data obtained in the macrophage infection assay it is most likely to be a surface antigen that interacts with a host, much like the LgMip which was shown to be a moonlighting protein and facilitated interaction with a host cell via collagen IV.

PPIases have already been evidenced as molecular chaperones with a plethora of roles in protein folding and regulation [247]. In the PaMip knockout backgrounds, if a transcription factor protein for example, is not correctly folded, the promotion / repression of its target gene(s) will be affected. Furthermore, transcriptional factors ensure the timing of gene expression at a given point in the cell cycle or in response to environmental conditions [249]. An example of this drawn from the RNA seq data is for genes involved in the biosynthesis of alginate and psl. Although the genes responsible for their biosynthesis are heavily downregulated, the gene encoding one of their positive transcriptional regulators, AlgT, is still transcribed at levels comparable of the WT [250, 251]. This suggests there may be an issue with the transcriptional factor AlgT itself, possibly because of the PaMip gene deletions. For example, AlgT may not be correctly folded, and therefore will not bind to the promoters of the alginate and psl biosynthesis genes resulting in their observed down regulation (Figure 3.23).

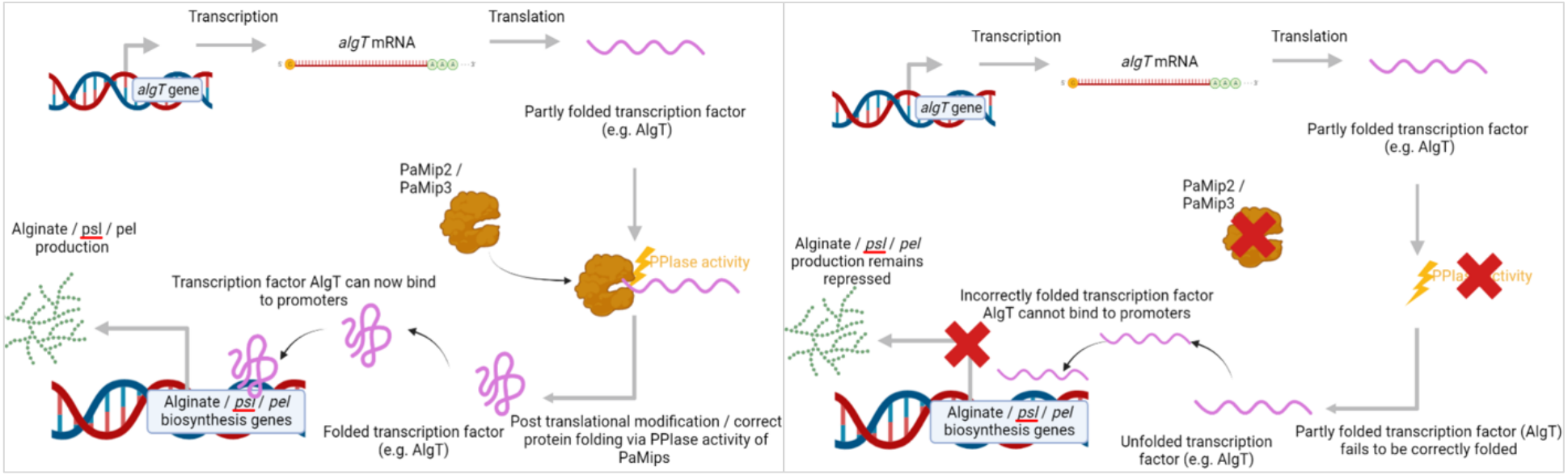


Figure 3.22. Diagram illustrating the proposed mechanism by which PaMips are involved in virulence. In the left image, transcription of *algT* transcription factor, is transcribed in strains I and J to levels comparable of the WT strain (Table 3.5). The resulting mRNA is translated however the resulting protein requires further post translational modification to become a functional AlgT transcription factor protein. In this case the post translational modification comes in the form of PPIase activity from a PaMip. This modification allows AlgT to become fully functional and bind directly to DNA and promote the transcription of genes involved in alginate/ psl / pel biosynthesis. The image on the right describes the same process however in strain I or J which lack PaMips. Without the PaMips the partly folded AlgT will not undergo further post translational modification. Without correct folding, AlgT cannot fulfil its role as a transcription factor and will no longer bind DNA and therefore cannot promote the expression of alginate / psl / pel biosynthesis genes. Without the expression of these genes biofilm formation is decreased, a phenotype observed in both strains I and J.

Alternatively, the PaMips may be involved in the correct folding of proteins directly involved in virulence mechanisms. For example, genes responsible for structures such as the type IV and flagella are normally transcribed in strains I and J based on the analysis of the RNA seq data. However, such structures are either physically missing or their production has been impaired suggesting a translational issue. Without the type IV pili structures, twitching motility is lost, along with further consequences such as impaired biofilm formation due to the inability of cells to perform type IV pili mediated attach to a solid surface. Similarly, the gene responsible for haemolysis, *plcH,* remains transcribed to WT levels, however haemolytic activity is lost in strains I and J.

One example of a PPIase in nature that behaves in a similar fashion is the parvulin Pin1. Pin1 is a PPIase that has a specific set of client proteins and plays an important role in post phosphorylation control in regulating protein function by catalysing the *cis-trans* isomerisation of proline-peptidyl bonds resulting in an alteration of a client protein's structure and function [252]. Pin1 has been shown to exert a pleotropic role in human physiology including; regulating protein functions such as catalytic activity, the phosphorylation status of a protein, protein interactions and subcellular localisations. If Pin1 and its PPIase activity is lacking, the correct regulation and folding of proteins involved in cancer initiation and progression are affected, leading to Pin1 being annotated as a proto-oncogene where mutations in the gene can lead to cancer. Furthermore, Pin1 is indirectly involved in transcription as it isomerizes two transcription factors, NF-KB and STAT3, this action of Pin1 on these two transcription factors results in the transcription of genes encoding cyclin D1.

Relating this back to PaMips, they may potentially have their own specific set of virulence-related client proteins such as type IV pili biosynthesis proteins or PLC-H. Due to the absence of PPIase activity, type IV pili biosynthesis proteins containing proline-peptidyl bonds that require *cis-trans* isomerisation may not be correctly folded resulting in the failure to assemble type IV pili structures. This would lead to a loss of twitching motility and other related phenotypes (Figure 3.24). In parallel, PaMips also act like Pin1 with regards to their indirect involvement in transcription by the isomerization of transcription factors such AlgT. These proposed different mechanisms would begin to elucidate as to how PaMips exert such a pleiotropic effect on PA01 virulence.

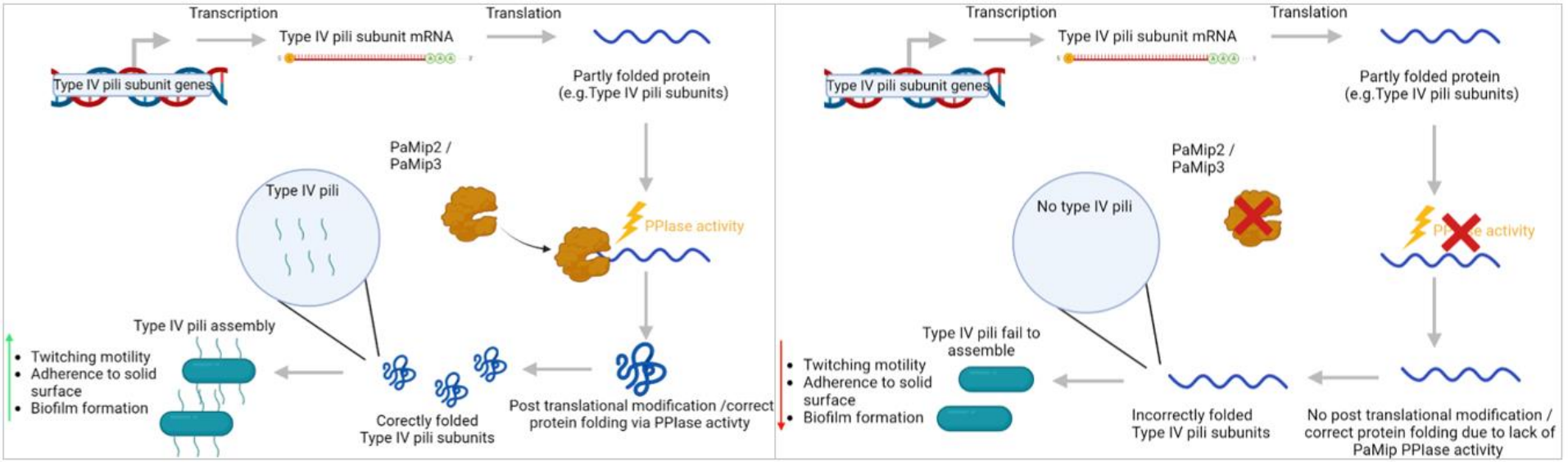


Figure 3.23. Diagram illustrating a second proposed mechanism by which PaMips are involved in virulence. In the left image, transcription of type IV pili biosynthesis genes are transcribed in strains I and J to levels comparable of the WT strain into mRNA (Table 3.5). The mRNA is translated and partly folded. To become fully folded type IV pili subunits capable of assembling into type IV pili appendages, the subunits must be folded further and post translationally modified courtesy of PaMip PPIase activity. Once correctly folded and modified, the type IV pili subunits are then able to assemble together to form the type IV pili structures responsible for surface attachment, initiating biofilm formation and twitching motility. The image on the right is the same process however in strain I or J, whom lack PaMips. Without the PaMips the partly folded type IV pili subunits will not undergo further post translational modification and will remain only partly folded. Without correct folding, the subunits are not able to assemble resulting in no type IV pili appendages on the cell surfaces, as was seen in the SEM images for strains I and J. Without the presence of these appendages, cell surface attachment is not initiated resulting in attenuation in biofilm formation and twitching motility lost.

A second example of a protein that plays a similar role to that of Pin1 in bacteria and is a major facilitator of virulence processes is DsbA [253, 254]. DsbA is an enzyme primarily responsible for disulphide bond formation between two amino acid side chains in proteins that are secreted. For DsbA to function, a key proline residue within its structure must be in the correct isomer. If this isomer is not converted from the *cis* to *trans* configuration, DsbA is unable to form disulphide bonds in the tertiary structure of a protein. Thus, a PPIase within a bacteria must facilitate this isomerisation. The exact PPIase in PA01 which is responsible for this isomerisation of the proline residue in DsbA is unknown [255]. Upon translation of a secreted protein into the periplasm, DsbA mediates disulphide bond formation ensuring its correct assembly. Without functioning DsbA these proteins would fail to assemble correctly and be unable to exert their desired function. Again, the mechanism by which DsbA is involved in virulence is a reasonable model by which to explain how PaMips may be exerting their role in virulence using their PPIase activity and its necessity in folding virulence associated proteins / transcription factors.

Although I have established that PaMips are essential for PA01 virulence, how they mediate this virulence remains unsolved and this is something which is yet to be answered for all of the Mips described in the literature. How we will address this will be discussed in my closing remarks. To ensure that all of the results described in this chapter are a direct consequence of PaMip gene deletions, complementation of PaMips was the next important step and will be discussed in the next chapter.

Chapter 4:

Complementation of Mip-gene deletions to restore *in vivo* and *in vitro* virulence phenotypes, whole genome sequencing and the prophage problem

Chapter 4. Complementation of Mip-gene deletions to restore *in vivo* and *in vitro* virulence phenotypes, whole genome sequencing and the prophage problem

4.1 Introduction

Following the results presented in Chapter 3, to definitively conclude these now annotated Mip genes (*pamip1*, *pamip2*, *pamip3*) in PA01 were essential for virulence and ensuring the results obtained were a direct consequence of the intended Mip gene deletions and not downstream deleterious effects or genomic changes as a result of the genetic manipulation of the Pa01 strain, complementing the genes to restore phenotypes was essential. If, upon complementation of the Mip-genes, phenotypes were restored to that of the WT PA01 strain, it would be safe to conclude the attenuated virulence phenotypes observed were a direct consequence of the Mip gene deletions, thus overall concluding that PA01 does possess Mip-genes essential for full virulence.

Given that only double knockout strains I and J presented with significant phenotypic differences to the WT in plate-based virulence assays, coupled with lab time constraints due to Covid-19, they were the only strains to be complemented along with the single mutant strains from which they were made. Furthermore, considering the single knockout strains D (Δ*pamip1*), E (Δ*pamip2*), and F (Δ*pamip*3) presented with no real phenotypic differences in virulence assays to those of the WT, the double knockout strains I (Δ*pamip1*/Δ*pamip2*) and J (Δ*pamip2*/Δ*pamip3*), comprised of combinations of gene knockouts in Δ*pamip1*, Δ*pamip2* and Δ*pamip3*, complementation of a single gene in either I or J, effectively restoring them to a single deletion genotype, should hypothetically restore their phenotypes to that of the WT, D, E and F strains.

Several complementation strategies were attempted. The first and most favourable strategy involved the use of a chromosomally integrative suicide plasmid, pUC18-mini-TN7T-Gm, which inserts a single copy of a gene, under the control of its native promoter at a neutral attTn7 site in the PA01 genome, downstream of the *glmS* gene [188]. This approach was used to replicate, as closely as possible, normal transcript levels. The second complementation strategy was performed using the plasmid pJH10Ts created by previous lab member Tom Scott, which expresses a gene under the control of a constitutive promoter on the plasmid [187]. The final approach was using the pME6032 plasmid which expresses an inserted gene under an inducible *lac* promoter.

In parallel to complementation, and to ensure no other genomic changes had occurred as a result of genetic manipulation, all strains were sent away for full genome sequencing to MicrobesNG (University of Birmingham). This sequencing revealed a concerning artefact for the version of strain I created by Silke Alt, a 3275 bp region of the genome was missing and another 8450 bp region was duplicated. Given these genomic changes, phenotypes and results obtained for Silke's mutant I were now compromised, and I could no longer say for certain these results were a direct consequence of the Mip gene deletions. How these issues were addressed will be

described below. It is worth noting that the results for mutant I presented in Chapter 3 were derived using my version of mutant I rather than Silke's. However, the phenotypes for both versions were found to be the same across all assays.

This chapter will discuss all complementation strategies with their success initially evaluated using *in vitro* plate based assays before moving to *in vivo* experiments. The first two complementation attempts failed *in vitro* so *in vivo* analysis was not performed. At this point whole-genome sequencing was performed which revealed the missing 3275 bp region. How this issue was addressed using several approaches will be discussed. This issue was ultimately resolved by recreating strain I, ensuring the region remained in the genome via whole-genome sequencing and repeating all assays described in chapter 3 to ensure all phenotypes of my strain I were identical to Silke Alt's version. This rigorous testing confirmed that deletions in *pamip1* / *pamip2* were directly responsible for the virulence attenuation in all assays. Following this confirmation, the final complementation strategy was undertaken. This final complementation strategy used the pME6032 vector housing either a single Mip gene (returning double knockout strains I and J back to single knockout strains they were created from i.e. strain D, E or F), or two Mip genes which returned the strain back to the WT genotype. These complementation strains were initially assessed *in vitro* which showed some degree of success in returning phenotypes of strains I and J *in vitro.* With some degree of success *in vitro*, *in vivo* analysis was performed in the *G. mellonella* infection model which interestingly showed all complementation vectors restored killing of the larvae by strains I and J.

4.2 Complementation using pUC18-mini-TN7T-Gm based plasmids

The use of transposon mini-TN7 vectors to complement genes in gram-negative bacteria by inserting themselves into the chromosome in a site-specific manor is common practice. Transposons are mobile genetic elements that are able to 'exchange places' (transpose), into genomes at specific target sites. Transposon Tn7 and its relatives have been shown to insert themselves into the chromosome of several gram-negative bacteria at high frequency and with excellent site-specificity. Gram-negative bacteria possess a TN7 specific attachment site termed *attTN7* that is always located immediately downstream of the *glmS* gene, most bacteria only possess a single copy of the *glmS* gene [188]. Having only a single specific site to integrate into made this TN7 transposon an ideal candidate to create a TN7-mini-based vector for the insertion of DNA into a chromosome. The creation of mini-TN7-based vectors has become a useful genetic tool for gene complementation in gram-negative bacteria [256].

The genome of PA01 contains only a single *attTN7* site (54 bp in size) making the use of a TN7 based vector an attractive candidate to complement Mip gene deletions [257]. Here, the vector pUC18-mini-TN7T-Gm was used to complement PaMip genes. Although this vector allows the insertion of DNA into a chromosome at a specific site, a promoter is still required. To achieve native levels of transcription, the native transcriptional promoter for each Mip gene was predicted bioinformatically and upstream regions for genes *pamip2* (84 bp) and *pamip3* (125 bp) were taken. Although these upstream regions appear small, they were used as these two Mip genes do not appear to comprise part of an operon and were predicted to possess their own transcriptional promoter (Figure 4.1). The construct to complement Mip gene *pamip1* was designed differently as the gene makes up part of a large operon, sitting next to the penultimate gene in the operon. Here, 280 bp upstream of the first gene in the operon was taken, with the transcriptional promoter for the operon, predicted to be encoded in this stretch (Figure 4.1 and Figure 4.2). This 280 bp region was effectively 'cut and paste' next to the beginning of the *pamip1* gene in the pUC18-mini-TN7T-Gm construct. Constructs were designed using Geneious Prime software and were ordered from GenScript (Table 2.3).

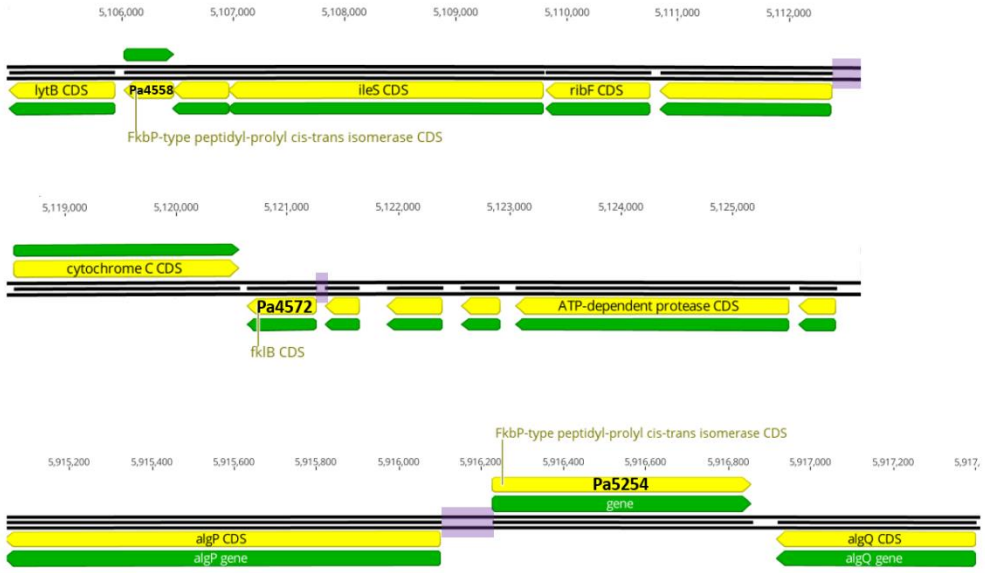


Figure 4.1. Designing the Complementation constructs for *pamip2* and *pamip3*. This Image was obtained from Geneious Prime used to represent the chromosome of PA01 and the genomic context of each Mip gene. Each coloured arrow denotes an annotated or hypothetical gene. The top image is for the gene encoding *pamip1* (*pa4558*) which is adjacent to the penultimate gene in the operon. Highlighted in purple are the regions predicted to house the transcriptional regulator for each Mip gene, for *pamip1*, this purple region is located at the beginning of the operon (Figure 4.2). For the genes encoding PaMip2 and PaMip3 plasmids were ordered from Genscript by taking the entire gene coding region as well as the highlighted with restriction sites added either end for the cloning into pUC18-mini-TN7T-Gm.

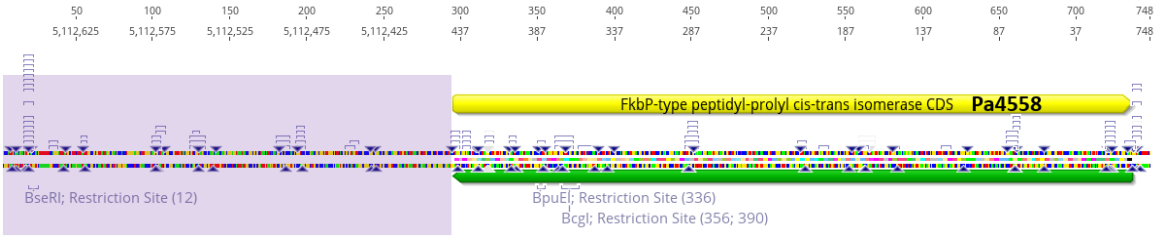


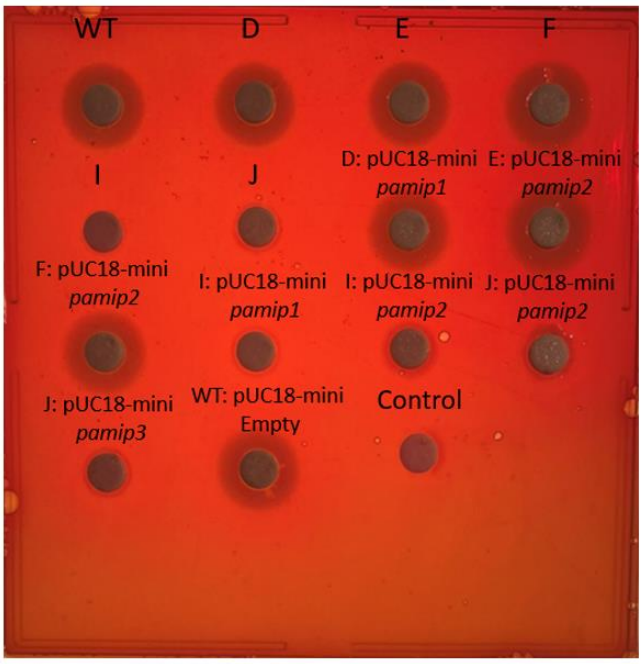
Figure 4.2. Designing the Complementation construct for PaMip1. This image was created to depict the 'cut and paste' 280bp upstream region of the PaMip1 encoding operon where the native transcriptional promoter was predicted to be encoded. This construct was designed in Geneious Prime with restriction sites designed to clone into the MCS of pUC18-mini-Gm-TN7T and ordered from Genscript who sent the linear construct in a cloning vector pUC57.

All constructs ordered from GenScript arrived in pUC57 housing vector which were sent for sequencing to ensure their fidelity. Once confirmed, the synthesised fragments for each Mip gene (promoter region + *pamip1* / promoter region + *PaMip2* / promoter region + *PaMip3*) were digested and ligated into pUC18-mini-TN7T-Gm, with correct assembly confirmed by colony PCR and sequencing. The assembled pUC18-mini-TN7T-Gm vectors housing each Mip gene were then transformed into the relevant strains via electroporation and co-transformed with the helper plasmid pTNS2. Strains D, E, F, I and J were all complemented. The single knockout strains D, E and F were complemented with their relevant Mip-gene deletion. Double knockouts I and J were complemented using a single Mip gene independently, thus effectively returning double knockouts to single knockouts; I (Δ*pamip1*/Δ*pamip2*) : pUC18-mini-TN7T-Gm promoter region + *pamip1*, I(Δ*pamip1*/Δ*pamip2*):pUC18-mini-TN7T-Gm promoter region + *pamip2*, J (Δ*pamip2*/Δ*pamip3*):pUC18-mini-TN7T-Gm promoter region + *pamip2* and J (Δ*pamip2*/Δ*pamip3*):pUC18-mini-TN7T-Gm promoter region + *pamip3*. To ensure the pUC18-mini-TN7T-Gm constructs had integrated into the chromosome at the correct *attTN7* site, colony PCR using primers up and downstream of the *attTN7* site were used. If the insertion was incorrect the colony PCR would reveal a band at around 325 bp in size, if the insertions were correct the bands would be around 1000 bp. Once confirmed, these strains were used in assays to assess whether complementation was successful.

4.2.1 Haemolysis and protease assays using pUC18-mini-TN7T-Gm complementation constructs

To assess whether complementation was successful, strains confirmed to have pUC18-mini-TN7T-Gm constructs successfully integrated into the chromosome at the correct *attTN7* site were used. Given all complementation constructs were designed to express complemented Mip genes under the control of their predicted native promoter, no inducement was required while also, theoretically, the complemented genes should have been transcribed at the normal WT levels which would prevent any unusual phenotypes observed as a result of over or under expression. To assess whether complementation was successful in the restoration, or at least in part, of phenotypes, haemolysis and protease assays were used. These assays were ideal as they provide unambiguous measurable and observable phenotypes. Furthermore, these assays are straight forward and highly repeatable. If complementation failed to restore phenotypes of mutant strains to those of the WT in these assays, it was unlikely that the complementation would have restored phenotypes in other more complicated, and time-consuming assays such as the *G. mellonella* infection model.

1



2

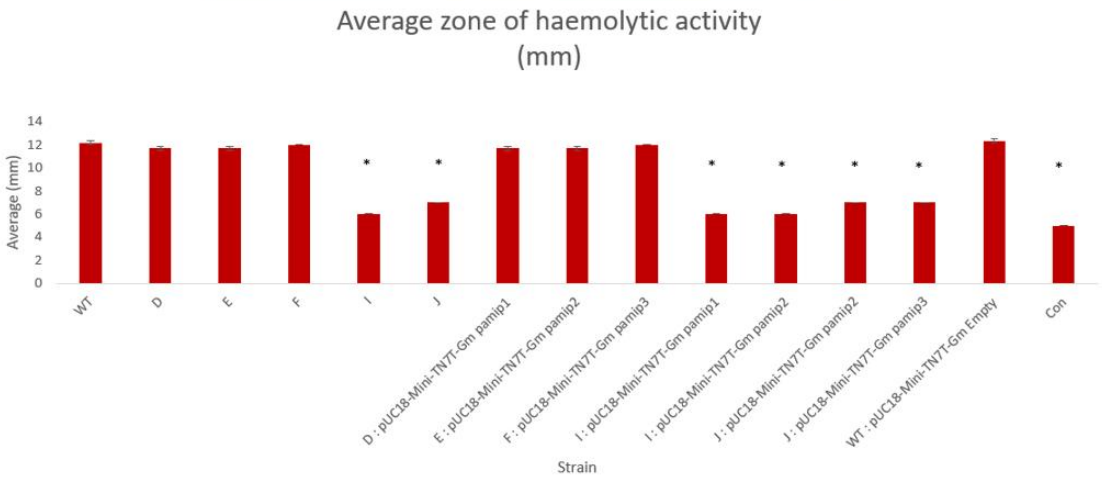
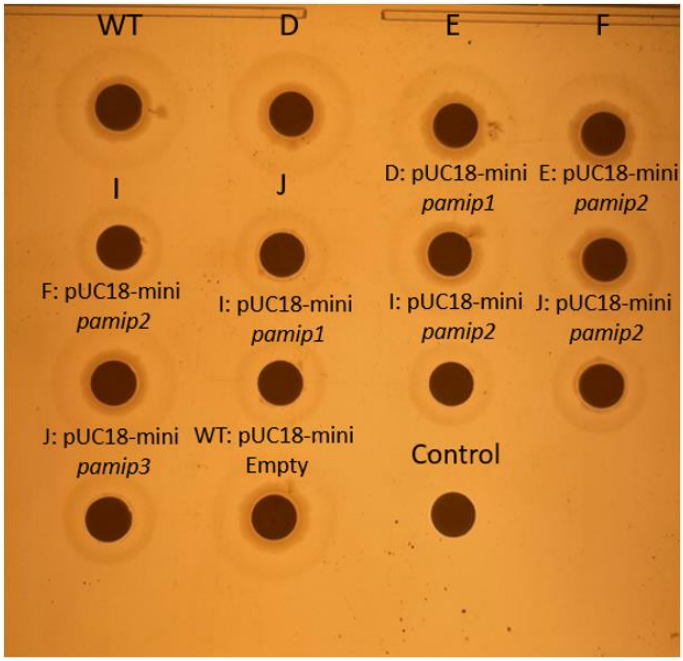


Figure 4.3. Haemolytic activity of PaMip mutant strains transformed with complementation vector pUC18-mini-TN7T-Gm; Panel 1 shows blood agar plates with strains showing clear zones of haemolytic activity compared to the control using sterile water. The WT and single Mip knockout strains D, E and F have large zones of activity compared to those of strains I and J, both of which have a clear, reduced halo. Transforming strains D, E and F with complementation vectors did not change their phenotypes as expected. Both strains I and J when transformed with the complementation vectors housing individual PaMip genes failed to complement the reduced haemolytic activity. The WT transformed with empty pUC18-Mini-TN7T-Gm vector behaved normally and still demonstrated haemolytic activity. The bar chart in panel 2 shows the average zones of haemolytic activity measured from plates. Along the X axis are the Mip knockout strains described in Table 3.1. Standard error variance was calculated and plotted as seen by the error bars comparing the difference between the mean with the amount of scatter within the groups, error bars that overlap are not statistically significant. The asterisks above columns indicate ($P<0.05$, $n=3$) whereby a strain demonstrates significantly different haemolytic activity when compared to the WT (see methods).

1



2

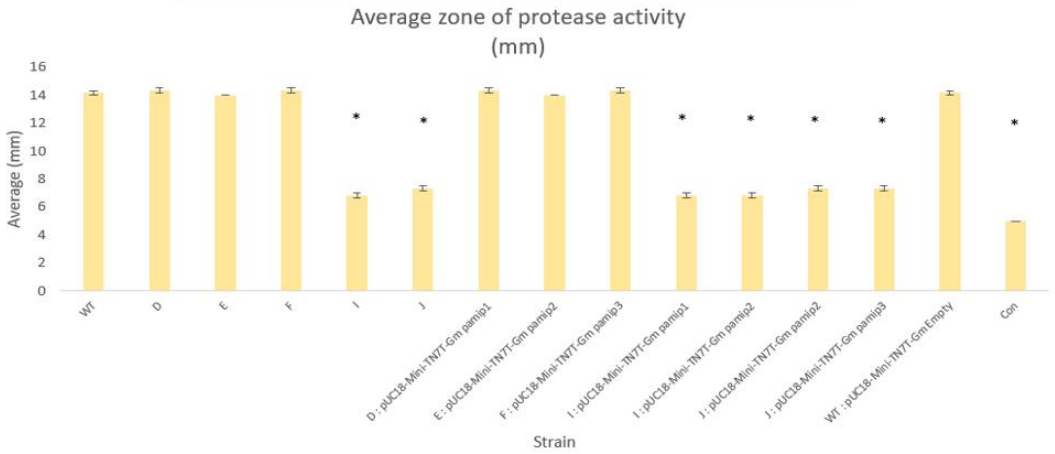


Figure 4.4. Proteolytic activity of PaMip mutant strains transformed with complementation vector pUC18-mini-TN7T-Gm; Panel 1 shows lactose-free whey protein agar plates with strains showing clear zones of proteolytic activity compared to the control using sterile water. The WT and single Mip knockout strains D, E and F have large zones of activity compared to those of strains I and J, both of which have a clear, reduced halo. Transforming strains D, E and F with complementation vectors did not change their phenotypes as expected. Both strains I and J when transformed with the complementation vectors housing individual PaMip genes failed to complement the reduced proteolytic activity. The WT transformed with empty pUC18-Mini-TN7T-Gm vector behaved normally and still demonstrated proteolytic activity The bar chart in panel 2 is the average zones of proteolytic activity measured from plates. Along the X axis are the Mip strains as seen in Table 3.1. Standard error variance was calculated and plotted as seen by the error bars comparing the difference between the mean with the amount of scatter within the groups, error bars that overlap are not statistically significant. The asterisks above columns indicate ($P<0.05$, $n=3$) where by a strain demonstrates significantly different haemolytic activity when compared to the WT (see methods).

For both haemolysis and protease assays no changed phenotypes were restored to that of the WT. The WT strain was transformed with empty pUC18-mini-TN7T-Gm vector to ensure the WT phenotypes remained unaffected. All single mutants D, E and F were complemented with their respective deleted Mip gene and presented with identical phenotypes to the WT, as they had done in the initial assays. These results provided information that the complementation constructs and subsequent genetic manipulation of these strains were of no detriment to the strains which behaved as expected. The double Mip knockout strains I and J were complemented with each of the two Mip genes removed (effectively restoring strain I back to either single knockout strain D or E, and J to E or F, depending on which Mip gene was complemented). Given single Mip gene strains D, E and F had no observably different phenotypes to the WT, restoring a double Mip knockout strain (I or J) to a single Mip knockout strain via complementation should theoretically partly or fully restore phenotypes.

No phenotypes were restored for strains I and J, with the transformed, complemented strains still producing significantly less haemolytic and proteolytic activity compared to the WT ($P<0.05$, $n=3$) (Figures 4.3 and 4.4). Furthermore, there were no statistical differences in the amount of haemolytic and proteolytic activity between strains I and J and their complemented counterparts (Figures 4.4 and 5.5). Each construct was sequenced to ensure it was correct, as well as performing colony PCR for each Strain transformed with a complementation vector to ensure the correct insertion of the pUC18-mini-TN7T-Gm vector into the *attTN7* site, with bands subsequently also sequenced. Complementation using this method may have failed due to issues around uncovering and using the correct transcriptional promoter for each PaMip gene. Theoretically these constructs should have provided complementation as they were designed by taking stretches of DNA upstream of each PaMip gene that were predicted to house each gene's transcriptional promoter. These predictions however, were based on 'best guesses' and may not have contained the correct transcriptional promoter required for optimal expression of each PaMip gene. Even if the correct transcriptional promoter was captured, other factors around the optimal level of expression of the PaMip genes may have been a limiting factor at either a transcriptional or translational level.

In an attempt to negate this native transcriptional promoter issue, but while still using the preferred complementation strategy using the integrative pUC18-mini-TN7T-Gm vector, I attempted to clone a Ptac promoter with the lac operator via Gibson Assembly into the pUC18-mini-TN7T-Gm vector just before the MCS began where the PaMip genes were cloned in to. This would ensure the presence of a known promoter and theoretically allow expression of the genes cloned into the MCS. Unfortunately, this approach failed at the first hurdle with the amplification

of the Ptac promoter and lac operator proving unsuccessful after several attempts and several different primer combinations and PCR cycles.

Given the ambiguities around using the native promoter for each gene, the second complementation approach using the pJH10TS vector which was performed in parallel to this strategy. Given two complementation strategies were being implemented, analytical techniques to assess whether the complementation constructs were in fact expressing the Mip genes to those of the WT levels, such as Western Blot, were not performed.

4.3 Complementation using pJH10TS

Following the unsuccessful attempt to complement WT phenotypes in Mip gene knockouts using the pUC18-mini-TN7T-Gm strategy, a second approach was undertaken. The potential issue with the previous strategy may have been down to the failure to encapsulate the correct transcriptional regulator, given the sequences taken were only predicted to house them. To negate this potential issue, the second approach used a broad range IncQ expression vector with an IPTG-inducible Ptac promoter that is constitutively expressed. This vector, called pJH10TS, was constructed by previous lab member Tom Scott who used it extensively to complement *P. fluorescens* mutants of the obafluorin biosynthetic pathway. The pJH10TS vector has been used by several other lab members to varying success for complementing gene knockouts in *Pseudomonas* species and *E. coli*.

This vector can be difficult to digest and ligate into, so Gibson Assembly was used to create complementation constructs. The pJH10TS vector was linearised using the restriction enzyme XbaI and genes encoding PaMip1, PaMip2 and PaMip3 were amplified using primers containing 20 bp overlapping arms identical to 20bp upstream and 20 bp downstream of the XbaI restriction site in the pJH10TS MCS. Once confirmed via sequencing, the amplified genes and the linearised vector were assembled via Gibson Assembly. The mix was transformed into *E.coli* and resultant colonies were checked by colony PCR and sequencing. Once confirmed, these vectors were extracted and transformed into the relevant strain. Here, the WT strain was transformed with empty pJH10TS to ensure the WT phenotypes remained unaffected and presented the same as the untransformed WT due to the introduction of the pJH10TS vector which had primarily been used in *P. fluorescens*. Single knockout strains D, E and F were complemented with their respective Mip gene that had been deleted. Double knockout strains I and J were complemented with each of the single Mip genes as described above for the pUC18mini-TN7T-Gm approach, restoring strain I back to either single knockout strain D or E, and strain J back to E or F.

4.3.1 Haemolysis and protease assays using pJH10TS complementation constructs

As was done for the pUC18-mini-TN7T-Gm strategy, to assess whether complementation was successful in the restoration, or at least in part, of phenotypes, haemolysis and protease assays were used.

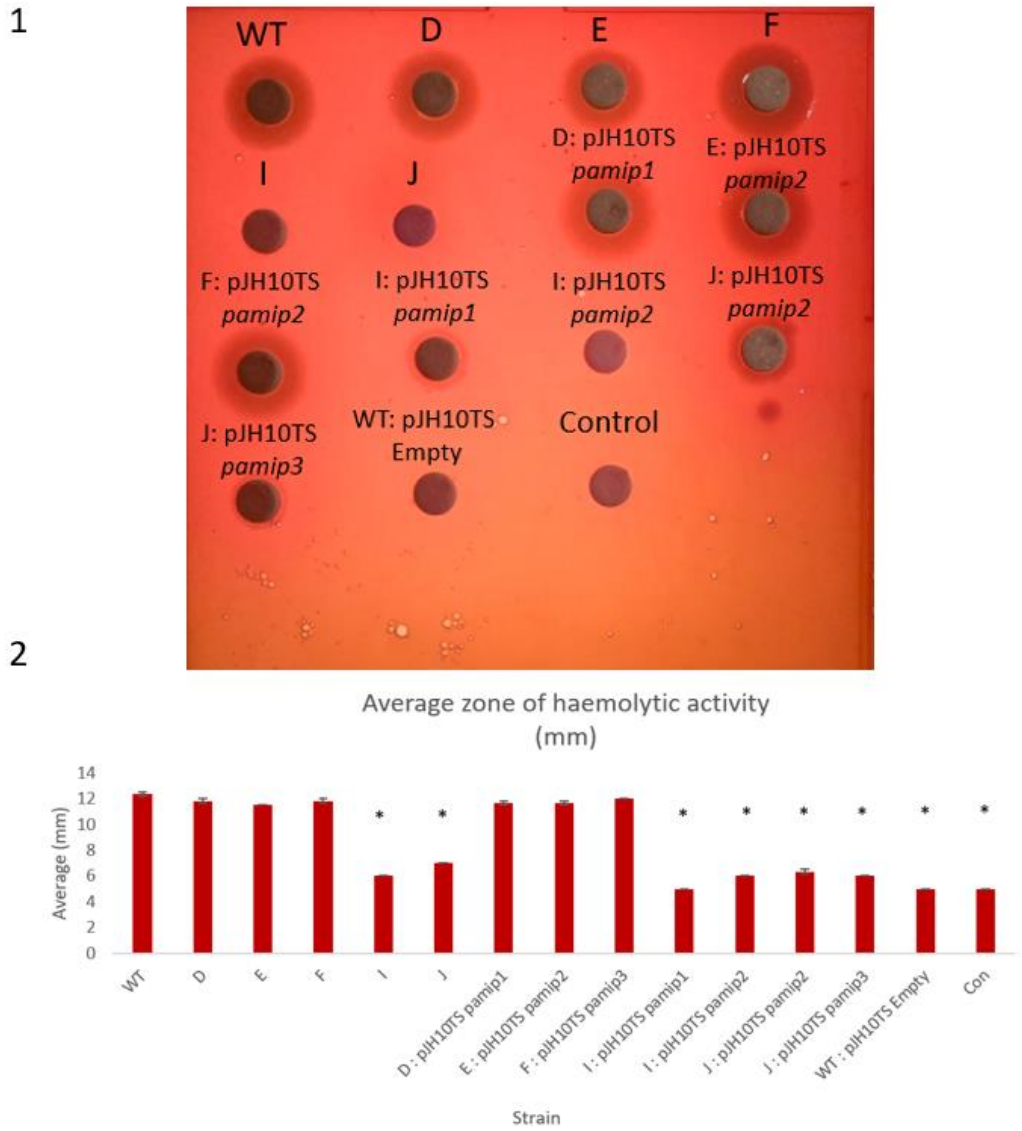


Figure 4.5. Haemolytic activity of PaMip mutant strains transformed with complementation vector pJH10TS; Panel 1 shows blood agar plates with strains showing clear zones of haemolytic activity compared to the control using sterile water. The WT and single Mip knockout strains D, E and F have large zones of activity compared to those of strains I and J, both of which have a clear, reduced halo. Transforming strains D, E and F with complementation vectors did not change their phenotypes as expected. Both strains I and J when transformed with the complementation vectors housing individual PaMip genes failed to complement haemolytic activity. The WT transformed with empty pJH10TS vector demonstrated a complete loss of haemolytic activity across several different WT strains housing the plasmid suggesting this vector was unsuitable to continue any complementation studies. The bar chart in panel 2 is the average zones of haemolytic activity measured from plates. Along the X axis are the Mip knockout strains as seen in Table 3.1. Standard error variance was calculated and plotted as seen by the error bars comparing the difference between the mean with the amount of scatter within the groups, error bars that overlap are not statistically significant. The asterisks above columns indicate ($P<0.05$, $n=3$) where by a strain demonstrates significantly different haemolytic activity when compared to the WT (see methods).

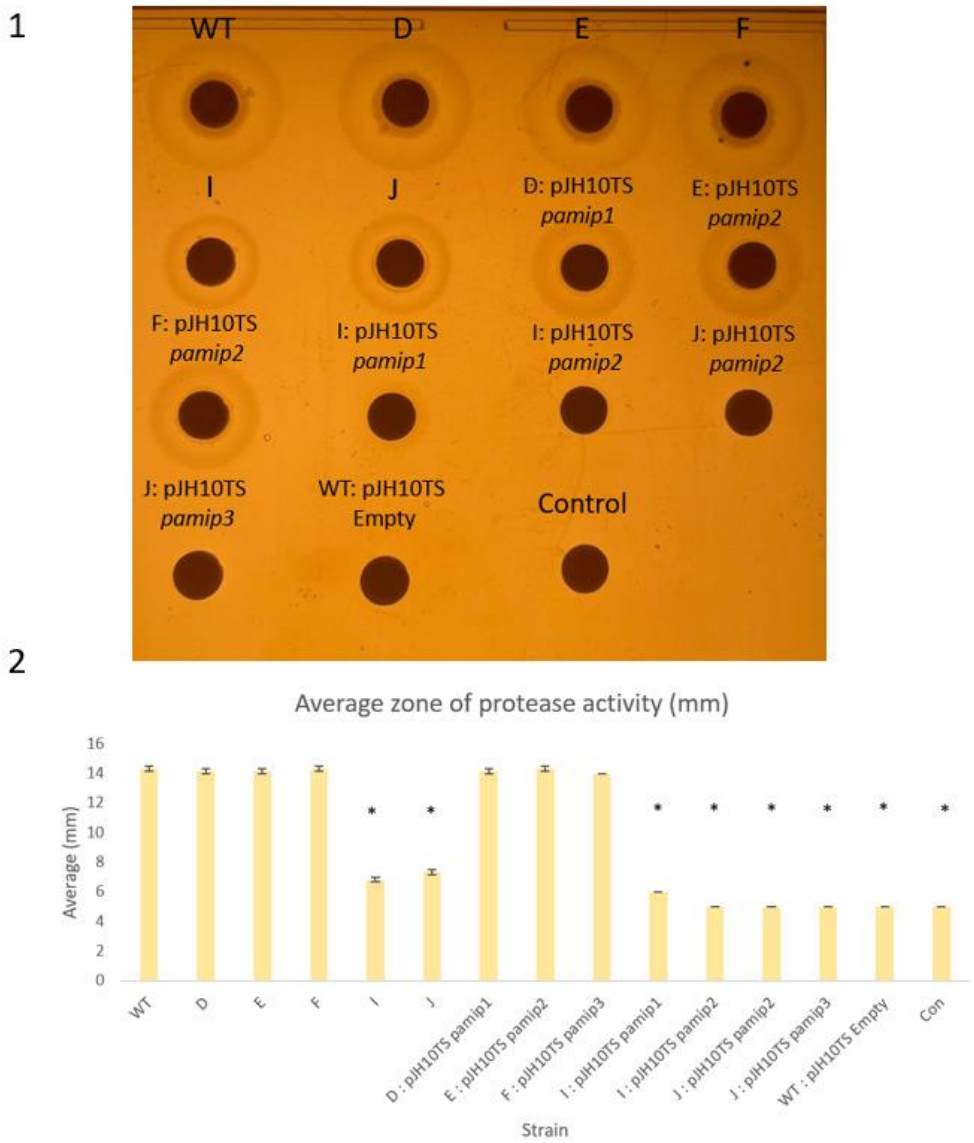


Figure 4.6. Proteolytic activity of PaMip mutant strains transformed with complementation vector pJH10TS; Panel 1 shows lactose-free whey protein agar plates with strains showing clear zones of proteolytic activity compared to the control using sterile water. The WT and single Mip knockout strains D, E and F have large zones of activity compared to those of strains I and J, both of which have a clear, reduced halo. Transforming strains D, E and F with complementation vectors did not change their phenotypes as expected. Both strains I and J when transformed with the complementation vectors housing individual PaMip genes failed to complement proteolytic activity. The WT transformed with empty pJH10TS vector demonstrated a complete loss of haemolytic activity across several different WT strains housing the plasmid suggesting this vector was unsuitable to continue any complementation studies. The bar chart in panel 2 is the average zones of proteolytic activity measured from plates. Along the X axis are the Mip strains as seen in Table 3.1 Standard error variance was calculated and plotted as seen by the error bars comparing the difference between the mean with the amount of scatter within the groups, error bars that overlap are not statistically significant. The asterisks above columns indicate ($P<0.05$, $n=3$) where by a strain demonstrates significantly different haemolytic activity when compared to the WT (see methods).

As was the case for the pUC18-mini-TN7T-Gm strategy, both the haemolysis and protease assays using strains transformed with pJH10TS complementation vectors produced no observable and measurable phenotypes, failing to restore any activity. Unfortunately, the control experiment in which the WT was transformed with empty vector lost both haemolytic and proteolytic activity (Figures 4.5 and 4.6). To ensure this was not an isolated occurrence, I repeated the transformation of empty pJH10TS vector into the WT strain several independent occasions, picking positive colonies from independent transformations and repeated the assays which exhibited the same loss of WT activity.

Not only did the WT strains lose activity, strains D, E and F housing pJH10TS complementation vectors saw a reduction in proteolytic activity. Complementation of PaMip genes in I and J failed to restore any haemolytic or proteolytic activity. This failure to restore any phenotype was not a surprise given the WT strain transformed with pJH10TS lost all activity. Previous works using this vector and its parent vectors had predominantly been to complement genes in *P. fluorescens*, with these results suggesting pJH10TS is not entirely compatible with *P. aeruginosa* strains and may exert some negative effect. These results indicated that the pJH10TS constructs and the strains complemented with these constructs were not suitable for further study.

4.4 Whole Genome Sequencing and the missing pre-prophage region

Following the failure of the first two complementation strategies in restoring any observable virulence phenotypes for strains I and J, I began to question the validity of the results obtained for the Mip-gene knockout strains. One would assume that if the Mip-gene deletions were the only variable and sole cause of the reduced virulence phenotypes observed in assays presented in Chapter 3, then reintroducing these genes should fully, or in part, return WT phenotypes. Given this was not the case for the first two strategies, it was decided to send all strains for whole genome sequencing to ensure there were no other unintended genomic changes that would account for the phenotypes observed for the mutant strains as a result of the genetic manipulation and deletion of Mip genes in the WT strain.

Although the first two complementation strategies failed, they were not checked to determine whether they were optimally expressing their encoded Mip-gene product. If these expression vectors were not expressing their housed gene, this would explain why phenotypes were not restored to those similar to the WT. Should the vectors have been fully expressing their encoded Mip-gene, one would assume this would have resulted in the complementation of phenotypes. This could have been checked by western blot analysis which would have confirmed whether the complementation vectors were expressing the Mip-genes. Given time constraints in the lab, the next best alternative was to send strains away for whole-genome sequencing which could

be analysed in the background as I continued other complementation strategies. The WT strain and mutant strains D, E, F, I and J were all sent for whole genome sequencing and the returned data was analysed by Dr Govind Chandra and then compared to the published PA01 genome.

4.4.1 Whole Genome Sequencing returned and the pre-prophage problem

The genome sequencing of the WT strain and mutant strains D, E, F and J returned with no anomalies when compared to the known PA01 genome, other than those anticipated in the form of PaMip gene deletions. Alarmingly, upon the return and analysis of the whole genome sequence data for strain I, a 3275 bp region was found to be missing, with a second ~8450 bp region found to have been duplicated. To ensure this result was reliable, and not due to poor sequence coverage at that particular region, I performed a series of colony PCRs (Figure 4.7). These colony PCRs further confirmed that the first region was indeed missing. At this point, until this unexpected, potentially hypothesis changing, issue was addressed and resolved, the complementation of genes in these strains was halted.

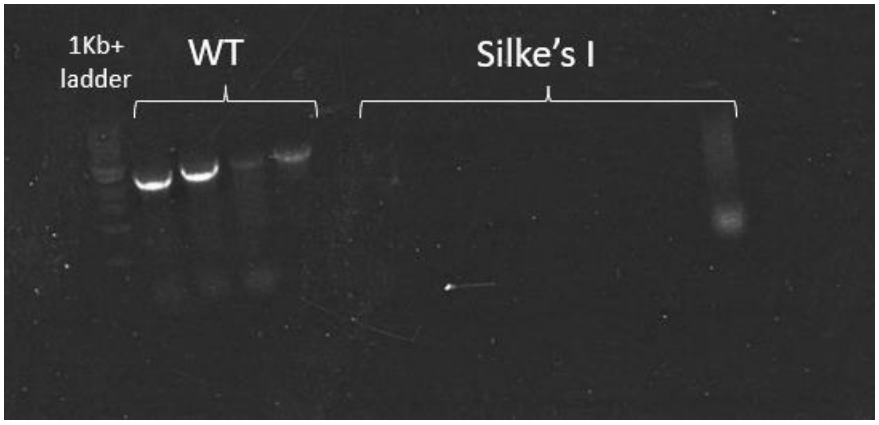


Figure 4.7. Image of an analytical agarose gel ran with PCR product assessing the presence of the pre-prophage region in strain I. Image of an agarose gel ran with colonies picked for colony PCR to assess the 3275 bp pre-prophage region. This region has been lost in Silke Alt's version of strain I.

The missing 3275 bp region is situated upstream of a region encoding a filamentous Pf1 prophage and is centred around a highly repetitive codon region that includes genes *pa0715* and *pa0716* (Figure 4.8). The duplicated ~8450 bp region was immediately downstream of this 3275 bp missing region starting at the p*a0717* gene and ending at *pa0728*. Interestingly, the ~8450 bp pf1 prophage coding region found to be duplicated in Silke's strain I houses the same genes shown to be highly upregulated in the RNA seq for strain J as previously discussed in Chapter 3.7. Furthermore, as previously mentioned, literature reports have linked the prophage coding region from a virulent / pathogenic lifestyle to a more non-invasive, sessile lifestyle. One might assume given strain I possessed this virulence enhancing region in duplicate, it would display an increase in processes associated with the sessile lifestyle such as biofilm formation. However, in all assays this was not the case.

Due to this unexpected output, results obtained for Silke's strain I were questionable, as the phenotypes observed could not be attributed as a direct consequence of *pamip1*/*pamip2* gene deletions. For the purposes of scientific rigor, and to further assess this observation, two approaches were taken: 1) to recreate strains D, E and I myself, and to subject these new strains to virulence assays ensuring the same phenotypes were observed to those of Silke's strain I, and then to send my strain I would be sent for full genome. 2) The second approach performed in parallel, was to attempt deleting this 3275 bp region in the WT and perform virulence assays using the resulting strain to determine if any phenotypic variation was observed when missing this region.

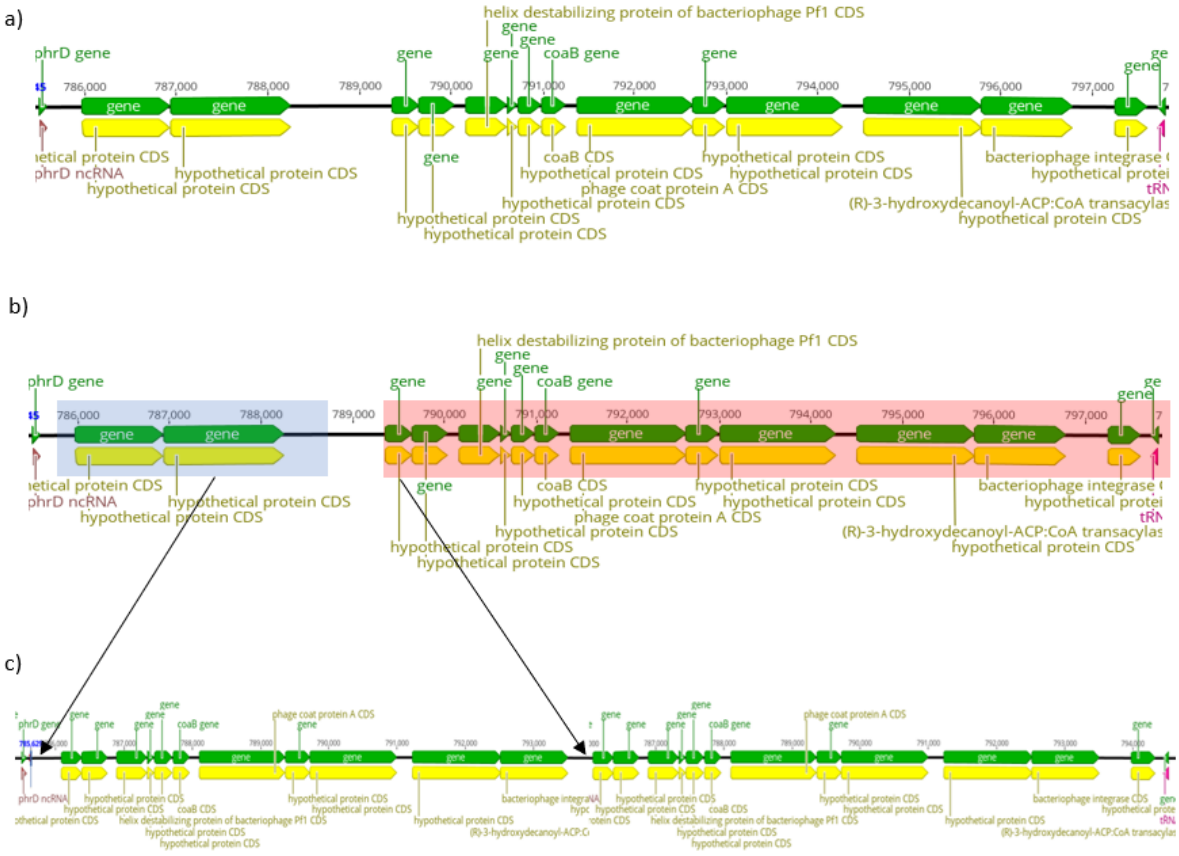


Figure 4.8. Diagram highlighting the changes in strain I's genome from the whole-genome sequencing. Panel a) shows the entire Pf1 prophage region as it should appear in the WT PA01 strain. Panel b) highlighted in blue is the 3275 bp region that is missing from Silke Alt's version of strain I, while highlighted in red is the entire replicated region in this strain. The arrows map onto panel c) which shows what the returned sequencing revealed and how strain I's chromosome appeared, missing the 3275 bp pre-prophage region and containing the duplicated ~8450 bp prophage region.

4.4.2 Recreation of strains D and E

As mentioned, single knockout strains D and E whose single Mip gene deletions, when combined, make up strain I, possessed no unintended genomic changes. This potentially suggested that when the second gene was deleted in a single strain, the unintended genomic changes occurred. Given D and E make up strain I, for completeness I recreated both strains D and E as well as strain I. For the recreation of mutant strains D, E and I, I designed the constructs in an identical manner to that used by Silke Alt, using the same primers, restriction sites, homologous up and downstream regions taken for Fragment A and Fragment B, thus also leaving the same codons in place at the beginning and end of both *pamip1* and *pamip2* genes (Materials & Methods 2.5). This method results in a truncated gene deletion rather than end to end gene deletion which was preferential to negate potential deleterious polar effects.

I then followed the same *P. aeruginosa* gene knockout protocol step-by-step as used previously by Silke. Producing a knockout in the *pamip2* gene, thus creating mutant strain E, proved efficient (5 days in total) with 50% of colonies showing a successful double-crossover event. Mutant bands were cut from gels and sent for sequencing to ensure the mutation was correct. Glycerol stocks were prepared from these positively identified mutants. Producing a single knockout in the *pamip1* gene proved difficult. The process from start to finish was repeated five times, with 100 colonies (taken from different plates) checked by colony PCR for each attempt (Figure 4.9). On the fifth attempt only 2 out of the 100 bands presented with the correct mutant band sizes (Figure 4.10). These bands were cut from gels and sent for sequencing to ensure the mutation was correct and to confirm the creation of a single knockout in the *pamip1* gene, thus creating mutant strain D. Glycerol stocks were prepared from these positively identified mutants.

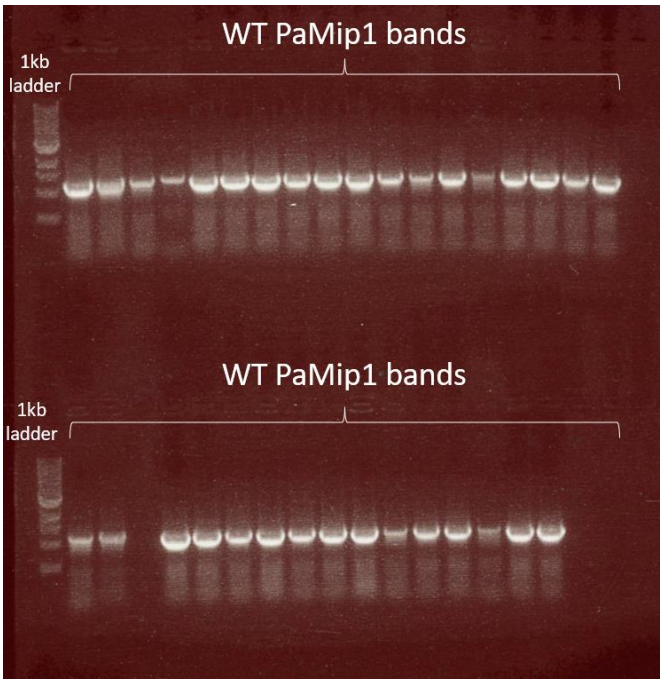


Figure 4.9. Image of an analytical agarose gel image assessing double cross-over event for *pamip1* knockout. Example image of an agarose gel for colonies picked for colony PCR when searching for a successful double-cross event resulting in mutation of the *pamip1* gene. All colonies present with the WT band size for the *pamip1* gene around 1290 bp in size with no mutants obtained which would be expected at around 400 bp in size. This image is representative as to what was seen for the first four knockout attempts having screened over 400 colonies not a single mutant band was found.

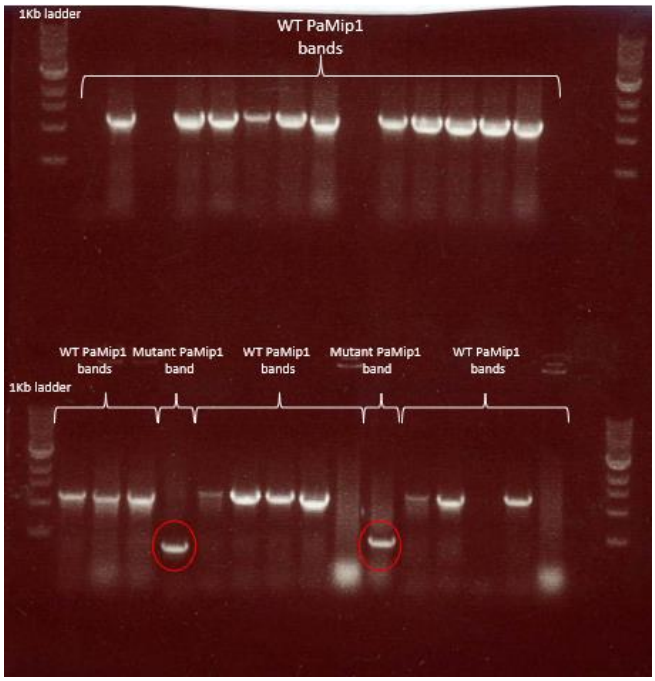


Figure 4.10. Image of an analytical agarose gel image confirming double cross-over event for *pamip1* knockout. Image of an agarose gel for colonies picked for colony PCR when searching for a successful double-cross event resulting in mutation of the *pamip1* gene. Most colonies presented with the WT band size for the *pamip1* gene around 1290 bp however on the fifth knockout attempt there were two colonies presenting with mutant band at around 400 bp in size, circled in red. These two colonies were taken as glycerol stocks. Each band was extracted and sent for sequencing to confirm the correct sequence for the mutation.

4.4.3. Presence of pre-prophage region in newly made strains D and E

To ensure at this stage that the 3275 bp region had not been lost, colony PCR was performed from plates streaked with the newly made strains D and E using primers sitting outside the 3275 bp region and primers sitting within this region. The primers sitting outside the 3275 bp region struggled to reliably amplify the region (Figure 4.11). This could have been due to the GreenGo Taq polymerase losing efficiency in trying to amplify such a large stretch of DNA; alternatively, this may have been a potential consequence of the region possessing highly repetitive sequence stetches at the beginning and the end of the 3275 bp sequence. The internal primers were designed to amplify a 1000 bp region sitting directly at the centre of the 3275 bp region; thus, should the overall 3275 bp region be missing, these primers would fail to bind and no new bands would be seen. These primers produced reliable results across all colonies evidencing that the 3275 bp region was still present in these newly created versions of strains deletion D and E (Figure 4.12).

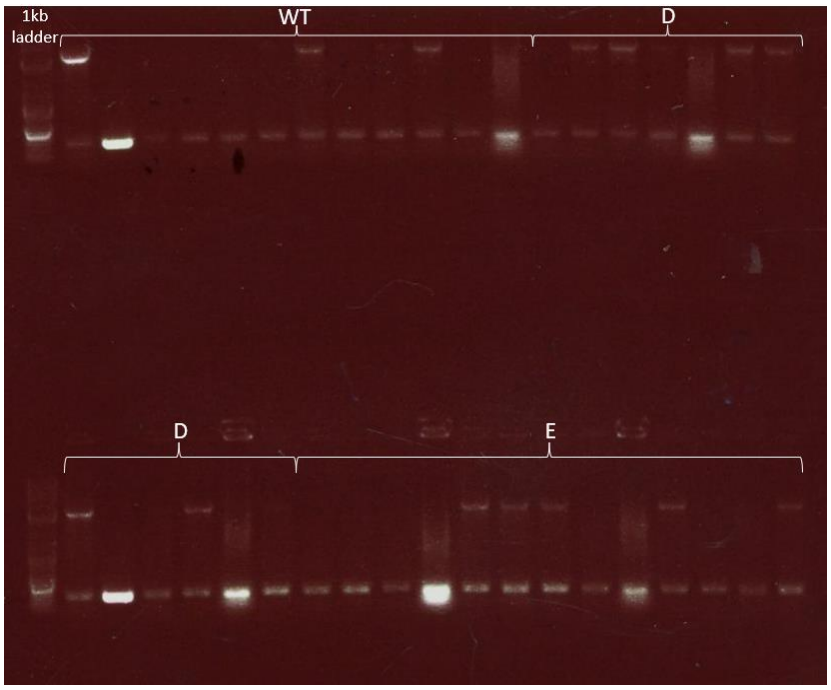


Figure 4.11. Image of an analytical agarose gel image assessing the reliability of primers amplifying a 3275 bp region. Example image of an agarose gel ran with colonies picked for colony PCR to ensure the external colony PCR primers were highly efficient in amplifying the 3275 bp pre-prophage region in the WT strain and newly created mutant strains D and E. All bands for this region are faint and appear to be present in only 1 in 2 colonies suggesting these primers were less than optimal for conclusive analytic purposes.

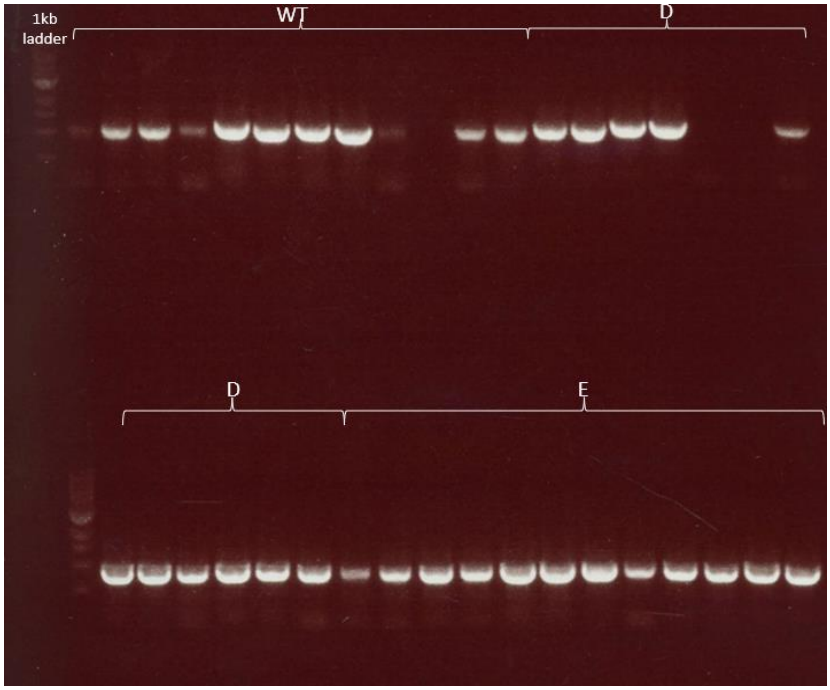


Figure 4.12. Image of an analytical agarose gel image confirming reliability of second set of primers amplifying internal region of the 3275 bp. Example image of an agarose gel ran with colonies picked for colony PCR to ensure the internal colony PCR primers were highly efficient in amplifying the shorter internal 1000 bp pre-prophage region in the WT strain and newly created mutant strains D and E. These primers consistently amplified the region and were used moving forward to assess the presence of the pre-prophage region.

4.4.4. Recreation of strain I

The methodology applied by Silke Alt in creating strain I was again successfully applied. Thus, strain D (Δ*pamip1*) was used as a starting point from which *pamip2* was to be deleted, leading to a new version of strain I. Additionally, I attempted to recreate strain I with strain E (Δ*pamip2*) as a starting point from which *pamip1* was to be deleted. However, this second approach proved unsuccessful despite several attempts, which is unsurprising giving the difficulty faced in knocking out *pamip1* to create strain D.

With the successful recreation of strain I, assessing whether or not it still possessed the 3275 bp prophage region, was essential. This was done using a series of colony PCRs along with sending the new strains D, E and I away for whole genome sequencing. Using the internal colony PCR primers to reliably amplify the 1000 bp region that lies in the centre of the 3275 bp region revealed my new strain I still possessed the region (Figure 4.13). The presence of this 3275 bp region was further confirmed upon the return of the whole genome sequencing from MicrobesNG, which also indicated no evidence of a duplication event having occurred for the ~8450bp region. Following this exhaustive confirmation process, assays were performed to ensure the new strain displayed phenotypes similar to those of Silke's original version. This was shown to be the case, thus we could conclude the phenotypes observed were independent of

any issues associated pre-prophage region and were a direct consequence of the intended *pamip1* and *pamip2* gene deletions.

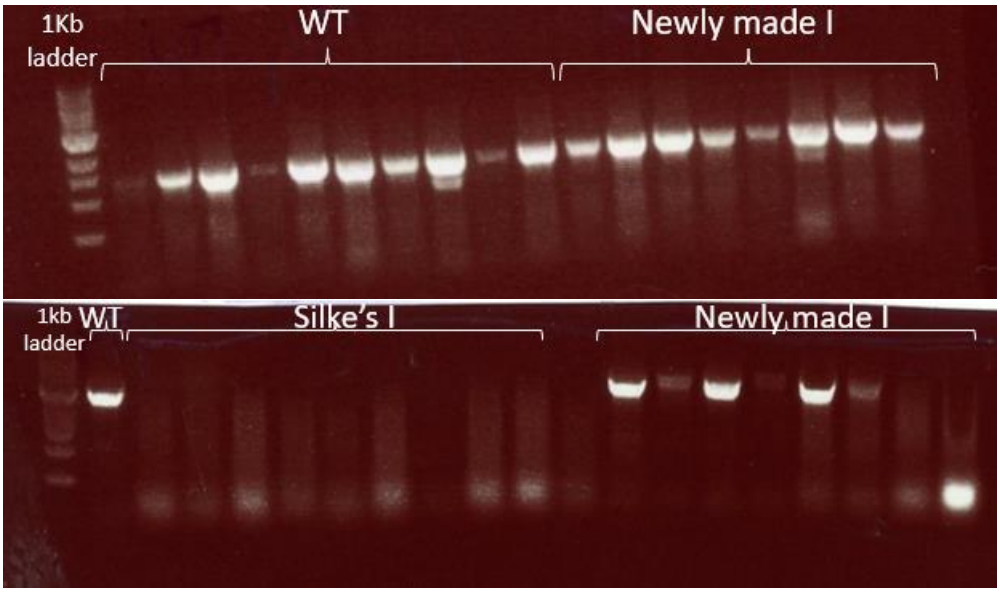


Figure 4.13. Image of an analytical gel ran with colony PCR products to assess whether the newly created strain I contained the 3275 bp region. This used internal primers that amplified the internal 1000 bp region found at the centre of the 3275 bp region missing in Silke's I strain. The WT was used as a positive control to ensure the PCR reaction was successful in amplifying the target region known to be still present in the WT strain which allowed a comparison of my newly made I. Silke's strain I was used as a negative control as no bands should appear in lanes using colonies from this strain given it was confirmed this strain did not have this region. The image shows clearly the newly created strain I contains the prophage region confirming it had not been lost as a consequence of the *pamip1* and *pamip2* gene deletions.

Given the phenotypes for both versions of strain I were identical, I concluded the phenotyping results obtained for both strains were valid. Subsequently, when performing further assays with these strains, I periodically plated the O/N LB culture for the purposes of colony PCR to continually ensure the genomic integrity of the strains. Interestingly I discovered that, over time, the prophage region began to delete from the new version of strain I (Figure 4.14). Given that in the literature this prophage and its coding region has not been subject to in-depth behavioural studies, it is difficult to explain this phenomenon. Given what we know about prophages as a whole and their ability to integrate into a genome with ease, it is plausible they are able to jump out of a genome under 'unfavourable' conditions [258]. Based on my observations, it seems reasonable to suggest that the strain I genetic background, the loss the *pamip1* and *pamip2* is somehow unfavourable for the retention of the prophage region.

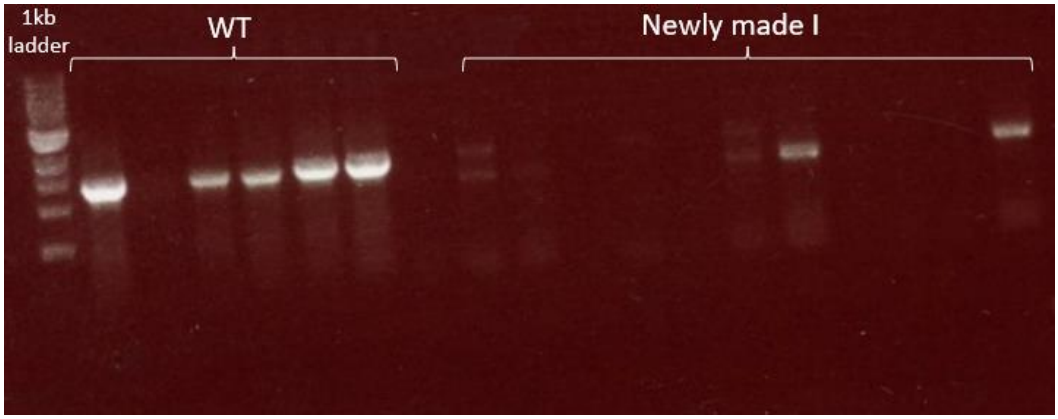


Figure 4.14. Image of an analytical gel ran with colony PCR products to assess whether the newly created strain I continually contained the 3275 bp region when compared to the WT strain. Colony PCRs were ran using internal primers that amplified the internal 1000 bp region found at the centre of the 3275 bp region. This image was taken several months after the creation of the newly made strain I which demonstrates the WT Pa01 strain to reliably possess the prophage region with clear bright bands, compared to the WT, the newly made strain I possess faint bands for the prophage region with around only 1 in 2 colonies appearing to still possess the region. Despite initial colony PCRs for this region in the newly made mutant I demonstrating clear, bright bands (Figure 4.13), coupled with the whole genome sequencing confirming the presence of this region, it appears, overtime, this region is lost in the Δ*pamip1*/Δ*pamip2* gene knockout background.

4.4.5 Targeted deletion of the pre-prophage region

In parallel to recreating strain I, I attempted to intentionally delete the 3275 bp pre-prophage region. My hypothesis was that, if the deletions of *pamip1* and *pamip2* in strain I were the sole cause of the reduced virulence observed, then a strain lacking just the 3275bp pre-prophage region should present with the same phenotypes as the WT. However, if in these assays the 3275 bp mutant evidenced the same phenotypes as those seen in Silke's strain I, it was more likely that this 3275 bp region was responsible for the phenotypes and reduction in virulence observed in strain I, not the PaMip gene deletions. The deletion of this region was performed using the same gene deletion strategy that has been described and used for all genes thus far (homologous recombination using the plasmid pTS1).

Having followed the gene deletion protocol from start to finish on seven separate occasions, and picking over 2000 colonies across numerous different plates, a deletion in this region proved impossible to achieve. Initial rounds of colony PCR using primers external and internal to the 3275 bp region identified mutant bands which initially proved promising, these strains were taken as glycerol stocks (Figure 4.15). However, when streaked out again and analysed by colony PCR to verify the purity of the glycerol stock, a WT band was found. This happened on numerous occasions that a mutant was thought to be identified (Figure 14.6).

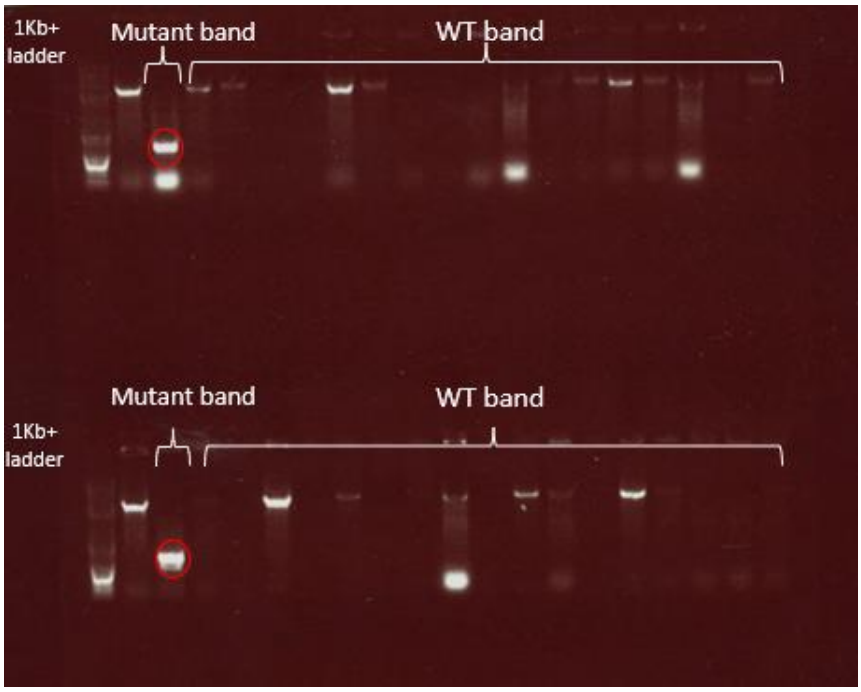


Figure 4.15. Image of an analytical gel ran with colony PCR products when searching for a successful double-cross event resulting in deletion of the 3275 bp pre-prophage region. The WT bands are around 3275 bp in size, whereas a successful double-cross over event resulting in a deletion of the 3275 bp region would appear around 600bp in size.

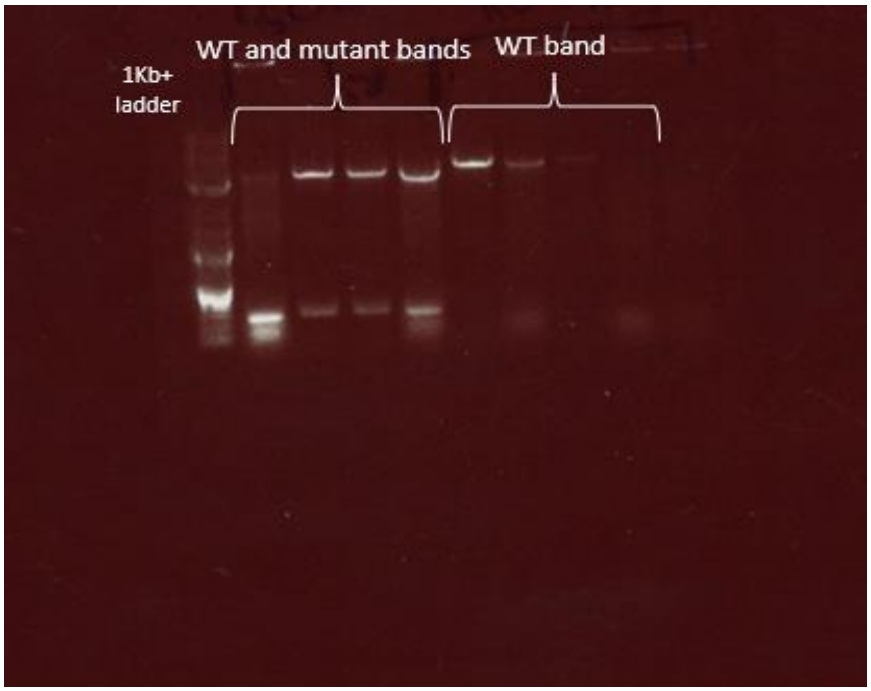


Figure 4.16. Image of an analytical gel ran representative as to what was seen when checking colony PCR products to assesses whether the single colonies taken as glycerol stocks that initially presented with the correct 3275 bp pre-prophage region deletion (evidenced in Figure 4.15.) were pure and retained the intended deleted region when grown in liquid culture and streaked out. When streaked out across plates and colony PCR performed on single colonies, both the WT and mutant bands were seen suggesting mixed populations existed even within a single colony.

This phenomenon was observed in each independent experiment performed. All manner of troubleshooting was performed; however, no mutant was created that reliably possessed and retained the 3275 bp pre-prophage deletion. As previously mentioned, this bizarre phenomenon may be partly explained due to the nature of non-lytic filamentous prophages. When under less-than ideal conditions in the strain, the region may exist in quiescent and is excised from the genome. Fortunately, the experiments targeting the recreation of strain I described above were successful and able to clear up any ambiguities surrounding the missing pre-prophage region and its potential effect on the phenotypes observed for strain I. Thus, work on deletion of this pre-prophage region was halted. However, this region might be of future interest given its unpredictable behaviour. Although this may sound a long, arduous process, it was essential research to be performed, without which, any definitive conclusions around the necessity of PaMip genes for the optimal virulence processes of PA01, could not have been drawn.

4.5 Complementation using pME6032

The final complementation strategy was to use an expression vector, pME6032 [189]. This plasmid was initially derived from a pVS1 shuttle vector found in a *Pseudomonas* species. The pVS1 shuttle vector and its recombinant vector derivatives were shown to exhibit stable characteristics when transformed into PA01 suggesting these vectors were ideal candidates to edit for expression purposes in PA01 [189]. The pME6032 vector was designed by removing the kanamycin resistance cassette found in its origin pVS1 vector and replacing this with a tetracycline resistance cassette along an expanded MCS sitting downstream of the strong Ptac promoter and its *lacIq* repressor [259]. Thus, expression of genes from pME6032 is induced by the addition of IPTG, however the promoter is leaky and does not always require IPTG to induce low level expression.

The PaMip genes were cloned into pME6032 as single PaMip gene constructs, and, for the first time, as double PaMip constructs. Strains I and J were transformed with both the singular PaMip gene vectors (effectively restoring them to a single knockout strain), and with vectors containing both PaMip genes, effectively restoring them to the WT background (Table 4.1). Again, theoretically, given that single Mip knockout strains D, E and F demonstrated no clear phenotypic differences for most assays when compared to the WT, the restoration of a singular PaMip gene into the double knockout strains I and J should restore their phenotypes to those observed for their single knockout parent strains, which were essentially the WT phenotypes. As done previously, to initially assess whether complementation was successful both haemolysis and protease assays were performed.

Strain	Complementation construct	Expected phenotypic return
WT	pME6032: Empty	WT
I (Δ*pamip1* / Δ*pamip2*)	pME6032: *pamip1*	D / WT
I (Δ*pamip1* / Δ*pamip2*)	pME6032: *pamip2*	E / WT
I (Δ*pamip1* / Δ*pamip2*)	pME6032: *pamip1*/*pamip2*	D / E / WT
I (Δ*pamip1* / Δ*pamip2*)	pME6032: *pamip2*/*pamip1*	D / E / WT
I (Δ*pamip1* / Δ*pamip2*)	pME6032: Empty	I
J (Δ*pamip2*/ Δ*pamip3*)	pME6032: PaMip2	E / WT
J (Δ*pamip2*/ Δ*pamip3*)	pME6032: PaMip3	F / WT
J (Δ*pamip2*/ Δ*pamip3*)	pME6032: *pamip2*/*pamip3*	E / F / WT
J (Δ*pamip2*/ Δ*pamip3*)	pME6032: *pamip3*/*pamip2*	E / F / WT
J (Δ*pamip2*/ Δ*pamip3*)	pME6032: Empty	J

Table 4.1. Table laying out what strains were complemented with which vectors and the phenotype that would be expected to return. Complementation of either double knockout strain I or J with a single Mip gene would theoretically restore the phenotype observed for singular Mip knockout strains or to those phenotypes observed in the WT strain.

4.5.1 Haemolysis and protease assays using pME6032 complementation constructs

Once complementation constructs were confirmed by colony PCR and sequenced to check their correct assembly, they were transformed into their corresponding PaMip mutant strain. As was done for the pUC18-mini-TN7T-Gm and pJH10TS strategies, to assess whether complementation was successful in the restoration, or at least in part, of phenotypes, haemolysis and protease assays were used.

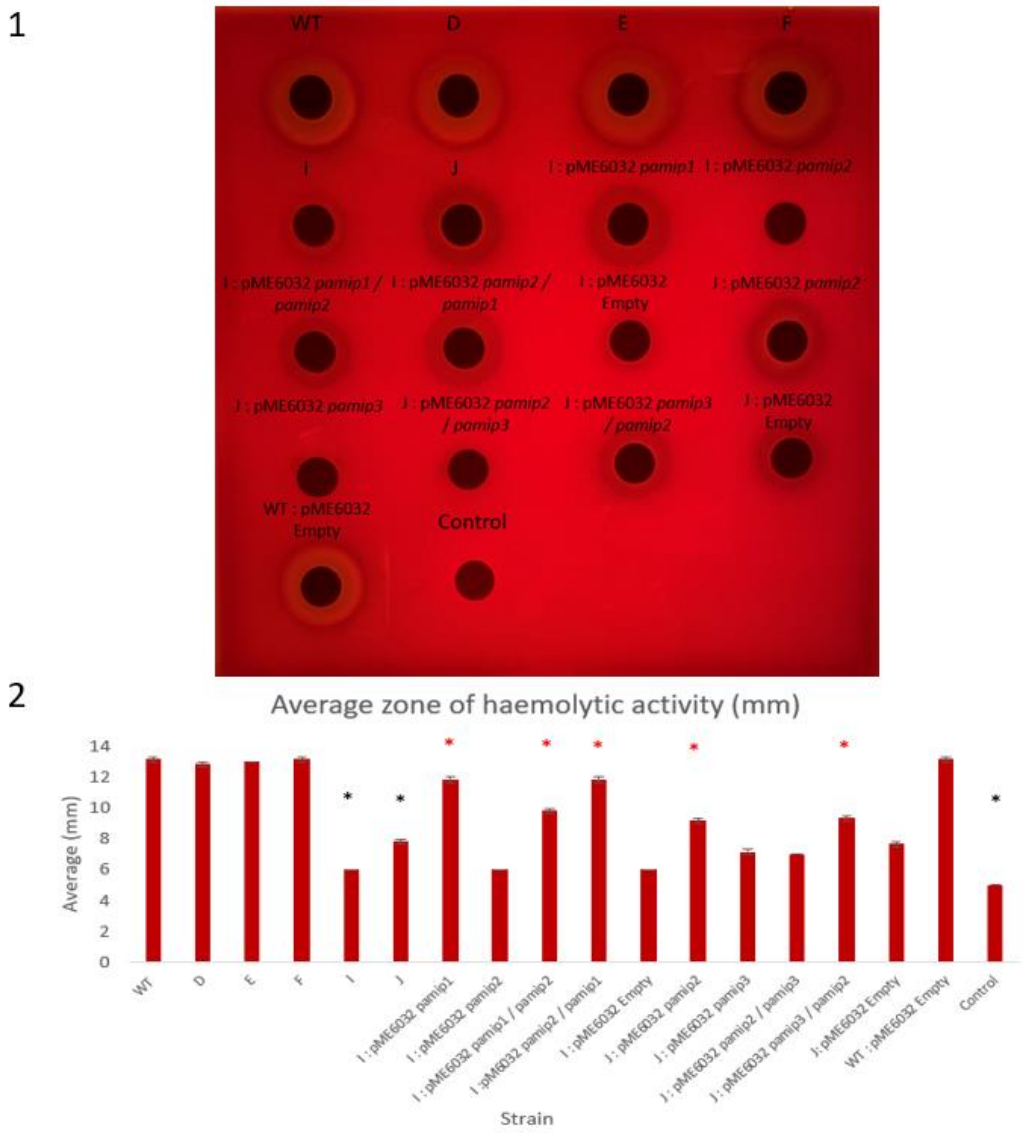
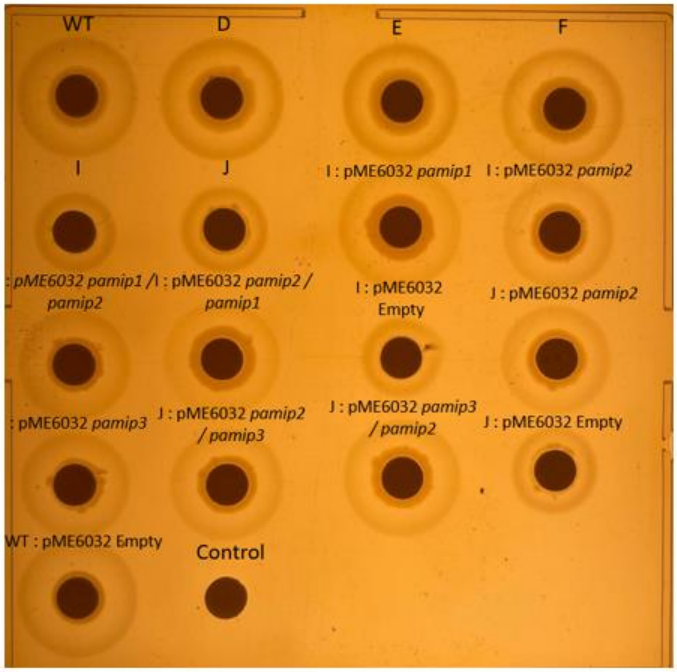


Figure 4.17. Haemolytic activity of PaMip mutant strains transformed with various pME6032 complementation vectors; Panel 1 shows blood agar plates with strains showing clear zones of haemolytic activity compared to the control using sterile water. The WT and single Mip knockout strains D, E and F have large zones of activity compared to those of strains I and J, both of which have a clear, reduced halo. Transforming strains Both strains I and J when transformed with the complementation vectors housing individual PaMip partly restored haemolytic activity in select cases. Transforming strains with pME6032 vector housing both PaMip genes also partly restored haemolytic activity in select cases. The WT transformed with empty pME6032 vector demonstrated no difference in haemolytic activity compared to the non-transformed WT strain. The bar chart in panel 2 is the average zones of haemolytic activity measured from plates. Along the X axis are the Mip knockout strains as seen in Table 3.1 and Table 4.1. Standard error variance was calculated and plotted as seen by the error bars comparing the difference between the mean with the amount of scatter within the groups, error bars that overlap are not statistically significant. The black asterisks above columns indicate ($P<0.05$, $n=3$) where a strain demonstrates significantly different haemolytic activity when compared to the WT (see methods). The red asterisks above columns indicate ($P<0.05$, $n=3$) where a complementation strain has been compared to its double knockout parent strain and exhibited significantly more haemolytic activity than its parent, thus confirming complementation was successful.

1



2

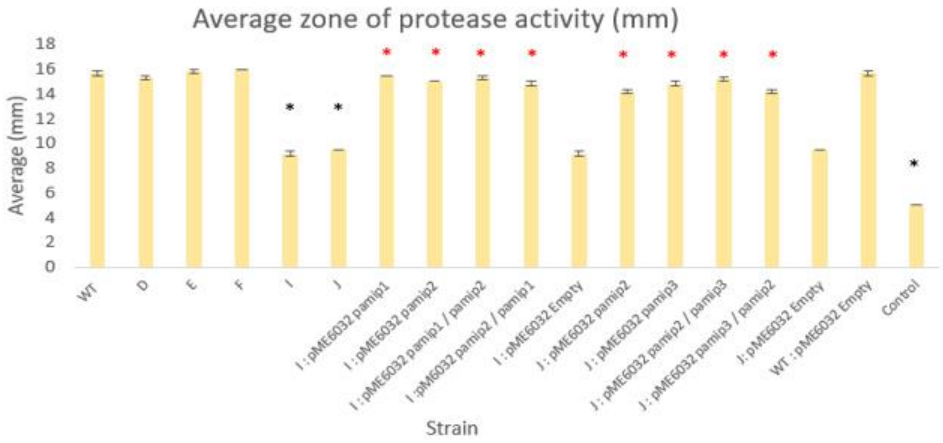


Figure 4.18. Proteolytic activity of PaMip mutant strains transformed with various pME6032 complementation vectors; Panel 1 shows lactose-free whey protein agar plates with strains showing clear zones of proteolytic activity compared to the control using sterile water. The WT and single Mip knockout strains D, E and F have large zones of activity compared to those of strains I and J, both of which have a clear, reduced halo. Transforming strains Both strains I and J when transformed with the complementation vectors housing individual PaMip partly restored haemolytic activity in select cases. Transforming strains with pME6032 vector housing both Mip genes also partly restored haemolytic activity in select cases. The WT transformed with empty pME6032 vector demonstrated no difference in haemolytic activity compared to the non-transformed WT strain. The graph in panel 2 is the average zones of haemolytic activity measured from plates. Along the X axis are the Mip knockout strains as seen in Table 3.1 and Table 4.1. Standard error variance was calculated and plotted as seen by the error bars comparing the difference between the mean with the amount of scatter within the groups, error bars that overlap are not statistically significant. The black asterisks above columns indicate ($P<0.05$, $n=3$) where a strain demonstrates significantly different haemolytic activity when compared to the WT (see methods). The red asterisks above columns indicate ($P<0.05$, $n=3$) where a complementation strain has been compared to its double knockout parent strain and exhibited significantly more haemolytic activity than its parent, thus confirming complementation was successful.

As with assays presented in 3.3.1 and 3.3.2, the WT and strains D, E, F, I and J all performed as expected, with I and J attenuated in both haemolytic and proteolytic activity (Figure 4.17 and 4.18). However, for the first time, complementation of phenotypes were observed. Strain I was complemented with singular PaMip genes, *pamip1* and *pamip2* independently (Table 4.1). Re-introducing the single *pamip1* gene (effectively returning this strain back to strain E like genotype), restored both haemolytic and proteolytic activity (clearance zones from 6mm to 11.83mm and 9.1mm to 15.5mm, respectively ($P<0.05$, $n=3$)); these levels were comparable to those of the WT (Figure 4.17 and 4.18). This result supports our initial hypothesis that effectively returning a double knockout strain to a single knockout strain would restore haemolytic and proteolytic activity. However, this was not the case for the reintroduction of the single *pamip2* gene into the strain I background; this did not return haemolytic activity, but it did return proteolytic activity. This variation in outcomes point towards how each PaMip of PA01 may exert its own independent role in virulence.

For strain J, reintroducing *pamip2* on pME6032 plasmid slightly increased haemolytic activity by around 1.3mm ($P<0.05$, $n=3$) compared with non-complemented strain J, although not to levels comparable to the single mutants or the WT. As with strain I, re-introducing *pamip2* to strain J restored complete proteolytic activity from 9.5 mm to 14.1 mm ($P<0.05$, $n=3$). Complementing *pamip3* into strain J failed to restore phenotypes and proved to hinder haemolytic activity. Strain J ordinarily displayed partial haemolysis with an average activity zone of 7.8mm, but when complemented with *pamip3* it exhibited visibly less haemolytic activity. This may suggest the overexpression of the *pamip3* gene negatively affects processes involved in haemolytic activity. In contrast, re-introduction of *pamip3* to strain J restored its proteolytic activity to levels comparable to the WT strain (from an average zone of activity of 9.5 mm to 14.8 mm when complemented). Transforming these strains with empty pME6032 had no effect on phenotypes, with the transformed WT still performing both haemolytic and proteolytic activity comparable to the untransformed WT strain, while transformed strains I and J remained attenuated in both haemolytic and proteolytic activity, comparable to their untransformed parent strain.

To further extend these observations and results, I took the single Mip gene pME6032 plasmids that had failed in complementing certain phenotypes, and ligated the second Mip gene, that was successful in complementing phenotypes into the vector creating a double complementation construct. These constructs should effectively return double knockout strains I and J to the WT genotype, not solely back to the genotypes of single PaMip deletion strains D, E and F as described above. Given complementation of the PaMip1 gene successfully restored both haemolytic and proteolytic activity in strain I (Figure 4.17 and 4.18), it was thought adding this *pamip1* gene into the *pamip2* pME6032 construct that failed to complement phenotypes in

strain I (thus creating a double Mip gene complementation construct pME6032: *pamip2*/*pamip1*), would restore haemolytic activity. Incredibly, this was found to be the case. As described above, the pME6032: *pamip2* vector transformed into strain I failed to increase any haemolytic activity, with inhibition zones (6mm) identical to the parent strain I. Upon the introduction of *pamip1* into the pME6032 *pamip2* vector (making it pME6032 *pamip2*/*pamip1*) and transformed into strain I, restoring I to the same genotype as the WT, full haemolytic activity was restored with average haemolytic activity zones restored to 11.8mm (Figure 4.17). Surprisingly, if the reverse to the above was performed, with *pamip2* introduced into the successful complementation construct pME6032: *pamip1* (creating pME6032: *pamip1* / *pamip2*), then transformed into strain I, haemolytic activity that was previously restored to levels comparable to those of the WT using the pME6032: *pamip1* construct, were reduced. Proteolytic activity remained unaffected and remained fully complemented.

A similar observation was made when using double complementation constructs in strain J. Transforming strain J with pME6032: *pamip3* failed to restore any observable phenotype in the haemolysis assays. Introducing *pamip2* into the unsuccessful pME6032: *pamip3* construct (creating pME6032: *pamip3*/*pamip2*), restoring J to the same genotype as the WT, zones of haemolytic activity increased by 2mm, similar activity zones to that observed in strain J complemented with pME6032: *pamip2* ($P<0.05$, $n=3$). Additionally, when *pamip3* was introduced into the successful complementation construct pME6032: *pamip2* (creating pME6032: *pamip2*/*pamip3*), a construct that had previously shown a partial increase of haemolytic activity in strain J, caused all haemolytic activity to be lost. This phenomenon was only observed in the haemolysis assay, with all complementation constructs for both I and J were able to restore proteolytic activity.

4.5.2 Biofilm formation assay using pME6032 complementation constructs

Following the partial success in complementation of both haemolysis and protease activity, assessing whether other phenotypes had been partially or fully complemented using these constructs was necessary. As previously mentioned, biofilm formation is a key virulence determinant of PA01 and is responsible for its characterisation as a chronic infection following its transition from the planktonic lifestyle.

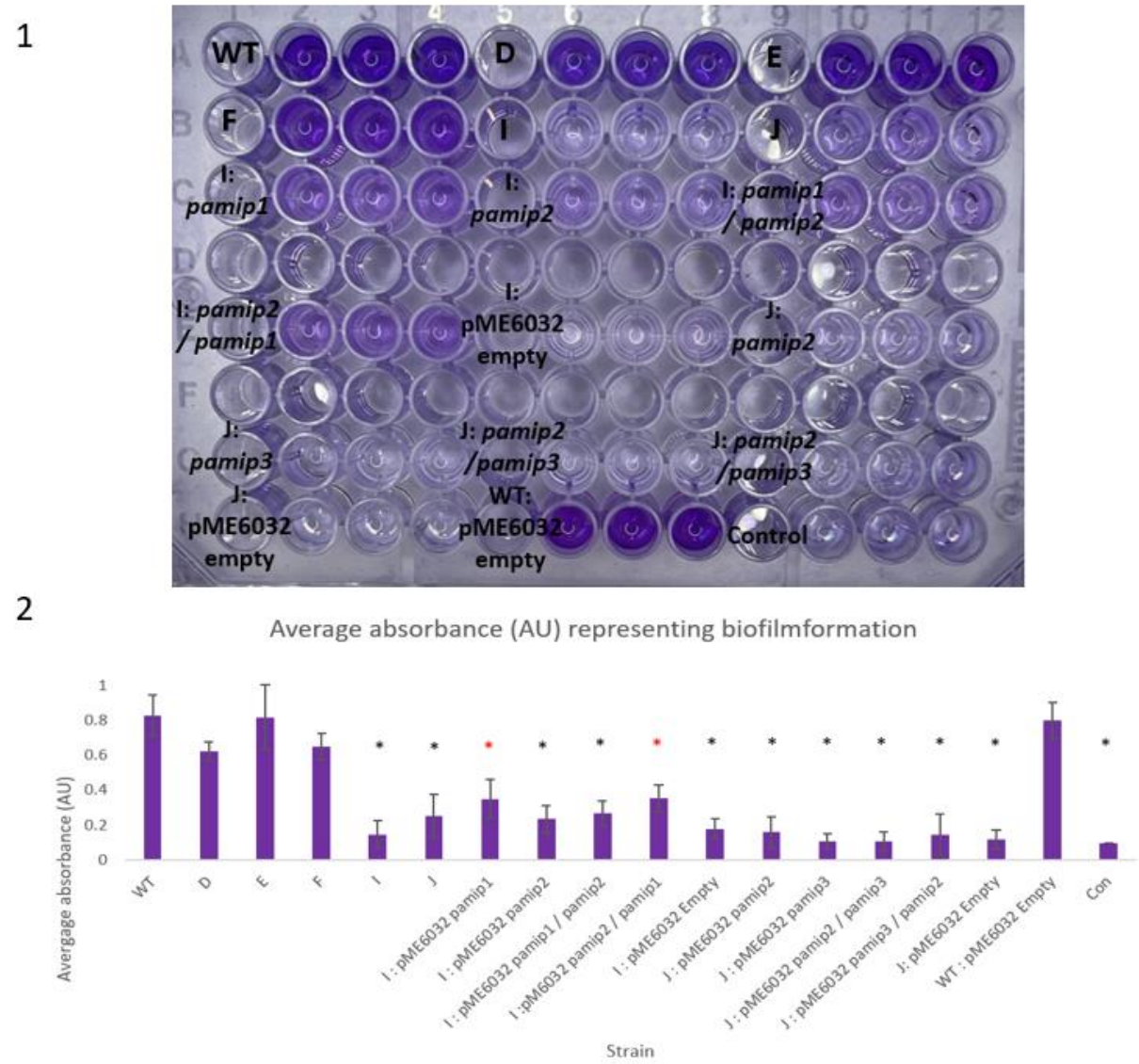


Figure 4.19. Biofilm formation assay using PaMip mutant strains transformed with the various pME6032 complementation vectors; Panel 1 Biofilm dissolved in acetic acid displaying different purple intensities across all strains carried out in triplicate. Example plate where strains have been grown in a 96-well plate then stained with crystal violet. Wells with a much lighter colour compared to the WT and other strains have a lower-level biofilm density. Strains are labelled from the WT to strain WT : pME6032 Empty vector, with each strain grown in triplicate. The biofilm was quantified by UV and plotted in the bar chart in panel 2 Standard error variance was calculated and plotted as seen by the error bars comparing the difference between the mean with the amount of scatter within the groups, error bars that overlap are not statistically significant. The black asterisks above columns indicate ($P<0.05$, $n=3$) where a strain demonstrates significantly different biofilm formation when compared to the WT (see methods). The red asterisks above columns indicate ($P<0.05$, $n=3$) where a complementation strain has been compared to its double knockout parent strain and exhibited significantly more biofilm formation than its parent, thus confirming complementation was successful.

The WT and strains D, E, F, I and J all performed as expected, with I and J attenuated in biofilm formation (Figure 4.19). To ensure the pME6032 vector had no effect the WT and strains I and J were transformed with empty pME6032 but showed no significant changes in biofilm formation (Figure 4.19).

Complementation with the various plasmids then led to similar outcomes as observed for the haemolysis and proteolysis assays. Complementation of strain I with *pamip1* led to a partial restoration of phenotype (Figure 4.19). Complementing strain I with *pamip2* failed to increase the amount of biofilm formed in strain I (increased absorbance). As described in the haemolysis and proteolysis assays, introducing *pamip1* into the pME6032: *pamip2* vector partially restored biofilm formation with a small, but significant increase. Transforming empty pME6032 vector into strain I had no effect on biofilm formation density, with levels comparable to untransformed strain I.

All attempts to restore any biofilm density and successfully complement the phenotype in strain J failed (Figure 4.19). As previously mentioned, failure to complement a certain phenotype could be attributed to many contributing factors including transcriptional and translation problems or expression levels of the Mip genes are far from what would be normal in the WT strain with genes under the control of the Ptac promoter and not their own chromosomal native promoter. If a gene is not expressed optimally, it may explain as to why some phenotypes are not returned.

4.5.3 Twitching motility assay using pME6032 complementation constructs

Biofilm formation partially relies on the expression of type IV pili and their mediated attachment to solid surfaces. Furthermore, the presence of type IV pili is required for the twitching movement. Given that both strains I and J exhibited reduced biofilm formation along with the loss of twitching motility, coupled with the data from the previous section (**4.5.2**) highlighting the varying success of complementation constructs to restore increase in biofilm formation, assessing whether the twitching motility had returned was necessary. It would be expected that a complementation vector successful in partially restoring biofilm formation would likely return the twitch motility phenotype, whereas complementation strains that failed to increase biofilm formation when compared to their uncomplemented parent strain would also fail to restore twitching motility.

Figure 4.20. Twitching motility plates using various pME6032 complementation vectors in strain I; Images comparing twitching motility of the WT strain, single knockout strains D and E against double knockout strain I and all complemented I strains. The WT and single knockout strains show twitching motility, the cloudy halo, this is the growth outwards along the solid plastic petri dish as a result of twitching, the other growth is on the surface of the agar from the initial inoculum. Some complemented I strains exhibit a restored twitching phenotype, while others do not.

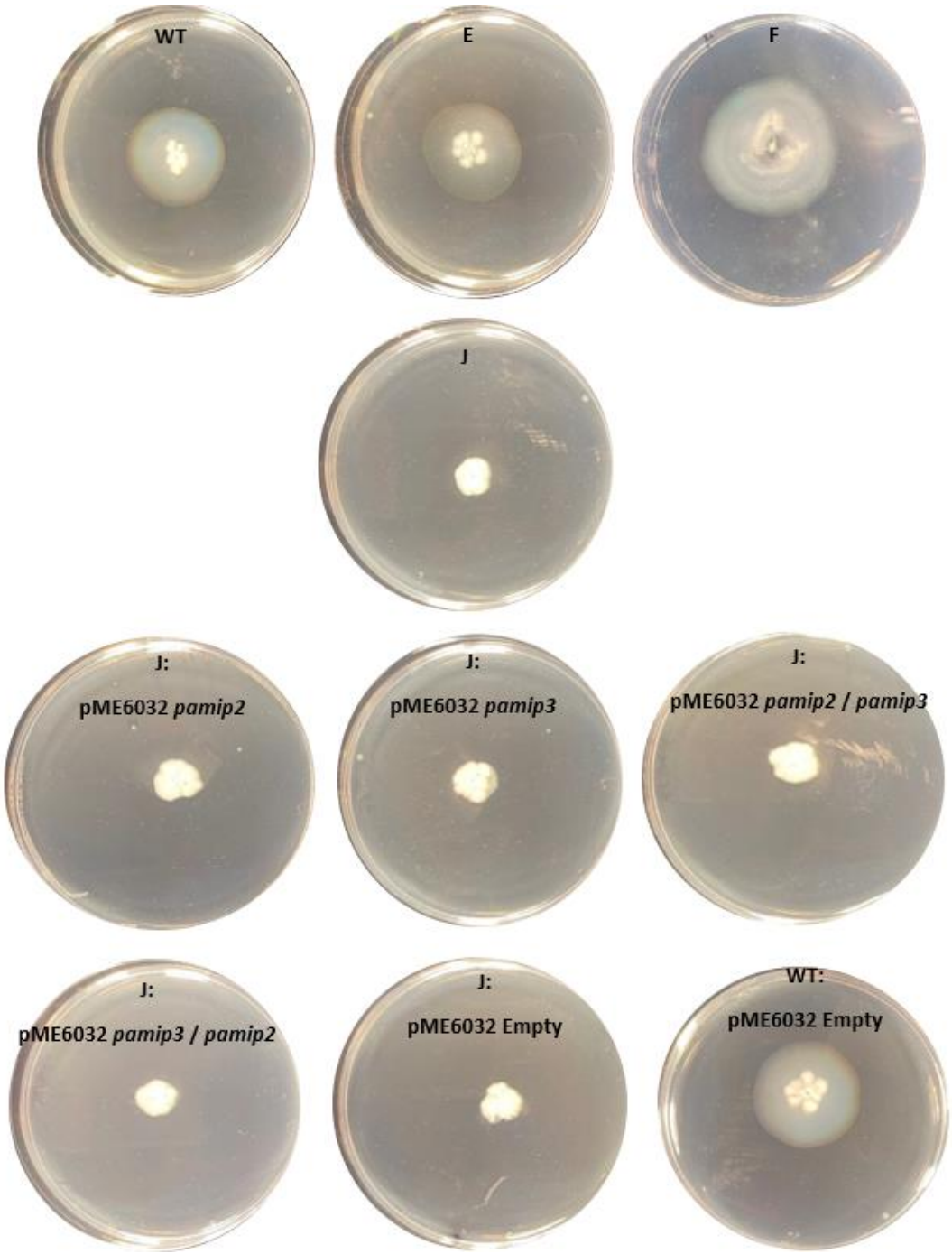


Figure 4.21. Twitching motility plates using various pME6032 complementation vectors in strain J; Images comparing twitching motility of the WT strain, single knockout strains E and F against double knockout strain J and all complemented I strains. The WT and single knockout strains show twitching motility, the cloudy halo, this is the growth outwards along the solid plastic petri dish as a result of twitching, the other growth is on the surface of the agar from the initial inoculum. All complemented J strains failed to restore the twitching phenotype.

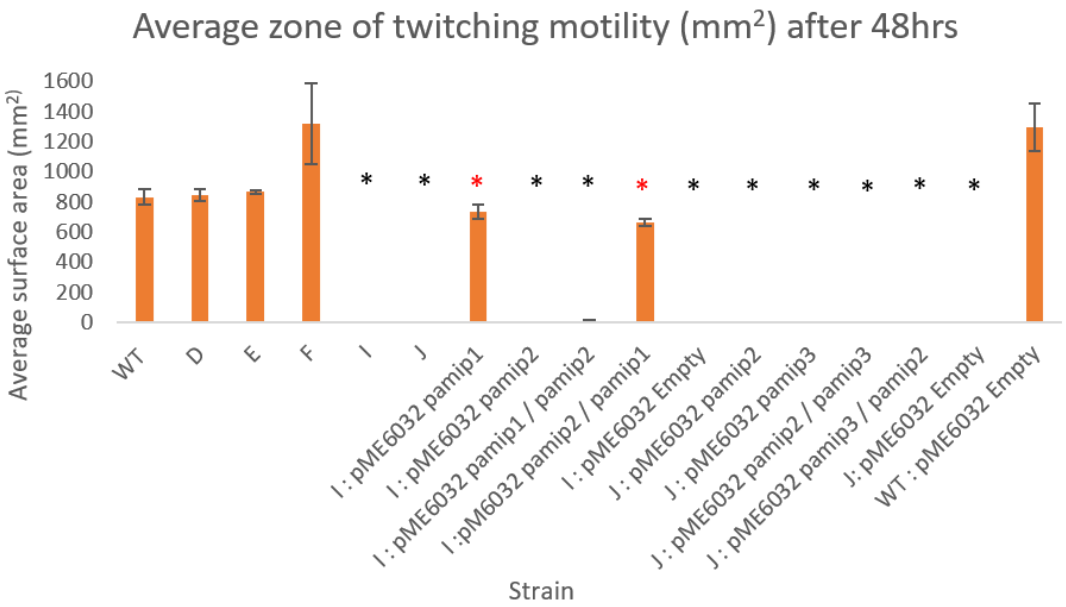


Figure 4.22. Bar chart plotting the average zone of twitching motility measured after 48h for PaMip knockout strains and strains I and J transformed with various pME6032 complementation vectors. Twitching motility zones were calculated using the equation; ½a x ½b x π (where 'a' is the longest diameter and 'b' is the shortest diameter). Standard error variance was calculated and plotted as seen by the error bars comparing the difference between the mean with the amount of scatter within the groups, error bars that overlap are not statistically significant. The black asterisks above columns indicate ($P<0.05$, n=2) (see methods). The red asterisks above columns indicate ($P<0.05$, n=2) where a complementation strain has been compared to its double knockout parent strain and exhibited significantly more biofilm formation than its parent, thus confirming complementation was successful.

The WT strain and mutant strains D, E and F all displayed twitching, strains I and J did not, consistent with our previous observations (**3.3.4**). The empty pME6032 construct introduced into the WT strain and mutant strains I and J did not affect their twitching motility. For strain I only complementation with *pamip1* had any effect, with twitching restored to almost WT or single mutant strain levels; introducing *pamip1* into the pME6032: *pamip2* construct (creating pME6032: *pamip2*/*pamip1*) gave an identical result, however, did not yield an increase to WT levels. Curiously, transformation of strain I with the pME6032: *pamip1*/*pamip2* construct, in which the order of genes is revered, failed to restore any level of twitching motility. All complementation constructs transformed into strain J failed to restore any twitching phenotype. Overall, these results are consistent with those for the biofilm assay above, thus, our hypothesis that complementation strains whose biofilm formation and density was partially restored would also exhibit a returned twitch motility phenotype was proven true.

4.5.4 Swarming motility assay using pME6032 complementation constructs

Following the return of twitching motility in some complemented strains, assessing the effect on swarming motility was examined. Swarming phenotypes are subjective and based on observation rather than any calculable measurement. As per results in **3.3.5** the WT strains D, E and F displayed the expected twitching motility, with strains I and J exhibiting abnormal phenotypes. Swarming by strain I is fairly uniform with an outward circular movement pattern, but extends to a greater distance than for the WT strain or the single Mip mutant strains (Figures 3.8 and 4.23). Strain J exhibits an unstructured outward movement with what appear to look like tendril projections, much like the swarming patterns associated with PA14 (Figures 3.8 and 4.24).

The WT strain and double mutant strain I and J were all transformed with empty pME6032 vector to control for plasmid effects and no effect was observed (Figures 4.23 and 4.24). Similar to the observations above for other pME6032 complementation assays, when strain I was complemented with pME6032: *pamip1* and pME6032: *pamip2*/*pamip1* constructs the phenotype returned to that of the WT (Figure 4.23). Introducing the remaining complementation vectors had no effect on the swarming phenotype for strain I.

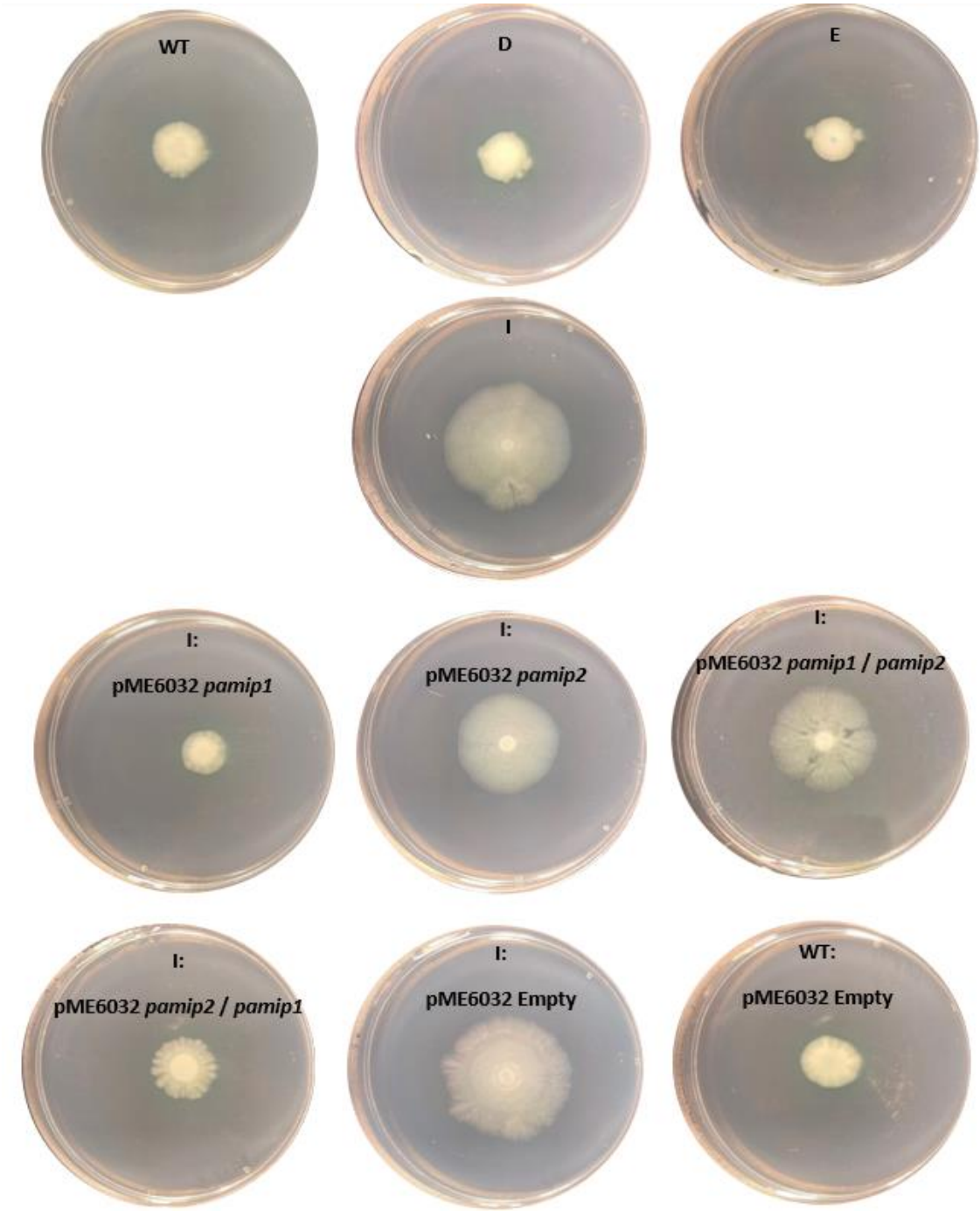


Figure 4.23. Swarming motility plates using various pME6032 complementation vectors in strain I; Images comparing swarming motility of the WT strain, single knockout strains D and E and the double knockout strain I, and all complemented versions of strain I. The WT and single knockout strains show the expected swarming motility compared to strain I, whose phenotype differs by moving outward in a larger, more uniformed manner. Some complemented versions of strain I exhibit a restored swarming phenotype, while others do not.

Strain J has the most striking phenotype with loss of the circular, uniform outward movement characteristic of the WT PA01 strain, and instead has the appearance of tendril projections in a flower-shaped pattern, a phenotype more associated with PA14. Introducing the pME6032: *pamip2* construct appears to slightly reduce the tendril projections and flower shaped pattern, whereas the pME6032: *pamip3* construct has no effect. Furthermore, when PaMip3 is inserted into the pME6032: *pamip2* vector, the partial restoration of the swarming phenotype previously was lost. However, when *pamip2* is introduced to the pME6032: *pamip3* construct, which had no effect on phenotype, the tendril projections appear to have disappeared along with the flower-like pattern, with a more circular, uniform appearance returning which is quite similar to the WT.

Once again, as seen in the haemolysis assay, introducing pME6032: *pamip2* into strain I did not restore and phenotype, and in fact acted almost as an antagonist when added to the pME6032: *pamip1* vector. However, introducing the pME6032: *pamip2* vector into strain J, in part, returned some phenotypic resemblance to that of the WT. As to why this occurs is difficult to say with a variety of factors contributing, as previously mentioned, the most likely cause will be at a transcriptional or translational level. Overall, the swarming phenotype was complemented to some degree in both stain I and J using certain vectors.

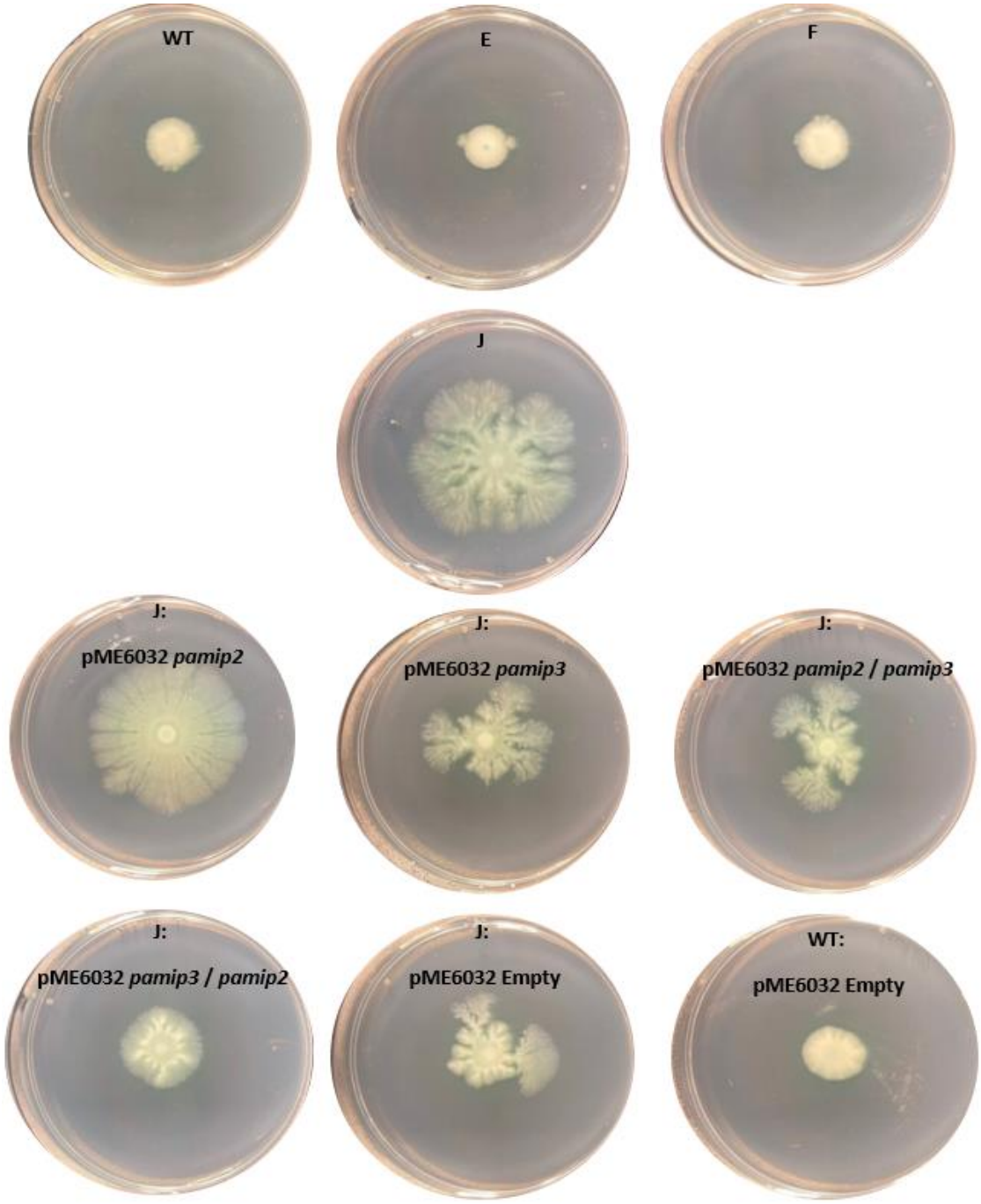


Figure 4.24. Swarming motility plates using various pME6032 complementation vectors in strain J; Images comparing swarming motility of the WT strain, single knockout strains E and F against double knockout strain J and all complemented I strains. The WT and single knockout strains show ordinary swarming motility compared to strain J, whose phenotype is drastically different moving outward in a pattern resembling a flower head and a phenotype similar to the swarming motility of the PA14 *P. aeruginosa* strain.

4.5.5. Assessing the return of killing of *Galleria mellonella* using strains complemented with pME6032 Mip gene constructs

With complementation constructs varying in their success at returning plate-based phenotypes, it was prudent to assess how they performed *in vivo G. mellonella* model. To infect and kill any host, a strain must optimally express a variety of virulence mechanisms. Given the above results, one would hypothesise only strains whose phenotypes were in some part restored through complementation would exhibit an increased ability to kill *G. mellonella* larvae. Conversely, strains whose phenotypes were not restored might be able to demonstrate an improved ability to kill *G. mellonella* larvae.

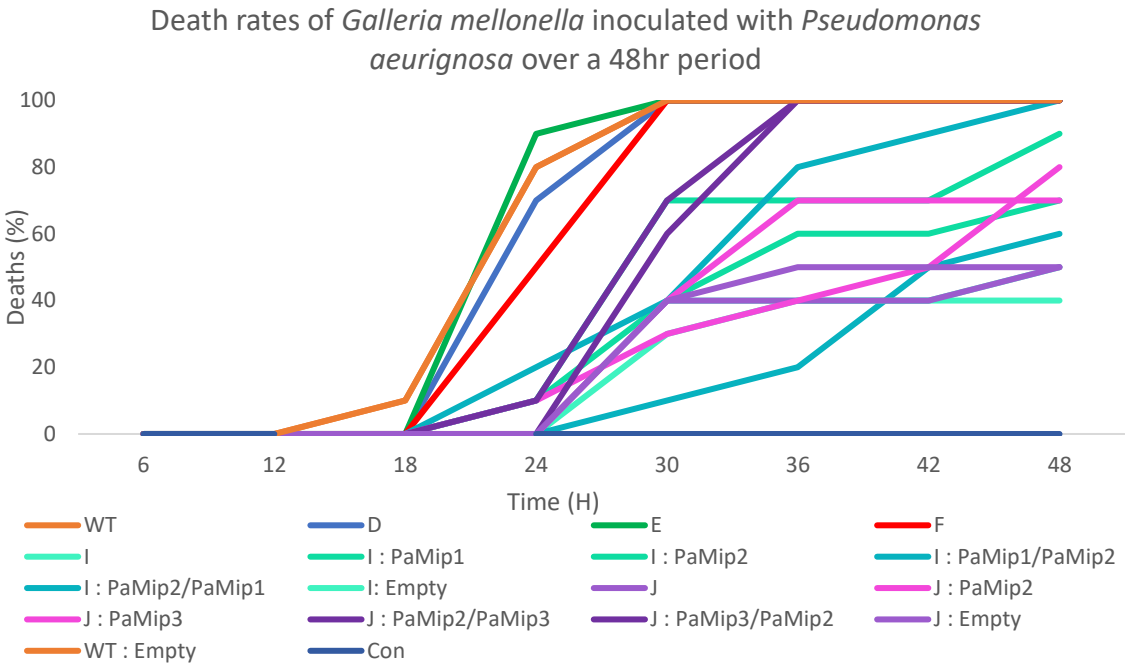


Figure 4.25. Line graph plotting deaths of *G. mellonella* inoculated with PaMip knockout strains and strains I and J transformed with various pME6032 complementation vectors; Percentage of deaths in *G. mellonella* wax moth larvae for all mutant and complementation strains over a 48 h period with recordings taken at 6 h intervals. This experiment was performed twice with results pooled. Each time the experiment was performed 10 larvae were used per strain. Final deaths were recorded as a percentage of deaths. Single Mip complementation constructs across strains I and J have been given the same colour, while double complementation constructs have been given a darker shade, this allows a comparison of how single constructs performed vs the double constructs. The single PaMip constructs for strain I improve killing of *G. mellonella* compared to uncomplemented strain I. The double PaMip constructs completely restore killing of *G. mellonella*. The same trend is seen for the different complementation constructs in strain J, single constructs improve a strains ability to kill the larvae, whereas the double constructs completely restore a strains ability to kill the larvae.

Results for the WT and various deletion strains mimicked those presented in Chapter **3.5.1**, with the WT strain performing as expected and begins killing *G. mellonella* larvae at around the 18 h timepoint reaching 100% deaths between 24-30 h timepoints. Strain E was able kill the larvae at a similar rate to the WT, beginning at 18 h, reaching 100% larval deaths after 30 h. Both strains D and F exhibited a delayed killing effect with both recording their first larval death 6h after the WT at 24 h; however, they still reached 100% of larvae deceased after 30 h like the WT and strain E. Strains I and J, are both significantly attenuated in *G. mellonella* larvae killing. Strain I only began to kill larvae after around 30 h, a 12 h delay when compared to the WT strain. Final death rates after 48 h for strain I only reach 30% compared to the WT of 100%. Strain J followed a similar pattern to strain I, with deaths of larvae presenting around 30 h post inoculation and never reaching 100%. Introducing the empty pME6032 vector into the WT and strains I and J had no effect on virulence. The control larvae group inoculated with sterile PBS resulted in zero deaths, suggesting good technique when injecting the larvae providing assurances that it was unlikely any deaths were as a result of puncture wounds. This experiment was carried out twice, at different times, and for each experiment 10 larvae were used per strain.

Introducing the pME6032: *pamip1* vector into strain I, previously successful in restoring a variety of *in vitro* phenotypes, led to a partial restoration of virulence with an increase of the death rate of 30% to 70% after 48 h. Furthermore, killing of *G. mellonella* larvae was brought forward by 6 h, now beginning at 24 h compared to the uncomplemented strain I whose first recorded death was after 30 h. Unexpectedly, for the first time, introducing pME6032: *pamip2* into strain I restored a phenotype similar to that of the WT with killing of *G. mellonella* starting after 18-24 h, and displaying a death rate of 90% dead larvae after 48 h (from 30% for strain I). Both of the double Mip complementation vectors also restored virulence from 30% deaths to 60% and 100% deaths at 48 h for pME6032: *pamip1*/*pamip2* and pME6032: *pamip2*/*pamip1*, respectively. Unlike previous assays, the introduction of *pamip2* into the pME6032: *pamip1* construct did not hinder any phenotypic restoration, but in fact accelerated it with pME6032: *pamip1* alone reaching 70% dead larvae which then further increased to 100% dead larvae following the introduction of *pamip2* into the construct. Furthermore, this construct restored the timepoint at which the first larvae was recorded dead to that of the WT.

Introducing pME6032: *pamip2* construct into strain J partially restored the killing of larvae to that of the WT strain with an increase in percentage of larvae killed from 50% to 70% at 48 h post inoculation although at a fairly steady rate with no drastic increases in deaths between each timepoint. Similar results were observed in strain J complemented with pME6032: *pamip3* with an increase in the percentage of larvae killed from 50% to 80% at 48 h post inoculation. Although both single constructs did not return 100% killing of larvae after 48 h, there is a

significant increase when compared to the uncomplemented parent strain J. The virulence of strain J was returned to that of the WT when either of the vectors containing both genes were introduced; 100% of larvae were dead after 36 h, the quickest of any complementation constructs.

The WT strain housing empty pME6032 vector follows the same death trajectory as the untransformed WT strain, with death timepoints the same and reaching 100% deaths in the same time period, thus the pME6032 vector has no effect on WT PA01 ability to kill *G. mellonella* larvae. Furthermore, empty pME6032 vector transformed into strains I and J did not increase their ability to kill the larvae, with both strains I and J housing pME6032 exhibiting similar final death counts to their untransformed parent strains. The control larvae group inoculated with sterile PBS resulted in zero deaths, suggesting good technique when injecting the larvae proving assurances that it was unlikely any deaths were as a result of puncture wounds. Once more, this experiment was carried out twice, each time 10 larvae were used per strain to try and rule out any deaths from initial inoculation trauma.

These results in an *in vivo* system, which provides a holistic readout on the virulence mechanisms of PA01strongly suggest that the various PaMip gene deletions can be complemented in full as introducing two PaMip genes back into double knockouts completely restored their ability to kill *G. mellonella* larvae when compared to their uncomplemented parent strain, improving both their effectiveness and speed at which the larvae were killed during the time course.

4.6 Discussion

The results described above demonstrate varying success of the complementation constructs in their ability to fully restore phenotypes of the double PaMip knockout strains I and J. The results obtained here in assessing which PaMip gene complementation constructs were able to restore any phenotype, in part or fully, were fairly representative with all plate-based virulence assays exhibiting a familiar pattern. This pattern highlighted that some single PaMip constructs were more successful than others in complementation of phenotypes in double knockout strains I and J, somewhat restoring their phenotypes back to those of their single knockout strain parents (D, E and F). For strain I, transforming it with the single PaMip construct pME6032: *pamip1* fully restored many phenotypes. However, when transformed with the single PaMip construct pME6032: *pamip2*, no phenotypes were returned for *in vitro* assays. When *pamip1* was introduced into the pME6032: *pamip2* construct (making it pME6032: *pamip2*/*pamip1*) and transformed into strain I, phenotypes were restored. Whereas when *pamip2* was introduced into the pME6032: *pamip1* construct (making it pME6032: *pamip1*/*pamip2*), complementation

of phenotypes was in fact dampened from what it was prior to the introduction of PaMip2 into the construct.

A similar theme occurred for complementation in strain J. Transforming strain J with the single Mip construct pME6032: *pamip2* partially restored some phenotypes. However, when transformed with the single Mip construct pME6032: *pamip3*, no phenotypes returned. When *pamip2* was introduced into the pME6032: *pamip3* construct (making it pME6032: *pamip3*/*pamip2*) and transformed into strain J, some phenotypes were partially restored. Whereas when *pamip3* was introduced into the pME6032: *pamip2* construct (making it pME6032: *pamip2*/*pamip3*), complementation of phenotypes was abolished from what it was prior to the introduction of *pamip3* into the construct. This could have been a construct issue by adding the second gene in causing an effect on the transcription of both. Furthermore, this could suggest a gene dosage and an expression titre response in that the overexpression of some genes, transcribed to levels higher than the WT, could lead to dysregulation and cause a strain to exhibit similar phenotypes as the original knockouts. Equally, this same effect would be seen if the gene is transcribed at lower levels compared to the WT. This could have been checked by RT-qPCR to show levels of expression compared to the WT.

The results described here suggest that the complementation of certain PaMip genes may promote certain biological activities more than others. As seen for *pamip1*, when reintroduced into strain I either as a single complementation or ligated into the pME6032: *pamip2* construct, both haemolytic and proteolytic activity are restored to levels comparable to those of the WT strain. Haemolytic activity as previously mentioned is a result of the expression of a phospholipase C (PLC-H), while proteolytic activity can be attributed to several protease enzymes. Analysing the RNA-seq data in the mutant I background for *plch* gene (*pa0844*) and several protease producing genes, levels of expression when compared to the WT were unchanged, suggesting these gene are expressed normally and the loss of haemolysis and proteolysis is not the result of any transcriptional changes in these genes in the double mutant I. Re-introducing the *pamip1* gene returns both haemolytic and proteolytic activity in strain I, and given the nature of FKBPs, it is possible that PaMip1 acts in some part as a chaperone required for the correct folding of PLC-H phospholipase and proteases following their transcription.

Given PaMip2 fails to complement haemolytic activity, but does restore proteolytic activity, PaMip2 may somehow be more involved with the chaperoning of proteolytic enzymes. A secondary explanation could be attributed to poor expression levels given that in the pME6032 vector genes are under the control of the Ptac promoter and not its own chromosomal native promoter. If the *pamip2* gene is not expressed optimally it may explain as to why some

phenotypes are not returned. When *pamip2* was introduced into the pME6032: *pamip1* complementation construct, restoration of haemolysis is damped, however this could be due to some transcriptional and translational issues inserting a secondary gene into the vector under the control of a non-native promoter. A western blot analysis could be used in future to ensure the correct expression of both genes in the construct.

The results presented for strain J further suggest that PaMips of PA01 may exert their own independent role in virulence, rather than complete redundancy between one another. However, when a certain combination of PaMips were deleted, a more broad, pleiotropic attenuation in virulence is observed. Furthermore, the failure of the pME6032: *pamip3* complementation construct to return any observable phenotype in plate-based assays gives further credence that the *pamip3* gene and its product are more likely located in the outer membrane acting as a surface antigen / moonlighting protein where it is able to exert its virulence affects, much like the LgMip, rather than act as a traditional cytoplasmic FKBP involved in protein folding and as a chaperone. This is further evidenced in that the *G. mellonella* infection model reintroducing pME6032: *pamip3*, for the first time, partly restored a phenotype, with the death of a larvae occurring at the same time point as the WT, rather than the typical delay in killing strain J usually exhibited.

When assessing the complementation strains in an *in vivo* model, the pattern described above in that certain complementation constructs were continually successful in restoring plate-based virulence phenotypes of strains I and J, while others failed, was incorrect (Table 4.2). As mentioned, the hypothesis that given strains I and J, demonstrated clear, attenuated phenotypes in all plate-based virulence assays, these virulence mechanisms would also remain attenuated in an *in vivo* model which would result in an attenuation in *G. mellonella* killing. Complementation constructs able to restore phenotypes of I and J in plate-based virulence assays to those of the WT would theoretically also return a strains *in vivo* virulence and no longer be attenuated in the killing of *G. mellonella*. Conversely, complementation constructs that failed to restore plate-based virulence phenotypes would also fail to return a strains *in vivo* virulence and remain attenuated in the killing of *G. mellonella*. In fact, what was observed was all complementation constructs housing either a single or double PaMip gene somewhat returned the killing of *G. mellonella* when compared to the uncompleted parent strain.

So why did some complementation constructs fail to return plate-based virulence phenotypes of strains I and J, while *in vivo* these same strains housing the same constructs returned the *in vivo* killing of *G. mellonella*? The answer is most likely attributed to the pleotropic nature the PaMips play in the virulence of PA01. Using plate-based virulence assays provided an insight and allowed us to unpick as to what virulence mechanisms may be hindered as a consequence of

PaMip-gene knockouts, and subsequently which of these virulence mechanisms may not be optimally performing during an *in vivo* infection assay, contributing towards a strain's virulence attenuation in the *G. mellonella* infection model. However, these plate-based assays were specific in assessing a particular virulence phenotype in each assay rather than a holistic virulence overview. These specific phenotypes rely on many intertwining regulatory systems that turn on and switch off specific virulence mechanisms at precise times and environmental queues, which to date are still not fully understood. If genes involved in these regulatory systems, for example *pamip1-3* genes, and at a transcriptional or translational level, are inhibited or deleted, particular virulence mechanisms will remain switched off or not optimally expressed resulting in the loss of a given phenotype. Given the complexities involved in both transcriptional and translational regulation of virulence associated mechanisms, if a given gene and its product involved in these regulatory pathways are not expressed at the optimal level, a given phenotype will be lost. This may explain the failure of some complementation constructs in resorting a only specific phenotype such as haemolytic activity or twitching motility.

The complementation constructs that failed in restoring a specific phenotype of a strain to WT levels in plate-based assays may also have been the result of the construct expressing its housed PaMip gene to either suboptimal or excessive levels. Given the intricate nature of how PA01 regulates different virulence mechanisms and their expression, anything other than the native level and timing of expression could fail to restore each individual phenotype in plate-based assays. This is seen for some complementation constructs that were unable to complement a single specific phenotype but were able to restore the *in vivo* killing of *G. mellonella*. A potential explanation for these results may be that although complementation constructs may not have returned optimal expression of a PaMip gene(s), these genes were at least returned and expressed in some capacity, returning several overarching, interconnected virulence processes, not just assessing one specific phenotypic trait as with plate-based assays. A complementation experiment that would have solidified this hypothesis would have been to use the macrophage infectivity assay, however given time constraints this experiment was unable to be performed in time for the writing of this thesis.

Any concerns around the missing 3275 bp pre-prophage region (and duplication of prophage structural genes) of strain I were resolved through rigorous investigation by starting by recreating strains D, E and I, and verifying their correct nature through whole-genome sequencing and PCR analysis of any progeny strains. These strains were then tested in phenotypic assays and all presented identical phenotypes to the original version produced by Silke Alt. Why this region was unstable overtime in both versions of strain I remains unknown. Future work examining this phenomenon would prove interesting, especially if the entire

prophage cluster could be deleted cleanly and the resulting strain assessed for fitness and virulence.

In future experiments it would be illuminating to perform Western Blot analysis for strains housing the different complementation constructs. This would show if the PaMip proteins were produced, if they are not it could explain some of the negative complementation results for *in vitro* assays. Equally, this could also have further elucidated whether the expression levels of PaMips contributed to the success or failure of the complementation constructs. If the complementation strategies and the constructs used were expressing the PaMips, however failed to restore any phenotypes, it would have begun to evidence whether virulence effects of PaMips were dependant on their expression levels, with levels differing from those of the WT, either too low or high, hindering successful phenotypic complementation.

Future work that will be interesting to carry out will be to conclusively obtain SEM images of type IV pili in complementation strains I and J that saw any return of twitching motility and biofilm formation, both of which are type IV pili associated phenotypes. Given the SEM images obtained in chapter 3 for strains I and J both lacking any observable type IV pili, coupled with the lack of type IV pili associated phenotypes, complemented strains that saw a return of any type IV pili associated phenotypes would theoretically possess imageable type IV pili. Further future work involving strains housing pME6032 complementation constructs will look at whether these complementation construct restore the ability of strains F, I and J to optimally enter, colonise and replicate within macrophages to levels comparable of the WT strain. Given these constructs were successful at restoring each strain's ability to kill *G. mellonella*, it would be assumed they too would partly or fully restore phenotypes to that of the WT strain in a macrophage infection assay.

Summary of complementation construct success in restoring phenotype for each strain in each phenotypic assay (omplete restoration = ✓) (No restoration = X) (Partial restoration = ~)						
	G. mellonella	**Haemolysis**	**Protease**	**Biofilm**	**Twitching**	**Swarming**
I	X	X	X	X	X	X
I : pME6032 *pamip1*	✓	✓	✓	✓	~	✓
I : pME6032 *pamip2*	✓	X	✓	X	X	X
I : pME6032 *pamip1 / pamip2*	✓	✓	✓	X	X	X
I : pME6032 *pamip2 / pamip1*	✓	✓	✓	✓	✓	✓
I : pME6032 Empty	X	X	X	X	X	X
J	X	X	X	X	X	X
J : pME6032 *pamip2*	✓	~	✓	X	X	✓
J : pME6032 *pamip3*	✓	X	✓	X	X	X
J : pME6032 *pamip2 / pamip3*	✓	X	✓	X	X	X
J : pME6032 *pamip3 / pamip2*	✓	✓	✓	X	X	✓
J : pME6032 Empty	X	X	X	X	X	X

Table 4.2. A summary of the work described in this chapter in which complementation strains were successful in returning phenotypes to strains I and J. The table highlights a pattern in which complementation constructs, for the most part, were successful and others that were not.

Chapter 5:

Assessing the role the Type Six Secretion System (T6SS) hcp-island III (HSI-III) plays in PA01 virulence

Chapter 5. Assessing the role the Type Six Secretion System (T6SS) hcp-island III (HSI-III) plays in PA01 virulence

5.1 Introduction

As presented in Chapter 3 (**3.7.3**) analysis of the RNA seq datasets, there was a significant downregulation of a cluster of genes encoding an understudied, and potentially important virulence determinant, the Type Six Secretion System (T6SS) hcp-island III (HSI-III) – hereafter this will be named T6SS HSI-III. PA01 encodes three different T6SSs which are annotated HSI-I, HSI-II and HSI-III, with the majority of published literature only reporting work on the HSI-I and HSI-II islands. Furthermore, as described in Chapter 3 (**3.7.3**), the T6SS HSI-I and HSI-II islands have both been associated with delivering toxins and effectors to other prokaryotic cells in a bid to outcompete them in a given environment (Figure 5.1). Furthermore, only the Type Three Secretion System (T3SS) had been described as a key system by which PA01 delivers effectors and toxins directly into eukaryotic cells. Out of the 14 genes comprising the Type Six Secretion System (T6SS) Island III (HSI-III), 11 displayed a significant fold change in expression when compared to the wildtype (WT).

Once expressed, T6SS is assembled and acts as a puncturing mechanism by which PA01 can deliver toxins and effectors to neighbouring foreign cells. Sitting atop the T6SS structure is the spike protein, VgrG, which punctures the target cell membrane and allows the delivery of toxins and effectors that are loaded onto its tip [1, 260]. This process requires a large amount of energy which is provided by the AAA^+ (**A**TPases **A**ssociated with diverse cellular **A**ctivities) ATPase, ClpV [261]. Each time the system fires it must be disassembled before being reassembled again to fire once more, effectively ‘reloading’.

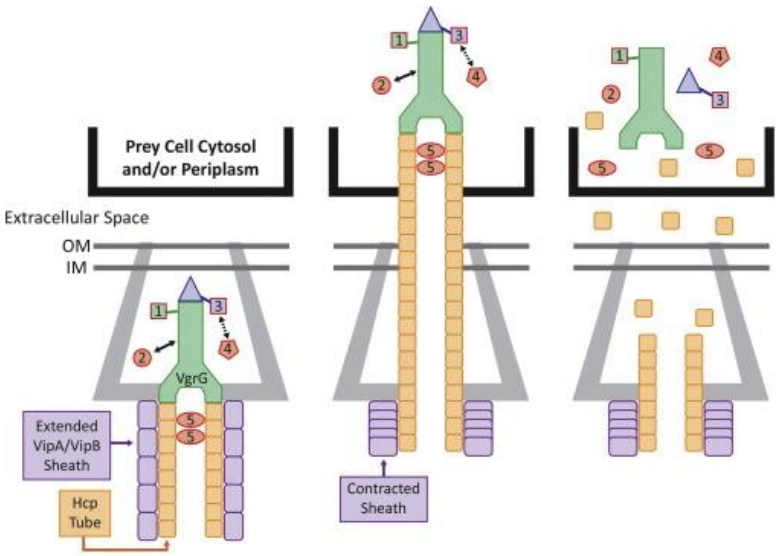


Figure 5.1. Diagrammatic representation of the extension and contraction of the T6SS taken from Ho *et al* [1]. Effectors and toxins to be delivered to the target (2,4) are loaded to the tip of VgrG or within the end of Hcp (5). Sheath contraction (VipA/VipB) leads to extension of the Hcp nanotube and the delivery of VgrG, along with any associated effectors, into the target periplasm. Once delivered, Hcp is no longer needed so is disassembled. VipA/VipB remain contracted and require disassembling by the action of ClpV.

The expression of the T6SS is associated with the transition to the sessile / chronic infectious lifestyle that is mediated by the GacS / RetS two-component system as described in the introduction (**1.5.1**). With both PA01 strains I and J are attenuated in phenotypes associated with the sessile / chronic lifestyle (biofilm formation significantly down, type IV pili expression non-existent), in addition with this significant T6SS HSI-III downregulation identified in the RNA seq data; this potentially suggests that PaMips may partly exert their role in virulence by interacting with genes and proteins of regulation pathways such as the GacA/RetS two component-system or QS system, which regulate virulence factors. In the absence of the PaMips, several virulence mechanisms may not be expressed, or may not be correctly assembled leading to the overall reduction in virulence observed for mutant strains I and J (Figure 5.2).

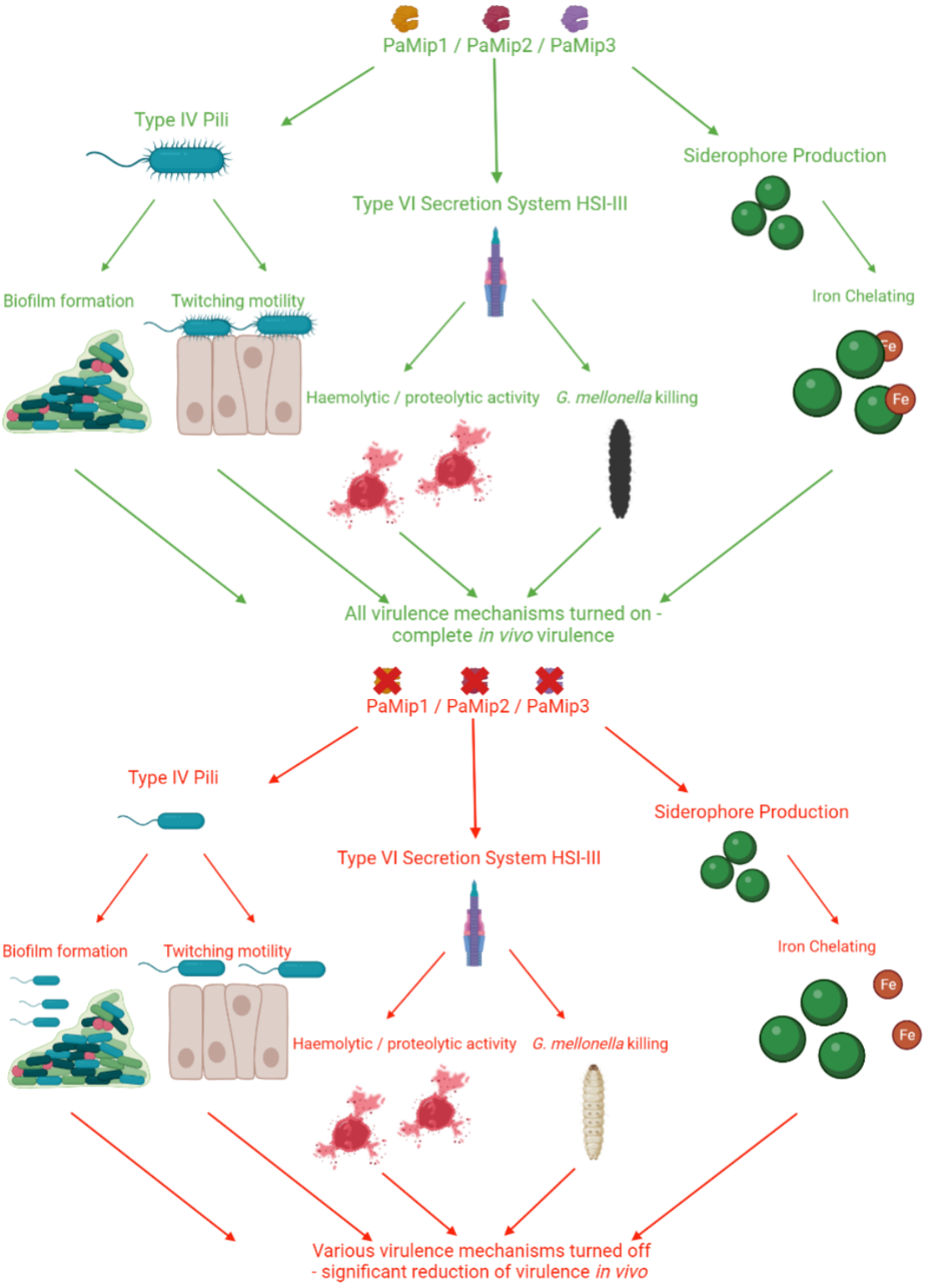


Figure 5.2. Flow diagram demonstrating the proposed pleiotropic virulence effects associated with PaMips. Panel 1) the flow diagram demonstrates that when PaMip1 / PaMip2 /PaMip3 are present all of the virulence mechanisms pictured are expressed and functioning optimally; this is indicated by green arrows. Expression of all these mechanisms contributes to complete virulence of PA01 *in vivo* evidenced by killing of *G. mellonella* and colonisation of macrophages described in Chapter 3. Panel 2) demonstrates that in PaMip mutant strains, several virulence processes are either not expressed or their components are not assembled correctly; indicated by red arrows. This results in a reduction of virulence and reduced *in vivo* infectivity evidenced by data in the *G. mellonella* and macrophages assays described in Chapter 3.

As noted already, the role of PaMip's is pleiotropic and it is thus unlikely that downregulation of the T6SS HSI-III is responsible for the loss of all the phenotypes described for strains I and J in Chapter 3, however, it may be involved in the loss of haemolysis and proteolysis activities along with the attenuation of killing *G. mellonella*. For example, given that secretion systems are responsible for the delivery of toxins and effectors, haemolysins and proteases may represent such effector proteins delivered by the T6SS HSI-III. Therefore, if the T6SS HSI-III is a virulence determinant associated with haemolysis and proteolysis, when genes encoding components of the T6SS HSI-III are deleted in PA01, similar phenotypes to those of strains I and J should be observed in selected assays. However, it seems unlikely that the T6SS HSI-III is directly involved in biofilm formation, so when assayed, a T6SS HSI-III mutant should still exhibit biofilm formation. Therefore, T6SS HSI-III knockouts should affect certain phenotypes in plate-based and *in vivo* virulence assays, but not all (Figure 5.3).

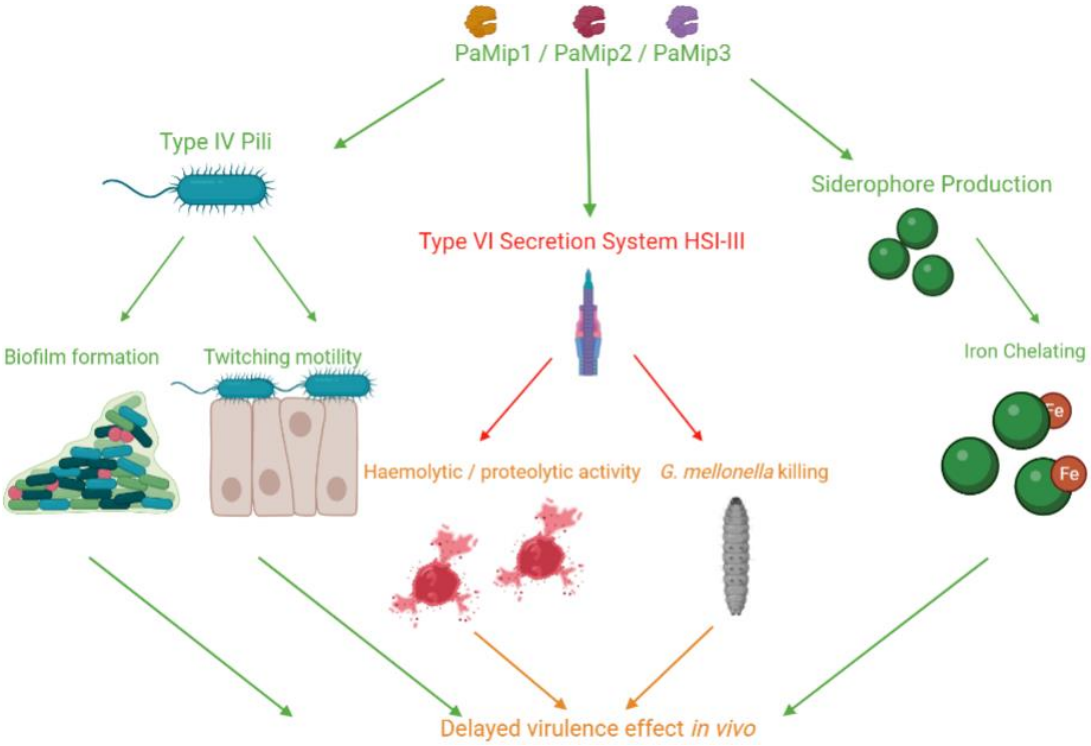


Figure 5.3. Flow diagram demonstrating the proposed pleiotropic virulence effects of PaMips and the effect of a T6SS HSI-III knockout. This demonstrates how when PaMips are expressed in PA01, they are able to play a role in virulence by aiding in the type IV pili expression and the production of siderophores. Here the PaMips also encourage the expression of the T6SS as would be expected in our hypothesis. However, in this strain components of the T6SS HSI-III are deleted which results in attenuation of haemolysis and proteolytic activity which ultimately contributes to a delayed *in vivo* virulence affect. Other virulence processes in the T6SS HSI-III mutant strains are theoretically still expressed so overall virulence is not lost, and overall virulence is reduced. This would lead to a delayed virulence effect *in vivo* and intermediate phenotypes between the WT and PaMip mutants, unlike the complete loss of virulence as was in PaMip mutant strains I and J.

Based on the results in Chapter 3, it is my hypothesis that all T6SS islands interact with prokaryotic hosts, but that the T6SS HSI-III also plays a role in delivering toxins and effectors into eukaryotic host cells, like the T3SS. The aim of this chapter was to investigate the role of the T6SS HSI-III in virulence. To achieve this, genes and their products in each HSI island were selected for deletion; the AAA^+ ATPases ClpV1, ClpV2 and ClpV3, and the VgrG1, VgrG and VgrG3 spike proteins were all chosen as each HSI encodes their own copy of these genes which means the effects of individual gene deletions can directly be compared. These genes all share sequence conservation, with *clpv* gene products sharing approximately 60% amino acid sequence similarity, and the *vgrg* gene products sharing approximately 53% amino acid sequence similarity. These two gene families were selected given their importance to the overall function of each T6SS island.

As mentioned above, ClpV is essential and is required for the assembly and disassembly of the entire T6SS (Figure 5.3) [1]. VgrG was chosen as it is required for the puncturing of target cells following forming a complex with proline-alanine-alanine-arginine repeats (PAAR repeat-containing protein).This complex forms a rigid needle structure enabling the puncturing of target cell membranes and it houses effector proteins such as Tse1-3 [262]. The various *clpV* and *vgrG* genes were deleted using the same protocol performed for the knockout of PaMip genes using the pTS1 plasmid, and creates a truncated version of each gene rather than a start to stop codon deletion; this was designed in multiples of 3bp to ensure the reading frame remained intact. The resultant knockout strains were checked by colony PCR and sequenced (Table 5.1).

Genotype	Strain name
PA01 Δ*pa2371* (Clpv3)	Δ*clpv3*
PA01 Δ*pa2373* (VgrG3)	Δ*vgrg3*
PA01 Δ*pa1662* (ClpV2)	Δ*clpv2*
PA01 Δ*pa1511* (VgrG2)	Δ*vgrg2*
PA01 Δ*pa0090* (Clpv1)	Δ*clpv1*
PA01 Δ*pa0091* (VgrG1)	Δ*vgrg1*
PA01 Δ*pa2371* (Clpv3) : pME6032 ClpV3	Δ*clpv3* : *clpv3*
PA01 Δ*pa2371* (Clpv3) : pME6032 Empty	Δ*clpv3* : pME6032 Empty

Table 5.1. List of T6SS mutant strains created to test the hypothesis presented in this introduction. Single knockout strains possessing an in-frame gene deletion of either *clpv* or *vgrg* in each T6SS island from HSI I to III. One strain, ClpV3, was complemented with the full-length gene as well as transformed with empty pME6032 vector as a control.

With the creation of T6SS knockout strains (confirmed by colony PCR and sequencing), if our hypothesis is correct, upon repeating the assays described in Chapter 3, strains containing deletions in HSI-III proteins ClpV3 and VgrG3 should demonstrate a loss of haemolytic and proteolytic activity with an overall delay *in vivo* killing of *G. mellonella*. However, these strains should theoretically also still be able to form biofilms and perform twitching and swarming motilities. Furthermore, any deletions in the HSI-I proteins ClpV1 / VgrG1 and HSI-II proteins ClpV2 / VgrG2 should have no effect on these phenotypes as literature reports describe them as primarily being responsible for prokaryotic cell interaction, not eukaryotic, and in these HSI-I / HSI-II mutant strains, the HSI-III should remain functional. To test this hypothesis, the relevant T6SS mutant strains (Table 5.1) were created and subjected to phenotypic assays, the results of which will be discussed in this chapter. Given time constraints in the lab, only one strain, Δ*clpv3*, was able to be complemented and will be included in the below results. Furthermore, the Δ*clpv3* strain was sent for RNA seq to aid in elucidating any observed phenotypes. All raw data for the assays described below can be found in the Appendix.

5.2 *In vitro* plate based virulence assays using T6SS HSI I-III mutant strains

In vitro assays are experimental biology research tools that allow the assessment of cell health after experimental treatment or manipulation. Using *in vitro* assays to assess virulence is useful to help understand which infectious processes may have been affected by gene knockouts and the potential pleotropic affects they may have. Furthermore, they can elucidate which virulence mechanisms may be impaired when in a host, resulting in loss of infectivity and whether moving to an *in vivo* model is worthwhile. Prior to *in vivo* assays, having an idea of what virulence processes are attenuated *in vitro*, will help us decide whether to undertake *in vivo* studies and if so, aid in interpreting the *in vivo* data.

5.2.1 Haemolysis assay using T6SS HSI I-III mutant strains

Given both HSI-I and HSI-II islands are responsible for the delivery of toxins and effectors directly into target prokaryotic cells, if our hypothesis presented above is correct, it could be assumed that the T6SS HSI-III may be involved in delivering effectors such as PLC-H, which is responsible for haemolytic activity of PA01 [222]. As described above, it is hypothesised for strains possessing knockouts in genes of the T6SS HSI-III would exhibit phenotypes similar to those seen for PaMip knockout strains I and J, while the corresponding genes in islands HSI-I and HSI-II should phenocopy the WT. To test this, strains with deletions in key components of each T6SS islands were examined in haemolysis assays (Table 5.1 and Figure 5.4). These assays were performed in an identical manner to those described in Chapters 3 and 4.

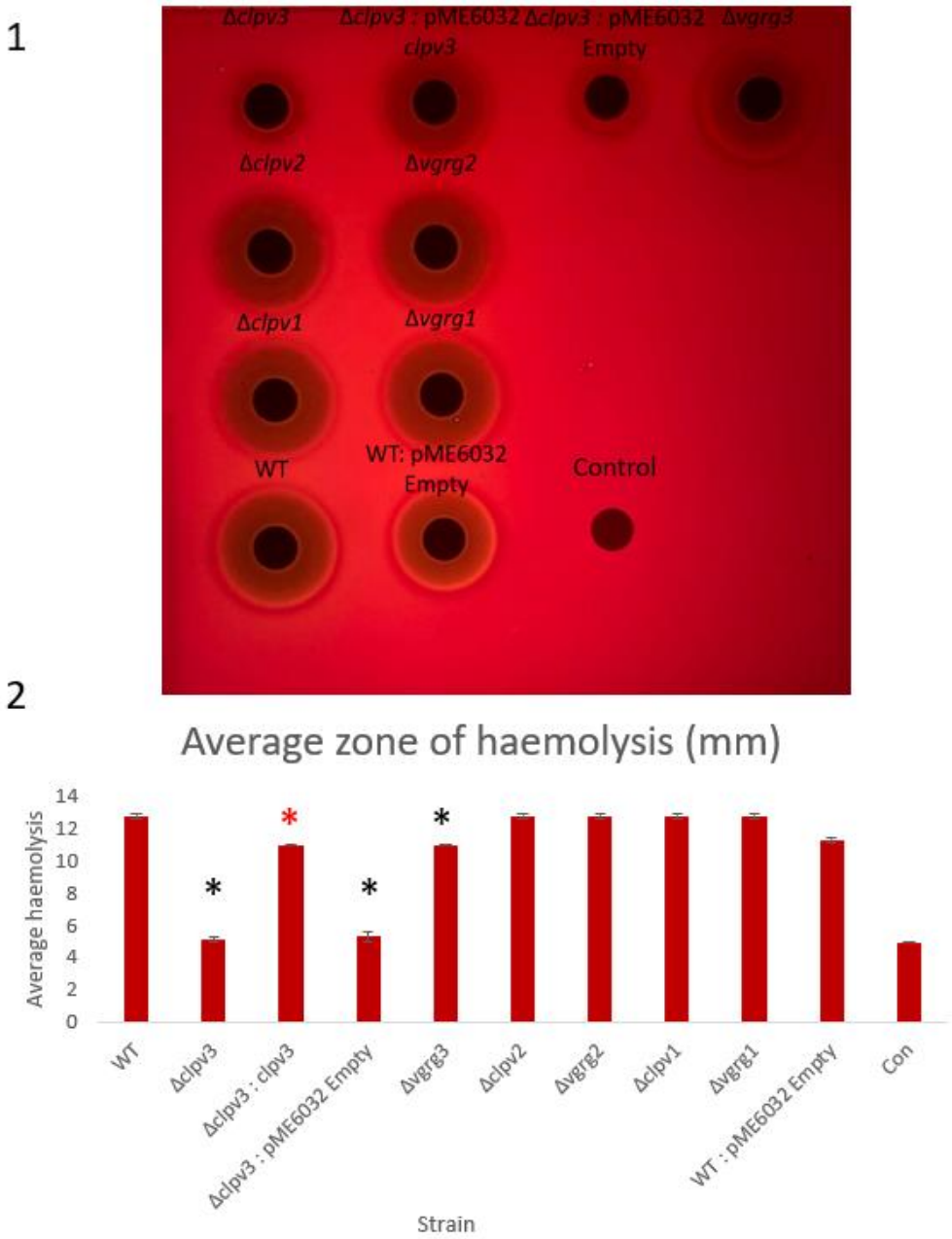


Figure 5.4. Haemolytic activity of T6SS deletion mutants. Panel 1 shows blood agar plates spotted with deletion mutants of T6SS components for each T6SS island. Haemolysis is indicated by clear inhibition zones where blood cells have been broken down. The WT has a large activity zone with mutants in the T6SS HSI I-II clusters displaying comparable activity. The control using sterile water exhibited no haemolytic activity as expected. The ΔClpV3 strain has a clear, reduced halo with less haemolytic activity which was returned upon complementation. The bar chart in panel 2 shows the average zones of haemolytic activity measured from this and replica plates. Along the X axis are the T6SS knockout strains as described in Table 5.1. Standard error variance was calculated and plotted as seen by the error bars comparing the difference between the mean with the amount of scatter within the groups, error bars that overlap are not statistically significant. The asterisks above columns indicate ($P<0.05$, $n=3$) (see methods).

As expected, the WT was able to perform haemolytic activity. Strains with gene deletions in either the T6SS HSI-I and HSI-II islands were also able to perform haemolysis in the assay to levels comparable of the WT strain (Figure 5.4). Unexpectedly, a gene deletion in VgrG3 of the T6SS HSI-III was able to exhibit haemolytic activity, although statistically slightly lower than the WT ($P<0.05$, n=3). The PA01 strain possessing a gene deletion in ClpV3 displayed significantly less haemolytic activity when compared to the WT strain or its corresponding ClpV gene in either HSI-I or HSI-II. This activity was returned upon complementation with the pME6032:ClpV3 vector, but not the empty pME6032 vector control, suggesting it is the deletion in the ClpV3 encoding gene that is the direct cause for the loss of haemolytic activity. It was predicted that gene deletions in any T6SS HSI-III gene would result in the loss of haemolytic activity. However, this result suggests that VgrG3 may not be as important as hypothesised for such activity (or that one of VgrG1 or VgrG2 is able to complement the loss). It is evident, however that deletions in either of the HSI-I or HSI-II islands does not impact the ability of PA01 to perform haemolysis. Furthermore, these results suggest ClpV3 is required for haemolytic activity, which could explain why both strains I and J exhibited a significant loss in haemolytic activity given ClpV3 and the entire HSI-III was significantly downregulated in both strains.

5.2.2 Protease assay using T6SS HSI I-III mutant strains

As with the hypothesis for the haemolysis assay, strains possessing knockouts in genes of the T6SS HSI-III should be attenuated in performing proteolytic activity and therefore exhibit phenotypes similar to those seen for PaMip knockout strains I and J, while deletion of the corresponding genes in islands HSI-I and HSI-II should still phenocopy the WT.

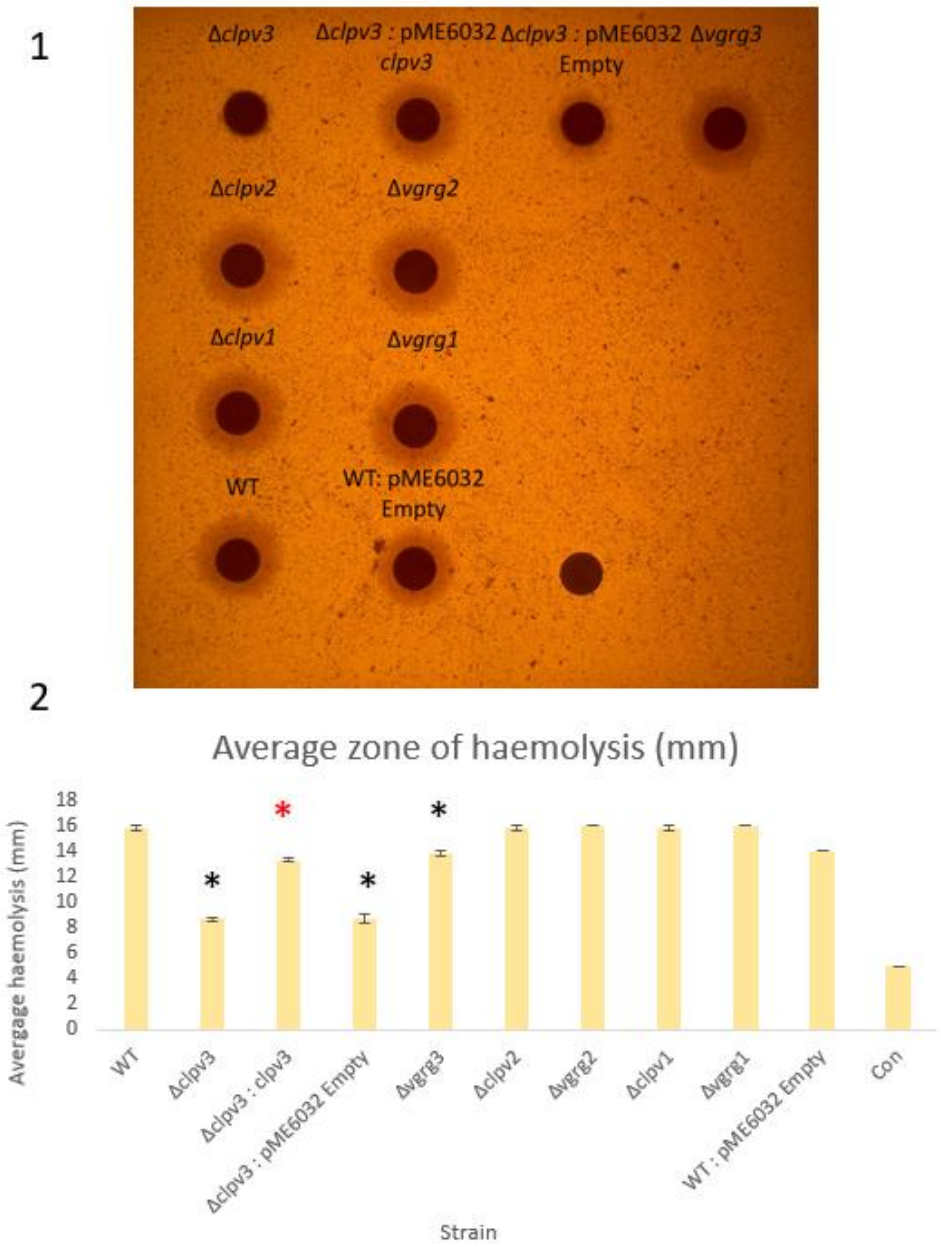


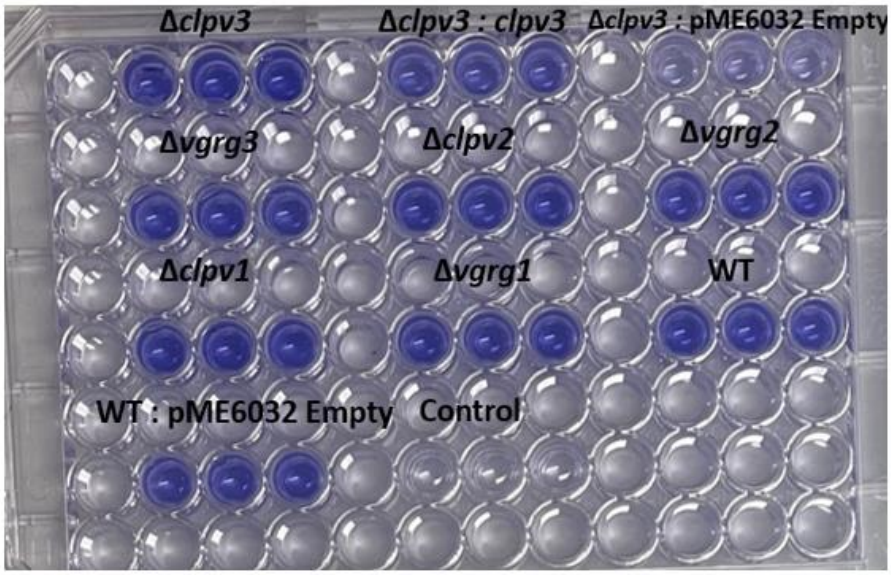
Figure 5.5. Proteolytic activity of T6SS deletion mutants. Panel 1 shows lacto-free whey agar plates agar plates spotted with deletion mutants of T6SS components for each T6SS island. Proteolysis is indicated by clear inhibition zones where the whey protein has been broken down. The WT has a large activity zone with mutants in the T6SS HSI I-II clusters displaying comparable activity. The control using sterile water exhibited no haemolytic activity as expected. The ΔClpV3 strain has a clear, reduced halo with less proteolytic activity which was returned upon complementation. The bar chart labelled in panel 2 shows the average zones of proteolytic activity measured from plates. Along the X axis are the T6SS knockout strains as described in Table 5.1. Standard error variance was calculated and plotted as seen by the error bars comparing the difference between the mean with the amount of scatter within the groups, error bars that overlap are not statistically significant. The asterisks above columns indicate ($P<0.05$, $n=3$) (see methods).

Consistent with the haemolytic activity assays above, strains with gene deletions in either the T6SS HSI-I and HSI-II islands exhibited proteolytic activity at levels comparable to the WT strain (Figure 5.5). Contrary to our hypothesis, a VgrG3 knockout was able to perform proteolytic activity at levels comparable to the WT suggesting VgrG3 is not essential for proteolytic activity in PA01. The PA01 strain possessing a deletion in the *clpv3* gene performed significantly less proteolytic activity when compared to the WT strain or the corresponding *clpv* gene deletions in either HSI-I or HSI-II. This activity was returned upon complementation with the pME6032: *clpv3* vector, but not the empty pME6032 vector control, suggesting it is the deletion in *clpv3* gene that is the direct cause for the loss of proteolytic activity. Again, as described above, these results suggest ClpV3 is in some way required for proteolytic activity, which could explain why both strains I and J exhibited a loss in proteolytic activity given ClpV3 and the entire HSI-III was significantly downregulated in both strains as evidenced in the RNA seq. This result I supported further by a recent publication showing several effectors secreted by the T6SS HSS-III are phospholipases [263].

5.2.3 Biofilm formation assay using T6SS HSI I-III mutant strains

In the last two sections I examined virulence phenotypes associated with mechanisms that depend upon effector delivery, something that would logically be associated with the T6SS HSI-III. In contrast we would not expect biofilm formation to be associated with the T6SS. On this basis we anticipated the VgrG and ClpV deletion strains would behave as the WT if tested in the biofilm assay.

1



2

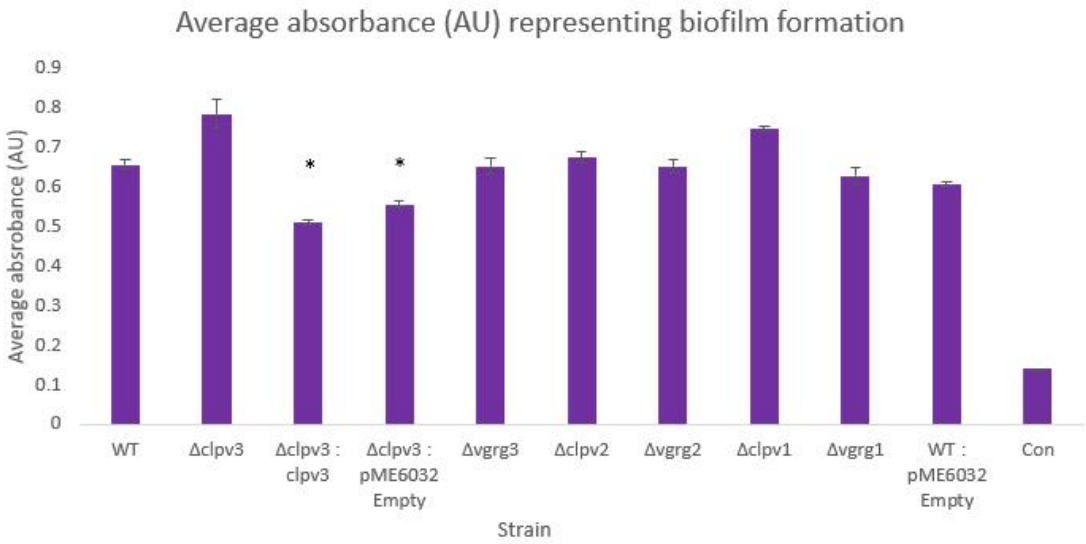


Figure 5.6. Biofilm assay using T6SS knockout strains. Biofilm retained dye dissolved in acetic acid displaying different purple intensities across all strains carried out in triplicate (panel 1). Example plate where strains have been grown in a 96-well plate then stained with crystal violet. Wells with a much lighter colour compared to the WT and other strains have produced less biofilm. Strains are labelled from the ΔClpV3 strain to the control well, each strain was grown in triplicate. The retained dye dissolved in acetic acid was quantified by UV and plotted in the bar chart in panel 2. Standard error variance was calculated and plotted as seen by the error bars comparing the difference between the mean with the amount of scatter within the groups, error bars that overlap are not statistically significant. A control well of uninoculated LB was used. The asterisks above columns indicate ($P<0.05$, $n=3$) (see methods).

As expected, all T6SS knockout strains were able to form biofilms to similar densities of the WT strain, suggesting neither *vgrg* nor *clpv* genes and their products in any T6SS island are required for biofilm formation. Unexpectedly the ClpV3 complementation strain mutant and the ClpV3 strain housing empty pME6032 both produced statistically less dense biofilms than the WT or parent strains (Figure 5.6). In contrast the WT transformed with empty pME6032 vector was shown to form a biofilm as well as its untransformed WT counterpart, thus the addition of the

pME6032 vector does not appear to affect biofilm formation and the result obtained for the ClpV3 strains is obscure and should be repeated.

5.2.4 Twitching motility assay using T6SS HSI I-III mutant strains

Given twitching motility is a type IV pili dependant form of movement, I hypothesised that it would be unlikely a deletion of any gene in any of the T6SS HSI islands would result in the loss of, or, affect significantly twitching motility in the T6SS mutants. Strains with gene deletions in either the T6SS HSI-I and HSI-II islands as well as the Δ*vgrg3* mutant were able to twitch to levels comparable of the WT (Figure 5.7 and 5.8). Surprisingly, the Δ*clpv3* mutant was unable to perform twitching motility to any degree. Upon complementation of the *clpv3* gene, the twitching motility was returned. When transformed with the empty pME6032 vector control no twitching motility was returned in the Δ*clpv3* strain. This data surprisingly suggests ClpV3 is somehow required for twitching motility. How it is involved remains unknown, however a recent publication has linked the expression of T6SS to the number of type IV pili expressed [263]. The suggestion made by the authors was that, as the T6SS is a weapon that fires when in contact with a neighbouring cell, to move toward or away from a target cell it requires movement in the form of twitching; thus type IV pili are expressed more highly [263]. Furthermore, both the T6SS HSI-III and type IV pili are under the control of the GacS / RetS two component system [264]. A deletion in in the *clpv3* gene may have some genetic link to how the type IV pili are expressed, perhaps as ClpV3 is an ATPase it is somehow required the movement of type IV pili which is also ATP dependant. Alternatively, there could be a feedback mechanism at work as the failure to initiate T6SS HSI-III mediated eukaryotic cell puncturing results in the downregulation of other systems which would be rendered pointless without the puncturing of a target cell. Based on the current data, these are all assumptions and require much more in-depth research to determine the functional link between these systems.

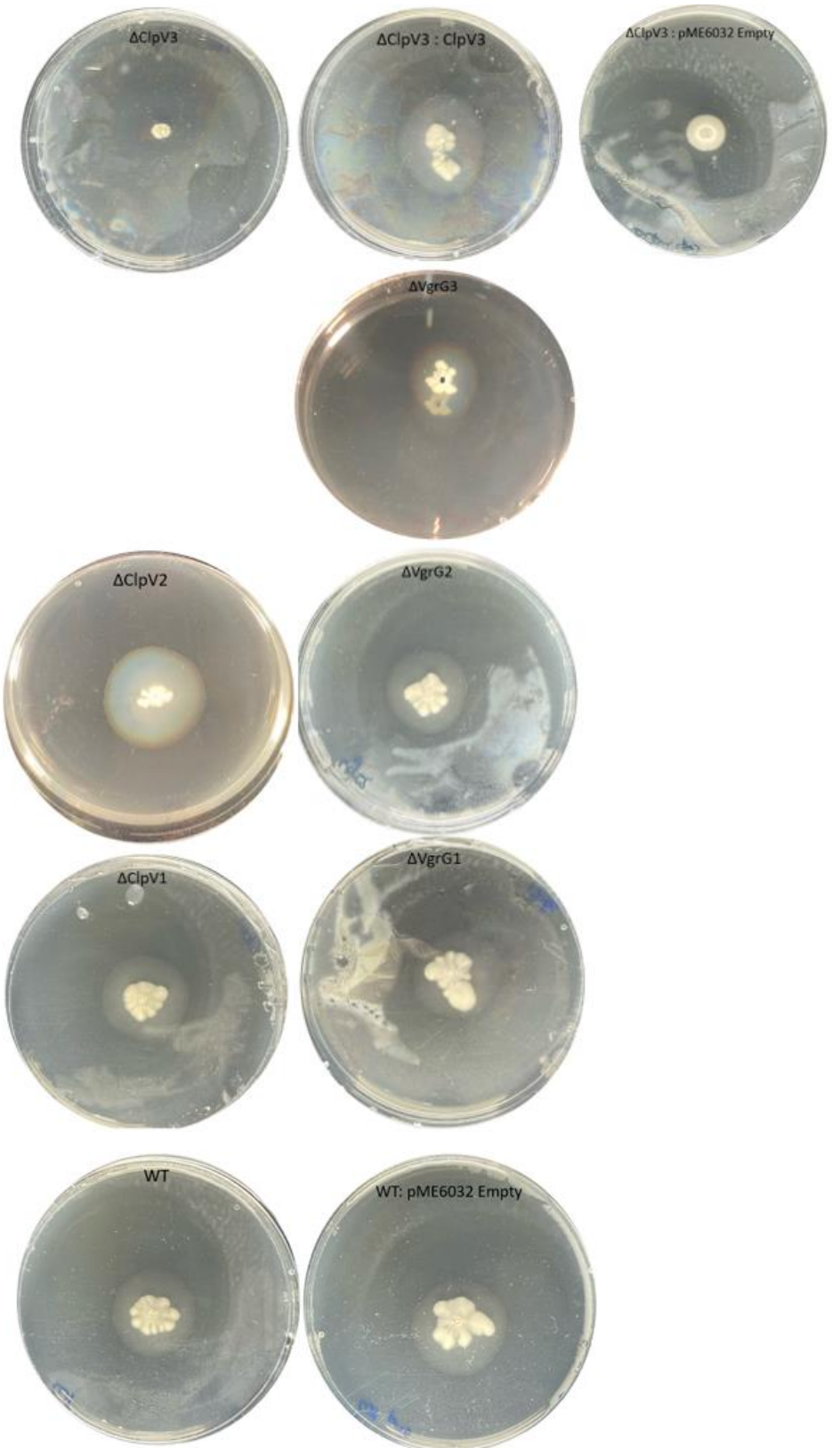


Figure 5.7. Twitching motility assay using T6SS knockout strains. The WT has a very faint white cloud, this is the growth outwards along the solid plastic petri dish as a result of twitching; the denser, more defined, growth is on the surface of the agar arising from the initial inoculum. The WT has a clear twitching motility with mutants in the T6SS HSI I-II clusters displaying comparable activity. The Δ*clpv3* strain has no twitching motility which was returned upon complementation.

Figure 5.8. Bar chart plotting the average zone of twitching motility measured after 48h for T6SS knockout strains. Twitching motility zones were calculated using the equation; ½a x ½b x π (where 'a' is the longest diameter and 'b' is the shortest diameter). Standard error variance was calculated and plotted as seen by the error bars comparing the difference between the mean with the amount of scatter within the groups, error bars that overlap are not statistically significant. The asterisks above columns indicate ($P<0.05$, n=2) (see methods).

5.2.5 Swarming motility assay using T6SS HSI I-III mutant strains

As with the twitching assay hypothesis, given swarming is a flagella dependant form of movement, I hypothesised that it would be unlikely a deletion of any gene in any of the T6SS HSI islands would result in the loss of or affect swarming motility. Strains with gene deletions in either the T6SS HSI-I and HSI-II islands as well as the Δ*vgrg3* mutant were able to swarm to levels comparable of the WT (Figure 5.9). Once again, the Δ*clpv3* strain presented an unusual phenotype and appeared to swarm less than the WT or other T6SS knockout strains. Swarming plates are subjective and visual with no statistical analysis performed. However, it does appear that for the Δ*clpv3* mutant plate, swarming movement is performed to a lesser degree than other strains, visualised by a smaller circle. The movement appears to be uniform and circular as with the WT strain, however to smaller diameter. When complemented with ClpV3, the Δ*clpv3*:*clpv3* strain swarms to an abnormal extent with a phenotype similar to strain I (Figure 3.8), whereas the Δ*clpv3* strain transformed with empty pME6032 swarmed similarly to the untransformed Δ*clpv3* strain. Why this phenotype is observed it is difficult to say. However, a recent pre-print reported a similar phenotype in a swarming assay for a Δ*clpv3* mutant [265]. This result will be discussed in detail shortly.

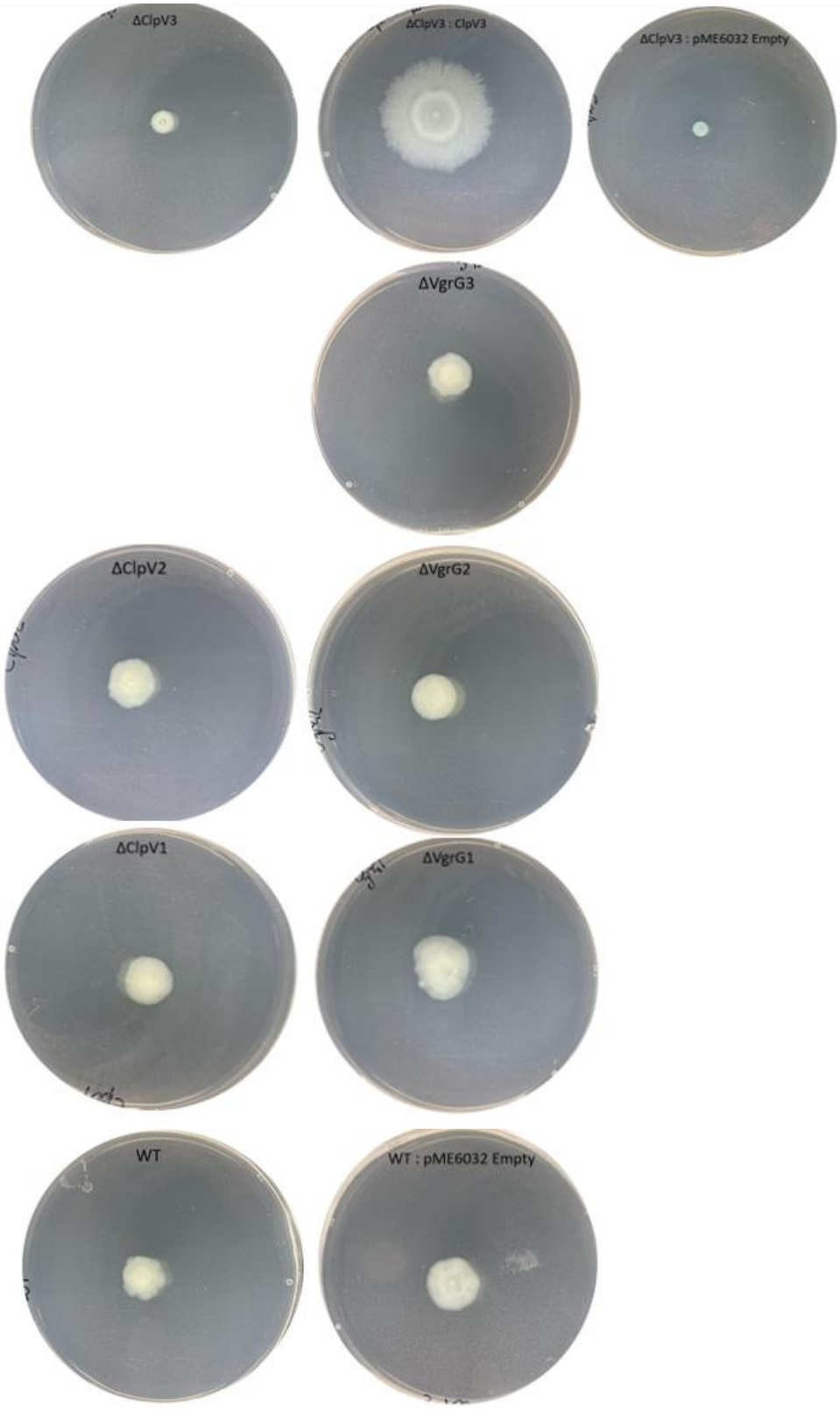


Figure 5.9. Swarming motility plates for the T6SS knockout strains. The WT has a circular uniform circle outward with mutants in the T6SS HSI I-II clusters displaying comparable motility. The Δ*clpv3* strain has reduced swarming motility which was greatly increased upon complementation to levels surpassing the WT.

5.3 *In vivo* infection assays

5.3.1 *Galleria mellonella* infection model using T6SS HSI I-III mutant strains

Thus far, mutants with deletions in the genes encoding HSI-I ClpV1 / VgrG1 and HSI-II ClpV2 / VgrG2 display no effect on the plate-based virulence phenotypes assessed above, this is consistent with literature reports which describe them as being responsible for interactions with prokaryotic cells, not eukaryotic. In all of these HSI-I / HSI-II mutant strains, the HSI-III will remain functional. Therefore, based on our hypothesis that the HSI-III is responsible for eukaryotic cell interactions; the HSI-I / HSI-II mutant strains should still be able to kill *G. mellonella* as effectively as the PA01 WT strain. Furthermore, given the unexpected results described for the Δ*vgrg3* strain showing no difference to the WT in plate-based virulence phenotypes, I hypothesised that this strain should still be able to kill *G. mellonella*. Finally, given the observed plate-based virulence phenotypes of Δ*clpv3* strain, it would be expected in an *in vivo model* to exhibit a delayed killing of *G. mellonella*, however killing should not be completely abolished as was observed for PaMip mutant strain I, as, theoretically other important virulence systems responsible for eukaryotic cellular interaction, such as the T3SS, should still be expressed.

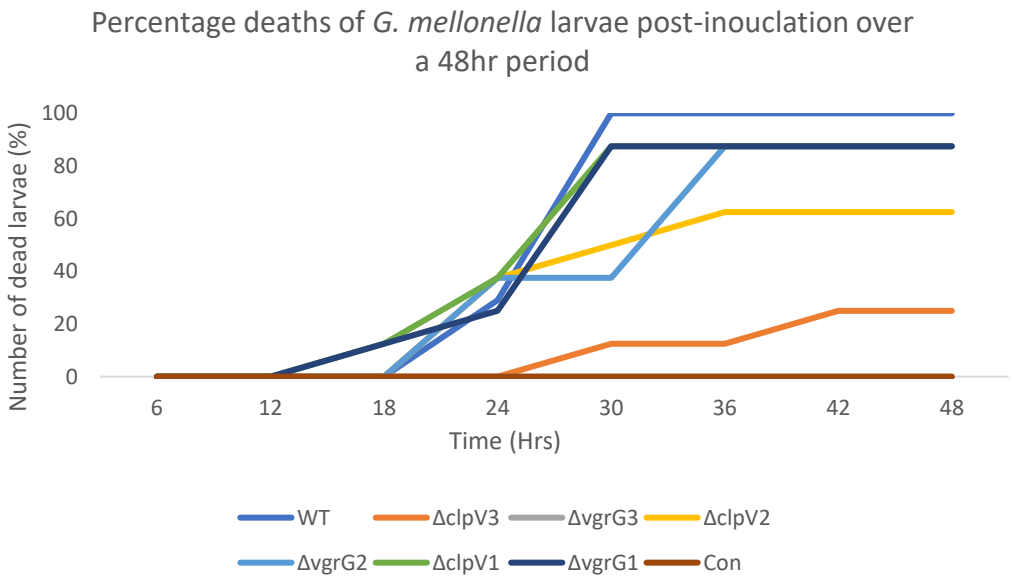


Figure 5.10. Line graph plotting deaths of *G. mellonella* larvae inoculated with T6SS knockout strains. Percentage of deaths in larvae for all T6SS mutant strains over a 48 h period with recordings taken at 6 h intervals. This experiment was performed twice with results pooled. 10 larvae were used per strain experiment. Final deaths were recorded as a percentage.

As with *G. mellonella* assays described in previous chapters, the WT starts killing larvae at around the 18 h timepoint with 100% larval deaths after 30 h. The Δ*clpv1* and Δ*vgrg1* strains began killing *G. mellonella* at the same time point as the WT, while also reaching 100% dead larvae at a comparable time to the WT. The Δ*clpv2* and Δ *vgrg2* knockout strains began killing of *G. mellonella* at the same timepoint as the WT, but took longer to kill the larvae, only reaching 90%

deaths after 36 h for Δ*vgrg2* strain and 60% deaths after 48 h for Δ*clpv2* knockout strain as opposed to the usual 30 h time point. Given the *in vitro* results presented thus far for the Δ*clpv3* knockout strain with a reduction or loss of several virulence related phenotypes, it was predicted *in vivo* virulence would also be attenuated. Indeed, the *clpv3* mutant strain performed poorly in the killing assay, beginning to kill *G. mellonella* after 24 h and only killing 20% of inoculated larvae over the duration of the assay. The most unexpected result was observed in the Δ*vgrg3* mutant strain which was unable to kill any *G. mellonella* larvae in either experimental run. Given no phenotypic changes were observed for the mutant in any of the *in vitro* assay, I had predicted it would still be able to kill *G. mellonella* at levels comparable to the WT. However, the complete opposite was recorded suggesting that VgrG3 is involved and required for PA01 virulence. Exactly why virulence is only observed *in vivo* for this mutant remains unclear. One hypothesis I would propose is that given, *vgrg* genes encode spike proteins involved in puncturing host cells (although previously not known to puncture eukaryotic cells), and house different effectors and toxins, it is plausible that VgrG3 of PA01 is a spike protein required for the delivery of effectors and toxins specific to eukaryotic host cells, something that has not yet been described in literature. Investigation of this hypothesis will require further experimentation and study which I will discuss in my final chapter.

5.4 RNA Sequencing analysis of the Δ*clpv3* mutant strain

Given the strong phenotypes observed for the *in vitro* assays, the Δ*clpv3* knockout strain is likely to express some genes significantly different compared to the WT. RNA seq analysis may therefore aid in elucidating why some phenotypes of these virulence assays were observed. In part, only the Δ*clpv3* (*pa2371*) deletion strain was sent for RNA sequencing due to the cost involved and the phenotypes observed with only the Δ*clpv3* knockout strain exhibiting significant phenotypic changes (plus the decision was made prior to the *in vivo* assays being performed). As with the RNA sequencing described in Chapter 3, three biological replicates were used for the Δ*clpv3* knockout strain which was run alongside three replicates of the WT. As with Chapter 3, the strains were grown on agar plates as the basis for RNA extraction rather than liquid culture to mimic the conditions of most plate-based virulence assays described thus far in attempt to recreate as closely as possible the transcriptome of the Δ*clpv3* knockout strain in a virulence assay.

As with previously described RNA seq, Dr Govind Chandra analysed the data starting with the WT strain which had RNA extraction performed in triplicate taking an average of all three transcript readings for each gene and subjected to statistical analysis ($P<0.05$, $FDR<0.05$) to establish a base transcript level for every gene in the PA01 genome. A P value of <0.05 would indicate the RNA seq result was of statistical significance, the FDR (false discovery rate) value,

also set at a significance level of FDR<0.05, further ensures statistical significance by correcting for multiple testing by giving the proportion of tests above a threshold will be false positives. The *clpv3* gene sits at the end of an operon, so its deletion and genetic manipulation would unlikely have unintended consequences on the transcription of surrounding genes. Furthermore, to minimise these possible polar effects of a gene deletion, 9-51 bp at the beginning and end of the gene were left, creating a truncated version of the gene to be deleted in multiples of 3 to ensure the reading frame remained intact.

The ClpV3 operon begins at *pa2365* and ends with *clpv3* (*pa2371*) as noted (Figure 5.11). The first gene in the ClpV3 operon is *pa2365* encoding ***hsiB3,*** a hypothetical protein of the T6SS HSI-III predicted to make up tubules of the contractile sheath [266]. For Δ*clpv3*, every gene in this operon in the RNA seq is downregulated up to four-fold when compared to the WT, which could mean two things, the first is that the genetic manipulation of the *clpv3* gene has somehow affected the transcription of the entire operon (Table 5.2). Secondly, and more likely, given the *clpv3* gene has been deleted, the entire assembly of the T6SS HSI-III apparatus is affected as ClpV3 forms the baseplate from which the structure builds while also provides an important source of energy for the system, potentially acting as a feedback mechanism. However, our observation is consistent with a recent paper which will be discussed below, in which RNA seq data for a Δ*clpv3* knockout strain also revealed that all genes involved the T6SS HSI-III system were significantly downregulated [265]. Furthermore, genes involved in the T6SS HSI-III that do not sit on the same operon as ClpV3 including VgrG3, which are under divergent control of RpoN, are also significantly downregulated (Table 5.2). The T6SS HSI-III is believed to be activated by QS and under the control of RpoN, which has divergent control of the three operons via the *las* and *rhl* systems. RpoN activates the expression of the HSI-III left operon, while unexpectedly is thought to repress the right operon. The two HSI-III operons are induced during the growth stage transition by LasR and RhlR regulators, and are fully expressed in the late stationary phase. Furthermore, the HSI-III genes belong to a group of genes that are induced when in contact with eukaryotic cells and are more highly expressed at 37°C [243].

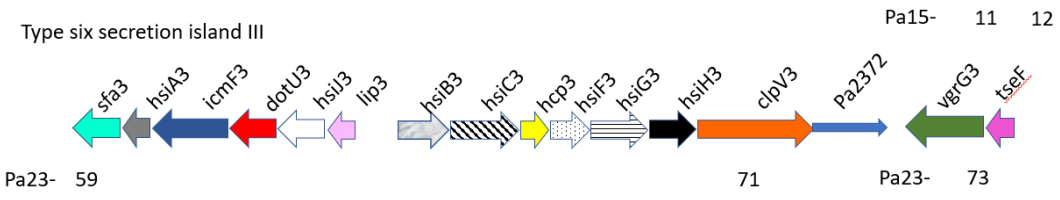


Figure 5.11. Diagram of the T6SS HSI-III operons in the *P. aeruginosa* genome. Abbreviations for each gene are as follows; *ppk*, *Pseudomonas* protein kinase; *ppp*, *Pseudomonas* protein phosphatase; *icm*, intracellular multiplication; *dot*, defect in organelle trafficking; *hsi*, hcp secretion island; *lip*, lipoprotein; *fha*, forkhead-associated protein; *hcp*, haemolysin coregulated protein; *sfa*, sigma activator factor; *stk*, serine/threonine kinase; *stp*, serine/threonine phosphatase; *vgrG*, valine-glycine repeats protein.

			Δ*clpV3*		
Gene No	Gene name	Description	LogFC	p Value	FDR
Haemolytic and Proteolytic genes					
pa0844	*plcH*	Haemolytic phospholipase C precursor	0.497	0.195	0.308
pa0766	*mucD*	Serine protease MucD precursor	0.832	0.058	0.118
pa1871	*lasA*	Protease precursor	-1.147	0.009	0.026
pa2862	*lipA*	lipA, lactonizing lipase	0.142	0.536	0.653
pa5134	*ctpA*	Carboxyl-terminal protease	0.860	0.000	0.001
GacS / RetS Two-Component system genes					
pa0527	*rsmY*	Regulatory RNA	-0.463	0.035	0.079
pa0905	*rsmA*	Regulatory RNA	-1.440	0.004	0.013
pa0928	*gacS*	Sensor/response regulator hybrid	-0.007	0.965	0.979
pa2586	*gacA*	Response Regulator	-0.655	0.094	0.175
pa3621.1	*rsmZ*	Regulatory RNA	1.931	0.001	0.000
pa3974	*ladS*	Lost Adherence Sensor	-1.135	0.001	0.001
pa4856	*retS*	Regulator of Exopolysaccharide and Type III Secretion	-1.199	0.001	0.001
Alginate /*psl* / *pel* mucoidy synthesis genes					
pa0762	*algT*	Positive regulator alginate biosynthesis	0.177	0.704	0.790
pa0763	*mucA*	Anti-sigma factor MucA	0.647	0.225	0.342
pa0764	*mucB*	Negative regulator alginate biosynthesis	0.724	0.135	0.232
pa4033	*MucE*	Positive regulator alginate biosynthesis	1.762	0.000	0.002
pa3649	*mucP*	Positive regulator alginate biosynthesis	0.051	0.811	0.869
pa5253	*algP*	Alginate biosynthesis regulatory protein	0.094	0.812	0.869
pa3541	*alg8*	Alginate biosynthesis	0.032	0.935	0.958
pa3542	*alg44*	Alginate biosynthesis	0.966	0.102	0.186
pa3543	*algK*	Alginate biosynthesis	1.259	0.033	0.075
pa3544	*algE*	Alginate biosynthesis	0.566	0.152	0.255
pa3545	*algG*	Alginate biosynthesis	-0.735	0.002	0.007
pa3546	*algX*	Alginate biosynthesis	0.594	0.342	0.470
pa3547	*algL*	Alginate biosynthesis	0.607	0.346	0.473
pa3548	*algI*	Alginate biosynthesis	0.114	0.782	0.847
pa3549	*algJ*	Alginate biosynthesis	0.782	0.292	0.417
pa3550	*algF*	Alginate biosynthesis	-0.461	0.306	0.431
pa3551	*algA*	Alginate biosynthesis	-0.191	0.467	0.594
pa2232	*pslB*	pslB, biofilm formation protein PslB	-1.616	0.0001	0.0001
pa2233	*pslC*	pslC, biofilm formation protein PslC	-1.441	0.0001	0.0001
pa2234	*pslD*	pslD, biofilm formation protein PslD	-1.423	0.0001	0.001
pa2235	*pslE*	pslE, biofilm formation protein PslE	-1.677	0.0001	0.0001
pa2236	*pslF*	pslF, biofilm formation protein PslF	-1.616	0.0001	0.0001
pa2237	*pslG*	pslG, biofilm formation protein PslG	-1.493	0.0001	0.0001
pa2238	*pslH*	pslH, biofilm formation protein PslH	-1.169	0.0001	0.001
pa2239	*pslI*	pslI, biofilm formation protein PslI	-1.199	0.0001	0.0001
pa2240	*pslJ*	pslJ, biofilm formation protein PslJ	-1.167	0.0001	0.0001

pa2241	*pslK*	pslK, biofilm formation protein PslK	-0.796	0.000	0.001
pa2242	*pslL*	pslL, hypothetical protein	-0.780	0.005	0.018
pa2243	*pslM*	pslM, FAD-binding dehydrogenase	1.042	0.009	0.026
pa2244	*pslN*	pslN, hypothetical protein	-0.706	0.090	0.169
pa3058	*pelG*	pelG, pellicle/biofilm biosynthesis	0.371	0.323	0.449
pa3059	*pelF*	pelF, pellicle/biofilm biosynthesis	0.182	0.685	0.777
pa3060	*pelE*	pelE, pellicle/biofilm biosynthesis	0.257	0.592	0.701
pa3061	*pelD*	pelD, pellicle/biofilm biosynthesis	0.618	0.287	0.412
pa3062	*pelC*	pelC, pellicle/biofilm biosynthesis	0.677	0.312	0.437
pa3063	*pelB*	pelB, pellicle/biofilm biosynthesis	1.095	0.091	0.171
pa3064	*pelA*	pelA, hypothetical protein	-0.078	0.864	0.910
***las / rhl* Quorum sensing (QS) system genes**					
pa1432	*lasI*	Autoinducer synthesis protein	-0.191	0.562	0.674
pa1430	*lasR*	Transcriptional regulator	-1.213	0.007	0.023
pa3476	*rhlI*	Autoinducer synthesis protein RhlI	0.588	0.123	0.216
pa3477	*rhlR*	Transcriptional regulator	-0.773	0.054	0.111
Type IV Pili assocaited genes					
pa4462	*rpoN*	Sigma factor	0.137	0.631	0.732
pa4525	*pilA*	Type IV pili synthesis	-0.137	0.788	0.852
pa4526	*pilB*	Type IV pili synthesis	2.535	0.0001	0.0001
pa4528	*pilD*	Type IV pili synthesis	0.732	0.0001	0.001
pa4546	*pilS*	Type IV pili synthesis	0.730	0.001	0.005
pa4547	*pilR*	Type IV pili synthesis	0.226	0.211	0.326
pa4549	*fimT*	Type IV pili synthesis	1.224	0.003	0.012
pa4550	*fimU*	Type IV pili synthesis	-0.182	0.621	0.725
pa4551	*pilV*	Type IV pili synthesis	-0.422	0.245	0.365
pa4552	*pilW*	Type IV pili synthesis	-0.640	0.091	0.170
pa4553	*pilX*	Type IV pili synthesis	-0.580	0.111	0.200
pa4554	*pilY*	Type IV pili synthesis	0.065	0.859	0.905
pa4555	*pilY2*	Type IV pili synthesis	-0.748	0.009	0.027
pa4556	*pilE*	Type IV pili synthesis	-1.034	0.0001	0.001
pa4959	*fimX*	Phosphodiesterase involved in twitching	0.155	0.414	0.540
Cyclic-di-GMP synthesis genes					
pa1120	*tpbB*	diguanylate cyclase	-0.069	0.725	0.805
pa4108		Cyclic di-gmp phosphodiesterase - degradation of cyclic di-gm	-1.028	0.0001	0.0001
pa4781		Cyclic di-gmp phosphodiesterase	-1.619	0.001	0.003
pa3702	wspR	directs formation of c-di-GMP	-0.150	0.417	0.543
pa3703	wspF	directs formation of c-di-GMP	-0.509	0.021	0.054
Flagella assembly and Rhamnolipid biosynthesis genes					
pa1100	fliE	fliE, flagellar hook-basal body complex	1.541	0.0001	0.0001
pa1101	fliF	fliF, flagellar MS-ring protein	1.741	0.0001	0.0001
pa1102	fliG	fliG, flagellar motor switch protein FliG	0.884	0.001	0.003
pa1103	Hypo	flagellar assembly protein FliH	-0.172	0.249	0.370
pa1104	fliL	fliI, flagellum-specific ATP synthase	0.200	0.153	0.256
pa1105	fliJ	fliJ, flagellar biosynthesis chaperone	0.127	0.543	0.659
pa1454	fleN	flagellar synthesis regulator	-0.079	0.798	0.859

pa3387	*rhlG*	Rhamnolipid biosynthesis	2.737	0.000	0.000
pa3476	*rhlL*	Rhamnolipid biosynthesis	0.588	0.123	0.216
pa3478	*rhlB*	Rhamnolipid biosynthesis	-1.173	0.001	0.004
pa3479	*rhlA*	Rhamnolipid biosynthesis	-1.214	0.004	0.015
Type Six Secretion System HSI-III genes					
pa2359	*sfa3*	Probable transcriptional regulator	-0.156	0.725	0.805
pa2360	*hsiA3*	T6SS HSI-III hypothetical protein	-1.567	0.001	0.004
pa2361	*Icmf3*	T6SS HSI-III hypothetical protein	-1.193	0.015	0.041
pa2362	*dotU3*	T6SS HSI-III hypothetical protein	-1.435	0.006	0.020
pa2363	*hsiJ3*	T6SS HSI-III hypothetical protein	-1.265	0.002	0.009
pa2364	*lip3*	T6SS HSI-III hypothetical protein	-1.705	0.000	0.000
pa2365	*hsiB3*	T6SS HSI-III hypothetical protein	-2.385	0.000	0.001
pa2366	*hsiC3*	T6SS HSI-III hypothetical protein	-2.233	0.000	0.001
pa2367	*hcp3*	Secreted protein Hcp	-1.238	0.013	0.035
pa2368	*hsiF3*	T6SS HSI-III hypothetical protein	-1.715	0.003	0.012
pa2369	*hsiG3*	T6SS HSI-III hypothetical protein	-0.701	0.174	0.284
pa2370	*hsiH3*	T6SS HSI-III hypothetical protein	0.883	0.077	0.149
pa2371	*ClpV3*	Probable ClpA/B-type protease	-2.040	0.000	0.001
pa2373	*VgrG3*	T6SS HSI-III hypothetical protein	-0.335	0.270	0.394
pa2374	*tseF*	T6SS HSI-III hypothetical protein	-1.329	0.000	0.000
Type Six Secretion System HSI-II genes					
pa1511	*VgrG2*	T6SS HSI-II hypothetical protein	-0.504	0.076	0.148
pa1512	*HCP2*	Secreted protein Hcp	-0.217	0.477	0.603
pa1656	*hsiA2*	T6SS HSI-II hypothetical protein	0.745	0.002	0.007
pa1657	*hsiB2*	T6SS HSI-II hypothetical protein	0.251	0.358	0.485
pa1658	*hsiC2*	T6SS HSI-II hypothetical protein	-0.656	0.027	0.065
pa1659	*hsif3*	T6SS HSI-II hypothetical protein	-0.721	0.103	0.188
pa1660	*hsiG2*	T6SS HSI-II hypothetical protein	-0.282	0.473	0.599
pa1661	*hsiH2*	T6SS HSI-II hypothetical protein	-0.502	0.218	0.334
pa1662	*ClpV2*	Probable ClpA/B-type protease	-0.780	0.033	0.075
pa1663	*sfa2*	Probable transcriptional regulator	-0.263	0.614	0.720
pa1664	*fha2*	T6SS HSI-II hypothetical protein	-0.786	0.359	0.486
pa1665	*lip2*	T6SS HSI-II hypothetical protein	-0.197	0.728	0.807
pa1666	*hsij2*	T6SS HSI-II hypothetical protein	-0.533	0.336	0.463
pa1667	*dotu3*	T6SS HSI-II hypothetical protein	-0.769	0.001	0.005
pa1668	*icmF2*	T6SS HSI-II hypothetical protein	-0.516	0.205	0.320
pa1669	*stp1*	T6SS HSI-II hypothetical protein	-0.351	0.238	0.358
pa1670	*stk1*	T6SS HSI-II hypothetical protein	-0.178	0.559	0.672
Type Six Secretion System HSI-I genes					
pa0074	*ppkA*	Serine threonine protein kinase PpkA	0.165	0.738	0.815
pa0075	*pppA*	serine/threonine phosphatase	0.128	0.821	0.876
pa0076	*pppB*	T6SS HSI-I hypothetical protein	-0.166	0.763	0.834
pa0077	*icmF1*	Type VI secretion protein IcmF	-0.023	0.958	0.973
pa0078	*dotU1*	T6SS HSI-I hypothetical protein	-0.310	0.495	0.617
pa0079	*hsiJ1*	T6SS HSI-I hypothetical protein	-0.766	0.041	0.089
pa0080	*lip1*	T6SS HSI-I hypothetical protein	-0.959	0.027	0.065
pa0081	*fha1*	Fha domain-containing protein	-0.946	0.000	0.000

pa0082	*hsiA1*	T6SS HSI-I hypothetical protein	-0.114	0.662	0.759
pa0083	*hsiB1*	T6SS HSI-I hypothetical protein	-0.571	0.172	0.281
pa0084	*hsiC1*	T6SS HSI-I hypothetical protein	-0.632	0.113	0.202
pa0085	*hcp1*	Secreted protein Hcp	-1.059	0.006	0.019
pa0086	*hsiE1*	T6SS HSI-I hypothetical protein	0.427	0.307	0.432
pa0087	*hsif1*	T6SS HSI-I hypothetical protein	0.684	0.217	0.333
pa0088	*hsiG1*	T6SS HSI-I hypothetical protein	0.330	0.503	0.625
pa0089	*hsiH1*	T6SS HSI-I hypothetical protein	0.352	0.544	0.659
pa0090	*clpv1*	Probable ClpA/B-type protease	-0.030	0.940	0.962
pa0091	*Vgrg1*	T6SS HSI-I hypothetical protein	0.604	0.078	0.151
Type Three Secretion System genes					
pa1697	*pscN*	Type III SS related	1.773	0.012	0.033
pa1706	*ppcrV*	Type III SS related	0.978	0.000	0.001
pa1708	*popB*	Type III SS related	0.242	0.203	0.318
pa1709	*popD*	Type III SS related	0.353	0.110	0.198
pa1710	*exsC*	Type III SS related	0.112	0.619	0.723
pa1711	*exsE*	Type III SS related	0.208	0.248	0.369
pa1712	*exsB*	Type III SS related	0.582	0.005	0.018
pa1713	*exsA*	Type III SS related	2.298	0.000	0.000
pa1714	*exsD*	Type III SS related	1.169	0.000	0.000
pa1716	*pscC*	Type III SS related	0.784	0.045	0.095
pa1719	*pscF*	Type III SS related	0.756	0.126	0.220
pa1723	*PscJ*	Type III SS related	1.035	0.022	0.054
pa1725	*pscL*	Type III SS related	0.415	0.483	0.608

Table 5.2. A table of RNA sequence data for the ΔClpV3 knockout strain looking at specific groups of genes involved in a variety of virulence processes. The left column labelled gene number is the annotated gene number as it appears in the chromosome, the column labelled gene name is the name given to a particular gene in literature and what it may be referred to as. The Log Fold Change (LogFC) column is the log-ratio of a gene's or a transcript's expression values compared to the WT strain. The LogFC value for each gene was obtained taking the average of three transcript readings for each gene in the Pa01 genome from which statistical anaylsis could be applied. To get the overall fold change in real terms compared to the WT, 2 is raised to the power of the LogFC value. The P value column indicated if the LogFC change of a given gene was of statistical significance if P<0.05. This was further statistically challenged by looking at the FDR value, also set at a significance level of FDR<0.05. Using both values ensured the LogFC value was statistically significant and not down to chance or false positives. Boxes highlighted in green call attention to a gene that has an increased LogFC change when compared to the WT and are statistically significant (P<0.05, FDR<0.05) suggesting the gene is more highly transcribed and upregulated compared to the WT. Boxes highlighted in red call attention to a gene that has a decrease LogFC change when compared to the WT and are statistically significant (P<0.05, FDR<0.05) suggesting the gene is more lowly transcribed and upregulated compared to the WT. Boxes that remain unhighlighted show genes whose LogFC change is not statistically significant when compared to the WT, therefore one could assume transcript levels of a given gene between the WT and a given strain are no different, neither up or down regulated.

With the results for Δ*clpv3* strain in the *G. mellonella* assay, the most interesting read out from the RNA seq data is for genes involved in the T3SS. As discussed previously, in literature, the T3SS has been annotated as the primary weapon of PA01 that is the only known system to translocate secreted toxins and effectors directly into the cytoplasm of eukaryotic host cells. In the Δ*clpv3* mutant strain, the RNA seq reveals that all genes involved in the T3SS are significantly upregulated compared to the WT. Given that in literature the T3SS has been attributed as a significant virulence determinant that it is required for the optimal killing of *G. mellonella*, its significant upregulation in the Δ*clpv3* knockout strain coupled with the current literature assumptions that the T6SS HSI-III is likely only important in prokaryotic cell interaction, a deletion in the *clpv3* gene should theoretically not have affected PA01's ability to kill *G. mellonella* [267]. This was in fact the opposite to what was observed, in that the Δ*clpv3* mutant was unable to kill *G. mellonella* larvae to levels of the WT despite the T3SS being significantly upregulated. Furthermore, although there is no RNA seq for the Δ*vgrg3* strain to assess what levels the T3SS are transcribed at, the Δ*vgrg3* strain was unable to kill any larvae across both experimental runs. This data together suggests the T6SS HSI-III of PA01 is required for the optimal killing of *G. mellonella* and therefore must, to some extent, interact and deliver toxins directly to eukaryotic cells like the T3SS.

5.5 Discussion

Some aspects of our initial hypothesis regarding the role of the T6SS HSS-III in PA01 viruelnce are supported by the results presented in this chapter. The HSI-III island is important for the haemolytic and proteolytic activity of PA01, with a deletion in the ClpV3 gene resulting in a loss of both activities. Unexpectedly, this was not the case for the Δ*vgrg3* knockout strain. Furthermore, deletions in the T6SS HSI-I and HSI-II did not affect PA01's ability to perform haemolysis and proteolysis or affect biofilm formation, twitching and swarming motilities. The hypothesis that all T6SS knockout strains should still be able to form biofilms was shown to be correct with all strains able to form biofilms to densities comparable of the WT strain.

An aspect of my initial hypothesis shown to be incorrect is that a T6SS HSI-III knockout, specifically the Δ*clpv3* strain, would only affect certain phenotypes in plate-based and *in vivo* virulence assays (Figure 5.2), and that it would only exhibit phenotypes somewhere between that of the PA01 WT strain and those of the PaMip strains I and J (based on its significant downregulation in both strains). Surprisingly, a deletion in *clpv3* does cause an effect on twitching and swarming motilities.

At the time my work was undertaken, and my hypothesis formulated, there were no significant reports in the literature assessing the importance of the T6SS HSI-III in PA01 virulence. As

mentioned previously, the HSI-I and HSI-II islands have been extensively studied and their functions annotated. These reports describe them as primarily being responsible for interaction with prokaryotic cells, in particular of killing off neighbouring cells. There has been no report of PA01 T6SS functioning to interact with eukaryotic cells or that is important for delivering effectors to eukaryotic cells.

Since starting my work, two papers have been published that investigate the potential importance and contribution that the T6SS HSI-III plays in PA01 virulence. One of these papers (by *Li et al.*) described the creation of a Δ*clpv3* knockout in PA01 and testing it in a series of *in vitro* and *in vivo* assays to assess the effect on virulence [265]. They described very similar phenotypes to those described in this thesis i.e. the Δ*clpv3* mutant strain was attenuated in its ability to perform proteolytic activity and also displayed decreased pathogenicity in *G. mellonella* assays. Furthermore, they also reported that their Δ*clpv3* strain exhibited reduced swarming motility and a lack of twitching motility, much to their surprise. The data presented here, coupled with the recently published papers suggests the T6SS HSI-III plays a far more important role in virulence than first thought. Both the research conducted here, and that reported in recently published paper, suggests the T6SS HSI-III secretion system island plays a far more diverse role in virulence than previously recognised and its expression affects many more genes other than those just in the T6SS HSI-III, suggesting a surprisingly complex connectivity with other virulence related loci.

In summary, on the basis of the results presented here and in literature, I propose that the T6SS HSI-III is the secretion island expressed by PA01 during the sessile lifestyle associated with chronic infections and interactions with eukaryotic cells. During the acute infection, the T3SS is the only annotated secretion system of PA01 that has been shown to be responsible for delivering toxins and effectors directly into the cytoplasm of eukaryotic cells [268]. Furthermore, this T3SS system has been shown to be inactive during the sessile lifestyle associated with chronic infections [67, 268]. During the lifestyle transition mediated by the GacS /RetS two component system, T3SS genes are switched off, while all T6SS HSI genes are switched on. It is believed that the HSI-I and HSI-II are switched on to compete with neighbouring bacterial cells in a nutrient limiting environment while PA01 is stationary and forming a biofilm. It is entirely plausible that, even during this sessile lifestyle, PA01 would need to express a secretion system responsible for the delivery of toxins and effectors directly into eukaryotic host cells (e.g. during a chronic infection process: this requirement could be fulfilled by the T6SS HSI-III. Evidence to support this hypothesis is provided by both my *G. mellonella* data and that presented in the paper by *Li et al*., which both show, independently, that Δ*clpv3* knockout strains, plus my Δ*vgrg3* knockout strain, were severely attenuated at killing *G. mellonella*. Furthermore, my study

expanded this investigation by comparing T6SS HSI-III knockout strains to strains housing equivalent deletions in genes in the T6SS HSI I and II all of which displayed no attenuation at killing *G. mellonella*. This confirms the HSI-I and II are not required for / are not involved in pathogenicity against eukaryotic cells, while providing further evidence that the T6SS HSI-III is a secretion system required by PA01 for virulence against eukaryotic cells, most likely required during chronic infections when the T3SS is inactive.

Chapter 6:

Discussion, future work and concluding remarks

6. Discussion

6.1 Summary of results

Based on the combined results presented in Chapters 3 and 4, I believe it is reasonable to conclude that PA01 does indeed possess Mip genes (that we have called *pamip1*, *pamip2* and *pamip3*), and that these are required, in some part, for PA01 virulence. Furthermore, it is evident that these PaMips exert their virulence effects in a pleiotropic manner based on results presented in both the *in vitro* virulence assays and, more importantly, *in vivo* infection models. In some capacity the *in vitro* phenotypes were able to be complemented; however, the complementation was construct dependant (Table 4.2). In contrast, and most importantly, the attenuated killing of *G. mellonella* by strains I and J was restored to that of the WT upon genetic complementation. These data, coupled with the RNA seq data described in Chapter 3 confirming no change in expression of genes surrounding PaMip gene knockouts ensuring there were no downstream deleterious effects as a consequence of PaMip gene deletions, along with whole-genome sequencing data ensuring no genomic changes occurred in my version of strain I and J, it is conclusive PA01 possess Mip genes required for the full virulence of PA01. Thus, the work presented in this thesis is the first instance whereby a front line, clinical ESKAPE pathogen, which is placed at the top of the Who Health Organisation's critical list for requiring new antimicrobial therapies, has been found to possess Mip genes.

Several published reports regarding the involvement of Mips in the virulence of other gram-negative pathogens have shown their activity can be inhibited through the use of small molecules including rapamycin, FK506 and other simpler synthetic compounds inspired by these complex natural products. By exploration, I suggest that the PaMip's are thus promising targets for the development of future anti-virulence drugs. The importance of identifying new potential drug targets in the field of Microbiology is becoming increasingly paramount. In other gram-negative organisms possessing this virulence-associated subgroup of FKPBs, Mips, that were capable of performing PPIase activity, when assayed in the presence of FK506 or rapamycin, exhibited a significant reduction in virulence *in vivo*. However, many of these studies did not further investigate the mechanisms through which Mips may exert their virulence effects. In this study not only have I identified that PA01 possesses Mips required for virulence *in vivo*, but I have also begun to further our understanding into their pleiotropic nature and to dissect how these important virulence factors may function by using *in vitro* assays and RNA sequencing data.

From the analysis of RNA seq data, the most striking and significant observation was that for a cluster of genes encoding the T6SS HSI-III; the majority of the HSI-III genes were highly unregulated in both strains I and J, with 11 of these amongst the 20 most highly unregulated

genes in strain I. As discussed in Chapter 5, the phenotypes observed for mutants carrying deletions in key T6SS HSI-III genes suggest it is important for virulence, presenting similar phenotypes to those of strains I and J, both of which exhibited a significant downregulation of this T6SS HSI-III. This suggests a link between PaMips and the T6SS HSI-III, and that some of the phenotypes observed for strains I and J could be attributed to the downregulation of the HSI-III.

Given the T6SS HSI-III is expressed upon the transition to the sessile lifestyle that is associated with chronic infections, and when combined with the phenotypes observed in strains I and J which suggest an attenuation in this transition (lack of biofilm formation, expression of type IV pili and twitching motility), it is possible that PaMips may be involved in the lifestyle transitions, without which, PA01 cells remain planktonic as evidenced by the hyper-swarming motility of strains I and J. Furthermore, based on similar observations made for T6SS HSI-III knockout strains and PaMip strains I and J, in the *G. mellonella* model, I propose that the T6SS HSI-III is the system expressed during the sessile lifestyle and chronic infections, when the T3SS is inactive, and as such is responsible for the delivery of toxins and effectors intro eukaryotic cells.

As described in Chapter 3 based on observed phenotypes and RNA seq data, I propose PaMips may mediate this transition by several means. Firstly, given the PPIase activity of PaMip2 and PaMip3 (predicted from sequence analysis), it is possible they are involved in the correct translation of proteins and/or in their correct folding, or in other posttranslational processes. More specifically, if, for example, a transcription factor protein, is not correctly folded the promotion / repression of its target gene(s) will be affected. An example of this from our data is for AlgT, a transcriptional regulator of alginate and psl biosynthesis genes. The transcription of *algT* is transcribed to WT levels, however, the alginate and psl biosynthesis genes are significantly downregulated, suggesting there may be an issue with the transcriptional factor AlgT. These might include, for example, incorrect protein folding, due to the lack of PaMip's, meaning AlgT will not bind to the promoters of the alginate and psl biosynthesis genes, resulting in their observed down regulation (Figure 6.1). Alternatively, the PaMips may be involved in the correct folding of proteins directly involved in virulence mechanisms. For example, the genes responsible for structures such as the type IV pili and flagella are transcribed at levels equivalent to the WT in strains I and J. However, such structures are either physically missing or their production has been impaired suggesting a translational issue. Without the type IV pili structures, twitching motility is lost, along with further consequences such as impaired biofilm formation due to the inability of cells to perform type IV pili mediated attachment to a solid surface (Figure 6.2). Similarly, the gene responsible for haemolysis, *plcH,* remains transcribed to WT levels, however haemolytic activity is lost in strains I and J.

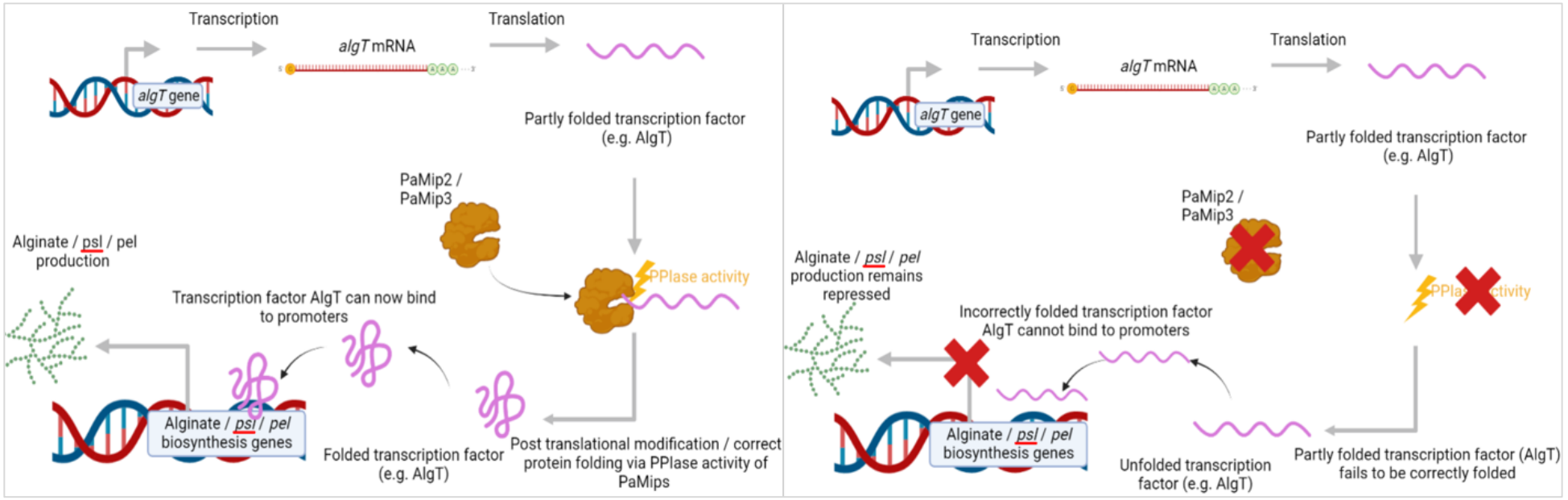


Figure 6.1 Diagram illustrating the proposed mechanism by which PaMips are involved in virulence. In the left image, the *algT* transcription factor encoding gene, is transcribed in strains I and J at levels comparable of the WT strain (Table 3.5). The resulting mRNA is translated, however the resulting protein requires further post translational modification to become a functional AlgT transcription factor protein. In this case the post translational modification comes in the form of PPIase activity from a PaMip. This modification allows AlgT to become fully functional and bind directly to DNA and promote the transcription of genes involved in alginate/ psl / pel biosynthesis. The image on the right describes the same process however in strain I or J which lack PaMips. Without the PaMips the partly folded AlgT will not undergo further post translational modification. Without correct folding, AlgT cannot fulfil its role as a transcription factor and will no longer bind DNA and therefore cannot promote the expression of alginate / psl / pel biosynthesis genes. Without the expression of these genes biofilm formation is decreased, a phenotype observed in both strains I and J.

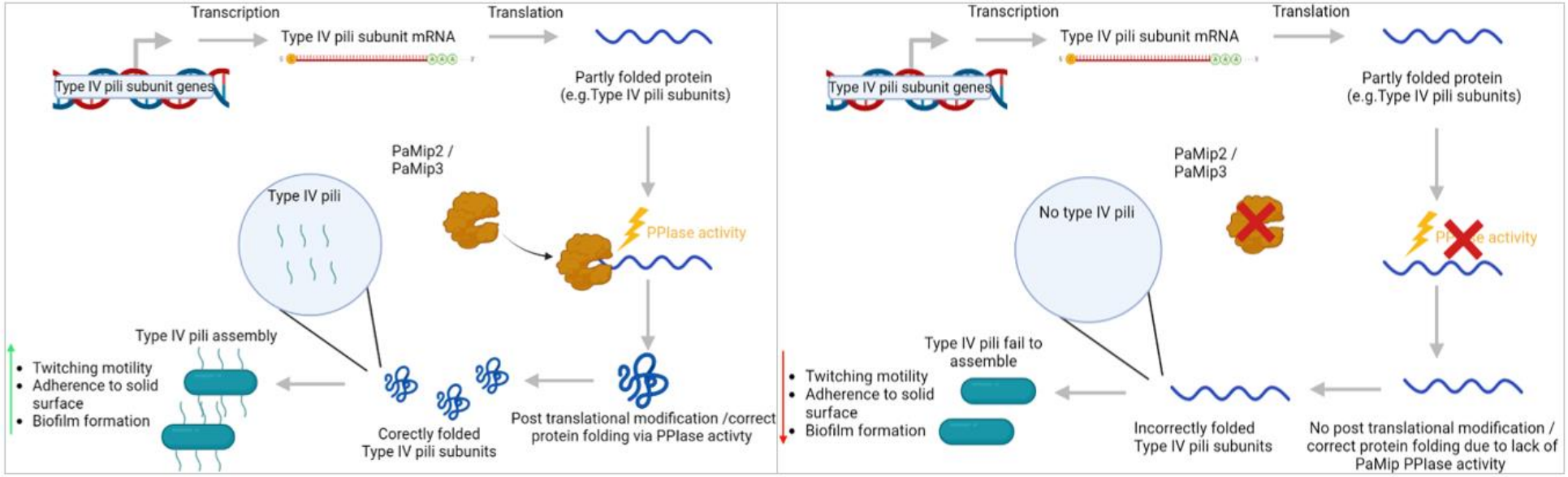


Figure 6.2 Diagram illustrating a second proposed mechanism by which PaMips are involved in virulence. In the left image, transcription of type IV pili biosynthesis genes are transcribed in strains I and J to levels comparable of the WT strain into mRNA (Table 3.5). The mRNA is translated and resulting proteins partly folded. To become fully folded type IV pili subunits capable of assembling into type IV pili appendages, the subunits must be folded further and post translationally modified courtesy of PaMip PPIase activity. Once correctly folded and modified, the type IV pili subunits are then able to assemble and to form the type IV pili structures responsible for surface attachment, initiating biofilm formation and twitching motility. The image on the right is the same process however in strain I or J, which lack PaMips. Without the PaMips the partly folded type IV pili subunits will not undergo further post translational modification and will remain only partly folded. Without correct folding, the subunits are not able to assemble resulting in no type IV pili appendages on the cell surfaces, as was seen in the SEM images for strains I and J. Without the presence of these appendages, cell surface attachment is not initiated resulting in attenuation in biofilm formation and twitching motility lost.

Based on its sequence, PaMip1 is unlikely to perform PPIase activity as it is missing two essential amino acid residues in its active site; a similar observation has been made for the *B. pseudomelli* Mip BPSL0918 which also lacks these key residues and does not display PPIase activity [177]. Both PaMip1 and the BPSL0918 Mip, have been shown to be involved in virulence. It is likely PaMip1 exerts its virulence effects in the same manner proposed for the BPSL0918 Mip as a chaperone involved in the maturation of outer membrane proteins, much like a SurA mutant *E. coli* strain which lacked PPIase activity [269]. PaMip2 is predicted to possess PPIase activity, and to be localised to the cytoplasm, and it is likely involved in virulence by mediating the correct folding of virulence factors. One virulence factor that requires a specific proline residue to be in the correct isomeric form (*trans*) to function is DsbA (whose function will be discussed shortly). DsbA requires a PPIase to perform this isomerisation, without which DsbA is unable to form disulphide bonds in the tertiary structure of virulence proteins [270-272]. It is reasonable to postulate that PaMip2 may perform such PPIase activity on virulence related client proteins that possess a proline in their structure requiring isomerisation. Finally, given PaMip3 is predicted to be localised to the outer membrane, it likely exerts its virulence role in a similar fashion to its LgMip homologue which was found to interact with collagen IV in the lung and initiate interaction with a target host cell [158]. Thus, PaMip3 may contribute to PA01's ability to migrate across cell membranes, this will be discussed in more detail in the following section. In all, each PaMip is required for virulence and likely plays their own unique role in virulence.

6.2 Future work

6.2.1 Short term aims and experiments

The most immediate experiments will focus on the macrophage infectivity assays using the complemented Mip mutant strains. Given complementation of strains I and J was successful in the *G. mellonella* infection model (**4.5.5**), assessing these strains in macrophages assays would be extremely useful. One would expect the complemented strains to have restored the ability to enter and colonise macrophage cells given the results obtained in the *G. mellonella* model. Out of all the complement strains to assess, strain F (Δ*pamip3*), would be the most intriguing to analyse, given that in the macrophage assay strain F was the only single mutant to display significant attenuation compared to the WT (Figure 3.18). The *pamip3* gene is the only PaMip predicted to localise on the outer membrane, and such is the closest homologue of the LgMip. LgMip, as discussed in the introduction (**1.12.5)**, is required for the interaction with collagen IV to mediate *L. pneumophilia'*s transmigration across the lung epithelial cell barrier via the C-domain of the LgMip protein [158]. Given their homology, and the result observed for strain F in the macrophage assay, the virulence role of PaMip3 may exert can potentially be explained by the extrapolation of the LgMip mechanism. To assess this hypothesis, complementing strain

F with *pamip3* (as well as strain J which also houses a *pamip3* deletion) should restore its ability to enter and colonise macrophages.

To further use the macrophage model, we could assess whether the addition of PPIase activity inhibitors such as FK506 or rapamycin into the assay hinders the WT or PaMip knockout strains ability to colonise macrophages. This is a straightforward assay that only requires an extra 30-minute incubation step to be incorporated into the methodology. Given in literature reports that the addition of either FK506 or rapamycin reduced the virulence of a gram-negative bacteria (housing a Mip protein) in an infection model, replicating this with our PA01 Mip mutant strains should provide similar results [178]. For the WT strain we would expect to see a slight decrease in virulence and a reduced number of cells recovered from the macrophages. Where we might expect to see the most significant decrease in recovered cells would be in the single PaMip knockout strains D and E. Under normal conditions, strains D and E are able to enter and colonise macrophages to comparable levels as the WT. Co-incubating these strains in the presence of FK506 or rapamycin should theoretically inhibit the remaining PaMips expressed resulting in the phenotypes of D and E shifting to those more similar to strains I and J.

Preliminary data that has not been discussed in this thesis has shown that in the presence of rapamycin, the WT strain exhibits a 43% reduction in macrophage infectivity, while single deletion strains D and E, which normally infect macrophages to comparable levels of the WT, also exhibit a reduction in infectivity of macrophages in the presence of rapamycin (27.2% and 55% respectively ($P<0.05$, $n=10$)). Interestingly, this reduction in the presence of rapamycin for the WT, D and E strains mimics the decrease in macrophage infectivity observed for strain F (57% ($p<0.05$, $n=10$)) compared to the WT under untreated conditions. Strain F houses a deletion in PaMip3, which is the only PaMip predicted to be localised to the outer membrane and is homologous to the LgMip, which known to facilitate interaction and migration across target cells and collagen IV. This result suggests that PaMip3 is required for full infectivity of macrophages and that it can be inhibited pharmacologically to reduce PA01 infectivity of macrophages. This experiment supports the idea that PaMips are an attractive target for the development of anti-virulence drugs. Moreover, it would be extremely interesting to study the importance of PaMips in a mammalian host, for example using a mouse infection model.

So far, the characteristic PPIase activity known for FKBPS has been assumed for our PaMips based on the sequence alignments to other known FKBPs and Mips with PPIase activity. These sequence alignments revealed that PaMip2 and PaMip3 both possess the nine residues involved in PPIase activity (Appendix image 1.4), two of which (aspartate and tyrosine) are absolutely essential for PPIase activity to take place [97]. PaMip1 possess neither of these residues suggesting it would not perform PPIase activity, like the BSPL0918 Mip homologue which also

lacks these residues. In work not discussed in this thesis, these three PaMips were all purified to conduct a PPIase assay. Initial attempts to capture the *cis-trans* isomerisation through UV absorbance assays failed. This was, in part, due to sub-optimal equipment with which to measure this assay as this reaction happens almost instantly and requires auto-injectors. This PPIase assay is conducted by using α-chymotrypsin to catalyse the cleave of the Suc-ala-phe-pro-phe-4-nitroaniline; this results in liberation of the 4-nitro group allowing a measurable UV reading to be taken. However, α-chymotrypsin can only perform this cleavage reaction when the pro-phe bond is in the *trans* configuration, not the *cis*. To assess whether our PaMips possess PPIase activity, there should have been a visible increase in fluorescence with an increased level of isomerisation occurring, with α-chymotrypsin able to cleave 4-nitroaniline at an increased rate. Carrying out this assay led to inconclusive results as mixing the substrate, protein and α-chymotrypsin was done manually and this reaction occurred too rapidly for me to measure manually. Recently, the Norwich Research Park acquired a Varioskan LUX multimode microplate reader that possesses auto-injectors to start the reactions that begins taking measurements instantly. This equipment will provide an excellent platform to repeat the PPIase activity assays for the PaMips. Uncovering whether PaMips possesses PPIase activity will help further solidify the potential mechanisms by which PaMips exert their pleiotropic effects on virulence as proposed above and illustrated in Figures 3.22 and 3.23.

6.2.2 Long term aims and future work

With this work highlighting that PaMips exert their effect on virulence in a pleiotropic manner, uncovering exactly how this is achieved remains a major and exciting unanswered question. Numerous literature reports describe how Mips are required for a number of gram-negative pathogens to exert their virulence. However, there have been no concrete studies confirming the molecular basis for how they achieve this. Several studies have begun to elucidate and demonstrate that for the majority of Mips the active site must remain intact, as when they are inhibited by FK506 or rapamycin, virulence is attenuated. Exactly why this active site is required, has not yet been revealed except for LgMip when mediating interaction with collagen IV. As hypothesised above, they may exert their virulence effects through a particular set of virulence-associated client proteins which may possess a proline in their secondary structure that requires isomerisation, and without which the protein remains partially / incorrectly folded or unable to function correctly, or even failing to become phosphorylated. As briefly mentioned above, one such example of a virulence protein in bacteria known to require a proline amide bond in a specific orientation is DsbA. As mentioned previously (**3.8**), DsbA forms disulphide bonds between two amino acid side chains in proteins. If a specific proline residue in the secondary structure of DsbA is not converted from the *cis* to *trans* isomer, DsbA is unable to form disulphide

bonds in the tertiary structure of a protein. If our PaMips interact with proteins such as DsbA, a series of different experiments can be performed, including a mass spectrometry (MS) based proteomics approach, a co-immunoprecipitation pull-down assay (Co-IP), and the use of a bacterial-adenylate cyclase two hybrid (BACTH) library. Following these global approaches, depending on what interacting targets are identified, specific functional assays can be performed such as a DsbA functional assay, with and without the presence of a cognate PaMip; this could be extended by labelling DsbA with C^{13} or N^{15} labelled proline and tracking the effects of co-incubation with a PaMip using NMR.

Taking a mass spectrometry (MS) based proteomics approach would help verify the hypothesis described in Figures 6.1 and 6.2 for the action of PaMips. In a similar analysis to that of the RNA seq, performing a proteomics approach would elucidate whether virulence related proteins are quantifiable to levels comparable to the WT in strains I and J and whether they have the correct structural modifications to become functional proteins. For example, from our RNA seq data, we know the genes required for type IV pili biosynthesis are upregulated. However, from our twitching motility assay and SEM images we know the type IV pili structures are missing. As hypothesised, these type IV pili structures may fail to assemble as a result of incorrect protein folding or lack of chaperone activity in our PaMip knockout strains. The proteomics platform would help elucidate whether these type IV pili biosynthesis proteins are present at quantifiable levels and how that relates to the WT strain, and also whether they possess the required post-translational modifications (folding, phosphorylation etc.) to be able to function. It is here that we would most likely observe changes in virulence proteins that have been transcribed to WT levels in strains I and J, but which lack activity or are unable to fulfil their biological function due to the reduction in post-translation modification / chaperoning activity as a consequence of PaMip deletions, resulting in the loss of *in vitro* virulence phenotypes. This approach would further give us an idea of what client proteins are likely assocaited with PaMips [273].

In tandem, using a Co-IP would help further illuminate potential client proteins that PaMips interact with and how they are involved in virulence. By flag-tagging our PaMips and growing the strain overnight, our PaMips should be associated with proteins they naturally interact with. The following day the cells would be lysed, the lysate would then be run down a specifically prepared solid immobilised support (magnetic beads) modified with flag-tag specific antibodies; the flag-tag antibody would then capture the PaMip, along with any other proteins that have formed a complex with them. Once bound to the antibody on the column, the PaMip-protein complex should remain bound following wash steps. The resulting complex could then be given to the Proteomics platform to perform mass spectrometry and elucidate the identity of any proteins that PaMips interact with [274]. This should hopefully further aid us in understanding

how PaMips exert their virulence effects and what their client proteins are. In previous work not discussed in this thesis (performed by Dr Eleftheria Trampari,) this Co-IP procedure was performed once, with preliminary results suggesting that PaMip2 forms a complex with DsbA. Given what has been described about DsbA in this study, and its known role in virulence, it provides a clue as to how PaMips exert their virulence effects.

The final approach to identify proteins that interact with PaMips involves using a bacterial-adenylate cyclase two hybrid library (BACTH). This library has a guaranteed 40X genomic coverage/sample and has already been constructed for the lab by the Canadian company BioS&T. This library can be used to uncover protein-protein interactions using features from the adenylate cyclase toxin in *Bordetella pertussis*. This adenylate cyclase has two distinct domains, T25 and T18 and is required for the synthesis of cyclic adenosine monophosphate (cAMP), an important secondary messenger required in multiple biological processes. For the cyclase to function both domains must be present to catalyse the conversion of ATP to cAMP, which is then able to bind CAP (**C**atabolite **A**ctivator **P**rotein), a transcriptional regulator whose affinity for DNA increases when bound to cAMP allowing the transcription of genes, in this case reporter genes such as *lacZ* (Figure 6.3).

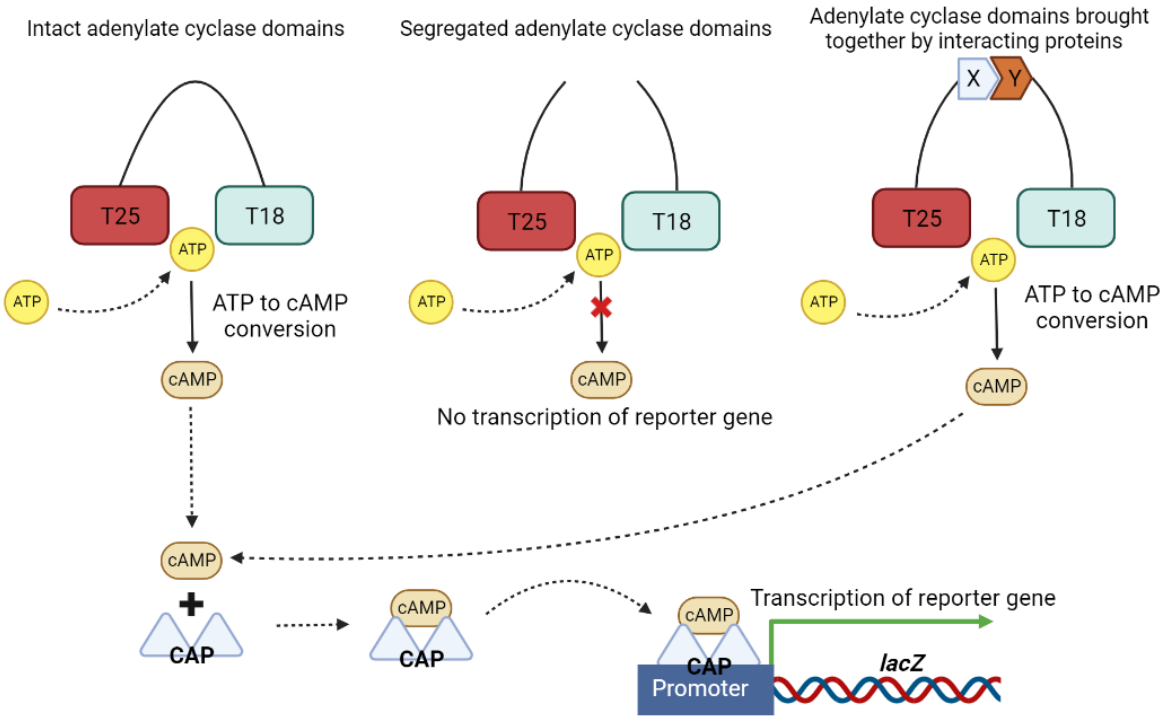


Figure 6.3. Principle of the bacterial adenylate two-hybrid system (BACTH). Ordinarily, the two adenylate cyclase domains T25 (red) and T18 (green) are intact and linked which catalyses the conversion of ATP to cAMP by the displacement of pyrophosphate. cAMP binds to CAP (**C**atabolite **A**ctivator **P**rotein) creating a complex. This complex promotes the binding of CAP to DNA, promoting the transcription of certain reporter genes, in this case, *lacZ*. When the adenylate cyclase domains T25 and T18 are segregated, the ability to convert ATP to cAMP is lost. This ability to convert ATP to cAMP is restored using the BACTH system whereby a 'bait' protein (here labelled 'X' representing PaMips) is linked to domain T25, interacts with an unknown protein, labelled 'Y', linked to domain T18. This brings the T25 and T18 domains together in close proximity to create a functional adenylate cyclase which in turn converts ATP to cAMP, ultimately allowing the transcription of reporter genes leading to characteristic phenotypes on indicator plates (Figure 6.4). The restoration of functional adenylate cyclase suggests that proteins 'X' and 'Y' interact, this can elude towards what a bait protein of interest (here PaMips) may interact with in a PA01 cell.

These adenylate cyclase domains can be used for the study of protein-protein interactions by cloning either domain into a plasmid (pKT25 or pUT18C) into which PaMips are then cloned as 'bait' proteins and are linked to either the T25 or T18 domain. BioS&T constructed a plasmid library of PA01 genes by shearing PA01 genomic DNA by sonication and isolating fragment sizes of 500-2000 bp. These fragments were cloned into either the pKT25 or pUT18C plasmids which, respectively, possess either the T25 or T18 domain of adenylate cyclase. This cloning process, in theory, creates two plasmid libraries capturing every gene in the PA01 genome linked to either the T18 or T25 domains. Given the 40X genomic coverage, essentially every gene and its protein product in the genome should be captured in these plasmids [275]. The 'bait' vectors (pKT25 containing either; *pamip1 / pamip2 / pamip3* and pUT18C containing either *pamip1 / pamip2 / pamip3*) are transformed into BTH101, a chemically competent *E.coli* strain. Once transformed with a bait vector, the BTH101 cells are prepared to become electrocompetent which allows the cells to be transformed with either the pKT25 or pUT18C plasmid library (Figure 6.4),

transformants are then plated on indicator plates where characteristic phenotypes can be observed. This BACTH system takes advantage of the BTH101 *E. coli* strain which lacks a functional adenylate cyclase, thus loses the ability to produce cAMP, which in turn, affects the expression of many genes under the control of the *lac* or *mal* operon. The synthesis of cAMP is restored when a functional adenylate cyclase is formed when a 'bait' vector housing a PaMip linked to either the T25 or T18 adenylate cyclase domain interacts with an unknown PA01 protein housed on a vector from the PA01 plasmid library linked to the missing adenylate cyclase domain (T18 or T25 respectively). With the restoration of cAMP synthesis, CAP is able to bind DNA and promote the transcription of reporter genes such as *lacZ*. The transcription of *lacz* results in β-galactosidase activity which can be visualised on reporter plates supplemented with IPTG whereby colonies turn blue indicating functional adenylate cyclase has been restored, indicating a PaMip-protein interaction. These colonies can be picked, subjected to colony PCR using T18C or T25-specific primers and the resultant bands sent for sequencing to uncover the sequence of the associated gene. This can help us to unpick protein-protein interactions and elucidate what PaMips may interact with ordinarily in PA01. Understanding PaMip-protein interactions will help us determine what client proteins they may possess and how they exert their role in PA01 virulence.

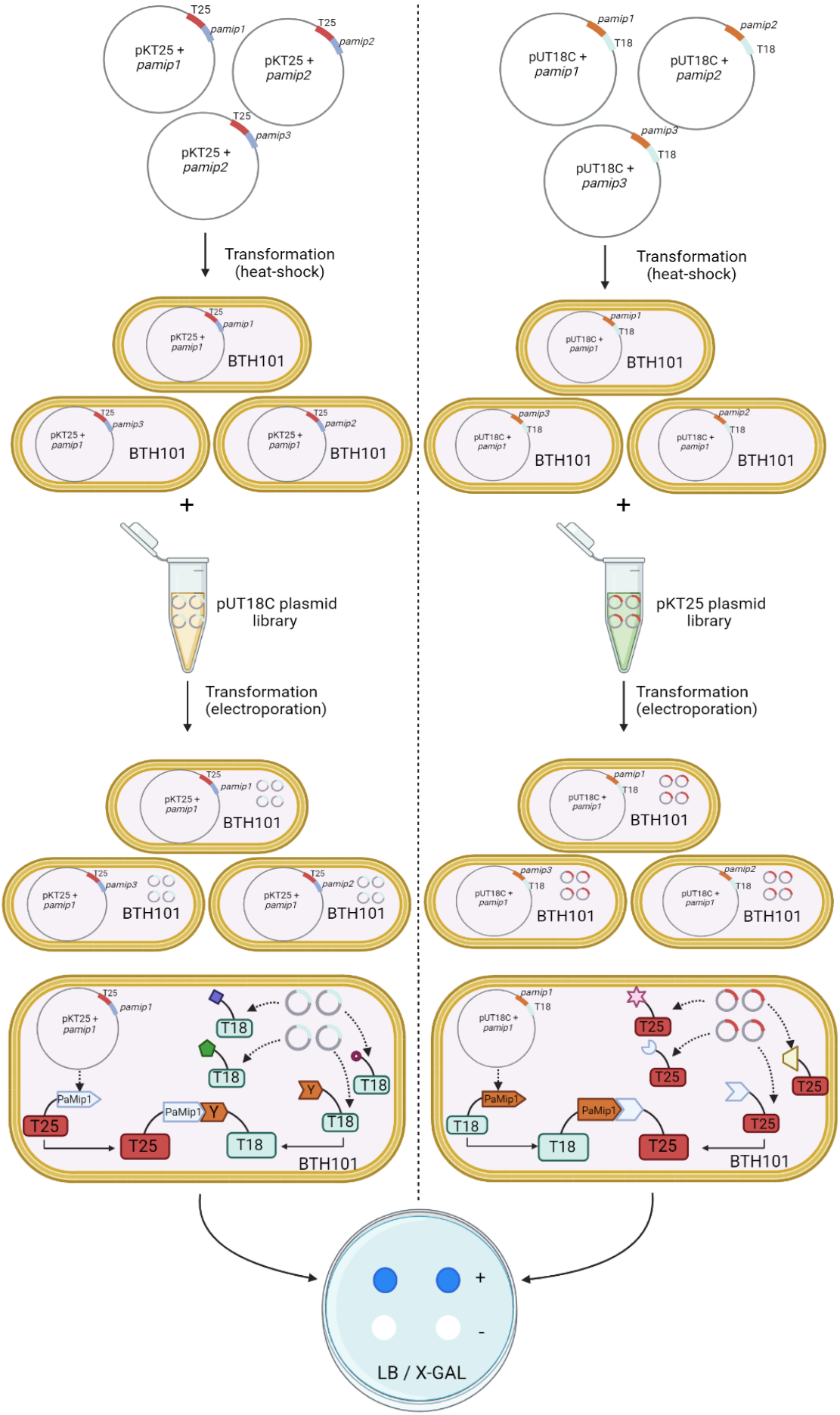

T25
pamip1
pKT25 + pamip1
T25
pamip2
pKT25 + pamip2
T25
pamip3
pKT25 + pamip2
pamip1
T18
pUT18C + pamip1
pamip2
T18
pUT18C + pamip2
pamip3
T18
pUT18C + pamip3
Transformation (heat-shock)
Transformation (heat-shock)
BTH101
pUT18C plasmid library
pKT25 plasmid library
Transformation (electroporation)
Transformation (electroporation)
PaMip1
Y
LB / X-GAL
+
-

Figure 6.4. Stepwise overview of the bacterial two-hybrid procedure. The genes, and their product of interest (*pamip1*, *pamip2*, *pamip3*) are cloned into the bait vectors pKT25 and pUT18C and transformed via heat-shock into BTH101 cells which lack adenylate cyclase activity and are unable to synthesise cAMP. Genes cloned into the pKT25 vector are linked to the T25 domain of adenylate cyclase, whereas genes cloned into the pUT18C vector are linked to the T18 domain of adenylate cyclase. These cells now housing either pKT25 + *pamip1* / *pamip2*/ *pamip3* and pUT18C + *pamip1* / *pamip2* / *pamip3* are prepared to become electrocompetent. Once prepared cells housing the pKT25 + *pamip1* / *pamip2*/ *pamip3* are co-transformed with a small volume of the pUT18C PA01 plasmid library (prepared by BioS&T) which should capture and cover every gene in the PA01 genome which each linked to the T18 domain of adenylate cyclase. Cells housing the pUT18C + *pamip1* / *pamip2*/ *pamip3* are co-transformed with a small volume of the pKT25 PA01 plasmid library (prepared by BioS&T) which should capture every gene in the PA01 genome with each linked to the T25 domain of adenylate cyclase. At this point in the cells, the PaMip bait vector will express our PaMip protein fused to either T25 orT18 of adenylate cyclase, while the plasmid library transformed into the cell should theoretically express all the genes in the PA01 genome fused to either T18 or T25, respectively. As seen in the diagram, the PaMip protein has a specific 'shape' that will only bind a protein of a certain 'shape' that ordinarily it would interact with, this interaction with the unknown protein will bring the T25 / T18 domain into close contact with its missing adenylate cyclase counterpart T18 / T25 respectively, returning the synthesis of cAMP. This restoration of cAMP synthesis can be visualised by the cells being plated on reporter plates e.g. supplemented with IPTG. Any PaMip-protein interactions can be observed on the plates indicated by blue colonies as a result of the restoration in adenylate cyclase activity. Any white colonies on the plates are negative and no positive PaMip-protein interactions have occurred. These colonies can be taken forward and used to uncover proteins that PaMips interact with and aid in the elucidation of how they exert their virulence roles.

6.3 Closing remarks

This work here highlights that PA01 possess Mip genes required for virulence. Moreover, this is the first instance where an ESKAPE pathogen has been evidenced to possess this group of virulence related proteins. The preliminary data using rapamycin in macrophage infectivity assays provides evidence that these PaMips are inhibitable, and, when inhibited, exhibit a drastic reduction in virulence. The significance of this work cannot be understated and has the potential to provide a basis for new ways to tackle the increasing antimicrobial resistance threat posed by gram-negative pathogens. Given several gram-negative pathogens have been evidenced to possess Mip proteins, an anti-virulence drug targeting approach might be applicable beyond PA01. Interestingly, upon further bioinformatic analysis, other clinical strains of *P. aeruginosa*, including the highly virulent PA14, possess identical sequences for *pamip1*, *pamip2* and *pamip3*.

Uncovering the set of virulence-related client proteins associated with PaMips will help us to elucidate how PaMips exert their effect on virulence in PA01, further solidify their candidacy for the development of anti-virulence drugs. Understanding the mechanisms by which PaMips exert their virulence will provide an understanding of the cellular processes that would be attenuated in the presence of an anti-virulence drug, which in turn would allow us to assess how this would affect the progression of a PA01 infection in a patient, improving outcomes (e.g. by preventing biofilm formation it would significantly reduce the chance of a chronic infection developing).

Furthermore, we aim to be the first group to uncover the exact molecular basis by which PaMips are involved in PA01 virulence. I believe targeting bacterial virulence is an important area of research that will become increasingly studied in the coming years as the issues associated with anti-microbial resistance increase, and the stagnation in the discovery and development of new antibiotics.

Bibliography

Bibliography

1. Ho, Brian T., Tao G. Dong, and John J. Mekalanos, *A View to a Kill: The Bacterial Type VI Secretion System.* Cell Host & Microbe, 2014. **15**(1): p. 9-21.
2. Koonin, E.V. and Y.I. Wolf, *Genomics of bacteria and archaea: the emerging dynamic view of the prokaryotic world.* Nucleic Acids Res, 2008. **36**(21): p. 6688-719.
3. Romalde, J.L., S. Balboa, and A. Ventosa, *Editorial: Microbial Taxonomy, Phylogeny and Biodiversity.* Frontiers in Microbiology, 2019. **10**.
4. Bergan, T., *1 Classification of Enterobacteriaceae*, in *Methods in Microbiology*, T. Bergan, Editor. 1984, Academic Press. p. 1-41.
5. Coico, R., *Gram staining.* Curr Protoc Microbiol, 2005. **Appendix 3**: p. Appendix 3C.
6. Silhavy, T.J., D. Kahne, and S. Walker, *The bacterial cell envelope.* Cold Spring Harb Perspect Biol, 2010. **2**(5): p. a000414.
7. Becerra, S.C., et al., *An optimized staining technique for the detection of Gram positive and Gram negative bacteria within tissue.* BMC Research Notes, 2016. **9**(1): p. 216.
8. Clarridge, J.E., 3rd, *Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases.* Clin Microbiol Rev, 2004. **17**(4): p. 840-62, table of contents.
9. Fernández, M., et al., *Analysis of the pathogenic potential of nosocomial Pseudomonas putida strains.* Frontiers in Microbiology, 2015. **6**.
10. Iwashita, Y., et al., *Arbekacin treatment of a patient infected with a Pseudomonas putida producing a metallo-beta-lactamase.* J Intensive Care, 2013. **1**(1): p. 3.
11. Hotta, Y., et al., *Classification of Genus Pseudomonas by MALDI-TOF MS Based on Ribosomal Protein Coding in S10–spc–alpha Operon at Strain Level.* Journal of Proteome Research, 2010. **9**(12): p. 6722-6728.
12. Gomila, M., et al., *Phylogenomics and systematics in Pseudomonas.* Frontiers in Microbiology, 2015. **6**.
13. Moore, E.R.B., et al., *The Determination and Comparison of the 16S rRNA Gene Sequences of Species of the Genus Pseudomonas (sensu stricto and Estimation of the Natural Intrageneric Relationships.* Systematic and Applied Microbiology, 1996. **19**(4): p. 478-492.
14. Yabuuchi, E., et al., *Proposal of Burkholderia gen. nov. and Transfer of Seven Species of the Genus Pseudomonas Homology Group II to the New Genus, with the Type Species Burkholderia cepacia (Palleroni and Holmes 1981) comb. nov.* Microbiology and Immunology, 1992. **36**(12): p. 1251-1275.
15. Mulet, M., et al., *Concordance between whole-cell matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry and multilocus sequence analysis approaches in species discrimination within the genus Pseudomonas.* Systematic and Applied Microbiology, 2012. **35**(7): p. 455-464.
16. Mulet, M., E. García-Valdés, and J. Lalucat, *Phylogenetic affiliation of Pseudomonas putida biovar A and B strains.* Research in Microbiology, 2013. **164**(4): p. 351-359.
17. Silby, M.W., et al., *Pseudomonas genomes: diverse and adaptable.* FEMS Microbiology Reviews, 2011. **35**(4): p. 652-680.
18. Spiers, A.J., A. Buckling, and P.B. Rainey, *The causes of Pseudomonas diversity.* Microbiology (Reading), 2000. **146 (Pt 10)**: p. 2345-2350.
19. Haas, D. and G. Défago, *Biological control of soil-borne pathogens by fluorescent pseudomonads.* Nature Reviews Microbiology, 2005. **3**(4): p. 307-319.
20. Arnold, D.L. and G.M. Preston, *Pseudomonas syringae: enterprising epiphyte and stealthy parasite.* Microbiology, 2019. **165**(3): p. 251-253.
21. Pereira, C., et al., *Kiwifruit bacterial canker: an integrative view focused on biocontrol strategies.* Planta, 2021. **253**(2): p. 49.
22. Xin, X.F., B. Kvitko, and S.Y. He, *Pseudomonas syringae: what it takes to be a pathogen.* Nat Rev Microbiol, 2018. **16**(5): p. 316-328.

23. Hou, S., et al., *Damage-Associated Molecular Pattern-Triggered Immunity in Plants.* Frontiers in Plant Science, 2019. **10**.
24. Valencia-Botín, A.J. and M.E. Cisneros-López, *A Review of the Studies and Interactions of Pseudomonas syringae Pathovars on Wheat.* International Journal of Agronomy, 2012. **2012**: p. 692350.
25. Wakelin, S.A., et al., *Cost of root disease on white clover growth in New Zealand dairy pastures.* Australasian Plant Pathology, 2016. **45**(3): p. 289-296.
26. Sun, D., et al., *Identification of a Pseudomonas putida as biocontrol agent for tomato bacterial wilt disease.* Biological Control, 2017. **114**: p. 45-50.
27. Agisha, V.N., et al., *Broad-spectrum antimicrobial activity of volatile organic compounds from endophytic Pseudomonas putida BP25 against diverse plant pathogens.* Biocontrol Science and Technology, 2019. **29**(11): p. 1069-1089.
28. Molina, L., et al., *Chapter Four - The versatility of Pseudomonas putida in the rhizosphere environment*, in *Advances in Applied Microbiology*, G.M. Gadd and S. Sariaslani, Editors. 2020, Academic Press. p. 149-180.
29. Behnsen, J. and M. Raffatellu, *Siderophores: More than Stealing Iron.* mBio, 2016. **7**(6): p. e01906-16.
30. Shimao, M., *Biodegradation of plastics.* Current Opinion in Biotechnology, 2001. **12**(3): p. 242-247.
31. Saminathan, P., et al., *Biodegradation of plastics by Pseudomonas putida isolated from garden soil samples.* J Adv Bot Zool, 2014. **1**(3): p. 34-38.
32. Roohi, et al., *Microbial Enzymatic Degradation of Biodegradable Plastics.* Curr Pharm Biotechnol, 2017. **18**(5): p. 429-440.
33. Cook, T.B., et al., *Genetic tools for reliable gene expression and recombineering in Pseudomonas putida.* J Ind Microbiol Biotechnol, 2018. **45**(7): p. 517-527.
34. Mozejko-Ciesielska, J., *Chapter 10 - Pseudomonas putida–based cell factories*, in *Microbial Cell Factories Engineering for Production of Biomolecules*, V. Singh, Editor. 2021, Academic Press. p. 165-181.
35. Poblete-Castro, I., et al., *Industrial biotechnology of Pseudomonas putida and related species.* Appl Microbiol Biotechnol, 2012. **93**(6): p. 2279-90.
36. Horii, T., H. Muramatsu, and Y. Iinuma, *Mechanisms of resistance to fluoroquinolones and carbapenems in Pseudomonas putida.* Journal of Antimicrobial Chemotherapy, 2005. **56**(4): p. 643-647.
37. Fass, R.J., et al., *In vitro activities of quinolones, beta-lactams, tobramycin, and trimethoprim-sulfamethoxazole against nonfermentative gram-negative bacilli.* Antimicrob Agents Chemother, 1996. **40**(6): p. 1412-8.
38. Kim, S.E., et al., *Nosocomial Pseudomonas putida Bacteremia: High Rates of Carbapenem Resistance and Mortality.* Chonnam Med J, 2012. **48**(2): p. 91-5.
39. Hoffland, E., J. Hakulinen, and J.A. Van Pelt, *Comparison of systemic resistance induced by avirulent and nonpathogenic Pseudomonas species.* Phytopathology, 1996. **86**(7): p. 757-762.
40. Choudhary, D.K., A. Prakash, and B.N. Johri, *Induced systemic resistance (ISR) in plants: mechanism of action.* Indian journal of microbiology, 2007. **47**(4): p. 289-297.
41. Luján, A.M., P. Gómez, and A. Buckling, *Siderophore cooperation of the bacterium Pseudomonas fluorescens in soil.* Biology Letters, 2015. **11**(2): p. 20140934.
42. Ganeshan, G. and A. Manoj Kumar, *Pseudomonas fluorescens, a potential bacterial antagonist to control plant diseases.* Journal of Plant Interactions, 2005. **1**(3): p. 123-134.
43. Thomashow, L.S. and D.M. Weller, *Role of a phenazine antibiotic from Pseudomonas fluorescens in biological control of Gaeumannomyces graminis var. tritici.* J Bacteriol, 1988. **170**(8): p. 3499-508.

44. Thomashow, L.S., et al., *Production of the antibiotic phenazine-1-carboxylic Acid by fluorescent pseudomonas species in the rhizosphere of wheat.* Appl Environ Microbiol, 1990. **56**(4): p. 908-12.
45. Voisard, C., et al., *Cyanide production by Pseudomonas fluorescens helps suppress black root rot of tobacco under gnotobiotic conditions.* The EMBO journal, 1989. **8**(2): p. 351-358.
46. Biaggini, K., et al., *The pathogenic potential of Pseudomonas fluorescens MFN1032 on enterocytes can be modulated by serotonin, substance P and epinephrine.* Arch Microbiol, 2015. **197**(8): p. 983-90.
47. Donnarumma, G., et al., *Effect of Temperature on the Shift of Pseudomonas Fluorescens from an Environmental Microorganism to a Potential Human Pathogen.* International Journal of Immunopathology and Pharmacology, 2010. **23**(1): p. 227-234.
48. Gershman, M.D., et al., *Multistate Outbreak of Pseudomonas fluorescens Bloodstream Infection after Exposure to Contaminated Heparinized Saline Flush Prepared by a Compounding Pharmacy.* Clinical Infectious Diseases, 2008. **47**(11): p. 1372-1379.
49. Stover, C.K., et al., *Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen.* Nature, 2000. **406**(6799): p. 959-964.
50. Cloeckaert, A., M.S. Zygmunt, and B. Doublet, *Editorial: Genetics of Acquired Antimicrobial Resistance in Animal and Zoonotic Pathogens.* Front Microbiol, 2017. **8**: p. 2428.
51. Sun, E., et al., *Broad-Spectrum Adaptive Antibiotic Resistance Associated with Pseudomonas aeruginosa Mucin-Dependent Surfing Motility.* Antimicrob Agents Chemother, 2018. **62**(9).
52. Ojewumi, M.E., et al., *Bioremediation: Data on Pseudomonas aeruginosa effects on the bioremediation of crude oil polluted soil.* Data Brief, 2018. **19**: p. 101-113.
53. Lam, J., et al., *Genetic and Functional Diversity of Pseudomonas aeruginosa Lipopolysaccharide.* Frontiers in Microbiology, 2011. **2**.
54. Chevalier, S., et al., *Structure, function and regulation of Pseudomonas aeruginosa porins.* FEMS Microbiology Reviews, 2017. **41**(5): p. 698-722.
55. Richardot, C., et al., *Carbapenem resistance in cystic fibrosis strains of Pseudomonas aeruginosa as a result of amino acid substitutions in porin OprD.* International journal of antimicrobial agents, 2015. **45**(5): p. 529-532.
56. James, C.E., et al., *How beta-lactam antibiotics enter bacteria: a dialogue with the porins.* PLoS One, 2009. **4**(5): p. e5453.
57. Moradali, M.F., S. Ghods, and B.H.A. Rehm, *Pseudomonas aeruginosa Lifestyle: A Paradigm for Adaptation, Survival, and Persistence.* Frontiers in Cellular and Infection Microbiology, 2017. **7**.
58. Rampioni, G., et al., *Effect of efflux pump inhibition on Pseudomonas aeruginosa transcriptome and virulence.* Sci Rep, 2017. **7**(1): p. 11392.
59. Poole, K., *Pseudomonas aeruginosa: resistance to the max.* Front Microbiol, 2011. **2**: p. 65.
60. Guttenplan, S.B. and D.B. Kearns, *Regulation of flagellar motility during biofilm formation.* FEMS Microbiol Rev, 2013. **37**(6): p. 849-71.
61. Kearns, D.B., *A field guide to bacterial swarming motility.* Nature Reviews Microbiology, 2010. **8**(9): p. 634-644.
62. Feldman, M., et al., *Role of Flagella in Pathogenesis of Pseudomonas aeruginosa Pulmonary Infection.* Infection and Immunity, 1998. **66**(1): p. 43-51.
63. Craig, L., M.E. Pique, and J.A. Tainer, *Type IV pilus structure and bacterial pathogenicity.* Nature Reviews Microbiology, 2004. **2**(5): p. 363-378.
64. Barken, K.B., et al., *Roles of type IV pili, flagellum-mediated motility and extracellular DNA in the formation of mature multicellular structures in Pseudomonas aeruginosa biofilms.* Environmental Microbiology, 2008. **10**(9): p. 2331-2343.

65. Hrabak, E.M. and D.K. Willis, *The lemA gene required for pathogenicity of Pseudomonas syringae pv. syringae on bean is a member of a family of two-component regulators.* J Bacteriol, 1992. **174**(9): p. 3011-20.
66. Mulcahy, H., et al., *Pseudomonas aeruginosa RsmA plays an important role during murine infection by influencing colonization, virulence, persistence, and pulmonary inflammation.* Infect Immun, 2008. **76**(2): p. 632-8.
67. Jimenez, P.N., et al., *The multiple signaling systems regulating virulence in Pseudomonas aeruginosa.* Microbiol Mol Biol Rev, 2012. **76**(1): p. 46-65.
68. Goodman, A.L., et al., *Direct interaction between sensor kinase proteins mediates acute and chronic disease phenotypes in a bacterial pathogen.* Genes Dev, 2009. **23**(2): p. 249-59.
69. Fuqua, C. and E.P. Greenberg, *Self perception in bacteria: quorum sensing with acylated homoserine lactones.* Curr Opin Microbiol, 1998. **1**(2): p. 183-9.
70. Fuqua, W.C., S.C. Winans, and E.P. Greenberg, *Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators.* J Bacteriol, 1994. **176**(2): p. 269-75.
71. Reimmann, C., et al., *The global activator GacA of Pseudomonas aeruginosa PAO positively controls the production of the autoinducer N-butyryl-homoserine lactone and the formation of the virulence factors pyocyanin, cyanide, and lipase.* Mol Microbiol, 1997. **24**(2): p. 309-19.
72. Pena, R.T., et al., *Relationship Between Quorum Sensing and Secretion Systems.* Frontiers in Microbiology, 2019. **10**.
73. Kay, E., et al., *Two GacA-dependent small RNAs modulate the quorum-sensing response in Pseudomonas aeruginosa.* J Bacteriol, 2006. **188**(16): p. 6026-33.
74. Pessi, G., et al., *The global posttranscriptional regulator RsmA modulates production of virulence determinants and N-acylhomoserine lactones in Pseudomonas aeruginosa.* J Bacteriol, 2001. **183**(22): p. 6676-83.
75. Moscoso, J.A., et al., *The Pseudomonas aeruginosa sensor RetS switches type III and type VI secretion via c-di-GMP signalling.* Environ Microbiol, 2011. **13**(12): p. 3128-38.
76. Irie, Y., et al., *Self-produced exopolysaccharide is a signal that stimulates biofilm formation in Pseudomonas aeruginosa.* Proc Natl Acad Sci U S A, 2012. **109**(50): p. 20632-6.
77. Valentini, M. and A. Filloux, *Biofilms and Cyclic di-GMP (c-di-GMP) Signaling: Lessons from Pseudomonas aeruginosa and Other Bacteria.* J Biol Chem, 2016. **291**(24): p. 12547-12555.
78. Kang, C.-I., et al., *Pseudomonas aeruginosa Bacteremia: Risk Factors for Mortality and Influence of Delayed Receipt of Effective Antimicrobial Therapy on Clinical Outcome.* Clinical Infectious Diseases, 2003. **37**(6): p. 745-751.
79. Palmer, K.L., et al., *Cystic fibrosis sputum supports growth and cues key aspects of Pseudomonas aeruginosa physiology.* J Bacteriol, 2005. **187**(15): p. 5267-77.
80. Ozer, E., et al., *An inside look at a biofilm: Pseudomonas aeruginosa flagella biotracking.* Sci Adv, 2021. **7**(24).
81. Thi, M.T.T., D. Wibowo, and B.H.A. Rehm, *Pseudomonas aeruginosa Biofilms.* Int J Mol Sci, 2020. **21**(22).
82. Luzar, M.A., M.J. Thomassen, and T.C. Montie, *Flagella and motility alterations in Pseudomonas aeruginosa strains from patients with cystic fibrosis: relationship to patient clinical condition.* Infect Immun, 1985. **50**(2): p. 577-82.
83. Pamp, S.J., et al., *Tolerance to the antimicrobial peptide colistin in Pseudomonas aeruginosa biofilms is linked to metabolically active cells, and depends on the pmr and mexAB-oprM genes.* Mol Microbiol, 2008. **68**(1): p. 223-40.
84. Soto, S.M., *Role of efflux pumps in the antibiotic resistance of bacteria embedded in a biofilm.* Virulence, 2013. **4**(3): p. 223-9.

85. Yum, H.K., et al., *Recurrent Pseudomonas aeruginosa Infection in Chronic Lung Diseases: Relapse or Reinfection?* Tuberc Respir Dis (Seoul), 2014. **77**(4): p. 172-7.
86. Pang, Z., et al., *Antibiotic resistance in Pseudomonas aeruginosa: mechanisms and alternative therapeutic strategies.* Biotechnology Advances, 2019. **37**(1): p. 177-192.
87. Heras, B., M.J. Scanlon, and J.L. Martin, *Targeting virulence not viability in the search for future antibacterials.* Br J Clin Pharmacol, 2015. **79**(2): p. 208-15.
88. Iwasaki, J., et al., *Broad-spectrum in vitro activity of macrophage infectivity potentiator inhibitors against Gram-negative bacteria and Leishmania major.* Journal of Antimicrobial Chemotherapy, 2022. **77**(6): p. 1625-1634.
89. Rasch, J., C.M. Unal, and M. Steinert, *Peptidylprolyl cis-trans isomerases of Legionella pneumophila: virulence, moonlighting and novel therapeutic targets.* Biochem Soc Trans, 2014. **42**(6): p. 1728-33.
90. He, Z., L. Li, and S. Luan, *Immunophilins and Parvulins. Superfamily of Peptidyl Prolyl Isomerases in Arabidopsis.* Plant Physiology, 2004. **134**(4): p. 1248-1267.
91. Fischer G, B.H., Mech C, *Determination of enzymatic catalysis for the cis-trans-isomerization of peptide binding in proline-containing peptides.* Biomed Biochem Acta, 1984. **43**(10): p. 1101-11.
92. Göthel, S.F. and M.A. Marahiel, *Peptidyl-prolyl cis-trans isomerases, a superfamily of ubiquitous folding catalysts.* Cellular and Molecular Life Sciences CMLS, 1999. **55**(3): p. 423-436.
93. Brandts, J.F., Halvorson, H.R, Brennan, M, *Consideration of the Possibility That the Slow Step in Protein Denaturation Reactions Is Due to Cis-Trans Isomerism of Proline Residues.* Biochemistry 1975. **14**(22): p. 4953-4963.
94. Fischer, G., Wittmann-Liebold, B, Lang, K, Kiefhaber, T, Schmid, Franz, *Cyclophilin and peptidyl-prolyl cis-trans isomerase are probably identical proteins.* Nature, 1989. **337**: p. 476.
95. Joseph, A.P., Srinivasan, N., de Brevern, A.G., *Cis-trans peptide variations in structurally similar proteins.* Amino acids, 2012. **43**(3): p. 1369-1381.
96. Stewart, D.E., A. Sarkar, and J.E. Wampler, *Occurrence and role of cis peptide bonds in protein structures.* J Mol Biol, 1990. **214**(1): p. 253-60.
97. Hacker, J., Fischer, G.,, *Immunophilins: structure—function relationship and possible role in microbial pathogenicity.* Molecular Microbiology, 1993. **10**(3): p. 445-456.
98. Rüegger A, K.M., Lichti H, Loosli HR, Huguenin R, Quiquerez C, von Wartburg A., *Cyclosporin A, a Peptide Metabolite from Trichoderma polysporum (Link ex Pers.) Rifai, with a remarkable immunosuppressive activity.* helvetica Chimica Acta, 1976. **59**(4): p. 1075-92.
99. Handschumacher, R.E., et al., *Cyclophilin: a specific cytosolic binding protein for cyclosporin A.* Science, 1984. **226**(4674): p. 544-7.
100. Takahashi, N., T. Hayano, and M. Suzuki, *Peptidyl-prolyl cis-trans isomerase is the cyclosporin A-binding protein cyclophilin.* Nature, 1989. **337**(6206): p. 473-5.
101. Liu, J., et al., *Inhibition of T cell signaling by immunophilin-ligand complexes correlates with loss of calcineurin phosphatase activity.* Biochemistry, 1992. **31**(16): p. 3896-901.
102. Kawashima, H., Y. Fujino, and M. Mochizuki, *Effects of a new immunosuppressive agent, FK506, on experimental autoimmune uveoretinitis in rats.* Investigative Ophthalmology & Visual Science, 1988. **29**(8): p. 1265-1271.
103. Matsuda, S. and S. Koyasu, *Mechanisms of action of cyclosporine.* Immunopharmacology, 2000. **47**(2-3): p. 119-25.
104. Harding, M.W., et al., *A receptor for the immuno-suppressant FK506 is a cis–trans peptidyl-prolyl isomerase.* Nature, 1989. **341**: p. 758.
105. Siekierka, J.J., et al., *FK-506, a potent novel immunosuppressive agent, binds to a cytosolic protein which is distinct from the cyclosporin A-binding protein, cyclophilin.* The Journal of Immunology, 1989. **143**(5): p. 1580-1583.

106. Standaert, R.F., et al., *Molecular cloning and overexpression of the human FK506-binding protein FKBP.* Nature, 1990. **346**: p. 671.
107. Harrison, R.K. and R.L. Stein, *Substrate specificities of the peptidyl prolyl cis-trans isomerase activities of cyclophilin and FK-506 binding protein: evidence for the existence of a family of distinct enzymes.* Biochemistry, 1990. **29**(16): p. 3813-6.
108. Maki, N., et al., *Complementary DNA encoding the human T-cell FK506-binding protein, a peptidylprolyl cis-trans isomerase distinct from cyclophilin.* Proceedings of the National Academy of Sciences, 1990. **87**(14): p. 5440-5443.
109. Ohara, O., R.L. Dorit, and W. Gilbert, *One-sided polymerase chain reaction: the amplification of cDNA.* Proceedings of the National Academy of Sciences of the United States of America, 1989. **86**(15): p. 5673-5677.
110. Loh, E., *Anchored PCR: Amplification with single-sided specificity.* Methods, 1991. **2**(1): p. 11-19.
111. Perry, A.C., I.J. Nicolson, and J.R. Saunders, *Neisseria meningitidis C114 contains silent, truncated pilin genes that are homologous to Neisseria gonorrhoeae pil sequences.* Journal of bacteriology, 1988. **170**(4): p. 1691-1697.
112. Sampson, B.A. and E.C. Gotschlich, *Neisseria meningitidis encodes an FK506-inhibitable rotamase.* Proceedings of the National Academy of Sciences of the United States of America, 1992. **89**(4): p. 1164-1168.
113. Pearlman, E., N.C. Engleberg, and B.I. Eisenstein, *Identification of protein antigens of Legionella pneumophila serogroup 1.* Infect Immun, 1985. **47**(1): p. 74-9.
114. Jung, A.L., et al., *Legionella pneumophila Outer Membrane Vesicles: Isolation and Analysis of Their Pro-inflammatory Potential on Macrophages.* J Vis Exp, 2017(120).
115. Cianciotto, N.P., Eisenstein, B.L, Mody, C.H, Toews, G.B, Engleberg, N.P,, *A Legionella pneumophila Gene Encoding a Species-Specific Surface Protein Potentiates Initiation of Intracellular Infection.* Infection and Immunity, 1989. **57**(4): p. 1255-1262.
116. Engleberg, N.C., Pearlman, E and Eisenstein, B.I, *Legionella pneumophila Surface Antigens Cloned and Expressed in Escherichia coli Are Translocated to the Host Cell Surface and Interact with Specific Anti-Legionella Antibodies.* JOURNAL OF BACTERIOLOGY, 1984. **16**: p. 199-203.
117. Engleberg, N.C., D.J Drutz, B.I Eisenstien *Cloning and Expression of Legionella pneumophila Antigens in Escherichia coli.* Infection and Immunity, 1984. **44**(2): p. 222-227.
118. Pearlman, E., et al., *Growth of Legionella pneumophila in a human macrophage-like (U937) cell line.* Microb Pathog, 1988. **5**(2): p. 87-95.
119. Engleberg, N.C., Carter, C, Weber, D.R, Cianciotto, N.P, Eisenstein, B.I, *DNA sequence of mip, a Legionella pneumophila gene associated with macrophage infectivity.* Infection and Immunity, 1989. **57**(4): p. 1263-1270.
120. Lundemose, A.G., et al., *Characterization and identification of early proteins in Chlamydia trachomatis serovar L2 by two-dimensional gel electrophoresis.* Infect Immun, 1990. **58**(8): p. 2478-86.
121. Lundemose, A.G., et al., *Chlamydia trachomatis contains a protein similar to the Legionella pneumophila mip gene product.* Molecular Microbiology, 1991. **5**(1): p. 109-115.
122. Tropschug, M., et al., *Isolation and sequence of an FK506-binding protein from N. crassa which catalyses protein folding.* Nature, 1990. **346**: p. 674.
123. Lane, W.S., et al., *Complete amino acid sequence of the FK506 and rapamycin binding protein, FKBP, isolated from calf thymus.* J Protein Chem, 1991. **10**(2): p. 151-60.
124. Jin, Y.J., et al., *Molecular cloning of a membrane-associated human FK506- and rapamycin-binding protein, FKBP-13.* Proc Natl Acad Sci U S A, 1991. **88**(15): p. 6677-81.
125. Wiederrecht, G., et al., *FKB1 encodes a nonessential FK 506-binding protein in Saccharomyces cerevisiae and contains regions suggesting homology to the cyclophilins.* Proc Natl Acad Sci U S A, 1991. **88**(3): p. 1029-33.

126. Pahl, A. and U. Keller, *FK-506-binding proteins from streptomycetes producing immunosuppressive macrolactones of the FK-506 type.* Journal of bacteriology, 1992. **174**(18): p. 5888-5894.
127. Echenique-Rivera, H., et al., *Transcriptome analysis of Neisseria meningitidis in human whole blood and mutagenesis studies identify virulence factors involved in blood survival.* PLoS Pathog, 2011. **7**(5): p. e1002027.
128. Leuzzi, R., et al., *Ng-MIP, a surface-exposed lipoprotein of Neisseria gonorrhoeae, has a peptidyl-prolyl cis/trans isomerase (PPIase) activity and is involved in persistence in macrophages.* Molecular Microbiology, 2005. **58**(3): p. 669-681.
129. Fischer, G., et al., *Mip protein of Legionella pneumophila exhibits peptidyl-prolyl-cis/trans isomerase (PPIase) activity.* Mol Microbiol, 1992. **6**(10): p. 1375-83.
130. Engleberg, N.C., Eisenstein, B.I, *Progress in the pathogenesis of Legionella pneumophila.* Microbial Pathogenesis, 1991. **10**: p. 11-13.
131. Lundemose, A.G., J.E. Kay, and J.H. Pearce, *Chlamydia trachomatis Mip-like protein has peptidyl-prolyl cis/trans isomerase activity that is inhibited by FK506 and rapamycin and is implicated in initiation of chlamydial infection.* Mol Microbiol, 1993. **7**(5): p. 777-83.
132. Siekierka, J.J. and N.H. Sigal, *FK-506 and cyclosporin A: immunosuppressive mechanism of action and beyond.* Curr Opin Immunol, 1992. **4**(5): p. 548-52.
133. Bonner, J.M. and G.L. Boulianne, *Diverse structures, functions and uses of FK506 binding proteins.* Cell Signal, 2017. **38**: p. 97-105.
134. R. L. Vicente, S.M., J. R. Valverde, C. Palomino, R. P. Mellado and S. Gullón, *Functional identification of a Streptomyces lividans FKBP-like protein involved in the folding of overproduced secreted proteins.* Open Biology 2019(9).
135. Horne, S.M., et al., *Decreased intracellular survival of an fkpA mutant of Salmonella typhimurium Copenhagen.* Infection and immunity, 1997. **65**(2): p. 806-810.
136. Horne, S.M. and K.D. Young, *Escherichia coli and other species of the Enterobacteriaceae encode a protein similar to the family of Mip-like FK506-binding proteins.* Arch Microbiol, 1995. **163**(5): p. 357-65.
137. Roof, W.D. and R. Young, *Phi X174 lysis requires slyD, a host gene which is related to the FKBP family of peptidyl-prolyl cis-trans isomerases.* FEMS Microbiol Rev, 1995. **17**(1-2): p. 213-8.
138. Bernhardt, T.G., W.D. Roof, and R. Young, *The Escherichia coli FKBP-type PPIase SlyD is required for the stabilization of the E lysis protein of bacteriophage phi X174.* Mol Microbiol, 2002. **45**(1): p. 99-108.
139. Arié, J.-P., N. Sassoon, and J.-M. Betton, *Chaperone function of FkpA, a heat shock prolyl isomerase, in the periplasm of Escherichia coli.* Molecular Microbiology, 2001. **39**(1): p. 199-210.
140. Mo, Y.Y., N.P. Cianciotto, and L.P. Mallavia, *Molecular cloning of a Coxiella burnetii gene encoding a macrophage infectivity potentiator (Mip) analogue.* Microbiology, 1995. **141 (Pt 11)**: p. 2861-71.
141. Moro, A., et al., *Secretion by Trypanosoma cruzi of a peptidyl-prolyl cis-trans isomerase involved in cell infection.* Embo j, 1995. **14**(11): p. 2483-90.
142. Cianciotto, N.P., et al., *Detection of mip-like sequences and Mip-related proteins within the family Rickettsiaceae.* Curr Microbiol, 1995. **30**(3): p. 149-53.
143. Magdolen, U., et al., *Selection of specific gene probes by combined use of low-stringency PCR amplification and Southern-blot hybridization.* Current Genetics, 1995. **27**(4): p. 390-392.
144. Masuzawa, T., et al., *Relationship between pathogenicity of Coxiella burnetii isolates and gene sequences of the macrophage infectivity potentiator (Cbmip) and sensor-like protein (qrsA).* FEMS Microbiology Letters, 1997. **154**(2): p. 201-205.
145. Pereira, P.J., et al., *Trypanosoma cruzi macrophage infectivity potentiator has a rotamase core and a highly exposed alpha-helix.* EMBO Rep, 2002. **3**(1): p. 88-94.

146. Buchmeier, N.A. and F. Heffron, *Intracellular survival of wild-type Salmonella typhimurium and macrophage-sensitive mutants in diverse populations of macrophages.* Infect Immun, 1989. **57**(1): p. 1-7.
147. Gahring, L.C., et al., *Invasion and replication of Salmonella typhimurium in animal cells.* Infect Immun, 1990. **58**(2): p. 443-8.
148. Kontinen, V.P. and M. Sarvas, *The PrsA lipoprotein is essential for protein secretion in Bacillus subtilis and sets a limit for high-level secretion.* Mol Microbiol, 1993. **8**(4): p. 727-37.
149. Cron, L.E., et al., *Surface-associated lipoprotein PpmA of Streptococcus pneumoniae is involved in colonization in a strain-specific manner.* Microbiology (Reading), 2009. **155**(Pt 7): p. 2401-2410.
150. Hermans, P.W., et al., *The streptococcal lipoprotein rotamase A (SlrA) is a functional peptidyl-prolyl isomerase involved in pneumococcal colonization.* J Biol Chem, 2006. **281**(2): p. 968-76.
151. Marjuki, H., et al., *Genetic Similarity of Gonococcal Homologs to Meningococcal Outer Membrane Proteins of Serogroup B Vaccine.* mBio, 2019. **10**(5): p. e01668-19.
152. Hung, M.C., et al., *The Neisseria meningitidis macrophage infectivity potentiator protein induces cross-strain serum bactericidal activity and is a potential serogroup B vaccine candidate.* Infect Immun, 2011. **79**(9): p. 3784-91.
153. Dunkelberger, J.R. and W.-C. Song, *Complement and its role in innate and adaptive immune responses.* Cell Research, 2010. **20**(1): p. 34-50.
154. Weerts, H.P., et al., *A High-throughput Shigella-specific Bactericidal Assay.* J Vis Exp, 2019(144).
155. Ünal, C.M. and M. Steinert, *Microbial peptidyl-prolyl cis/trans isomerases (PPIases): virulence factors and potential alternative drug targets.* Microbiol Mol Biol Rev, 2014. **78**(3): p. 544-71.
156. Riboldi-Tunnicliffe, A., et al., *Crystal structure of Mip, a prolylisomerase from Legionella pneumophila.* Nat Struct Biol, 2001. **8**(9): p. 779-83.
157. Ceymann, A., et al., *Solution structure of the Legionella pneumophila Mip-rapamycin complex.* BMC Struct Biol, 2008. **8**: p. 17.
158. Ünal, C., et al., *Collagen IV-derived peptide binds hydrophobic cavity of Legionella pneumophila Mip and interferes with bacterial epithelial transmigration.* Cell Microbiol, 2011. **13**(10): p. 1558-72.
159. Budiman, C., et al., *Engineering of monomeric FK506-binding protein 22 with peptidyl prolyl cis-trans isomerase.* The FEBS Journal, 2009. **276**(15): p. 4091-4101.
160. Köhler, R., et al., *Biochemical and functional analyses of the Mip protein: influence of the N-terminal half and of peptidylprolyl isomerase activity on the virulence of Legionella pneumophila.* Infect Immun, 2003. **71**(8): p. 4389-97.
161. Wintermeyer, E., et al., *Influence of site specifically altered Mip proteins on intracellular survival of Legionella pneumophila in eukaryotic cells.* Infect Immun, 1995. **63**(12): p. 4576-83.
162. Helbig JH, K.B., Knospe H, Bubert B, Yu C, Lück CP, Riboldi-Tunnicliffe A, Hilgenfeld R, Jacobs E, Hacker J, Fischer G., *The PPIase active site of Legionella pneumophila Mip protein is involved in the infection of eukaryotic host cells.* Biological Chemistry, 2003. **384**(1): p. 125-37.
163. Wagner, C., et al., *Collagen binding protein Mip enables Legionella pneumophila to transmigrate through a barrier of NCI-H292 lung epithelial cells and extracellular matrix.* Cellular Microbiology, 2007. **9**(2): p. 450-462.
164. Rasch, J., et al., *Novel Cycloheximide Derivatives Targeting the Moonlighting Protein Mip Exhibit Specific Antimicrobial Activity Against Legionella pneumophila.* Frontiers in Bioengineering and Biotechnology, 2015. **3**.
165. Terada, L.S., et al., *Pseudomonas aeruginosa hemolytic phospholipase C suppresses neutrophil respiratory burst activity.* Infect Immun, 1999. **67**(5): p. 2371-6.

166. Debroy, S., et al., *Legionella pneumophila Mip, a surface-exposed peptidylproline cis-trans-isomerase, promotes the presence of phospholipase C-like activity in culture supernatants.* Infect Immun, 2006. **74**(9): p. 5152-60.
167. Ünal, C.M. and M. Steinert, *Microbial Peptidyl-Prolyl cis/trans Isomerases (PPIases): Virulence Factors and Potential Alternative Drug Targets.* Microbiology and Molecular Biology Reviews, 2014. **78**(3): p. 544-571.
168. Zang, N., et al., *Requirement of a mip-like gene for virulence in the phytopathogenic bacterium Xanthomonas campestris pv. campestris.* Mol Plant Microbe Interact, 2007. **20**(1): p. 21-30.
169. Dow, J.M., et al., *Biofilm dispersal in Xanthomonas campestris is controlled by cell-cell signaling and is required for full virulence to plants.* Proceedings of the National Academy of Sciences, 2003. **100**(19): p. 10995-11000.
170. Chan, J.W. and P.H. Goodwin, *The molecular genetics of virulence of Xanthomonas campestris.* Biotechnol Adv, 1999. **17**(6): p. 489-508.
171. Wuthiekanun, V. and S.J. Peacock, *Management of melioidosis.* Expert Rev Anti Infect Ther, 2006. **4**(3): p. 445-55.
172. Oren, A. and G.M. Garrity, *List of new names and new combinations previously effectively, but not validly, published.* Int J Syst Evol Microbiol, 2016. **66**(7): p. 2463-2466.
173. Gilad, J., D. Schwartz, and Y. Amsalem, *Clinical features and laboratory diagnosis of infection with the potential bioterrorism agents burkholderia mallei and burkholderia pseudomallei.* Int J Biomed Sci, 2007. **3**(3): p. 144-52.
174. Dance, D., *Treatment and prophylaxis of melioidosis.* Int J Antimicrob Agents, 2014. **43**(4): p. 310-8.
175. Schweizer, H.P., *Mechanisms of antibiotic resistance in Burkholderia pseudomallei: implications for treatment of melioidosis.* Future Microbiology, 2012. **7**(12): p. 1389-1399.
176. Pilatz, S., et al., *Identification of Burkholderia pseudomallei Genes Required for the Intracellular Life Cycle and In Vivo Virulence.* Infection and Immunity, 2006. **74**(6): p. 3576-3586.
177. Norville, I.H., et al., *A novel FK-506-binding-like protein that lacks peptidyl-prolyl isomerase activity is involved in intracellular infection and in vivo virulence of Burkholderia pseudomallei.* Microbiology, 2011. **157**(Pt 9): p. 2629-38.
178. Norville, I.H., et al., *A Burkholderia pseudomallei macrophage infectivity potentiator-like protein has rapamycin-inhibitable peptidylprolyl isomerase activity and pleiotropic effects on virulence.* Infect Immun, 2011. **79**(11): p. 4299-307.
179. Norville, I.H., et al., *The structure of a Burkholderia pseudomallei immunophilin-inhibitor complex reveals new approaches to antimicrobial development.* Biochem J, 2011. **437**(3): p. 413-22.
180. Dubin, G., et al., *Bacterial proteases in disease–role in intracellular survival, evasion of coagulation/fibrinolysis innate defenses, toxicoses and viral infections.* Current pharmaceutical design, 2013. **19**(6): p. 1090-1113.
181. Reffuveille, F., et al., *Involvement of peptidylprolyl cis/trans isomerases in Enterococcus faecalis virulence.* Infect Immun, 2012. **80**(5): p. 1728-35.
182. Overweg, K., et al., *The putative proteinase maturation protein A of Streptococcus pneumoniae is a conserved surface protein with potential to elicit protective immune responses.* Infect Immun, 2000. **68**(7): p. 4180-8.
183. Unal, C.M. and M. Steinert, *FKBPs in bacterial infections.* Biochim Biophys Acta, 2015. **1850**(10): p. 2096-102.
184. Qin, S., et al., *Pseudomonas aeruginosa: pathogenesis, virulence factors, antibiotic resistance, interaction with host, technology advances and emerging therapeutics.* Signal Transduction and Targeted Therapy, 2022. **7**(1): p. 199.
185. Podnecky, N.L., et al., *Mechanisms of Resistance to Folate Pathway Inhibitors in Burkholderia pseudomallei: Deviation from the Norm.* mBio, 2017. **8**(5): p. e01357-17.

186. Roulová, N., et al., *Antibiotic resistance of Pseudomonas aeruginosa isolated from hospital wastewater in the Czech Republic.* Journal of Water and Health, 2022. **20**(4): p. 692-701.
187. Scott, T.A., et al., *An L-threonine transaldolase is required for L-threo-β-hydroxy-α-amino acid assembly during obafluorin biosynthesis.* Nature Communications, 2017. **8**(1): p. 15935.
188. Choi, K.-H., et al., *A Tn7-based broad-range bacterial cloning and expression system.* Nature Methods, 2005. **2**(6): p. 443-448.
189. Heeb, S., et al., *Small, stable shuttle vectors based on the minimal pVS1 replicon for use in gram-negative, plant-associated bacteria.* Mol Plant Microbe Interact, 2000. **13**(2): p. 232-7.
190. Coffey BM, A.G., *Biofilm formation in the 96-well microtiter plate.* Methods in molecular Biology, 2014. **1149**: p. 631-641.
191. Ha, D.G., S.L. Kuchma, and G.A. O'Toole, *Plate-based assay for swarming motility in Pseudomonas aeruginosa.* Methods Mol Biol, 2014. **1149**: p. 67-72.
192. Turnbull, L. and C.B. Whitchurch, *Motility assay: twitching motility.* Methods Mol Biol, 2014. **1149**: p. 73-86.
193. Hoegy, F., G.L. Mislin, and I.J. Schalk, *Pyoverdine and pyochelin measurements.* Methods Mol Biol, 2014. **1149**: p. 293-301.
194. Hoegy, F. and I.J. Schalk, *Monitoring iron uptake by siderophores.* Methods Mol Biol, 2014. **1149**: p. 337-46.
195. Gudrun Koch, P.N.-J., Robbert H. Cool, and Wim J. Quax, *Assessing Pseudomonas Virulence with Nonmammalian Host: Galleria mellonella.* Methods in Molecular Biology, 2014.
196. Ali, N.M., et al., *Pseudomonas aeruginosa associated pulmonary infections and in vitro amplification virulent rhamnolipid (rhlR) gene.* Braz J Biol, 2021. **82**: p. e228009.
197. Sage, A.E. and M.L. Vasil, *Osmoprotectant-dependent expression of plcH, encoding the hemolytic phospholipase C, is subject to novel catabolite repression control in Pseudomonas aeruginosa PAO1.* Journal of Bacteriology, 1997. **179**(15): p. 4874-4881.
198. Minandri, F., et al., *Role of Iron Uptake Systems in Pseudomonas aeruginosa Virulence and Airway Infection.* Infection and Immunity, 2016. **84**(8): p. 2324-2335.
199. Hamdani, S., et al., *Isolation and identification of proteolytic bacteria from pig sludge and protease activity determination.* IOP Conference Series: Earth and Environmental Science, 2019. **230**.
200. Seo, J. and A.J. Darwin, *The Pseudomonas aeruginosa periplasmic protease CtpA can affect systems that impact its ability to mount both acute and chronic infections.* Infect Immun, 2013. **81**(12): p. 4561-70.
201. Dubin, G., et al., *Bacterial proteases in disease - role in intracellular survival, evasion of coagulation/ fibrinolysis innate defenses, toxicoses and viral infections.* Curr Pharm Des, 2013. **19**(6): p. 1090-113.
202. Persat, A., et al., *Type IV pili mechanochemically regulate virulence factors in Pseudomonas aeruginosa.* Proc Natl Acad Sci U S A, 2015. **112**(24): p. 7563-8.
203. Zolfaghar, I., D.J. Evans, and S.M. Fleiszig, *Twitching motility contributes to the role of pili in corneal infection caused by Pseudomonas aeruginosa.* Infect Immun, 2003. **71**(9): p. 5389-93.
204. Miller, R.M., et al., *Pseudomonas aeruginosa twitching motility-mediated chemotaxis towards phospholipids and fatty acids: specificity and metabolic requirements.* Journal of bacteriology, 2008. **190**(11): p. 4038-4049.
205. Xicohtencatl-Cortes, J., et al., *Intestinal adherence associated with type IV pili of enterohemorrhagic Escherichia coli O157:H7.* J Clin Invest, 2007. **117**(11): p. 3519-29.
206. Overhage, J., et al., *Swarming of Pseudomonas aeruginosa Is a Complex Adaptation Leading to Increased Production of Virulence Factors and Antibiotic Resistance.* Journal of Bacteriology, 2008. **190**(8): p. 2671-2679.

207. Caiazza, N.C., R.M. Shanks, and G.A. O'Toole, *Rhamnolipids modulate swarming motility patterns of Pseudomonas aeruginosa.* J Bacteriol, 2005. **187**(21): p. 7351-61.
208. Toutain, C.M., M.E. Zegans, and G.A. O'Toole, *Evidence for two flagellar stators and their role in the motility of Pseudomonas aeruginosa.* J Bacteriol, 2005. **187**(2): p. 771-7.
209. Abdel-Mawgoud, A.M., F. Lépine, and E. Déziel, *Liquid chromatography/mass spectrometry for the identification and quantification of rhamnolipids.* Methods Mol Biol, 2014. **1149**: p. 359-73.
210. Cox, C.D. and P. Adams, *Siderophore activity of pyoverdin for Pseudomonas aeruginosa.* Infect Immun, 1985. **48**(1): p. 130-8.
211. Pereira, T.C., et al., *Recent Advances in the Use of Galleria mellonella Model to Study Immune Responses against Human Pathogens.* J Fungi (Basel), 2018. **4**(4).
212. Taciak, B., et al., *Evaluation of phenotypic and functional stability of RAW 264.7 cell line through serial passages.* PLoS One, 2018. **13**(6): p. e0198943.
213. Mittal, R., et al., *Otopathogenic Pseudomonas aeruginosa Enters and Survives Inside Macrophages.* Front Microbiol, 2016. **7**: p. 1828.
214. Zumerle, S., et al., *Intercellular Calcium Signaling Induced by ATP Potentiates Macrophage Phagocytosis.* Cell Rep, 2019. **27**(1): p. 1-10.e4.
215. Jackson, A.A., et al., *Global regulator Anr represses PlcH phospholipase activity in Pseudomonas aeruginosa when oxygen is limiting.* Microbiology (Reading), 2014. **160**(Pt 10): p. 2215-2225.
216. Kukurba, K.R. and S.B. Montgomery, *RNA Sequencing and Analysis.* Cold Spring Harb Protoc, 2015. **2015**(11): p. 951-69.
217. McAteer, S., et al., *The lytB Gene of Escherichia coli Is Essential and Specifies a Product Needed for Isoprenoid Biosynthesis.* Journal of Bacteriology, 2001. **183**(24): p. 7403-7407.
218. Potter, S., et al., *Occurrence of homologs of the Escherichia coli lytB gene in gram-negative bacterial species.* J Bacteriol, 1998. **180**(7): p. 1959-61.
219. Gustafson, C.E., S. Kaul, and E.E. Ishiguro, *Identification of the Escherichia coli lytB gene, which is involved in penicillin tolerance and control of the stringent response.* J Bacteriol, 1993. **175**(4): p. 1203-5.
220. Wang, X., et al., *Bacterial Respiratory Chain Diversity Reveals a Cytochrome c Oxidase Reducing O2 at Low Overpotentials.* Journal of the American Chemical Society, 2019. **141**(28): p. 11093-11102.
221. Konyecsni, W.M. and V. Deretic, *DNA sequence and expression analysis of algP and algQ, components of the multigene system transcriptionally regulating mucoidy in Pseudomonas aeruginosa: algP contains multiple direct repeats.* J Bacteriol, 1990. **172**(5): p. 2511-20.
222. König, B., M.L. Vasil, and W. König, *Role of Haemolytic and Non-Haemolytic Phospholipase C from Pseudomonas Aeruginosa in Interleukin-8 release from Human Monocytes.* Journal of Medical Microbiology, 1997. **46**(6): p. 471-478.
223. Coin, D., et al., *LasA, alkaline protease and elastase in clinical strains of Pseudomonas aeruginosa: quantification by immunochemical methods.* FEMS Immunology & Medical Microbiology, 1997. **18**(3): p. 175-184.
224. Kessler, E., et al., *Elastase and the LasA protease of Pseudomonas aeruginosa are secreted with their propeptides.* J Biol Chem, 1998. **273**(46): p. 30225-31.
225. Martin, D.W., et al., *Mechanism of conversion to mucoidy in Pseudomonas aeruginosa infecting cystic fibrosis patients.* Proc Natl Acad Sci U S A, 1993. **90**(18): p. 8377-81.
226. Qiu, D., et al., *Regulated proteolysis controls mucoid conversion in Pseudomonas aeruginosa.* Proc Natl Acad Sci U S A, 2007. **104**(19): p. 8107-12.
227. Wozniak, D.J. and D.E. Ohman, *Transcriptional analysis of the Pseudomonas aeruginosa genes algR, algB, and algD reveals a hierarchy of alginate gene expression which is modulated by algT.* Journal of Bacteriology, 1994. **176**(19): p. 6007-6014.

228. Colvin, K.M., et al., *The Pel and Psl polysaccharides provide Pseudomonas aeruginosa structural redundancy within the biofilm matrix.* Environ Microbiol, 2012. **14**(8): p. 1913-28.
229. Wilder, C.N., S.P. Diggle, and M. Schuster, *Cooperation and cheating in Pseudomonas aeruginosa: the roles of the las, rhl and pqs quorum-sensing systems.* The ISME Journal, 2011. **5**(8): p. 1332-1343.
230. Glessner, A., et al., *Roles of Pseudomonas aeruginosa las and rhl Quorum-Sensing Systems in Control of Twitching Motility.* Journal of Bacteriology, 1999. **181**(5): p. 1623-1629.
231. Burrows, L.L., *Pseudomonas aeruginosa twitching motility: type IV pili in action.* Annual review of microbiology, 2012. **66**(1): p. 493-520.
232. Strom, M.S. and S. Lory, *Structure-function and biogenesis of the type IV pili.* Annu Rev Microbiol, 1993. **47**: p. 565-96.
233. Jenal, U., A. Reinders, and C. Lori, *Cyclic di-GMP: second messenger extraordinaire.* Nature Reviews Microbiology, 2017. **15**(5): p. 271-284.
234. Dasgupta, N., S.K. Arora, and R. Ramphal, *fleN, a gene that regulates flagellar number in Pseudomonas aeruginosa.* J Bacteriol, 2000. **182**(2): p. 357-64.
235. Mougous, J.D., et al., *A virulence locus of Pseudomonas aeruginosa encodes a protein secretion apparatus.* Science (New York, N.Y.), 2006. **312**(5779): p. 1526-1530.
236. Hachani, A., et al., *Type VI secretion system in Pseudomonas aeruginosa: secretion and multimerization of VgrG proteins.* The Journal of biological chemistry, 2011. **286**(14): p. 12317-12327.
237. Chen, L., et al., *Composition, function, and regulation of T6SS in Pseudomonas aeruginosa.* Microbiol Res, 2015. **172**: p. 19-25.
238. Hood, R.D., et al., *A type VI secretion system of Pseudomonas aeruginosa targets a toxin to bacteria.* Cell Host Microbe, 2010. **7**(1): p. 25-37.
239. Russell, A.B., et al., *Type VI secretion delivers bacteriolytic effectors to target cells.* Nature, 2011. **475**(7356): p. 343-7.
240. Sana, T.G., et al., *The second type VI secretion system of Pseudomonas aeruginosa strain PAO1 is regulated by quorum sensing and Fur and modulates internalization in epithelial cells.* J Biol Chem, 2012. **287**(32): p. 27095-105.
241. Russell, A.B., et al., *Diverse type VI secretion phospholipases are functionally plastic antibacterial effectors.* Nature, 2013. **496**(7446): p. 508-12.
242. Lesic, B., et al., *Quorum sensing differentially regulates Pseudomonas aeruginosa type VI secretion locus I and homologous loci II and III, which are required for pathogenesis.* Microbiology (Reading, England), 2009. **155**(Pt 9): p. 2845-2855.
243. Sana, T.G., et al., *Divergent control of two type VI secretion systems by RpoN in Pseudomonas aeruginosa.* PLoS One, 2013. **8**(10): p. e76030.
244. Waite, R.D., et al., *Clustering of Pseudomonas aeruginosa transcriptomes from planktonic cultures, developing and mature biofilms reveals distinct expression profiles.* BMC Genomics, 2006. **7**: p. 162.
245. Secor, P.R., et al., *Pf Bacteriophage and Their Impact on Pseudomonas Virulence, Mammalian Immunity, and Chronic Infections.* Frontiers in Immunology, 2020. **11**.
246. Secor, P.R., et al., *Filamentous Bacteriophage Produced by Pseudomonas aeruginosa Alters the Inflammatory Response and Promotes Noninvasive Infection In Vivo.* Infect Immun, 2017. **85**(1).
247. Dunyak, B.M. and J.E. Gestwicki, *Peptidyl-Proline Isomerases (PPIases): Targets for Natural Products and Natural Product-Inspired Compounds.* J Med Chem, 2016. **59**(21): p. 9622-9644.
248. Huang, H., et al., *An integrated genomic regulatory network of virulence-related transcriptional factors in Pseudomonas aeruginosa.* Nature Communications, 2019. **10**(1): p. 2931.

249. Martin, E.W. and M.H. Sung, *Challenges of Decoding Transcription Factor Dynamics in Terms of Gene Regulation.* Cells, 2018. **7**(9).
250. Cross, A.R., et al., *Overproduction of the AlgT Sigma Factor Is Lethal to Mucoid Pseudomonas aeruginosa.* J Bacteriol, 2020. **202**(20).
251. Hershberger, C.D., et al., *The algT (algU) gene of Pseudomonas aeruginosa, a key regulator involved in alginate biosynthesis, encodes an alternative sigma factor (sigma E).* Proc Natl Acad Sci U S A, 1995. **92**(17): p. 7941-5.
252. Chen, Y., et al., *Prolyl isomerase Pin1: a promoter of cancer and a target for therapy.* Cell Death & Disease, 2018. **9**(9): p. 883.
253. Joly, J.C. and J.R. Swartz, *Protein folding activities of Escherichia coli protein disulfide isomerase.* Biochemistry, 1994. **33**(14): p. 4231-6.
254. Smith, R.P., et al., *Targeting Bacterial Dsb Proteins for the Development of Anti-Virulence Agents.* Molecules, 2016. **21**(7).
255. Ha, U.H., Y. Wang, and S. Jin, *DsbA of Pseudomonas aeruginosa is essential for multiple virulence factors.* Infect Immun, 2003. **71**(3): p. 1590-5.
256. Jittawuttipoka, T., et al., *Mini-Tn7 vectors as genetic tools for gene cloning at a single copy number in an industrially important and phytopathogenic bacteria, Xanthomonas spp.* FEMS Microbiol Lett, 2009. **298**(1): p. 111-7.
257. Choi, K.-H. and H.P. Schweizer, *mini-Tn7 insertion in bacteria with single attTn7 sites: example Pseudomonas aeruginosa.* Nature Protocols, 2006. **1**(1): p. 153-161.
258. Fortier, L.C. and O. Sekulovic, *Importance of prophages to evolution and virulence of bacterial pathogens.* Virulence, 2013. **4**(5): p. 354-65.
259. Heeb, S., C. Blumer, and D. Haas, *Regulatory RNA as mediator in GacA/RsmA-dependent global control of exoproduct formation in Pseudomonas fluorescens CHA0.* J Bacteriol, 2002. **184**(4): p. 1046-56.
260. Hachani, A., et al., *Type VI secretion system in Pseudomonas aeruginosa: secretion and multimerization of VgrG proteins.* J Biol Chem, 2011. **286**(14): p. 12317-27.
261. Pietrosiuk, A., et al., *Molecular basis for the unique role of the AAA+ chaperone ClpV in type VI protein secretion.* J Biol Chem, 2011. **286**(34): p. 30010-21.
262. Wettstadt, S., et al., *Delivery of the Pseudomonas aeruginosa Phospholipase Effectors PldA and PldB in a VgrG- and H2-T6SS-Dependent Manner.* Frontiers in Microbiology, 2019. **10**(1718).
263. Rudzite, M., R. Endres, and A. Filloux, *Effectiveness of Pseudomonas aeruginosa type VI secretion system relies on toxin potency and type IV pili-dependent interaction.* bioRxiv, 2022: p. 2022.11.30.518499.
264. Brencic, A., et al., *The GacS/GacA signal transduction system of Pseudomonas aeruginosa acts exclusively through its control over the transcription of the RsmY and RsmZ regulatory small RNAs.* Molecular Microbiology, 2009. **73**(3): p. 434-445.
265. Li, Y., et al., *ClpV3 of the H3-Type VI Secretion System (H3-T6SS) Affects Multiple Virulence Factors in Pseudomonas aeruginosa.* Frontiers in Microbiology, 2020. **11**.
266. Lossi, N.S., et al., *The HsiB1C1 (TssB-TssC) complex of the Pseudomonas aeruginosa type VI secretion system forms a bacteriophage tail sheathlike structure.* J Biol Chem, 2013. **288**(11): p. 7536-7548.
267. Miyata, S., et al., *Use of the Galleria mellonella Caterpillar as a Model Host To Study the Role of the Type III Secretion System in Pseudomonas aeruginosa Pathogenesis.* Infection and Immunity, 2003. **71**(5): p. 2404-2413.
268. Hauser, A.R., *The type III secretion system of Pseudomonas aeruginosa: infection by injection.* Nat Rev Microbiol, 2009. **7**(9): p. 654-65.
269. Behrens, S., et al., *The SurA periplasmic PPIase lacking its parvulin domains functions in vivo and has chaperone activity.* Embo j, 2001. **20**(1-2): p. 285-94.
270. Charbonnier, J.B., et al., *On the role of the cis-proline residue in the active site of DsbA.* Protein Sci, 1999. **8**(1): p. 96-105.

271. Kadokura, H., L. Nichols, and J. Beckwith, *Mutational Alterations of the Key cis Proline Residue That Cause Accumulation of Enzymatic Reaction Intermediates of DsbA, a Member of the Thioredoxin Superfamily.* Journal of Bacteriology, 2005. **187**(4): p. 1519-1522.
272. Schmidt, M., et al., *Francisella tularensis subsp. holarctica DsbA homologue: a thioredoxin-like protein with chaperone function.* Microbiology (Reading), 2013. **159**(Pt 11): p. 2364-2374.
273. Park, A.J., et al., *A temporal examination of the planktonic and biofilm proteome of whole cell Pseudomonas aeruginosa PAO1 using quantitative mass spectrometry.* Mol Cell Proteomics, 2014. **13**(4): p. 1095-105.
274. Rehman, Z.U., et al., *Insights into the assembly of the alginate biosynthesis machinery in Pseudomonas aeruginosa.* Appl Environ Microbiol, 2013. **79**(10): p. 3264-72.
275. Gallagher, K.A., et al., *c-di-GMP Arms an Anti-σ to Control Progression of Multicellular Differentiation in Streptomyces.* Mol Cell, 2020. **77**(3): p. 586-599.e6.

Appendix

Appendix

1. Raw data for results presented in Chapter 3.

	Haemolysis Plate 1 (mm)	Haemolysis Plate 2 (mm)	Haemolysis Plate 3 (mm)
WT	12	11.5	12
B	12	12	12
C	12	12	12
D	10.5	11	11
E	11	11	11
F	12	12	12
G	12	12	12
H	12	11	12
I	6	6	6
J	7	7	7
K	12	12	12
L	12	11.5	12
M	12	12.5	12
N	11.5	12	12
O	10.5	11	11
P	11.5	12	12
Q	11.5	11	11.5
Control	5	5	5

Appendix Table 1.1. Raw data for haemolysis assays

	Protease Plate 1 (mm)	Protease Plate 2 (mm)	Protease Plate 3 (mm)
WT	14	14.5	14
B	14	14.5	14.5
C	14.5	14.5	14.5
D	14	14	14.5
E	14.5	14.5	14.5
F	14.5	14	14.5
G	14.5	14	14
H	14	14.5	14
I	9	8.5	9
J	9.5	10	9.5
K	14	14.5	14.5
L	14	14	14
M	14.5	14	14
N	14	14.5	14.5
O	14	14.5	14
P	14	14	14.5
Q	14	14	14.5
Control	5	5	5

Appendix Table 1.2. Raw data for proteolysis assays

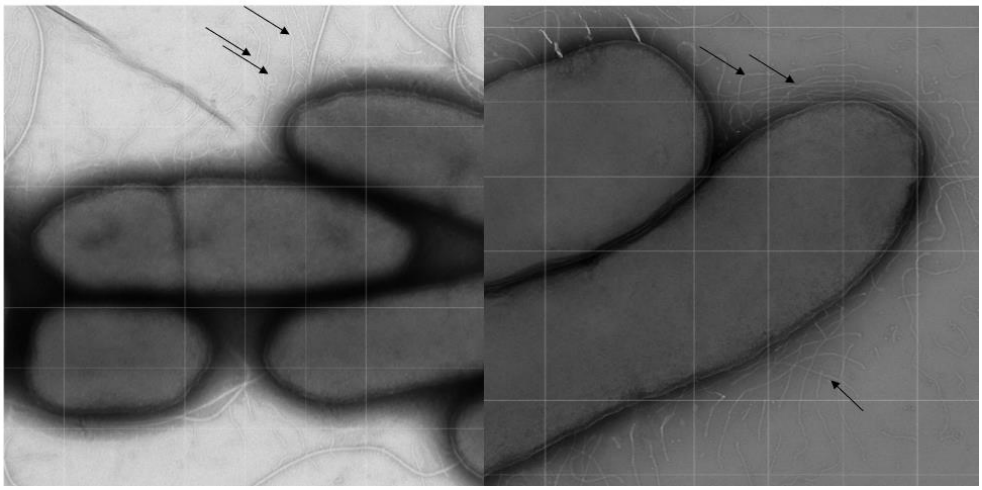
	WT			B			C	
0.494	0.422	0.401	0.527	0.492	0.517	0.391	0.538	0.47
	D			E			F	
0.413	0.477	0.456	0.472	0.471	0.432	0.527	0.494	0.498
	G			H			I	
0.431	0.515	0.608	0.53	0.556	0.543	0.325	0.308	0.304
	J			K			L	
0.376	0.442	0.358	0.583	0.573	0.545	0.974	0.65	0.686
	M			N			O	
0.532	0.611	0.584	0.442	0.489	0.523	0.534	0.6	0.59
	P			Q			Con	
0.58	0.519	0.574	0.555	0.582	0.533	0.094	0.085	0.09

Appendix Table 1.3. Raw data for biofilm formation assays

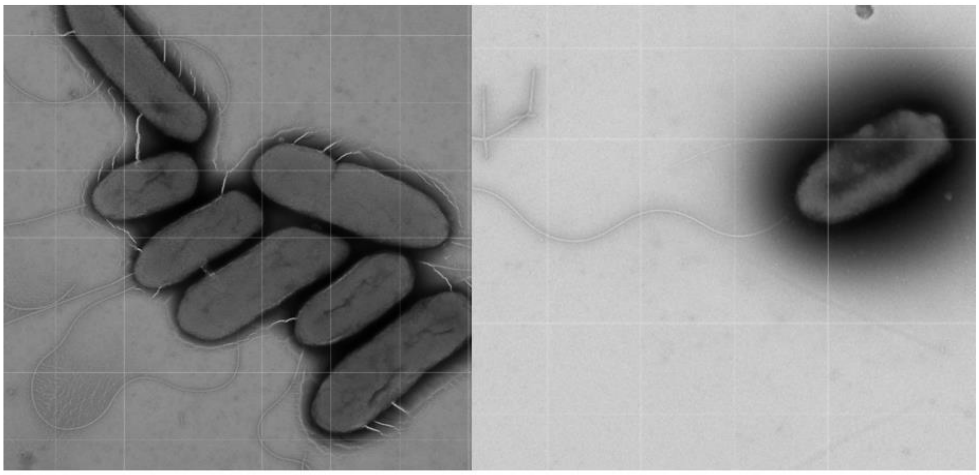
	Plate 1 (mm)	Plate 2 (mm)	Plate 1 (mm)	Plate 2 (mm)
WT	200	176	379.2	343.2
D	175.9	240.2	343.2	471.15
E	226.9	200.2	414.6	379.3
F	113	175.9	361.2	361.2
I	0	0	0	38.5
J	0	0	0	0

Appendix Table 1.4. Raw data for twitching assays

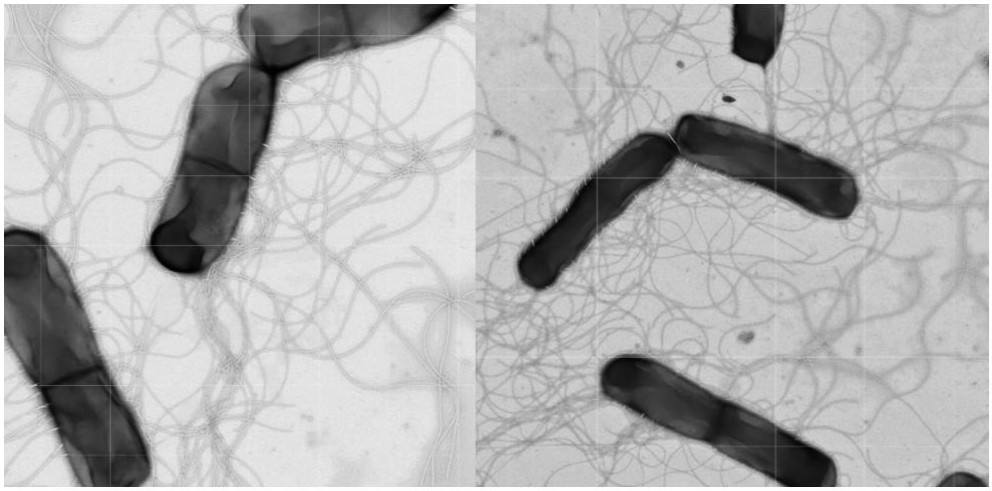
WT



I

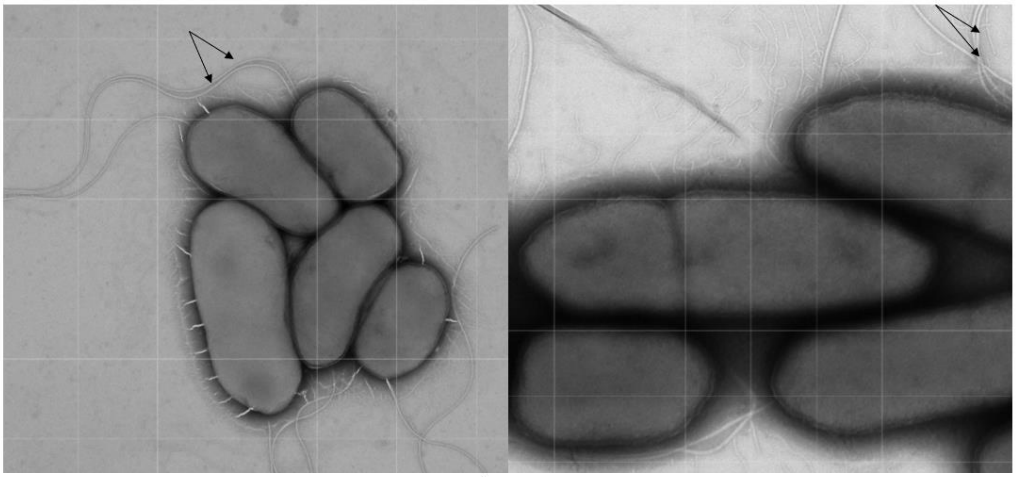


J

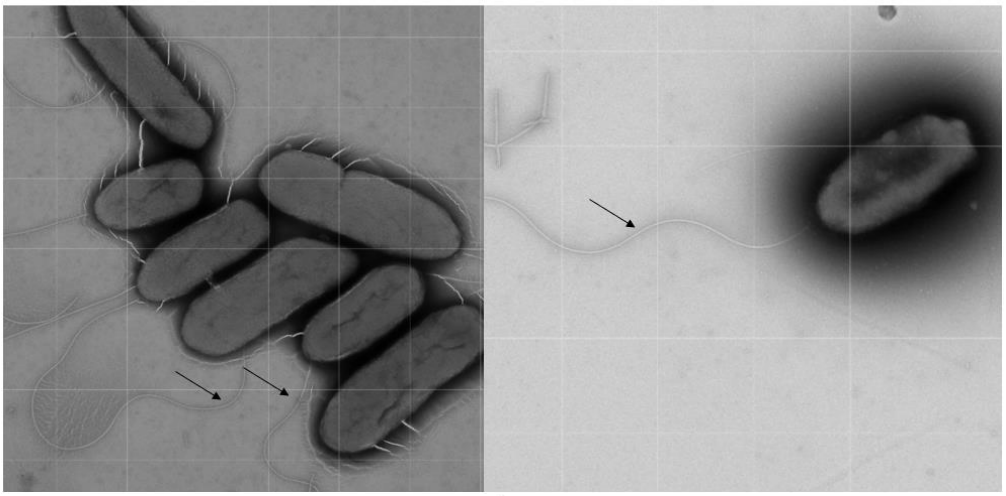


Appendix Image 1.1. Further images displaying type IV pili in the WT in the top images. The middle image further shows that all strain I cells lack type IV pili. The bottom image further shows that all strain J cells lack type IV pili.

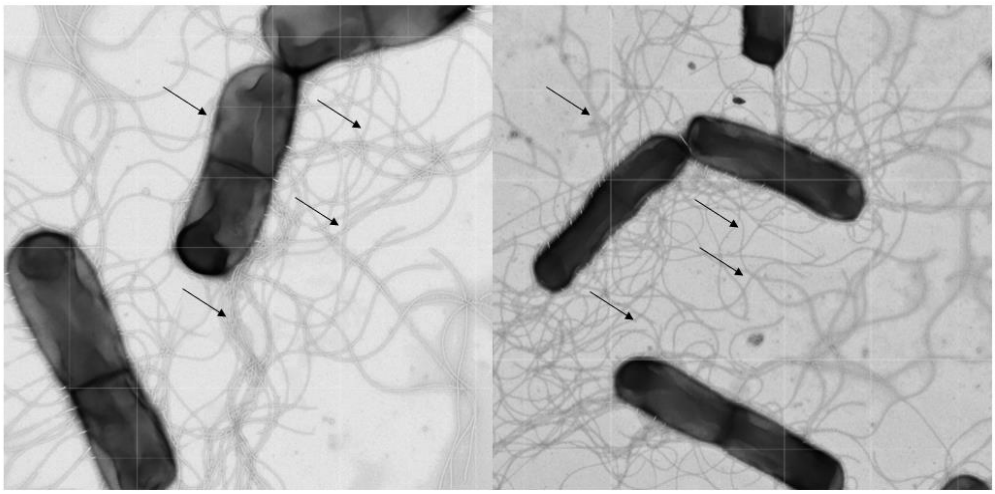
WT



I



J



Appendix Image 1.2. Further images displaying twinning flagella in the WT in the top images. The middle image further shows that all strain I cells lack twinning flagella with only one flagellum. The bottom image further shows that all strain J cells have hyper flagellation.

	6hr	12hr	18hr	24hr	30hr	36hr	42hr	48hr
WT	0	0	0	0	10	17	20	20
B	0	0	0	0	8	16	20	20
C	0	0	0	0	12	16	20	20
D	0	0	0	0	0	16	80	90
E	0	0	0	0	4	16	80	20
F	0	0	0	0	10	20	20	20
G	0	0	0	0	12	18	20	20
H	0	0	0	0	10	18	20	20
I	0	0	0	0	0	0	0	2
J	0	0	0	0	0	18	18	20
K	0	0	0	0	6	12	18	20
L	0	0	0	0	8	14	18	20
M	0	0	0	0	16	18	18	20
N	0	0	0	0	14	20	20	20
O	0	0	0	0	16	20	20	20
P	0	0	0	0	18	20	20	20
Q	0	0	0	0	18	20	20	20
Control	0	0	0	0	0	0	0	0

Appendix Table 1.5. Raw data for the number of *G. mellonella* larvae counted as dead across two experiments. Each experiment was performed with 10 larvae per strain. This table is the pooled raw data across the two experiments from which statistics and conversion to a final percentage was performed that is presented in Figure 3.15.

	Plate 1 colonies counted	Plate 2 colonies counted	Plate 3 colonies counted	Plate 4 colonies counted
WT	194	174	156	74
D	184	226	212	232
E	186	192	172	154
F	76	66	82	94
I	14	44	52	92
J	82	106	116	132
Control	0	0	0	0

Appendix Table 1.6. Raw data for MOI 10 plates and colonies counted per plate in replicates of 4 plates per strain

	Plate 1 colonies counted	Plate 2 colonies counted	Plate 3 colonies counted	Plate 4 colonies counted
WT	194	174	156	74
D	184	226	212	232
E	186	192	172	154
F	76	66	82	94
I	14	44	52	92
J	82	106	116	132
Control	0	0	0	0

Appendix Table 1.7 Raw data for MOI 50 plates and colonies counted per plate in replicates of 4 plates per strain

			I			J		
Gene Nº	Gene name	Description	Transcription level	p Value	FDR	Transcription level	p Value	FDR
Haemolytic and Proteolytic genes								
pa0844	*plcH*	Haemolytic phospholipase C precursor	0.940	0.000	0.006	0.410	0.001	0.004
pa0766	*mucD*	Serine protease MucD precursor	0.886	0.045	0.134	0.661	0.125	0.172
pa1871	*lasA*	Protease precursor	-0.658	0.000	0.005	-1.214	0.000	0.000
pa2862	*lipA*	lipA, lactonizing lipase	-0.557	0.009	0.079	0.674	0.003	0.008
pa5134	*ctpA*	Carboxyl-terminal protease	0.729	0.001	0.006	1.820	0.000	0.000
GacS / RetS Two-Component system genes								
pa0527	*rsmY*	Regulatory RNA	-0.357	0.348	0.546	-1.016	0.013	0.028
pa0905	*rsmA*	Regulatory RNA	-0.679	0.130	0.281	-0.756	0.094	0.136
pa0928	*gacS*	Sensor/response regulator hybrid	-0.040	0.790	0.900	1.242	2.2x10-8	8.6x10-7
pa2586	*gacA*	Response Regulator	-0.690	0.008	0.050	-0.510	0.030	0.080
Ppa621.1	*rsmZ*	Regulatory RNA	1.665	0.000	0.000	1.090	0.001	0.002
pa3974	*ladS*	Lost Adherence Sensor	-0.690	7.5x10-5	0.020	-1.170	2.51x10-7	5.15x10-6
pa4856	*retS*	Regulator of Exopolysaccharide and Type III Secretion	-0.622	0.004	0.023	-0.700	0.001	0.010
Alginate /*psl* / *pel* mucoidy synthesis genes								
pa0762	*algT*	Positive regulator alginate biosynthesis	0.339	0.469	0.659	-0.596	0.211	0.269
pa0763	*mucA*	Anti-sigma factor MucA	0.841	0.120	0.266	0.080	0.877	0.898
pa0764	*mucB*	Negative regulator alginate biosynthesis	0.848	0.084	0.207	0.308	0.513	0.572
pa4033	*MucE*	Positive regulator alginate biosynthesis	-0.002	0.993	0.997	-0.847	0.006	0.016
pa3649	*mucP*	Positive regulator alginate biosynthesis	-0.053	0.840	0.927	-0.659	0.005	0.013
pa5253	*algP*	Alginate biosynthesis regulatory protein	0.425	0.287	0.482	-1.165	0.008	0.018
pa3541	*alg8*	Alginate biosynthesis	-0.240	0.549	0.720	1.508	0.001	0.004
pa3542	*alg44*	Alginate biosynthesis	0.246	0.669	0.812	2.692	0.000	0.002
pa3543	*algK*	Alginate biosynthesis	0.318	0.573	0.740	3.066	0.000	0.001
pa3544	*algE*	Alginate biosynthesis	0.041	0.916	0.957	2.475	0.408	0.474
pa3545	*algG*	Alginate biosynthesis	-0.480	0.028	0.095	0.169	0.408	0.474
pa3546	*algX*	Alginate biosynthesis	-0.329	0.603	0.762	2.792	0.000	0.002
pa3547	*algL*	Alginate biosynthesis	-0.120	0.853	0.922	2.917	0.000	0.002
pa3548	*algI*	Alginate biosynthesis	0.054	0.897	0.948	1.553	0.001	0.004
pa3549	*algJ*	Alginate biosynthesis	-0.185	0.806	0.898	3.399	0.000	0.002
pa3550	*algF*	Alginate biosynthesis	-0.622	0.174	0.342	1.806	0.000	0.003
pa3551	*algA*	Alginate biosynthesis	-0.389	0.150	0.310	1.247	0.000	0.001
pa2232	pslB	pslB, biofilm formation protein PslB	-0.257	0.083	0.254	-1.508	0.000	0.000
pa2233	pslC	pslC, biofilm formation protein PslC	0.073	0.621	0.807	-0.796	0.000	0.001
pa2234	pslD	pslD, biofilm formation protein PslD	-0.068	0.622	0.807	-0.472	0.004	0.011
pa2235	pslE	pslE, biofilm formation protein PslE	-0.404	0.007	0.053	-1.258	0.000	0.000
pa2236	pslF	pslF, biofilm formation protein PslF	-0.331	0.054	0.193	-1.190	0.000	0.000
pa2237	pslG	pslG, biofilm formation protein PslG	-0.447	0.002	0.021	-1.173	0.000	0.000
pa2238	pslH	pslH, biofilm formation protein PslH	-0.222	0.259	0.507	-1.340	0.000	0.000
pa2239	pslI	pslI, biofilm formation protein PslI	-0.497	0.004	0.033	-0.946	0.000	0.000
pa2240	pslJ	pslJ, biofilm formation protein PslJ	-0.214	0.120	0.323	-1.312	0.000	0.000

pa2241	pslK	pslK, biofilm formation protein PslK	-0.135	0.630	0.812	-1.570	0.000	0.000
pa2242	pslL	pslL, hypothetical protein	0.069	0.652	0.825	-0.891	0.000	0.000
pa2243	pslM	pslM, FAD-binding dehydrogenase	0.450	0.224	0.466	0.656	0.086	0.151
pa2244	pslN	pslN, hypothetical protein	1.294	0.002	0.019	0.686	0.053	0.103
pa3058	pelG	pelG, pellicle/biofilm biosynthesis	0.795	0.042	0.127	1.353	0.001	0.005
pa3059	pelF	pelF, pellicle/biofilm biosynthesis	0.204	0.650	0.796	1.933	0.000	0.002
pa3060	pelE	pelE, pellicle/biofilm biosynthesis	0.080	0.868	0.933	1.720	0.002	0.006
pa3061	pelD	pelD, pellicle/biofilm biosynthesis	-0.040	0.945	0.970	2.683	0.000	0.002
pa3062	pelC	pelC, pellicle/biofilm biosynthesis	-0.098	0.886	0.943	3.038	0.000	0.002
pa3063	pelB	pelB, pellicle/biofilm biosynthesis	0.460	0.469	0.658	3.212	0.000	0.001
pa3064	pelA	pelA, hypothetical protein	-0.665	0.162	0.327	1.720	0.001	0.004
***las / rhl* Quorum sensing (QS) system genes**								
pa1432	*lasI*	Autoinducer synthesis protein	0.198	0.360	0.610	-0.240	0.260	0.370
pa1430	*lasR*	Transcriptional regulator	0.275	0.308	0.560	-0.976	0.002	0.007
pa3476	*rhlI*	Autoinducer synthesis protein RhlI	-0.279	0.218	0.458	0.513	0.033	0.069
pa3477	*rhlR*	Transcriptional regulator	0.272	0.174	0.403	-0.627	0.006	0.016
Type IV Pili assocaited genes								
pa4462	*rpoN*	Sigma factor	0.321	0.268	0.461	-0.714	0.020	0.038
pa4525	*pilA*	Type IV pili synthesis	0.796	0.040	0.160	-0.400	0.240	0.350
pa4526	*pilB*	Type IV pili synthesis	-0.090	0.730	0.870	2.930	7.78x10-8	2.09x10-6
pa4528	*pilD*	Type IV pili synthesis	-0.014	0.900	0.950	0.180	0.160	0.250
pa4546	*pilS*	Type IV pili synthesis	-0.050	0.710	0.860	0.690	0.000	0.001
pa4547	*pilR*	Type IV pili synthesis	-0.401	0.020	0.110	1.010	0.000	0.001
pa4549	*fimT*	Type IV pili synthesis	0.451	0.011	0.070	1.910	1.26x10-8	5.77x10-7
pa4550	*fimU*	Type IV pili synthesis	0.440	0.012	0.070	0.460	0.010	0.020
pa4551	*pilV*	Type IV pili synthesis	0.313	0.020	0.120	0.240	0.070	0.130
pa4552	*pilW*	Type IV pili synthesis	0.040	0.780	0.900	-0.390	0.030	0.070
pa4553	*pilX*	Type IV pili synthesis	0.380	0.015	0.080	-0.010	0.890	0.920
pa4554	*pilY*	Type IV pili synthesis	0.350	0.182	0.410	0.460	0.080	0.150
pa4555	*pilY2*	Type IV pili synthesis	0.314	0.130	0.346	-0.410	0.050	0.100
pa4556	*pilE*	Type IV pili synthesis	0.250	0.140	0.360	-0.910	8.18x10-5	0.001
pa4959	*fimX*	Phosphodiesterase involved in twitching	0.294	0.068	0.223	0.283	0.077	0.140
Cyclic-di-GMP synthesis genes								
pa1120	*tpbB*	diguanylate cyclase	-0.095	0.628	0.779	0.065	0.738	0.778
pa4108		Cyclic di-gmp phosphodiesterase - degradataion of cyclic di-gmp	0.236	0.169	0.397	-0.440	0.018	0.041
pa4781		Cyclic di-gmp phosphodiesterase	-0.370	0.004	0.036	-0.369	0.004	0.013
pa3702	wspR	directs formation of c-di-GMP	0.056	0.761	0.870	-0.001	0.995	0.996
pa3703	wspF	directs formation of c-di-GMP	-0.144	0.309	0.560	-0.696	0.000	0.001
Flagella assembly and Rhamnolipid biosynthesis genes								
pa1100	fliE	fliE, flagellar hook-basal body complex	1.005	0.000	0.000	-0.880	0.000	0.001
pa1101	fliF	fliF, flagellar MS-ring protein	1.265	0.000	0.000	-0.084	0.648	0.699
pa1102	fliG	fliG, flagellar motor switch protein FliG	0.819	0.001	0.010	-0.580	0.016	0.031
pa1103	Hypo	flagellar assembly protein FliH	-0.277	0.071	0.185	-0.196	0.191	0.247
pa1104	fliL	fliI, flagellum-specific ATP synthase	-0.014	0.918	0.957	0.433	0.004	0.011
pa1105	fliJ	fliJ, flagellar biosynthesis chaperone	-0.093	0.656	0.802	0.433	0.046	0.075
pa1454	fleN	flagellar synthesis regulator	-0.210	0.500	0.682	-0.048	0.693	0.771

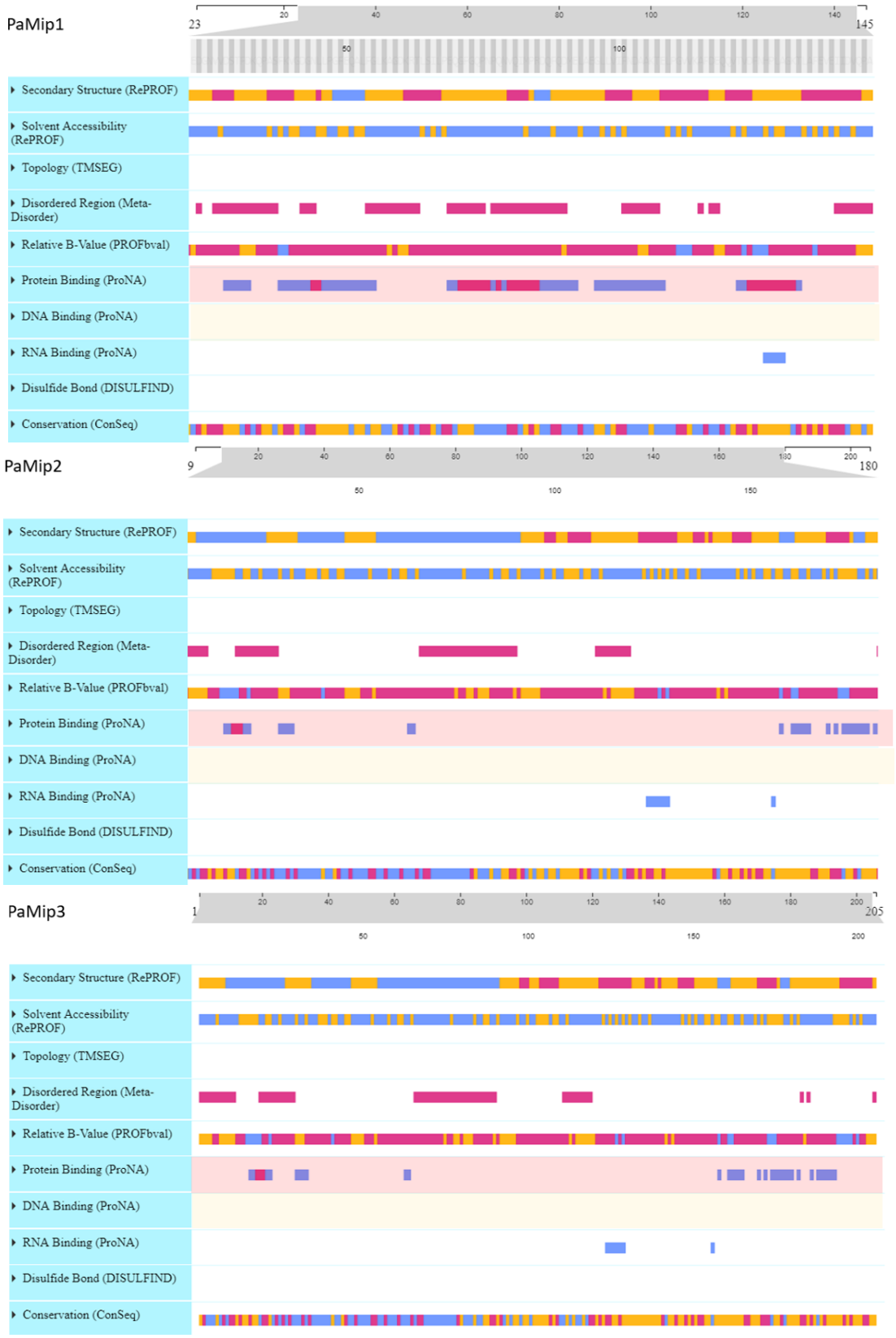
pa3387	*rhlG*	Rhamnolipid biosynthesis	0.390			1.980	0.000	0.002
pa3476	*rhlL*	Rhamnolipid biosynthesis	0.510	0.030	0.060	-1.066	0.009	0.020
pa3478	*rhlB*	Rhamnolipid biosynthesis	-0.360	0.010	0.080	0.550	0.001	0.003
pa3479	*rhlA*	Rhamnolipid biosynthesis	-0.090	0.510	0.730	-0.450	0.008	0.021

Appendix Table 1.8. A table of RNA sequence data obtained by Silke Alt that corresponds to Table 3.5. in Chapter 3. The table of RNA sequence data looking at specific groups of genes involved in a variety of virulence processes that have been discussed for both strains I and J. The left column labelled gene number is the annotated gene number as it appears in the chromosome, the column labelled gene name is the name given to a particular gene in literature and what it may be referred to as. The Log Fold Change (LogFC) column is the log-ratio of a gene's or a transcript's expression values compared to the WT strain. The LogFC value for each gene was obtained taking the average of three transcript readings for each gene in the Pa01 genome from which statistical anaylsis could be applied. To get the overall fold change in real terms compared to the WT, 2 is raised to the power of the LogFC value. The P value column indicated if the LogFC change of a given gene was of statistical significance if $P<0.05$. This was further statistically challenged by looking at the FDR value, also set at a significance level of $FDR<0.05$. Using both values ensured the LogFC value was statistically significant and not down to chance or false positives. Boxes highlighted in green call attention to a gene that has an increased LogFC change when compared to the WT and are statistically significant ($P<0.05$, $FDR<0.05$) suggesting the gene is more highly transcribed and upregulated compared to the WT. Boxes highlighted in red call attention to a gene that has a decrease LogFC change when compared to the WT and are statistically significant ($P<0.05$, $FDR<0.05$) suggesting the gene is more lowly transcribed and upregulated compared to the WT. Boxes that remain unhighlighted show genes whose LogFC change is not statistically significant when compared to the WT, therefore one could assume transcript levels of a given gene between the WT and a given strain are no different, neither up or down regulated.

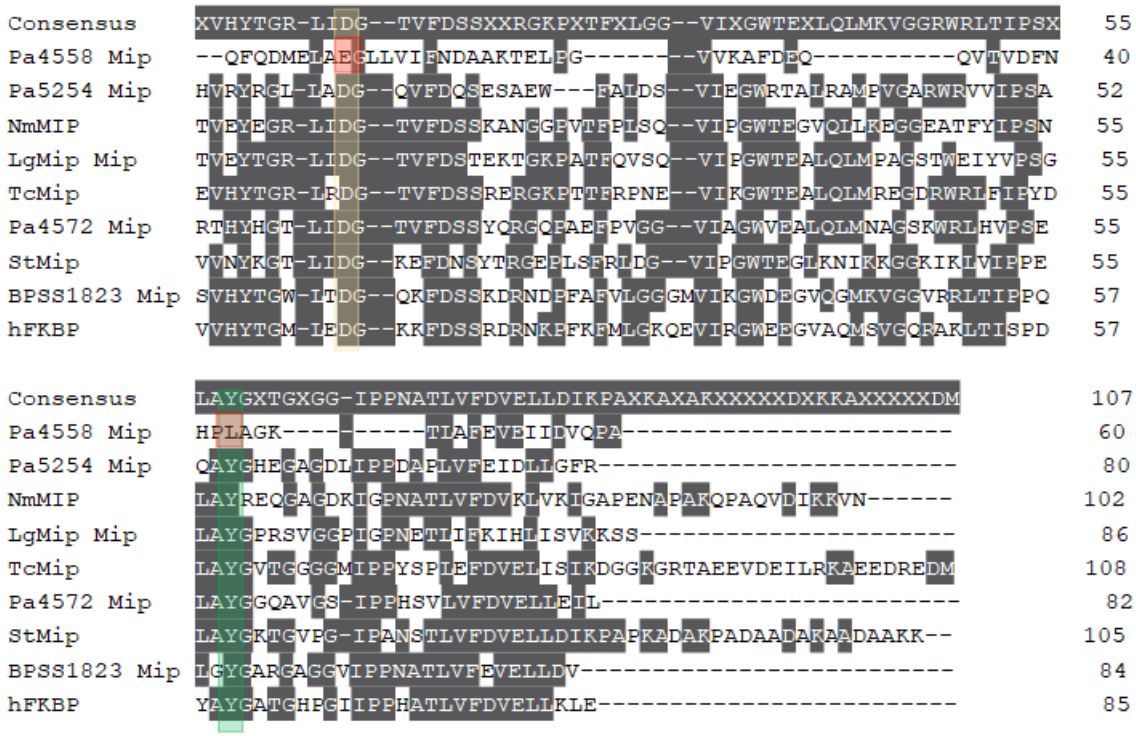
			I			J		
Gene No	Gene name	Description	Transcription level	p Value	FDR	Transcription level	p Value	FDR
Type Six Secretion System HSI-III genes								
pa2359	*sfa3*	Probable transcriptional regulator	-0.356	0.428	0.622	-0.250	0.240	0.347
pa2360	*hsiA3*	T6SS HSI-III hypothetical protein	-1.046	0.017	0.065	-2.787	0.000	0.000
pa2361	*Icmf3*	T6SS HSI-III hypothetical protein	-0.914	0.054	0.153	-1.520	0.000	0.000
pa2362	*dotU3*	T6SS HSI-III hypothetical protein	-1.629	0.003	0.017	-2.956	0.000	0.000
pa2363	*hsiJ3*	T6SS HSI-III hypothetical protein	-1.139	0.005	0.027	-2.413	0.000	0.000
pa2364	*lip3*	T6SS HSI-III hypothetical protein	-1.150	0.002	0.014	-2.340	0.000	0.000
pa2365	*hsiB3*	T6SS HSI-III hypothetical protein	-1.456	0.010	0.047	-2.624	0.000	0.000
pa2366	*hsiC3*	T6SS HSI-III hypothetical protein	-1.623	0.003	0.019	-1.436	0.000	0.000
pa2367	*hcp3*	Secreted protein Hcp	-0.374	0.411	0.606	-1.297	0.000	0.000
pa2368	*hsiF3*	T6SS HSI-III hypothetical protein	-1.505	0.008	0.038	-2.195	0.000	0.002
pa2369	*hsiG3*	T6SS HSI-III hypothetical protein	-1.432	0.011	0.048	-3.414	0.000	0.000
pa2370	*hsiH3*	T6SS HSI-III hypothetical protein	-1.301	0.015	0.060	-3.451	0.000	0.000
pa2371	*ClpV3*	Probable ClpA/B-type protease	-0.695	0.096	0.228	-2.193	0.000	0.000
pa2373	*VgrG3*	T6SS HSI-III hypothetical protein	-0.064	0.829	0.910	-0.624	0.000	0.001
pa2374	*tseF*	T6SS HSI-III hypothetical protein	-0.639	0.013	0.055	-1.087	0.002	0.006
Type Six Secretion System HSI-II genes								
pa1511	*VgrG2*	T6SS HSI-II hypothetical protein	0.141	0.293	0.543	2.338	0.001	0.000
pa1512	*Hcp2*	Secreted protein Hcp	0.088	0.703	0.854	0.039	0.897	0.914
pa1656	*hsiA2*	T6SS HSI-II hypothetical protein	0.645	0.005	0.040	-1.057	0.000	0.001
pa1657	*hsiB2*	T6SS HSI-II hypothetical protein	-0.135	0.404	0.648	-1.117	0.001	0.003
pa1658	*hsiC2*	T6SS HSI-II hypothetical protein	-0.180	0.190	0.423	-0.758	0.012	0.025
pa1659	*hsif3*	T6SS HSI-II hypothetical protein	-0.548	0.040	0.158	0.451	0.288	0.350
pa1660	*hsiG2*	T6SS HSI-II hypothetical protein	-0.062	0.709	0.858	1.154	0.007	0.016
pa1661	*hsiH2*	T6SS HSI-II hypothetical protein	-0.106	0.510	0.727	0.784	0.057	0.089
pa1662	*ClpV2*	Probable ClpA/B-type protease	0.196	0.187	0.420	0.901	0.013	0.028
pa1663	*sfa2*	Probable transcriptional regulator	-0.520	0.034	0.141	1.599	0.005	0.014
pa1664	*fha2*	T6SS HSI-II hypothetical protein	-1.067	0.196	0.432	1.956	0.020	0.038
pa1665	*lip2*	T6SS HSI-II hypothetical protein	0.321	0.319	0.572	1.808	0.004	0.011
pa1666	*hsij2*	T6SS HSI-II hypothetical protein	-0.079	0.718	0.863	1.270	0.027	0.048
pa1667	*dotu3*	T6SS HSI-II hypothetical protein	0.229	0.062	0.211	0.513	0.019	0.036
pa1668	*icmF2*	T6SS HSI-II hypothetical protein	0.152	0.501	0.721	1.126	0.009	0.020
pa1669	*stp1*	T6SS HSI-II hypothetical protein	-0.016	0.926	0.966	1.101	0.001	0.004
pa1670	*stk1*	T6SS HSI-II hypothetical protein	-0.180	0.783	0.902	0.802	0.013	0.027
Type Six Secretion System HSI-I genes								
pa0074	*ppkA*	Serine threonine protein kinase PpkA	-0.124	0.802	0.895	2.106	0.000	0.002
pa0075	*pppA*	serine/threonine phosphatase	-0.305	0.597	0.757	2.253	0.001	0.004
pa0076	*pppB*	T6SS HSI-I hypothetical protein	-0.890	0.126	0.275	1.897	0.002	0.007
pa0077	*icmF1*	Type VI secretion protein IcmF	-0.174	0.694	0.829	1.941	0.000	0.002
pa0078	*dotU1*	T6SS HSI-I hypothetical protein	-0.345	0.449	0.643	1.704	0.001	0.004
pa0079	*hsiJ1*	T6SS HSI-I hypothetical protein	-0.730	0.051	0.145	1.126	0.004	0.010
pa0080	*lip1*	T6SS HSI-I hypothetical protein	-0.597	0.151	0.312	0.803	0.054	0.085
pa0081	*fha1*	Fha domain-containing protein	-0.268	0.096	0.227	-0.057	0.711	0.753

pa0082	*hsiA1*	T6SS HSI-I hypothetical protein	0.329	0.213	0.392	0.350	0.184	0.239
pa0083	*hsiB1*	T6SS HSI-I hypothetical protein	-0.056	0.890	0.944	0.219	0.587	0.642
pa0084	*hsiC1*	T6SS HSI-I hypothetical protein	-0.294	0.447	0.641	0.209	0.582	0.638
pa0085	*hcp1*	Secreted protein Hcp	-0.487	0.165	0.330	-0.558	0.111	0.156
pa0086	*hsiE1*	T6SS HSI-I hypothetical protein	0.442	0.291	0.487	1.598	0.001	0.004
pa0087	*hsif1*	T6SS HSI-I hypothetical protein	0.639	0.250	0.440	2.493	0.000	0.002
pa0088	*hsiG1*	T6SS HSI-I hypothetical protein	-0.118	0.811	0.901	2.494	0.000	0.001
pa0089	*hsiH1*	T6SS HSI-I hypothetical protein	-0.038	0.948	0.971	2.608	0.000	0.002
pa0090	*clpv1*	Probable ClpA/B-type protease	0.168	0.673	0.816	1.746	0.000	0.002
pa0091	*Vgrg1*	T6SS HSI-I hypothetical protein	0.212	0.523	0.699	1.640	0.000	0.001
Type Three Secretion System genes								
pa1697	*pscN*	Type III SS related	0.500	0.097	0.281	2.920	1.88x10-7	4.24x10-6
pa1706	*ppcrV*	Type III SS related	0.770	0.008	0.050	0.370	0.150	0.240
pa1708	*popB*	Type III SS related	0.580	0.050	0.180	0.160	0.560	0.660
pa1709	*popD*	Type III SS related	0.770	0.030	0.020	0.240	0.270	0.370
pa1710	*exsC*	Type III SS related	0.620	0.010	0.080	-0.330	0.160	0.250
pa1711	*exsE*	Type III SS related	0.560	0.030	0.137	-0.050	0.800	0.850
pa1712	*exsB*	Type III SS related	0.500	0.014	0.080	-0.110	0.540	0.650
pa1713	*exsA*	Type III SS related	0.810	0.010	0.080	0.530	0.080	0.150
pa1714	*exsD*	Type III SS related	0.200	0.210	0.440	0.020	0.900	0.930
pa1716	*pscC*	Type III SS related	0.090	0.625	0.808	0.730	0.002	0.008
pa1719	*pscF*	Type III SS related	0.480	0.094	0.270	0.490	0.080	0.150
pa1723	*PscJ*	Type III SS related	0.110	0.450	0.690	1.080	4.01x10-6	4.58x10-5
pa1725	*pscL*	Type III SS related	0.040	0.880	0.940	0.010	0.950	0.970
pf1 bacteriophage genes								
pa0715		Hypothetical protein	0.589	0.106	0.297	-1.496	0.001	0.003
pa0716		Hypothetical protein	0.190	0.592	0.789	-1.277	0.001	0.003
pa0718		hypothetical protein of bacteriophage Pf1	-1.157	0.006	0.046	1.833	0.000	0.001
pa0719		hypothetical protein of bacteriophage Pf1	-1.774	0.002	0.023	1.692	0.000	0.002
pa0720		helix destabilizing protein of bacteriophage Pf1	-2.729	0.000	0.003	1.034	0.008	0.019
pa0721		hypothetical protein of bacteriophage Pf1	-0.970	0.013	0.075	0.963	0.012	0.025
pa0722		hypothetical protein of bacteriophage Pf1	-2.773	0.000	0.001	0.246	0.514	0.573
pa0723	*coaB*	coat protein B of bacteriophage Pf1	-2.891	0.000	0.001	0.192	0.655	0.705
pa0724		phage coat protein A	-0.099	0.701	0.853	1.643	0.001	0.003
pa0725		hypothetical protein of bacteriophage Pf1	-0.591	0.026	0.119	1.460	0.000	0.002
pa0726		hypothetical protein of bacteriophage Pf1	-0.324	0.287	0.536	1.463	0.001	0.005
pa0727		Pf replication initiator protein	0.140	0.541	0.750	0.413	0.168	0.222
pa0728		probable bacteriophage integrase	-0.065	0.816	0.916	0.442	0.158	0.211
pa0729		PfiT antitoxin, PfiA	0.463	0.118	0.320	-0.818	0.049	0.078

Appendix Table 1.9. A table of RNA sequence data obtained by Silke Alt that corresponds to Table 3.6. in Chapter 3. It looks at clusters of genes that are most highly up and downregulated in strains I and J. The left column labelled gene number is the annotated gene number as it appears in the chromosome, the column labelled gene name is the name given to a particular gene in literature and what it may be referred to as. The Log Fold Change (LogFC) column is the log-ratio of a gene's or a transcript's expression values compared to the WT strain. The LogFC value for each gene was obtained taking the average of three transcript readings for each gene in the Pa01 genome from which statistical anaylsis could be applied. To get the overall fold change in real terms compared to the WT, 2 is raised to the power of the LogFC value. The P value column indicated if the LogFC change of a given gene was of statistical significance if $P<0.05$. This was further statistically challenged by looking at the FDR value, also set at a significance level of $FDR<0.05$. Using both values ensured the LogFC value was statistically significant and not down to chance or false positives. Boxes highlighted in green call attention to a gene that has an increased LogFC change when compared to the WT and are statistically significant ($P<0.05$, $FDR<0.05$) suggesting the gene is more highly transcribed and upregulated compared to the WT. Boxes highlighted in red call attention to a gene that has a decrease LogFC change when compared to the WT and are statistically significant ($P<0.05$, $FDR<0.05$) suggesting the gene is more lowly transcribed and upregulated compared to the WT. Boxes that remain unhighlighted show genes whose LogFC change is not statistically significant when compared to the WT, therefore one could assume transcript levels of a given gene between the WT and a given strain are no different, neither up or down regulated.



Appendix Image 1.3. Prediction of what each PaMip most likely bind to using their amino acid sequences. No prediction tools predicted that any PaMip would bind DNA directly, evidencing they would likely not act directly as transcription factors, the lack of DNA binding is indicated in the yellow box whereby no part of the amino acid sequence possess DNA binding domains. All PaMips are predicted to somewhat bind to other proteins, something essential to their PPIase activity, these regions of the PaMip sequences likely to bind other proteins are coloured in blue and pink boxes which are overall highlighted in red.



Appendix Image 1.4. Protein alignment using the C-terminus of known FKBPs and Mips highlighting the required aspartate (yellow box) and tyrosine (green box) residues required for PPIase activity. These residues are lacking in the Pa4558 Mip (highlighted in the red boxes). As this Mip lacks these residues it is unlikely it would be able to perform PPIase activity.

2. Raw data for results presented in Chapter 4

	Haemolysis Plate 1 (mm)	Haemolysis Plate 2 (mm)	Haemolysis Plate 3 (mm)
WT	12	12.5	12
D	11.5	11.5	12
E	11.5	12	11.5
F	12	12	12
I	6	6	6
J	7	7	7
D : pUC18-Mini-TN7T-Gm *pamip1*	11.5	11.5	12
E : pUC18-Mini-TN7T-Gm *pamip2*	11.5	12	11.5
F : pUC18-Mini-TN7T-Gm *pamip3*	12	12	12
I : pUC18-Mini-TN7T-Gm *pamip1*	6	6	6
I : pUC18-Mini-TN7T-Gm *pamip2*	6	6	6
J : pUC18-Mini-TN7T-Gm *pamip2*	7	7	7
J : pUC18-Mini-TN7T-Gm *pamip3*	7	7	7
WT : pUC18-Mini-TN7T-Gm Empty	12.5	12.5	12
Control	5	5	5

Appendix Table 2.1. Raw data for haemolysis assays using pUC18-mini-TN7T-Gm complementation vectors.

	Protease Plate 1 (mm)	Protease Plate 2 (mm)	Protease Plate 3 (mm)
WT	14	14.5	14
D	14.5	14	14.5
E	14	14	14
F	14.5	14.5	14
I	7	6.5	7
J	7.5	7.5	7
D : pUC18-Mini-TN7T-Gm *pamip1*	14.5	14	14.5
E : pUC18-Mini-TN7T-Gm *pamip2*	14	14	14
F : pUC18-Mini-TN7T-Gm *pamip3*	14.5	14.5	14
I : pUC18-Mini-TN7T-Gm *pamip1*	7	6.5	7
I : pUC18-Mini-TN7T-Gm *pamip2*	7	6.5	7
J : pUC18-Mini-TN7T-Gm *pamip2*	7.5	7.5	7
J : pUC18-Mini-TN7T-Gm *pamip3*	7.5	7.5	7
WT : pUC18-Mini-TN7T-Gm Empty	14	14.5	14
Control	5	5	5

Appendix Table 2.2. Raw data for proteolysis assays using pUC18-mini-TN7T-Gm complementation vectors.

	Haemolysis Plate 1 (mm)	Haemolysis Plate 2 (mm)	Haemolysis Plate 3 (mm)
WT	12.5	12.5	12
D	12	11.5	12
E	11.5	11.5	11.5
F	11.5	12	12
I	6	6	6
J	7	7	7
D : pJH10TS *pamip1*	11.5	11.5	12
E : pJH10TS *pamip2*	11.5	12	11.5
F : pJH10TS *pamip3*	12	12	12
I : pJH10TS *pamip1*	5	5	5
I : pJH10TS *pamip2*	6	6	6
J : pJH10TS *pamip2*	6.5	6.5	6
J : pJH10TS *pamip3*	6	6	6
WT : pJH10TS Empty	5	5	5
Control	5	5	5

Appendix Table 2.3. Raw data for haemolysis assays using pJH10TS complementation vectors.

	Protease Plate 1 (mm)	Protease Plate 2 (mm)	Protease Plate 3 (mm)
WT	14	14.5	14.5
D	14.5	14	14
E	14	14	14.5
F	14.5	14.5	14
I	7	6.5	7
J	7.5	7.5	7
D : pJH10TS *pamip1*	14.5	14	14
E : pJH10TS *pamip2*	14.5	14	14.5
F : pJH10TS *pamip3*	14	14	14
I : pJH10TS *pamip1*	6	6	6
I : pJH10TS *pamip2*	5	5	5
J : pJH10TS *pamip2*	5	5	5
J : pJH10TS *pamip3*	5	5	5
WT : pJH10TS Empty	5	5	5
Control	5	5	5

Appendix Table 2.4. Raw data for proteolysis assays using pJH10TS complementation vectors

	Haemolysis Plate 1 (mm)	Haemolysis Plate 2 (mm)	Haemolysis Plate 3 (mm)
WT	13	13	13.5
D	12.5	13	13
E	13	13	13
F	13.5	13	13
I	6	6	6
J	8	8	7.5
I : pME6032 *pamip1*	12	11.5	12
I : pME6032 *pamip2*	6	6	6
I : pME6032 *pamip1* / *pamip2*	10	9.5	10
I : pME6032 *pamip2* / *pamip1*	12	12	11.5
I : pME6032 Empty	6	6	6
J : pME6032 *pamip2*	9.5	9	9
J : pME6032 *pamip3*	7	7	7.5
J : pME6032 *pamip2* / *pamip3*	7	7	7
J : pME6032 *pamip3* / *pamip2*	9.5	9.5	9
J: pME6032 Empty	7.5	8	7.5
WT : pME6032 Empty	13	13	13.5
Control	5	5	5

Appendix Table 2.5. Raw data for haemolysis assays using pME6032 complementation vectors

	Protease Plate 1 (mm)	Protease Plate 2 (mm)	Protease Plate 3 (mm)
WT	15.5	15.5	16
D	15.5	15	15.5
E	15.5	16	16
F	16	16	16
I	9	9.5	9
J	9.5	9.5	9.5
I : pME6032 *pamip1*	15.5	15.5	15.5
I : pME6032 *pamip2*	15	15	15
I : pME6032 *pamip1* / *pamip2*	15	15.5	15.5
I : pME6032 *pamip2* / *pamip1*	15	15	14.5
I : pME6032 Empty	9	9.5	9
J : pME6032 *pamip2*	14	14.5	14
J : pME6032 *pamip3*	15	15	14.5
J : pME6032 *pamip2* / *pamip3*	15.5	15	15
J : pME6032 *pamip3* / *pamip2*	14	14.5	14
J: pME6032 Empty	9.5	9.5	9.5
WT : pME6032 Empty	15.5	15.5	16
Control	5	5	5

Appendix Table 2.6. Raw data for proteolysis assays using pME6032 complementation vectors

	WT			D			E	
0.873	0.816	0.789	0.626	0.612	0.628	0.715	0.827	0.914
	F			I			J	
0.668	0.641	0.636	0.144	0.167	0.132	0.204	0.296	0.257
	I: pME6032 *pamip1*			I: pME6032 *pamip2*			I: pME6032 *pamip1* / *pamip2*	
0.34	0.321	0.391	0.255	0.228	0.223	0.285	0.259	0.257
	I: pME6032 *pamip2* / *pamip1*			I: pME6032 Empty			J: pME6032 *pamip2*	
0.335	0.353	0.37	0.191	0.179	0.172	0.142	0.15	0.187
	J: pME6032 *pamip3*			J: *pamip2* / *pamip3*			J: *pamip3* / *pamip2*	
0.108	0.101	0.112	0.119	0.112	0.102	0.11	0.135	0.192
	J: pME6032 Empty			WT: pME6032 Empty			Control	
0.124	0.11	0.126	0.766	0.796	0.831	0.098	0.098	0.098

Appendix Table 2.7. Raw data for biofilm formation assays using pME6032 complementation vectors

	Plate 1 (mm^2)	Plate 2 (mm^2)
WT	881.05	778.968
D	881.05	804.096
E	879.48	854.352
F	1586.9	1045.953
I	0	0
J	0	0
I : pME6032 *pamip1*	778.968	683.1675
I : pME6032 *pamip2*	0	0
I : pME6032 *pamip1* / *pamip2*	15.7	0
I : pME6032 *pamip2* / *pamip1*	683.1	637.623
I : pME6032 Empty	0	0
J : pME6032 *pamip2*	0	0
J : pME6032 *pamip3*	0	0
J : pME6032 *pamip2* / *pamip3*	0	0
J : pME6032 *pamip3* / *pamip2*	0	0
J: pME6032 Empty	0	0
WT : pME6032 Empty	1451.92	1133.901

Appendix Table 2.8. Raw data for twitching assays using pME6032 complementation vectors

6	12	18	24	30	36	42

WT	0	0	2	16	20	20	20
D	0	0	0	14	20	20	20
E	0	0	0	18	20	20	20
F	0	0	0	10	20	20	20
I	0	0	0	0	6	8	8
I : pME6032 *pamip1*	0	0	0	2	8	12	12
I : pME6032 *pamip2*	0	0	0	2	14	14	14
I : pME6032 *pamip1* / *pamip2*	0	0	0	4	8	16	18
I : pME6032 *pamip2* / *pamip1*	0	0	0	0	2	4	10
I : pME6032 Empty	0	0	0	0	8	10	10
J : pME6032 *pamip2*	0	0	0	0	10	10	10
J : pME6032 *pamip3*	0	0	0	0	8	14	14
J : pME6032 *pamip2* / *pamip3*	0	0	0	2	4	8	10
J : pME6032 *pamip3* / *pamip2*	0	0	0	0	12	20	20
J: pME6032 Empty	0	0	0	2	14	20	20
WT : pME6032 Empty	0	0	0	0	0	12	14
I : pME6032 *pamip1*	0	0	2	16	20	20	20
Control	0	0	0	0	0	0	0

Table 2.9. Raw data for the number of *G. mellonella* larvae counted as dead across two experiments. Each experiment was performed with 10 larvae per strain. This table is the pooled raw data across the two experiments from which statistics and conversion to a final percentage was performed that is presented in Figure 4.25.

3. Raw data for results presented in Chapter 5.

	Haemolysis Plate 1 (mm)	Haemolysis Plate2 (mm)	Haemolysis Plate 3 (mm)
WT	12.5	13	13
Δ*clpv3*	5	5	5.5
Δ*clpv3* : *clpv3*	11	11	11
Δ*clpv3:* pME6032 E	6	5	5
Δ*vgrg3*	11	11	11
Δ*clpv2*	13	13	12.5
Δ*vgrg2*	13	12.5	13
Δ*clpv1*	12.5	13	13
Δ*vgrg1*	12.5	13	13
WT pME6032	11.5	11.5	11
Control	5	5	5

Appendix Table 3.1. Raw data for haemolysis assays using T6SS knockout strains

	Protease Plate 1 (mm)	Protease Plate 2 (mm)	Protease Plate 3 (mm)
WT	16	15.5	16
Δ*clpv3*	8.5	9	8.5
Δ*clpv3* : *clpv3*	13.5	13	13.5
Δ*clpv3:* pME6032 E	9	9	8
Δ*vgrg3*	14	13.5	14
Δ*clpv2*	16	16	15.5
Δ*vgrg2*	16	16	16
Δ*clpv1*	15.5	16	16
Δ*vgrg1*	16	16	16
WT pME6032	14	14	14
Control	5	5	5

Appendix Table 3.2. Raw data for proteolysis assays using T6SS knockout strains

	Δ*clpv3*			Δ*clpv3* : *clpv3*			Δ*clpv3*: pME6032 E	
0.715	0.798	0.837	0.505	0.5	0.522	0.555	0.569	0.543
	Δ*vgrg3*			Δ*clpv2*			Δ*vgrg2*	
0.658	0.686	0.613	0.696	0.647	0.679	0.627	0.651	0.678
	Δ*clpv1*			Δ*vgrg1*			WT	
0.756	0.745	0.736	0.637	0.586	0.657	0.677	0.652	0.638
	WT pME6032			Control				
0.593	0.608	0.613	0.14	0.139	0.139			

Appendix Table 3.3. Raw data for biofilm formation assays using T6S knockout strains

	Plate 1 (mm^2)	Plate 2 (mm^2)
WT	678.32	649.43
Δ*clpv3*	0	0
Δ*clpv3* : *clpv3*	701.23	686.23
Δ*clpv3:* pME6032 E	58.59	34.23
Δ*vgrg3*	665.78	654.48
Δ*clpv2*	657.89	643.76
Δ*vgrg2*	645.78	639.45
Δ*clpv1*	697.34	681.42
Δ*vgrg1*	703.46	682.34
WT pME6032	698.27	676.1
Con	0	0

Appendix Table 3.4. Raw data for twitching assays using T6SS knockout strains

	6	12	18	24	30	36	42	48
WT (14)	0	0	4	14	14	14	14	14
Δ*clpv3* (16)	0	0	0	0	2	2	4	4
Δ*vgrg3* (16)	0	0	0	0	0	0	0	0
Δ*clpv2* (14)	0	0	0	6	8	10	10	5
Δ*vgrg2* (16)	0	0	0	6	6	14	14	14
Δ*clpv1* (16)	0	0	2	6	14	14	14	14
Δ*vgrg1* (18)	0	0	2	4	16	16	16	16
Con (14)	0	0	0	0	0	0	0	0

Table 3.5. Raw data for the number of *G. mellonella* larvae counted as dead across two experiments. Each experiment was performed with 'x' number larvae per strain (x = number next to the gene name). 10 larvae per strain per experiment was not possible so the larvae were split accordingly. This table is the pooled raw data across the two experiments from which statistics and conversion to a final percentage was performed that is presented in Figure 5.10.