

# **Post-translational modification of proteins secreted by *Pseudomonas aeruginosa* and its role in bacterial physiology**

**Suzanne Forrest**

Corpus Christi College

Department of Biochemistry  
University of Cambridge

March 2021

This thesis is submitted for the degree of *Doctor of Philosophy*





“The first step of any project is to grossly underestimate its complexity and difficulty.”

- Nicoll Hunt



# Declaration

This thesis is the result of my own work conducted at the University of Cambridge and includes nothing which is the outcome of work done in collaboration except as declared in the preface and specified in the text. It is not substantially the same as any work that I have submitted before for any degree or other qualification except as declared in the preface and specified in the text. Parts of the work presented in the Introduction chapter of this dissertation have been adapted from a previously published review in Science Progress; “Arming the troops: Post-translational modification of extracellular bacterial proteins” by S Forrest and M Welch, 2020. This review was authored by myself and edited by my supervisor, Dr Martin Welch. The length of this thesis does not exceed the prescribed word limit for the Biology Degree Committee.

S. Forrest

March 2021



# Abstract

## Post-translational modification of proteins secreted by *Pseudomonas aeruginosa* and its role in bacterial physiology

Suzanne L. Forrest

*Pseudomonas aeruginosa* is a multi-drug resistant, opportunistic pathogen capable of causing serious and potentially life-threatening infections worldwide. An important component of its infection strategy is the secretion of an extensive arsenal of virulence factors. These virulence factors cause considerable tissue damage and are also involved in attenuation of the host immune response, as well as inter-kingdom defences. Maintaining control over the production, secretion and activity of these proteins is therefore essential and is likely to lie not only at the transcriptional level, but also through post-translational modification (PTM). By understanding the nature of PTM in extracellular proteins and how these may affect protein function and stability, a future platform for alternative anti-virulence interventions can be developed.

In this dissertation I first looked at how environmental factors affect protein secretion in *Pseudomonas aeruginosa*. Although factors including nutrient availability, temperature, growth stage and strain type were found to modulate the overall proteomic output, the presence of charge variant protein isoforms proved to be a prominent and consistent biological phenomenon. Several key virulence factors presenting as charge variants were then selected for further analysis, and the nature of their PTMs was characterised by LC-MS/MS. All protein isoforms investigated were found to be multi-modified by acetylation and methylation, with a considerable level of overlap between these modifications in many cases. In the final part, site-directed mutagenesis of residues that are modified in the *P. aeruginosa* elastase enabled initial insights into how PTM may affect this protein's function. Phenotypic analyses revealed that all mutants displayed reduced elastinolytic activity, as well as a reduced ability to swarm for two of these mutants. In particular, mutation of arginine 226 appeared to have the most prominent effect, with potential consequences in the processing of other secreted virulence factors.

In conclusion, several principal virulence factors produced by *P. aeruginosa* were found to be consistently secreted as charge variable isoforms. These isoforms displayed extensive modification by acetylation and methylation and the potential roles these PTMs may have is discussed.



# Acknowledgements

I would firstly like to thank my supervisor Dr Martin Welch for the opportunity to work in his group and for his support and encouragement throughout the project. In addition, a huge thank you to the BBSRC and the Evelyn Trust for funding this research. Thank you to Corpus Christi College for being such a welcoming and friendly college and for their generous travel grant to allow attendance at the Pseudomonas Conference 2019. My gratitude also goes out to everyone at Page Medical, particularly Debbie Cockayne, for taking me on and for providing me with such an enjoyable PIPs internship.

This project would not have been possible without the expert assistance from the Cambridge Centre for Proteomics, in particular thanks to Renata Feret and Mike Deery for all your help with the proteomics experiments. Also thanks to our visiting researcher Sidra Abbas for providing me with your clinical isolates to test.

I would like to thank all of the members of the Welch group, both past and present, who provided a fun working environment and for your helpful advice when things felt like they would never work: Yassmin Abdelhamid, Larson Grimm, Rory Triniman, Alyssa McVey, Andre Wijaya, Stephen Dolan, Eve Maunders, Audrey Crousilles, Stephen Trigg, Tom O'Brien, Isaac On, Meng Wang, Wendy Figueroa-Chavez, Bruno Toscano, Natasha Palmer, Veena Mohan and Sean Bartlett. My gratitude also goes out to all members of the Salmond group for their insightful comments and suggestions around the lab and during group meetings.

To my friends, especially Ruby and Ella, for the endless amounts of fun and necessary stress relief, and also for being great listeners when all I needed to do was rant. Thanks to Nick, Sally, Tom and Laura for your encouragement and for looking after me so well during the tough final push. To my family, I would not be where I am today without your unwavering support and love, you guys are the best. In particular, thank you to my mum for inspiring me and sparking my interest in microbiology, and also for your excellent proof-reading skills. To Will for your tough love and for supporting me every single day. I could not have gotten through this without you. Also, for not getting too annoyed with me for my constant tech-related questions, you're a star. Finally, to my Dad. You may not have been here physically, but I know that your love and support will always be there. I hope I have made you all proud.





# Table of Contents

<b>Chapter 1: Introduction .....</b>	<b>1</b>
1.1 <i>Pseudomonas aeruginosa</i> .....	1
1.1.1 Acute infection with <i>P. aeruginosa</i> .....	2
1.1.2 Chronic infection with <i>P. aeruginosa</i> .....	4
1.1.2.1 Cystic fibrosis .....	4
1.1.3 Antibiotic resistance .....	6
1.1.3.1 Intrinsic resistance mechanisms .....	7
1.1.3.2 Acquired resistance mechanisms .....	8
1.1.4 Current treatments .....	9
1.2 Virulence factors .....	9
1.2.1 Cell-associated factors .....	10
1.2.2 Secreted factors .....	11
1.2.2.1 Rhamnolipids .....	12
1.2.2.2 Toxins.....	13
1.2.2.3 Exoproteases .....	14
1.2.2.4 Lipases.....	17
1.2.2.5 Pyocins.....	17
1.2.2.6 Other virulence factors .....	19
1.3 Secretion systems.....	21
1.3.1 Type I Secretion System .....	22
1.3.2 Type II Secretion System .....	22
1.3.3 Type III Secretion System .....	24
1.3.4 Type V Secretion System.....	24
1.3.5 Type VI Secretion System.....	25
1.3.6 Alternative secretion mechanisms .....	26
1.4 Post-translational modifications .....	27
1.4.1 Methylation.....	28
1.4.2 Acylation.....	31
1.4.3 Phosphorylation .....	34
1.4.4 Experimental constraints.....	36
1.5 Project Aims.....	38
<b>Chapter 2: Materials and Methods .....</b>	<b>39</b>
2.1 Bacterial strains .....	39
2.2 Growth conditions .....	40
2.2.1 Overnight culture .....	40
2.2.2 Planktonic culture .....	42

2.2.3	Growth on solid media .....	42
2.2.4	Measuring planktonic growth .....	42
2.2.5	Antibiotics .....	43
2.3	Proteomic analysis .....	43
2.3.1	Secretome extraction for 2D gels .....	43
2.3.2	Secretome column concentration .....	44
2.3.3	Quantification of proteins .....	44
2.3.4	1D SDS-PAGE .....	44
2.3.4.1	Coomassie blue staining .....	44
2.3.5	Western blotting .....	46
2.3.5.1	In-gel Western .....	46
2.3.6	2D gel electrophoresis .....	46
2.3.6.1	Silver staining .....	47
2.3.6.2	2D-DiGE .....	48
2.3.7	Tandem Mass Spectrometry (LC-M/SMS) .....	49
2.3.8	MASCOT database searches .....	50
2.3.9	Protein structure prediction and modelling .....	50
2.3.10	Bioinformatics .....	51
2.4	DNA manipulation techniques .....	51
2.4.1	DNA extraction and quantification .....	51
2.4.2	Polymerase chain reaction (PCR) .....	52
2.4.2.1	Colony PCR .....	53
2.4.3	Agarose gel electrophoresis .....	55
2.4.4	Restriction digest and DNA ligation .....	55
2.4.5	Transformation of <i>E. coli</i> by electroporation .....	56
2.4.5.1	Transformation of <i>P. aeruginosa</i> by electroporation .....	56
2.4.6	Confirmation of the transposon-insertion site .....	56
2.4.7	Construction of pLasB and p0622 complementation vectors .....	56
2.4.8	Site directed mutagenesis (SDM) .....	57
2.4.9	DNA sequencing .....	58
2.4.10	Expression of plasmid-encoded constructs .....	58
2.5	Phenotypic assays .....	58
2.5.1	Protease production .....	58
2.5.2	Elastase activity .....	58
2.5.3	Staphylolysin activity .....	59
2.5.4	Motility assays .....	59
2.5.4.1	Swarming motility .....	59
2.5.4.2	Swimming motility .....	59
2.5.5	Biofilm formation .....	60
2.5.6	Statistical analysis .....	60
2.6	<i>Caenorhabditis elegans</i> infection model .....	60
2.6.1	Synchronisation of the worm culture .....	60
2.6.2	Slow killing assay .....	61

2.7	Pyocin assays .....	62
2.7.1	Pyocin induction .....	62
2.7.2	Pyocin susceptibility testing .....	62

### **Chapter 3: Factors affecting the *P. aeruginosa* secretome ..... 63**

3.1	Introduction .....	63
3.2	Visualisation of the <i>P. aeruginosa</i> exoproteome .....	64
3.3	Validation of charge heterogeneity as a biological phenomenon .....	66
3.4	The effect of carbon source on charge variation .....	68
3.5	The effect of growth temperature on charge variation .....	71
3.6	The effect of growth stage on charge variation.....	73
3.7	Secretomes from clinical isolates .....	76
3.8	Discussion.....	79

### **Chapter 4: Acetylation of *P. aeruginosa* secreted proteins ..... 85**

4.1	Introduction .....	85
4.2	Identification of charge variant proteins from 2D gels.....	86
4.2.1	Hypothetical reasoning for two ImpA charge trains.....	90
4.3	Rationale for the selection of PTM to investigate .....	91
4.4	Acetylation of ImpA .....	93
4.4.1	Background of ImpA.....	93
4.4.2	Differences in ImpA proteoform abundance.....	93
4.4.3	Prediction of lysine acetylation.....	94
4.4.4	MS/MS identification of ImpA acetylation.....	95
4.4.5	Structural mapping of ImpA acetylation .....	99
4.4.6	Sequence conservation of ImpA acetylated lysines.....	102
4.4.7	ImpA acetylation motif analysis.....	103
4.4.8	ImpA K-ac secondary structure preference.....	104
4.5	Acetylation of LasB .....	105
4.5.1	Background of LasB.....	105
4.5.2	Differences in LasB proteoform abundance.....	105
4.5.3	Prediction of LasB lysine acetylation.....	106
4.5.4	MS/MS identification of LasB acetylation .....	107
4.5.5	Structural mapping of LasB acetylation.....	111
4.5.6	Sequence conservation of LasB acetylated lysines.....	111
4.5.7	LasB acetylation motif analysis .....	112
4.5.8	LasB K-ac secondary structure preference.....	113
4.6	Acetylation of PA0622 .....	114
4.6.1	Background of PA0622 .....	114
4.6.2	Differences in PA0622 proteoform abundance .....	115
4.6.3	Prediction of PA0622 lysine acetylation .....	116
4.6.4	MS/MS identification of PA0622 acetylation .....	117

4.6.5	Structural mapping of PA0622 acetylation .....	118
4.6.6	Sequence conservation of PA0622 acetylated lysines .....	120
4.6.7	PA0622 acetylation motif analysis .....	121
4.7	Acetylation of OprD .....	123
4.8	Discussion.....	124

## **Chapter 5: Methylation of *P. aeruginosa* secreted proteins..... 129**

5.1	Introduction .....	129
5.2	Methylation of ImpA .....	130
5.2.1	Prediction of ImpA methylation .....	130
5.2.2	MS/MS identification of ImpA methylation.....	132
5.2.3	Structural mapping of ImpA methylation .....	135
5.2.4	Methylation at the ImpA active site .....	138
5.2.5	ImpA methylation motif and sequence conservation.....	138
5.2.6	Methylation of an “unprocessed” ImpA sample .....	140
5.3	Methylation of LasB .....	142
5.3.1	Prediction of LasB methylation.....	142
5.3.2	MS/MS identification of LasB methylation .....	143
5.3.3	Structural mapping of LasB methylation .....	146
5.3.4	LasB methylation motif and sequence conservation.....	146
5.4	Methylation of PA0622.....	147
5.4.1	MS/MS identification of PA0622 methylation .....	147
5.4.2	Structural mapping of PA0622 methylation.....	147
5.4.3	PA0622 methylation motif and sequence conservation .....	150
5.5	Methylation of other secreted proteins.....	150
5.6	Immunoaffinity confirmation of methylation .....	152
5.7	Overview of PTM overlap.....	154
5.8	Discussion.....	155

## **Chapter 6: Initial insights into the functional consequences of PTM.. 163**

6.1	Introduction .....	163
6.2	Site-directed mutagenesis of key residues .....	164
6.2.1	Effect of LasB SDM on protein structure and pl.....	165
6.3	Secretion of the recombinant SDM proteins .....	166
6.4	Effect of LasB SDM on protease activity .....	168
6.4.1	Effect of LasB SDM on elastinolytic activity .....	170
6.5	Effect of LasB SDM on staphylytic activity .....	171
6.6	Effect of LasB SDM on biofilm formation .....	173
6.7	Effect of LasB SDM on motility .....	174
6.7.1	Swarming motility .....	174
6.7.2	Swimming motility .....	176
6.8	<i>Caenorhabditis elegans</i> infection model.....	177

6.9	Effect of SDM on R2 pyocin activity.....	179
6.10	Discussion.....	181
<b>Chapter 7: Final discussion.....</b>		<b>187</b>
<b>8</b>	<b>References .....</b>	<b>193</b>
<b>9</b>	<b>Appendices .....</b>	<b>220</b>



# List of Figures

## Chapter 1

Figure 1.1 Prevalence of organisms within the lungs of CF patients .....	5
Figure 1.2 Mechanisms of antibiotic resistance encoded by <i>P. aeruginosa</i> .....	7
Figure 1.3 Virulence factors secreted by <i>P. aeruginosa</i> .....	12
Figure 1.4 Overview of the LasB secretion pathway and proteolytic cascade .....	15
Figure 1.5 The secretion systems of <i>P. aeruginosa</i> and their respective substrates .....	21
Figure 1.6 The structure of small post-translational modifications (PTMs) .....	29

## Chapter 3

Figure 3.1 2-DE of the <i>P. aeruginosa</i> secretome .....	65
Figure 3.2 2-DE comparison of secretome extraction techniques .....	67
Figure 3.3 2D-DiGE of PAO1 secretome samples grown in different media .....	69
Figure 3.4 2D-DiGE of PAO1 secretome samples grown at different temperatures .....	72
Figure 3.5 2D-DiGE of PAO1 secretome samples from different time points .....	74
Figure 3.6 2D-DiGE of <i>P. aeruginosa</i> clinical isolate secretomes .....	78

## Chapter 4

Figure 4.1 Bottom-up proteomic strategy .....	86
Figure 4.2 Protein spot coding .....	87
Figure 4.3 Sequence coverage of ImpA .....	91
Figure 4.4 Close up image of charge trains A and B .....	94
Figure 4.5 Representative MS/MS spectra of ImpA acetyl-peptides .....	98
Figure 4.6 Structural mapping of acetylated lysines in spots A1-A5 .....	100
Figure 4.7 Structural mapping of acetylated lysines in spots B1-B6 .....	101
Figure 4.8 Sequence alignment of ImpA .....	102
Figure 4.9 ImpA acetylation motif analysis .....	103
Figure 4.10 Secondary structure at each K-ac site in ImpA .....	104
Figure 4.11 Close up image of charge train D .....	106

Figure 4.12 Representative MS/MS spectra of LasB acetyl-peptides .....	109
Figure 4.13 Structural mapping of acetylated lysines in spots D1-D4.....	110
Figure 4.14 Sequence alignment of LasB .....	112
Figure 4.15 LasB acetylation motif analysis.....	113
Figure 4.16 Secondary structure at each K-ac site in LasB .....	114
Figure 4.17 R2-pyocin operon and structure.....	115
Figure 4.18 Close up image of charge train C .....	116
Figure 4.19 Structural mapping of acetylated lysines in spots C1-C4.....	119
Figure 4.20 Structural mapping of PA0622 K-ac sites onto the pyocin trunk.....	120
Figure 4.21 Sequence alignment of PA0622 .....	121
Figure 4.22 PA0622 acetylation motif analysis .....	122
Figure 4.23 Secondary structure at the PA0622 K-ac sites .....	122
Figure 4.24 Structural mapping of OprD K-ac.....	123

## Chapter 5

Figure 5.1 Structural mapping of methylated residues identified in spots A1-A5 ....	136
Figure 5.2 Structural mapping of methylated residues identified in spots B1-B6 ....	137
Figure 5.3 Residues at the ImpA active site.....	138
Figure 5.4 ImpA methylation motif analysis .....	139
Figure 5.5 Structural mapping of methylated residues identified in spots D1 to D4	145
Figure 5.6 LasB methylation motif analysis.....	146
Figure 5.7 Structural mapping of methylated residues identified in spots C1 to C4	149
Figure 5.8 PA0622 methylation motif analysis .....	150
Figure 5.9 Methylation of other secreted proteins.....	152
Figure 5.10 Immunoaffinity confirmation of methylation.....	153
Figure 5.11 Evidence of overlap of PTM at the same residue .....	154
Figure 5.12 Overview of the PTMs identified for three <i>P. aeruginosa</i> secreted proteins .....	155
Figure 5.13 Change of thinking with respect to protein charge variation .....	161



## Chapter 6

Figure 6.1 Confirmation of LasB site-directed mutagenesis .....	164
Figure 6.2 Overlay of SDM predictions on the structure of LasB .....	166
Figure 6.3 Confirmation of LasB complementation and secretion .....	167
Figure 6.4 Protease activity of LasB and the SDM mutants.....	169
Figure 6.5 Elastin-Congo red (ECR) assay of LasB SDM mutants.....	171
Figure 6.6 Staphylytic activity of LasB and the SDM mutants.....	172
Figure 6.7 Biofilm formation of the LasB and SDM expressing strains .....	174
Figure 6.8 Swarming motility of LasB SDM mutants .....	175
Figure 6.9 Swimming motility of LasB SDM mutants .....	176
Figure 6.10 <i>Caenorhabditis elegans</i> infection model survival curve .....	178
Figure 6.11 PA0622 SDM mutants .....	180
Figure 6.12 Pyocin activity of PA0622 SDM mutants.....	180



# List of Tables

## Chapter 1

Table 1.1 Overview of key <i>P. aeruginosa</i> virulence factors .....	20
--	----

## Chapter 2

Table 2.1 List of bacterial strains.....	39
Table 2.2 Growth media .....	41
Table 2.3 Antibiotic solutions.....	43
Table 2.4 SDS-PAGE solutions and buffers.....	45
Table 2.5 2-DE parameters .....	47
Table 2.6 2-DE solutions and buffers .....	48
Table 2.7 Plasmids used in this research.....	52
Table 2.8 PCR reagents.....	53
Table 2.9 PCR parameters.....	53
Table 2.10 Oligonucleotide primers.....	54
Table 2.11 Restriction digestions and DNA ligation reaction volumes .....	55

## Chapter 4

Table 4.1 Protein spot identifications from LC-MS/MS data.....	88
Table 4.2 Mass and pI changes from typical PTMs .....	92
Table 4.3 Spot intensities of charge trains A and B .....	94
Table 4.4 ProAcePred prediction of K-ac in ImpA.....	95
Table 4.5 Acetylated lysines identified in spots A1 to A5 .....	96
Table 4.6 Acetylated lysines identified in spots B1 to B6 .....	97
Table 4.7 Spot intensities of charge train D .....	106
Table 4.8 ProAcePred prediction of lysine acetylation in LasB .....	107
Table 4.9 Acetylated lysines identified in spots D1 to D4.....	107
Table 4.10 Spot intensities of charge train C .....	116
Table 4.11 ProAcePred prediction of lysine acetylation in PA0622 .....	117

Table 4.12 Acetylated lysines identified in spots C1 to C4 .....	117
--	-----

## Chapter 5

Table 5.1 Prediction of ImpA methylation .....	131
Table 5.2 Methylated K/R/N/H identified in spots A1-A5.....	133
Table 5.3 Methylated K/R/N/H identified in spots B1 to B6.....	134
Table 5.4 Comparison of ImpA methylation from centrifugal concentration.....	141
Table 5.5 Prediction of LasB methylation.....	142
Table 5.6 Methylated K/R/N/H identified in spots D1 to D4 .....	144
Table 5.7 Methylated K/R/N/H identified in spots C1 to C4 .....	148

## Chapter 6

Table 6.1 Proteolytic zone diameter measurements.....	169
Table 6.2 Swimming migration of LasB SDM mutants.....	176
Table 6.3 Average lethal time in the <i>C. elegans</i> infection assay.....	178
Table 6.4 Overview of the phenotypic assay results.....	181

# Abbreviations

2-DE	2-Dimensional gel electrophoresis
2D-DiGE	2-Dimensional difference gel electrophoresis
ADP	Adenosine diphosphate
AGSY	Alanine, glycerol, salt, yeast medium
AIDS	Acquired immune deficiency syndrome
ARDS	Acute respiratory distress syndrome
ASB-14	Amidosulfo betaine-14
ATP	Adenosine triphosphate
BSA	Bovine serum albumin
CAIs	Community acquired infections
cAMP	Cyclic adenosine monophosphate
Cb <sup>r</sup>	Carbenicillin resistance
CCP	Cambridge Centre for Proteomics
CF	Cystic fibrosis
CFTR	Cystic fibrosis transmembrane conductance regulator
CIF	CFTR inhibitory factor
D	Aspartic acid
COPD	Chronic obstructive pulmonary disease
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DSSP	Dictionary of secondary structure of proteins
DTT	Dithiothreitol
E	Glutamic acid
ECL	Enhanced chemiluminescence
ECR	Elastin-Congo red
EDTA	Ethylenediaminetetracetic acid
eEF-2	Eukaryotic elongation factor 2
EV	Empty vector
H	Histidine
HAP	Hospital acquired pneumonia
HCD	Higher energy collisional dissociation
IEF	Isoelectric focusing
IPG	Immobilized pH gradient
K	Lysine
K-ac	Acetylated lysine
LB	Luria broth
LBA	Luria broth agar
LC-MS/MS	Liquid chromatography tandem mass spectrometry
LPS	Lipopolysaccharide
MIC	Minimum inhibitory concentration
MW	Molecular weight
N	Asparagine
NDK	Nucleoside diphosphate kinase
NGM	Nematode growth medium
OD	Optical density
OMP	Outer membrane protein
OMV	Outer membrane vesicle
ORF	Open reading frame

PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDB	Protein Data Bank
pI	Isoelectric point
PQS	Pseudomonas Quinolone Signal
PSGL-1	P-selectin glycoprotein ligand 1
PTM	Post-translational modification
PVDF	Polyvinylidene fluoride
R	Arginine
RND	Resistance-nodulation-division
SAM	S-adenosylmethionine
SD	Shine Dalgarno
SDM	Site-directed mutagenesis
SDS	Sodium dodecyl sulfate
SDS-PAGE	SDS-Polyacrylamide gel electrophoresis
SMA	Skimmed milk agar
T4P	Type IV pili
TAE	Tris-acetate-EDTA
TCA	Trichloroacetic acid
TEMED	Tetramethylethylenediamine
Tet <sup>r</sup>	Tetracycline resistance
TLR	Toll-like receptor
TxSS	Type x secretion system
UTI	Urinary tract infection
VAP	Ventilator-associated pneumonia
WHO	World Health Organisation
WT	Wild-type

# Chapter 1

## 1 Introduction

### 1.1 *Pseudomonas aeruginosa*

*Pseudomonas aeruginosa* is a Gram-negative rod-shaped bacterium of the *Gammaproteobacteria* class of organisms. First isolated in 1882 by Gessard, it can be found ubiquitously throughout the environment where it inhabits a range of different niches (Gessard, 1984). Due to its metabolic versatility and minimal nutritional requirements, *P. aeruginosa* thrives in soil and aquatic environments as well as colonising nematodes, plants and animals (Frimmersdorf *et al.*, 2010). The ability of *P. aeruginosa* to infect a whole host of species implicates it not only as a pathogen of considerable medical importance, but also an organism with critical environmental significance (Walker *et al.*, 2004). Whilst it normally maintains the role of a harmless microorganism in healthy human individuals, certain conditions can induce a switch to pathogenicity, leading to potentially fatal opportunistic infections (Bassetti *et al.*, 2018). As a result of high infection rates globally, it persists as a considerable clinical and economic burden worldwide (Nathwani *et al.*, 2014).

*P. aeruginosa* is motile by means of a single polar flagellum, which promotes swimming in liquid environments, as well as type IV pili which are necessary for solid-surface twitching motility (Murray and Kazmierczak, 2008; Khan *et al.*, 2020). The bacterium is also able to swarm across semi-solid surfaces in a coordinated colony effort that requires cell to cell signalling. This motility is typically in response to changes in nutritional availabilities and is dependent on bacterial cell density and surfactant production (Kohler *et al.*, 2000; Overhage *et al.*, 2008). Motility is important for pathogenicity as it allows microbial colonisation in different environments as well as dispersal to occupy other niches. It also allows the bacterium to attach to surfaces and form robust biofilms (Yeung *et al.*, 2009).

Biofilm formation involves the transition from a planktonic (“free swimming”) state into a sessile aggregative form. Within the biofilm, microbial communities are protected

by an extracellular polymeric matrix made up of polysaccharides, proteins and DNA. This matrix acts as a reservoir of nutritional metabolites and allows the exchange of DNA and chemical messengers for community cooperation (Mann and Wozniak, 2012). It also forms a physical shield to protect cells from environmental stressors as a means of enhancing population stability (Mulcahy *et al.*, 2014). Although the expression of virulence-associated genes is typically downregulated in a biofilm, this state promotes increased resistance to antibiotics and enables evasion from the host immune response (Klausen *et al.*, 2003; Sharma *et al.*, 2014; Vital-Lopez *et al.*, 2015). The biofilm microenvironment is therefore thought to be a major player in chronic infection and antibiotic treatment failure.

Infections from *P. aeruginosa* are common in the hospital setting due to the organism's proclivity to grow and form biofilms on medical equipment. Devices with high levels of moisture, such as catheters and ventilators, are particularly susceptible to colonisation (Maurice *et al.*, 2018). The hardness of the biofilm environment means that they are difficult to eradicate, favouring persistence of the pathogen and facilitating transmission between patients. Its resilience and ubiquity are so extensive that it has even been isolated from hospital supplies of distilled water and antiseptic solutions (Favero *et al.*, 1971). As such, it has been designated as a key nosocomial pathogen and is one of the leading causes of infection-associated hospital morbidities (Weinstein *et al.*, 2005; Hidron *et al.*, 2008; Shi *et al.*, 2019).

### **1.1.1 Acute infection with *P. aeruginosa***

The shift from harmless organism to virulent pathogen tends to occur where there is a breach of host barriers to infection (Babich *et al.*, 2020). This is a common problem in patients with extensive burn wounds, where the exposed deep epidermal layers become contaminated. Colonisation and proliferation of *P. aeruginosa* within the damaged tissue is often extensive and can lead to further serious complications including pseudomonal bacteraemia and subsequent sepsis (Lyczak *et al.*, 2000; Bassetti *et al.*, 2018). Deterioration often occurs rapidly and can lead to death within a short period of time (Turner *et al.*, 2014). Similarly, *P. aeruginosa* is also commonly associated with surgical wound site and other skin and soft tissue infections (Ruffin and Brochiero, 2019).



The use of invasive medical equipment is associated with higher incidences of infection. Urinary tract infections (UTIs) from the use of indwelling catheters are the most common hospital-acquired infections, of which *P. aeruginosa* is the third leading cause (Mittal *et al.*, 2009; Weiner *et al.*, 2016). In contrast, community-acquired UTIs are rarely caused by this pathogen (Gupta *et al.*, 2001). Respiratory tract infections by *P. aeruginosa* are the most frequent and arguably the most serious (Gellatly and Hancock, 2013). Acute ventilator-associated pneumonia (VAP) is habitually a consequence of direct trauma to the airway epithelium from endotracheal intubation. The development of biofilms on ventilation tubing and internally on mucosal surfaces enables bacterial persistence and makes both immune clearance and treatment particularly challenging (Williams *et al.*, 2010; Pericolini *et al.*, 2018). Alongside hospital-acquired pneumonia (HAP), these *Pseudomonas*-derived pneumonias are consistently correlated with poor prognoses and high fatality. The development of secondary complications, such as bacteraemia or acute respiratory distress syndrome (ARDS), further exacerbates the threat to life (Weiner *et al.*, 2016).

Disruption of the immune system in immunocompromised patients is another significant risk factor for acute infection with *P. aeruginosa*. Individuals with conditions such as acquired immune deficiency syndrome (AIDS), neutropenic patients or those receiving pharmacological immunosuppression (e.g. cancer chemotherapy or transplant patients) are of particular concern (Mendelson *et al.*, 1994). There is normally a strong neutrophilic response to infection and therefore neutrophil depletion predisposes patients to an increased risk of serious bacteraemia (Skopelja-Gardner *et al.*, 2019). Where bacteraemia does occur, mortality can be as high as 61% (Kim *et al.*, 2014).

Problematically, immunocompetency is not a fail-safe against *P. aeruginosa* associated morbidity. Acute community-acquired infections (CAIs) in healthy individuals are relatively common, usually occurring if the bacterial inoculum is sufficient for pathology. Water-dwelling *P. aeruginosa* is the cause of hot tub folliculitis, and it is also frequently implicated in otitis externa, or “swimmer’s ear” (Lutz and Lee, 2011; Guida *et al.*, 2016). This infection of the external ear canal causes inflammation and can temporarily result in hearing loss (Sundström *et al.*, 1996). Otitis externa

arising from *P. aeruginosa* infection is also a prevalent problem in dogs (Pye, 2018). Additionally, insufficient hygienic practice when handling contact lenses can often lead to painful ulcerative keratitis caused by *P. aeruginosa* contamination. The bacterium is able to survive on the lens surface as well as within the storage solution (Stapleton *et al.*, 1995; Fleiszig and Evans, 2010; Wu *et al.*, 2015). Corneal ulceration as a result of ocular colonisation by *P. aeruginosa* is typically more severe than from other aetiological bacterial agents. This leads to more serious visual disruption and is typically more difficult to treat (Sy *et al.*, 2012).

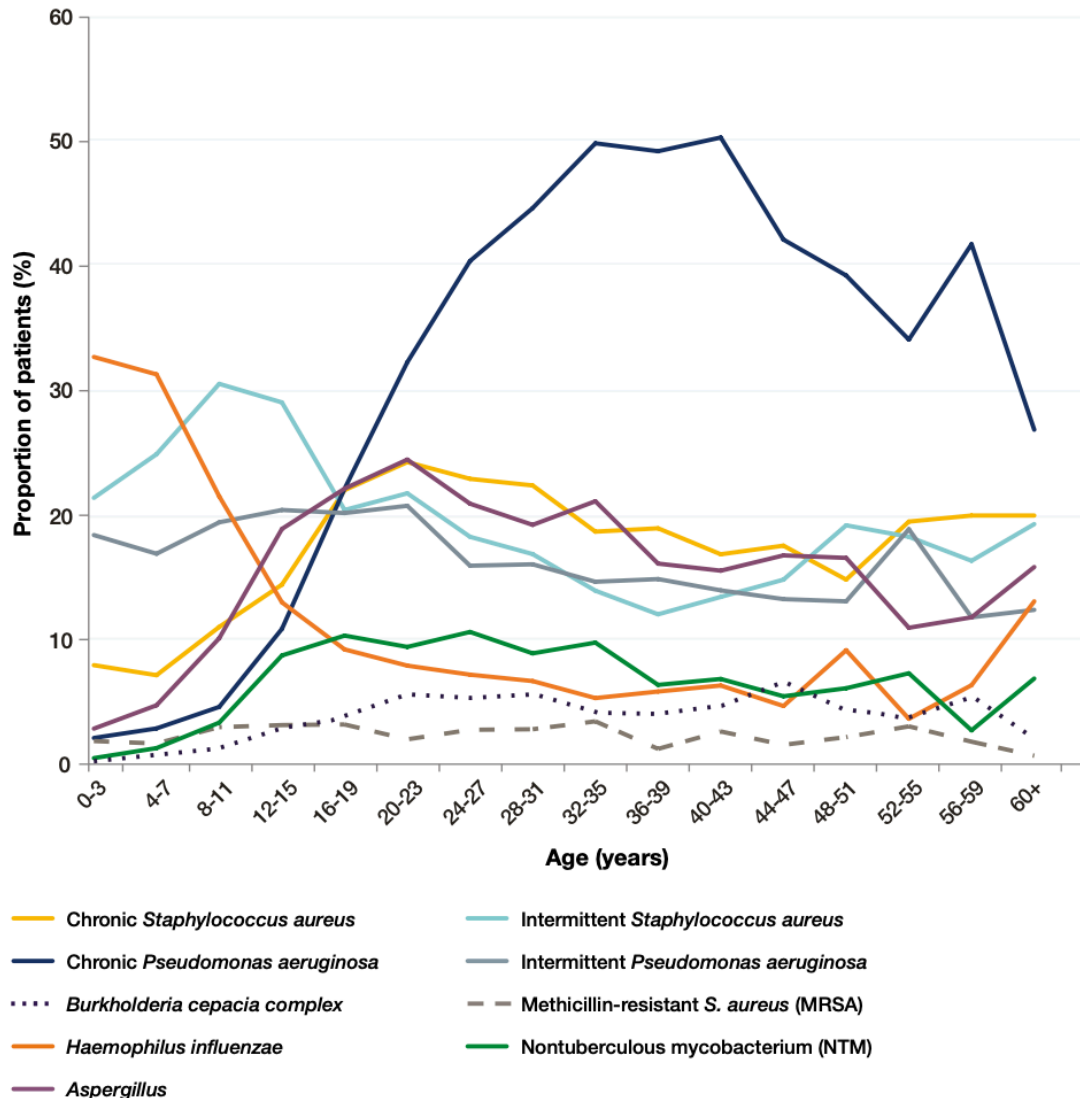
### **1.1.2 Chronic infection with *P. aeruginosa***

When host defences are unable to eradicate *P. aeruginosa*, and invasive acute infection is not fatal, development into the chronic phase of infection can occur (Balasubramanian *et al.*, 2014). Long-term colonisation tends to arise principally in the airways and is most common in individuals with underlying respiratory pathoses, such as bronchiectasis or chronic obstructive pulmonary disease (COPD) (Angrill *et al.*, 2002; Faure *et al.*, 2018). In both conditions, the accumulation of mucus in the airways provides a favourable warm and moist environment for the establishment of *P. aeruginosa* biofilms (Lorenz *et al.*, 2019). Persistence within the airways is maintained through a complex interplay between host and pathogen but is also associated with a poorer overall prognosis. By far the most widely researched respiratory condition associated with *P. aeruginosa* is cystic fibrosis.

#### **1.1.2.1 Cystic fibrosis**

Cystic fibrosis (CF) is a genetic disease of autosomal recessive inheritance that is most commonly found in Caucasian populations, affecting around 1 in 3,000 births in Europe (Farrell, 2008). Mutations in both copies of the cystic fibrosis transmembrane conductance regulator (CFTR) gene cause distinct and debilitating pathologies; mutation in only one allele confers healthy carrier status (Sosnay *et al.*, 2013). The CFTR gene encodes a chloride ion channel with a variety of functions (Bhagirath *et al.*, 2016). There is widespread expression of the CFTR protein throughout the body, including in the pancreas, liver, and gastrointestinal system, resulting in the vast range of clinical symptoms. Although life expectancy has dramatically increased in CF in the

last few decades, a significant proportion of patients still die at a relatively young age due to bacterial lung infection and its associated pulmonary tissue damage (Keogh *et al.*, 2018; Malhotra *et al.*, 2019a).



**Figure 1.1 Prevalence of organisms within the lungs of CF patients**

Prevalence of organisms within the lungs of CF patients by age group, n = 10070. During early infection, *S. aureus* and *H. influenzae* are common, but by the late teenage years, chronic infection with *P. aeruginosa* is dominant. Data taken from the UK Cystic Fibrosis Registry Annual Data Report 2019.

Bacterial colonisation in CF usually occurs at an early age, and whilst the lung microenvironment remains consistently polymicrobial, there are stark fluctuations in the abundances of certain species over time (McDaniel *et al.*, 2015). *Staphylococcus*

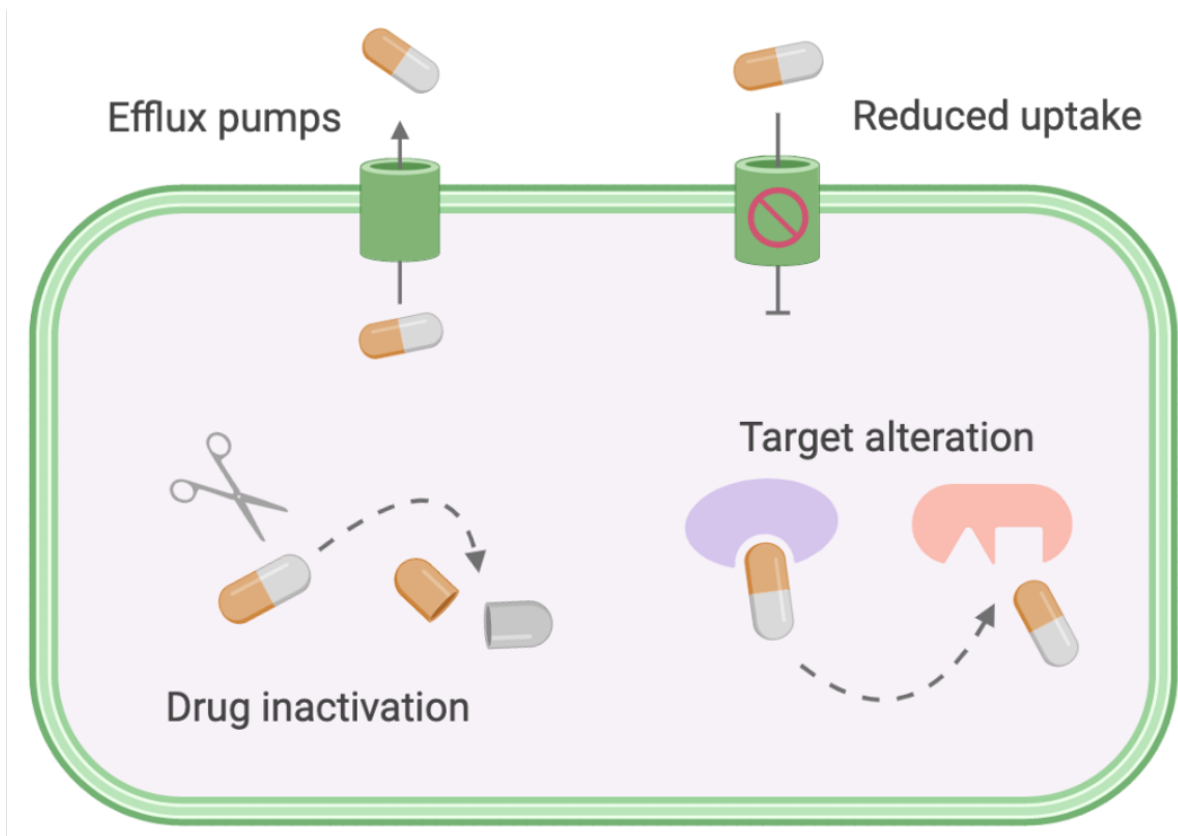
*aureus*, *Haemophilus influenzae* and *Aspergillus* are commonly isolated, yet *P. aeruginosa* is usually the dominant species by the late teenage years (Figure 1.1). In fact, at least 80% of adult CF patients harbour *P. aeruginosa* within their lungs (Lyczak *et al.*, 2002). Intermittent highly destructive acute infections tend to give way to chronic microbial establishment within biofilms (Faure *et al.*, 2018). Concerningly, a previous study indicated that the age of onset of chronic *P. aeruginosa* infection positively correlated with age of death (Frederiksen *et al.*, 1996; Clark *et al.*, 2015). The presence of lipids, proteins and ions within the thick and sticky mucus provides a nutrient rich environment for bacterial propagation, and also promotes comprehensive phenotypic diversity within the constituent species (Gellatly and Hancock, 2013).

Pulmonary infection by *P. aeruginosa* dramatically accelerates the deterioration of lung function (Henry *et al.*, 1992). Infection instigates a strong inflammatory response that ultimately results in pronounced tissue damage. In addition, bacterial invaders secrete an array of virulence factors which can enhance this destruction (Tingpej *et al.*, 2007). Whilst the development of mucoid biofilms during the chronic phase is typically accompanied by reduced virulence and a lower (yet chronic) immune response, pulmonary exacerbations in which symptoms dramatically worsen become much more frequent with age (Malhotra *et al.*, 2019b). Prolonged prophylactic antibiotic therapy is typically prescribed to control infectious exacerbations, however this not only prolongs adverse drug side effects but also drives the emergence of drug resistant mutations (Stickland *et al.*, 2010).

### **1.1.3 Antibiotic resistance**

The 20<sup>th</sup> century has been dubbed the “antibiotic era” due to the rapid expansion of available chemotherapies to treat bacterial infections. Unfortunately, due to misuse, over prescription and gratuitous use in the agricultural industry, resistance to these drugs is now widespread. A lack of discovery of new and effective antibiotics in the last 30 years, combined with an increasing aged and immunocompromised population, is cause for serious concern. *P. aeruginosa* is one of the ESKAPE “super bug” pathogens, so named due to their widespread multi-drug resistance and clinical significance (Santajit and Indrawattana, 2016; Mulani *et al.*, 2019). In 2017, the World Health Organisation (WHO) ranked *P. aeruginosa* in the top priority “critical” category

for which new antibiotics are urgently required (Tacconelli *et al.*, 2018) . The WHO further stated that this pathogen is one of the greatest bacterial threats to human health (World Health Organisation, 2017). One reason that *P. aeruginosa* is so formidable is because of its multiple intrinsic and acquired resistance mechanisms (Figure 1.2).



**Figure 1.2 Mechanisms of antibiotic resistance encoded by *P. aeruginosa***

*P. aeruginosa* typically harbours multiple intrinsic and acquired resistance mechanisms, including the expression of efflux pumps, reduced uptake of drugs through downregulation or inactivation of porins, inactivation of drugs and the alteration of drug targets so that drugs are no longer effective.

#### 1.1.3.1 Intrinsic resistance mechanisms

Intrinsic antibiotic resistance is an organism's innate chromosomally-encoded mechanisms by which it evades the action of antibiotics (Pang *et al.*, 2019). Low outer

membrane permeability is one such structural mechanism whereupon *P. aeruginosa* increases its resistance to various antibiotics. Studies have shown that the permeability of the *P. aeruginosa* outer membrane is up to 100-fold lower than that of *E. coli* (Hancock and Brinkman, 2002). *P. aeruginosa* also encodes mechanisms by which it can expel toxic compounds out of the cell. Twelve efflux pumps of the resistance-nodulation-division (RND) family have so far been identified in *P. aeruginosa*, four of which are involved in antibiotic resistance (Dreier and Ruggerone, 2015; Daury *et al.*, 2016). These efflux pumps extrude a broad range of antibiotic classes, including  $\beta$ -lactams, quinolones and aminoglycosides (Masuda *et al.*, 2000). Clinically isolated strains of *P. aeruginosa* have been found to upregulate the expression of these drug efflux pumps (Llanes *et al.*, 2004).

Another intrinsic mechanism employed by *P. aeruginosa* involves the production of enzymes capable of modifying or breaking down antibiotic compounds so that they are no longer active. The most well-known group of enzymes are  $\beta$ -lactamases. These enzymes hydrolyse the  $\beta$ -lactam ring of  $\beta$ -lactam antibiotics, including penicillins, cephalosporins and cephamycins (Ullah *et al.*, 2017). The main  $\beta$ -lactamase produced by *Pseudomonas* is encoded by the inducible *ampC* gene (Balasubramanian *et al.*, 2014). *P. aeruginosa* also produces aminoglycoside-modifying enzymes as an additional mechanism to prevent the activity of the aminoglycoside class of antibiotics. Drug inactivation occurs via phosphorylation, acetylation or adenylation of the amino or hydroxyl groups within the compound structure (Poole, 2005).

#### **1.1.3.2 Acquired resistance mechanisms**

Alongside intrinsic mechanisms of antibiotic resistance, *P. aeruginosa* is also able to acquire resistance through mutational changes or through the acquisition of new resistance genes (Hwang and Yoon, 2019). There are countless mutations that can occur to reduce bacterial susceptibility to antibiotics (Henrichfreise *et al.*, 2007). Mutations can serve to change the expression levels of key resistance proteins, such as downregulating outer membrane porins, or overexpressing efflux pumps and deactivating enzymes. Mutation can also induce a change in the function of proteins or even change the target site of a drug such that it is no longer able to bind

(Fernández and Hancock, 2012). Under selective pressures, such as nutrient depletion or antibiotic attack, the rate of mutation increases to enhance the chances of survival. This is a particular problem when chronic therapy is required, for example in CF patients (Oz *et al.*, 2014). Bacteria can also assimilate novel resistance cassettes in the form of plasmids, transposons, integrons or prophages. Horizontal transfer of genetic elements occurs either through conjugation, transformation or transduction from the same or different bacterial species. The acquisition of mutations or external DNA helps to reinforce and support natural intrinsic mechanisms and can foster resistance to multiple drug types (Pang *et al.*, 2019).

#### **1.1.4 Current treatments**

Due to the widespread multi-drug resistance of *P. aeruginosa*, it is currently recommended that prescribed therapies are guided by an antibiogram if possible. The determination of isolate sensitivities improves patient outcomes compared to when purely empirical treatment is sought. Patients should be monitored regularly throughout the course of antibiotic therapy to detect the emergence of resistance and the choice of antibiotic kept under close review. Monitoring of any side effects is also essential to provide best care (Horcajada *et al.*, 2019). Recommended antibiotics vary depending on local guidelines, but a combination therapy of a  $\beta$ -lactam and an aminoglycoside is commonly advised (Mulani *et al.*, 2019). The additional use of  $\beta$ -lactam inhibitors such as sulbactam should also be encouraged to improve the efficacy of this therapy regimen.

### **1.2 Virulence factors**

*P. aeruginosa* has one of the largest genomes of all sequenced bacteria at between 5.5-7 Mbp encoded on a single circular chromosome. Size variation occurs due to the chromosomal accessory genome and acquisition of different extrachromosomal plasmids (Stover *et al.*, 2000; Klockgether *et al.*, 2011). The 5,700+ open reading frames (ORFs) include numerous genes that are important for pathogenicity. The production of an extensive arsenal of virulence factors is vital for *P. aeruginosa* survival and colonisation, as well as the establishment and maintenance of infection (Ballok and O'Toole, 2013). They are also an important component of interkingdom

defences against competing organisms. For this reason, virulence factors are increasingly being explored as potential drug targets.

The expression of virulence-associated genes is generally not constitutive but is regulated in response to changes in the environment. More specifically, virulence genes tend to be upregulated in stressful conditions, such as in low nutrient environments or upon external attack (Breidenstein *et al.*, 2011). This regulation of gene expression is controlled by the hierarchical and interconnected quorum sensing network, composed of the *las*, *rhl* and PQS systems. Perception of population densities allows community coordination and occurs through the secretion of autoinducer chemical messengers which activate cognate gene regulators once a critical threshold is reached (Lee and Zhang, 2015). Virulence is most prominent during acute infection when planktonic growth is favoured, and then subsequent downregulation is observed during chronic infection and within the biofilm microenvironment. This phenotypic switch converts the aggressive, invasive strategy into a method of immune evasion. As the production of virulence factors is also metabolically costly, downregulation preserves vital nutrient stores for infection maintenance (Smith *et al.*, 2006; Moradali *et al.*, 2017).

### **1.2.1 Cell-associated factors**

Cell-associated virulence determinants are components which remain within or attached to the cell. They are typically bound to the outer membrane and are most often involved in cell adhesion and motility. The monotrichous flagellum, for example, is primarily used for swimming motility and chemotaxis. It is composed of external glycosylated flagellin (FliC) proteins, anchored by a protein hook to a membrane embedded motor protein complex (Hayakawa, 2012). Flagella enable the binding to mucosa via host mucins and glycolipids, which are enriched in repairing epithelia (Arora *et al.*, 1998; Bucior *et al.*, 2012). They are also strongly immunogenic through recognition by TLR-5 and TLR-2 of innate immune cells. Interestingly, the TLR-5 binding domain is not exposed in the intact flagellar shaft, therefore suggesting that monomeric release of flagellin subunits stimulates inflammation (Campodónico *et al.*, 2010; Patankar *et al.*, 2013). Aflagellated *P. aeruginosa* are defective in acute



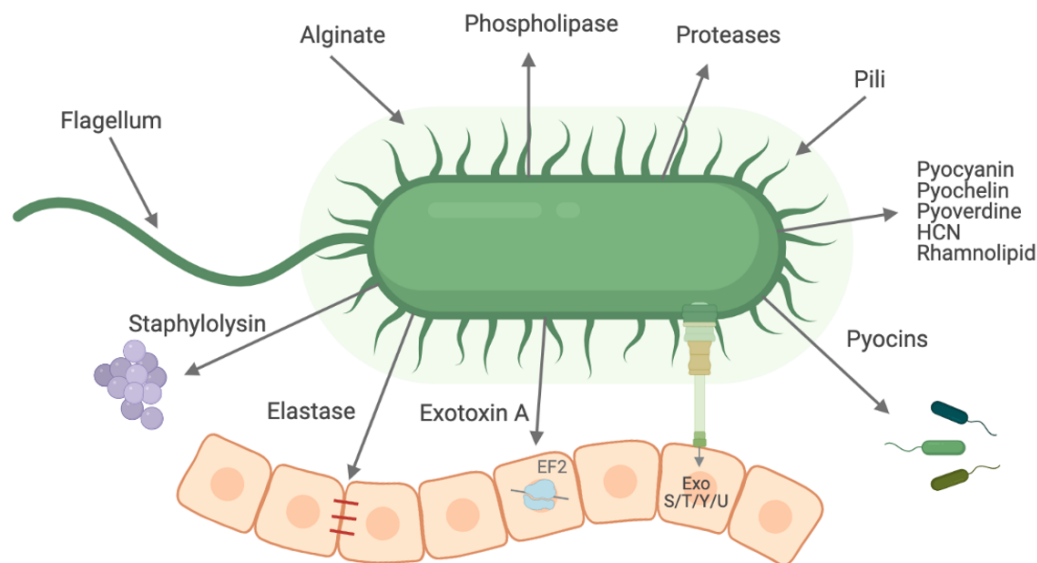
infection strategies, demonstrating the pathogenic importance of this appendage (Garcia *et al.*, 2018; Lorenz *et al.*, 2019).

Similarly, type IV pili (T4P) are also major surface adhesins responsible for motility and more selective attachment to host cells. As well as initiating destructive inflammatory responses, they can also promote cell aggregation and the formation of concentrated microcolonies which are more resistant to immune attack (Gellatly and Hancock, 2013; Nguyen *et al.*, 2015). Antigenic variation of the pilin subunit additionally helps to reduce host immune memory and promote evasion. Pilial detection of surface contact furthermore triggers a signalling cascade that promotes the expression of other virulence-associated genes and fosters positive feedback for continued cell attachment (Persat *et al.*, 2015).

The major endotoxin of *P. aeruginosa* is the complex lipopolysaccharide (LPS) that forms a major part of the external outer membrane. LPS is comprised of a lipid A anchor, a core oligosaccharide and an O-antigen polysaccharide branch. Each domain can be highly variable, yet the specific composition of the O-antigen is used for strain serotyping. There are pleiotropic effects associated with the presence of LPS, including acting as a physical barrier for the exclusion of certain molecules and moderating interactions with antibiotic compounds. Most notably however, LPS is strongly pro-inflammatory, which can be rapidly devastating by initiating septic shock in infected patients (King *et al.*, 2009; Huszczyński *et al.*, 2020).

### **1.2.2 Secreted factors**

*P. aeruginosa* secretes a wealth of proteins and compounds through its various secretion systems (Figure 1.3 and Figure 1.5). These secreted virulence factors play integral roles in all stages of infection and survival and are an important component of the bacterium's fundamental biology that must be thoroughly understood.



**Figure 1.3 Virulence factors secreted by *P. aeruginosa***

*P. aeruginosa* secretes a wealth of proteins and compounds into the extracellular environment to aid successful colonisation, including toxins, various proteases and lipases and small compound siderophores.

### 1.2.2.1 Rhamnolipids

Rhamnolipids are secreted virulence factors that have a many different functions. The glycolipids act as biosurfactants by lowering surface tension and allowing swarming motility of cells (Caiazza *et al.*, 2005). Their amphipathic composition permits integration into the host cell membrane, causing disruption of tight junctions and impairment of the epithelial barrier. This in turn induces ciliostasis and also allows *P. aeruginosa* paracellular access for adherence and then internalisation which can prove fatal to the cell. Rhamnolipids can additionally inhibit the action of phagocytes as well as cause the necrotic death of several key immune mediators (Zulianello *et al.*, 2006; Jensen *et al.*, 2007).

Membrane intercalation can likewise occur in other prokaryotes and simple eukaryotes, causing pore formation and subsequent lysis. This antimicrobial activity has been shown against a range of Gram-positive and Gram-negative organisms, as well as fungi and viruses. Alongside disruption of other species' biofilms, these actions confer a strong competitive advantage to *P. aeruginosa* (Haba *et al.*, 2003; Wood *et*

*al.*, 2018). Rhamnolipids can also promote *P. aeruginosa* microcolony formation as well as seed dispersal from mature biofilms and are important for the maintenance of internal channels to allow the flow of oxygen and nutrients deep within the biofilm matrix (Davey *et al.*, 2003; Newman *et al.*, 2017).

#### 1.2.2.2 Toxins

The two main classical exotoxins of *P. aeruginosa* are exotoxin A and exoenzyme S. Exotoxin A is a secreted A/B protein composed of an enzymatic domain (A) and a cell binding domain (B), similar to toxins produced by *Corynebacterium diphtheriae* and *Vibrio cholerae*. Following endocytosis into the host cell, exotoxin A inactivates eukaryotic elongation factor 2 (eEF-2) through ADP-ribosylation. This halts protein synthesis, causing cell cycle arrest and swift apoptosis. Local tissue necrosis at the site of infection is therefore inevitable and the toxin can mediate development into systemic disease. As such, it is one of the most lethal weapons within the *P. aeruginosa* armoury and has been investigated as an anti-cancer agent (Morlon-Guyot *et al.*, 2009; Michalska and Wolf, 2015).

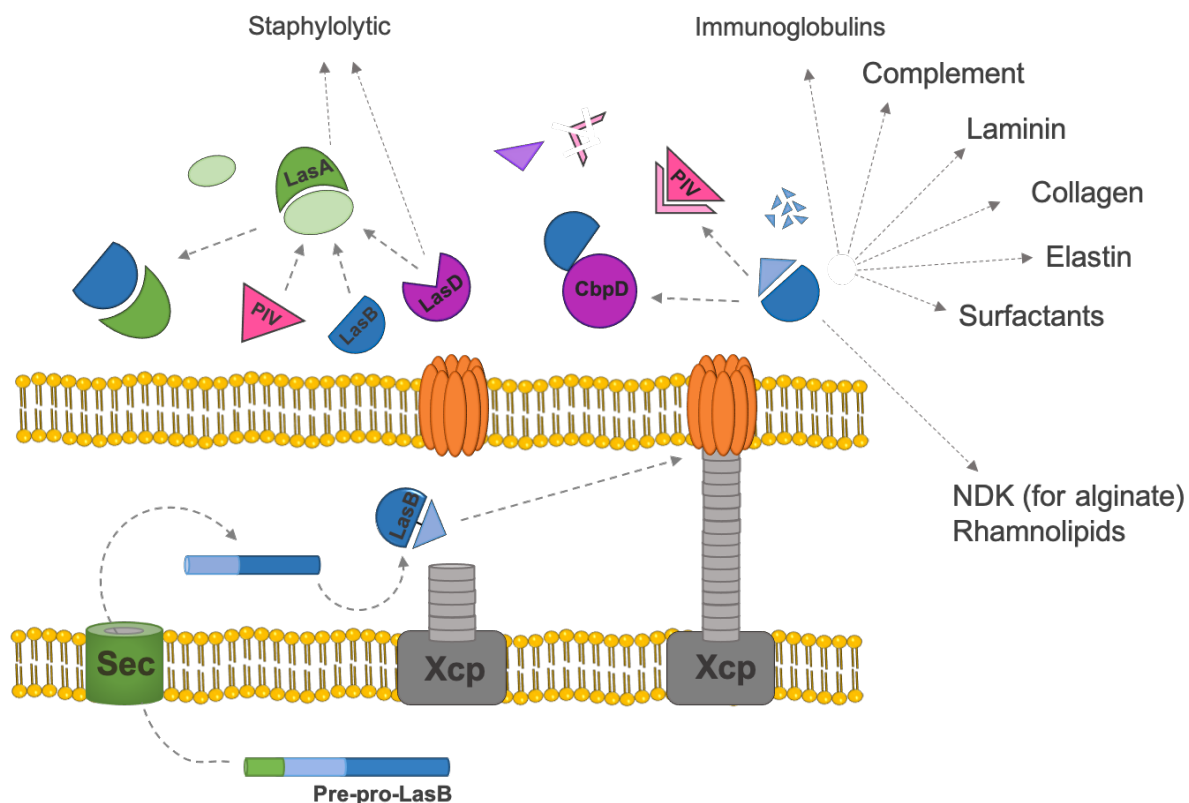
Exoenzyme S (ExoS) is actually one of four known exotoxins which are injected directly into the host cytoplasm via the type three secretion system (T3SS) of *P. aeruginosa*. There is distinct heterogeneity in production of these four effectors, the combination of which confers phenotypic strain diversity. Whilst most strains encode both ExoT and ExoY, ExoU and ExoS appear to be mutually exclusive in the genome. ExoS and ExoT are bifunctional cytotoxins that both contain an N-terminal GTPase activating domain and a C-terminal ADP-ribosyl transferase domain. The difference is grounded in the ADP-ribosylation targets, where the range of ExoT substrates is more restricted (Engel and Balachandran, 2009). The action of these two domains triggers actin cytoskeleton depolymerisation, inhibiting phagocytosis by immune cells and causing cell detachment. ExoS in particular is associated with slow cytotoxicity and cell death (Soong *et al.*, 2008). Conversely, ExoU provokes rapid necrosis by rupturing eukaryotic membranes with its phospholipase activity. In this way it is also able to promote pronounced immunosuppression for advanced bacterial infiltration. ExoY, an adenylate cyclase, is less well characterised but contributes to heightened endothelial permeability and cytoskeleton disruption through increased intracellular

cAMP. The expression of any of the T3SS effectors is associated with advanced and severe pathology in infected individuals (Hauser, 2009; Kloth *et al.*, 2018).

### 1.2.2.3 Exoproteases

One of the most extensive classes of virulence factors is the exoproteases. Approximately 3% of the *P. aeruginosa* genome encodes proteases, which have a variety of different biological functions and pathological effects (Stover *et al.*, 2000). There are several recognised proteases that are secreted during infection, including alkaline protease (AprA), protease IV (PrpL), large extracellular protease A (LepA), small protease (PASP), aminopeptidase (PAAP) and Immunomodulating metalloprotease (ImpA). In general, these proteolytic enzymes are relatively promiscuous in terms of their substrate specificities, capable of breaking down a diverse range of host proteins throughout the body (Moradali *et al.*, 2017; Faure *et al.*, 2018). This includes various components of the innate immune system such as immunoglobulins and interferons, as well as structural and functional tissue proteins including laminin and fibrinogen (Marquart, Caballero, *et al.*, 2005; Laarman *et al.*, 2012). As such, these enzymes have established roles in the *P. aeruginosa* infection strategy through tissue damage and immune evasion. *In vivo* infection studies have supported how important these virulence factors are for *P. aeruginosa* colonisation and its disruption of numerous biological niches (Döring *et al.*, 1983; Guillon *et al.*, 2017; O'callaghan *et al.*, 2019).

Elastase B, also known as pseudolysin or LasB, is perhaps the most well-known of the secreted proteases. It is a zinc metalloprotease of the M4 thermolysin peptidase family of enzymes. Encoded by the *lasB* gene, elastase is produced as a pre-pro-protein with an N-terminal signal peptide for secretion via the T2SS (Figure 1.4). Upon translocation from the cytoplasm across the inner membrane by the Sec machinery, the signal peptide is cleaved. Autocleavage of the pro-domain then occurs in the periplasm following protein folding. This pro-domain remains non-covalently bound to the mature protein to chaperone secretion via the Xcp machinery into the extracellular environment whereupon it dissociates and degrades. By maintaining an interaction with the mature elastase, the pro-domain inhibits any intracellular proteolytic activity (Braun *et al.*, 1998; McIver *et al.*, 2004).



**Figure 1.4 Overview of the LasB secretion pathway and proteolytic cascade**

Overview of the LasB secretion pathway and subsequent proteolytic activation cascade. The LasB pre-pro-protein is transported into the periplasm through the Sec translocon via an encoded signal peptide. Following protein folding, the pro-domain is cleaved and remains non-covalently bound to the mature protein. Secretion through the T2SS Xcp machinery then occurs, after which the pro-domain is degraded. In the extracellular environment, mature LasB cleaves CbpD into LasD and activates Protease IV (PIV), all of which play a role in the activation of LasA. Mature LasA is then able to enhance the elastolytic activity of LasB. There are many other host-derived substrates of LasB activity, indicated by the arrows on the right. LasB also has intracellular roles in regulating rhamnolipid and alginate production. Adapted from (Forrest and Welch, 2020).

There is an abundance of elastase targets that have so far been discovered, and research still continues to highlight novel substrates. As the name suggests, a notable target of elastase is elastin. This protein is a key component of the extracellular matrix that provides elasticity to connective tissues throughout the body, including the skin, major internal organs and the vascular system (Cathcart *et al.*, 2011). It can also degrade type III and IV collagen, fibrin, and disrupt tight junctions between cells through the debilitation of occludins, claudins and tricellulin (Golovkine *et al.*, 2018). This leads to profuse tissue damage, a loss of epithelial and endothelial integrity and

increased permeability to invading bacteria. Through the inactivation of many key components of the innate and adaptive immune defences, elastase also enhances *P. aeruginosa*'s immune evasion tactics. The degradation of both immunoglobulins (A and G) and surfactant proteins (A and D) prevents bacterial opsonisation and thus deescalates the threat of phagocytosis. It can also downregulate complement molecules, interleukins and cytokines (Kuang *et al.*, 2011; Bastaert *et al.*, 2018). Finally, elastase and alkaline protease together remove exogenous flagellin monomers to reduce host immune recognition (Casilag *et al.*, 2015). The significance of elastase in infection and virulence therefore cannot be understated.

On top of activating host proteases to initiate self-injury, LasB also activates other *P. aeruginosa* proteases. LasA, or staphylolysin, is a further zinc metalloprotease but of the M23 family. Activation occurs through an elegant cleavage cascade in which LasB cleaves the pro-peptide of PrpL, and then together these mature active proteases cleave LasA into its active form (Oh *et al.*, 2017; Li and Lee, 2019). LasB also activates a secreted chitin binding protein (CbpD) which is cleaved into another staphylolysin, LasD. Whilst LasA possesses its own intrinsic elastolytic activity, it mainly acts to enhance the more powerful elastase activity of LasB (Peters and Galloway, 1990). LasA can also hydrolyse a range of glycine-containing proteins, particularly the penta-glycine bridges in the peptidoglycan cell wall of Staphylococci. Rapid lysis of *S. aureus* cells through action of LasA can occur, and it is therefore an important defence mechanism against competing bacteria (Kessler *et al.*, 1992; Spencer *et al.*, 2010).

A novel protease was recently identified that primarily interferes with the host immune response; it was consequently named the immunomodulating metalloprotease (ImpA). This protease cleaves P-selectin glycoprotein ligand 1 (PSGL-1), CD43 and CD44 expressed on the majority of leukocytes. These receptors interact with P- and E-selectin which are upregulated on endothelial cells following recognition of foreign bacterial structures and the production of pro-inflammatory mediators. Cleavage of these receptors by ImpA prevents leukocyte rolling over endothelial cells and subsequent extravasation. This ultimately limits leukocyte homing to the site of infection and can inhibit phagocytosis by neutrophils and macrophages (Bardoel *et al.*, 2012). Interestingly, despite being secreted through the T2SS, ImpA is under the

regulation of the T3SS transcriptional activator ExsA, indicating that protein secretion is an intricately interconnected process (Tian *et al.*, 2019).

#### 1.2.2.4 Lipases

*P. aeruginosa* furthermore exports many different lipase and phospholipase enzymes. The T2SS is used for the secretion of three phospholipases C; PlcB, PlcH (haemolytic) and PlcN (non-haemolytic). Through hydrolysis of the phosphodiester bond in the phospholipid backbone, PlcH and PlcN can target eukaryotic cell membranes and cause cell lysis (Terada *et al.*, 1999). This mechanism can also be used to break down lung surfactant. Other eukaryotic substrates, including phosphatidylcholine, sphingomyelin and phosphatidylserine, are found within the erythrocyte membrane, marking these cells as a lethal target (Cota-Gomez *et al.*, 1997). It is not currently known why PlcN is non-haemolytic despite its homology to PlcH. PlcH is upregulated in the majority of CF isolates, and has been linked with a decline in lung function with increased exacerbations (Wargo *et al.*, 2011; Newman *et al.*, 2017). Another set of phospholipases are exported through the T6SS. The eukaryotic-like phospholipases D (PldA), alongside PldB, degrade phosphatidylethanolamine in the bacterial membrane and have therefore been implicated in inter-kingdom bacterial warfare (Barker *et al.*, 2004; Russell *et al.*, 2013). PldA and PldB are injected directly into other prokaryotes whilst the producing strain harbours cognate immunity proteins for protection (Jiang *et al.*, 2014).

#### 1.2.2.5 Pyocins

Pyocins are largely underappreciated virulence factors produced by *P. aeruginosa* that are deployed for antagonistic niche occupation. They are a type of bacteriocin (or protein antibiotic) with a narrow killing spectrum of the same or closely related species. These munitions are produced by more than 90% of *P. aeruginosa* isolates and enable control of population dynamics within heterogeneous communities, particularly under anaerobiosis (Heo *et al.*, 2007; Waite and Curtis, 2009). There are four main families of pyocin which differ in their morphology and mode of action. These are the R-type (rod), F-type (flexuous), S-type (soluble) and L-type (lectin), however there is a common lack of consistency in the understanding of their evolutionary biology which

results in inconsistent nomenclature (Lemieux *et al.*, 2016). Production is induced during the SOS response. The RecA repair protein mediates autocleavage of the negative pyocin regulator PrtR in response to DNA damage. This in turn derepresses PrtN, the positive regulator, which can then bind to a P-box consensus motif upstream of the pyocin biosynthetic gene clusters and initiate pyocin expression (Penterman *et al.*, 2014). Their bactericidal activity has been of clinical interest for many years and has proven more efficacious in a murine lung infection model than inhaled tobramycin, as well as extensively protective in a murine peritonitis model (Matsui *et al.*, 1993; Scholl and Martin, 2008; McCaughey *et al.*, 2016).

The R- and F-type pyocins, commonly referred to as tailocins, are morphologically and genetically related to the *Myoviridae* and *Siphoviridae* bacteriophage families respectively. Nakayama *et al.* ascertained their ancestral relationship to the P2 and  $\lambda$  phages. These pyocins are not defective phages however, as has been reported, but have evolved into a specialised *P. aeruginosa* defence mechanism in their own right (Nakayama *et al.*, 2000; Behrens *et al.*, 2017). Five R-type and nine F-type pyocins have been documented so far based on their producing strain, amino acid sequence and killing spectrum. In laboratory strain PAO1, R2 and F2 are encoded adjacently in a gene cluster between *trpE* and *trpG* of the tryptophan operon. The pyocin loci are composed of open reading frames for structural, regulatory and chaperone genes, as well as a lysis cassette for release of the pyocin particles (Scholl and Martin, 2008; Buth *et al.*, 2018).

R-type pyocins are non-flexible, contractile multi-protein complexes that kill susceptible cells through pore formation. Binding to the cell via the tail fibres causes contraction of the pyocin sheath, which inserts the core needle-like structure through the cell envelope. This dissipates the proton motive force, depolarises the membrane and finally results in cell lysis (Redero *et al.*, 2018). The C-terminal domain of the tail fibres mediates binding to LPS as a receptor, which determines their strain specificity (Williams *et al.*, 2008). Immunity to the action of R-pyocins is thought to be due to alteration of the LPS O-antigen residues (Köhler *et al.*, 2010). R-pyocins have been shown to be important for *P. aeruginosa* survival and dominance in infection models and competition assays and are also able to disrupt established biofilms (Oluyombo *et al.*, 2019). Comparatively, F-pyocins are vastly understudied. They are structurally



flexible, non-contractile rods with fewer tail fibres than R-pyocins. Although the mechanism of action has not been elucidated, it is assumed that they also cause membrane depolarisation. R-pyocins show superior bactericidal activity as only one or two pyocins are required to kill one bacterium, compared with 300 for the F-pyocins (Heo *et al.*, 2007; Redero *et al.*, 2018).

#### **1.2.2.6 Other virulence factors**

The synthesis and secretion of hydrogen cyanide by *P. aeruginosa* occurs under microaerobic conditions. This lethal asphyxiant can diffuse through tissues where it impairs cellular respiration; levels of HCN in the CF lung can be as high as 130  $\mu\text{M}$ , well within the toxic range (Anderson *et al.*, 2010). Furthermore, replication of CF pathology is possible through secretion of a CFTR inhibitory factor (CIF). The epoxide hydrolase activity of CIF promotes ubiquitin-mediated degradation of eukaryotic ABC transporters, including CFTR and P-glycoprotein (Bomberger *et al.*, 2011).

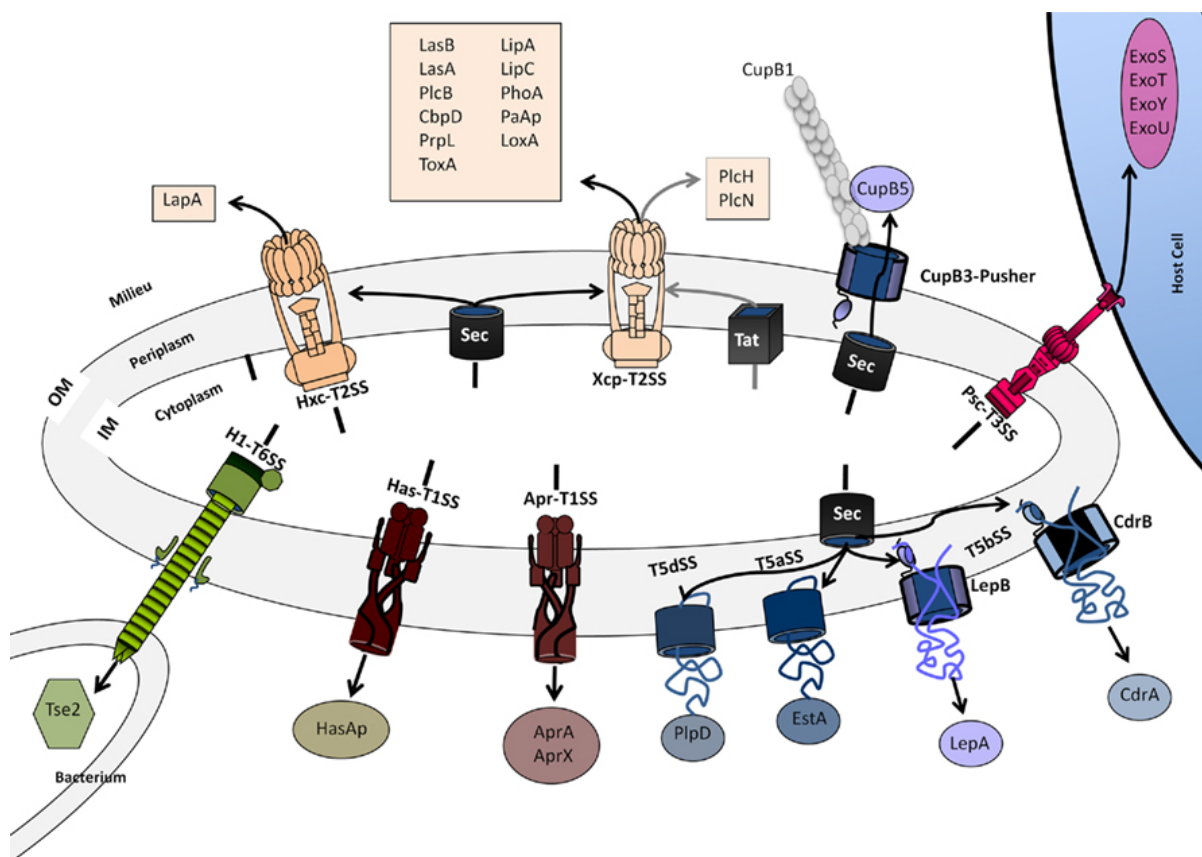
*P. aeruginosa* secretes small molecule chelators called siderophores to scavenge essential iron from the surrounding environment; pyochelin and pyoverdine are the major siderophores (Luscher *et al.*, 2018). It can also uptake metabolically costly siderophores produced by other bacterial species and sequester iron for competitive advantage (Schiessl *et al.*, 2017). A switch to phenazine-based iron uptake system is induced under microaerobic and anaerobic conditions (Cornelis and Dingemans, 2013). Pyocyanin, the most well studied phenazine, is largely responsible for the distinctive blue-green colouration of *Pseudomonas* cultures. The removal of iron from mitochondria results in significant damage, mitophagy and ultimately cell necrosis (Kirienko *et al.*, 2015; Kang *et al.*, 2018). Pyocyanin can also generate reactive oxygen species which affects calcium homeostasis and depletes host cyclic adenosine monophosphate (cAMP) and adenosine triphosphate (ATP) (Ho Sui *et al.*, 2012; Hall *et al.*, 2016). They therefore play a critical role in infection (Cezairliyan *et al.*, 2013; Schiessl *et al.*, 2019). The extensive array of known virulence factors produced by *P. aeruginosa* is nonetheless expanding, with novel determinants frequently being unearthed (Zrieq *et al.*, 2015).

**Table 1.1 Overview of key *P. aeruginosa* virulence factors**

<b>Virulence factor</b>	<b>Biological activities</b>
Flagellum	Motility, chemotaxis, and binding to host mucosa
Type IV pili	Cell surface adhesion, twitching motility, biofilm formation and immune evasion
LPS	Pro-inflammatory resulting in endotoxic shock
Rhamnolipid	Biosurfactant for motility, tight junction disruption, immune cell destruction, disruption of ciliary function, antimicrobial activity and involved in biofilm formation
Exotoxin A	Inhibition of eukaryotic protein synthesis through inactivation of eEF-2, cytotoxic, involved in tissue damage and invasion
LasA	Staphylolysin activity and enhancement of LasB elastinolytic activity
LasB (Elastase)	Protease and elastase activity, degradation of a wide range of host proteins (including collagen and fibrin), including lung surfactant, inactivation of complement, proteolysis of immunoglobulins and activation of other virulence factors
Protease IV	Degradation of immunoglobulins, complement and surfactant and activation of other virulence factors
Alkaline protease	Broad spectrum protease involved in immune evasion
ImpA	Inhibition of leukocyte homing via cleavage of PSGL-1, CD43 and CD44
Phospholipase C	Eukaryotic membrane disruption, breakdown of lung surfactant and lysis of erythrocytes
ExoS	Actin cytoskeleton depolymerisation, phagocytosis inhibition, cell detachment, apoptosis and necrosis
ExoT	Actin cytoskeleton depolymerisation, phagocytosis inhibition, cell detachment, apoptosis and necrosis
ExoU	Phospholipase activity, rupture of epithelial barriers and cell necrosis
ExoY	Disruption of eukaryotic cell actin cytoskeletons and increase in endothelial permeability
Siderophores (Pyoverdine and Pyochelin)	Iron acquisition and sequestration, and inhibition of mitochondrial function
Phenazines (e.g. Pyocyanin)	Redox activity, oxidative damage to host cells, cytotoxic and immune disruptive
Phospholipase C	Cytolytic, destruction of cell membranes and tissue invasion
Hydrogen cyanide	Suppression of aerobic respiration and cytotoxicity
Pyocins (R-, F-, S- and L-type)	Bacteriocin with bactericidal activity, involved in antagonistic niche occupation

### 1.3 Secretion systems

It is essential for bacteria to be able to transport substrates to different cellular compartments, into the extracellular environment or inject them directly into other prokaryotic or eukaryotic cells. This distribution ensures optimal substrate activity and is necessary for the majority of cellular processes. For proteins to be fully secreted, they must traverse at least two phospholipid membranes. *P. aeruginosa* encodes five of the eight classes of known bacterial secretion systems, through which approximately 20% of the proteome is secreted (Figure 1.5) (Bendtsen *et al.*, 2005; Lewenza *et al.*, 2005). These molecular nanomachines are responsible for the secretion of specific consignments of substrates, which are predominantly virulence factors. The tight regulation of secretion and its machinery lies in highly complex networks, including the integration of the two-component and quorum sensing systems (Zhu *et al.*, 2016).



**Figure 1.5** The secretion systems of *P. aeruginosa* and their respective substrates  
Adapted from (Filloux, 2011)

The type 1, 3 and 6 secretion systems involve the direct one-step delivery of cargo proteins to the extracellular environment. In comparison, proteins secreted by the type 2 and 5 secretion systems must first be transported from the cytoplasm into the periplasm before then being exported via their dedicated outer membrane machinery (Costa *et al.*, 2015). The first step is accomplished using the general secretory (Sec) pathway for unfolded proteins, or the twin-arginine translocation (Tat) pathway for folded proteins; both are resident in the inner membrane. Similar to the other secretion systems, the Sec and Tat translocases are membrane-bound protein complexes driven by the activity of specialised ATPases. Pre-proteins are targeted through these membrane channels by an N-terminally encoded signal sequence of approximately 20 residues in length. Signal peptide cleavage following translocation into the periplasm releases cargo proteins for assisted onward transport across the outer membrane (Green and Mecsas, 2016).

### **1.3.1 Type I Secretion System**

The T1SS is a simple mechanism for one-step direct export of proteins out of the cell. It is formed of three main components, an inner membrane ATP-binding cassette (ABC) transporter, an outer membrane protein (OMP) and a periplasmic adaptor protein that connects the two. A protein's non-cleavable C-terminal secretion signal directs interaction with the ABC transporter which in turn hydrolyses ATP to provide energy for transport. The unfolded protein is then transported out of the cell via the outer membrane pore (Bleves *et al.*, 2010). In *P. aeruginosa* there are currently three characterised T1SSs, with other putative systems suggested through genome mining (Filloux, 2011). The AprA and AprX proteases and the HasA haemophore (for iron scavenging) are secreted through the Apr and Has systems respectively (Létoffé *et al.*, 1998; Duong *et al.*, 2001). Recently, the system responsible for the secretion of small GTPase-inhibiting virulence factor TesG was elucidated. Export via the TesA/B/C T1SS apparatus is associated with the development of chronic lung infection through host immune suppression (Zhao *et al.*, 2019).

### 1.3.2 Type II Secretion System

T2SSs consist of complex multiprotein nanomachines for the secretion of several different large proteins. These systems utilise the two-step mechanism of export, initially involving exoprotein translocation across the inner membrane via the Sec or Tat pathway, as outlined above. The T2SS architecture is composed of several proteins known indiscriminately as general secretory proteins (GSPs). In *P. aeruginosa* the two main T2SSs are the Xcp and Hxc systems, formed by the Xcp and Hxc GSPs respectively (Figure 1.5). Although these systems are structurally homologous, Hxc is only utilised in phosphate-limiting conditions and exclusively secretes alkaline phosphatase LapA (Ball *et al.*, 2002).

The Xcp proteins make up four main subassemblies of the secretion system, or secreton (Robert *et al.*, 2005). The XcpPSYZ inner membrane complex has various roles, principally as a controlled gateway for exoprotein export and to mediate between various system components (Fulara *et al.*, 2018). The cytoplasmic XcpR hexameric ATPase closely interacts with the inner membrane complex and provides energy through ATP hydrolysis to drive the export process. An outer membrane secretin is formed by 12 monomers of XcpQ, assembling into a  $\beta$ -barrel pore through which proteins are translocated (Douzi *et al.*, 2017). Finally, the pseudopilus is constructed from the major pseudopilin XcpT, with four minor pilins XcpUVWX at the tip (Y. Zhang *et al.*, 2018). These pilin-like proteins are N-terminally processed by the peptidase XcpA (also known as PilD) for functional activation.

Through the concerted effort of these GSPs, proteins are pushed out of the cell in a piston-like motion. After folding in the periplasm (if Sec-translocated), proteins can then enter the secretion machinery. A secretion motif that targets periplasmic substrates to the T2SS has not been found, yet some studies have instead suggested that the signal for secretion lies in the protein's conformation (Douzi *et al.*, 2012). Assembly of the pseudopilus through polymerisation of XcpT subunits then forces the protein to traverse through the secretin and into the extracellular milieu (Durand *et al.*, 2011). Subsequent disassembly resets the piston to allow the expulsion of other proteins. The T2SS is employed in the secretion of many different Sec-dependent virulence factors, including proteases such as LasB and LasA, exotoxin A, lipases

and phosphatases. The phospholipase C enzymes also use the Xcp system but are transported pre-folded into the periplasm through the Tat transporter (Filloux, 2011).

### 1.3.3 Type III Secretion System

The T3SS is arguably the most renowned of the *P. aeruginosa* secretion systems due to its well-established roles in acute infection. T3S involves the one-step passage of effector proteins across both bacterial membranes and the eukaryotic membrane for direct injection into the host cell. Its activity is triggered by contact with host cells or by low environmental calcium concentrations. This is regulated through a complex signalling network overseen by the master transcriptional regulator ExsA (Williams McMackin *et al.*, 2019). Four T3SS effectors have so far been characterised, ExoT, ExoY, ExoS and ExoU; these have been explored in greater detail earlier in the chapter (section 1.2.2.2). There are, however, novel putative effectors now coming to light (Burstein *et al.*, 2015).

The injectisome apparatus is structurally similar to the flagellar machinery and is made up of over 20 different proteins encoded across five clustered operons (Hauser, 2009). The basal body is a multimeric complex that spans across both bacterial membranes and functions as a sorting platform for structural and effector proteins. This is in close association with the cytosolic ATPase PscN which provides energy for the transport of effectors and also to enable effector uncoupling from their respective pre-secretion chaperones (Halder *et al.*, 2019). The basal body encompasses a hollow needle-like filament intracellularly. This needle is a polymer of PscF protein subunits and is the vector through which cargo proteins transit across membranes. Additionally, the secretin PscC assembles into a membrane channel through which the needle can protrude from the cell surface. At the tip of the needle is PcrV, a protein which controls the folding and function of the translocation complex proteins PopB and PopD (Sato *et al.*, 2011). These PopB/D translocon proteins form a pore in the target eukaryotic membrane to allow effector migration through the needle and into the host cytosol (Tang *et al.*, 2018).

#### 1.3.4 Type V Secretion System

The T5SS is unique in that the secretion machinery is encoded contiguously with the effector protein, typically within a single polypeptide chain. This two-step system initially relies on native polypeptide transport into the periplasm via the Sec translocon. From here, there are three broad mechanisms by which proteins are then secreted (Leo *et al.*, 2012; Meuskens *et al.*, 2019). The simplest is the autotransporter system, T5aSS, T5cSS or T5eSS. Proteins exported in this way are synthesised in a modular fashion with a C-terminal translocator domain and an N-terminal passenger domain connected by a linker. The translocator forms a  $\beta$ -barrel channel in the outer membrane through which the functional passenger domain can exit the cell. Following export, the passenger domain either remains bound to the cell surface or is proteolytically cleaved and released. This is the case in the secretion of esterase EstA and aminopeptidase AaaA (Wilhelm *et al.*, 2007; Luckett *et al.*, 2012). The T5bSS, also known as two-partner secretion, is similar except that the passenger domain (TpsA) and translocator  $\beta$ -barrel (TpsB) are separate proteins. Protease LepA and adhesin CdrA are secreted in this way via their cognate translocators LepB and CdrB, respectively. T5dSS is a hybrid of the 5a and 5b systems and is responsible for the lipolytic PlpD protein (Salacha *et al.*, 2010). The final mechanism involves two accessory proteins; an usher ( $\beta$ -barrel) and a chaperone to facilitate pre-secretion folding.

#### 1.3.5 Type VI Secretion System

The most recently discovered secretion system in *P. aeruginosa* is the T6SS (Mougous *et al.*, 2006). This versatile system is capable of directly injecting a wide range of effector proteins into both prokaryotic and eukaryotic cells. As a result of the diverse effector activities, this system has been implicated in many biological scenarios, including virulence, microbial warfare, nutrient scavenging and bacterial communication and cooperation (Russell *et al.*, 2014). *P. aeruginosa* encodes three distinct T6SS islands, designated H1 - H3-T6SS, which differ in their repertoire of cargo proteins and their respective targets (Hood *et al.*, 2010). These systems can be expressed simultaneously and share structural and functional similarities with bacteriophage tails. Effector proteins, such as Tse1-3 and phospholipase D, are

typically encoded alongside an immunity protein for protection against self-injury (Wettstadt *et al.*, 2019). Their secretion can be triggered by membrane damage, for example through the action of antibiotics or attempts at bacterial conjugation. Control over the T6SS occurs at many levels, including post-translationally through the action of threonine kinase PpkA and its antagonistic phosphatase PppA (Chen *et al.*, 2015; Han *et al.*, 2019).

### **1.3.6 Alternative secretion mechanisms**

Aside from secretion via dedicated membrane bound machinery, proteins can also be exported through non-classical methods. Currently non-classical secretion pathways are not well understood, and the term is generally used when there is an absence of an encoded secretion signal and the export pathway is unknown (Bendtsen *et al.*, 2005; Wang *et al.*, 2013). For example, several conventionally cytoplasmic proteins have been shown to have ‘moonlighting’ activity outside of the cell, such as EF-Tu, but their secretion mechanism and its regulation remains unclear (Prezioso *et al.*, 2019).

One well documented alternative method of secretion is the release of outer membrane vesicles (OMVs). OMVs are small, spherical, membrane enclosed sacs that can contain an assortment of bacterial components; most notably virulence-associated proteins. They are formed by blebbing of the outer membrane, a process which is tightly regulated. OMV release is commonly a result of activation of the SOS response to stressful environmental conditions. Sensing of host derived factors, such as tear fluid in the eye, also triggers release (Metruccio *et al.*, 2016). Interestingly, they have been shown to be involved in the secretion of proteins which are known to be secreted via other mechanisms, such as elastase, phospholipase C and alkaline protease (MacDonald and Kuehn, 2013). As a result of their membranous composition, OMVs can more readily disseminate throughout host tissues and can fuse with host membranes to deliver effectors. They can also destructively trigger the host immune response (Ellis *et al.*, 2010). Finally, presumed altruistic suicide by triggering of cell lysis is used as a means of releasing the large protein complexes of R-type pyocins and generating OMVs (Turnbull *et al.*, 2016) .



## 1.4 Post-translational modifications

Regulation of the proteome is essential for bacterial viability, especially in volatile and competitive environments where nutrients are scarce. Control is tightly maintained at the transcriptional, translational and post-translational levels. At this latter stage, proteins can be modified by dedicated enzymes which chemically alter their amino acid sequence. This post-translational modification (PTM) frequently involves the covalent addition or removal of a functional group on a distinct amino acid residue, or at the extreme N- and C- terminal ends of a polypeptide (Figure 1.6). These modifications can range in size from small chemical moieties, such as methyl groups, to large and complex molecules such as lipids and polypeptides. Whilst these additions are commonly reversible, PTM can also involve irreversible chemical conversion of residues or peptide cleavage (André *et al.*, 2017). Through variable modification of their proteins, bacteria vastly increase the complexity of their proteome and are able to rapidly respond to internal or external stimuli. They can therefore remain dynamic and versatile entities despite only harbouring relatively small genomes.

Characterisation of post-translational modifications has principally been focused in eukaryotic research where it was generally regarded as a domain-centric phenomenon. There are currently more than 200 known eukaryotic PTMs, with certain modifications appearing much more common (or perhaps more readily discovered). Despite suggestions of a more limited capacity for PTMs in 'simpler' bacterial systems, research has illustrated that PTM in prokaryotes is likely to be just as important and as diverse. The number of recognised bacterial PTMs is steadily increasing and their adoption can have extensive effects on the functioning of proteins.

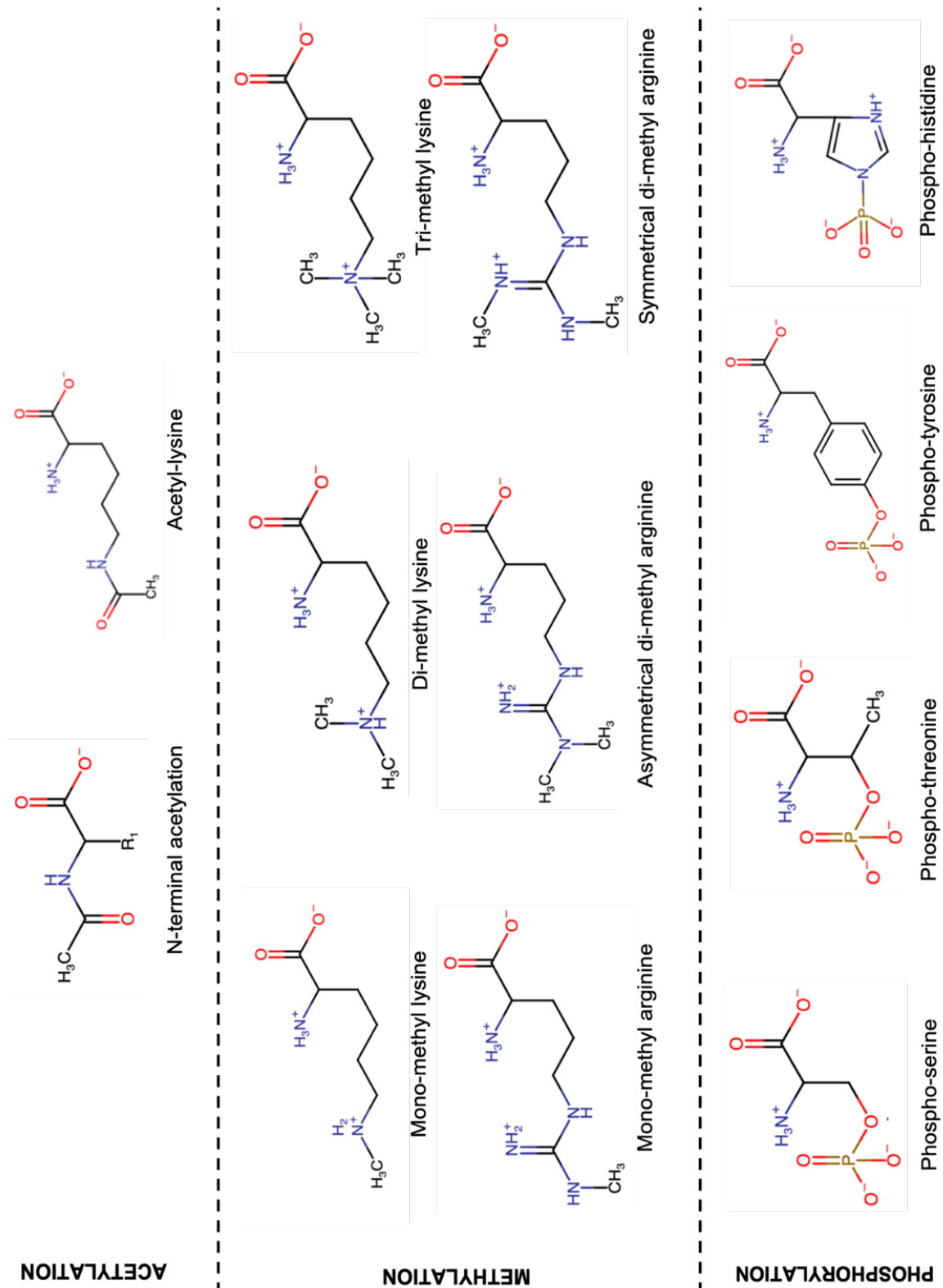
Modification of a protein alters its chemical properties and can therefore manipulate its conformation, stability and localisation within the cell, as well as interactions with other molecules. As such, PTMs can influence a variety of biological processes and are key to virulence. Nevertheless, the true purpose of many PTMs and how they interact is still unclear. Much still needs to be done to elucidate the fundamental role they play in an organism's success, which will also help to identify potential therapeutic drug targets. Research into the PTM status of secreted bacterial proteins

is hitherto relatively limited. This topic has been reviewed and published during the course of this research project (Forrest and Welch, 2020); see Appendix C.

#### 1.4.1 Methylation

Methylation is a small PTM that predominantly occurs on the side chains of arginine (R) and lysine (K) residues. It involves the addition of up to two (R) or three (K) methyl groups to the side chain terminal amine (Figure 1.6). Atypical protein methylation can also occur on asparagine (N), aspartic acid (D), histidine (H), glutamine (Q), glutamic acid (E) and cysteine (C) residues, however these are all vastly understudied (Zhang *et al.*, 2018). Attachment is catalysed by methyltransferase enzymes, with S-adenosyl-methionine (SAM) serving as the principal methyl group donor. Demethylase enzymes can also reverse this modification under certain circumstances. Each methyl group adds 14 Da in mass to the protein and does not affect the overall charge of positive or uncharged residues. However, O-methylation of the E/D side chain carboxylic acid neutralises its negative charge (Sprung *et al.*, 2008).

Despite the earliest observations of post-translational methylation occurring in *Salmonella enterica* flagellar proteins, the majority of research has migrated onto the eukaryotic platform (Ambler and Rees, 1959). It is best known as a method of transcriptional control through eukaryotic histone-mediated chromatin remodelling (Greer and Shi, 2012; Lanouette *et al.*, 2014). As such, the prevalence and implications of bacterial methylation are not well understood. Some large-scale studies have shown an enrichment of methylation targets within central metabolism and intracellular signal transduction pathways (Pang, Gasteiger and Wilkins, 2010; Zhang *et al.*, 2018). For example, in *P. aeruginosa*, the opposing actions of the methyltransferase CheR and the methylesterase CheB on the surface exposed methyl-accepting chemotaxis proteins (MCPs) modulate an intracellular signal transduction cascade. This in turn controls flagella-mediated chemotaxis in response to extracellular stimuli (Schmidt *et al.*, 2011; García-Fontana *et al.*, 2013)



**Figure 1.6 The structure of small post-translational modifications (PTMs)**

The structure of small post-translational modifications (PTMs) on their substrate amino acids, including acetylation (N $\alpha$  and N $\epsilon$ ), methylation (K/R) and phosphorylation (S/T/Y/H).

Studies of the spirochete *Leptospira interrogans* have identified many different methylated proteins. A large proportion are E or Q methylated and are involved in motility and chemotaxis. In contrast, targets of K or R methylation appear to be devoid of a distinct functional class (Cao *et al.*, 2010). Interestingly, this methylation seems to occur somewhat in response to host signals (Nally *et al.*, 2017). The outer membrane protein LipL32 has been repeatedly shown to have a variable methylation status, particularly at epitope locations recognised by B-cells. The presence of these different modification isoforms is likely to mediate and dampen host immune recognition (Temmerman *et al.*, 2004; Eshghi *et al.*, 2012; Witchell *et al.*, 2014). This is further supported by various studies in *Rickettsia* that show methylation of OmpB contributes to phase variation by altering its surface antigenicity (Chao *et al.*, 2008; Abeykoon *et al.*, 2016). The degree of OmpB methylation also appears to correlate with strain pathogenicity (Abeykoon *et al.*, 2014). In general, outer membrane proteins appear to be extensively methylated in a range of different bacteria, aiding immune evasion strategies and enabling successful infection (Yang *et al.*, 2017).

An intriguing role of methylation in protein stability and protection against proteolysis has been suggested. A study of the simple eukaryote *Saccharomyces cerevisiae* found that methylated proteins tended to have a longer half-life than non-modified proteins, and that just under half of methylatable lysines were potential ubiquitination targets. This suggests that methylation may protect against ubiquitin-mediated proteolytic degradation (Pang *et al.*, 2010). A similar protective role has also been found for methylation in a few *Mycobacterium tuberculosis* adhesins, which are much more stable in their methylated forms (Pethe *et al.*, 2002).

Furthermore, it has been postulated that methylation of *S. enterica* flagellin may protect the bacterial appendage within the external environment. Whilst this is experimentally unverified, flagellin methylation by FliB has proven to be definitively important for adhesion in both *S. enterica* and other members of the *Enterobacteriaceae* family. It is thought that this is partly due to the resulting increase in flagellar surface hydrophobicity; enhanced adhesion fundamentally enhances virulence (Bogomolnaya *et al.*, 2014; Horstmann *et al.*, 2020). Likewise, T4P subunits are methylated in *P. aeruginosa*, *Neisseria meningitidis* and *Synechocystis* sp, although the true function of this modification remains unconfirmed (Kim *et al.*, 2011;

Aly *et al.*, 2013). In addition, methyltransferase activity has been shown to be a necessary exploit for *Escherichia coli* viability in stationary phase and during heat stress (Li and Clarke, 1992). *Yersinia pseudotuberculosis* also depends on the capabilities of its methyltransferase VagH for virulence, particularly with regards to T3SS functioning (Garbom *et al.*, 2007).

Focusing specifically on *P. aeruginosa* proteins, the secreted virulence factors CbpD and LasB have recently been identified as possessing multiple methylated lysines. Consistent with many other bacterial proteins, the utilisation of this modification is remarkably variable. Discrete lysines can be mono-, di- or tri-methylated, or even unmodified, with several different isoforms present in the same sample (Gaviard *et al.*, 2019). In contrast, the trimethylation of *P. aeruginosa* elongation factor EF-Tu is thought to be highly stable. This archetypal cytoplasmic protein actually moonlights on the cell surface where it functions as an environmental sensor and adhesin (N'Diaye *et al.*, 2019). Exclusive trimethylation at K5 does not impact its canonical functions in translation but instead enhances epithelial attachment by mimicking a phosphorylcholine molecule (Barbier *et al.*, 2013; Prezioso *et al.*, 2019). Intriguingly, the responsible methyltransferase EftM is thermoregulated, showing greater activity at 25°C compared to 37°C. This suggests that methylation may be important in the early stages of bacterial infiltration, or for survival in the external environment (Owings *et al.*, 2016). EftM has homologues in several species, and their functions have proven indispensable for many pathogens (Widjaja *et al.*, 2017).

Overall, investigations into bacterial methylation are insufficient, the purpose of the modification is largely ambiguous and there is a distinct lack of methylated arginine residues identified despite the prevalence of this modification in eukaryotes. It is unclear whether this is as a result of methodological limitations or that it is a less significant mechanism of regulation in prokaryotes.

#### **1.4.2 Acylation**

Acetylation is another small modification that occurs either on the  $\epsilon$ -amine of the lysine side chain, or at the protein N-terminus (Figure 1.6). Lysine acetylation is reversible through the action of acetyltransferases and antagonistic deacetylases. This reaction

can also occur non-enzymatically, usually at an alkaline pH. Incorporation of an acetyl group requires acetyl-CoA or acetyl-phosphate as a donor. This modification adds 42 Da in mass to the protein whilst neutralising the positive charge of the lysine residue.

In bacteria, acetylation is a prominent PTM. Approximately 90% of the central metabolic proteins encoded by *S. enterica* are lysine acetylated (Wang *et al.*, 2010). It is thought that this is a regulatory mechanism which mediates cellular adaptation to changing environments. This comes from observations of varying levels of modification under different carbon sources and at different stages of growth (Kosono *et al.*, 2015; Gaviard, Broutin, *et al.*, 2018). There is a plethora of studies uncovering the acetylome of many bacterial species under numerous different growth conditions (Christensen *et al.*, 2019). Most of these investigations have highlighted several metabolic pathways, as well as transcription and translation, as the main targets of acetylation (Thao and Escalante-Semerena, 2011; Liao *et al.*, 2014; Pan *et al.*, 2015; Yu *et al.*, 2016). Nonetheless, analysis of the *P. aeruginosa* acetylome is currently fairly limited, particularly with regards to the secretome (Ouidir *et al.*, 2015).

Whilst dissection of the breadth of acetylation in an organism is insightful, a dearth of information regarding the impact these modifications have on protein function is a frequent drawback. To examine the function acetylation plays, the majority of investigators use genetic manipulation to knock-out the responsible modifying enzymes. This indirect route can show the global phenotypic effects of this PTM but has many caveats and leads to considerable inconsistencies when comparing organisms. For this reason, there is minimal veritable data on the consequences of acetylating specific proteins, and proposed mechanisms are usually speculative. For example, several studies have looked at the role acetylation plays in different stress responses using this method, including temperature, salt, acid and ROS (Ma and Wood, 2011; Ren *et al.*, 2015). A clear conclusion cannot be drawn from comparing these studies that pertains to different bacterial species. It can only be generally concluded that this PTM plays a role in the stress response, but with different positive and negative outcomes for each organism (Christensen *et al.*, 2019).

This complexity is further evident in the links between acetylation and protein stability. In *E. coli*, acetylation of RNase R promotes its degradation, however this PTM

enhances stability in the transcriptional regulator HilD of *S. typhimurium* (Liang *et al.*, 2011; Sang *et al.*, 2017). It is likely that such inconsistencies lie in the variation between bacterial proteomes and the interconnected relationships between the different PTMs utilised. This further supports the notion that the function of PTM is strain, and even protein, specific.

In addition to methylation, several proteins involved in controlling *P. aeruginosa* flagellar rotation and chemotaxis are acetylated (Barak and Eisenbach, 2001). Flagellin monomers and proteins which make up T4P are similarly modified, indicating a role of this PTM in motility (Ouidir *et al.*, 2015). Acetylation of multiple components of certain *P. aeruginosa* secretion systems has also been found. As with methylation, modification of these proteins is inconsistent, leading to heterogeneity in the protein subpopulation. Typically, the purpose of these PTMs remains elusive. It has been suggested that acetylation can inhibit the activity of enzymes through disruption of molecular interactions within the active site (Zhang *et al.*, 2009; Crosby *et al.*, 2012; Kim *et al.*, 2013). This is true in *M. tuberculosis* where the active site of the secreted phosphatase PtpB is inaccessible following acetylation, preventing its dephosphorylation of host proteins (Singhal *et al.*, 2015). In addition, acetylation of the *M. tuberculosis* heat shock protein X (HspX) blocks its immunogenicity and is therefore associated with infection latency (Liu *et al.*, 2014).

Lysine residues are also subject to other forms of acylation, including propionylation (+56 Da), crotonylation (+68 Da), butyrylation (+70 Da), malonylation (+86 Da) and succinylation (+100 Da). These PTMs also neutralise the positive charge of the lysine, except for malonylation and succinylation which introduce a negative charge of -1 (Zhang *et al.*, 2011; Alleyn *et al.*, 2018). As these other acylations were only recently discovered, there is not an abundance of data regarding their presence in prokaryotes. However, early observations suggest similar roles to acetylation in controlling central metabolism and protein synthesis (Yang *et al.*, 2015; Qian *et al.*, 2016). Intriguingly, it seems that these other types of acylation can be more abundant than the established acetylations (Xu *et al.*, 2018).

Succinylation has been shown to be prevalent in a range of bacterial species. There appears to be significant overlap between acetylation and succinylation, as the same

lysine residues have been found to harbour both modifications (Weinert *et al.*, 2013; Kosono *et al.*, 2015; Gaviard *et al.*, 2019). The overlap between these modifications is biochemically confusing considering the differences in their mass, structural bulkiness and charge modifying properties, which would imply significantly different functions. Some proteins have been identified as having ten or more succinylation sites, highlighting how important this modification is likely to be (Pan *et al.*, 2015; Xie *et al.*, 2015). A very recent study has shown the presence of a succinyl-lysine in the pro-peptide of LasB, and an increase in the number of acylations upon secretion. This may suggest a potential role in protein maturation or stability (Gaviard *et al.*, 2019). Further exploration of the acyl modifications is necessary to tease out the cross-talk between these PTMs and how relevant they are to prokaryotic function.

### 1.4.3 Phosphorylation

Phosphorylation is a relatively promiscuous PTM that occurs predominantly on the side chains of serine (S), threonine (T), tyrosine (Y) and histidine (H). Phosphorylation of arginine (R), lysine (K), aspartic acid (D), glutamic acid (E) and cysteine (C) have, however, also been identified. Whilst S/T/Y phosphorylation is most common in eukaryotes, there is some debate about the most frequent phospho-residue in prokaryotes (Fortuin *et al.*, 2015). Addition of a phosphate group, predominantly from ATP as a phosphate donor, is catalysed by kinases and is removed by phosphatases (André *et al.*, 2017). This PTM adds 80 Da to a protein and introduces an overall negative charge. Phosphorylation is typically rapid, reversible and transient, and therefore has been principally implicated in bacterial signal transduction systems (Tiwari *et al.*, 2017). As such, the most well-defined phosphoproteins are intracellular and membrane localised proteins (Lai *et al.*, 2017).

As with most PTMs, research has focused on constructing the global phosphoproteome of an organism, rather than its specific functions in distinct proteins. Where the roles of phosphorylation have been elucidated, it can be reasoned that this PTM has pleiotropic impacts on modified proteins. A general role as a regulatory mechanism has been found in several organisms. Arginine phosphorylation of multiple stress response regulators by kinase McsB mediates *B. subtilis* stress resistance. This includes proteins such as the ClpCP protease and GroEL chaperonin



which are involved in protein quality control systems (Schmidt *et al.*, 2014; Fuhrmann *et al.*, 2016). Additionally, phosphorylation functions as a degradation tag for Clp protease targeting (Trentini *et al.*, 2016). The McsB homologue in *S. aureus* is critical to virulence, suggesting that phosphorylation is essential not only for a coordinated and efficient response to stressors, but also successful infection. This is further supported by findings that kinase activity supports *S. aureus* virulence traits, including proteolysis, haemolysis and biofilm formation (Wozniak *et al.*, 2012).

Phosphorylation is involved in many aspects of protein secretion. Several components of the *M. tuberculosis* ESX-1 T7SS machinery are phosphorylated, as are several virulence factors (Fortuin *et al.*, 2015). This includes EsxB and PtpA which are secreted from the cell and trigger an immunogenic response in the host. Furthermore, tyrosine phosphorylation of PtpA is essential for its phosphatase activity in inhibiting phagosome maturation (Zhou *et al.*, 2015). An extracellular kinase of *Acinetobacter baumannii* is also phosphorylated, indicating the potential importance of this modification for secreted enzymes (Soares *et al.*, 2014). In *Helicobacter pylori*, phosphorylation of the secreted toxins CagA and VacA is used to control their activity following injection into the gastric epithelium (Ge *et al.*, 2011).

Pseudomonads use phosphorylation to control numerous intracellular processes. Significantly, the number of phosphotyrosines seems to correlate with *P. aeruginosa* strain virulence (Ge and Shan, 2011). In terms of protein secretion, there have also been some interesting developments. Studies highlighting phosphorylation of FlhC flagellin subunits have shown that this modification does not impact motility as hypothesised, but instead influences Type 2 secretion. This modification is growth phase dependent and leads to reduced levels of and activity of Type 2 secreted proteins (Kelly-Wintenberg *et al.*, 1993; Suriyanarayanan *et al.*, 2016). In addition, flagellin phosphorylation positively moderates biofilm formation and dispersal. This is consistent with observations in *S. aureus* whereby phosphorylation controls the biofilm phenotype, except through a flagellar-independent mechanism (Wozniak *et al.*, 2012). Furthermore, *P. aeruginosa* TbpA phosphatase activity reduces biofilm formation by negatively regulating c-di-GMP in response to quorum sensing molecules, which in turn diminishes exopolysaccharide production (Ravichandran *et al.*, 2009; Ueda and Wood, 2009). This shows a glimpse into the complex regulatory

network controlling the sessile biofilm phenotype, in which phosphorylation is a central player.

Direct phosphorylation of *P. aeruginosa* secreted effectors has also been demonstrated. One study identified 28 phospho-exoproteins of strain PA14, with varying degrees of modification for each (Ouidir *et al.*, 2014). Further to this, it has been shown that intracellular elastase is non-phosphorylated, whereas the secreted form carries 19 phosphorylations (Gaviard *et al.*, 2019). Exactly how and why this modification occurs has not yet been elucidated, however it has been postulated that this PTM may be an additional secretion signal. Looking at the T6SS, export of the virulence factor Hcp1 is dependent on threonine phosphorylation which directs system assembly (Mougous *et al.*, 2007).

#### **1.4.4 Experimental constraints**

The last decade has galvanised a surge in research attempting to understand prokaryotic post-translational modification, largely thanks to significant improvements in the proteomic techniques used to identify them. Such research is now reaching its adolescence and despite the invaluable leaps in knowledge acquired, there are still many teething problems that must be overcome. Whilst early studies have historically addressed individual proteins or singular PTMs, the focus has now tended to move to a more “omic” approach through global modification studies in a wide range of organisms (Cao *et al.*, 2010; Kosono *et al.*, 2015; Brown *et al.*, 2017). Unfortunately, the vast majority of these studies fail to consider the potential importance of secreted proteins, and this protein fraction is commonly removed prior to analysis. Bacterial secretome research is therefore relatively limited. This further constrains the capacity for understanding these modifications in the bacterial context, as well as understanding how they may contribute to the pathogenicity of an organism through its modified extracellular effectors.

A more complex issue to resolve is in comprehending the extent to which growth conditions affect PTM. Several studies have shown that variables such as carbon source, temperature and growth phase influence how modified a protein is (Kosono *et al.*, 2015; Prezioso *et al.*, 2019). Attempting to examine all possible combinations

of these factors and how the results apply to the real-world setting is a huge undertaking in terms of cost, resources and manpower. Hopefully, technological and bioinformatic advances will help such experiments to become less demanding. In addition, it is critical to understand the consistencies across several bacterial strains, or even species, in order to remove strain-bias and provide a platform for the development of targeted therapies using PTM-directed drugs.

The gold standard method for detecting PTMs is tandem mass spectrometry (MS/MS). This technique allows strong resolution of proteins and their modifications, typically with high sensitivity and accuracy of detection (Witze *et al.*, 2007). However, unequivocal limitations in probing the proteome in this way still persist. There are considerable difficulties in detecting phosphoproteins due to the labile nature of the phosphate bond. Preferential cleavage of this bond over the peptide bond during fragmentation leads to loss of the phosphate group, particularly when collision-induced dissociation is employed. This issue is exaggerated when detecting phosphohistidine as the P-N bond is even more unstable under the acidic conditions used during MS (Cain *et al.*, 2014). The lack of complete fragmentation of a peptide can also make discerning the location of a modification within the peptide difficult if multiple modifiable residues are present. In addition, it is common for the overall protein sequence coverage to be relatively low (<50%), meaning that there is likely to be a significant proportion of the protein that has not been analysed and therefore several modifications overlooked.

To discriminate between certain PTMs, a mass spectrometer with a high resolving power is required. For example, the molecular mass changes from trimethylation (42.04695 Da) and acetylation (42.01056 Da) are very close and therefore the mass accuracy must be below 5 ppm in order to determine which modification is present (Gaviard, Jouenne, *et al.*, 2018). Further to this, there are software limitations in how MS/MS data can be reliably analysed. Each modification that is added into the search parameters increases the bioinformatic search space and the number of iterative searches the probabilistic software must make. This greatly increases the time taken for analysis and rate of false PTM discoveries (Ahrné *et al.*, 2010). There are also typically limits on the mass shift window available for searches, and therefore larger modifications are frequently excluded (Brown *et al.*, 2017).

High levels of quantity and purity of protein are generally necessary for sufficient MS sequence coverage and resolution (Silva *et al.*, 2013). This can be challenging when proteins are expressed at low concentrations and is particularly problematic in identifying modified residues. PTMs typically have low stoichiometries, whereby only a small proportion of proteins from a subpopulation are modified. They can therefore fall below the threshold for MS identification or be drowned out by non-modified proteins. To mitigate this issue, antibody enrichment of modified peptides can be utilised prior to MS analysis. This itself comes with obstacles due to commonly low specificities and binding efficiencies of the antibodies which are used (Zhao and Jensen, 2009). There are also limitations on the availability of commercial antibodies, for this purpose, especially if the modification is atypical.

## **1.5 Project Aims**

Previous research in the laboratory utilised 2D gel technology for separating the *P. aeruginosa* secretome. It was noted that a large proportion of the proteins are present as lateral “charge trains” with single protein isoforms presenting with variable isoelectric points but consistent molecular weights, however this was never further investigated.

My aims were therefore to firstly characterise how changes in culture growth conditions may affect secreted protein charge isoform presentation. Secondly, to understand which post-translational modifications may be present on *P. aeruginosa* secreted proteins with respect to small chemical modifications, which had been hitherto unexplored. Finally, to perform preliminary investigations into the role that identified modifications may have on secreted proteins.

# Chapter 2

## 2 Materials and Methods

### 2.1 Bacterial strains

The bacterial strains used in this work are listed in Table 2.1. All strains were stored in 25% (v/v) glycerol solution at -80°C. When required, strains were streaked out on LB agar plates (Table 2.2), incubated at 37°C for approximately 16 hours and subsequently stored at 4°C for up to two weeks. Clinical isolates were selected to allow collaborative work with a visiting researcher.

**Table 2.1** List of bacterial strains

Strain	Description	Reference
<i>Pseudomonas aeruginosa</i>		
PAO1	Wild type	B. Iglewski, University of Rochester, USA
PAK	Phage sensitive strain	(Takeya and Amako, 1966)
B5	Clinical isolate from a medical implant in Pakistan	Gifted by S. Abbas
NB1	Clinical from a medical implant in Pakistan	Gifted by S. Abbas
PW7302	<i>lasB</i> ::ISphoA/hah derivative of PAO1, Tet <sup>R</sup>	(Jacobs <i>et al.</i> , 2003)
PW2131	<i>PA0622</i> ::ISlacZ/hah derivative of PAO1, Tet <sup>R</sup>	(Jacobs <i>et al.</i> , 2003)
LasB-EV	PW7302 containing pUCP20 empty vector, Tet <sup>R</sup> Cb <sup>R</sup>	This study
LasB-WT	PW7302 containing pLasB, Tet <sup>R</sup> Cb <sup>R</sup>	This study
LasB-K211A	PW7302 containing pLasB-211, Tet <sup>R</sup> Cb <sup>R</sup>	This study
LasB-R226A	PW7302 containing pLasB-226, Tet <sup>R</sup> Cb <sup>R</sup>	This study
LasB-K378A	PW7302 containing pLasB-378, Tet <sup>R</sup> Cb <sup>R</sup>	This study

Table 2.1 continued		
PA0622-EV	PW2131 containing pUCP20 empty vector, Tet <sup>R</sup> Cb <sup>R</sup>	This study
PA0622-WT	PW2131 containing p0622, Tet <sup>R</sup> Cb <sup>R</sup>	This study
PA0622-K171A	PW2131 containing p0622-171, Tet <sup>R</sup> Cb <sup>R</sup>	This study
PA0622-N172A	PW2131 containing p0622-172, Tet <sup>R</sup> Cb <sup>R</sup>	This study
PA0622-K176A	PW2131 containing p0622-176, Tet <sup>R</sup> Cb <sup>R</sup>	This study
PA0622-K312A	PW2131 containing p0622-312, Tet <sup>R</sup> Cb <sup>R</sup>	This study
Other		
<i>Escherichia coli</i> DH5 $\alpha$	(F-) <i>supE44</i> $\Delta$ <i>lacU169</i> ( $\phi$ 80 <i>lacZ</i> $\Delta$ <i>M15</i> ) $\Delta$ <i>argF</i> <i>hsdR17</i> <i>recA1</i> <i>endA1</i> <i>thi-1</i> <i>relA1</i>	(Bryant, 1988)
<i>Escherichia coli</i> OP50	Uracil auxotroph for laboratory maintenance of <i>C. elegans</i>	(Brenner, 1974)
<i>Staphylococcus aureus</i> Newman	Laboratory strain	(Duthie and Lorenz, 1952)

## 2.2 Growth conditions

The growth media used in this research are listed in Table 2.2. Unless otherwise stated, bacteria were routinely cultured in liquid LB media and on solid LBA plates at 37°C aerobically. All media constituents and glassware were autoclaved at 115°C for 15 minutes or filtered-sterilised using 0.22  $\mu$ m filters (Millipore) prior to use.

### 2.2.1 Overnight culture

A single bacterial colony from a streak plate was used to inoculate 10 mL of growth medium in a universal tube. The culture was incubated on a rotating wheel at 37°C overnight. When necessary, appropriate antibiotics (Table 2.3) were added to the medium prior to incubation.

**Table 2.2** Growth media

<b>Growth Media</b>	<b>Ingredients per litre</b>
Luria Burtani Broth (LB)	10 g Tryptone 5 g NaCl 5 g Yeast extract
Luria Burtani Agar (LBA)	As above with 1.5% (w/v) agar
Alanine, glycerol, salts and yeast extract (AGSY) pH 7.0	5 g L-alanine 5 g NaCl 3 g K <sub>2</sub> HPO <sub>4</sub> 3 g Yeast extract
Added post-autoclaving	10 mL 50% glycerol 10 mL 1M MgSO <sub>4</sub> 1 mL 5 mM FeCl <sub>2</sub> 1 mL 100 mM CaCl <sub>2</sub> 1 mL 7.5 mM ZnSO <sub>4</sub>
MOPS Minimal Media	1X MOPS media (LaBauve and Wargo, 2012) 20 mM Glucose or Acetate
Skim Milk Agar (SMA)	50 g Tryptic soy broth 15 g Agar
Autoclaved separately on skim milk cycle	200 ml 5% (w/v) skimmed milk
Swarming agar	1.07 g NH <sub>4</sub> Cl 1.7 g Na <sub>2</sub> HPO <sub>4</sub> 3 g KH <sub>2</sub> PO <sub>4</sub> 0.5 g NaCl 1.98 g Glucose 5 g Casein hydrolysate 5 g Agar
Added post-sterilisation	1 ml 1M MgSO <sub>4</sub> 1 ml 1M CaCl <sub>2</sub>

<b>Table 2.2 continued</b>	
Swimming agar	10 g Tryptone
	5 g NaCl
	3 g Agar

### **2.2.2 Planktonic culture**

An aliquot of overnight culture was used to inoculate growth media in suitable conical flasks to an initial optical density (OD<sub>600</sub>) of 0.05, unless otherwise stated. Antibiotics were added to the media when necessary (Table 2.3). Sub-cultures were incubated at 37°C either in a shaking water bath or an orbital shaker at 220 rpm to enable sufficient aeration until the required final OD<sub>600</sub> was reached. Aliquots of the culture were taken every hour to monitor the growth (section 2.2.4).

### **2.2.3 Growth on solid media**

Bacterial colonies were grown in 10 cm round or 12 cm square Petri-dishes containing 25 mL or 60 mL of solid growth media, respectively. A sterile inoculation loop or glass spreader was used to distribute cells across the agar to isolate single colonies when necessary.

### **2.2.4 Measuring planktonic growth**

The growth and cell density of planktonic cultures was assessed by measuring the OD<sub>600</sub> of a culture aliquot using a spectrophotometer (Eppendorf BioPhotometer). For growth curves, 100 µl of a fresh bacterial sub-culture at OD<sub>600</sub> 0.05 was added in triplicate to a sterile 96-well microtiter plate and covered with a gas-permeable plate seal (Breathe-Easy®). The plate was incubated in a FLUOstar Omega microplate reader (37°C, 250 rpm shaking) and the OD<sub>600</sub> was measured every 10 mins for 16 hours.



## 2.2.5 Antibiotics

The antibiotics used in this study are listed in Table 2.3. All antibiotic solutions were prepared in 50% (v/v) ethanol, filter sterilised through 0.22 µm membrane filters and stored at -20°C in aliquots.

**Table 2.3** Antibiotic solutions

Stock (mg/ml)		Final concentration (µg/ml)	
		<i>P. aeruginosa</i>	<i>E. coli</i>
Carbenicillin	50	250	50
Tetracycline	10	50	-
Ciprofloxacin	0.5	0.05	-

## 2.3 Proteomic analysis

### 2.3.1 Secretome extraction for 2D gels

PAO1 was grown in AGSY medium, as detailed in section 2.2.2, until the cultures reached the late exponential/early stationary phase of growth (~ 6.5 hours), unless otherwise stated. Cell-free culture supernatants were harvested via centrifugation at 15,000 x *g* for 20 minutes at 4°C (Beckman Coulter Avanti J-26 XPI), and then membrane filtration of the supernatant to remove residual bacteria (Millipore Stericup, 0.22 µm). The proteins were precipitated by adding 12% (w/v) trichloroacetic acid (TCA) and incubating at 4°C overnight. Precipitated proteins were pelleted by centrifugation at 21,000 x *g* for 40 minutes at 4°C and the pellet was subsequently washed twice with ice-cold 80% acetone to remove any traces of TCA. Following a final sedimentation, the pellet was left to air dry at room temperature and then resuspended in AUT buffer (Table 2.4), aliquoted and stored at -80°C.

### **2.3.2 Secretome column concentration**

Cultures were grown to the early stationary phase of growth and cell-free supernatants were harvested as described in Section 2.3.1. Following filtration, the supernatants were then concentrated 10-fold in a Viviaspin® centrifugal concentrator column (10 kDa MWCO, Sartorius) at 3,200 x g at 4°C. To remove excess salt for 2D analysis, samples were dialysed in dialysis tubing suspended in excess dialysis buffer overnight at 4°C. If necessary, samples were further purified with a 2D Clean-Up Kit (GE Healthcare).

### **2.3.3 Quantification of proteins**

The protein concentration in secretome preparations was determined using the Bio-Rad Protein Assay according to the manufacturer's instructions. Bovine serum albumin (BSA) was used as a protein standard (concentrations ranging from 0 – 1.5 mg/ml) to generate a standard curve. BSA standards were prepared in the sample buffer to allow accurate sample quantitation from standard linear regression analysis. Measurements were taken spectrophotometrically (Eppendorf BioPhotometer or FLUOstar Omega) at 595 nm following incubation of the Bradford reagent with the sample for 10 minutes at room temperature.

### **2.3.4 1D SDS-PAGE**

Protein samples were prepared with 4X loading dye and heated to 95°C for 10 minutes. The samples were then electrophoretically separated in 12% SDS polyacrylamide gels submerged in running buffer (Table 2.4) at 120V for approximately 2 hours. Precision Plus Protein Standard (Bio-Rad) was also loaded as a molecular weight marker. Gels were subsequently stained with Coomassie Brilliant Blue or used for western blot analysis.

#### **2.3.4.1 Coomassie blue staining**

Following SDS-PAGE, gels were incubated in Coomassie Brilliant Blue stain overnight at room temperature with gentle rotation. The stain was then replaced with

Destain I solution for 1 hour, followed by Destain II solution until protein bands could be clearly visualised. All solutions are detailed in Table 2.4.

**Table 2.4** SDS-PAGE solutions and buffers

<b>Solution</b>	<b>Components</b>
AUT buffer	7 M Urea 2 M Thiourea 2% (w/v) amido sulphobetaine 14 (ASB-14)
Dialysis buffer (pH 7.5)	100 mM NaCl 50 mM Tris-HCl 5% (v/v) glycerol
4X Loading buffer	50 mM Tris-HCl 2% (w/v) SDS 10% (v/v) Glycerol 0.1% (v/v) Bromophenol blue 10 mM DTT
6% Stacking gel	1.98 ml 30% Bis-Acrylamide 3.78 ml 0.5M Tris-HCl (pH 6.8) 150 µl 10% SDS 9 ml dH <sub>2</sub> O 15 µL tetramethylethylenediamine (TEMED) 75 µL 10% (w/v) ammonium persulphate
12% Resolving gel	6 ml 30% Bis-acrylamide 3.75 ml 1.5M Tris-HCl (pH 8.8) 150 µl 10% SDS 5.03 ml dH <sub>2</sub> O 7.5 µL tetramethylethylenediamine (TEMED) 75 µL 10% (w/v) ammonium persulphate
SDS running buffer (pH8.3)	25 mM Tris 0.19 M Glycine 0.1% (w/v) SDS
Coomassie stain	1 g/l Coomassie Brilliant Blue G 50% (v/v) methanol 10% (v/v) acetic acid
Destain I	50% (v/v) methanol 7% (v/v) acetic acid
Destain II	10% (v/v) methanol 7% (v/v) acetic acid)

### **2.3.5 Western blotting**

Proteins were transferred from the SDS-PAGE gel to a polyvinylidene difluoride (PVDF) membrane (0.45  $\mu\text{m}$ , Millipore) following membrane activation with methanol. Transfer was achieved in 10 minutes using the Trans-Blot Turbo Transfer System (Bio-Rad) with the accompanying transfer buffer and filter stacks. Membranes were then incubated with blocking buffer (5% (w/v) skimmed milk in PBS) for a minimum of 2 hours and subsequently incubated overnight at 4°C with anti-mono methyl arginine antibody (Cell Signalling Technology). All subsequent steps were performed at room temperature. The membranes were washed four times in wash buffer (PBS with 0.1% (v/v) Tween-20) for 5 minutes with vigorous shaking before the addition of HRP-conjugated goat anti-rabbit secondary antibodies diluted (1:15,000) in blocking buffer with 0.1% (v/v) Tween-20 and incubating for 1 hour. The washing steps were repeated and Clarity Western ECL Substrate (Bio-Rad) was added to the membranes for 5 minutes. Protein bands were then visualised with a ChemiDoc imaging system using optimal auto-exposure parameters.

#### **2.3.5.1 In-gel Western**

For immunoblotting of a 2D gel, an In-Gel Western protocol was utilised as an alternative to traditional membrane blotting due to the size of the gel. After electrophoresis, the gel was incubated in 50% isopropanol and 5% acetic acid for 15 minutes followed by a wash step with  $\text{dH}_2\text{O}$ . The gel was then incubated with anti-mono methyl arginine antibody (Cell Signalling Technology) overnight at 4°C (1:1000). Subsequent washing and incubation steps were conducted as detailed above. The secondary antibody used was Goat anti-rabbit IRDye 800CW and the gel was scanned using an Odyssey CLx imaging system (Li-Cor).

### **2.3.6 2D gel electrophoresis**

Two-dimensional PAGE (2-DE) was used as a method of further separating secreted proteins whereby proteins were initially separated by their isoelectric point in the first dimension, followed by their molecular weight in the second. Protein samples (~200  $\mu\text{g}$ ) were mixed with 2X sample buffer and DeStreak Rehydration Solution (GE

Healthcare) and were added to 13 cm pH 3-10 non-linear Immobiline IPG strips. The strips were sealed with paraffin oil and placed on to an IPGphor Isoelectric Focusing Unit where proteins were focused in the first dimension under the parameters detailed in Table 2.5. The strips were then incubated first with reducing equilibration buffer and then alkylating equilibration buffer for 15 minutes each. Following a wash in SDS running buffer, the strip was loaded onto a prepared 12% SDS-PAGE gel (Table 2.4) and secured in place by adding sealing agarose solution. The second-dimension separation was then run in SDS running buffer in a Hoefer SE 600 Ruby system under standard running parameters (Table 2.5). The gel was stained with silver stain (Section 2.3.6.1) and scanned using a flatbed scanner. All solutions are detailed in Table 2.6.

**Table 2.5** 2-DE parameters

	Step	Parameter	Time (hours)
1 <sup>st</sup> Dimension	1	200 vH at 20 V	10
	2	500 vH at 500 V	1
	3	1000 vH at 1000 V	1
	4	4800 vH at 8000 V	6
2 <sup>nd</sup> Dimension	1	20 mA / gel	0.25
	2	50 mA / gel	3.5

### 2.3.6.1 Silver staining

Following the second-dimension separation, gels were fixed in fixing solution overnight with gentle agitation. The next day gels were incubated at room temperature, first with sensitisation solution for 1 minute, then ice-cold silver stain for 20 minutes and finally development solution until protein spots were visible; gels were washed three times in dH<sub>2</sub>O after each step. Blocking solution was then added and the gels were incubated for 10 minutes before gel image scanning with a flatbed scanner (Epson Expression 10000 XL). Gel spots were quantified and analysed using ImageJ and SameSpots (Total Labs) software, respectively.

**Table 2.6** 2-DE solutions and buffers

<b>Solution</b>	<b>Components</b>
2X sample buffer	AUT buffer 2% (w/v) DTT 2% (v/v) IPG buffer
Equilibration stock	6 M Urea 75 mM Tris (pH 8.8) 30% (v/v) Glycerol 2% (w/v) SDS
Equilibration buffer - reducing	Equilibration stock 1% (w/v) DTT
Equilibration buffer - alkylation	Equilibration stock 2.5% Iodoacetamide
Sealing agarose	0.5% agarose 0.002% Bromophenol blue SDS running buffer
Fixing solution	45% (v/v) methanol 10% (v/v) acetic acid
Sensitisation solution	0.2 g/L Na <sub>2</sub> S <sub>2</sub> O <sub>3</sub>
Silver stain	0.4 g/L AgNO <sub>3</sub> 0.2% (w/v) Formaldehyde
Developing solution	7.5 g Na <sub>2</sub> CO <sub>3</sub> 0.5 mg Na <sub>2</sub> S <sub>2</sub> O <sub>3</sub> 0.05% (w/v) Formaldehyde
Blocking solution	5% (v/v) acetic acid

### 2.3.6.2 2D-DiGE

Two-dimensional Difference Gel Electrophoresis (2D-DiGE) is an extension of 2D-PAGE whereby protein samples are labelled with different fluorescent CyDyes such that multiple protein samples can be separated and visualised on the same gel. This allows direct comparison of samples without any gel-to-gel variation. Protein samples (50 µg) were labelled with 250 pM of Cy2, Cy3 or Cy5 minimal dyes for 30 minutes

in the dark. The labelling reaction was then quenched by the addition of 10 mM lysine and a further incubation for 10 minutes. The samples were then pooled and subjected to the standard 2D-PAGE protocol (Section 2.3.6). Instead of silver staining, the gels were scanned (pixel size 100  $\mu\text{m}$ ) using a Typhoon 9400 scanner at excitation wavelengths 488 nm (Cy2), 532 nm (Cy3) and 633 nm (Cy5). ImageQuant and SameSpots were used for image overlay and analysis.

### **2.3.7 Tandem Mass Spectrometry (LC-MS/MS)**

Gel spots were picked and sent for LC-MS/MS analysis at the Cambridge Centre for Proteomics (CCP). Spots were cut into 1mm<sup>2</sup> pieces, destained, reduced (DTT) and alkylated (iodoacetamide) and subjected to enzymatic digestion with sequencing grade trypsin (Promega, Madison, WI, USA) overnight at 37°C. After digestion, the supernatant was pipetted into a sample vial and loaded onto an autosampler for automated LC-MS/MS analysis.

All LC-MS/MS experiments were performed using a Dionex Ultimate 3000 RSLC nanoUPLC (Thermo Fisher Scientific Inc, Waltham, MA, USA) system coupled to a Q Exactive Orbitrap mass spectrometer (Thermo Fisher Scientific Inc, Waltham, MA, USA) or a nanoAcquity UPLC (Waters Corp., Milford, MA) system coupled to a LTQ Orbitrap Velos hybrid ion trap mass spectrometer (Thermo Scientific, Waltham, MA). Separation of peptides was performed by reverse-phase chromatography at a flow rate of 300 nL/min using a Thermo Scientific reverse-phase nano Easy-spray column (Thermo Scientific PepMap C18, 2 $\mu\text{m}$  particle size, 100A pore size, 75  $\mu\text{m}$  i.d. x 50cm length). Peptides were initially loaded onto a pre-column (Thermo Scientific PepMap 100 C18, 5 $\mu\text{m}$  particle size, 100A pore size, 300  $\mu\text{m}$  i.d. x 5mm length) from the Ultimate 3000 autosampler with 0.1% formic acid for 3 minutes at a flow rate of 10  $\mu\text{L}/\text{min}$ . After this period, the column valve was switched to allow elution of peptides from the pre-column onto the analytical column. Solvent A was water + 0.1% formic acid and solvent B was 80% acetonitrile, 20% water + 0.1% formic acid. The linear gradient employed was 2-40% B in 30 minutes. Further wash and equilibration steps gave a total run time of 60 minutes. The LC eluant was sprayed into the mass spectrometer by means of an Easy-Spray source (Thermo Fisher Scientific Inc.). All *m/z* values of eluting ions were measured in an Orbitrap mass analyzer, set at a

resolution of 70000 and was scanned between  $m/z$  380-1500. Data dependent scans (Top 20) were employed to automatically isolate and generate fragment ions by higher energy collisional dissociation (HCD, normalised collision energy: 25%) in the HCD collision cell and measurement of the resulting fragment ions was performed in the Orbitrap analyser, set at a resolution of 17500. Singly charged ions and ions with unassigned charge states were excluded from being selected for MS/MS and a dynamic exclusion window of 20 seconds was employed.

### **2.3.8 MASCOT database searches**

Raw data files were converted to mgf files and peak lists were submitted to the MASCOT search algorithm (Matrix Science) and searched against the UniProt\_Paeruginosa or Paeruginosa\_20170607 database (5563 sequences; 1857270 residues) and a common contaminant sequences database containing non-specific proteins such as keratins and trypsin (116 sequences; 38548 residues). Database searches were performed with the following parameters: two missed trypsin cleavage sites allowed; precursor ion and fragment ion mass tolerances of 20 ppm and 0.1 Da respectively; fixed modifications: cysteine carbamidomethylation; variable modifications: asparagine and glutamine deamidation, methionine oxidation, arginine mono-methylation, arginine di-methylation, lysine mono-methylation, lysine di-methylation, lysine tri-methylation, lysine acetylation, asparagine methylation and histidine methylation. Variable modifications were added in iterative searches to reduce the bioinformatic search space and the incidence of false identifications. A significance threshold value of  $p < 0.05$  and a peptide cut-off score of 20 were also applied. For each identification, the peptide ion score significance threshold was calculated, and this was used for validation. A peptide rank of 1 and an expectation value below 0.05 were also used. All spectra were manually checked to ensure a valid distribution of MS<sup>2</sup> ions and to confirm the modification site localisation.

### **2.3.9 Protein structure prediction and modelling**

Protein structural models were retrieved from the RCSB Protein Data Bank (PDB); LasB (1EZM), ImpA (5KDV), PA0622 (6PYT), PA0620 (6CL6), LasA (3IT7) and OprD (3SY7). For proteins where the structure had not yet been solved, the amino acid



sequence was retrieved from the Pseudomonas Genome Database and submitted to I-TASSER for structural prediction (J. Yang *et al.*, 2014). The predictive model with the highest confidence score was then selected for further modelling. The molecular graphics programme CCP4mg was used to map modified residues onto the protein structures.

### **2.3.10 Bioinformatics**

For prediction of acetylation sites, web-based algorithms Ensemble-PAIL (Xu *et al.*, 2010) and ProAcePred (Chen *et al.*, 2018) were used; the latter using the *E. coli* training dataset. For prediction of methylation sites, MASA (Shien *et al.*, 2009), PRmePRed (Kumar *et al.*, 2017) and iMethyl-PseAAC (Qiu *et al.*, 2014) were used. For sequence conservation analysis, amino acid sequences were BLAST and alignments were visualised with BoxShade. For sequence motif analysis, the ten residues up- and down-stream of each PTM site were listed and submitted to WebLogo to visualise the common amino acids surrounding each site (Crooks *et al.*, 2004). Protein secondary structure at each PTM site was assigned using the DSSP (Kabsch and Sander, 1983). The effects of amino acid substitution on the structure of LasB were predicted using Missense3D (Ittisoponpisan *et al.*, 2019).

## **2.4 DNA manipulation techniques**

### **2.4.1 DNA extraction and quantification**

Genomic DNA (gDNA) was extracted from PAO1 overnight cultures using a GeneJET Genomic DNA Purification Kit (Thermo Scientific) and was subsequently used as a template for gene amplification by polymerase chain reaction (PCR) (see section 2.4.2). Plasmid DNA was similarly extracted using the GeneJET Plasmid Miniprep Kit (Thermo Scientific); plasmids used in this research are shown in Table 2.7. The concentration and quality of all DNA preparations were determined using a NanoDrop ND-1000 spectrophotometer and they were then stored at -20°C.

**Table 2.7** Plasmids used in this research

Plasmid	Description	Reference
pUCP20	<i>Escherichia</i> to <i>Pseudomonas</i> shuttle vector, Cb <sup>R</sup>	West 1994
pLasB	pUCP20 containing the wild-type <i>lasB</i> gene preceded by its endogenous Shine-Dalgarno sequence, Cb <sup>R</sup>	This study
pLasB-211	pUCP20 containing <i>lasB</i> gene as above with a lysine to alanine codon substitution at K211, Cb <sup>R</sup>	This study
pLasB-226	pUCP20 containing <i>lasB</i> gene as above with an arginine to alanine codon substitution at R226, Cb <sup>R</sup>	This study
pLasB-378	pUCP20 containing <i>lasB</i> gene as above with a lysine to alanine codon substitution at K378, Cb <sup>R</sup>	This study
p0622	pUCP20 containing the wild-type <i>PA0622</i> gene preceded by a Shine-Dalgarno sequence, Cb <sup>R</sup>	This study
p0622-171	pUCP20 containing <i>PA0622</i> gene with a lysine to alanine codon substitution at K171, Cb <sup>R</sup>	This study
p0622-172	pUCP20 containing <i>PA0622</i> gene with an asparagine to alanine codon substitution at K172, Cb <sup>R</sup>	This study
p0622-176	pUCP20 containing <i>PA0622</i> gene with a lysine to alanine codon substitution at K176, Cb <sup>R</sup>	This study
p0622-312	pUCP20 containing <i>PA0622</i> gene with a lysine to alanine codon substitution at K312, Cb <sup>R</sup>	This study

#### 2.4.2 Polymerase chain reaction (PCR)

The reagents and parameters used for PCR are detailed in Tables 2.8 and 2.9, respectively. Reaction conditions were modified appropriately depending on the length of the amplicon and the optimal annealing temperature of the primers used. Oligonucleotide primers used in this research were purchased from Sigma-Aldrich and are shown in Table 2.10. Optimal reaction conditions were determined using gradient PCR and variable concentrations of Dimethyl Sulfoxide (DMSO). All PCR reactions were performed using a Veriti Thermal Cycler.

**Table 2.8** PCR reagents

Reagent	Vol (μl)
Template DNA (≤500 ng)	0.5
5X Q5 Reaction Buffer	10
Q5 HF DNA Polymerase	0.5
10 mM dNTPs	1
10 μM Forward primer	2.5
10 μM Reverse primer	2.5
DMSO	0 - 5
Nuclease free water	Up to 50 μl

**Table 2.9** PCR parameters

Step		Temperature (°C)	Time (sec)
Initial denaturation		98	30
35 cycles	Denaturation	98	10
	Annealing	55-70	20
	Extension	72	20 / kb amplicon
Final extension		72	5
Hold		4	∞

#### 2.4.2.1 Colony PCR

Colony PCR was used to confirm the presence of an insert following plasmid transformation (Section 2.4.5). A single colony was inoculated into 50 μl sterile dH<sub>2</sub>O, boiled at 95°C for 10 mins and then centrifuged (8,000 x *g*, 5 mins) to remove cell debris. An aliquot (1 μl) of the supernatant was used in place of the DNA template in the PCR mixture.

**Table 2.10** Oligonucleotide primers

N.B. Restriction sites are underlined, Shine-Dalgarno sequences are italicised and mutagenic codons are in lower case; F – forward, R - reverse.

Primer	Sequence (5' - 3')	Site	Product
<i>lasB</i> <sup>WT</sup> F	ACTTGCGAATTCAGGAGGTATCCAATGA AGAAGGTTTCTACGCTTGACC	<i>EcoRI</i>	pLasB
<i>lasB</i> <sup>WT</sup> R	ACAGCGAAGCTTTTACAACGCGCTCGG GC	<i>HindIII</i>	
<i>lasB</i> <sup>K211A</sup> F	GAAGATCGGCgcgTACACCTACGGTAGC GACTACGGTCCGC		pLasB-211
<i>lasB</i> <sup>K211A</sup> R	TGGTTGCCGCCGGGGCCG		
<i>lasB</i> <sup>R226A</sup> F	CGTCAACGACgcgTGCGAGATGGACGAC		pLasB-226
<i>lasB</i> <sup>R226A</sup> R	ATCAGCGGACCGTAGTCG		
<i>lasB</i> <sup>K378A</sup> F	TATGCGCGGCgcgAACGACTTCC		pLasB-378
<i>lasB</i> <sup>K378A</sup> R	TAGAACTCGGCAGCCTCG		
<i>PA0622</i> <sup>WT</sup> F	ACAGCGGAATTCAGGAGGTCGACGATG AGTTTCTTCCACGGCGTTAC	<i>EcoRI</i>	p0622
<i>PA0622</i> <sup>WT</sup> R	ACAGCGAAGCTTTTAGGCGACATCCAGA ACTTCGG	<i>HindIII</i>	
<i>PA0622</i> <sup>K171A</sup> F	CGCCTACGCCgcgAACTTCGGCAGCAAG C		p0622-171
<i>PA0622</i> <sup>K171A</sup> R	ACGGCCGCCTCGTCGGTG		
<i>PA0622</i> <sup>N172A</sup> F	CTACGCCAAGgcgTTCGGCAGCAAGCGC CTGTTTCATGG		p0622-172
<i>PA0622</i> <sup>N172A</sup> R	GCGACGGCCGCCTCGTCG		
<i>PA0622</i> <sup>K176A</sup> F	CTTCGGCAGCgcgCGCCTGTTCATG		p0622-176
<i>PA0622</i> <sup>K176A</sup> R	TTCTTGCGTAGGCGACG		
<i>PA0622</i> <sup>K312A</sup> F	GACCTACGTGgcgGATGTCACCGAGGG		p0622-312
<i>PA0622</i> <sup>K312A</sup> R	TTGGTGATGCCGCGGTCC		

### 2.4.3 Agarose gel electrophoresis

PCR products were separated by size using agarose gel electrophoresis. A 1% (w/v) agarose gel was prepared in 1X Tris-Acetate-EDTA (TAE) buffer with the addition of 0.5 µg/ml ethidium bromide and allowed to set. DNA was mixed with 6X DNA loading dye (Thermo Scientific) and subsequently loaded into the wells of the gel submerged in a tank containing 1X TAE buffer. A 1kb Hyperladder (Bioline) was also loaded as a DNA size reference. The DNA was electrophoresed at 80 V for 60-90 minutes, depending on the size of the fragment, and then visualised using a UV transilluminator. Bands of the correct size were excised from the gel and the DNA was extracted using the GeneJet Gel Extraction Kit (Thermo Scientific).

### 2.4.4 Restriction digest and DNA ligation

Gel-extracted DNA and plasmid preparations were digested with EcoRI and HindIII restriction enzymes at 37°C for 3 hours. Restriction digests were then purified using the GeneJet PCR Purification Kit and quantified as previously described. The cut amplicon and plasmids were subsequently ligated using a 3:1 DNA insert to plasmid molar ratio using T4 DNA ligase (Thermo Scientific). The reaction was incubated on ice for one hour then at room temperature for one hour. A reaction mixture without the insert was included as a negative control. All reaction volumes are detailed in Table 2.11.

**Table 2.11** Restriction digestions and DNA ligation reaction volumes

Restriction digest			Ligation	
	PCR amplicon	Plasmid DNA	Plasmid	50 ng
EcoRI	2 µl	2 µl	Amplicon	As calculated
HindIII	2 µl	2 µl	T4 DNA ligase	1 µl
PCR product	41 µl	-	T4 ligase buffer	2 µl
Plasmid	-	10 µl	dH <sub>2</sub> O	Up to 20 µl
CutSmart buffer	5 µl	5 µl		
dH <sub>2</sub> O	Up to 50 µl	Up to 50 µl		

#### **2.4.5 Transformation of *E. coli* by electroporation**

Overnight cultures of *E. coli* DH5 $\alpha$  were sub-cultured 1:100 into 10 ml fresh LB and grown with rotation at 37°C for 2-3 hours until an OD<sub>600</sub> = ~0.4 was reached. Cells were harvested by centrifugation at 3,200 x *g* at 4°C for 8 minutes and the pellet was then resuspended and re-pelleted three times in 10% (v/v) ice-cold sterile glycerol. Electrocompetent cells were finally resuspended in 500  $\mu$ l 10% glycerol and a 100  $\mu$ l aliquot was added to 2  $\mu$ l of ligation mixture in an electroporation cuvette and incubated on ice for 20 minutes. The cells were transformed by electroporation at 2.5 kV (5  $\mu$ F, 200  $\Omega$ , 5 ms time constant) in an Eppendorf electroporator. Immediately following electroporation, 1 ml of pre-warmed LB was added to the transformed cells, which were then transferred to a microcentrifuge tube and incubated with rotation at 37°C for 1 hour. Cells were then centrifuged at 10,000 x *g* for 5 minutes, resuspended in 100  $\mu$ l LB and inoculated onto selective LBA plates. The plates were incubated for 16-24 hours to allow the growth of transformed colonies.

##### **2.4.5.1 Transformation of *P. aeruginosa* by electroporation**

The same methodology as above was used for electroporation of *P. aeruginosa*, however all steps were carried out at room temperature.

#### **2.4.6 Confirmation of the transposon-insertion site**

Colony PCR was used to confirm the presence of the transposon in the *lasB* and PA0622 transposon-insertion mutants, PW7302 and PW2131 respectively (Table 2.1). The appearance of a large band or no band following gel electrophoresis of the PCR product indicated the presence of the transposon in the gene; PAO1 was used as a positive control.

#### **2.4.7 Construction of pLasB and p0622 complementation vectors**

The open reading frames of the *lasB* and PA0622 genes were PCR amplified from wild-type PAO1 gDNA using the LasB<sup>WT</sup> F/R and PA0622<sup>WT</sup> F/R primers respectively, using conditions described above. These primers also encode an upstream ribosome

binding site which is included in the PCR product; the endogenous SD sequence was used where possible. The amplicons were then gel purified and digested independently alongside the pUCP20 plasmid using EcoRI and HindIII restriction enzymes. The resulting fragments were then separately ligated into digested pUCP20 downstream of the *lac* promoter, generating expression vectors pLasB and p0622. Each construct was independently electroporated into *E. coli* DH5 $\alpha$  and plated on LBA containing carbenicillin to select for positive transformants. Colony PCR was used to screen for the presence of correct constructs, and the plasmids were then extracted for confirmation by DNA sequencing. Once confirmed, pLasB and p0622 were electroporated into their respective transposon-insertion mutants, PW7302 and PW2131, for wild-type gene complementation. Transformants were plated on dual selection LBA plates (supplemented with carbenicillin and tetracycline, Table 2.3) to select for presence of the complementation vector and the transposon resistance cassette. Plates were incubated for 16-24 hours at 37°C and colonies were screened for the presence of the gene by colony PCR.

#### **2.4.8 Site directed mutagenesis (SDM)**

Mutations of key post-translationally modified residues were generated using a Q5 Site-Directed Mutagenesis Kit (NEB). Non-overlapping oligonucleotide primers were designed using the online NEBaseChanger tool (Table 2.10) to substitute the modified residue codon for an alanine codon (GCG) in a process of alanine scanning. Calculated annealing temperatures were used for a modified PCR protocol, as detailed in the manufacturer's instructions. Expression vectors pLasB and p0622 were used as the template DNA. Following PCR amplification, reaction mixtures were treated with a kinase, ligase, DpnI enzyme mix to phosphorylate and circularise the PCR product and remove non-mutagenised template DNA. An aliquot (5  $\mu$ l) was then added to 50  $\mu$ l 5-alpha Competent *E. coli* (NEB) and transformed by heat shock at 42°C for 30 seconds. Transformed cells were recovered in SOC medium at 37°C with rotation for 1 hour. Cells were then plated on dual selection plates (as described in 2.4.7) and incubated at 37°C overnight. The presence of an alanine substitution was confirmed by Sanger sequencing of plasmid extractions from positive colonies.

### **2.4.9 DNA sequencing**

All constructs were confirmed by Sanger sequencing using the DNA Sequencing Facility at the Department of Biochemistry, University of Cambridge, as per instructions.

### **2.4.10 Expression of plasmid-encoded constructs**

The expression and secretion of all plasmid-encoded constructs, both wild-type and alanine substituted mutants, were confirmed through SDS-PAGE of strain secretomes (Sections 2.3.2 – 2.3.4). The presence of a protein band at the correct molecular weight and corresponding absence of a band in the respective transposon mutant indicated successful gene complementation, expression and protein secretion.

## **2.5 Phenotypic assays**

### **2.5.1 Protease production**

Protease production was detected using skim milk agar (SMA) plates (Table 2.2). Overnight cultures were normalised to  $OD_{600} = 1$  and 3  $\mu$ L was spotted onto the plates and left to dry. Plates were incubated at 37°C for 24 hours, after which a proteolytic halo surrounding the colony could be seen and measured. Alternatively, 50  $\mu$ L of concentrated cell free supernatant was added to 100  $\mu$ L 4% (w/v) skim milk in a 96-well microtiter plate and the absorbance at 595 nm was measured following 18 hours of incubation at room temperature.

### **2.5.2 Elastase activity**

Elastinolytic activity was assessed using Elastin-Congo red as a substrate. 10 mg of Elastin-Congo red was added to 1.05 ml reaction buffer (0.05 M Tris-HCl, 0.5 mM  $CaCl_2$ , pH 7.5) in a microcentrifuge tube and immediately vortexed. Concentrated cell free supernatants (50  $\mu$ L) from cultures grown overnight were added and the mixture was incubated on a rotator at 37°C for 18 hours. An equal volume of LB was used as a negative control. The reaction was stopped by adding 100  $\mu$ L of 0.12 M EDTA and



the tubes were centrifuged (8,000 x *g*, 10 minutes) to pellet any undigested insoluble substrate. The absorbance of the supernatants was then measured at 495 nm (Eppendorf BioPhotometer); all samples were measured with two biological replicates.

### **2.5.3 Staphylolysin activity**

Heat-killed *S. aureus* was prepared by boiling an aliquot of a *S. aureus* Newman overnight culture at 95°C for 15 minutes. Cells were pelleted by centrifugation and then resuspended in 20 mM Tris-HCl (pH 8). A 10 µL aliquot of concentrated *P. aeruginosa* cell-free supernatant was added to 100 µL of resuspended *S. aureus* cells in a 96-well microtiter plate and the cell lysis was measured at 595 nm following 16 hours of incubation (FLUOstar Omega). Alternatively, 8 µL of an overnight culture of *S. aureus* Newman was added to 8 ml of top agar (LB with 0.35% agar) and poured onto a square LBA plate. Wells were then punctured in the set agar using a sterile P1000 tip and a 5 µL aliquot of cell-free supernatant was dispensed into the well. The plate was incubated at 37°C for 24 hours; a halo of no growth surrounding the well indicated staphylolytic activity.

### **2.5.4 Motility assays**

#### **2.5.4.1 Swarming motility**

Overnight cultures were normalised to OD<sub>600</sub> = 1 and 5 µL was spotted onto swarming agar plates (Table 2.2). Plates were left to dry and incubated upright at 37°C for 24 hours.

#### **2.5.4.2 Swimming motility**

Overnight cultures were normalised to OD<sub>600</sub> = 1 and 5 µL was inoculated at the bottom of swimming agar plates (Table 2.2). Plates were left to dry and incubated upright at 37°C for 24 hours.

### 2.5.5 Biofilm formation

Overnight cultures were normalised to  $OD_{600} = 0.1$  in fresh LB media and 100  $\mu$ L was added in triplicate to a 96-well microtiter plate. Plates were sealed with a sterile gas-permeable film and incubated either statically or on a rotating platform (150 rpm) in a humid chamber at 37°C for 24 hours. All of the following steps were then carried out at room temperature. Planktonic cells were aspirated from the wells and they were then washed with sterile dH<sub>2</sub>O to fully remove unadhered cells. Crystal violet (100  $\mu$ L, 0.1% w/v) was added, and the plate was incubated for 15 minutes. The wells were then washed 3 times with 120  $\mu$ L dH<sub>2</sub>O and the plates were left to dry. Adherent cells were destained by adding 120  $\mu$ L of 30% (v/v) acetic acid and incubating for 15 minutes. Resolubilised crystal violet, and thus biofilm mass, was quantified by measuring the absorbance at 595 nm.

### 2.5.6 Statistical analysis

All graphs show the mean value with error bars representing the standard deviation. Where appropriate, statistical analysis was performed using a one-way ANOVA with Tukey's multiple comparison *post hoc* test (\* =  $P < 0.05$ , \*\* =  $P < 0.01$ , \*\*\* =  $P < 0.001$ , \*\*\*\* =  $P < 0.0001$ ).

## 2.6 *Caenorhabditis elegans* infection model

The nematode *Caenorhabditis elegans* was used as a simple model for studying the establishment and dynamics of infection by *P. aeruginosa* and the elastase site-directed mutants. *C. elegans* DH26 (genotype *fer15-(b26)*) was used for slow-killing assays as it is a temperature sensitive strain which is sterile at assay temperatures of 25°C. Worms were maintained at 15°C on thin lawns of *E. coli* OP50 grown on NGM plates. All media and reagents are detailed in Table 2.12.

### 2.6.1 Synchronisation of the worm culture

When there were sufficient gravid adults, the worm culture was synchronised by bleaching. Briefly, worms and eggs were washed from plates with 5 ml M9 media and

sedimented for 1 minute at 1,500 x g. Most of the supernatant was removed and 1 ml 40% bleaching solution was added; lysis was monitored under a dissecting microscope until the majority of worms had lysed. Eggs were subsequently washed in 14 ml M9 media and re-pelleted by centrifugation (1,500 x g for 2 minutes) for a total of 5 washes. The eggs were then resuspended in 14 ml M9 media and incubated at 25°C overnight in a shaking water bath (150 rpm) to allow the eggs to hatch. Pelleted young larvae were spotted onto NGM-OP50 lawns and grown until the worms had reached the L4 stage (approximately 36 hours).

### 2.6.2 Slow killing assay

*P. aeruginosa* lawns were prepared by spotting 20 µL of an overnight culture onto 3.5 cm slow killing (SK) agar plates; a bare-agar border surrounding the inoculum was maintained. Plates were left to dry and then incubated at 37°C for 24 hours before being moved to a 25°C incubator for 24 hours to allow the expression of virulence factors. L4 worms were picked from NGM-OP50 plates and placed onto the assay plates using a sterile platinum wire. Plates were then sealed with parafilm and incubated at 25°C. Five replicate plates for each strain with ten worms per plate were used; OP50 was used as a negative control. Plates were scored for live worms every 24 hours for 11 days; a worm was considered dead when it no longer responded to touch. Live worms were passaged onto fresh lawns every 24 hours after counting to avoid the inclusion of any newly hatched larvae, reduce plate contamination and to ensure sufficient bacteria were present for feeding.

**Table. 2.12** *C. elegans* slow killing assay media and reagents

Reagent	Components per litre
Phosphate buffer (pH 6)	108.2 g KH <sub>2</sub> PO <sub>4</sub> 35.6 g K <sub>2</sub> HPO <sub>4</sub>
M9 buffer	3 g KH <sub>2</sub> PO <sub>4</sub> 6 g Na <sub>2</sub> HPO <sub>4</sub>
Added post-autoclaving	1 ml 1M MgSO <sub>4</sub>
40% Bleaching solution	250 ml 4 M NaOH 400 ml Fresh bleach
Nematode Growth Medium (NGM)	2.5 g Peptone

<b>Table 2.12 continued</b>	
Added post-autoclaving	3 g NaCl
	20 g Agar
	1 ml 1 M MgSO <sub>4</sub>
	25 ml Phosphate buffer (pH 6)
	1 ml 1 M CaCl <sub>2</sub>
Slow killing agar (SK)	1 ml 5 mg/ml cholesterol (in ethanol)
	3.5 g Peptone
	3 g NaCl
	17 g Agar
	1 ml 1 M MgSO <sub>4</sub>
Added post-autoclaving	25 ml Phosphate buffer (pH 6)
	1 ml 1 M CaCl <sub>2</sub>
	1 ml 5 mg/ml Cholesterol (in ethanol)

## 2.7 Pyocin assays

### 2.7.1 Pyocin induction

PAO1, PW2131 and all PA0622 site-directed mutants were cultured in 50 mL LB until OD<sub>600</sub> = 0.5. A sub-inhibitory concentration of ciprofloxacin (0.5 X MIC; Table 2.3) was subsequently spiked into the cultures and they were grown for an additional 2.5 hours. Cells were removed by sedimentation and the culture supernatants were filter sterilised.

### 2.7.2 Pyocin susceptibility testing

*P. aeruginosa* PAK indicator strain (Table 2.1) was used to perform spot assays of PA0622 site-directed mutants. PAK was grown overnight and subsequently diluted to OD<sub>600</sub> = 1. An aliquot (8 µL) was then added to 8 mL top agar (LB with 0.35% agar) and poured onto square LBA plates. Cell-free supernatants from the induced cultures were then harvested and 5 µL was spotted onto the agar. Plates were incubated at 37°C for 24 hours. A clear lysis halo indicated pyocin activity.

# Chapter 3

## 3 Factors affecting the *P. aeruginosa* secretome

### 3.1 Introduction

*P. aeruginosa* is a significant burden throughout the healthcare setting. Several studies have therefore sought to understand how it adapts to different niches by defining the complex dynamics that occur within the bacterium's proteome. However, the extracellular sub-proteome is commonly excluded from analysis by the removal of culture supernatants (Lecoutere *et al.*, 2012). Secreted proteins constitute a significant proportion of the virulence mechanisms used by *P. aeruginosa* and must therefore be fully characterised in order to understand its physiology and versatility. Through a thorough understanding of its virulence factors, new biomarkers or targets for therapeutic intervention against *P. aeruginosa* can be determined.

The conditions encountered by *P. aeruginosa* during the infection process are highly varied and therefore different experimental culture conditions must be investigated to understand how it adapts. A few studies have already been conducted to elucidate how individual factors relating to the growth environment affect extracellular proteins, however significant inter- and intra-study variability renders an overall appreciation of the dynamics inconclusive (Upritchard *et al.*, 2008; Termine and Michel, 2009; Scott *et al.*, 2013). Review of these papers also highlights how the phenomenon of protein charge isoforms within the secretome has previously been grossly neglected during analysis.

#### 3.1.1 Aims

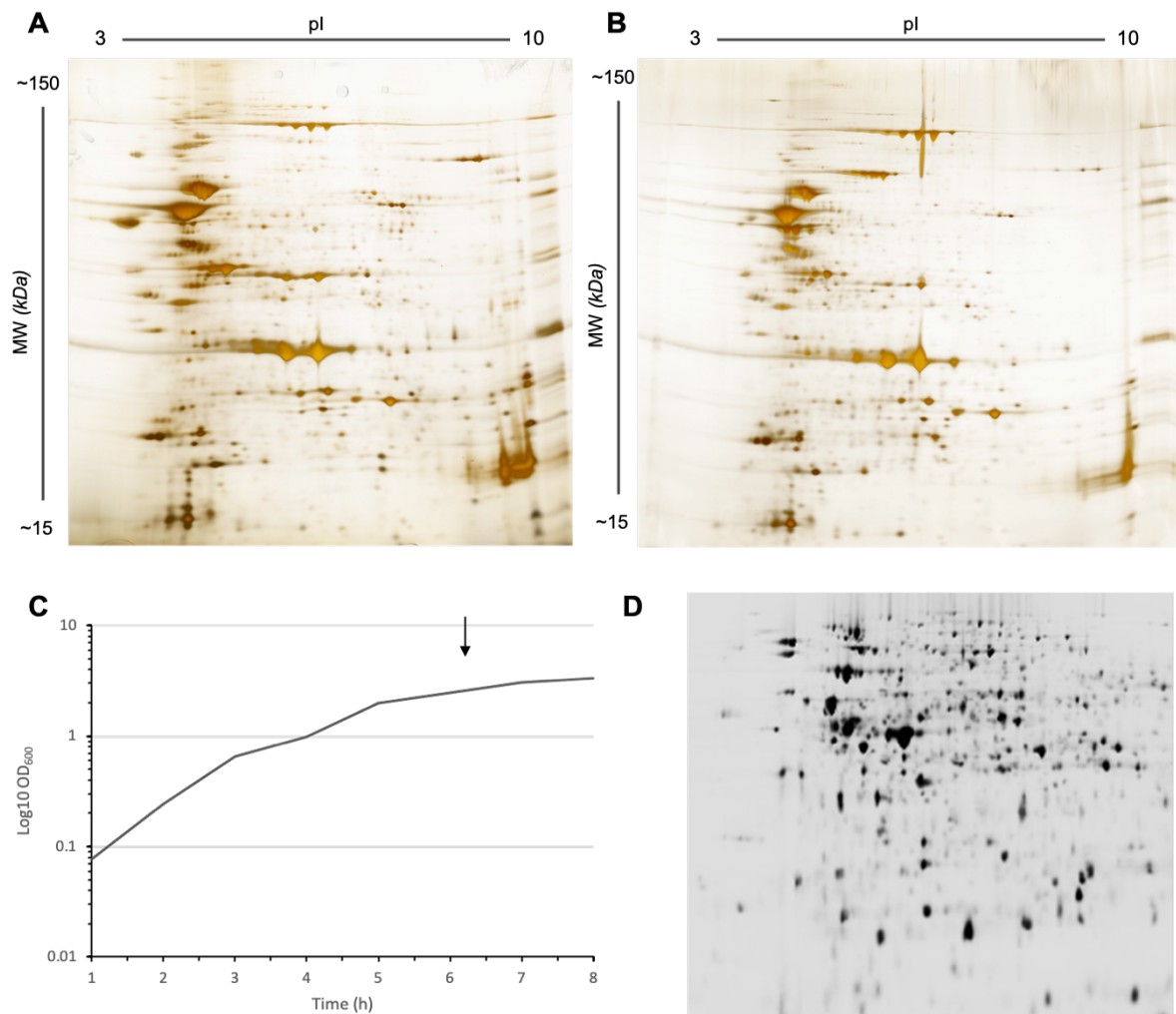
This Chapter aimed to investigate how changing individual components of the growth environment affects the secretome to further understand the fine-tuning of extracellular protein dynamics and charge variation. More specifically, the aims were to look at how changes in nutrient availability, temperature, growth phase (as a

measure of cell density) and strain specificity affect the *P. aeruginosa* secretome. Through the utilisation of 2-DE and 2D-DiGE technology, alterations in the secretome composition and protein charge heterogeneity could be accurately compared between culture conditions.

### **3.2 Visualisation of the *P. aeruginosa* exoproteome**

In order to visualise and confirm the separation pattern of the *P. aeruginosa* secretome, PAO1 was grown to the early stationary phase of growth in AGSY medium and the proteins were concentrated by precipitation from the culture supernatant. The choice of a rich medium and the time of sampling were selected to ensure sufficient protein secretion and better coverage of the available secretome. The secreted proteins were then separated in two dimensions, first by isoelectric point (pI) and second by molecular weight, on 2D acrylamide gels (Figure 3.1).

The gel images in Figure 3.1 show a good separation of the secretome, and this was validated through multiple biological and technical replicates displaying a nearly identical pattern of separation. This demonstrates that the sample preparation, storage and separation protocols are reliable and highly replicable. A small selection of spots (MW ~ 75 kDa, pI ~ 5.5-6) were only visible in one of the sample preparations (Fig 3.1B). Following further investigations of the protein identities using LC-MS/MS, they were determined to be isoforms of the quorum-sensing transcriptional regulator RhIR. Interpretation of the MS/MS data presented a high confidence score and 96% sequence coverage of the protein, suggesting that it was a valid positive identification. This result was highly surprising as RhIR is not known to be secreted from the cell, nor act as a moonlighting protein on the cell surface. Upon repetition, and in every subsequent sample preparation, these spots were absent and have therefore been classified as a contamination artifact from laboratory equipment and thus disregarded.



**Figure 3.1 2-DE of the *P. aeruginosa* secretome**

Proteins secreted by PAO1 grown at 37°C in AGSY medium until the late exponential phase of growth were subjected to two-dimensional gel electrophoresis (IEF pH 3-10 non-linear, 13cm) and were visualised with silver staining. The scanned gel images (A & B) are representative of multiple biological and technical replicates. (C) Growth curve of PAO1 with the time of sampling indicated by an arrow. (D) Intracellular PAO1 proteome separated by 2-DE, image from Dr Jane Gray (unpublished).

Overall, an average of 1187 spots were identified following image analysis. This is significantly more than several previous studies, although similar to more recent publications (Arevalo-Ferro *et al.*, 2003, 2004; Lecoutere *et al.*, 2012). This is likely to be due to improved sensitivity of the software used, however false positive and negative spot identifications are also likely. As expected, there appeared to be a significant number of protein species with similar molecular weights but variable isoelectric points presenting as lateral “charge trains”. These charge trains tend more toward the acidic side with a higher proportion of charge trains presenting with a lower

range of pIs. In terms of skew within a charge train, most spots appear to increase in abundance as the pI increases, with the exception of the spot with the highest pI. Previous analysis suggested that there are approximately 300 PAO1 exported proteins which, in combination with the gels presented here, suggests many of these proteins are present as multiple isoforms (Lewenza *et al.*, 2005). In *Mycobacterium tuberculosis*, an average of 8.4 protein species per protein were identified in the secretome, suggesting that secreted charge isoforms are prevalent within the bacterial domain (Lange *et al.*, 2014).

The number and separation of spots contained within a charge train was consistent in each secretome preparation, suggesting that this is a legitimate and conserved biological process. This is further supported by comparison with previous literature which showed similar protein isoform numbers from examination of the presented gels (Nouwens *et al.*, 2002). Any subtle differences in the number or intensity of visible spots that can be observed between gels is likely to be due to discrepancies in the length of time allocated for stain development, rather than differences in the protein composition of the samples. However, minor differences in protein abundance are to be expected due to natural biological variance (Termine and Michel, 2009). The extensive presence of charge trains within the secretome is in contrast to the intracellular proteome where they appear to be much less abundant (Figure 3.1D, courtesy of Dr Jane Gray). Such contrast is also evident in the protein profiles exhibited by Swatton *et al.* (Swatton *et al.*, 2016). This may suggest that charge heterogeneity is a necessary, or at least desirable, feature for proteins existing within the extracellular environment.

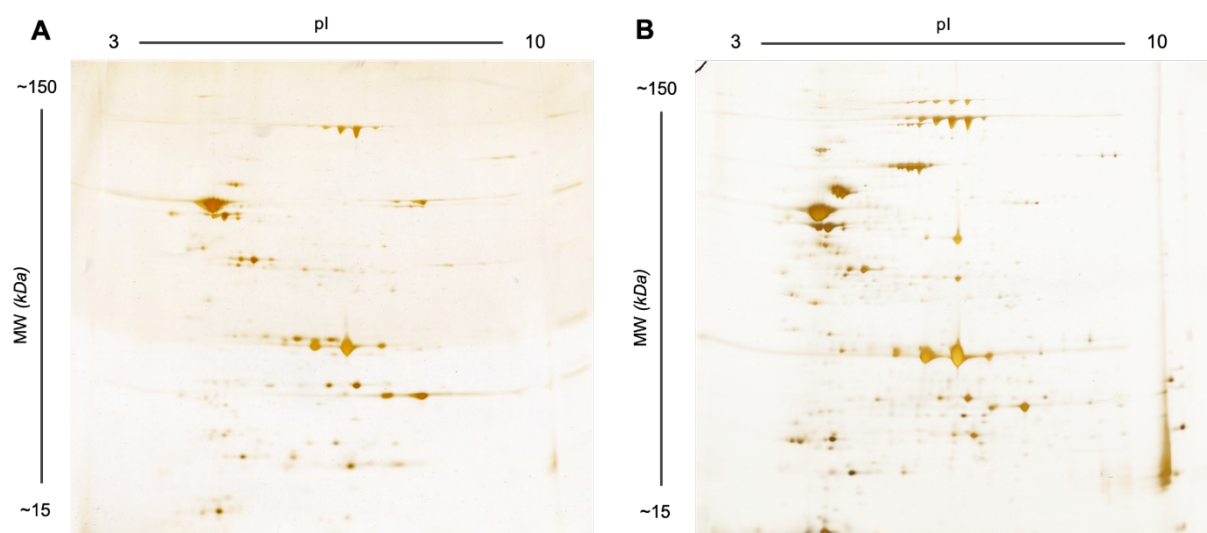
### **3.3 Validation of charge heterogeneity as a biological phenomenon**

Protein charge variation, as seen in 2D-PAGE, is not a novel phenomenon, however there is much debate in the literature regarding whether it is truly a biological occurrence or is merely the result of heterogeneous non-enzymatic modification during sample processing (Deng *et al.*, 2012). The use of certain reagents during both the extraction and solubilisation of proteins, as well as during isoelectric focusing (IEF), have been posited as contributing to changes in a protein's isoelectric point (Righetti, 2006). For example, it is well known that the use of urea as a solubilising agent can



cause the irreversible carbamylation of lysine and arginine residues. This removes the positive charge of the side chain, causing an acidic shift during IEF.

In order to verify that the observed charge heterogeneity is the result of biological activity and not an artifact of the sample preparation procedure, an alternative method of secretome extraction was performed with minimal processing steps or exogenous chemicals. Secreted proteins from PAO1 culture supernatants were concentrated using a Vivaspın® 20 centrifugal concentrator (10 kDa MWCO, Sartorius) in place of TCA precipitation. The samples were then dialysed to remove excess salts that were present in the culture medium. An initial attempt was made to separate these samples, however despite evidence of multiple charge trains as present in the original gels, the resolution of these samples was poor and migration patterns were disrupted (data not shown). This was largely due to the presence of interfering substances, such as pyocyanin, in the sample medium which caused high conductivity during IEF. This separation step uses high field strengths and therefore requires a low ionic strength sample buffer. To ameliorate this issue, the samples were ultimately prepared with a commercial 2-D Clean-Up Kit (GE Healthcare), which does not alter the spot separation pattern.



**Figure 3.2 2-DE comparison of secretome extraction techniques**

2-D gel electrophoresis comparison of the PAO1 secretome when extracted through centrifugal concentration with Vivapsin® columns (A) or TCA precipitation (B) of late exponential phase culture supernatants. Proteins were separated on 13cm IEF pH 3-10 non-linear strips and were visualised with silver staining.

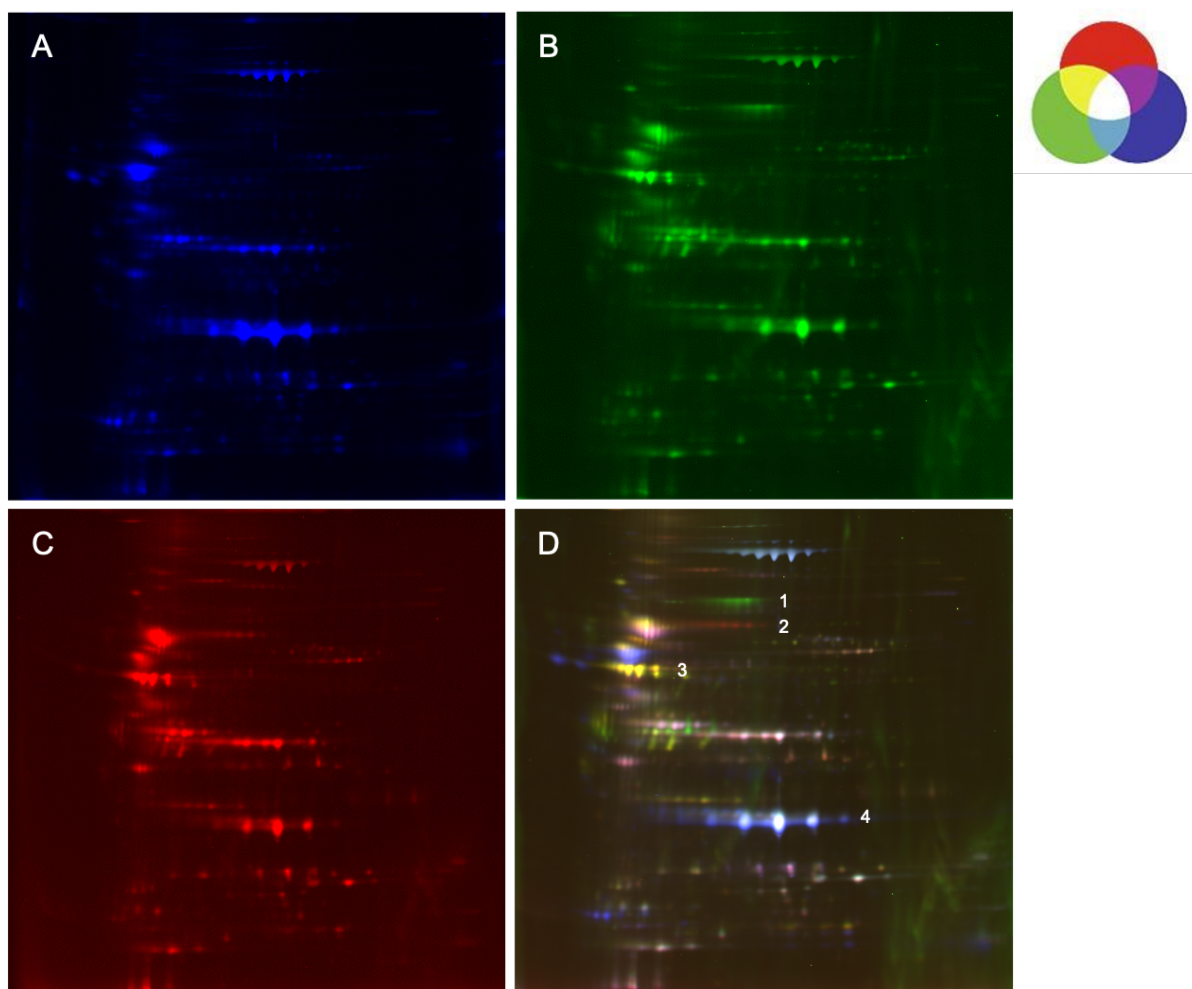
Upon comparison of the two methods of secretome extraction (Figure 3.2), the spot separation pattern remained consistent and charge trains were present in both samples. As previously mentioned, differences in the number or intensity of spots are likely to be due to minor differences in protein loading or the length of stain development. However, it must also be considered that TCA precipitation is beneficial for the extraction of basic proteins which may contribute to minor changes observed (Rabilloud and Lelong, 2011). These results ultimately show that the charge heterogeneity seen in the secretome is unlikely to be due to the sample preparation procedure and is in fact due to biological processing by the organism. The sample prepared by centrifugal concentration did not undergo solubilisation in a urea buffer, and therefore lysine carbamylation can be ruled out as the charge train causative agent. The use of other chemicals was also kept to a minimum, further suggesting that the use of standard reagents in sample preparation is unobjectionable for such studies.

### **3.4 The effect of carbon source on charge variation**

The choice of growth media when conducting any experiment with bacteria is critical to ensure a valid representation of the underpinning biology. Differences in the nutrient composition of the media can have significant effects on protein output (Marquart, Dajcs, *et al.*, 2005). This is particularly relevant for *P. aeruginosa* which demonstrates an extensive metabolic diversity in order to adapt to different environments. It's ability to survive and thrive on a variety of nutrient sources through metabolic plasticity translates to significant changes in the cell transcriptome (LaBauve and Wargo, 2012; Dolan *et al.*, 2020). Previous studies have shown that the presence or absence of certain media constituents, such as magnesium and phosphate, can affect the expression of multiple virulence factors, including secreted proteins (Ball *et al.*, 2016; Lotfy *et al.*, 2018).

To further examine the biological relevance of charge variation within the secretome, cultures grown in different media with different carbon sources were compared. A variation on 2-DE was used whereby the samples were differentially labelled with fluorescent dyes and simultaneously run on the same gel. Two-dimensional difference in-gel electrophoresis (2D-DiGE) allows direct comparison of protein

samples through the use of size and charge matched CyDyes, eliminating common issues associated with gel-to-gel variation. Following differential labelling, samples are pooled, electrophoresed and can be visualised independently under different wavelengths according to the respective dye excitation/emission spectrum. By overlaying the resulting gel images, changes in protein composition and abundance can be readily visualised without the need for potentially challenging spot matching between gels.



**Figure 3.3 2D-DiGE of PAO1 secretome samples grown in different media**

PAO1 was grown at 37°C in AGSY (A), MOPS glucose (B) or MOPS acetate (C) medium and the TCA-extracted secretome samples were labelled with different fluorescent dyes (Cy3, Cy5 and Cy2, respectively), pooled and run simultaneously on the same 2D gel. The gel was scanned with a Typhoon 9400 scanner at wavelengths 532 nm (A), 633 nm (B) and 488 nm (C). (D) An overlay of the three different culture condition images showing commonalities and discrepancies between the samples, as represented by the colour Venn diagram in the top right. Numbers are referred to in the text.

In this experiment, PAO1 was grown in AGSY as standard, as well as MOPS minimal medium supplemented with 20 mM of glucose or acetate. The AGSY, MOPS glucose and MOPS acetate grown samples were labelled with Cy3 (blue), Cy5 (green) and Cy2 (red), respectively. In the image overlay, any spots common to all three samples appear as white and overlapping spots between two growth conditions appear as represented by the colour Venn diagram, i.e. yellow for spots common to both MOPS conditions (Figure 3.3). Spots discrete to a specific growth medium present as the respective dye colour.

What is clearly and predominantly evident from the results presented in Figure 3.3 is that proteins are secreted by *P. aeruginosa* in a charge variable manner despite differences in nutrient composition and different carbon availabilities within the growth media. In all culture conditions, there are many charge trains present, with several abundantly expressed proteins showing overlapping migration patterns. In addition, where charge trains are present, the same number of isoforms are present in each culture condition, suggesting a consistent and important underpinning biological cause.

As expected, there are some proteins and charge trains which are unique to certain growth media, particularly with regard to the samples grown in minimal media. Two charge trains unique to each MOPS condition (Figure 3.3D 1 & 2) have the same isoform number and pI range, but distinct molecular weights. This may indicate a carbon source driven difference in proteolytic processing or large molecule PTM of a protein causing a molecular weight shift, without affecting other charge modifying PTMs. Growth in minimal media also caused an upregulation of one protein charge train (Figure 3.3D 3), supporting the notion that nutrient limitation can stimulate the secretion of certain virulence factors.

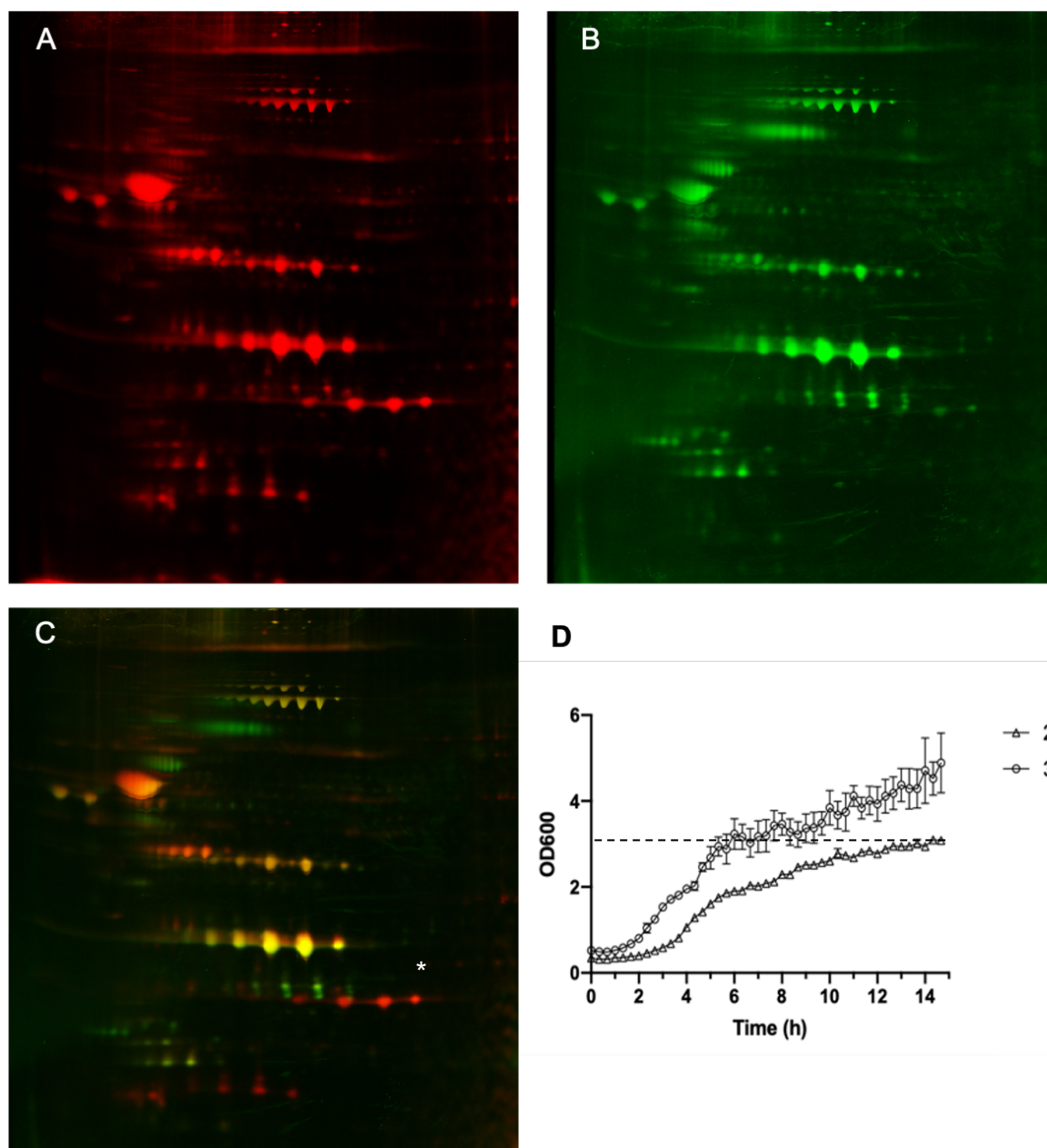
These results are somewhat surprising as previous studies have suggested that growth media has a more extreme and widespread effect on the secretome profile (Nouwens *et al.*, 2003; Scott *et al.*, 2013). In agreement with these findings, however, the charge train purported to be composed of LasB isoforms (Figure 3.3D 4) is more abundant in the rich growth media condition. The use of the CyDyes allowed greater sensitivity than the silver stain (Figure 3.1) for highlighting less abundant isoforms

within certain charge trains, exemplifying known differences in protein stain sensitivities.

### 3.5 The effect of growth temperature on charge variation

As an environmental organism and opportunistic pathogen, *P. aeruginosa* needs to be able to survive at a broad range of temperatures. Throughout the infection process there are marked changes in temperature from the external environment, through colonisation and ultimately during establishment within the human host. Its ability to infect a wide range of other species, such as plants and insects, also puts pressure on its capacity for adaption within variable climates. In addition, the temperature-variable phases of infection require the differential expression and secretion of proteins specifically designed to assist progression through each stage. In fact, 6.4% of the genome is transcriptionally thermoregulated, with particular emphasis on the transcription of exoproteins (Barbier *et al.*, 2014). Additionally, certain enzymes which modify proteins prior to secretion are controlled by structural thermoregulatory processes, promoting invasion tactics of the pathogen during the early stages of infection (Owings *et al.*, 2016).

An investigation into whether charge variation of secreted proteins is observed at different temperatures was therefore undertaken. In a similar fashion to the previous experiment, 2D-DiGE was utilised and the supernatants from cultures grown in AGSY at 25°C and 37°C were compared. Samples were taken from cultures at the same cell density ( $OD_{600} = 3$ ) (Figure 3.4D). Strikingly, there appeared to be marked differences between the secretome profiles of PAO1 grown at 25°C as opposed to 37°C (Figure 3.4). A few charge trains evident in previous experiments were consistently expressed at both temperatures with equal numbers of spots and equivalent abundance. However, many charge trains were only expressed at one of the temperatures. It was also found during sample preparation that the overall abundance of protein secreted at 25°C was 30% greater than at 37°C. This is consistent with transcriptomic upregulation of components of the T2SS at 28°C (Wu *et al.*, 2012). It is also explained by the prolonged growth time allowing a higher secretion titre.



**Figure 3.4 2D-DiGE of PAO1 secretome samples grown at different temperatures**

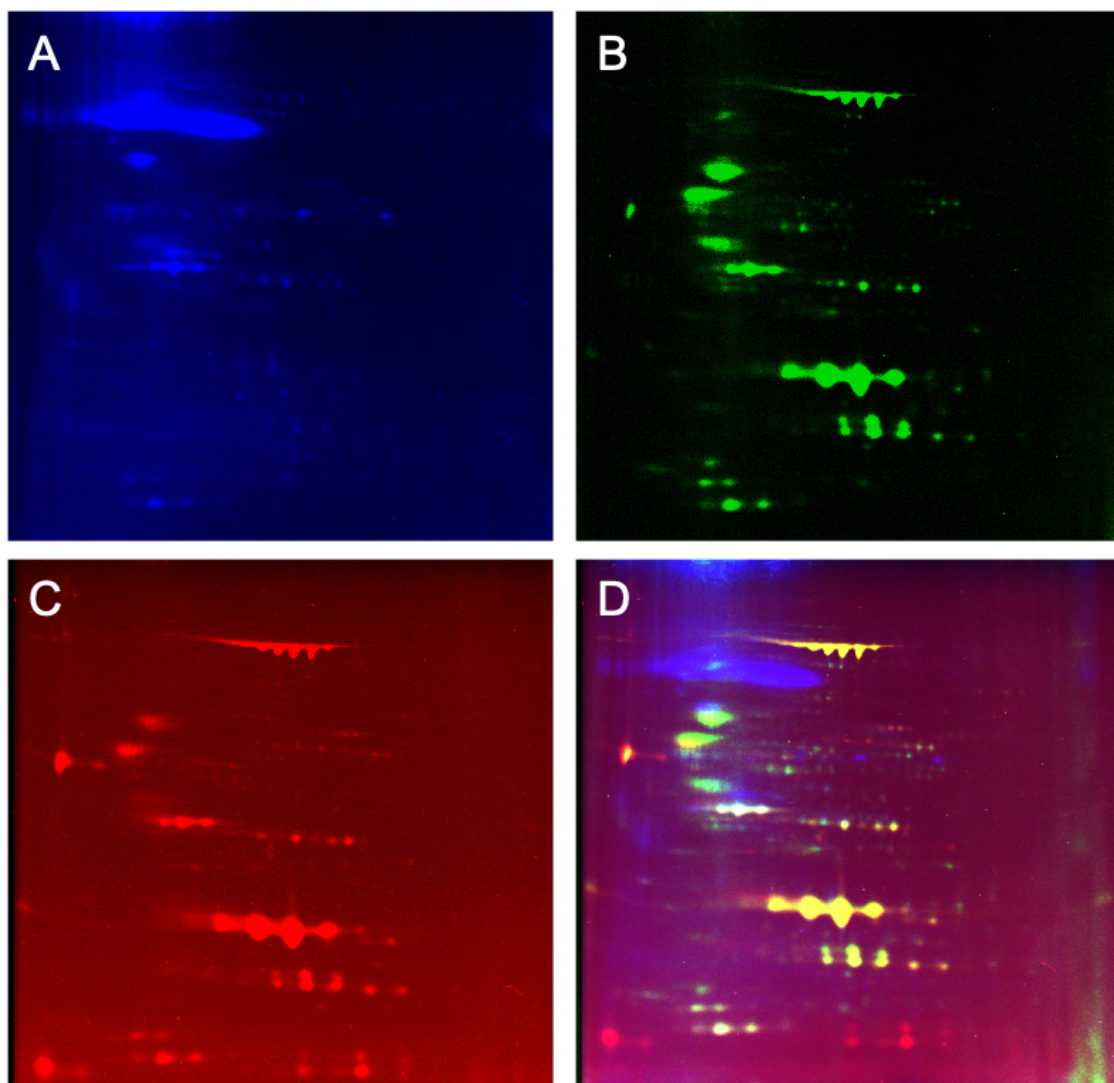
PAO1 was grown at 25°C (A) and 37°C (B) in AGSY medium and the extracted secretome samples were labelled with Cy2 and Cy5, respectively and run on the same 2D gel. The gel was scanned with a Typhoon 9400 scanner at wavelengths 488 nm (A) and 633 nm (B). (C) An overlay of the two gel images showing commonalities and discrepancies between the samples. (D) Growth curves of PAO1 grown at 25°C and 37°C with a dotted line showing the time of sampling; error bars represent the standard deviation of three replicates.

These results are in line with previous transcriptomic analyses which showed that genes encoding secreted factors are commonly up- or down-regulated at comparable experimental temperatures (Termine and Michel, 2009; Barbier *et al.*, 2014). A star has been used to indicate the likely location of the Protease IV charge train (Figure 3.4C), as inferred from comparisons with previous research (Ball *et al.*, 2016). This charge train was up-regulated at least 3-fold at 25°C upon image analysis. Unfortunately, it was not possible to select spots for sequencing due to technical issues, however it would be beneficial for future studies to investigate which proteins are preferentially secreted at each temperature. This would likely give more information about a virulence factor's biological role and predict when it may be important during the infection process. The key take away from these results is that charge variation of secreted proteins appears to be important both at environmental temperatures, as well as at internal host temperatures.

### **3.6 The effect of growth stage on charge variation**

In all experiments thus far, a specific time point for protein sampling was selected which ensured sufficient concentrations of proteins were secreted and allowed the separation of many charge variant protein species. However, this only represents a snapshot at one time point in the growth cycle and overlooks other potentially insightful sampling points in charge train development. Growth stage is a factor of cell density and, as such, community dynamics are likely to have an effect on protein secretion. Furthermore, it is known that protein secretion is somewhat governed by density-dependent quorum sensing. Therefore, in addition to the late exponential/early stationary phase time point, samples were also taken at the mid-exponential phase and mid-late stationary phases of growth. These samples were subjected to 2D-DiGE analysis and compared as previously detailed (Figure 3.5).





**Figure 3.5 2D-DiGE of PAO1 secretome samples from different time points**

PAO1 was grown at 37°C in AGSY medium and secretome samples were taken at 2.5 hours (A), 6.5 hours (B) and 24 hours (C) of growth. The samples were labelled with different fluorescent dyes (Cy3, Cy5 and Cy2, respectively), pooled and run simultaneously on the same 2D gel. The gel was scanned with a Typhoon 9400 scanner at wavelengths 532 nm (A), 633 nm (B) and 488 nm (C). (D) An overlay of the three culture condition images showing commonalities and discrepancies between the samples.

The results of the growth stage assay were not as clear and defined as hoped. Whilst the early entry and late stationary phases of growth showed a similar protein pattern, the mid-exponential sample displayed limited protein content and separation. This was not to be unexpected as protein secretion is known to be massively upregulated following the logarithmic phase of growth (Arevalo-Ferro *et al.*, 2003). This gel scan was also unfortunately slightly obscured by a presumed experimental artifact in the



top left corner. This issue would likely be ameliorated with additional replicates and fresh dyes. Several major charge trains were absent from the mid-exponential time point, including the LasB protein. As this exoprotease is under the control of the quorum sensing regulon, its expression is generally not induced until a critical cell threshold is reached, typically near stationary phase (Kuang *et al.*, 2020).

Intriguingly, a charge train that was later identified as isoforms of the R2 pyocin sheath protein (see Chapter 4) was clearly present in all three samples. This is surprising as these bactericidal complexes are typically released by cell lysis as a result of activation of the SOS response to stress (Turnbull *et al.*, 2016). One would therefore not expect their release until it is triggered in response to stressful environmental stimuli, such as nutrient limitation in the later stages of growth. Such triggers were unlikely to be encountered in the experimental conditions at this time point. Pyocin release could therefore be a precautionary measure of altruistic suicide during community expansion, or even during early host colonisation (suggested by up-regulation at 25°C, Figure 3.4), to enhance group survival and reduce competition. This does however require confirmation and extensive further investigation. More importantly, this protein appears to be variably charged from early in the growth curve, and this is consistent into the mid to late stages of growth.

It was also noted that the late-exponential/early-stationary sample contained some unique charge trains (shown as green spots). This is representative of the transient nature of some secreted proteins that are not expressed until later in the growth curve and are then subsequently degraded (Wehmhöner *et al.*, 2003). An abundance of unique very low molecular weight species in the mid-late stationary phase may be due to these degradation products. Once again, these results highlight that many proteins are secreted in a charge variable way throughout the growth phases and that they are stable into the late stationary phase. A more diverse set of early sampling times may highlight the progressive secretion of charge variants, as was seen in a study of *Leptospira biflexa* (Stewart *et al.*, 2016).

### 3.7 Secretomes from clinical isolates

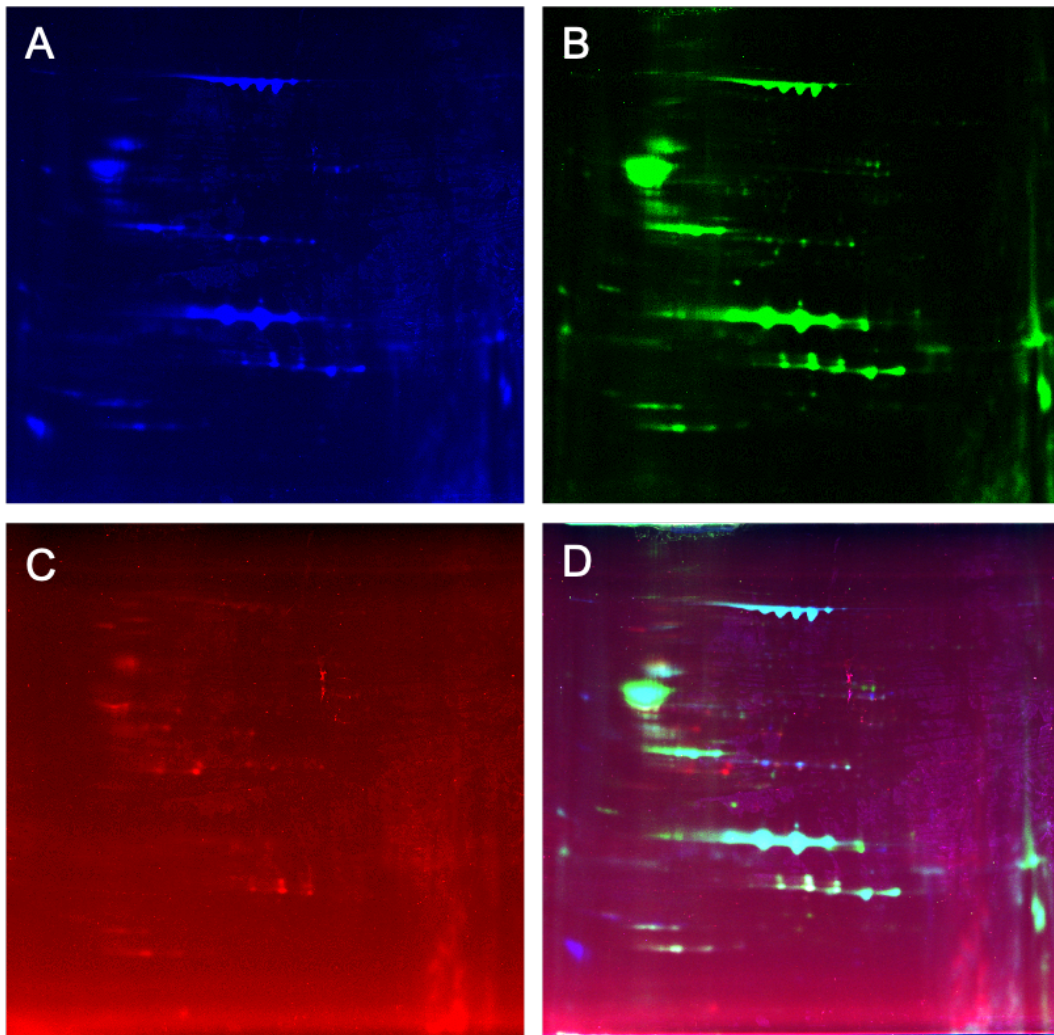
The secretomes of different strains of *P. aeruginosa* can be highly discrepant. In fact, Arevalo-Ferro et al stated that the extracellular sub-proteome is a sensitive measure of strain variation (Arevalo-Ferro *et al.*, 2004). This was further evident during analysis of my results in view of the fact that the PAO1 2D secretome pattern from previous publications were measurably different. This is somewhat due to the likely emergence of phenotypic and genotypic variants in different laboratories from differences in the maintenance and passaging of strains. This drives microevolution, resulting in notable differences in protein expression (Klockgether *et al.*, 2010).

In studying *P. aeruginosa*, it is necessary to understand the clinical relevance of any findings. PAO1, which was originally isolated from a wound, was therefore compared with other clinical isolates to determine their secretome profiles and presence of charge variants. The secretomes from *P. aeruginosa* strains isolated from medical implants in Pakistan by visiting researcher Sidra Abbas were analysed alongside PAO1 by 2D-DiGE.

There was a surprising level of overlap between the secretomes of PAO1 and one of the implant isolates (B5; Figure 3.6). This initially led me to assume that errors had been made in sample labelling and that these were potentially the same strain. However, upon consulting the literature, a considerable number of similarities in the secretomes between PAO1 and cystic fibrosis lung isolates have previously been identified (Sriramulu *et al.*, 2005; Scott *et al.*, 2013). This is in contrast to other researchers who have suggested that inter- and intra-clonal diversity lies within the exoproteome (Wehmhöner *et al.*, 2003). The data supporting the latter statement is, however, plagued by substantial differences in 2D separation between experiments, despite the same sample type being analysed. Their results must therefore be considered with caution.

Although the separation of isolate NB1 was poor, a notable absence of the prominent LasB charge train was observed. This strain was later found to be deficient in the LasR-LasI quorum sensing system which controls LasB expression. These results also correlated with a phenotypic reduction in protease activity in this strain (data not

shown). Nouwens et al also highlighted an absence of this protease in a clinical cytotoxic strain as compared with the more invasive PAO1 strain (Nouwens *et al.*, 2002). In all of the strain secretomes analysed, charge trains are clearly evident with consistent isoform numbers. This further supports the notion that charge variation is an important cross-strain phenomenon. In future experiments it would be interesting to further evaluate whether there are differences observed between clinical and environmental strains of *P. aeruginosa*.



**Figure 3.6 2D-DiGE of *P. aeruginosa* clinical isolate secretomes**

The standard laboratory strain of *P. aeruginosa* PAO1 (A) was grown at 37°C in AGSY medium alongside clinical isolates from Pakistan, B5 (B) and NB1 (C), and the culture secretomes were then extracted using TCA precipitation. The samples were labelled with different fluorescent dyes (Cy3, Cy5 and Cy2, respectively), pooled, and run simultaneously on the same 2D gel. The gel was scanned with a Typhoon 9400 scanner at wavelengths 532 nm (A), 633 nm (B) and 488 nm (C). (D) An overlay of the three strain images showing commonalities and discrepancies between the secretome samples.

### 3.8 Discussion

In understanding the biology of *P. aeruginosa* and the reasons why it is such a successful opportunistic pathogen, it is vital to study the factors which underpin its virulence. Proteins secreted into the extracellular environment drive infection and can have a critical influence on disease prognosis. The first full study of the *P. aeruginosa* secretome was conducted in 2002 by Nouwens *et al.*, however further developments since then have been few and far between (Nouwens *et al.*, 2002). In addition, differences in experimental parameters between studies have made comparison difficult. It was therefore necessary to conduct a series of investigations into factors which can affect the secretome profile of *P. aeruginosa*, whilst maintaining many consistent culture conditions. This ultimately supported a deeper understanding of the dynamics of secretome charge variation with respect to the growth environment. In analysing how changes in the culture medium, temperature, growth stage and experimental strain affected the secretome, it became clear that charge heterogeneity of secreted proteins is an indisputable and consistent phenomenon. Substantiating evidence showing that the divergence in protein isoform pI is unlikely to be caused by exogenous sample processing further highlights the importance of comprehending the biology behind it.

The use of 2-DE in this study provided valuable information on the status of proteins within the secretome that cannot be gathered through genetic, transcriptomic or bioinformatic analysis, as conducted previously (Termine and Michel, 2009; Barbier *et al.*, 2014). A lack of exact correlation between transcriptomic changes and changes at the protein level means that proteomic analysis is an essential component in understanding bacterial physiology (Hickey *et al.*, 2018). 2-DE works by separating proteins in the first dimension according to their isoelectric points (pI). A protein's pI is the pH at which the overall net charge on the molecule is zero, which in turn is a function of the primary structure. Proteins with minor discrepancies in amino acid sequence can therefore be visually separated where, at this level of scrutiny, their overall molecular weights are almost indistinguishable. This is particularly pertinent to the discovery of any potential post-translational modifications which can significantly alter pI.

Although several studies have explored the *P. aeruginosa* exoproteome, none so far have discussed the presence of charge isoforms beyond mere mention, instead choosing to focus on individual protein expression dynamics without taking isoforms into consideration. Outside of the bacterial context, suggestions have been made as to the cause of charge trains. One contentious assumption is that they are solely an artifact of sample processing. Urea is commonly used at high concentrations as a chaotropic agent in 2-DE sample preparation. Degradation of urea into isocyanic acid produces a compound which is highly reactive with the protein N-terminus and lysine or arginine side chains. This carbamylation reaction neutralises the side chain positive charge, thus reducing a protein's pI. It has been suggested that sequential carbamylation of a protein produces such charge train artifacts. However, subsequent studies have shown that carbamylation does not occur in the majority of samples which have been prepared and stored properly (McCarthy *et al.*, 2003). Additionally, it has also been shown that deamidation, which can also affect pI, does not spontaneously occur in appropriately handled samples, e.g. kept on ice (Righetti, 2006).

An alternative proposal is that charge trains are the result of conformers, i.e. identical protein species in different 3D conformations (Deng *et al.*, 2012). There are several issues with this theory in the context of most 2-DE studies, not least the common use of chaotropes and detergents in protein solubilisation which cause denaturation of the protein structure. Deng *et al.* further suggest that these conformers are transient and are dependent on experimental conditions, which is refuted by the replicability in the separation pattern presented here. One would additionally expect this effect to be abundant in all cell compartments if it were valid. Overall, these results show that the sample preparation procedure is unlikely to cause charge trains. A more likely suspect is versatile PTM of proteins; this will be further discussed in Chapters 4 and 5.

When evaluated in comparison to the current literature, the separation of the *P. aeruginosa* secretome presented here is of a much higher resolution than the majority of studies (Sriramulu *et al.*, 2005; Upritchard *et al.*, 2008; Scott *et al.*, 2013). In addition, the breadth of conditions explored is beyond other reports, which tend to focus on just one environmental factor. As previously mentioned, the lack of consistency in culture

conditions, as well as the evolutionary changes of PAO1, make it difficult to draw reliable conclusions from multiple studies. Furthermore, even within several publications there are vast inconsistencies between gels of the same sample type (Wehmhöner *et al.*, 2003). Not only was the characteristic spot pattern highly reproducible and consistent in all gels presented here, but the use of DiGE technology has also greatly reduced any potential for gel-to-gel variation and allowed for direct comparison of culture conditions. Some isoforms were also more easily visualised when using these dyes.

The results of the 2D-DiGE analysis of PAO1 grown in different media are somewhat surprising as there is a significant level of overlap between these samples. In contrast, Scott *et al* previously reported major differences between the secretomes of PAO1 grown in LBA and M9 minimal medium, however the resolution of the gels presented is questionable (Scott *et al.*, 2013). Although the experimental media used here is limited in scope, some interesting observations can be made. There appears to be a minor effect of sole carbon source on the secretome profile, as exemplified by the presence of a small number of protein spots unique to each growth condition. Carbon source, however, does not seem to impact the number of protein isoforms within a charge train, despite other reports that it can have an effect on the number PTM events which can alter protein charge (Gaviard, Broutin, *et al.*, 2018).

There is also evidence of an upregulation of one charge train in particular when *P. aeruginosa* was grown in a minimal medium as opposed to a nutrient rich medium (Figure 3.3D, charge train 3). The upregulation appears to be independent of growth temperature (Figure 3.4). This is in line with previous reports which show that limitation of some nutrients can stimulate the secretion of certain virulence factors (Marquart, Dajcs, *et al.*, 2005; Horsman *et al.*, 2012). It was not possible to sequence these spots, however comparison with previous publications suggests that they may be isoforms of FliC or aminopeptidase PA2939 (Nouwens *et al.*, 2002; Ball *et al.*, 2016). In the case of FliC, a flagellin monomer protein, this would be expected as flagella facilitate the acquisition of nutrients in low nutrient conditions (Feldman *et al.*, 1998). Alternatively, the increased abundance may be a function of protease downregulation and thus decreased degradation. FliC is known to be post-translationally modified by

phosphorylation, which may explain its presence as a charge train (Suriyanarayanan *et al.*, 2016).

Alongside differences in nutrient availability, differences in temperature are commonly encountered depending on the infection site, i.e. superficial wound or lung infections, or even at *ex-vivo* environmental locations. Microbial response to temperature fluctuation involves coordinated changes in gene expression to allow efficient adaptation (Kropinski *et al.*, 1987; Barbier *et al.*, 2014). Here I show a high degree of divergence in protein expression and secretion when the growth temperature was lowered from 37°C to 25°C (Figure 3.4). Thermoregulation of secreted proteins can also be independent of transcriptional regulation, but rather structural modification (Owings *et al.*, 2016).

Several abundant charge trains remained consistent, but many were only expressed at one temperature. Interestingly, there appeared to be an almost even number of unique proteins that were expressed at each temperature, suggesting that some secreted proteins are highly specialised to aid establishment within a particular niche. The upregulation at 25°C of one such protein, likely to be protease PrpL involved in corneal infections, is in accordance with transcriptomic analyses showing up to a 72-fold increase at this temperature (Termine and Michel, 2009; Barbier *et al.*, 2014). The abundance of LasB proteins at both temperatures is thus further explained as elastase is necessary for the post-secretional activation of PrpL and therefore must also be present at the lower temperature (Oh *et al.*, 2017).

A study in the spirochete *Leptospira biflexa* showed that as cultures advance towards the stationary phase of growth, the number of isoforms of membrane-bound proteins increases towards the acidic side. Interestingly this result was not replicated in cytosolic proteins (Stewart *et al.*, 2016). Unfortunately, differences in isoform number between sampling points were not identified here (Figure 3.5). This is either due to the limited number of samples taken, or because all isoforms are secreted at the same time during the transition to stationary growth. Interestingly, secreted EF-Tu is mono-methylated during log growth, but the number of methyl groups increases as cells enter stationary phase (Owings *et al.*, 2016). In *E. coli*, methylation of aspartic acid increases during stationary phase and is important for the stability of aging proteins;



this PTM would cause an acidic shift in pI (Li and Clarke, 1992). Soares et al also highlighted an increase in charge-modifying phosphorylation during the later stages of growth of *E. coli*, although only intracellular proteins were investigated (Soares *et al.*, 2013).

Growth in an artificial sputum media to replicate the lung microenvironment would have been beneficial to investigate, however the high concentration of serum proteins and mucins prevented analysis. In addition, exploration of secretome dynamics in other growth conditions, such as atmospheric variations and in competitive environments, would be valuable in future investigations. The identification of proteins that make up charge trains through comparison with published reports was challenging due to vast inconsistencies between studies and will therefore be investigated further in the next chapter, alongside the nature of PTMs of these charge variable protein species.

### **3.8.1 Conclusion**

In conclusion, the results ultimately show that differences in nutrient availability, temperature, growth phase and *P. aeruginosa* strain do not have a significant effect on secreted protein charge variation. The consistency of protein isoform number and distribution in all conditions tested suggests that this is a controlled and targeted biological phenomenon. It further highlights the importance of protein isoforms, and likely PTM, in a range of environmental and infection niches. PTM may therefore have a more critical and underappreciated role for proteins existing within commonly unstable extracellular environments.



## Chapter 4

### 4 Acetylation of *P. aeruginosa* secreted proteins

#### 4.1 Introduction

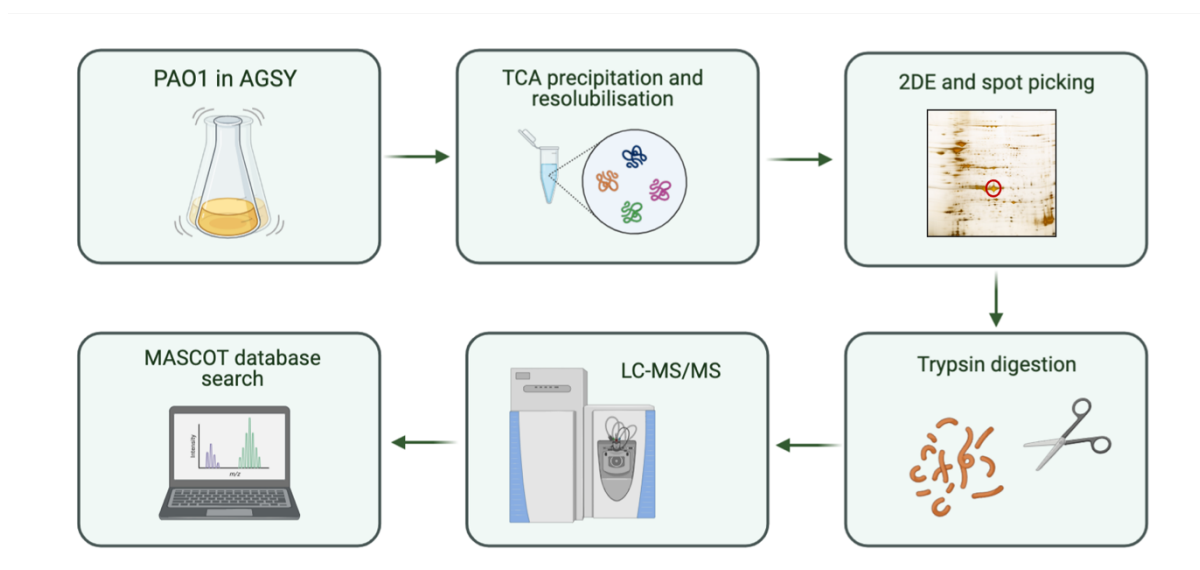
Post-translational modification is a sophisticated mechanism by which an organism can quickly and, in many cases, reversibly extend the chemical repertoire of its protein catalogue in response to environmental stimuli. PTMs act by changing the chemical properties of the amino acids which bear them and therefore have fundamental roles in protein conformation, stability, protein interactions and overall function (Cain *et al.*, 2014). Understanding of prokaryotic PTM systems is quickly developing, however the vast majority of studies fail to incorporate the extracellular protein portfolio (André *et al.*, 2017). This omission prevents deeper understanding of a bacterium's attack strategy through its virulence factors, which are potential targets for therapeutic intervention.

##### 4.1.1 Aims

The presence of protein charge variants secreted by *P. aeruginosa*, as discussed in Chapter 3, put forward the hypothesis that these proteins undergo variable but targeted charge modifying PTM. This Chapter therefore aimed to investigate whether a pattern of PTM could be uncovered that accounted for the trains of spots observed on 2D gels. Furthermore, these aims were focused on identifying the level of acetylation of three main secreted virulence factors, which are the most representative of charge variable isoforms, and providing an in-depth analysis of the nature of acetylation in these candidate proteins.

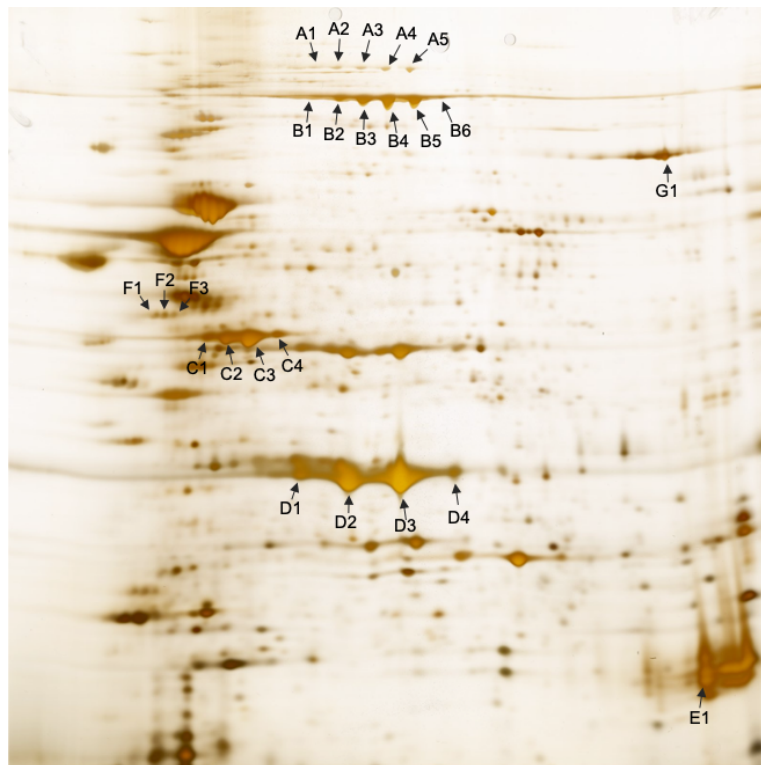
## 4.2 Identification of charge variant proteins from 2D gels

To first confirm that the “charge trains” observed in Chapter 3 were indeed formed by proteoforms of the same protein, a range of select protein spot candidates were picked and analysed by tandem mass spectrometry (Figure 4.1). The proteins were chosen to encompass both abundant and less abundant proteins at a range of molecular weights and isoelectric points. The coding assigned to the 24 respective spots can be seen in Figure 4.2 and their protein identities are outlined in Table 4.1.



**Figure 4.1 Bottom-up proteomic strategy**

PAO1 was grown at 37°C in AGSY medium, and the secreted proteins were extracted from cell-free culture supernatants by TCA precipitation. Resolubilised proteins were then run on 2D gels (13cm IEF pH 3-10 non-linear) and spots from selected charge trains were picked, digested with trypsin and analysed by LC-MS/MS. MASCOT software was then used to analyse the mass spectra and identify both the protein and specified variable PTMs.



**Figure 4.2 Protein spot coding**

2D gel electrophoresis (13 cm IEF pH 3-10 non-linear) of proteins secreted by PAO1 when grown at 37°C in AGSY medium until the late exponential phase of growth. Protein spots selected for LC-MS/MS analysis are indicated by an arrow; the identification of each coded spot is listed in Table 4.1

**Table 4.1 Protein spot identifications from LC-MS/MS data**  
The location of each spot on the 2D gel is detailed in Figure 4.2

Spot ID	Protein name	Locus tag <sup>a</sup>	Function	MW <sup>b</sup> (kDa)	pI <sup>c</sup>	Sequence coverage <sup>d</sup> (%)	No. significant peptide matches <sup>e</sup>	MASCOT score <sup>f</sup>
A1	ImpA	PA0572	Immunomodulating metalloprotease	100.1 (95.9)	6.52 (5.94)	38	68	2280
A2	ImpA	PA0572	Immunomodulating metalloprotease	100.1 (95.9)	6.52 (5.94)	43	114	2792
A3	ImpA	PA0572	Immunomodulating metalloprotease	100.1 (95.9)	6.52 (5.94)	45	119	2964
A4	ImpA	PA0572	Immunomodulating metalloprotease	100.1 (95.9)	6.52 (5.94)	75	641	4438
A5	ImpA	PA0572	Immunomodulating metalloprotease	100.1 (95.9)	6.52 (5.94)	67	386	4797
B1	ImpA	PA0572	Immunomodulating metalloprotease	100.1 (95.9)	6.52 (5.94)	66	196	3498
B2	ImpA	PA0572	Immunomodulating metalloprotease	100.1 (95.9)	6.52 (5.94)	60	214	3479
B3	ImpA	PA0572	Immunomodulating metalloprotease	100.1 (95.9)	6.52 (5.94)	68	412	4504
B4	ImpA	PA0572	Immunomodulating metalloprotease	100.1 (95.9)	6.52 (5.94)	77	562	5050
B5	ImpA	PA0572	Immunomodulating metalloprotease	100.1 (95.9)	6.52 (5.94)	78	840	5104
B6	ImpA	PA0572	Immunomodulating metalloprotease	100.1 (95.9)	6.52 (5.94)	66	283	3704
C1		PA0622	Pyocin sheath protein	41.2	5.08	68	120	1436
C2		PA0622	Pyocin sheath protein	41.2	5.08	69	173	1656
C3		PA0622	Pyocin sheath protein	41.2	5.08	85	118	1354
C4		PA0622	Pyocin sheath protein	41.2	5.08	94	178	2106

Table 4.1 continued								
D1	LasB	PA3724	Elastase	53.7 (33.1)	6.74 (5.87)	75	322	1892
D2	LasB	PA3724	Elastase	53.7 (33.1)	6.74 (5.87)	77	575	2074
D3	LasB	PA3724	Elastase	53.7 (33.1)	6.74 (5.87)	73	238	1244
D4	LasB	PA3724	Elastase	53.7 (33.1)	6.74 (5.87)	71	280	1704
E1	LasA	PA1871	Staphylolysin	45.5	9.05 (7.82)	17	84	422
F1	OprD	PA0958	Porin D	45.5	4.75 (4.87)	55	72	1122
F2	OprD	PA0958	Porin D	45.5	4.75 (4.87)	55	76	1242
F3*	OprD	PA0958	Porin D	45.5	4.75 (4.87)	54	44	982
G1		PA0620	Pyocin tail spike protein	72.3	8.39	33	37	755

<sup>a</sup> Locus tag obtained from the Pseudomonas Genome Database (PGD) (Winsor *et al.*, 2016)

<sup>b</sup> MW shown as the full protein; in brackets the MW of the mature protein

<sup>c</sup> pI as reported on PGD; in brackets the calculated mature protein pI (calculated using Prot pI - <https://www.protpi.ch/>)

<sup>d</sup> Sequence coverage of the mature secreted protein

<sup>e</sup> Number of peptide matches averaged between database searches

<sup>f</sup> Maximum MASCOT score obtained during database searches

All protein spots analysed were matched to a known *P. aeruginosa* protein with a strong confidence score and an excellent number of significant peptide spectral matches. These proteins included the well-known virulence factors LasB and LasA, as well as proteins which make up the bacteriocin protein complex known as R-pyocin (PA0620 and PA0622). An outer-membrane Porin D (OprD) that possesses serine protease activity was also identified, suggesting that it is in some way secreted from

the cell. A more recently discovered virulence factor, ImpA, was also identified in two of the charge trains analysed, forming at least 11 different spots upon 2D separation. Interestingly another study also demonstrated 11 isoforms of ImpA, however they were all of the same molecular weight (Ball *et al.*, 2016). The calculated molecular weights of all of the proteins were generally well matched with their position on the gel, whereas the predicted pI failed to account for the isoforms in each case, as expected. This further highlights the need to consider all proteoforms when investigating secreted virulence factors.

The protein sequence coverage obtained for the majority of the spots was very good compared to previous studies, which allowed greater confidence in the matches. The lower coverage rates achieved in less abundant spots was to be expected due to fewer peptides sequenced. It is likely that the number of spots analysed here do not encompass all of the proteoforms of each protein as less abundant proteoforms may be below the limit of detection for the stain and some may have vastly discrepant molecular weights or pI. Those proteoforms analysed, however, are clearly the most abundant and therefore probably the most biologically relevant.

#### **4.2.1 Hypothetical reasoning for two ImpA charge trains**

As shown in Figure 4.2 and Table 4.1, the protein ImpA appears to form two distinct charge trains at discrepant molecular weights; this observation was consistent in a range of growth environments (Chapter 3). This has not been reported before and the underlying molecular basis for it is not yet understood. It is tempting to suggest that the lower charge train (B) is a product of proteolytic processing of the upper charge train (A), as is the case for several other secreted proteins. However, upon comparison of the peptide sequence coverage obtained during mass spectrometry analysis, there does not appear to be a clear difference in the amino acid composition between the charge trains, and thus no evidence of processing at either the N- or C- terminus of the protein, beyond signal peptide cleavage (Figure 4.3). A hypothesis therefore arises that a sub-set of the ImpA proteins undergo a PTM that causes a shift in their molecular weight, for example glycosylation. Complex modifications such as glycosylation are known to mediate bacterial adhesion and virulence, as well as enhance protein



stability and functioning (André *et al.*, 2017). Alternatively, this protein may indeed undergo additional cleavage at the N-terminus following signal peptide release, as is that case for other secreted proteins, that was not elucidated by comparison of peptide coverage. Clearly, the lower charge train encompasses the more abundant proteoforms, bringing into question what the purpose of such modification may be; this would be an interesting avenue for future study.

A	Sequence coverage – A5					B	Sequence coverage – B5				
1	MSLSTTAFFS	LQGENMSRSP	IPHRALLAG	FCLAGALSAQ	AATQEEILDA	1	MSLSTTAFFS	LQGENMSRSP	IPHRALLAG	FCLAGALSAQ	AATQEEILDA
51	ALVSGDSSQL	TDSHLVALRL	QQQVERIRQT	RTQLLDGLYQ	NLSQAYDPGA	51	ALVSGDSSQL	TDSHLVALRL	QQQVERIRQT	RTQLLDGLYQ	NLSQAYDPGA
101	ASMNVLPANP	DNTLPFLIGD	KGRVLASLSL	EAGGRGLAYG	TNVLTLQSGT	101	ASMNVLPANP	DNTLPFLIGD	KGRVLASLSL	EAGGRGLAYG	TNVLTLQSGT
151	NAAHAPLLKR	AVQWLNVGDP	GAATKDFKV	SVVGVDKTA	LNGLKSAGLQ	151	NAAHAPLLKR	AVQWLNVGDP	GAATKDFKV	SVVGVDKTA	LNGLKSAGLQ
201	PADAACNALT	DASCASTSKL	LVLGNASAA	SLSATVRARL	QAGLPILFVH	201	PADAACNALT	DASCASTSKL	LVLGNASAA	SLSATVRARL	QAGLPILFVH
251	TNGWNQSSSTG	QQILAGLGLQ	EGPYGGNYWD	KDRVPSRTR	TRSVELGGAY	251	TNGWNQSSSTG	QQILAGLGLQ	EGPYGGNYWD	KDRVPSRTR	TRSVELGGAY
301	GQDPALVQQI	VDGSWRTDYD	WSKCTSYVGR	TTCDVPEGLS	DFSKRVDVLK	301	GQDPALVQQI	VDGSWRTDYD	WSKCTSYVGR	TTCDVPEGLS	DFSKRVDVLK
351	GALDAYNQKA	QNLFALPGTT	SLRLWLLWAD	AVRQNIIRYPM	DKAADTARFQ	351	GALDAYNQKA	QNLFALPGTT	SLRLWLLWAD	AVRQNIIRYPM	DKAADTARFQ
401	ETTFVADAIVG	YVREAGAAQK	ELGSYAGQRQ	QSMFVSGSEE	TLTLTLPSTAQ	401	ETTFVADAIVG	YVREAGAAQK	ELGSYAGQRQ	QSMFVSGSEE	TLTLTLPSTAQ
451	GFTAIGRMAA	PGKRLSIRIE	DAGQASLAVG	LNTQRIGSTR	LWNTRQYDRP	451	GFTAIGRMAA	PGKRLSIRIE	DAGQASLAVG	LNTQRIGSTR	LWNTRQYDRP
501	RFLKSPDIKL	QANQSVLVS	PYGGLQLVY	SGATPGQTVT	VKVTGAASQP	501	RFLKSPDIKL	QANQSVLVS	PYGGLQLVY	SGATPGQTVT	VKVTGAASQP
551	FLDIQFGEDS	SQAIADFIQA	LDADKADWLE	IRSGSVEVHA	KVEKVRGSID	551	FLDIQFGEDS	SQAIADFIQA	LDADKADWLE	IRSGSVEVHA	KVEKVRGSID
601	KDYGGDVQRF	IRELNEVFID	DAYTLAGFAI	PNQARTPAIQ	QECARGWDC	601	KDYGGDVQRF	IRELNEVFID	DAYTLAGFAI	PNQARTPAIQ	QECARGWDC
651	DSETLHKLP	TQHINVDQYA	QCGGCGSGNP	YDQTWGLNPR	GWGESHELGH	651	DSETLHKLP	TQHINVDQYA	QCGGCGSGNP	YDQTWGLNPR	GWGESHELGH
701	NLQVNRKLVY	GGRSGEISNQ	IFPLHKDWRV	LREFGQNLDD	TRVNYRNAYN	701	NLQVNRKLVY	GGRSGEISNQ	IFPLHKDWRV	LREFGQNLDD	TRVNYRNAYN
751	LIVAGRAEAD	PLAGVYKRLW	EDPGTYALNG	ERMAFYTQWV	HYWADLKNDP	751	LIVAGRAEAD	PLAGVYKRLW	EDPGTYALNG	ERMAFYTQWV	HYWADLKNDP
801	LQGWDIWILL	YLHQRQVDKS	DWDANKAALG	YGTYAQRPGN	SGDASSTDGN	801	LQGWDIWILL	YLHQRQVDKS	DWDANKAALG	YGTYAQRPGN	SGDASSTDGN
851	DNLLGLSLWL	TQRDQRPTFA	LWGIRTSAAA	QAQVAAYGFA	EQPAFFYANN	851	DNLLGLSLWL	TQRDQRPTFA	LWGIRTSAAA	QAQVAAYGFA	EQPAFFYANN
901	RTNEYSTVKL	LDMSQGSPAW	PPF			901	RTNEYSTVKL	LDMSQGSPAW	PPF		

**Figure 4.3 Sequence coverage of ImpA**

Sequence coverage of the full ImpA protein sequence obtained during LC-MS/MS analysis of spots A5 (A) and B5 (B); red highlight indicates MS/MS peptide coverage

### 4.3 Rationale for the selection of PTM to investigate

Changes to a protein's pI can be induced by several different post-translational modifications, including truncation through proteolytic processing. By cleaving a section of a protein, the ratio of charged to uncharged residues is likely to be disrupted, causing a shift in the overall accumulative protein charge. This was ruled out as the cause of the laterally distributed charge trains as an appreciable change in the protein's molecular weight would be expected. Instead, a range of small chemical PTMs were considered (Table 4.2). Investigation of deamidation of asparagine and glutamine would have been the obvious choice due to its introduction of a negative charge whilst inducing only minor mass alterations. However, profuse deamidation

artifacts during the sample preparation for mass spectrometry are a major problem for determining true deamidation rates (Hao *et al.*, 2017). In addition, whilst potentially still biologically relevant, the spontaneous non-enzymatic nature of deamidation does not explain the consistency and likely targeted phenomenon of secreted protein charge variation.

**Table 4.2 Mass and pI changes from typical PTMs**

( - indicates negative shift, 0 indicates no change, + indicates positive shift. \*Atypical methylation of E/D causes a positive pI shift)

Modification	Mass Change (Da)	$\Delta pI$
Deamidation	0.98402	-
Methylation	14.01565	0 (+)*
Hydroxylation	15.99491	0
Di-methylation	28.0313	0
Acetylation	42.01057	-
Tri-methylation	42.04695	0
Carboxylation	43.98983	0
Phosphorylation	79.5682	-

Acetylation of lysine (K-ac) was therefore selected as a key PTM of interest as it induces only a minor mass change (+42.01 Da) whilst removing the side chain positive charge. Lysine acetylation involves the addition of an acetyl group ( $\text{CH}_3\text{CO}$ ) to the side chain  $\epsilon$ -amine (Figure 1.6). In bacteria this occurs post-translationally through the action of acetyltransferases and is reversed by deacetylases, although non-enzymatic acetylation can also occur (D. G. Christensen *et al.*, 2019). Both pathways require an

acetyl donor, which is most commonly acetyl-CoA or acetyl-phosphate, to a lesser extent.

Each charge train was therefore investigated to determine differences in the acetylation status of the protein spots contained within it. The following is a deep analysis of the K-ac sites for each protein, with respect to localisation, evolutionary conservation, sequence motifs and secondary structure preference.

## **4.4 Acetylation of ImpA**

### **4.4.1 Background of ImpA**

Spots A1-A5 and B1-B6 were all identified to be proteoforms of the secreted protein ImpA. Until relatively recently, the function of this protein was unknown. In 2012, Bardoel *et al.* determined that it functions as an immunomodulating zinc metalloprotease (Bardoel *et al.*, 2012). During infection, host leukocyte homing and neutrophil extravasation occurs through the interaction of P-selectin glycoprotein ligand-1 (PSGL-1) on leukocytes with P-selectin expressed on endothelial cells. ImpA cleaves PSGL-1, as well as CD43 and CD44 also expressed on leukocytes, to prevent the recruitment of phagocytic immune cells to the site of infection. It has further been confirmed that O-linked glycosylation of substrates is a key determinant of ImpA substrate specificity (Noach *et al.*, 2017). Biochemically, ImpA is a relatively large protein of 923 residues encoding a signal peptide and an 882 residue mature protein. The secreted mature protein requires zinc as a cofactor for proper functioning.

### **4.4.2 Differences in ImpA proteoform abundance**

From the 2D gel it is possible to determine differences in relative proteoform abundance by measuring the intensity of the spots with each charge train. The intensity, and therefore relative abundance, of each spot as a percentage of the whole charge train was measured using image analysis software ImageJ; the tabulated percentages are the average of three biological replicates (Figure 4.4; Table 4.3).



**Figure 4.4 Close up image of charge trains A and B**

See Figure 4.2 for the full 2D gel image

**Table 4.3 Spot intensities of charge trains A and B**

Numbers represent the intensity of each spot as a percentage of the whole charge train (average of three biological replicates  $\pm$  standard deviation for train B) as determined by analysis with ImageJ software

Spot		1	2	3	4	5	6
% intensity	A	12.1	20.1	19.4	26.6	21.8	-
	B	8.6 $\pm$ 1.4	16.2 $\pm$ 1.3	20.5 $\pm$ 2.9	28.2 $\pm$ 2.9	22.6 $\pm$ 0.3	4 $\pm$ 1.9

In both the upper and lower ImpA charge trains, the fourth proteoform (A4 and B4) is the most abundant. The charge trains assume a slight skew to the alkaline pH range, with increasing abundance as the pI increases. The spots at the most extreme ends of the charge train were the least abundant. This therefore suggests that A4 and B4 are the “main” proteoforms produced by the bacterium. Standard deviations were not calculated for charge train A due to reduced abundance and staining in replicates.

#### 4.4.3 Prediction of lysine acetylation

Prior to PTM analysis, *in silico* prediction of lysine acetylation was carried out. Several online tools specialising in acetylation prediction have been developed, however only two were found to still be functional. Ensemble-PAIL predicts acetylation sites using a machine-learning algorithm. One drawback for this algorithm is that the training data sets used for its development were based on eukaryotic proteins (Xu *et al.*, 2010). Using this platform 12 acetylation sites were predicted, however a lack of correlation between the lysine position stated and actual lysine positions within the primary sequence prevented these results from being fully comprehended.

Alternatively, and more relevantly, a prokaryote specific algorithm was used called ProAcePred (Chen *et al.*, 2018). This tool utilises a combination of sequence-based, physicochemical and evolutionary information in its predictions. Using the *E. coli* training dataset, 18 lysine residues were predicted to be acetylated out of a total of 31 available (Table 4.4). This suggested that a high level of acetylation was likely and may account for the different proteoforms seen within the charge trains.

**Table 4.4 ProAcePred prediction of K-ac in ImpA**

The amino acid sequence of ImpA was submitted to ProAcePred for prediction of lysine acetylation sites; the predicted sites are highlighted in blue

Protein name	Position of site	Flanking residues	SVM Probability
Q9I5W4	121	DNTLPFLIGD- <b>K</b> -GRVLASLSLE	0.5
Q9I5W4	176	VNGDPGAATA- <b>K</b> -DFKVSVVGVVD	0.64907
Q9I5W4	179	DPGAATAKDF- <b>K</b> -VSVVGVVDKTA	0.79518
Q9I5W4	187	DFKVSVVGVVD- <b>K</b> -TAALNGLKSA	0.57189
Q9I5W4	195	VDKTAALNGL- <b>K</b> -SAGLQPADAA	0.54023
Q9I5W4	281	EGPYGGNYWD- <b>K</b> -DRVPSRTRT	0.83426
Q9I5W4	323	GSWRTDYDWS- <b>K</b> -CTSYYVGRITC	0.70808
Q9I5W4	344	DDVPGLSDFS- <b>K</b> -RVDVLKGALD	0.78956
Q9I5W4	350	SDFSKRVDVL- <b>K</b> -GALDAYNQKA	0.60754
Q9I5W4	359	LKGALDAYNQ- <b>K</b> -AQNLFALPGT	0.60682
Q9I5W4	463	TAIGRMAAPG- <b>K</b> -RLSIRIEDAG	0.5
Q9I5W4	504	TRQYDRPRFL- <b>K</b> -SPDIKLQANQ	0.59015
Q9I5W4	591	IRSGSVEVHA- <b>K</b> -VEKVRGSIDK	0.61883
Q9I5W4	594	GSVEVHAKVE- <b>K</b> -VRGSIDKDYG	0.76394
Q9I5W4	601	KVEKVRGSID- <b>K</b> -DYGGDVQRFI	0.87084
Q9I5W4	819	LLYLHQRQVD- <b>K</b> -SDWDANKAAL	0.65505
Q9I5W4	826	QVDKSDWDAN- <b>K</b> -AALGYGTYAQ	0.5
Q9I5W4	909	NNRTNEYSTV- <b>K</b> -LLDMSQGSPA	0.60229

#### 4.4.4 MS/MS identification of ImpA acetylation

Following identification of the proteins in each spot, the mass spectrometry data was then probed for the presence of K-ac. The inclusion of lysine acetylation as a variable modification during MASCOT database searches allowed the identification of many modification sites. Only site identifications with a confidence score above the identity confidence threshold and an expect value below 0.05, as well as a valid distribution of daughter ions for site localisation, were included. The acetylated lysines found in both the upper (A) and lower (B) ImpA charge trains are listed in Table 4.5 and Table 4.6, respectively. Representative MS/MS spectra of several acetylated peptides are shown in Figure 4.5.

**Table 4.5 Acetylated lysines identified in spots A1 to A5**

**Bold** residues were found in more than one spot, *italicised* residues were also found in the corresponding spot in charge train B (see Table 4.6)

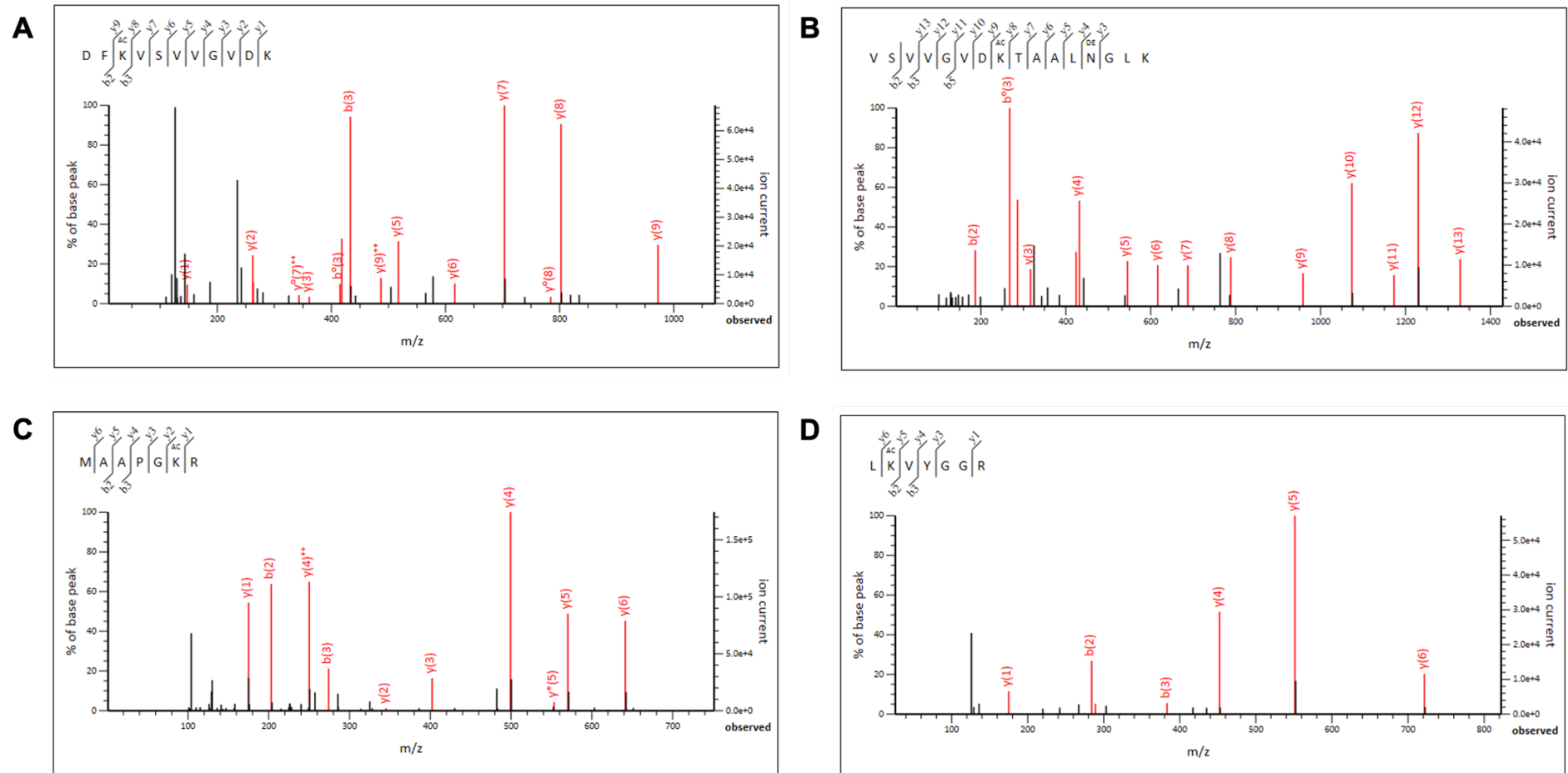
A1	A2	A3	A4	A5
	<b>K420</b>	<b>K420</b>	<i>K159</i>	<i>K463</i>
			K179	<b>K708</b>
			K187	
			<i>K195</i>	
			<i>K219</i>	
			K392	
			<b>K420</b>	
			<b>K463</b>	
			<i>K601</i>	
			<b>K708</b>	
			K726	

**Table 4.6 Acetylated lysines identified in spots B1 to B6**

**Bold** residues were found in more than one spot, *italicised* residues were also found in the corresponding spot in charge train A (see Table 4.5)

B1	B2	B3	B4	B5	B6
	<b><i>K420</i></b>	<b>K159</b>	<b><i>K159</i></b>	<b>K159</b>	<b>K187</b>
		<b>K195</b>	<b><i>K195</i></b>	K179	<b>K420</b>
		<b><i>K420</i></b>	<i>K219</i>	<b>K187</b>	
		<b>K463</b>	<b><i>K420</i></b>	<b>K195</b>	
		<b>K591</b>	<b><i>K463</i></b>	K350	
		<b>K708</b>	<b>K591</b>	K359	
			<i>K601</i>	K392	
				<b>K420</b>	
				<b><i>K463</i></b>	
				<b>K591</b>	
				<b><i>K708</i></b>	
				K726	
				K767	
				K819	
				K909	

It is clear from these results that, as predicted by the algorithms, this secreted virulence factor undergoes significant modification by acetylation. In fact, 10 out of the predicted 18 residues were found to be acetylated. Surprisingly, an increase in the number of K-ac was found from left to right across the charge train (with the exception of the spot with the highest pI; A5/B6). This was the opposite of what was expected as acetylation removes the positive charge and would therefore cause a decrease in the isoelectric point. As such, the opposite pattern of a decrease in acetylation from left to right was expected.



**Figure 4.5 Representative MS/MS spectra of ImpA acetyl-peptides**

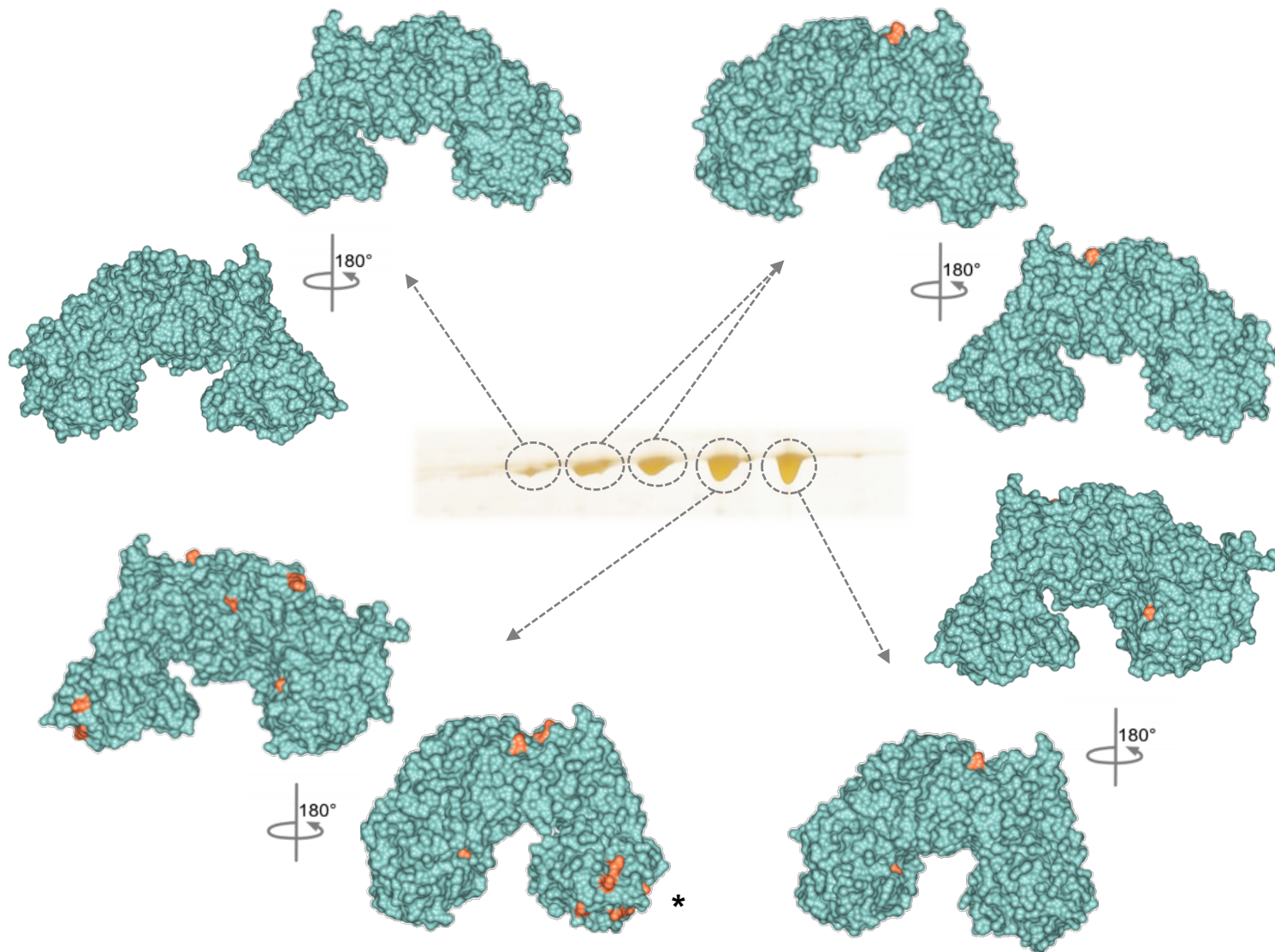
MS/MS spectra of ImpA peptides containing acetylated K179 (A), K187 (B), K463 (C) and K708 (D). Fragmentation of the peptide bond between residues during LC-MS/MS produces sequence specific fragment ions (denoted above by the b- and y-ions) with known mass/charge ( $m/z$ ) ratios. The presence and location of a PTM within a peptide, in this case acetylation, can be determined by interrogating the mass shift between ion peaks. An additional mass shift of +42.01 Da between residue ion peaks indicates the presence of an acetyl group on a lysine residue. The peptide b- and y-ions generated during LC-MS/MS and the location of the modified lysine within the peptide sequence (AC) are shown in the top right of each graph.



In Tables 4.5 and 4.6, residues which were found to be acetylated in more than one protein spot **within** a charge train are highlighted in bold; consistencies of parallel spots *between* the two ImpA charge trains are italicised (i.e., consistencies between A4 and B4). Differences in the number of K-ac found between spots is in part likely to be due to the differences in protein abundance and coverage within each sample (Table 4.1). For example, the peptide coverage of A1 was 38%, compared with 75% for A4. Furthermore, the emPAI value of a MS/MS sample gives an estimate of protein abundance based on the degree of protein coverage by the peptide matches. The emPAI values of spots B1 and B5 were 38.11 and 747.77 respectively, a significant difference. Fewer peptides sequenced in less abundant spots means that fewer acetylation events are likely to be uncovered. This is further supported by the fact that PTMs are typically sub-stoichiometric within a protein population, and therefore even less likely to be found in low abundant samples. It is important to note that out of the 17 K-ac identified, several lysines were found to be acetylated in multiple spots and in both charge trains. This therefore suggests that these residues, such as K420, are key targets for consistent modification.

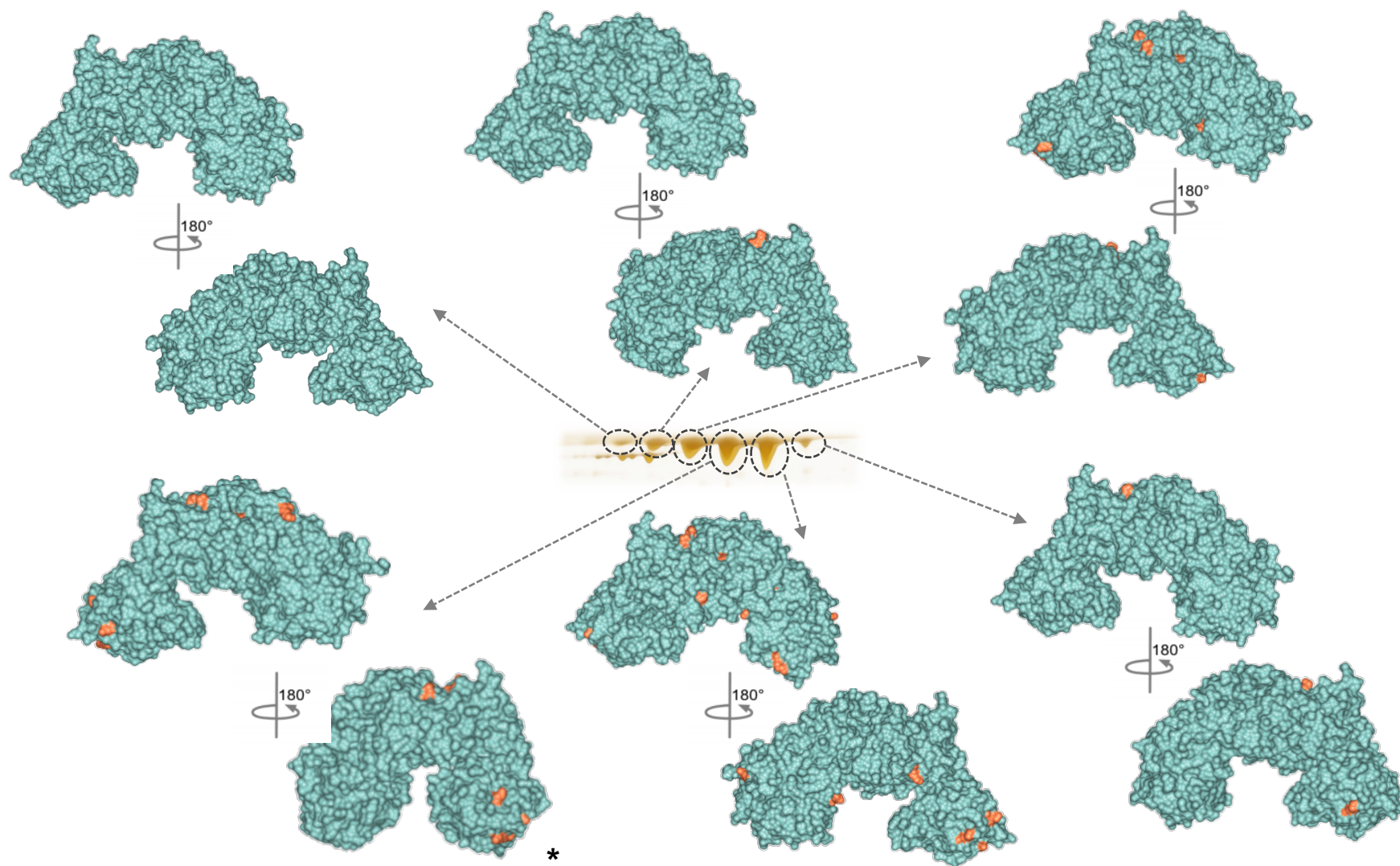
#### 4.4.5 Structural mapping of ImpA acetylation

The locations of the acetylated lysines were then mapped onto the 3D structure of the protein (PDB: 5KDV). It was noted during mapping that all but one (K726) of the modified residues are present on the surface of the protein; the models presented in Figure 4.6 (spots A1-A5) and Figure 4.7 (spots B1-B6) show the protein surface and the location of the modified lysines are highlighted in orange. This surface exposure is to be expected due to the hydrophilicity of lysine residues conferring preference at the protein surface. The distribution of acetylation sites across the protein surface appeared to be relatively even for the most modified spots (A4 and B5). In both charge trains, one of the “arms” on the protein does appear to be more heavily modified than the other (starred\* in the figures), however the reason for this is unclear.



**Figure 4.6 Structural mapping of acetylated lysines in spots A1-A5**

The crystal 3D crystal structure of ImpA (PDB: 5KDV) was analysed using CCP4mg software. The location of lysines identified as acetylated in the LC-MS/MS data of each spot from charge train A were mapped onto the surface of the protein structure; acetylated lysines are highlighted in orange and all other residues are coloured in teal. The models for each spot are identified by an arrow



**Figure 4.7 Structural mapping of acetylated lysines in spots B1-B6**

The crystal structure of ImpA (PDB: 5KDV) was analysed using CCP4mg software. The location of lysines identified as acetylated in the LC-MS/MS data of each spot from charge train B were mapped onto the surface of the protein structure; acetylated lysines are highlighted in orange and all other residues are coloured in teal. The models for each spot are identified by an arrow



#### 4.4.6 Sequence conservation of ImpA acetylated lysines

To check whether the modified lysines are conserved in other species, the protein sequence of ImpA was BLAST against the UniProtKB protein database. In their analysis, Bardoel et al found that ImpA is only present in *P. aeruginosa* where it is highly conserved and is absent in all other *Pseudomonas* species (Bardoel *et al.*, 2012). During the homology searching conducted here, one sequence in *P. fluorescens* showed 99.4% homology to ImpA, with all of the acetylated lysines identified being present. Aside from this, no sequences showed greater than 50% homology to *P. aeruginosa* ImpA, and the majority aligned to the peptidase M60 domain of the protein. Proteins from *Chitinimonas taiwanensis*, *Rhodofera sediminis*, *Curvibacter* sp. AEP1-3, *Undibacterium piscinae*, *Vibrio mytili*, *Gallaecimonas xiamenensis* and *Shewanella algae* were aligned to the ImpA sequence; the full alignment and BLAST outline can be found in Appendix A. The majority of these species were originally isolated from aquatic environments. Many of the lysines did not show strong conservation, including K420, however K159, K350, K392, K463, K708, K726 and K909 of ImpA were found in the corresponding aligned lysines from several of the species (Figure 4.8). Where these residues were not conserved, a substitution for another charged, modifiable residue appears to be common. This suggests that these residues, and their potential acetylation, may be important for the functioning of these proteins.

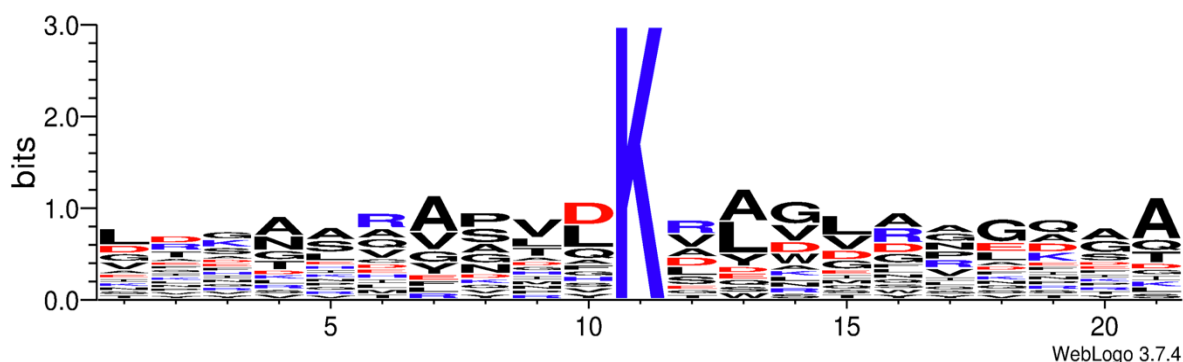
<i>P. aeruginosa</i>	127	SLSTLEAGGRGLAYGTNVLTOLSGT---	NAAHAPLIKRAVQWLVN	GP	GAATAKDFKVS	V
<i>C. taiwanensis</i>	127	SLSTLAQGGRSAGYGINVLDQFRKA---	NEAHMPAFQRLVSWLVV	GP	GATLPASIN	V
<i>R. sediminis</i>	139	SISVAAGGRSAGFGVOVLRFNGNQ---	LTAYRPAFKRLLAWLVR	GD	AAATLPATL	N
<i>C. AEP1-3</i>	139	SISVAGGRSAGFGVOVLSKFNANE---	LVAYRPAFKRLLAWLVR	GD	AAATLPATL	N
<i>U. piscinae</i>	129	SISTASNGRSAGYGANVLEOFRNNN---	NLNHAPAFKRLLSWLVL	GD	ATLALPASL	E
<i>V. mytili</i>	10	TIQGHDCVRAAAGGNGFLIQDNDNSYS	LDAYQKDEKRVLSWL	VS	GDANAPIPST	L
<i>G. xiamenensis</i>	202	IAGEVGGRYGGITGTHLFARFOAGE---	LLAMAPAEQILLAWLMQD---		ADLSQPHTL	A
<i>S. algae</i>	219	VAGEQHTGRFAAFGTHLFARFHAGE---	LTAMEPANDILLAWLLN	RAQA	AEELQQPITV	S
<i>P. aeruginosa</i>	344	KRVDVTKKALDAYNOKAQNLFALPGTTS	LRLLWLLWADAVR	ONIRYP	PMIKK	ADT--ARFOE
<i>C. taiwanensis</i>	347	TPAEQVRQINSFNAGRNLFSTPNTTL	LRLLTLWADTT	TRKDIT	YPLIK	KIKOP--AAFQR
<i>R. sediminis</i>	359	SPVDAIRSQIDVFENRAGONLFTADTT	VLRYLVWADVVR	RQLRYP	PMIK	KASTP--AAFEK
<i>C. AEP1-3</i>	359	SPVDAIRSQIDVFENKAGONLFTADTE	VLRYLVWADVVR	RQLRYP	PMIK	KASTP--AAFOK
<i>U. piscinae</i>	349	TPIESIRQIDNFTSRGNLFATPNTTSL	RLALLWADVQR	KQIVYP	VLK	ISKP--VEFOK
<i>V. mytili</i>	248	EPVEAIRQVFDSTYKSGTPLFSAPGNE	FYRLANIWADLN	RRSMSY	PMIK	KEEDMADYRFQK
<i>G. xiamenensis</i>	420	SGARSIKNALNRDLSQGLDLFAEPNLT	LLKLFVLLGDSY	RAQIHY	PMIK	KESTD-PNQFLA
<i>S. algae</i>	440	AGARAIKNSLNQLDSRGOMLFGNEGRR	LLKLFVLLGDL	YRADIAY	PMIK	KDTP-QGRFLA
<i>P. aeruginosa</i>	695	SHELGHNLQVNRIKVIYGGRSGEISNQ	IFPLEKQW	RVLREF	QONLDD	-TRVNYRNAYNLI
<i>C. taiwanensis</i>	694	SHELGHNLQOGMIVYGGRSGEVSNN	IFPLEKNWR	VLR	RELEDDR	DR-GRINYS
<i>R. sediminis</i>	708	SHEVGHNOQKGMKVIYDDRSGEVSNN	IFPLEKQWR	MLSEL	GYNSGD	-NRVAYOSAF
<i>C. AEP1-3</i>	708	SHEVGHNOQKGMKVIYDDRSGEVSNN	IFPLEKQWR	MLSEM	GYNTGD	-TRVSYL
<i>U. piscinae</i>	696	SHELGHNLQOGLINVIYGGRSGEVSNN	IFPLEMWR	VLR	HEKNT	-YVA-NHVGYS
<i>V. mytili</i>	603	SHELGHNLQYQEFKVDGGISGEVSNN	IFPLEKQGR	LYQDF	GVDLGS	-NKVDYEATF
<i>G. xiamenensis</i>	763	SHEIGHNLQSRRIKVIYGGRSGEVSNN	IFPLEKQW	ORFHD	SGERIES	CDRQDPALTY
<i>S. algae</i>	783	SHEIGHNLQQRNRIKVIYGGRSGEVSNN	IFPLEKQW	ORFKD	SGERIAS	CDRQSPOTTF

Figure 4.8 Sequence alignment of ImpA

Sequence alignment of ImpA with homologous proteins from other species, as identified by BLAST analysis; outlined in red are the key residues found to be conserved in several species.

#### 4.4.7 ImpA acetylation motif analysis

To determine whether there is a consensus sequence that may be recognised by a responsible acetyltransferase enzyme, motif analysis was conducted using WebLogo (Crooks *et al.*, 2004). The ten residues up- and down-stream of the 17 acetylated lysines were aligned and the frequency of amino acids around the modification site was graphically depicted (Figure 4.9); amino acid frequencies are represented by the size of their symbols. No specific sequence contexts were determined, although alanine appeared slightly more prevalent at the -4, +2 and +10 positions. Furthermore, the -4, -3 and +2 positions were slightly more dominated by uncharged hydrophobic residues. Previous studies in PA14 and various other prokaryotes noted the presence of aspartic acid (D) at the position preceding the lysine, this can be seen here (Chen *et al.*, 2018; Gaviard, Broutin, *et al.*, 2018). An increased frequency of leucine (L) at the +2 position, as seen here, was also previously highlighted in a study of *Corynebacterium glutamicum*, although it was a signal for the alternative acyl-PTM, succinylation (Mizuno *et al.*, 2016).

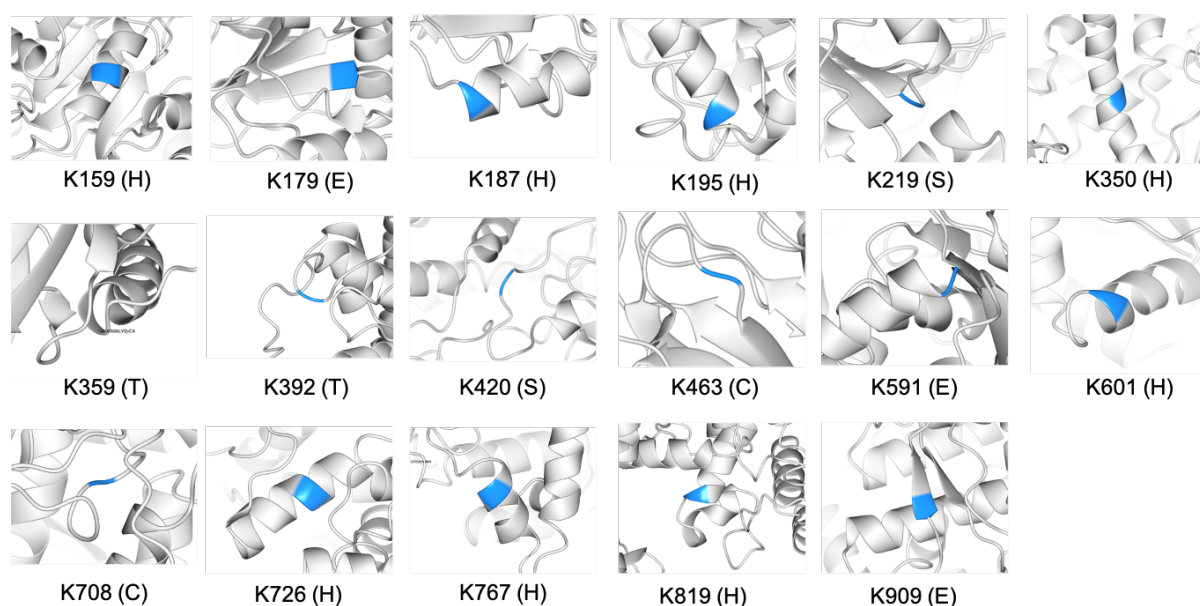


**Figure 4.9 ImpA acetylation motif analysis**

The frequency of residues surrounding every acetylated lysine ( $\pm 10$  residues) were analysed using WebLogo; the size of the symbol correlates to its frequency at that position. The colours represent the charge of the amino acid, positive (blue), negative (red), or uncharged (black)

#### 4.4.8 ImpA K-ac secondary structure preference

The secondary structure around each K-ac site was also investigated to determine whether preferential targeting may be due to the protein conformation instead of the primary sequence. The PDB file was run through DSSP software for standardised secondary structure assignment (Kabsch and Sander, 1983; Touw *et al.*, 2015) and the residues were also highlighted in the 3D structure (Figure 4.10). In agreement with previous findings, K-ac sites were preferentially found in ordered structures, in this case alpha helices at almost 50% (Lu *et al.*, 2011). However there does not appear to be a precise conserved secondary structure element that specifically confers amenability to lysine acetylation in this protein.



**Figure 4.10 Secondary structure at each K-ac site in ImpA**

The secondary structure surrounding the acetylated lysines in ImpA. The K-ac sites are highlighted in blue and the letter in brackets following the residue number indicates the DSSP structure assignment: H = alpha helix, E = beta sheet, S = bend, T = turn, C = coil

## 4.5 Acetylation of LasB

### 4.5.1 Background of LasB

Spots D1-D4 were all identified to be proteoforms of the secreted protein LasB. This protease is a key player in the *P. aeruginosa* infection strategy due to its extensive substrate specificity. LasB, or elastase, is able to degrade a myriad of host proteins, including elastin and collagen found abundantly in lung and vascular tissue, causing significant damage. It also subverts the host immune response by cleaving immunoglobulins, components of the complement system, surfactant proteins and proinflammatory cytokines (Hoge *et al.*, 2010; Sun *et al.*, 2020).

Endogenously, LasB is central to the activation of other important virulence factors. As outlined in Figure 1.4, an intricate extracellular cleavage cascade occurs following LasB secretion that results in that activation of LasA, LasD and Protease IV (Braun *et al.*, 1998; Oh *et al.*, 2017; Li and Lee, 2019). As an additional measure of immune evasion, LasB also degrades extracellular flagellin to prevent recognition by immune mediators (Casilag *et al.*, 2015). Intracellularly it processes nucleoside diphosphate kinase (NDK) to stimulate alginate synthesis and cause upregulation of the biofilm phenotype (Kamath *et al.*, 1998). Bioinformatic analyses have even suggested that 63% of the PAO1 proteome may be potential LasB (and LasA) substrates (Arevalo-Ferro *et al.*, 2003). LasB is therefore important to both the acute and chronic phases of infection.

Biochemically, LasB is produced as a 498 residue pre-pro-protein before being secreted through the T2SS as a 301 residue mature protein. It requires zinc and calcium as co-factors and belongs to the M4 thermolysin family of peptidases (Olson and Ohman, 1992).

### 4.5.2 Differences in LasB proteoform abundance

As discussed previously, differences in the relative proteoform abundance were measured by quantifying the intensity of each protein spot within the charge train. The

average percentage intensity of each spot from three biological replicates is shown in Table 4.7 with the respective spots shown in Figure 4.11.



**Figure 4.11 Close up image of charge train D**

See Figure 4.2 for the full 2D gel image

**Table 4.7 Spot intensities of charge train D**

Numbers represent the intensity of each spot as a percentage of the whole charge train (average of three biological replicates  $\pm$  standard deviation) as determined by analysis with ImageJ software

Spot	1	2	3	4
% intensity	16.9 $\pm$ 5.3	34.2 $\pm$ 3.1	39.8 $\pm$ 1.7	9.1 $\pm$ 2.0

The proteoforms in spots D2 and D3 were the most abundant, with the less abundant forms taking the positions at the highest and lowest pI. This follows a similar pattern to that seen with ImpA with a slight skew to the more alkaline range. Intriguingly, 2D gel separation of the PA14 secretome showed only three proteoforms, with the most abundant at the central position (Gaviard *et al.*, 2019). This suggests that PTM is likely to be slightly different between strains, resulting in minor differences in extracellular protein isoform populations. This discrepancy may also be a factor of stain sensitivity as a greater number of LasB isoforms were evident in the 2D-DiGE labelled gels presented in Chapter 3.

### 4.5.3 Prediction of LasB lysine acetylation

*In silico* prediction of LasB lysine acetylation was subsequently carried out. The mature LasB protein encodes 11 lysine residues. Once again, there was a lack of concordance between the stated acetylated lysine positions and the actual residue number in the primary sequence when Ensemble-PAIL was used. However, a total of 5 acetylation sites were predicted in the mature secreted protein. Using the *E.coli*



trained ProAcePred algorithm, four sites were predicted to be acetylated (Table 4.8). Although above the suggested stringency setting of 0.5, the predictions were not as probabilistically strong as those for ImpA. It is unclear whether there was a predictive overlap between the two algorithms.

**Table 4.8 ProAcePred prediction of lysine acetylation in LasB**

The amino acid sequence of LasB was submitted to ProAcePred for prediction of lysine acetylation sites; the predicted sites are highlighted in blue

Protein name	Position of site	Flanking residues	SVM Probability
P14756	282	DAHFFGGVVF- <b>K</b> -LYRDWFGTSP	0.55943
P14756	378	GEAAEFYMRG- <b>K</b> -NDFLIGYDIK	0.60242
P14756	389	NDFLIGYDIK- <b>K</b> -GSGALRYMDQ	0.50962
P14756	443	LANSPGWDTR- <b>K</b> -AFEVFVDANR	0.545

#### 4.5.4 MS/MS identification of LasB acetylation

The LC-MS/MS data was submitted to MASCOT for the identification of lysine acetylated peptides; acetylation sites for each protein spot are listed in Table 4.9 and representative MS/MS spectra are shown in Figure 4.12. Any K-ac found in more than one spot are highlighted in bold.

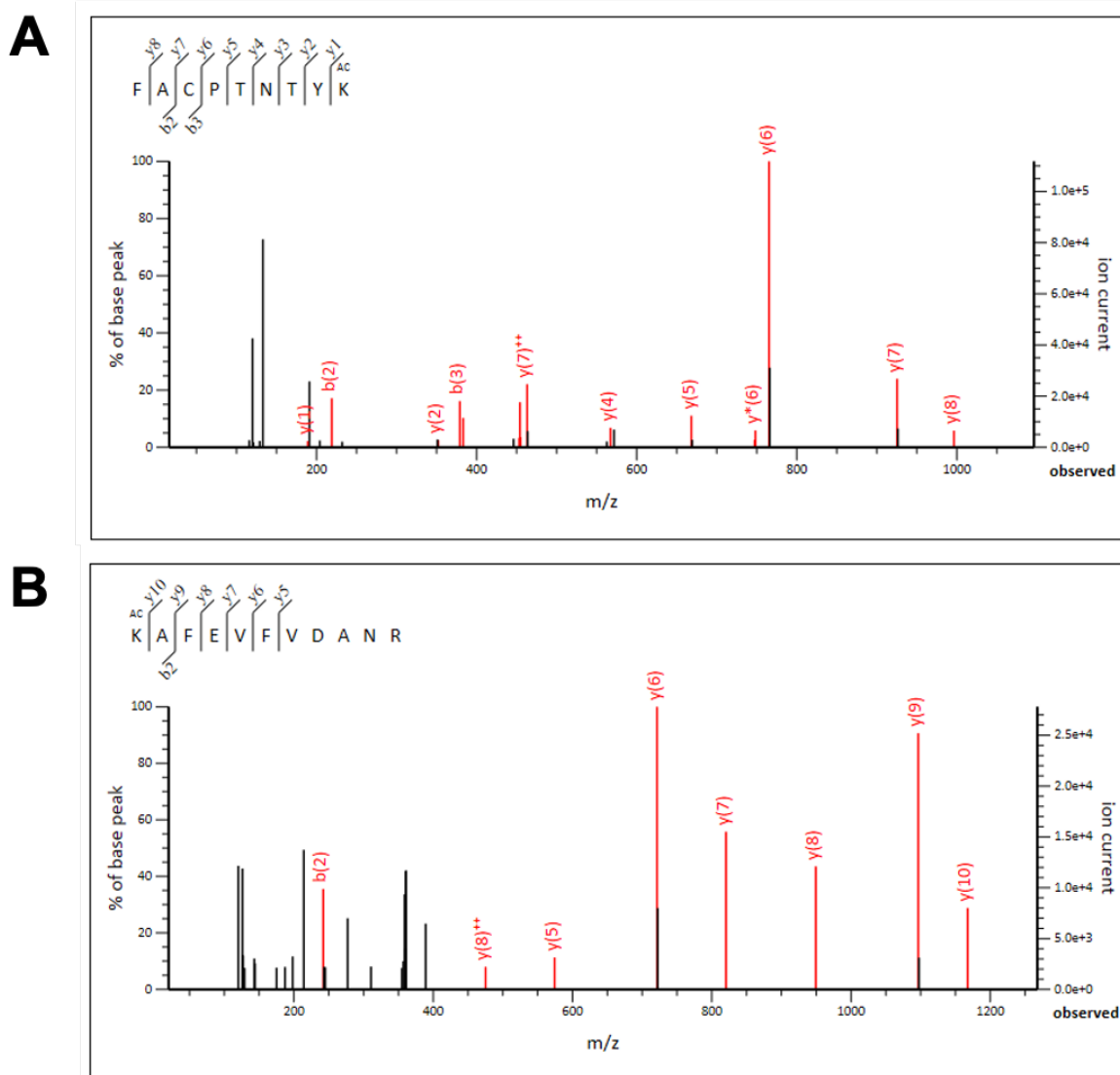
**Table 4.9 Acetylated lysines identified in spots D1 to D4.**

**Bold** residues were found in more than one spot.

D1	D2	D3	D4
<b>K211</b>	<b>K211</b>	<b>K443</b>	<b>K247</b>
<b>K247</b>	<b>K247</b>		<b>K261</b>
<b>K261</b>	<b>K261</b>		<b>K378</b>
<b>K378</b>	<b>K378</b>		
<b>K443</b>	K388		
	K389		

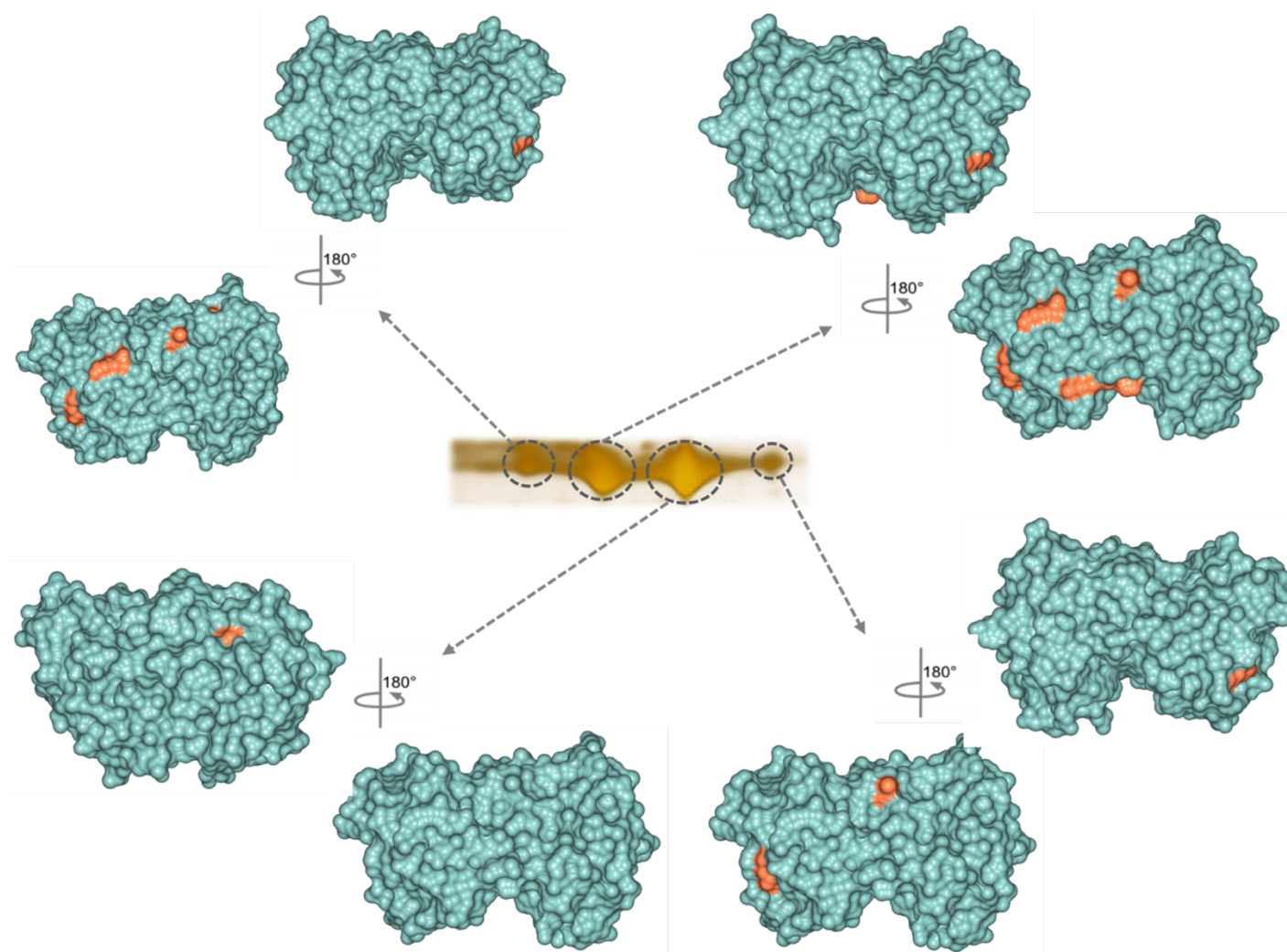
The results show that several acetylated lysines were identified in each of the spots analysed. In this case, there was considerable overlap in the K-ac residues identified between spots, therefore showing that these are important and consistent acetylation sites. Additionally, only one of the predicted sites (Table 4.8) was not realised in the data. Surprisingly, the most abundant spot (D3) had the least acetylation sites, however it must be noted that this spot was analysed on a different mass spectrometer from the other spots and as a result far fewer peptides were sequenced (as outlined in Table 4.1). It is likely therefore that many of the true sites in this spot were missed and are not truly absent. Once again it is likely that such differences in peptide abundance prevented a pattern of acetylation that could explain the charge train distribution phenomena being fully formulated. This is an intrinsic limitation of using bottom-up proteomic strategies for PTM identification as typical low peptide abundancies can prevent PTM identification.

After these results were obtained, a study in PA14 looking at PTM of LasB and CbpD (another secreted protein) was published. In agreement with my findings, the authors found acetylation of K211, K378, K388 and K443 of LasB (Gaviard *et al.*, 2019). This is an important validation of the results presented here and shows that these modifications occur in multiple *P. aeruginosa* strains. It is unclear whether there are differences in the scope of LasB acetylation between these strains or just experimental depth obtained. Nevertheless, the results presented here provide an extension of the number of known K-ac sites in LasB.



**Figure 4.12 Representative MS/MS spectra of LasB acetyl-peptides**

MS/MS spectra of LasB peptides containing acetylated K261 (A) and K443 (B). Fragmentation of the peptide bond between residues during LC-MS/MS produces sequence specific fragment ions (denoted above by the b- and y-ions) with known mass/charge (m/z) ratios. The presence and location of a PTM within a peptide, in this case acetylation, can be determined by interrogating the mass shift between ion peaks. An additional mass shift of +42.01 Da between residue ion peaks indicates the presence of an acetyl group on a lysine residue. The peptide b- and y-ions generated during LC-MS/MS and the location of the modified lysine within the peptide sequence (AC) are shown in the top right of each graph.



**Figure 4.13 Structural mapping of acetylated lysines in spots D1-D4**

The 3D crystal structure of LasB (PDB: 1EZM) was analysed using CCP4mg software. The location of lysines identified as acetylated in the LC-MS/MS data of each spot from charge train D were mapped onto the surface of the protein structure; acetylated lysines are highlighted in orange and all other residues are coloured in teal. The models for each spot are identified by an arrow

#### 4.5.5 Structural mapping of LasB acetylation

The locations of the acetylated lysines were mapped onto the 3D structure of the protein (Figure 4.13). All of the K-ac mapped to the surface of the protein, as before. There also appeared to be a preference for acetylation of one side of the protein compared to the other. This may be trivial or alternatively may infer a function of this PTM on the protein. For example, LasB is folded in the periplasm prior to being exported out of the cell. Acetylation at the protein surface and preferentially on one side could therefore be part of a secretion signal to drive interaction with the T2SS for export. This is purely speculative however, and no data is currently at hand to support this theory.

#### 4.5.6 Sequence conservation of LasB acetylated lysines

The conservation of the lysine acetylation sites in proteins homologous to LasB produced by other species was investigated. As expected, the highest hits returned following BLAST analysis mapped to homologous proteins produced by other *Pseudomonas aeruginosa* strains, with sequence identities upwards of 98%. Intriguingly, the UniProt database notes a natural variation in strain 569B where K211 is substituted for isoleucine. Beyond this, a range of bacterial species producing proteins with identities around 58-63% of LasB were aligned with its amino acid sequence. These strains included *Pseudomonas indica*, *Chromobacterium violaceum*, *Xenophilus* sp., *Pseudomonas alcaligenes*, *Burkholderiales bacterium* PBB2 and *Chitinimonas taiwanensis*; the BLAST outline and full alignment can be found in Appendix A. The strongest conservation was found for K261, K389 and K443, which were consistent in all sequences aligned; these residues were also found to be conserved in *Vibrio cholerae* during a later search (data not shown). These residues, and their PTM, may therefore be particularly important for protein function in many different species. Beltrao et al noted that residues tend to be more conserved when they have either a known regulator (i.e. modifying/de-modifying enzyme) or functional role (Beltrao *et al.*, 2013).

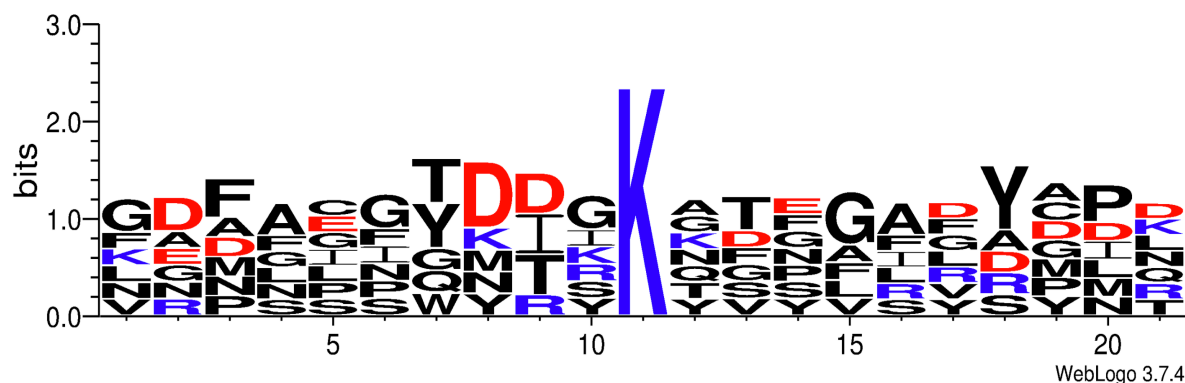
<i>P. aeruginosa</i>	163	YNVSYLI-PGEGLSRPHFVIDAKTGEVLDQWEGLAHA-EAGGPGGNQKICKYTYGSDYGP
<i>P. indica</i>	169	YVVSFYFL-PD-VPSRPHFLIDANDGSVLQQWEGLAHRTEASGPGGNEKTKGRVLYGSDYGP
<i>C. violaceum</i>	159	YLVSFVV-DGKEPSRPHLLIDANSQVLKQWEGLNHA-EAGGPGGNAKTQYTYGKDYG
<i>X. sp</i>	165	YLVSFVV-EGKEPSRPHLLIDANSQVLKQWEGLNHA-EAGGPGGNAKTQYTYGKDYG
<i>P. alcaligenes</i>	156	YVVSFFL-GGAKPSRPHFMTANSGETLQWEGLAHA-EASGPGGNSKTQYLYGTQYGP
<i>B. bacterium PBB2</i>	167	YLVSFLLTSHGDEPSRPHLLIDANSQVLEQWEGLNHA-NATGPGGNAKTCKYIEYGTTYGP
<i>C. taiwanensis</i>	166	YLVSFVIEDADHPSRPHFFIDANTGAVLERWEGLAHR-DATGPGGNTKTQYIEYGTDFGP
<i>P. aeruginosa</i>	221	LIVNDRCEMDDGNVITVDMNSSDTSKTPPFRFACPTNTYKQVNGAYSPLNDAHFFGGVV
<i>P. indica</i>	227	LVVTNDCEMNSGDVLTVDLHSTSNKRTPFQFACSYNDYKQINGAYSPLNDAHFFGNV
<i>C. violaceum</i>	217	LIVTSDCMDSGNVATVNLNGGT--SGTTPYKFFACPTNTYKQINGAYSPLNDAHFFGNV
<i>X. sp</i>	223	LIVTSDCMDSGNVATVNLNGGT--SGSAPYKFFACPTNTYKQINGAYSPLNDAHFFGNV
<i>P. alcaligenes</i>	214	LIVSRDCOMNSGDVITVNLNORYDNSSVTPFRFACPYNDYKQINGAYSPLNDAHFFGNV
<i>B. bacterium PBB2</i>	226	LVVTADCOMNSGNVITVDLKGGT--SGSTPFFKFTCPRNEYKLTNGAYSPLNDAHFFGNV
<i>C. taiwanensis</i>	225	LVVTDDCKMNSGNVITVNLNGGT--SGSTPFFQFTCPRNTYKQINGAYSPLNDAHFFGNV
<i>P. aeruginosa</i>	339	VSHGFTEQNSGLIYRGQSGGMNEAFSDMAGEAAEFYMRGKNDFLIGYDIKKGSGALRYMD
<i>P. indica</i>	345	VSHGFTEQNSNLOYYGQSGGMNEAFSDMAGEAAEFYMRGKNDWKVGFDIKKGTGALRYMD
<i>C. violaceum</i>	333	VSHGFTEQNSGLVYSGQSGGINEAFSDMAGEAAEFYMRGKNDVLVGAEIFKKKTGALRYFA
<i>X. sp</i>	339	VSHGFTEQNSGLVYSGQSGGINEAFSDMAGEAAEFYMRGKNDWLVGAEIFFKGSALRYFA
<i>P. alcaligenes</i>	334	VSHGFTEINSGLVYEGQSGGINEAFSDMAGEAAEFYMRGKNDWKVGADIFFKGNALRYMD
<i>B. bacterium PBB2</i>	342	VSHGFTEQNSGLVYSNOSGGMNEAFSDMAGEAAEFYMRGKNDWLVGSEIFFKASALRYFA
<i>C. taiwanensis</i>	341	VSHGFTEQNSGLVYSAOSGGMNEAFSDMAGEAAEFYMRGKNDWMVGAEIFKRSGALRYMD
<i>P. aeruginosa</i>	399	QPSRDGRSIDNASQYNGIDVHSSSGVYNRAFYLLANSFGWDTFKAFEVFDANRYYWTA
<i>P. indica</i>	405	DPTKDGRSIGHAADYRSGMNVHYSSGVYNKAFYLLANTAGWDTFKAFEVFDANRFYWTE
<i>C. violaceum</i>	393	DPTKDGRSIGNAKDYNGLDVHYSSGVYNKAFYLLATSEPNWNTFKAFEVFDANRLYWNA
<i>X. sp</i>	399	DPTRDGRSIGHAKDYNGLDVHYSSGVYNKAFYLLIATSAGWNTFKAFEVFDANRSLYWTA
<i>P. alcaligenes</i>	394	QPSRDGASIEHAQYHDGIDVHSSSGVYNRAFYLLSRTFGWNTFKAFEVFDANRFYWTE
<i>B. bacterium PBB2</i>	402	DPTRDGRSIDNASKYNGLDVHLSGGVYNKAFYLLATKPGWNTFKAFEVMVDANRLYWTA
<i>C. taiwanensis</i>	401	DPTRDGRSIGHASNYTSGMDVHYSSGVYNKAFYLLANKPGWNTFKAFEVMVDANRLHWTA

**Figure 4.14 Sequence alignment of LasB**

Sequence alignment of LasB with homologous proteins from other species, as identified by BLAST analysis; outlined in red are the acetyl-residues found to be conserved in several species

#### 4.5.7 LasB acetylation motif analysis

To determine whether there is a consensus sequence surrounding LasB acetylation sites, motif analysis was conducted. As before, the 10 residues up and downstream of the K-ac site were subjected to frequency mapping (Figure 4.15). The residues around each lysine appear to be quite variable with no clear preference for any amino acids at a specific site either up- or down-stream. This is not entirely unexpected as motif analysis generally uses large data sets comprising many protein species to determine clear consensus motifs. The data used here only encompasses the sequences for the seven LasB sites and is therefore not sufficient for patterns to become apparent. A signal motif for LasB acetylation can therefore not be elucidated from these results.



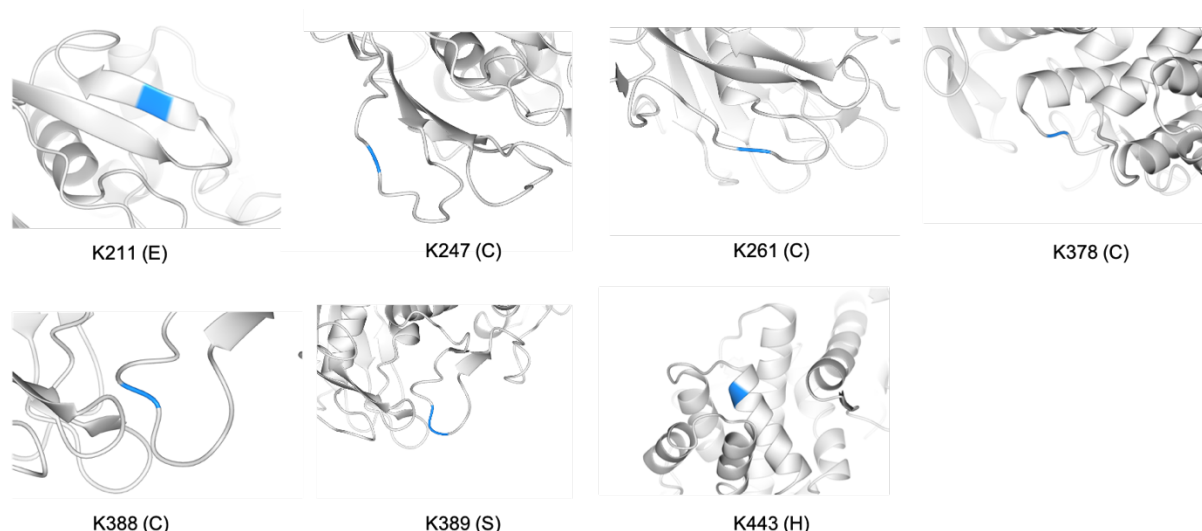
**Figure 4.15 LasB acetylation motif analysis**

The frequency of residues surrounding every acetylated lysine ( $\pm 10$  residues) were analysed using WebLogo; the size of the symbol correlates to its frequency at that position. The colours represent the charge of the amino acid, positive (blue), negative (red), or uncharged (black)

#### 4.5.8 LasB K-ac secondary structure preference

The secondary structure around each K-ac site was also investigated to determine whether PTM targeting may be due to the protein conformation instead of primary sequence. Through mapping of the acetylation sites onto the structures, it became clear that they mostly occur on less ordered and more flexible structures, such as coils and turns (as determined by DSSP analysis) (Figure 4.16). This is contrast to previous suggestions that K-ac sites are more common on ordered structures. Other studies have suggested that secondary structure at acetylation sites is highly species specific (Yang *et al.*, 2018). In comparison with the findings from ImpA where the K-ac sites were mostly located on alpha helices, it can be argued here that such specificity may actually go down to the protein level. Alternatively, the secondary structure may in fact be completely irrelevant to how the PTM is targeted or functions.





**Figure 4.16 Secondary structure at each K-ac site in LasB**

The secondary structure surrounding the acetylated lysines in LasB. The K-ac sites are highlighted in blue and the letter in brackets following the residue number indicates the DSSP structure assignment: H = alpha helix, E = beta sheet, S = bend, C = coil

## 4.6 Acetylation of PA0622

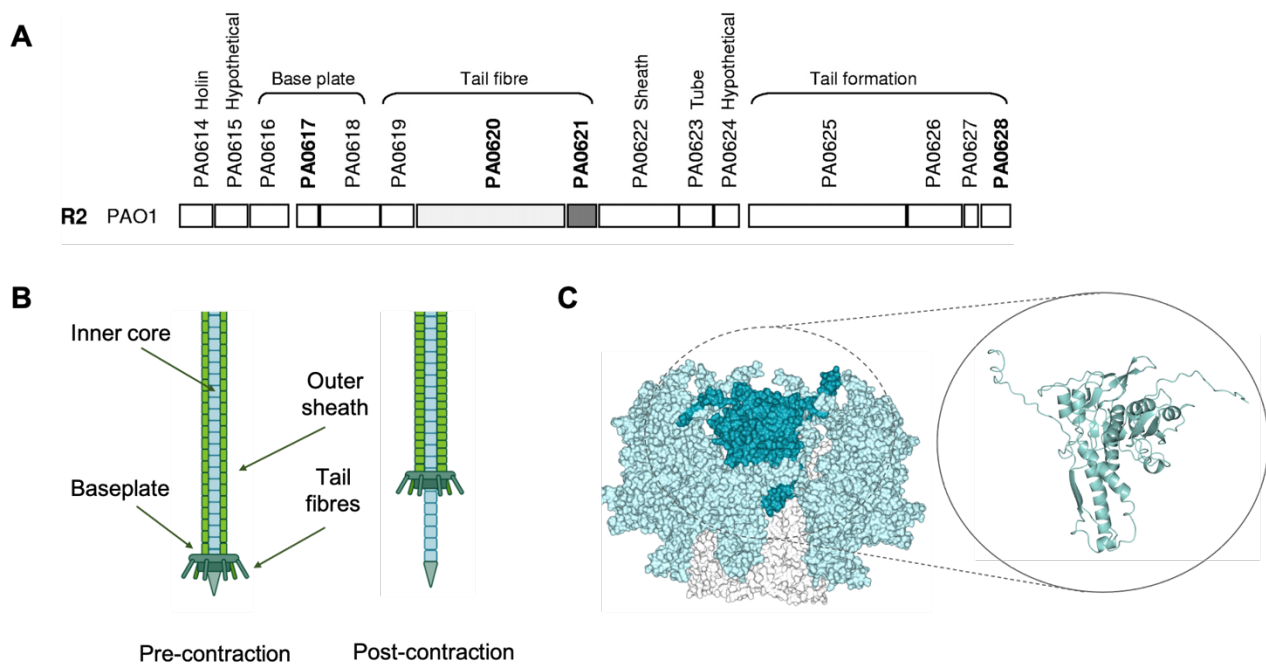
### 4.6.1 Background of PA0622

Spots C1-C4 were identified to be proteoforms of the pyocin sheath protein PA0622. In contrast to the other identified proteins, PA0622 does not function individually but actually forms part of the large macromolecular structure of the R2 pyocin. Pyocins are a type of bacteriocin (or more specifically tailocin), and the R2-type is evolutionarily related to the P2 bacteriophage (Nakayama *et al.*, 2000). As a result of this, there is vast confusion in the literature as to the true nature of the pyocin-encoding genes, with several studies misrepresenting them as symbiotic prophages instead of fundamental *P. aeruginosa* proteins (Lemieux *et al.*, 2016).

An R2 pyocin gene cluster encodes the structural and chaperone proteins necessary for pyocin production (Figure 4.17A). The pyocin nanomachine is principally composed of an inner core surrounded by an outer sheath and a baseplate from which tail fibres project (Figure 4.17B) (Buth *et al.*, 2018). The tail fibres are responsible for binding to target cells using LPS as a receptor; LPS serotype therefore mediates



pyocin susceptibility (Köhler *et al.*, 2010). After binding to a cell, the sheath contracts which drives the core structure through the cell envelope, causing depolarisation of the membrane, pore formation and ultimately cell death. The protein PA0622 makes up the outer sheath with a total of 162 subunits (6 subunits per level) weaved together to form an extensive mesh (Ge *et al.*, 2015) (Figure 4.17C). It is therefore a critical component of this potent nano-spear.

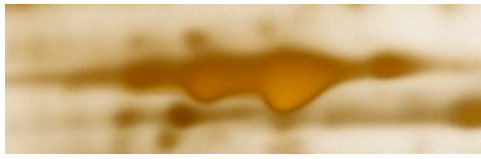


**Figure 4.17 R2-pyocin operon and structure**

(A) The R2-pyocin gene cluster with the known gene functions, adapted from (Köhler *et al.*, 2010); (B) R2-pyocin morphology in both the pre-contraction and post-contraction states; (C) 3D surface structure of the pyocin inner core (white) surrounded by a PA0622 sheath protein mesh (light blue), one subunit is highlighted in dark teal (PDB: 6PYT); the inset shows the ribbon structure of one subunit generated by I-TASSER structural prediction

#### 4.6.2 Differences in PA0622 proteoform abundance

Differences in the relative PA0622 proteoform abundance were measured. The average percentage intensity of each spot from three biological replicates is shown in Table 4.10 with the respective spots shown in Figure 4.18.



**Figure 4.18 Close up image of charge train C**

See Figure 4.2 for the full 2D gel image

**Table 4.10 Spot intensities of charge train C**

Numbers represent the intensity of each spot as a percentage of the whole charge train (average of three biological replicates  $\pm$  standard deviation) as determined by analysis with ImageJ software

Spot	1	2	3	4
% intensity	22.7 $\pm$ 5.3	33.8 $\pm$ 2.7	35.1 $\pm$ 3.7	8.4 $\pm$ 2.1

As with the two proteins investigated thus far, the spot at the second most alkaline pI (C3) proved most abundant. In all cases, the spot at the highest pI has shown to be the least abundant. The spots at the extreme ends are likely to be either the most or least modified causing the biggest shift in pI. This therefore suggests that a refined balance of modification may be preferable. Surprisingly, the study in PA14 only found one proteoform of PA0622 during 2D gel analysis (Gaviard *et al.*, 2019). As with LasB, this may suggest differences in PTM of these proteins between these strains, or experimental sensitivity.

#### 4.6.3 Prediction of PA0622 lysine acetylation

In contrast to ImpA and LasB, *in silico* prediction of PA0622 acetylation using Ensemble-PAIL did not return any predictive K-ac sites. This could be due to the even greater evolutionary discrepancy between eukaryotic training datasets and ancestral phage proteins. Conversely, ProAcePred predicted a total of 7 acetylation sites out of a total of 15 lysines. Once again this is a high level of predicted acetylation. Some sites, such as K312 and K176, demonstrated greater probability as compared with others.

**Table 4.11 ProAcePred prediction of lysine acetylation in PA0622**

The amino acid sequence of PA0622 was submitted to ProAcePred for prediction of lysine acetylation sites; the predicted sites are highlighted in blue

Protein name	Position of site	Flanking residues	SVM Probability
G3XD39	149	VATAMDGLAE- <b>K</b> -LRAIAILDGP	0.61907
G3XD39	171	STDEAAVAYA- <b>K</b> -NFGSKRLFMV	0.59442
G3XD39	176	AVAYAKNFGS- <b>K</b> -RLFMVDPGVQ	0.80201
G3XD39	277	WGNRTLSSDS- <b>K</b> -WAFVTRVRTM	0.58703
G3XD39	308	HKWAVDRGIT- <b>K</b> -TYVKDVTEGL	0.60789
G3XD39	312	VDRGITKTYV- <b>K</b> -DVTEGLRAFM	0.83795
G3XD39	326	EGLRAFMARDL- <b>K</b> -NQGAVINFEV	0.78353

#### 4.6.4 MS/MS identification of PA0622 acetylation

The PA0622 spot LC-MS/MS data was analysed to identify the presence of any acetylated peptides; the acetylation sites for each spot are listed in Table 4.12. Any K-ac found in more than one spot are highlighted in bold.

**Table 4.12 Acetylated lysines identified in spots C1 to C4**

**Bold** residues were found in more than one spot

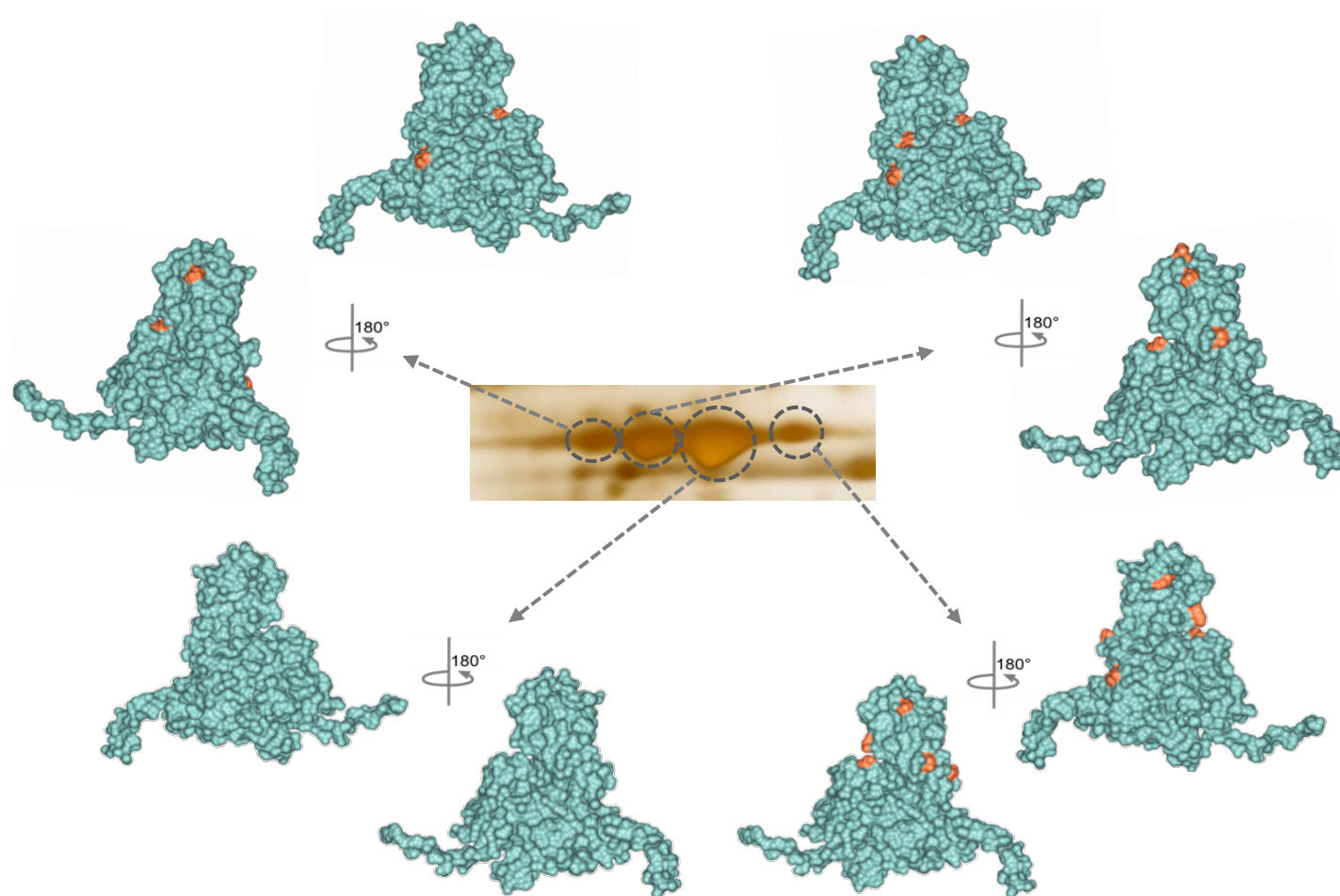
C1	C2	C3	C4
<b>K118</b>	<b>K118</b>		<b>K118</b>
<b>K171</b>	<b>K171</b>		<b>K171</b>
<b>K312</b>	K176		K277
	K308		K299
	<b>K312</b>		<b>K312</b>
	<b>K326</b>		<b>K326</b>
			K352

The results show that there are several PA0622 lysine residues that are acetylated in the extracellular environment; notably K118, K171, K312, K326, which demonstrated overlap between more than one spot. ProAcePred prediction of acetylation sites proved relatively accurate with the majority of predicted sites identified in the mass spectrometry data. Most surprisingly, the observation of increased acetylation with increased pI once more goes against the expected pattern. This can again be somewhat explained by the peptide sequence coverage and number of significant spectral matches, which were highest for spot C4 (Table 4.1). This in itself is unexpected as this protein spot was the least abundant when the spot intensities were measured (Section 4.6.2). Additionally, as with LasB, spot C3 was analysed on a different mass spectrometer to the rest of the spots. As very few modified peptides were found in C3 and D3 (which were analysed on the same mass spectrometer), it can be surmised that the true depth of acetylation was not realised in these samples. This is the first time, to my knowledge, that acetylation of a tailocin protein has been identified.

#### **4.6.5 Structural mapping of PA0622 acetylation**

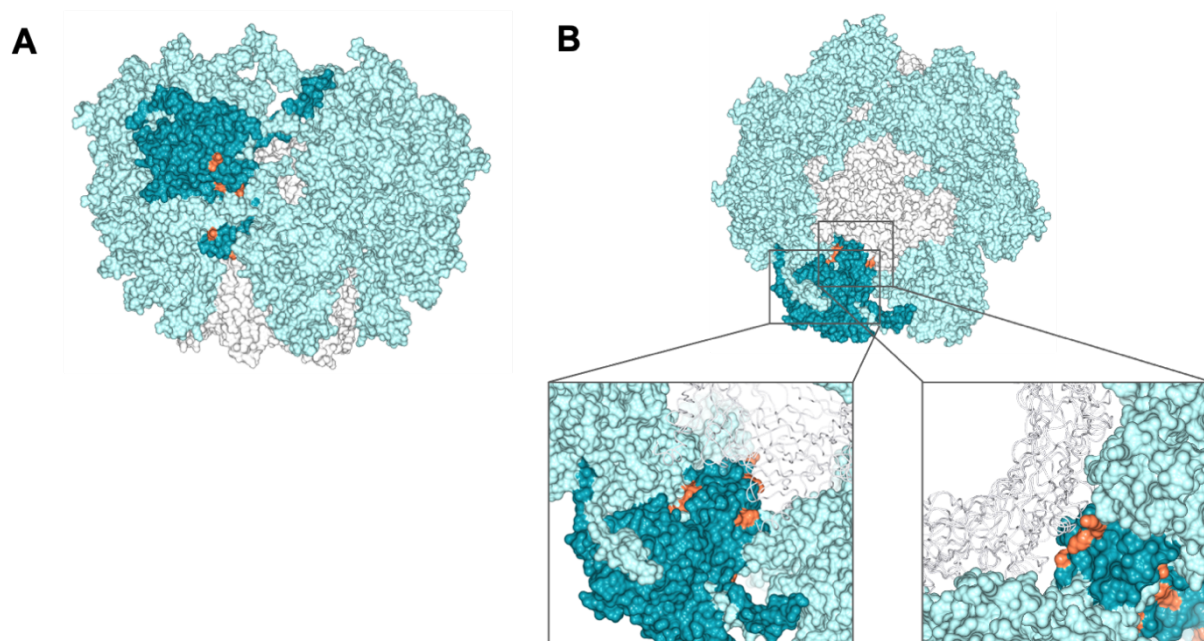
At the time of analysis, there were no deposited subunit structures of PA0622 in the Protein Data Bank onto which acetylation sites could be mapped. The amino acid sequence was therefore submitted to I-TASSER, an iterative threading programme for structural prediction. The output model had a high confidence score of 1.82 (where the range is -5 to 2) and was therefore used for structural mapping of K-ac (Figure 4.19); the subunit ribbon model can be seen in the inset of Figure 4.17C.

As with ImpA and LasB, all of the K-ac mapped to the surface of the protein. Additionally, there did not appear to be a preference for the outer “environment-facing” side of the protein or the inner side that interacts with the core structure. The full sheath is a complex mesh of subunits and, as such, it is also necessary to view the K-ac sites on the full structure to further elucidate potential protein interactions. For this, the PDB file (6PYT) composed of 12 subunits each of the core and sheath proteins was used for further PTM modelling (Ge *et al.*, 2020) (Figure 4.20).



**Figure 4.19 Structural mapping of acetylated lysines in spots C1-C4**

The likely 3D structure of PA0622, as predicted by I-TASSER, was analysed using CCP4mg software. The location of lysines identified as acetylated in the LC-MS/MS data of each spot from charge train C were mapped onto the surface of the protein structure; acetylated lysines are highlighted in orange and all other residues are coloured in teal. The models for each spot are identified by an arrow



**Figure 4.20 Structural mapping of PA0622 K-ac sites onto the pyocin trunk**

The full structure of the R2-pyocin trunk was obtained from the PDB (6PYT) and analysed using CCP4mg. The inner core proteins are coloured white and outer sheath PA0622 proteins are coloured light blue with one subunit highlighted in dark cyan; acetylated lysines are coloured orange. (A) Side view and (B) top view with insets showing close interactions between the core proteins and acetylated lysines

By mapping the acetylation sites onto a more complete structure of the pyocin, it is clear that several of the modified lysines are located close to other PA0622 subunits, as well as core structural proteins. It can therefore be hypothesised that acetylation may in some way either stabilise the structure or may be involved in protein interactions during pyocin assembly.

#### 4.6.6 Sequence conservation of PA0622 acetylated lysines

The conservation of the lysine acetylation sites in homologous proteins was investigated. The top hits when the amino acid sequence was BLAST against the UniProt protein database were mostly all *Pseudomonas* species, including *P. fluorescens*, *P. knackmussii* and *P. citronellolis* with 80%+ sequence homology. A protein from *Klebsiella pneumoniae*, another member of the ESKAPE pathogens that are highly virulent and antibiotic resistant, was also found to have 99% sequence identity. When these sequences were aligned, all but one of the acetylation sites were



conserved in all strains; K171 was substituted for glycine in *P. knackmussii* and *P. citronellolis*. These residues are likely to be key to the sheath structure. The full alignment can be found in Appendix A.

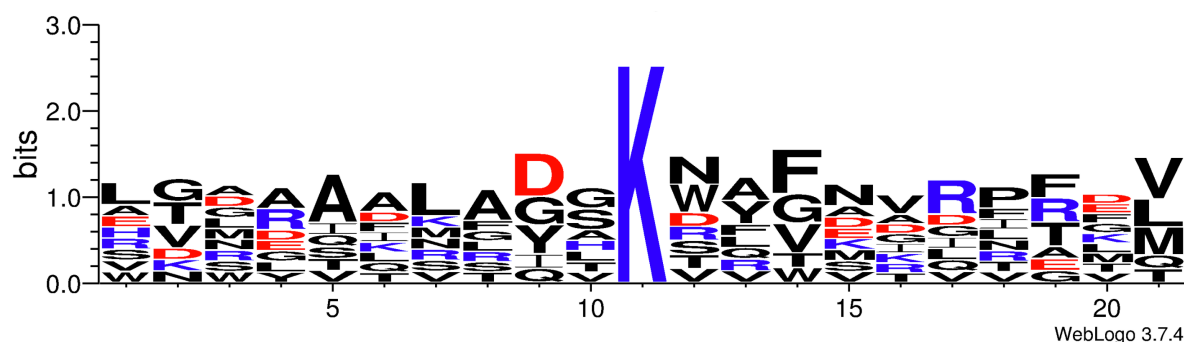
<i>P. aeruginosa</i>	61	IGSSIYLACEAIYNRAQAVIVAVGVEAETPEAQASAVIGGISAAGERTGLOALLDCKSR
<i>P. fluorescens</i>	61	IGSSIYLACEAIYNRAQAVIVAVGVEAETPEAQASAVIGGVSAAGERTGLOALLDCKSR
<i>K. pneumoniae</i>	61	IGSSIYLACEAIYNRAQAVIVAVGVEAETPEAQASAVIGGVSAAGERTGLOALLDCKSR
<i>P. knackmussii</i>	61	VDSAIYKSCVAIYTQSAVVAVGVKLAETPEMQASAVIGTVTASGORTGLOALLDCKSR
<i>P. citronellolis</i>	61	TGSAIYKACTAIFTQASAVVAVGVAEVEDPAQQTSAIGSVTESGORTGLOALLDCKSR
<i>P. aeruginosa</i>	121	FNAQPRLLVAPGHSAQQAVATAMDGLAEKLRAIAILDGPNSTDEAAVAYAKNFGSKRLFM
<i>P. fluorescens</i>	121	FNAQPRLLVAPGHSAQQAVATAMDGLAEKLRAIAILDGPNSTDEAAVAYAKNFGSKRLFM
<i>K. pneumoniae</i>	121	FNAQPRLLVAPGHSAQQAVATAMDGLAEKLRAIAILDGPNSTDEAAVAYAKNFGSKRLFM
<i>P. knackmussii</i>	121	FNTQPRLLVAPGHSSTQAVATAMDGLAEKLRAIAILDGPNSTDEDAVDYAGEFGSKRLFM
<i>P. citronellolis</i>	121	FNAQPRLLVAPKHSATEAVATAMDGLAGKLRAIAILDGPNSTDEAATAYAGEFGSKRVYL
<i>P. aeruginosa</i>	241	GDETCRANLLNNANIATIIIRDDGYRLWGNRTLSSDSKNVAFVTRVRTMDLVMDAILAGEKN
<i>P. fluorescens</i>	241	GDETCRANLLNNANIATIIIRDDGYRLWGNRTLSSDSKNVAFVTRVRTMDLVMDAILAGEKN
<i>K. pneumoniae</i>	241	GDETCRANLLNNANIATIIIRDDGYRLWGNRTLSSDSKNVAFVTRVRTMDLVMDAILAGEKN
<i>P. knackmussii</i>	241	GDAETCRANLLNNANIATIIIRDDGYRLWGNRTLSSDSKNVAFVTRVRTMDLVMDAILAGEKN
<i>P. citronellolis</i>	241	GDAETCRANLLNNANVTIIIRDDGYRLWGNRTLSSDSKNVAFVTRVRTMDLVMDAILAGEKN
<i>P. aeruginosa</i>	301	AVDRGITKTYVKDVTEGLRAFMRDLKNOGAVINFEVYADPDLNSASQLAQQKVYWNIRFT
<i>P. fluorescens</i>	301	AVDRGITKTYVKDVTEGLRAFMRDLKNOGAVINFEVYADPDLNSASQLAQQKVYWNIRFT
<i>K. pneumoniae</i>	301	AVDRGITKTYVKDVTEGLRAFMRDLKNOGAVINFEVYADPDLNSASQLAQQKVYWNIRFT
<i>P. knackmussii</i>	301	AVDRGITKTYVKDVTEGLRAFMRDLKNOGALIDFEVYADPDLNTSSQLAQQKVYWNIRFT
<i>P. citronellolis</i>	301	AVDRGITKTYVKDVTEGLRAFMRDLKNOGAVIDFEVYADAEELNTASQLAQQKVYWNIRFT

**Figure 4.21 Sequence alignment of PA0622**

Sequence alignment of PA0622 with homologous proteins from other species, as identified by BLAST analysis; outlined in red are the acetyl-residues found to be conserved in other species

#### 4.6.7 PA0622 acetylation motif analysis

The consensus sequences surrounding the PA0622 acetylation sites were assessed for motif analysis (Figure 4.22). As previously found with the other proteins analysed, there is no clear preference for an amino acid at any of the positions up- or downstream of the K-ac. Similar to LasB, aspartic acid is slightly more prevalent at the -2 position. Phenylalanine also appears slightly more frequently at the +3 site, however the number of K-ac sites compared is once again too low for any significance to be ascertained.

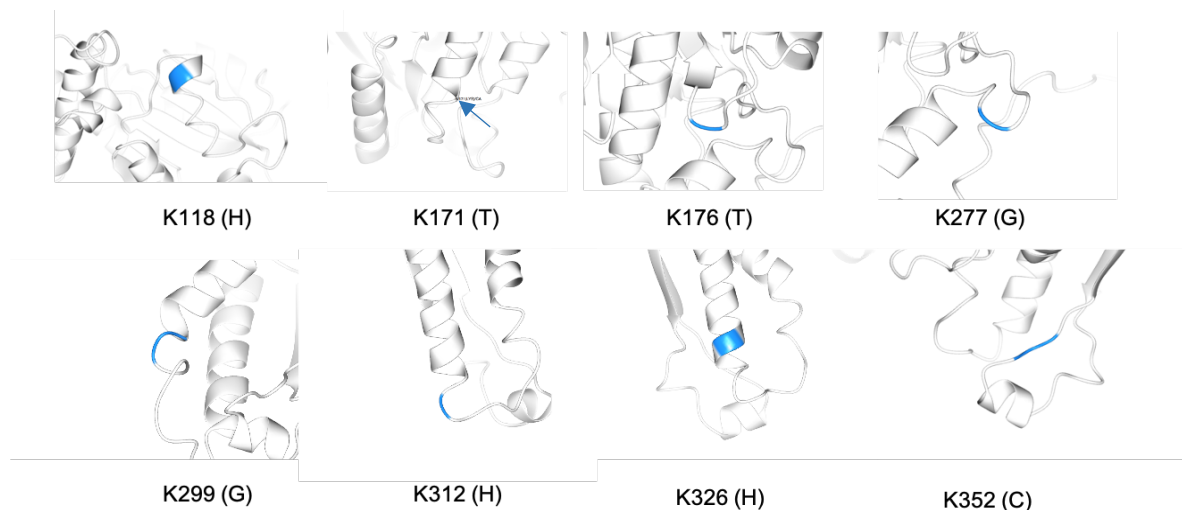


**Figure 4.22 PA0622 acetylation motif analysis**

The frequency of residues surrounding every acetylated lysine ( $\pm 10$  residues) were analysed using WebLogo; the size of the symbol correlates to its frequency at that position. The colours represent the charge of the amino acid, positive (blue), negative (red), or uncharged (black)

#### 4.5.8 Secondary structure preference

Due to the potentially stabilising or interactive nature of the PA0622 K-ac sites with surrounding residues of other proteins, it was questioned whether secondary structure could influence acetylation in this protein. The K-ac sites were thus highlighted in the subunit ribbon structure as well as DSSP assignment of the nature of the secondary structure (Figure 4.23). Acetylation seemed to preferentially occur either on/at the base of helices or in turns or coils. Beta sheets occur at a low frequency in this protein and so their absence here is expected. Overall, there is no preference for acetylation at any secondary structure.



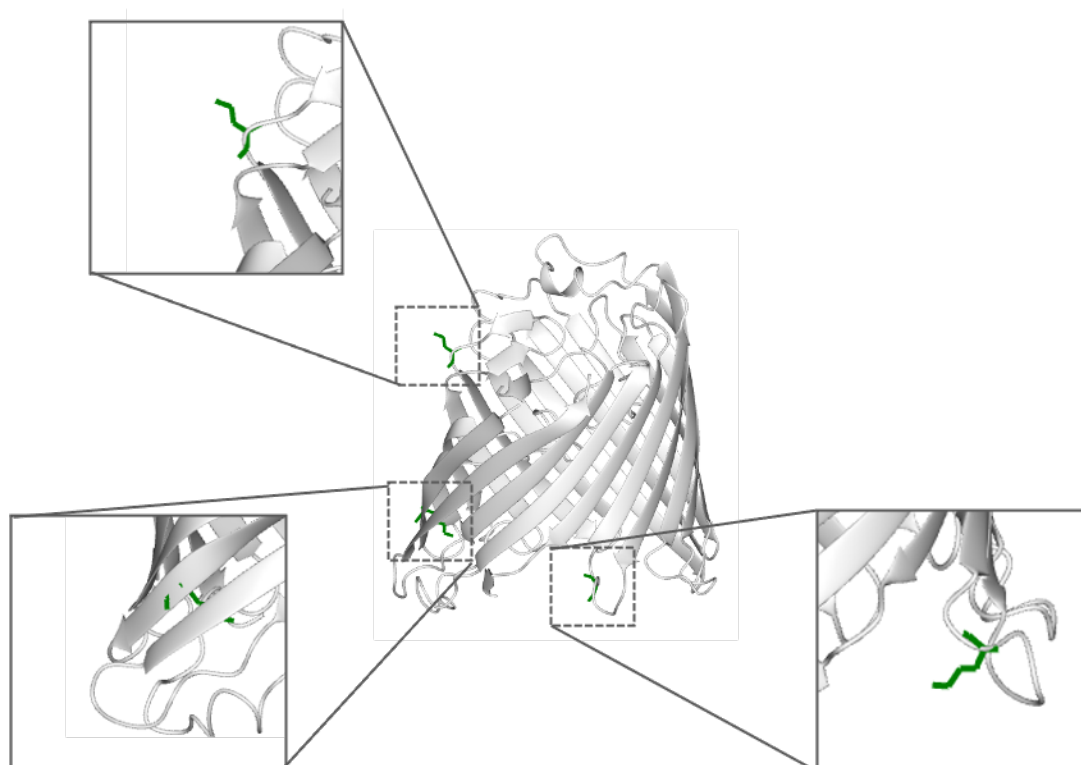
**Figure 4.23 Secondary structure at the PA0622 K-ac sites**

The predicted secondary structure surrounding the acetylated lysines in PA0622. The K-ac sites are highlighted in blue and the letter in brackets following the residue number indicates the DSSP structure assignment: H = alpha helix, G = 3-helix, T = turn, C = coil



## 4.7 Acetylation of OprD

In addition to the main proteins outlined above, three small protein spots at an acidic pI were also picked and analysed by LC-MS/MS (spots F1-F3, Figure 4.2). They were identified to be isoforms of the outer membrane porin protein OprD, alternatively known as OccD1. Further to this, each OprD charge isoform was found to be acetylated. All spots showed acetylation of K205, as well as K128 and K115 for spot F2 and K407 for spot F3. These acetylation sites were mapped onto the 3D structure of the protein, except K115 as this residue was missing from the published crystal structure (Figure 4.24). Lysine K205, the most commonly acetylated, is found on the extracellular side of the porin, whilst K128 and K405 are both on the periplasmic facing side. It is not yet clear from the position of these residues what the role of acetylation might be.



**Figure 4.24 Structural mapping of OprD K-ac**

Acetylated lysine residues (green) identified in LC-MS/MS analysis of OprD spots from a 2D gel were mapped onto the 3D structure (PDB: 3SY7) using CCP4mg. Top inset: K205, bottom left inset: K128, bottom right inset: K407

## 4.8 Discussion

The study of PTM in bacteria is no longer in its infancy, however the current state of knowledge is overpopulated by a subset of bacterial species (André *et al.*, 2017) . In addition, research into bacterial PTM commonly excludes the extracellular proteome from analysis, thus overlooking an important group of potentially virulence-defining proteins. So far only two studies have been published investigating PTMs of *P. aeruginosa* secreted proteins. Both studies concentrated on strain PA14, with one determining the phospho-exoproteome, and the other the acyl-exoproteome (with a focus on two virulence factors in particular) (Ouidir *et al.*, 2014; Gaviard *et al.*, 2019). The results presented here have not only confirmed those of the latter research in a different *P. aeruginosa* strain (PAO1) but have also extended these findings to include other *P. aeruginosa* secreted proteins, and a greater number of acetylation sites.

This study began by determining the identity of proteins which consistently and abundantly present as charge variants within the secretome. This included two potent proteases, LasB and ImpA, as well as components of the pyocin nanomachine. ImpA was identified at two distinct molecular weights for the first time. Although the reasons for this are unclear, it is hypothesised that a high molecular weight PTM, or potentially additional N-terminal cleavage, may be responsible; this would be an interesting avenue for further study. Any alternative proteoforms of these proteins which differ more significantly in their molecular weight and charge were not identified. For example, LasB has previously been sequenced from as many as 28 different spots at a variety of molecular weights and isoelectric points, exclusive of those analysed here (Scott *et al.*, 2013). These generally constitute the immature form and subsequent degradation products of LasB. As such, the spots selected for analysis in this research encompass proteoforms of the mature active protease, which is the most biologically and clinically relevant form.

The next question that stood to be answered was which charge modifying PTMs were present on the charge isoforms that could help to explain the charge train phenomenon. Deamidation was initially considered, however discrimination between the extent to which this PTM was biologically introduced or introduced through LC-MS/MS sample

preparation procedures was not possible (Hao *et al.*, 2017). Nonetheless, profuse deamidation was demonstrated in the data, some of which may contribute to the charge variation seen during 2D gel analysis (data not shown). Lysine acetylation was subsequently selected as a PTM of interest due to its relatively small size and charge reducing properties.

For all of the proteins analysed, clear and consistent lysine acetylation was observed. In the majority of cases, the acetylation sites were located on the surface of the protein. Whilst this is to be expected due to the hydrophilicity of the lysine substrate, it may also suggest that modification may occur after at least partial protein folding. This notion is supported by previous findings that the level of PTM of secreted proteins, such as LasB, significantly increases upon secretion (Gaviard *et al.*, 2019). ImpA, in particular, demonstrated a proclivity for modification on one “arm” of the protein. This protease has not been fully biochemically characterised; however, this surface interface may form an important domain for protein-protein interactions. As such, PTM at these sites may modulate such interactions within the extracellular environment. This is even more significant when the PA0622 K-ac sites are mapped onto the pyocin trunk, showing direct interaction of the acetylation sites with both core structures and other PA0622 subunits.

Unfortunately, no consensus sequence that may be a signal for acetylation could be determined. The relatively limited number of acetylation sites compared prevented any clear motifs from becoming apparent. However, even when the sites determined by acetylome analysis of a single species are aligned, it is common for as many as fifteen different motifs to emerge, suggesting quite broad target specificity for this PTM (Yang *et al.*, 2018). Additionally, comparison of the secondary structures at each K-ac site indicated that there were no clear structural preferences. Despite previous claims that acetylation at distinct secondary structures is likely to be species specific, these results suggest that such specificity may be at the single protein level. Alternatively, any correlations observed may be as a result of structural biases determined by the steric constraints of the amino acids. An alternative approach which deviates from analysis of linear alignment would be to consider the spatial (i.e. radial) amino acid composition from the PTM site, taking into account the tertiary structure of the protein (Su *et al.*, 2017). This approach would be particularly pertinent if PTM occurs post-folding.

Despite extensive acetylation being identified, the expected pattern of a decrease in acetylation across the charge train (as the spot pI increases) was not supported by the data. There are several likely reasons for this. Firstly, there were differences in the abundance of each charge isoform that generally translated to fewer peptides being sequenced during LC-M/SMS analysis for the less abundant spots. This in turn would result in fewer acetylated residues being identified in the data, especially considering that PTMs are typically sub-stoichiometric within a protein population (Hart-Smith *et al.*, 2012). Furthermore, signal suppression of modified peptides occurs during MS analysis due to the abundance of non-modified peptides, making PTM identification even more challenging (Temporini *et al.*, 2008). The less abundant spots constitute those at the lowest pI, and thus those postulated to harbour the most K-ac sites. Therefore, a lack of identification of increased acetylation in these isoforms does not mean these PTMs are indeed absent, but potentially just overlooked.

Secondly, by mining the data for one PTM in particular, other potential charge modifying PTMs are neglected. For example, mature LasB has previously been identified to undergo significant phosphorylation at serine, threonine and tyrosine residues (Ouidir *et al.*, 2014). Phosphorylation introduces a negatively charged phosphate group, therefore also contributing to a negative shift in pI. Additionally, it must be noted that there are several additional charge modifying acylations (e.g. propionylation, butyrylation, malonylation and crotonylation) that commonly occur on the same lysines as those that are acetylated (Gaviard *et al.*, 2019). In particular, it has been shown that there is a significant overlap between lysine acetylation and succinylation sites in several organisms (André *et al.*, 2017). This is surprising considering the differences in PTM chemistry. Acetylation (+42.01 Da) removes the lysine positive charge (+1 to 0), whilst succinylation (+100.02 Da) also introduces a negative charge (+1 to -1). These PTMs are therefore likely to have different effects on the protein, probably allowing the fine-tuning of an organism's responses to the environmental signals it perceives.

It is the combination of all PTMs which ultimately decides the overall protein charge and therefore they must all be considered collectively. This is a huge undertaking in terms of data acquisition and computational interpretation as hundreds of PTMs have so far been elucidated and many are yet to be discovered (Minguez *et al.*, 2012).

Current technologies are not well equipped to deal with such complex data. Moreover, proteins differ in their charge buffering capacity and so the amount of mobility shift due to PTM is proteoform, and even pH, specific (Zhu *et al.*, 2005; Lange *et al.*, 2014). Ultimately, it was concluded that the charge train phenomenon is unlikely to be deduced to modulation by a single PTM at this level of interrogation. This will be further discussed in Chapter 5 where the identification of another important PTM, methylation, is also explored.

The overlap of identified K-ac sites between spots (and between charge trains in the case of ImpA) not only provided replicate style validation, but also suggested that these are important and targeted sites for modification. Many of the sites were also pre-determined by predictive algorithms, supporting the proposal that in the absence of advanced technologies, the use of predictive algorithms for lysine acetylation can provide a starting platform for more targeted studies of potential bacterial PTMs (Chen *et al.*, 2018). Some of the K-ac sites were further confirmed by Gaviard *et al.* who also found that K211, K378, K388 and K443 of LasB, as well as K708 in ImpA, are acetylated in strain PA14 (Gaviard *et al.*, 2019). The additional residues and proteins highlighted in the research presented here therefore extend their findings. Differences in PTM identification are probably due to a combination of possible strain differences and differences in experimental resolution. On that note, Gaviard *et al.* used two-dimensional immunoaffinity enrichment of modified peptides, first isolating succinylated peptides and second acetylated peptides from the first isolation pool. This therefore excludes lysine residues which are not susceptible to succinylation.

Approximately fifty acetyltransferases have been annotated on the *Pseudomonas* Genome Database, however their respective substrates remain mostly elusive (Ouidir *et al.*, 2016). As a result of the similarities in acetylation sites, it can be speculated that PAO1 and PA14 encode a common acetyltransferase capable of catalysing the acetylation of these residues. Subject to confirmation of this, and elucidation of the role of these PTMs in protein function, the responsible enzyme could well become a new target for anti-virulence drug intervention.

In terms of the function of lysine acetylation, this PTM has been shown to have diverse roles in bacteria (Ren *et al.*, 2017; David G. Christensen *et al.*, 2019), including stress

resistance and stability, virulence and immunogenicity (Liang *et al.*, 2011; Ma and Wood, 2011; Liu *et al.*, 2014). However, the functional importance of acetylation with respect to a specific protein cannot be determined *a priori*. Due to the number of acetylation sites identified on the proteins presented here it is likely that some may be relatively functionally redundant, however several are likely to have significant roles. Initial investigations into the effects of lysine acetylation of LasB are detailed in Chapter 6.

#### **4.8.1 Conclusion**

In conclusion, a significant proportion of the *P. aeruginosa* secretome is present as a charge variable sub-population that is likely to be generated by inconsistent PTM. Whilst several of these secreted proteins harbour surface-exposed acetylated residues, which may account for some of the charge variation seen, an overall explanatory pattern of acetylation could not be deduced at this level of interrogation. Nonetheless, acetylation of these secreted proteins is clear and consistent and is therefore likely to have an important role in the protein, and thus bacterial, physiology.

## Chapter 5

### 5 Methylation of *P. aeruginosa* secreted proteins

#### 5.1 Introduction

Whilst protein methylation has been extensively studied in eukaryotes, particularly with regards to the regulation of histone functionality, it is vastly understudied in prokaryotes. Protein methylation involves the addition of a small methyl group (+14.0157 Da) to the amino acid side chain and predominantly occurs on lysine and arginine residues, although atypical methylation can also occur (M. Zhang *et al.*, 2018). S-adenosyl methionine (SAM) acts as a methyl donor, working in concert with methyltransferase and demethylase enzymes. There is a confusing inconsistency in the literature with regards to the effect of methylation on protein chemistry. Several researchers support the idea that methylation affects the charge of lysine and arginine residues (Frye *et al.*, 2006; Stewart *et al.*, 2016), although the general consensus appears to dispute this (Lanouette *et al.*, 2014). In *P. aeruginosa*, several intracellular or membrane-bound methyltransferase substrates are known, including chemotaxis related proteins and elongation factor EF-Tu (Schmidt *et al.*, 2011; Barbier *et al.*, 2013). Beyond this, there are no specific studies focusing on the methylation of secreted proteins.

##### 5.1.1 Aims

The aims of this Chapter were to develop upon the findings of widespread acetylation, as detailed in Chapter 4, by including methylation as a potential PTM of secreted proteins. More specifically, the defined location and degree of methylation of these proteins, as well as the presence of “atypical” methylation of asparagine and histidine were to be investigated.

## 5.2 Methylation of ImpA

### 5.2.1 Prediction of ImpA methylation

As with the investigations into secreted protein acetylation in Chapter 4, the first step in exploring methylation was to use *in silico* prediction to determine likely methylation sites. The prediction software used was MASA (Shien *et al.*, 2009), iMethyl-PseAAC (Qiu *et al.*, 2014) and PRmePRed (Kumar *et al.*, 2017). MASA and iMethyl both incorporate the protein physicochemical and structural information in their prediction of lysine and arginine methylation, with MASA placing greater emphasis on secondary structure and the accessible surface area. PRmePRed, meanwhile, focuses solely on arginine methylation, a relatively recent identified PTM in bacteria (Lassak *et al.*, 2019).

ImpA is a relatively large protein with a total of 82 (51 R, 31 K) methylatable lysine and arginine residues in the mature protein. Concerningly, there was very little overlap (shown in bold, Table 5.1) of the predicted methylation sites between the prediction software used. PRmePRed indicated extensive arginine methylation was likely with 21 sites predicted, however this was not supported by the predictions of iMethyl-PseAAC, which showed no overlap with any residues with the other software. A lack of concordance between them suggests that the computational prediction of protein methylation requires considerable further development. It is also questionable as to how valid these algorithms are when concerning bacterial proteins and PTM systems as they are generally derived from eukaryotic data.



**Table 5.1 Prediction of ImpA methylation**

The amino acid sequence of ImpA was submitted to several methylation prediction algorithms; MASA, PRmePRed and iMethyl-PseAAC. Residues predicted to undergo methylation by more than one algorithm are highlighted in **bold**

MASA	PRmePRed	iMethyl-PseAAC
<b>R135</b>	R69	K159
<b>R596</b>	R123	K281
R646	<b>R135</b>	K350
<b>R690</b>	R237	R373
	R239	K392
	R288	K504
	R290	K575
	R292	K591
	R330	K594
	R413	K635
	R429	K708
	R457	K726
	R464	K767
	R468	K797
	R485	K819
	<b>R596</b>	K826
	<b>R690</b>	
	R706	
	R713	
	R756	
	R837	

### 5.2.2 MS/MS identification of ImpA methylation

The data obtained during LC-MS/MS identification of charge train spots (Chapter 4) was further probed to allow the characterisation of protein methylation. Initially, methylation of lysine and arginine were investigated, however this was later extended to include atypical asparagine (N) and histidine (H) methylation. Due to the number of different variable modifications that needed to be included during MASCOT searching, a manual iterative approach was taken to reduce the bioinformatic search space and reduce the likelihood of false positive identification. Firstly mono-methyl K/R/N/H was pursued, followed by di-methyl K/R and tri-methyl K. Cross-referencing of spectra between searches was then undertaken to validate the results. The identified methylation sites for each spot in charge train A and B are listed in Table 5.2 and Table 5.3, respectively.

As predicted by the algorithms, a significant level of methylation was identified in ImpA (26 residues in total); this was mostly realised in the more abundant spots A4 and B3-B5. Intriguingly, the sites identified in spot A4 match better with those of B5, suggesting that there may be a pI shift involved in the responsible modification that causes the differences in molecular weight. Overall, it appears that the preliminary predictions of methylation made were not sufficient or relevant to this protein. K420 was found to undergo di-methylation and tri-methylation, but not mono-methylation, in all of the spots analysed. This residue was also found to be consistently acetylated in each spot, suggesting that PTM at this site is likely to be important. The spectra for the starred residues in Table 5.3 did not provide sufficient information for discrimination as to whether K726 and K729 were both mono-methylated or K726 di-methylated. There is, however, clear modification occurring at this location on the protein.

**Table 5.2 Methylated K/R/N/H identified in spots A1-A5**

**Bold** residues were found in more than one spot, *italicised* residues were also found in the corresponding spot in charge train B (see Table 5.3).

Me = mono-methyl, Me<sup>2</sup> = di-methyl, Me<sup>3</sup> = tri-methyl

A1			A2			A3			A4			A5		
Me	Me <sup>2</sup>	Me <sup>3</sup>	Me	Me <sup>2</sup>	Me <sup>3</sup>	Me	Me <sup>2</sup>	Me <sup>3</sup>	Me	Me <sup>2</sup>	Me <sup>3</sup>	Me	Me <sup>2</sup>	Me <sup>3</sup>
<b><i>K420</i></b>			<b><i>K420</i></b>			<b><i>K420</i></b> <b><i>K420</i></b>  <b><i>K601</i></b>			<i>H154</i>	K159	K187	<b>N207</b>	<b><i>K176</i></b>	
									K176	<b>K176</b>	K195	<b>R316</b>	<b><i>K195</i></b>	
									K179	K179	<b>K420</b>	<b>R646</b>	K323	
									<i>N192</i>	<b>K195</b>	K601		<i>K350</i>	
									<i>K195</i>	K359	K726		<b>K420</b>	
									<b>N207</b>	<b>K420</b>			<b>K601</b>	
									<b>R316</b>	<b>K726</b>			<b><i>K726</i></b>	
									<b>R646</b>					
									H663					
									N665					
									N679					
									R690					
									H696					
									H700					
									R729					
									<i>R901</i>					

**Table 5.3 Methylated K/R/N/H identified in spots B1 to B6**

**Bold** residues were found in more than one spot, *italicised* residues were also found in the corresponding spot in charge train A (see Table 5.2).

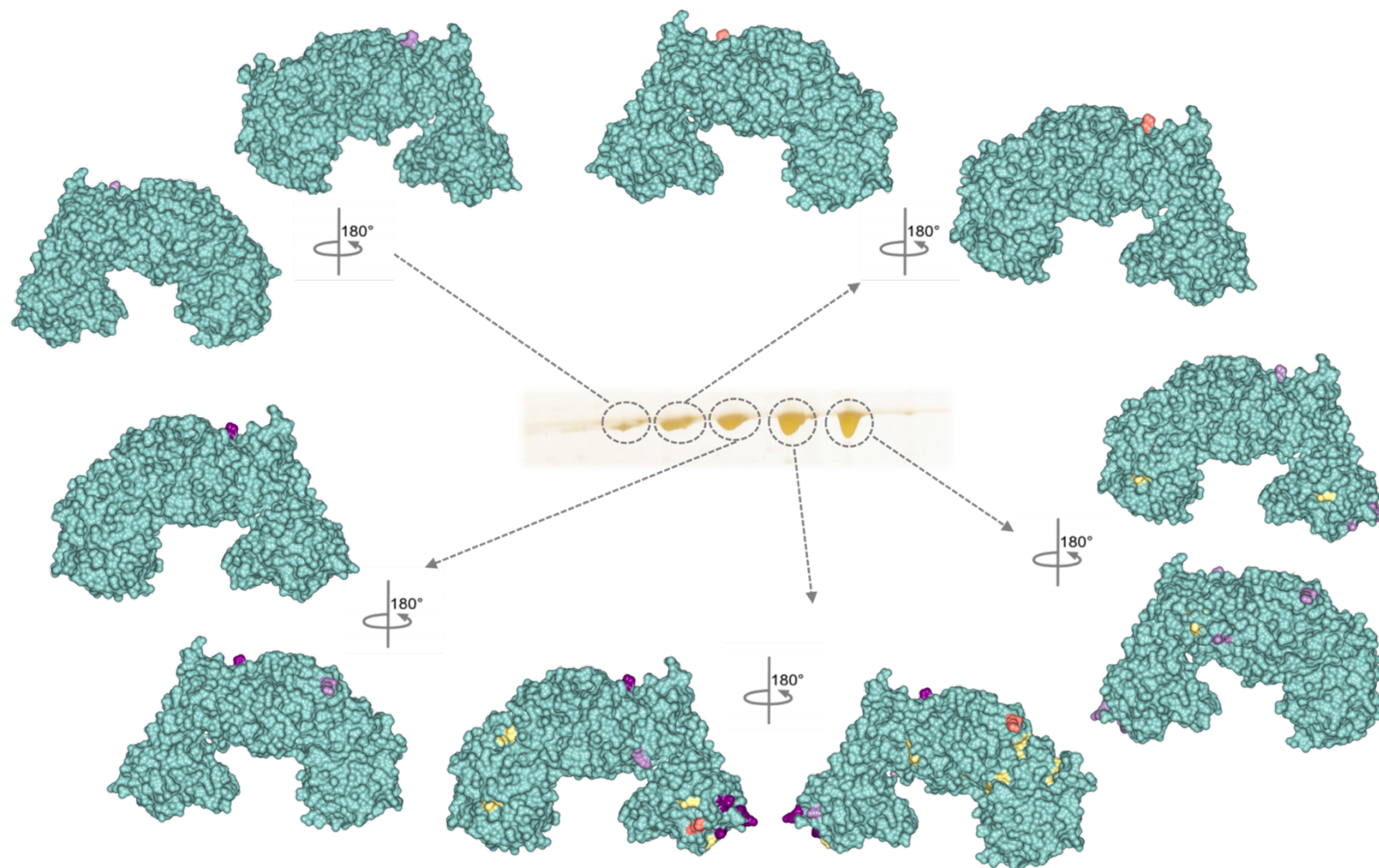
Me = mono-methyl, Me<sup>2</sup> = di-methyl, Me<sup>3</sup> = tri-methyl

B1			B2			B3			B4			B5			B6		
Me	Me <sup>2</sup>	Me <sup>3</sup>	Me	Me <sup>2</sup>	Me <sup>3</sup>	Me	Me <sup>2</sup>	Me <sup>3</sup>	Me	Me <sup>2</sup>	Me <sup>3</sup>	Me	Me <sup>2</sup>	Me <sup>3</sup>	Me	Me <sup>2</sup>	Me <sup>3</sup>
<b>K420</b>			<b>K420</b>	<b>K420</b>		<b>H154</b>	<b>K159</b>	<b>K195</b>	<b>R76</b>	<b>K159</b>	<b>K195</b>	<b>R76</b>	<b>K159</b>	<b>K420</b>	<b>K195</b>	<b>K176</b>	<b>K420</b>
						<b>R160</b>	<b>R160</b>	<b>K420</b>	<b>H154</b>	<b>R160</b>	<b>K420</b>	<b>H154</b>	<b>K176</b>	K726	<b>R646</b>	<b>K179</b>	
						<b>N192</b>	<b>K176</b>		<b>R160</b>	<b>K176</b>		<b>K187</b>	<b>K179</b>		<b>K726*</b>	<b>K726*</b>	
						<b>R646</b>	<b>K179</b>		<b>K187</b>	<b>K179</b>		<b>N192</b>	<b>K195</b>		R729*		
							<b>K187</b>		<b>N192</b>	<b>K187</b>		<b>K195</b>	K344				
							<b>K195</b>		<b>K195</b>	<b>K195</b>		<b>R316</b>	K350				
							<b>K420</b>		<b>N207</b>	<b>K420</b>		K350	K359				
									<b>R316</b>	<b>K726</b>		N357	R413				
									<b>R646</b>			K359	<b>K726</b>				
									<b>R901</b>			K601					
												<b>R646</b>					
												H663					
												N665					
												N679					
												H696					
												<b>K726</b>					
												N903					

Some residues, such as K195, were found to undergo all three levels of methylation, as well as acetylation. This observation of multi-modification of bacterial proteins has also been found in PA14, although not for this protein specifically (Gaviard *et al.*, 2019). A number of methyl-asparagine and histidine residues were also observed, indicating that these are underappreciated and biologically relevant PTMs. Significant consistencies both within the charge trains and between the charge trains suggests that these modifications are targeted and are therefore likely to serve some, as yet unidentified, purpose for the protein.

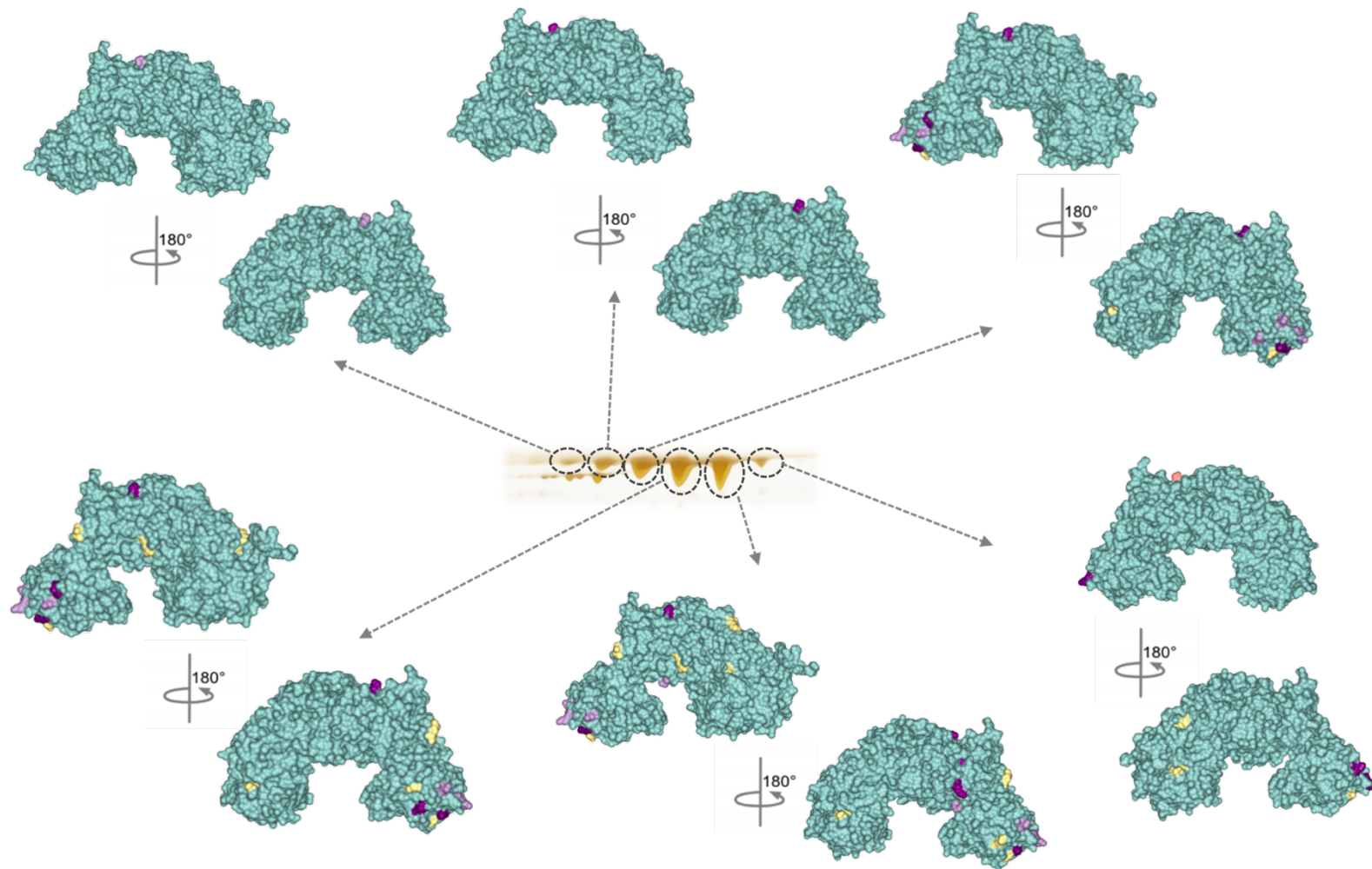
### **5.2.3 Structural mapping of ImpA methylation**

To gain a better understanding of where methylated residues are located on the protein, the methylation sites were mapped onto the ImpA 3D structure. All but three sites (H154, N665 and K726) could be mapped to the protein surface, and therefore 3D models of ImpA showing the protein surface were used to demonstrate methylation locations (Figure 5.1 and Figure 5.2). As was seen when the acetylated lysines were mapped onto the structure of ImpA, a preference for methylation occurred on one “arm” of the protein, which appeared more heavily modified than the other in both charge trains.



**Figure 5.1 Structural mapping of methylated residues identified in spots A1-A5**

The 3D crystal structure of ImpA (PDB: 5KDV) was analysed using CCP4mg software. The location of K/R/N/H residues identified as methylated in the LC-MS/MS data of each spot from charge train A were mapped onto the surface of the protein structure; mono-methyl = yellow, di-methyl = lilac, tri-methyl = orange, combination = dark purple, all other residues = teal. The models for each spot are identified by an arrow

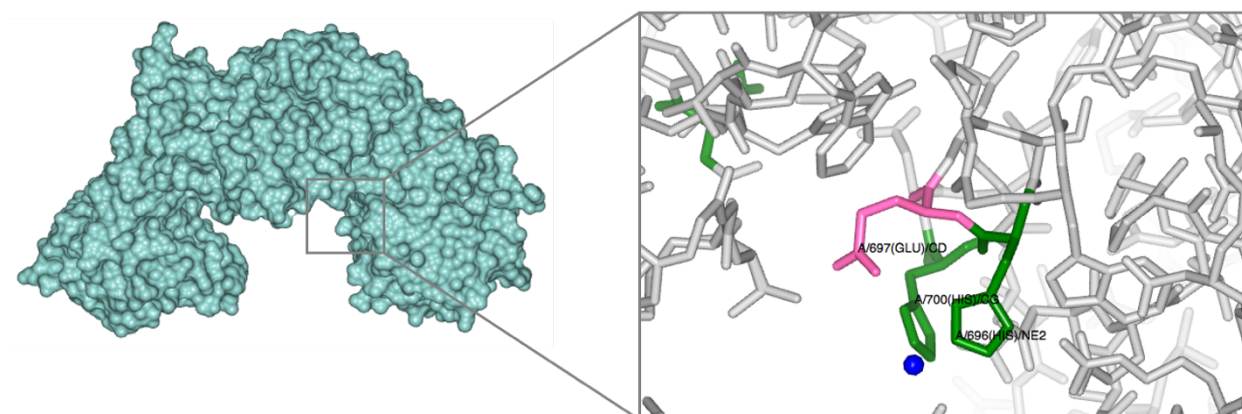


**Figure 5.2 Structural mapping of methylated residues identified in spots B1-B6**

The 3D crystal structure of ImpA (PDB: 5KDV) was analysed using CCP4mg software. The location of K/R/N/H residues identified as methylated in the LC-MS/MS data of each spot from charge train B were mapped onto the surface of the protein structure; mono-methyl = yellow, di-methyl = lilac, tri-methyl = orange, combination = dark purple, all other residues = teal. The models for each spot are identified by an arrow

### 5.2.4 Methylation at the ImpA active site

The number of methylated histidines identified was relatively low, consistent with their low abundance in the primary sequence. However, two histidine residues which are involved in the binding of zinc co-factor ions at the protease active site (H696 and H700) were found to be methylated in two spots (A4 and B5) (Figure 5.3). This PTM may be involved in the control of ImpA activity and may therefore have significant implications in the functioning of the protein in the extracellular environment. For example, methylation may be used to block the active site residues or co-factor binding to prevent off-target digestion of glycosylated proteins. Enzymatic de-methylation may then occur to alleviate this occlusion and allow controlled activity on specific substrates. The motif (HEXXH) which contains the methyl-histidines is highly conserved in zinc metalloproteases, and therefore the finding of methylation at these key residues may have further implications in other species (Hooper, 1994). This would be an interesting and important avenue for further investigation.



**Figure 5.3 Residues at the ImpA active site**

The crystal structure of ImpA (PDB: 5KDV) with a zoomed inset showing residue Q697 (highlighted in pink) which is a key active site residue, and H696 and H700 (highlighted in green) which are involved in zinc cofactor binding; the zinc ion is shown as a blue ball. All modelling was done using CCP4mg.

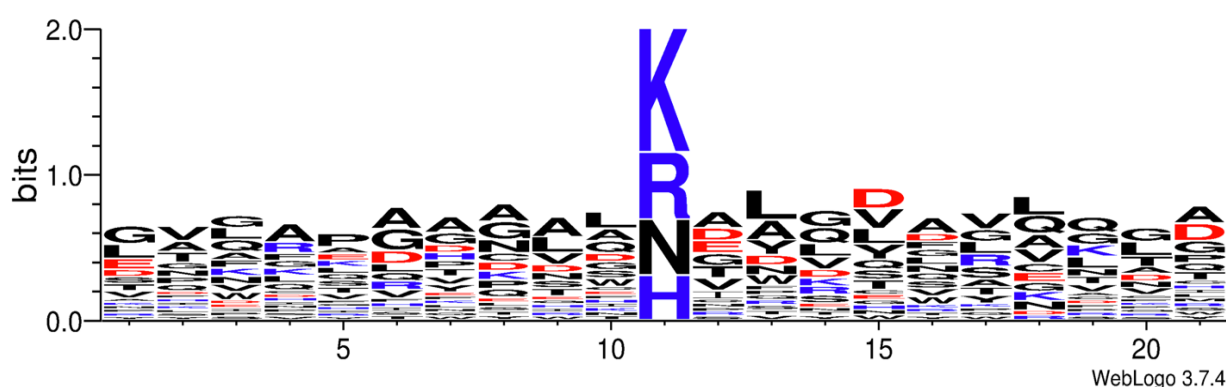
### 5.2.5 ImpA methylation motif and sequence conservation

In general, the residues found to undergo methylation in ImpA are not well conserved in homologues encoded by other species (see Section 4.4.6 for details of BLAST



analysis). However, the residues from N665 to K729 are very well conserved in all of the proteins that were aligned (see Appendix A). This range includes the residues within the vicinity of the active site which are likely to be key to protein function in multiple species. This gives greater weight to the notion that methylation may have a significant effect in this protein.

In eukaryotes, several protein methylation motifs are known, including RGG (and other glycine and arginine rich sequences), as well as PGM (proline, glycine, methionine) motifs (Guo *et al.*, 2014). In bacteria, methylation is far less studied and therefore no common motifs have been identified. Likewise, when the sequences ten residues up- and down-stream of the methylation site were aligned here, no clear consensus sequence for methylation was found (Figure 5.4). A recent study of *Salmonella* Typhimurium flagellar proteins also failed to identify a motif in these heavily lysine methylated proteins (Horstmann *et al.*, 2020). The researchers did, however, note the tendency to find small residues (such as alanine, glycine and valine) around the methyl-lysine, which is somewhat applicable here. The diversity and extent of methylation in ImpA is such that a methyl-determining secondary structure is highly unlikely.



**Figure 5.4 ImpA methylation motif analysis**

The frequency of residues surrounding every methylated K/R/N/H ( $\pm 10$  residues) were analysed using WebLogo; the size of the symbol correlates to its frequency at that position. The colours represent the charge of the amino acid, positive (blue), negative (red), or uncharged (black)

### 5.2.6 Methylation of an “unprocessed” ImpA sample

Previous literature has suggested that artificially induced *in vitro* methylation can occur during sample preparation for proteomic analysis through the use of chemicals such as methanol (Stewart *et al.*, 2016). Despite a lack of such methyl-containing chemicals being used in the analysis here, a spot was picked from a 2D gel run with a non-TCA precipitated, centrifugally concentrated “clean” secretome sample (Figure 3.2, Chapter 3). The spot picked was equivalent to the B5 spot from these analyses. This spot was also subjected to LC-MS/MS analysis and the data was searched to determine the degree of methylation. When compared to the spots generated from the TCA-precipitated samples, an even greater number of methylated residues were identified, and several previously identified methyl-residues were further confirmed (Table 5.4, in bold). The number of statistically significant spectral matches for this protein spot was 1175, likely contributing to the increased number of determined methylation sites. Additionally, this sample was analysed on a mass spectrometer with a greater mass accuracy and resolution. These results demonstrate that at least the majority, if not all, of the post-translational methylation shown here is derived from biological processes.

**Table 5.4 Comparison of ImpA methylation from centrifugal concentration**

Comparison of ImpA methylation from secretome samples concentrated through different experimental techniques: centrifugal concentration and TCA concentration (upper and lower charge train spots). Residues identified as methylated in both sample preparations are highlighted in **bold**.

Centrifugal		TCA – B5		TCA – A5	
N91	<b>H663</b>	R76	H696	H154	<b>R901</b>
K159	<b>N665</b>	H154	<b>K726</b>	<b>K176</b>	
R160	<b>N679</b>	K187	N903	<b>K179</b>	
<b>K176</b>	<b>R690</b>	<b>N192</b>		<b>N192</b>	
<b>K179</b>	H725	<b>K195</b>		<b>K195</b>	
<b>N192</b>	<b>K726</b>	<b>R316</b>		<b>N207</b>	
<b>K195</b>	<b>R729</b>	K350		<b>R316</b>	
<b>N207</b>	R756	<b>N357</b>		<b>R646</b>	
N255	K826	<b>K359</b>		<b>H663</b>	
<b>R316</b>	R837	K420		<b>N665</b>	
<b>N357</b>	N840	K601		<b>N679</b>	
<b>K359</b>	N852	<b>R646</b>		<b>R690</b>	
N632	N900	<b>H663</b>		H696	
K635	<b>R901</b>	<b>N665</b>		H700	
<b>R646</b>		<b>N679</b>		<b>R729</b>	

## 5.3 Methylation of LasB

### 5.3.1 Prediction of LasB methylation

Following on from the analysis of LasB acetylation, the methylation status was also investigated, starting with *in silico* prediction. When the primary sequence was submitted to MASA, PRmePRed and iMethyl-PseAAC, only one residue was predicted to be methylated by all three algorithms (R353). Once again, this suggests that these computational prediction algorithms are currently vastly inconsistent, lending doubt as to the validity of the predictions in a physiological context. Out of a total of 26 K/R residues, 14 were predicted to be methylation sites. Intriguingly, the lysine residues predicted to undergo methylation by iMethyl-PseAAC overlap strongly with those that were identified as acetylation sites by LC-MS/MS in Chapter 4.

**Table 5.5 Prediction of LasB methylation**

The amino acid sequence of LasB was submitted to several methylation prediction algorithms; MASA, PRmePRed and iMethyl-PseAAC. Residues predicted to undergo methylation by more than one algorithm are highlighted in **bold**

MASA	PRmePRed	iMethyl-PseAAC
<b>R353</b>	<b>R353</b>	K208
	R376	K211
	R395	K261
	R402	K282
	R405	K300
	R471	<b>R353</b>
		K378
		K388
		K443

### 5.3.2 MS/MS identification of LasB methylation

The methylation status of each LasB spot (D1-D4) was further determined experimentally by LC-MS/MS (Table 5.6). The results show that several residues are consistently methylated, and to a variable degree. This includes K211, R226 and K247, which were modified in all spots with mono-, di- and tri-methylations present. K211 and K247 were also found to be acetylated previously. Alternatively, certain residues such as K296 were consistently methylated to the same degree, in this case di-methylated. Once again, a respectable number of methyl-asparagine residues were identified, adding further weight to its importance in prokaryotic proteomes. Interestingly, no histidine residues were found to be methylated in this protein.

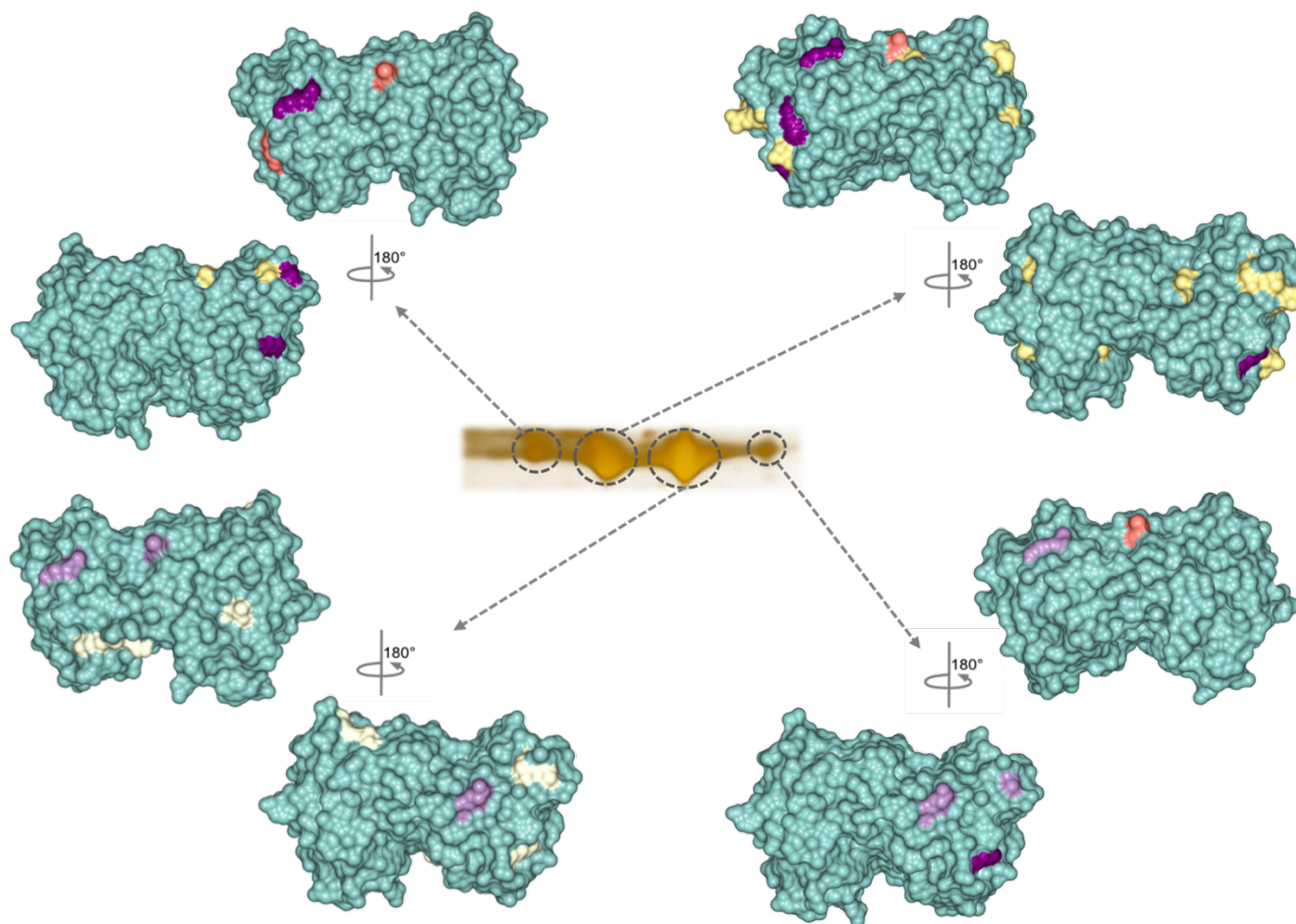
The use of an alternative mass spectrometer with a different mass accuracy and therefore resolution is likely to have affected the number of methylated residues identified in D3, in particular the number of asparagine residues. The mass tolerances allowed during secondary MASCOT searching were mistakenly kept constant for all spot samples, and therefore a higher level of stringency was used than recommended for spot D3 when asparagine methylation and di/tri-methyl K/R were added as variable modifications. Whilst the use of tighter mass tolerances causes many valid peptides to be missed during searching, it also increases discrimination and gives greater confidence in the peptides, and PTMs, that have been identified.

The study published recently by Gaviard et al also identified methylation of the LasB virulence factor (Gaviard *et al.*, 2019). Looking only at lysine residues, they found that K211, K378 and K443 are all modified by mono-, di- and tri-methylation. This is agreement with the findings here, except K443 was not identified during methylation analysis. By also including arginine and asparagine during data searches here, the number of known methylated residues of this protein have been extended. Together these results show that methylation is a clear and consistent PTM for this secreted protease.

**Table 5.6 Methylated K/R/N/H identified in spots D1 to D4**

**Bold** residues were identified in more than one spot. Me = mono-methyl, Me<sup>2</sup> = di-methyl, Me<sup>3</sup> = tri-methyl

D1			D2			D3			D4		
Me	Me <sup>2</sup>	Me <sup>3</sup>	Me	Me <sup>2</sup>	Me <sup>3</sup>	Me	Me <sup>2</sup>	Me <sup>3</sup>	Me	Me <sup>2</sup>	Me <sup>3</sup>
<b>N224</b>	<b>K211</b>	<b>K211</b>	K211	<b>K211</b>	<b>K211</b>	<b>N224</b>	<b>K211</b>			<b>K211</b>	<b>K247</b>
<b>R226</b>	<b>R226</b>	<b>K247</b>	<b>N224</b>	<b>K247</b>	<b>K247</b>	<b>R226</b>	<b>K296</b>			<b>R226</b>	<b>K378</b>
<b>N233</b>	<b>K247</b>	<b>K261</b>	<b>R226</b>	K261	<b>K261</b>	<b>K247</b>	K378			<b>K247</b>	
		<b>K378</b>	<b>N233</b>		<b>K378</b>	K378				<b>K296</b>	
			N240			K388					
			<b>K247</b>			K389					
			R252			R428					
			N360			R476					
			N379								
			N409								
			N461								
			R471								
			R485								



**Figure 5.5 Structural mapping of methylated residues identified in spots D1 to D4**

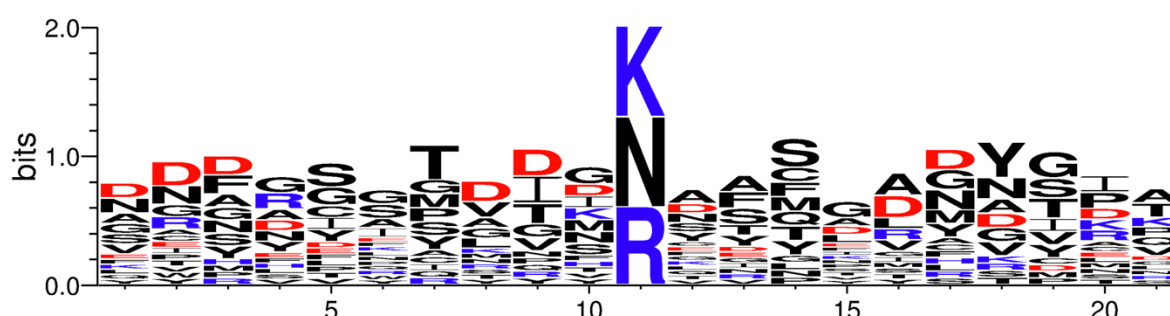
The 3D crystal structure of LasB (PDB: 1EZM) was analysed using CCP4mg software. The location of K/R/N/H residues identified as methylated in the LC-MS/MS data of each spot from charge train D were mapped onto the surface of the protein structure; mono-methyl = yellow, di-methyl = lilac, tri-methyl = orange, combination = dark purple, all other residues = teal. The models for each spot are identified by an arrow

### 5.3.3 Structural mapping of LasB methylation

The residues found to be methylated in LasB were mapped onto the 3D structure of the protein. All sites were found to have a relatively even distribution on the surface of the protein (Figure 5.5). None of the residues within the active site binding pocket were methylated. Residues N360 and K378/N379 mapped closely to the zinc and calcium binding sites respectively, however, upon further modelling, the modified side chains were not shown to interact with these ions.

### 5.3.4 LasB methylation motif and sequence conservation

When the LasB amino acid sequence was aligned with homologues from other species (see Section 4.5.6 for details of the BLAST analysis), it was surprising to note that asparagine was the most highly conserved among the methylated residues. Out of the 8 asparagines found to undergo methylation, 6 of them (75%) were conserved in all of the homologues (see Appendix A). This is in comparison to the lysine and arginine residues which were conserved at a rate of just 33% and 17% respectively. Furthermore, consensus sequence analysis proved inconclusive for specific motifs, although a prevalence of uncharged residues at positions -4 and +3 were clear (Figure 5.6). As with FljB and FliC in *S. Typhimurium*, aspartic acid (D) was found frequently around the methylation site (Horstmann *et al.*, 2020).



**Figure 5.6 LasB methylation motif analysis**

The frequency of residues surrounding every methylated K/R/N ( $\pm 10$  residues) were analysed using WebLogo; the size of the symbol correlates to its frequency at that position. The colours represent the charge of the amino acid, positive (blue), negative (red), or uncharged (black)



## 5.4 Methylation of PA0622

### 5.4.1 MS/MS identification of PA0622 methylation

A lack of consistency and accurate *in silico* prediction of methylation in ImpA and LasB discouraged the use of the algorithms for other proteins. Instead, attention was turned to the mass spectrometry identification of K/R/N/H methylation of the PA0622 protein spots (C1-C4), as detailed in Table 5.7. The results highlight many residues which were found to be consistently methylated in each spot, as well as several that undergo different degrees of methylation. Residues including K171, K176 and R246 fall into this category, suggesting that they are key, targeted modifications. Out of 58 methylatable residues in total, at least 16 were identified as being modified across the charge train. In general, most of the methylation occurred at lysine and arginine residues. Additionally, despite only encoding three histidines, one (H133) was methylated in two spots. An intriguing pattern emerged of more methylation sites becoming apparent with increasing pI of the protein isoform. This is surprising as methylation is not thought to modify the charge of these residues, although it is inconsistently cited to have an effect in some papers.

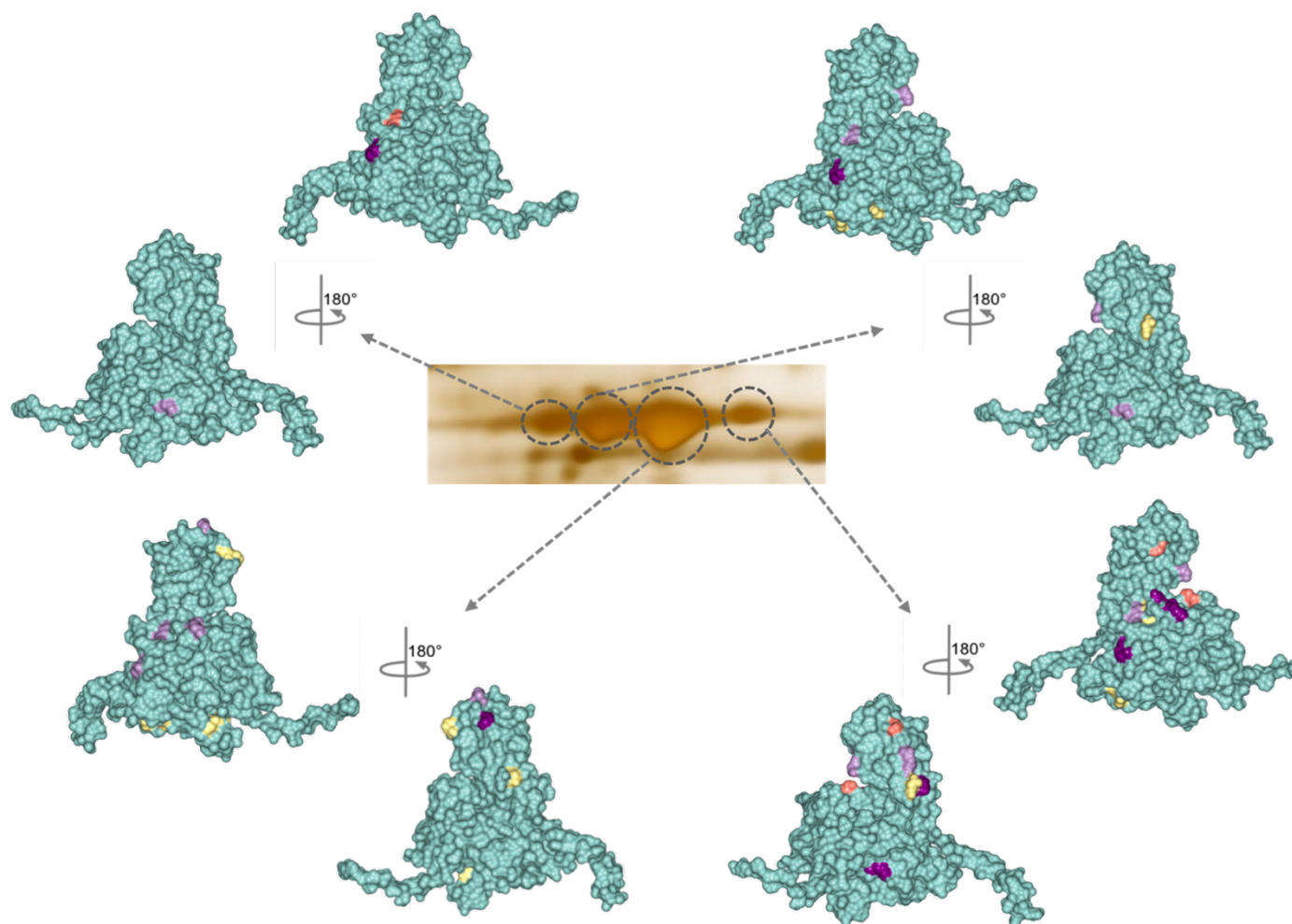
### 5.4.2 Structural mapping of PA0622 methylation

The methylation sites were mapped onto one subunit of the PA0622 structure (Figure 5.7). As before, all methylation sites were found to be at the surface of the structure. There appeared to be no preference for modification on either the core-facing or external-facing side of the protein, with several residues close to or at the interface between subunits, as with K-ac. This suggests that there may either be multiple roles for methylation of this protein (i.e. interaction with the core and stability of the sheath) or that some modification sites may be entirely redundant.

**Table 5.7 Methylated K/R/N/H identified in spots C1 to C4**

**Bold** residues were identified in more than one spot; spectra for starred\* residues contained insufficient information for discrepancy between mono-methyl K149 & R151 or di-methyl K149/1R151. Me = mono-methyl, Me<sup>2</sup> = di-methyl, Me<sup>3</sup> = tri-methyl

C1			C2			C3			C4		
Me	Me <sup>2</sup>	Me <sup>3</sup>	Me	Me <sup>2</sup>	Me <sup>3</sup>	Me	Me <sup>2</sup>	Me <sup>3</sup>	Me	Me <sup>2</sup>	Me <sup>3</sup>
N172	<b>K171</b>	<b>K171</b>	<b>H133</b>	<b>K171</b>	<b>K171</b>	<b>K149*</b>	<b>K149*</b>	<b>K312</b>	<b>H133</b>	<b>K149*</b>	K118
<b>R246</b>	<b>K227</b>	K176	N160	<b>K176</b>		<b>K151*</b>	<b>K151*</b>		<b>K149*</b>	<b>R151*</b>	<b>K171</b>
			<b>R246</b>	<b>K227</b>		R196	<b>K171</b>		<b>R151*</b>	<b>K171</b>	K227
			<b>N327</b>	<b>K299</b>		R234	<b>K176</b>		<b>R246</b>	<b>K176</b>	<b>K312</b>
						<b>R246</b>	K308		<b>N327</b>	<b>K227</b>	K326
						R304				<b>K299</b>	K352
						K312				R323	
						K326				K326	

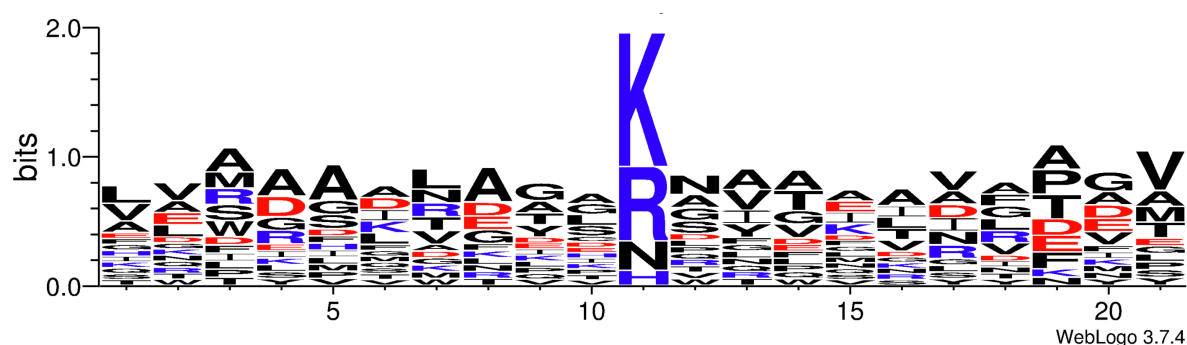


**Figure 5.7 Structural mapping of methylated residues identified in spots C1 to C4**

The likely structure of PA0622 was predicted using I-TASSER and then analysed using CCP4mg software. The location of K/R/N/H residues identified as methylated in the LC-MS/MS data of each spot from charge train C were mapped onto the surface of the protein structure; mono-methyl = yellow, di-methyl = lilac, tri-methyl = orange, combination = dark purple, all other residues = teal. The models for each spot are identified by an arrow

### 5.4.3 PA0622 methylation motif and sequence conservation

Upon alignment of the PA0622 sequence with similar proteins from other species (see Section 4.6.6 for details of BLAST analysis), it was found that all methylated residues were conserved by at least three species (including *K. pneumoniae*), with the majority conserved between all five proteins (see Appendix A). It would be interesting to determine whether PTM of these residues is also consistent across species. In agreement with findings by Horstmann et al, the locations surrounding the methylation site in PA0622 (Horstmann *et al.*, 2020) appear more populated by small amino acids such as alanine, valine and, to a lesser extent, threonine and glycine. Besides this, no clear consensus sequence was determined for methylation.



**Figure 5.8 PA0622 methylation motif analysis**

The frequency of residues surrounding every methylated lysine ( $\pm 10$  residues) were analysed using WebLogo; the size of the symbol correlates to its frequency at that position. The colours represent the charge of the amino acid, positive (blue), negative (red), or uncharged (black)

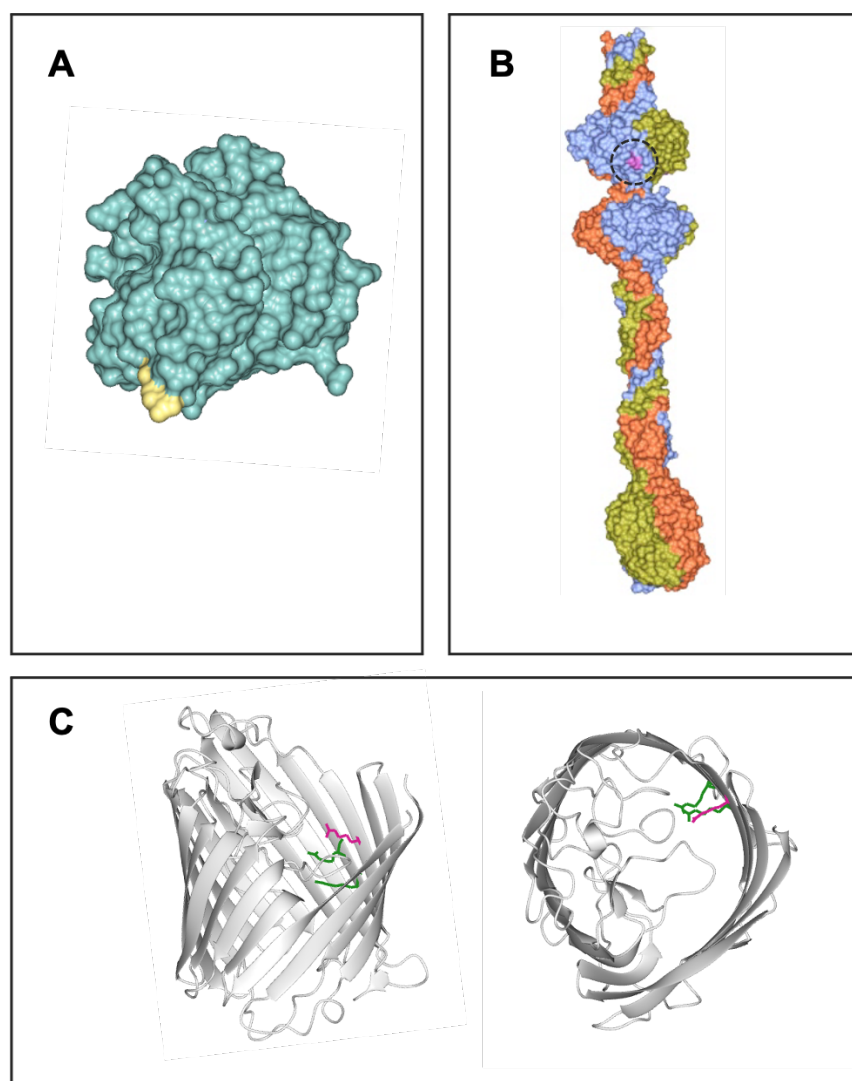
## 5.5 Methylation of other secreted proteins

In addition to the PTMs of the proteins outlined so far, the methylation status of several other extracellular proteins was also investigated. Firstly, the secreted protease LasA (Spot E1, Figure 4.2) was also found to be methylated at the surface exposed arginine R393 (Figure 5.9A). This protease enhances the elastinolytic activity of LasB as well as being a potent staphylolysin (Braun *et al.*, 1998). Sequence coverage of LasA was very low during mass spectrometry analysis (around 17%) and therefore several other PTM events were likely to have gone undetected. In addition, due to its highly basic

pI, any protein charge isoforms were not able to be separated at the pH ranges used, prohibiting their characterisation.

Another protein that forms part of the complex R2 pyocin particle was also identified as a methylation substrate (Spot G1, Figure 4.2). PA0620, which makes up the host-binding tail fibre, is methylated at R389. Each tail fibre is made up of a homotrimer of PA0620 proteins, of which the C-terminal domain determines target cell specificity through differential LPS recognition and binding (Buth *et al.*, 2018). The methylated R389 is found on the protein surface closer to the N-terminal base-plate binding domain of the tail fibre, rather than the highly variable C-terminal end (Figure 5.9B). It is not yet clear what the purpose of this modification is.

The outer membrane porin D (OprD) was identified in a train of three spots in the 2D gel (Spots F1-F3, Figure 4.2). The monomeric 18-stranded  $\beta$ -barrel structure sits in the outer membrane and acts as a channel for small molecule uptake. In addition to playing a major role in resistance to carbapenem antibiotics, OprD can also act as a serine protease (Li *et al.*, 2012). The LC-MS/MS data identified methylation of R433 in spot F1. When mapped onto the 3D structure, the arginine residue points into the lumen of the porin (Figure 5.9C). Significantly, this residue (R410 in the mature protein) constitutes part of a “basic ladder” of arginine residues within the barrel wall that forms an “electrophoretic conduit” capable of binding substrate carboxylate groups (Eren *et al.*, 2012). Mutation of this residue significantly decreases substrate passage, rendering it indispensable for porin activity (Samsudin and Khalid, 2019). It is also highly conserved in all OprD family channels. This porin residue is therefore vital to *P. aeruginosa*’s ability to thrive in nutrient poor environments and to prevent antibiotic uptake. Accordingly, it would be an important avenue of further study to determine what the role of methylation at this site may be. If, for example, it is involved in the modulation of substrate binding, this may be of key interest in future drug design.



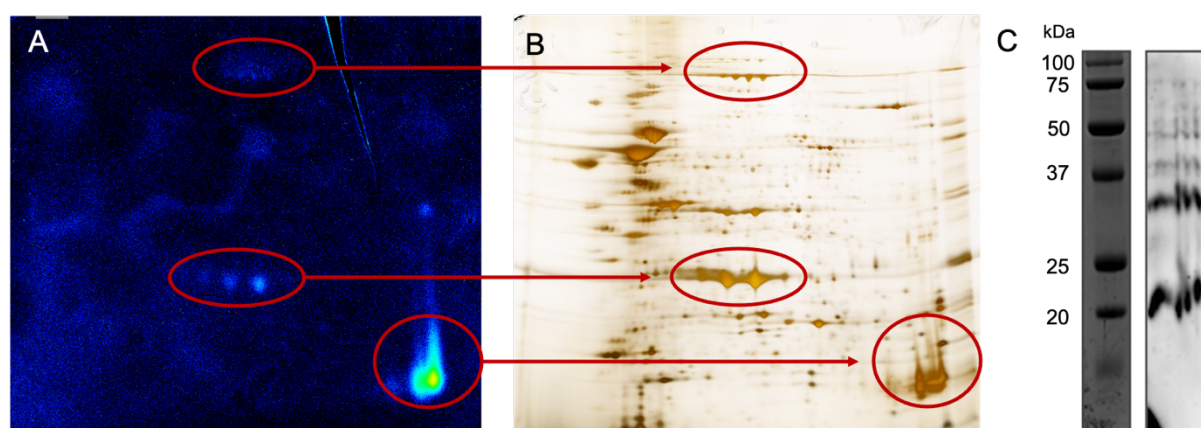
**Figure 5.9 Methylation of other PAO1 secreted proteins**

Several other spots on the PAO1 secretome 2D gel were analysed by LC-MS/MS analysis. The identified methylation sites were then mapped onto the 3D structures of these proteins using molecular modelling programme CCP4mg. (A) The location of methyl R93 (shown in yellow) on the surface of LasA (PDB: 3IT7), (B) The location of methyl R389 (shown in pink and encircled) on the PA0622 homotrimer of the R2 pyocin tail fibre (PDB: 6CL6), (C) The 3D structure of OprD (PDB: 3SY7) with methyl R443/410 (shown in pink) and two other major residues (R389 and R391) of the “basic ladder” (shown in green) pointing into the lumen, as viewed from the side and top

## 5.6 Immunoaffinity confirmation of methylation

To confirm that PAO1 secreted proteoforms are indeed methylated, an immunoaffinity probe of the 2D gel was performed. Due to the size of the 2D gel and consequent problems with membrane transfer, an in-gel Western approach was taken. Following

completion of the second dimension separation of the secretome, the gel was fixed and incubated with anti-mono methyl arginine (MMR) antibodies. Use of a near-infrared secondary antibody then allowed direct scanning of the gel and visualisation of methylated proteins. This technique also allowed comparison with a parallel gel visualised with silver stain (Figure 5.10 A & B).



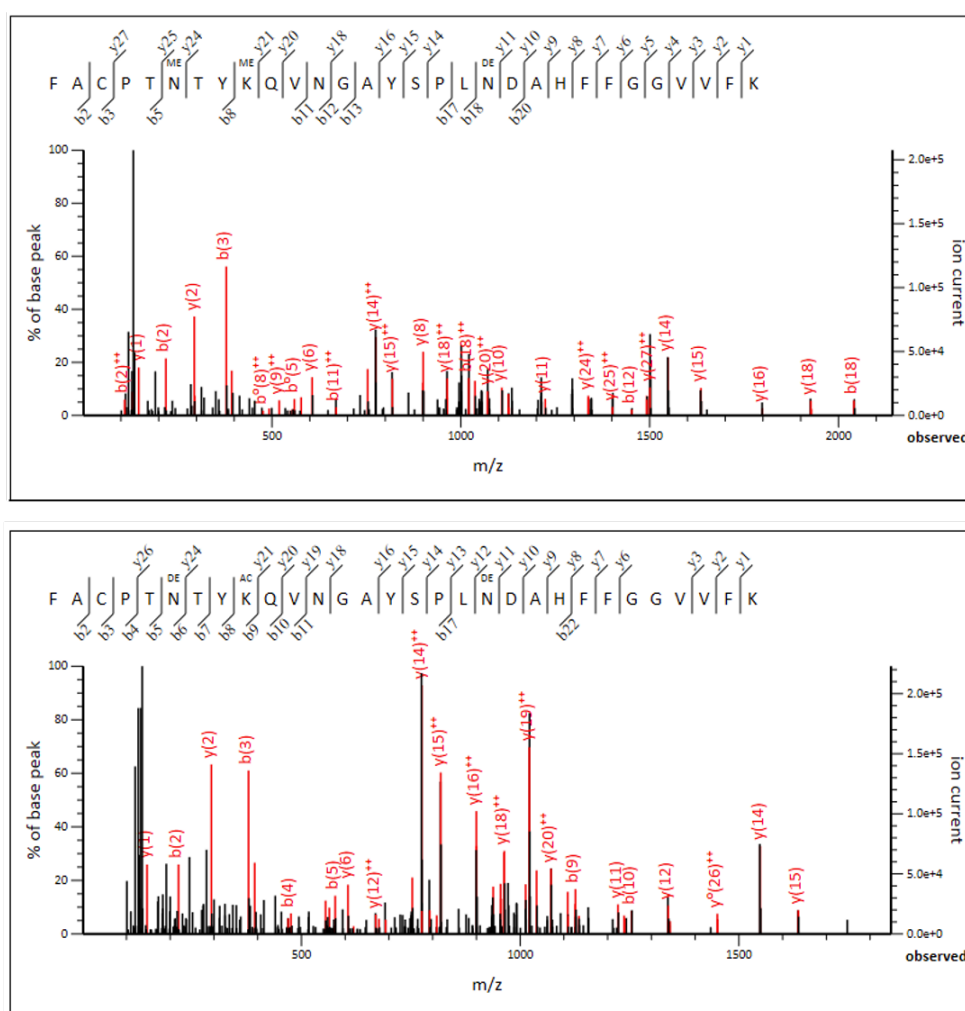
**Figure 5.10 Immunoaffinity confirmation of methylation**

A PAO1 secretome sample was run on parallel 2D gels which were either (A) probed directly with anti mono-methyl arginine (MMR) antibodies or (B) visualised with silver staining. Red circles and arrows indicate corresponding spots on each gel. (C) Western blot of the secretome separated on a 1D SDS-PAGE gel and subsequently transferred to PVDF and blotted with MMR antibodies (distortion in the bands is due to the high urea content of the sample buffer)

One drawback of direct antibody probing of a gel is the loss of sensitivity due to reduced antibody binding to proteins within the gel matrix. Less abundant proteins that form faint or small spots are therefore likely to go undetected. As shown in Figure 5.10, the charge train formed of LasB proteoforms was detected by the methyl-arginine antibodies, confirming the LC-MS/MS data identifying this PTM. Faint spots corresponding to the ImpA charge train are also evident. Surprisingly, the strongest signal came from the LasA spot(s), which was only identified to have one methylated arginine during mass spectrometry analysis. Saturation of this spot during scanning may also be the cause of fewer proteins being visualised. When the secretome was also run on a 1D SDS-PAGE gel, the bands for LasA (21 kDa) and LasB (33 kDa) showed a strong signal when detected with the MMR antibody following transfer to a PVDF membrane (Figure 5.10C). This suggests that there may potentially be more methylated arginines present in LasA than previously identified.

## 5.7 Overview of PTM overlap

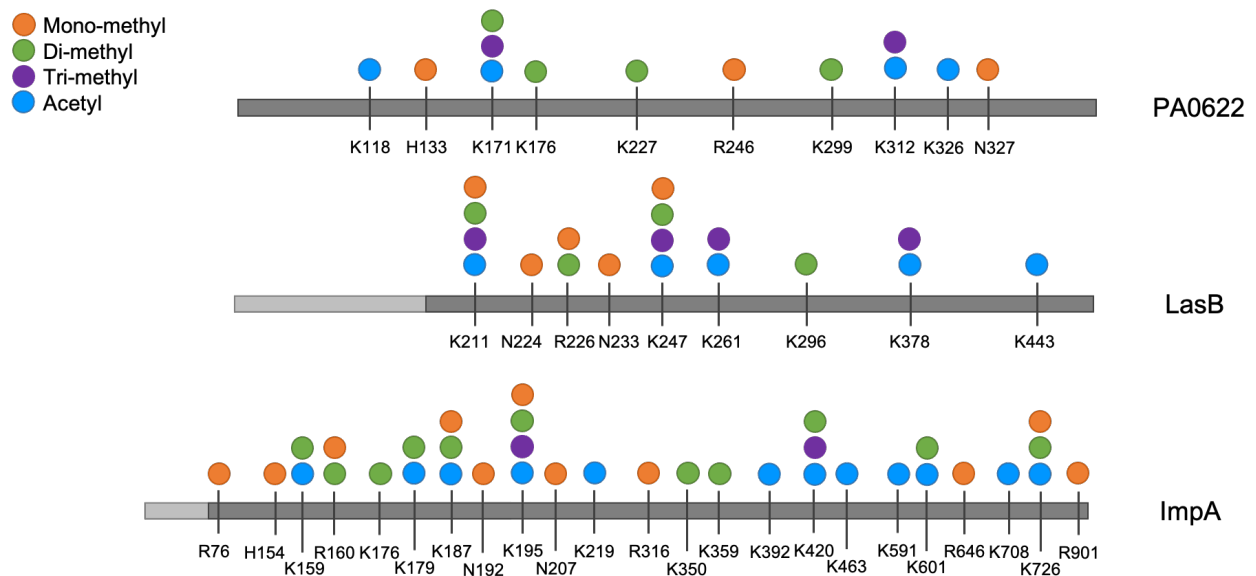
What has become evident during the course of this research is that several *P. aeruginosa* secreted factors are multi-modified by acetylation and methylation. In fact, there is evidence of significant overlap between the sites of these PTMs, suggesting that their biological application is a highly dynamic process (Figure 5.11 and Figure 5.12). In agreement with previous findings (Ouidir *et al.*, 2014), several serine, threonine and tyrosine residues were also found to be phosphorylated in these proteins (data not shown). Since lysine and arginine can also undergo phosphorylation, the true complexity of the PTM overlap is yet to be fully determined.



**Figure 5.11 Evidence of overlap of PTM at the same residue**

Two MS/MS spectra showing that LasB undergoes both methylation and acetylation at the same residue (K261) in the extracellular environment. Comparison of the graphs shows that the peptide 'FACPTNTYKQVNGAYSPLNDAAHFFGGVVFK' of LasB is both methylated (ME) and acetylated (AC) at K9 in the same sample. This indicates that PTM is both dynamic and inconsistent within a protein population.





**Figure 5.12 Overview of the PTMs identified for three *P. aeruginosa* secreted proteins**

PTMs found at different residues of the three main proteins (PA0622, LasB and ImpA) are shown by coloured circles; blue = acetylation, orange = mono-methylation, green = di-methylation and purple = tri-methylation. Only PTMs detected in more than one spot sample are depicted. This shows that there is significant overlap between the sites of different types of PTM on a protein.

## 5.8 Discussion

Following the identification of extensive acetylation of several virulence factors as described in Chapter 4, it was questioned whether other small PTMs are present that may have important roles in protein function. Considering the consistency in molecular weight of the charge isoforms (at the resolution afforded by 2-DE), methylation was selected for further investigation as its application only confers a modest increase in the molecular weight of a protein.

Both charge trains encompassing ImpA isoforms demonstrated widespread methylation with a significant overlap of methylation sites between spots. Certain residues, such as K195, were found to undergo multiple degrees of methylation (i.e. up to tri-methylation), as well as acetylation (Figure 5.12). Furthermore, certain residues, such as K420, were consistently modified in every spot, suggesting that such modifications are highly abundant within the protein population and are therefore

biologically significant. When the methylation sites were mapped onto the 3D surface of the protein, there was a dominance of sites on one end of the protein. This may be an important domain for functional protein-protein interactions, which are known to be mediated by PTMs such as methylation; experimental validation of this is warranted (Winter *et al.*, 2015). The role of PTM of non-surface expressed multi-modified residues, such as K726, should also be investigated. It was further shown that this methylation is indeed biologically driven and is unlikely to be artifactually introduced during sample preparation and separation.

When considering the methylation overlap between ImpA charge trains, there appears to be a slight pI shift to the acidic range for the upper charge train (or vice versa with the lower) that would mean better matching of methylation sites between spots. To explain this more clearly, the methylation sites identified in spot A4 appear to overlap better with those of B5 than B4, despite these spots migrating to different pIs. This could be explained by any potential truncations or additional complex PTMs (e.g. glycosylation) that cause a shift in pI as well as the evident molecular weight shift. It would be of keen interest to further characterise the molecular basis for this observation.

Arguably the most important finding from this set of data was the identification of methylation of the ImpA zinc binding histidine residues H696 and H700. As a zinc metalloprotease, ImpA requires the binding of zinc cofactors for proper functioning. Other studies in *Vibrio anguillarum* and *Bacteroides fragilis* found that mutation of these highly conserved histidine residues resulted in almost complete loss of proteolytic activity (Franco *et al.*, 2005; Yang *et al.*, 2007). Considering that ImpA acts on glycosylated substrates, it is hypothesised that methylation of the cofactor binding site may exert tight control over protease activity to prevent the digestion of off-target substrates, i.e. cleavage of glycosylated self-proteins such as flagella.

LasB, and its co-protease LasA, are also methylated on surface exposed residues. As with ImpA, a significant overlap between spots was seen and certain residues (K211 and K247) were shown to undergo all types of PTM investigated. In agreement with my findings, Gaviard *et al.* also showed that K211 and K378 are mono-, di- and trimethylated in PA14 (Gaviard *et al.*, 2019). Conversely, they also found multi-methylation of K443, whereas it was only identified as acetylated in PAO1 here. This

discrepancy is likely due to either strain or experimental differences. The role of PTM of several key modified residues will be explored further in Chapter 6.

Surprisingly, no histidine residues were found to be methylated, despite LasB also harbouring the conserved zinc metalloprotease motif found in ImpA (HEXXH). LasB is involved in the proteolytic activation of several other secreted proteins (Braun *et al.*, 1998; Li and Lee, 2019), and so may be somewhat constitutively “on” in the extracellular environment through the lack of methylation at the catalytic domain. Unfortunately, the very low sequence coverage of LasA obtained during mass spectrometry analysis will have prevented other methylation sites from being identified. Considering the strong signal obtained when the LasA spot was probed with anti-MMR antibodies (Figure 5.10), it is likely that several other residues are also methylated.

The two proteins that form the R2 pyocin sheath and tail fibres are also surface methylated. This is the first time that PTM of a tailocin has been demonstrated, despite morphologically similar homologues being expressed by many different bacterial species (Ghequire and De Mot, 2015; Ghequire *et al.*, 2015). As with the acetylation sites, many of the PA0622 methyl-residues are in close contact with other protein subunits, as well as the internal spear-like core structure. This may suggest that this PTM could have a role in protein stability or pyocin complex formation. Methylation of the tail fibre does not appear to occur at the highly variable target-determining end, further supporting the idea it is involved in the considerable *in vivo* stability of the pyocin macromolecular structure. Pyocins have been proposed as potential therapeutics for the targeted control of bacterial populations (Príncipe *et al.*, 2018). This is due to their striking potency whereby one particle is sufficient to kill a target cell (Scholl and Martin, 2008). As such, they have proven efficacy in several *in vivo* models (Gebhart *et al.*, 2015; Redero *et al.*, 2020). Further insight into pyocin assembly and function is therefore required, including further elucidation as to the role of the identified PTMs.

One of the most surprising, and potentially significant, findings of this study is the methylation of a key residue within the outer membrane porin D (OprD or OccD1) substrate recognition and binding site. R410 is part of a highly conserved basic ladder close to the porin constriction site and has been shown to bind substrates, including

positively charged amino acids, peptides and carbapenem antibiotics (Eren *et al.*, 2012, 2013). As a mechanism of carbapenem resistance, *P. aeruginosa* can downregulate or mutationally inactivate OprD to reduce membrane permeability (Li *et al.*, 2012). Reversible methylation may therefore serve as an alternative method of rapid and dynamic control, i.e., a post-translationally controlled on/off switch. This would also resolve any fitness costs associated with OprD loss.

The pore for OprD is substantially narrower than in other bacterial porins and may be further occluded by methylation, particularly as R410 sits at the narrowest “eyelet region” of the pore (Eren *et al.*, 2012; Parkin and Khalid, 2014). Indeed, other residues at the constriction site may also undergo methylation but were not identified during MS/MS analysis. Likewise, methylation may also be used as a method to control the rate of other substrate binding. It is unclear whether this would be through the inhibition of binding as a result of the disruption of molecular interactions or an increase in binding as a result of a larger charge distribution. Several other residues of OprD were found to be acetylated as detailed in Chapter 4, however the significance of this is unclear from their location on the protein. Ultimately, understanding the underlying biology of this channel could lead to the improvement of rational antibiotic design as a means of increasing drug permeation into the cell (Isabella *et al.*, 2015).

The lack of identification of a methylation sequence motif for any of the proteins is not surprising. Methyltransferases are commonly quite specific, particularly with regards to the target residue. Protein lysine methyltransferases (PKMTs) and protein arginine methyltransferases (PRMTs) are distinct types of enzyme. Their targeted activity is therefore likely to be signalled by different motifs, if present. These methyltransferase subtypes are further specified even to the single residue level. For example, EftM of *P. aeruginosa* is a thermo-regulated methyltransferase that specifically catalyses the tri-methylation of EF-Tu at K5 (Owings *et al.*, 2016; Prezioso *et al.*, 2019). Future work should therefore attempt to elucidate the modifying enzymes responsible for the PTMs identified here and further determine their substrate specificities and regulation.

This research has focused on the methylation of lysine, arginine, asparagine and histidine. Asparagine and histidine methylation are not well characterised and as such their biological relevance is not well understood. For all of the three main proteins

analysed, asparagine methylation was evident, as well as histidine methylation in ImpA and PA0622. This suggests that such an “atypical” PTM may actually be more typical in prokaryotes than is currently appreciated. This notion is further supported by recent findings in *E. coli* whereby the level of methylation of asparagine was on a par with that of lysine methylation; these findings were not replicated in the eukaryotic data set (M. Zhang *et al.*, 2018). To further advance the results presented here, a wider range of methyl-accepting amino acids should be investigated, including glutamine (Q), glutamic acid (E) and aspartic acid (D), to determine their relevance to *P. aeruginosa* secreted proteins. O-methylation of aspartic and glutamic acid neutralises their negative charge, thus having an effect of isoform pI (Sprung *et al.*, 2008; Lengqvist *et al.*, 2011). If present, they would therefore need to be incorporated into a complex model aimed at understanding charge train formation.

Furthermore, a greater range of methylation substrates should be investigated. An experiment was planned to explore this using immunoaffinity enrichment of secreted proteins with pan-methyl lysine antibodies. Mass spectrometry analysis of the enriched sample would have provided the first lysine-methylome analysis of *P. aeruginosa* proteins. The experiment was unfortunately prohibitively delayed due to the SARS-CoV-2 outbreak and was not able to be completed. It is hoped that this can be completed in the future to both confirm the methylation sites identified here, as well as provide a more global overview of methylation in the extracellular fraction.

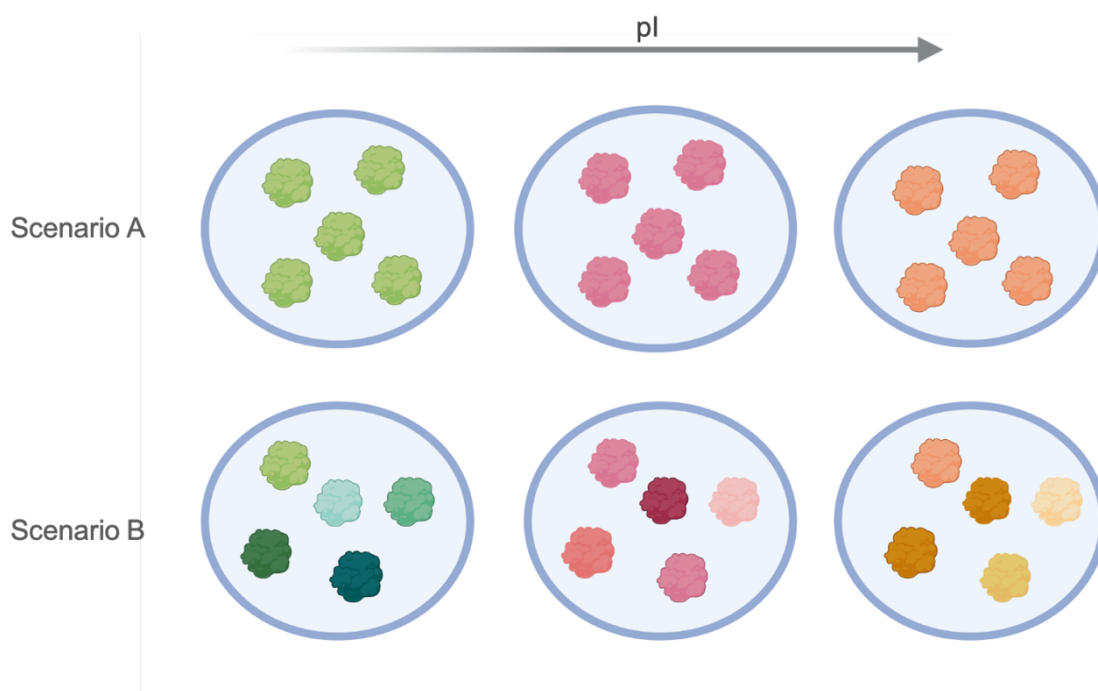
The comprehensive level of PTM identified is likely to have significant implications in the functioning of these proteins, and work must be done to determine their true biological role. Methylation is one of the most functionally versatile forms of PTM and has been shown to be essential for bacterial pathogenicity. For example, deletion of the methyltransferase VagH in *Yersinia pseudotuberculosis* prohibits secretion of the virulence determinant YopD, rendering it avirulent in an *in vivo* model (Garbom *et al.*, 2007). In applying these findings to the data presented here, methylation could potentially serve as a signal for the secretion of these virulence factors. However, due to the diverse functionality and varying degrees of methylation, as well as the overlap with acetylation sites (and potentially other PTMs), experiments into the role of PTM on individual proteins must be conducted rather than using *a priori* deduction.

The considerable overlap between acetylation and methylation sites may be an indication of reciprocal regulation or PTM crosstalk. For example, methylation may act to block the acetylation of certain lysine residues, in much the same way that methylation has been shown to block the ubiquitination of surface-expressed proteins (Pang *et al.*, 2010; Engström *et al.*, 2020). As a result, PTMs should be considered as integrated and co-regulated systems rather than separate entities (Lothrop *et al.*, 2013). This method of assessment was demonstrated by Lu *et al.*, who showed that *in silico* mutation of acetylation sites had an impact on the phosphorylation, methylation and ubiquitination status of a protein (Lu *et al.*, 2011). Such interdependencies between PTMs, even on distant regions on the protein or between proteins, have also been demonstrated experimentally (Lothrop *et al.*, 2013).

The initial aims of this research project were to determine a pattern of small charge modifying PTMs of extracellular proteins that could explain the charge train phenomenon. Acetylation unequivocally fits within this category, as does phosphorylation (which has been investigated in detail elsewhere) (Ouidir *et al.*, 2014). Whether K/R methylation also affects charge was a matter of initial confusion due to many conflicting statements in the literature that either supported a shift in pI (Bio-Rad; Frye *et al.*, 2006; Stewart *et al.*, 2016), or argued against it (Stallcup, 2001; Luo, 2018). The majority side with the latter, and so K/R methylation cannot be used to answer the original question posed. Nonetheless it is an important PTM, and extrapolation of these findings to include E/D methylation may help answer this question. Enigmatically, however, it is the combination of all charge modifying PTMs that ultimately decides the overall charge.

Inconsistencies in PTM site identification between samples are a common issue (Stewart *et al.*, 2016; Gaviard *et al.*, 2019). Figure 5.12 therefore only highlights the residues that were confirmed to be modified in more than one spot on the 2D gel. This does not, however, render those only identified in one spot invalid. They may form a group of rarer modifications, or alternatively were just missed during MS/MS analysis for some spots as the number of modified peptides identified tends to be proportional to the number of spectral *m/z* ratios used for database searches (Eshghi *et al.*, 2012). It cannot be ruled out that some of the overlap between acetylation and tri-methylation is a result of misidentification during MASCOT searching. The mass change from

these PTMs is very similar (+42.0106 for acetyl, +42.0470 for tri-methyl) and therefore any slight mass errors in detection may cause misidentification. However, not all K-ac and K-me<sup>3</sup> sites overlapped, indicating that the experimental sensitivity is sufficient enough for such discrepancy and PTM identification to be valid in the majority of cases.



**Figure 5.13 Change of thinking with respect to protein charge variation**

Initially, it was thought that each protein spot on a 2D gel contained one protein isoform with a set portfolio of PTMs (Scenario A). LC-MS/MS data showed many unmodified peptides were also sequenced alongside differentially modified peptides when a single spot was analysed. This therefore supports Scenario B whereby 2D gel spots contain multiple different protein isoforms that can be grouped, and separate together, according to their overall charge.

Until now, the spots within a charge train have been referred to as individual proteoforms. However, the evidence now suggests that multiple proteoforms are in fact present even within a single protein spot of a 2D gel (Figure 5.13). This complexity makes the characterisation of distinct proteoforms, as well as the determination of exactly what causes the charge variation seen, incredibly challenging. This is compounded by the fact that not all PTM sites are typically uncovered during LC-MS/MS analysis, and not all PTMs are commonly searched for in the data. Indeed, to retain consistency of molecular weight across a charge train, there must be some uniformity in the number of certain charge modifying PTMs within each spot. Without

more resolute data detailing all PTMs and their specific stoichiometries on a protein, alongside computational modelling of their combinatorial effects on pI, deduction of the cause of charge variation is unrealistic. Furthermore, the effect of PTM on pI is different for each protein, and therefore each charge train should be modelled individually.

### **5.8.1 Conclusion**

The results overall showed that methylation is another PTM that is widely and consistently found on the surface of *P. aeruginosa* secreted proteins. Furthermore, some of the methylated residues are known to have key roles in the functioning of the proteins, such as catalytic domains and substrate recognition sites. These results therefore potentially have significant implications for a deeper understanding of how these proteins function and for the realisation of possible drug targets to fight infection and improve chemotherapeutic outcomes.



## Chapter 6

### 6 Initial insights into the functional consequences of PTM

#### 6.1 Introduction

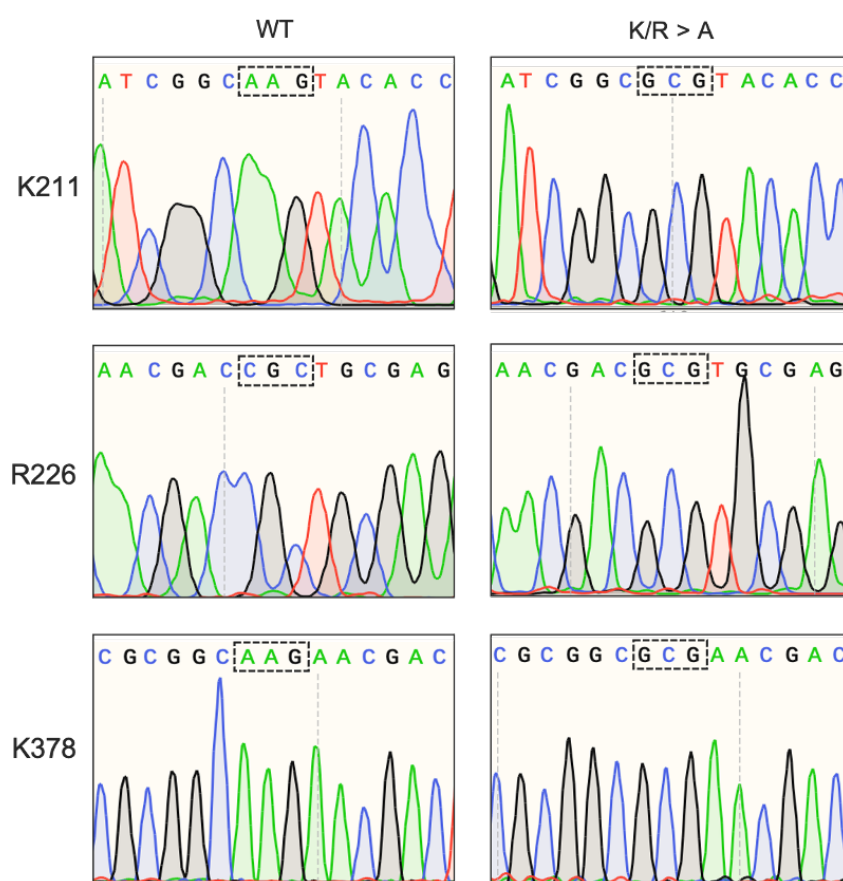
Elastase, or LasB, is a principal virulence factor of *P. aeruginosa* that contributes significantly to both the invasiveness and severity of infection (Galdino, Viganor, *et al.*, 2019; Li *et al.*, 2019). This is partly due to its involvement in the activation of, and synergistic activity with, several other virulence factors. It also has direct activity as a potent protease. Its broad substrate specificity facilitates rapid host colonisation by causing extensive tissue damage and weakening of the immune response. As such, the expression of LasB is consistent with a range of virulence-associated phenotypes, including proteolytic and staphylolytic activity, increased swarming motility and biofilm formation, and advanced pathogenicity in *in vivo* models (Galdino, de Oliveira, *et al.*, 2019).

##### 6.1.1 Aims

As highlighted in the previous chapters, LasB and PA0622 undergo extensive and consistent modification by methylation and acetylation. These PTMs can have major consequences on a protein's stability, location, activity and interactions. The aims in this chapter were therefore to undertake a preliminary investigation of the effects of mutating key modifiable residues within the LasB and, preliminarily, PA0622 sequences to provide initial insights into the role that methylation and acetylation may have in the functioning of these proteins.

## 6.2 Site-directed mutagenesis of key residues

To prefatorily investigate the role that methylation and/or acetylation might be playing in the functioning of LasB, site-directed mutagenesis (SDM) of key modified residues was undertaken. The rationale behind which residues were selected for mutagenesis lies in i) the most consistently modified residues across the charge train, ii) those also found in PA14 (Gaviard *et al.*, 2019) and iii) inclusion of both lysine and arginine residues. It was therefore decided to substitute K211, R226 and K378 for alanine. Alanine scanning mutagenesis is a widely used technique that is employed to explore the roles and stability of specific protein residues, or their PTMs (Suriyanarayanan *et al.*, 2016; Sakatos *et al.*, 2018). Substitution of most canonical amino acids to alanine maintains the preferred secondary structure but removes the reactive side chains.



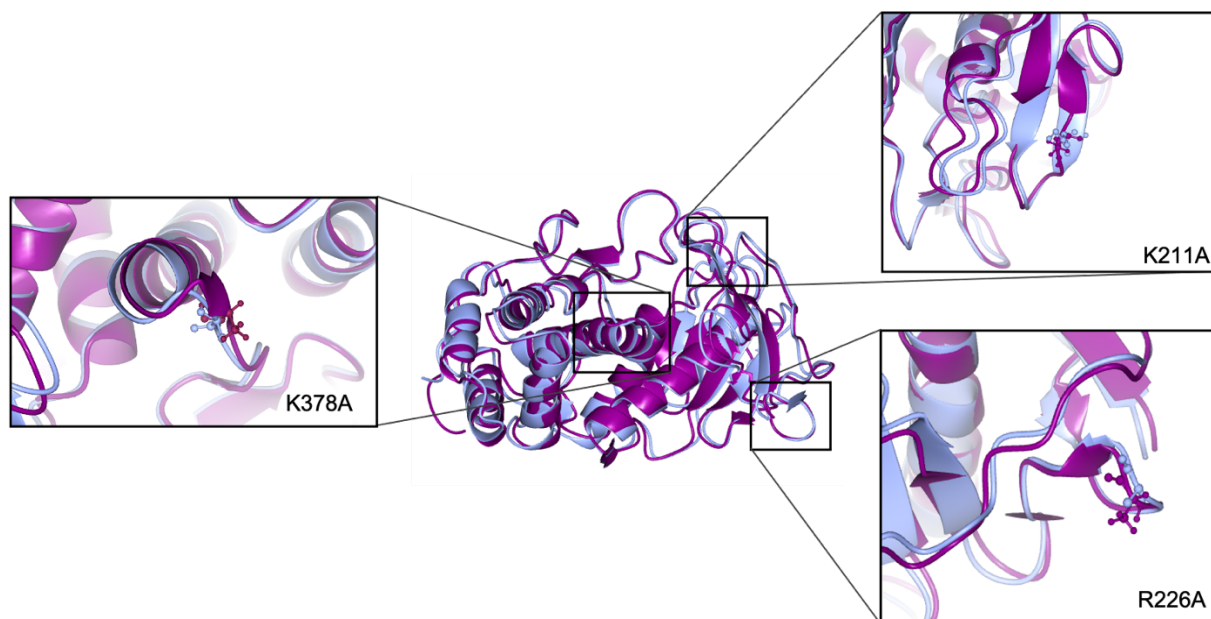
**Figure 6.1 Confirmation of LasB site-directed mutagenesis**

Sanger sequencing chromatogram confirmation of LasB WT and alanine substituted site-directed mutant constructs: K211A, R226A, K378A. The WT and mutated genes were encoded in pUCP20 as an expression vector.

The University of Washington PAO1 transposon mutant library was also utilised (Jacobs *et al.*, 2003). PW7302 hosts a transposon containing a tetracycline resistance cassette within the *lasB* ORF. This causes insertional inactivation of the gene and thus disruption of the protein. The presence of the ISphoA/hah transposon within the gene was confirmed by PCR. The wild-type *lasB* gene and its endogenous Shine-Dalgarno sequence was inserted into the pUCP20 plasmid and transformed into PW7302 for wild-type gene complementation *in trans*, generating LasB-WT. This construct was then used for SDM of the WT gene to substitute the codons for K211, R226 and K378 for the alanine codon GCG. Each SDM construct was validated by Sanger sequencing (Figure 6.1) and subsequently introduced independently into PW7302 alongside an empty vector control, generating LasB-K211A, LasB-R226A, LasB-K378A and LasB-EV, respectively.

### **6.2.1 Effect of LasB SDM on protein structure and pI**

The predicted effect of alanine substitution on the structure and isoelectric point of the protein was computationally interrogated to ensure that there were no major structural disruptions. The mutated primary sequences were submitted to I-TASSER for structural prediction and the outputs were superposed onto the X-ray defined structure of LasB (PDB: 1EZM) using the molecular imaging programme CCP4mg (Figure 6.2). As expected, substitution of the lysines or arginine did not have any major effects on protein structure, and the local secondary structure was generally maintained. A lack of structural damage was also confirmed using Missense3D (Ittisoponpisan *et al.*, 2019). Any off-target structural differences are arguably attributable to the use of an alternative PDB file as the top threading model in structural prediction, instead of 1EZM. In addition, the pI predictor Prot pi (<https://www.protpi.ch/>) showed that loss of the positively charged side chains only shifted the pI by -0.181.



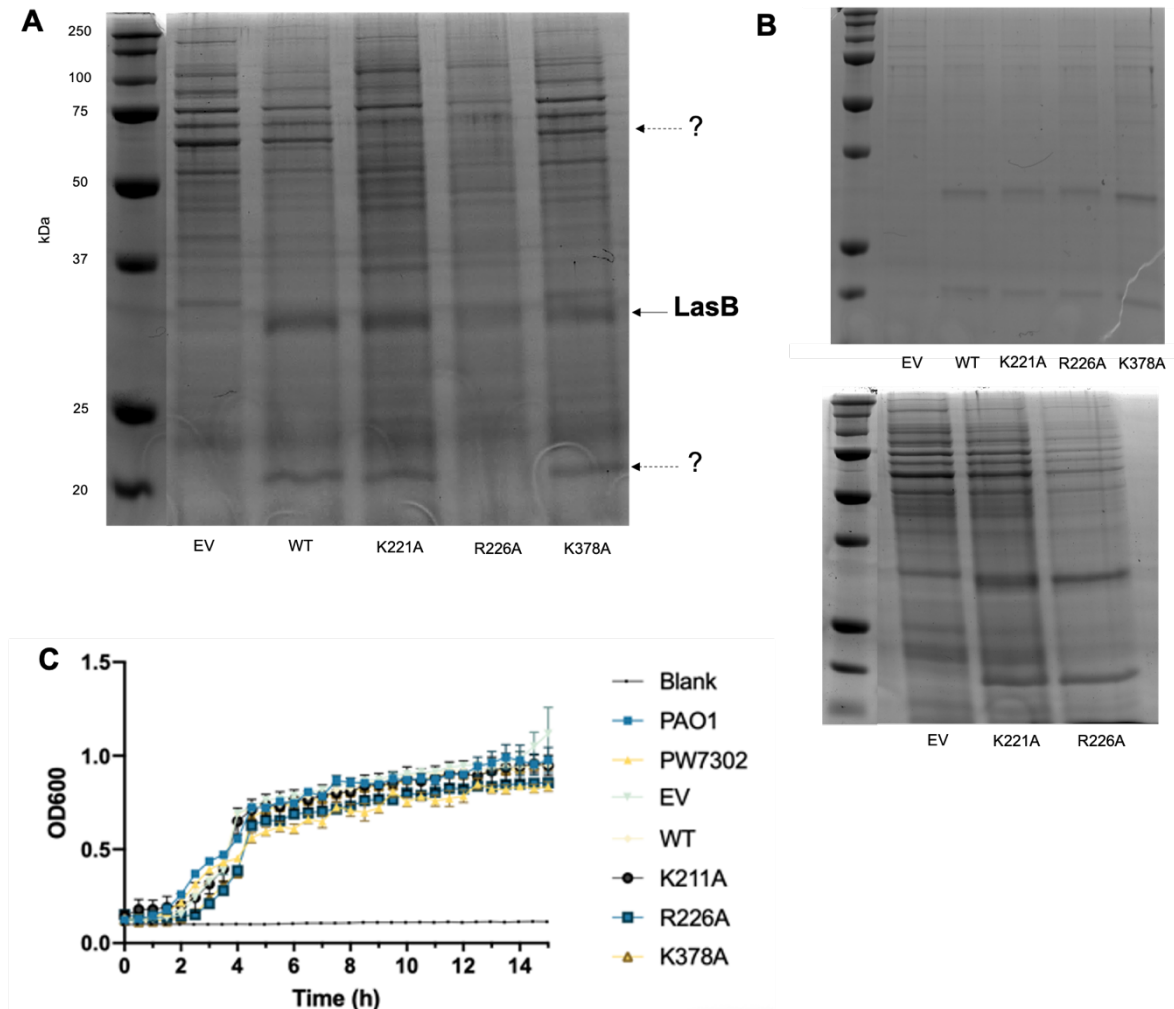
**Figure 6.2 Overlay of SDM predictions on the structure of LasB**

An overlay image of the LasB crystal structure (light blue) with the alanine substituted site-directed mutant predicted structures (magenta), with insets showing a close up of each mutated site: K211A, R226A and K378A. Mutated sequences for LasB, where the residue of interest was substituted for alanine, were submitted to I-TASSER for structural prediction and to Missense3D to bioinformatically determine likely structural changes. All site-directed mutant proteins were predicted to maintain the wild-type LasB secondary structure.

### 6.3 Secretion of the recombinant SDM proteins

Following confirmation of successful transformation of all constructs, it was important to ensure that the proteins were appropriately expressed and secreted from the cell. The use of a protein tag would have proved beneficial for both purification of the mutated proteins, as well as for immunoblotting to check secretion. However, an affinity tag was not included in the cloning process as attempts to purify recombinant LasB from PW7302 through the use of a poly-histidine tag indicated that inclusion of the tag at the C-terminus prevented tagged protein secretion (data not shown). This was problematic as although it remains unclear at which point in the secretion process PTM occurs, research has suggested major modifications may occur upon secretion (Ouidir *et al.*, 2014). To ensure successful secretion and potential modification of other residues, the tag was omitted. This also allowed the global *in vivo* effect of the mutations both within and outside of the cell to be studied, rather than only targeted

assays with exogenous purified protein. Cell-free supernatants from each strain were therefore concentrated and separated on a 1D-SDS PAGE gel (Figure 6.3).



**Figure 6.3 Confirmation of LasB complementation and secretion**

SDS-PAGE secretome separations of the LasB wild-type (WT) and SDM expressing strains (K211A, R226A and K378A), and the LasB-null empty vector (EV) control. (A) Initial gel separation with an arrow indicating the band for LasB and dotted arrows for protein bands that appear to be modulated in the SDM strains; (B) subsequent repeated separations. All gels were stained with Coomassie Brilliant blue. (C) Growth curves of all the strains indicating that the mutation and vector expression did not affect the growth rate of the strains, error bars represent the standard deviation of three replicates

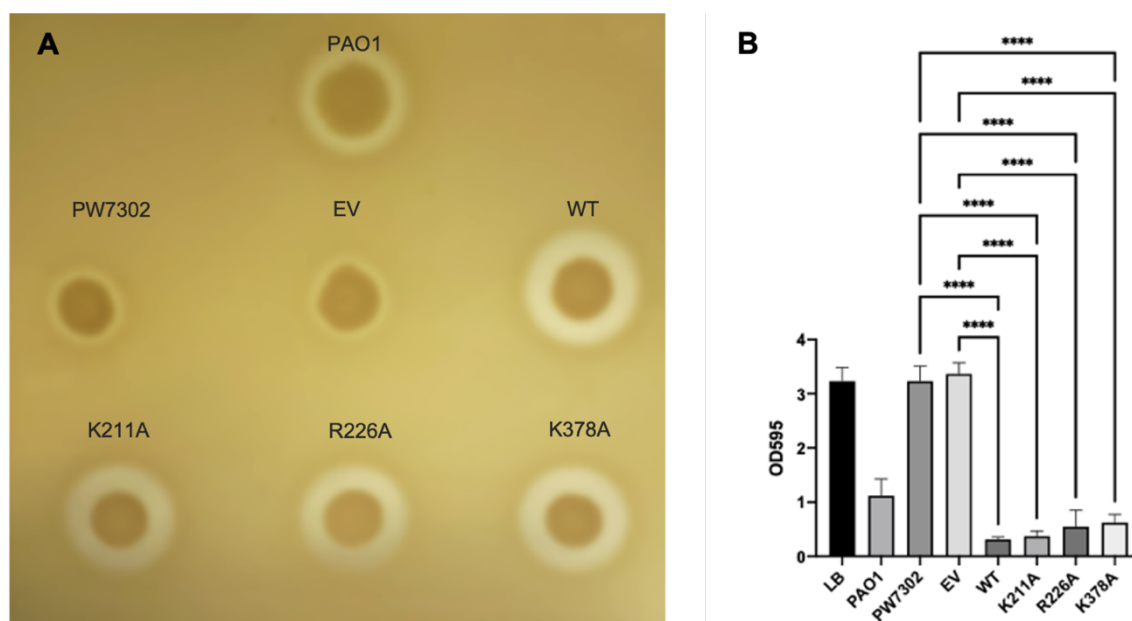
Figure 6.3A shows that LasB was effectively knocked out in PW7302 (with empty vector) and was successfully complemented and secreted from LasB-WT, LasB-K211A, LasB-R226A and LasB-K378A. Upon initial comparison of EV with the complemented strains, bands of presumed proteins that undergo activational

processing by LasB (i.e. LasA or CbpD) also appeared to be modulated (shown by dotted arrow). Initially it was thought that R226A (and K211A in the case of higher molecular weight modulations) may affect the processing of these virulence factors as no band was seen in the respective secretome. However, when this was repeated with additional biological replicates the observed differences were not consistent (Figure 6.3B). It is unclear what caused these changes, however inconsistent protein loading, or even differences in time of protein sampling, may be responsible. Absence of these bands in the initial gel may indicate differences in proteolytic degradation of these protein species. To ensure that any potential differences in protein output were not the result of differences in growth rate, the growth of all strains was measured over 15 hours and no significant differences were observed (Figure 6.3C).

#### **6.4 Effect of LasB SDM on protease activity**

The logical starting point from thereon was to investigate the protein's primary function as a protease. To screen for differences in proteolytic capacity, normalised cultures of each strain ( $OD_{600} = 1$ ) were spotted onto media containing skimmed milk (SMA) as a substrate. The presence and size of a hydrolytic zone of clearing around the culture indicated protease activity (Figure 6.4A). As expected, PW7302 and the empty vector control had reduced proteolytic activity as compared to the WT and SDM complemented strains. Interestingly, the R226A mutation showed slightly reduced activity when compared with the other strains in all four biological replicates (Table 6.1).

A skimmed milk liquid assay was also conducted to allow clearer quantitative, and cell-free, differences to be observed. Protein normalised culture supernatants were added to 4% skimmed milk solution in a 96-well microtiter plate and incubated for 16h. Decreases in the  $OD_{595}$  indicated degradation of casein. The results reflect those seen in the solid media experiment and significant differences were found between the LasB null and the complemented strains (Figure 6.4B). Whilst not statistically significant, both R226A and K378A displayed slightly reduced proteolytic activity as compared with WT and K211A, emulating the SMA results.



**Figure 6.4 Protease activity of LasB and the SDM mutants**

Skimmed milk protease assay for the detection of exoprotease production by LasB and its associated SDM point mutants on (A) solid milk-based media spotted with normalized cell culture (representative of four biological replicates, see Table 6.1 and Appendix B); or (B) in liquid form where normalized cell-free supernatants were added to skimmed milk and the OD<sub>595</sub> was measured following 16h incubation. The LasB-null strains demonstrated significantly reduced protease activity than the complemented strains; however no significant differences were found between the SDM and WT LasB expressing strains. The data shows the average of three biological and technical replicates each (n = 9) and error bars represent the standard deviation. Statistical significance between groups was assessed by one-way ANOVA with Tukey's multiple comparison *post hoc* test (\*\*\*\* = P < 0.0001).

**Table 6.1 Proteolytic zone diameter measurements**

Zone measurements for each strain spotted onto skimmed milk agar, as displayed in Figure 6.4A; numbers denote the average of three biological replicates  $\pm$  standard deviation where applicable.

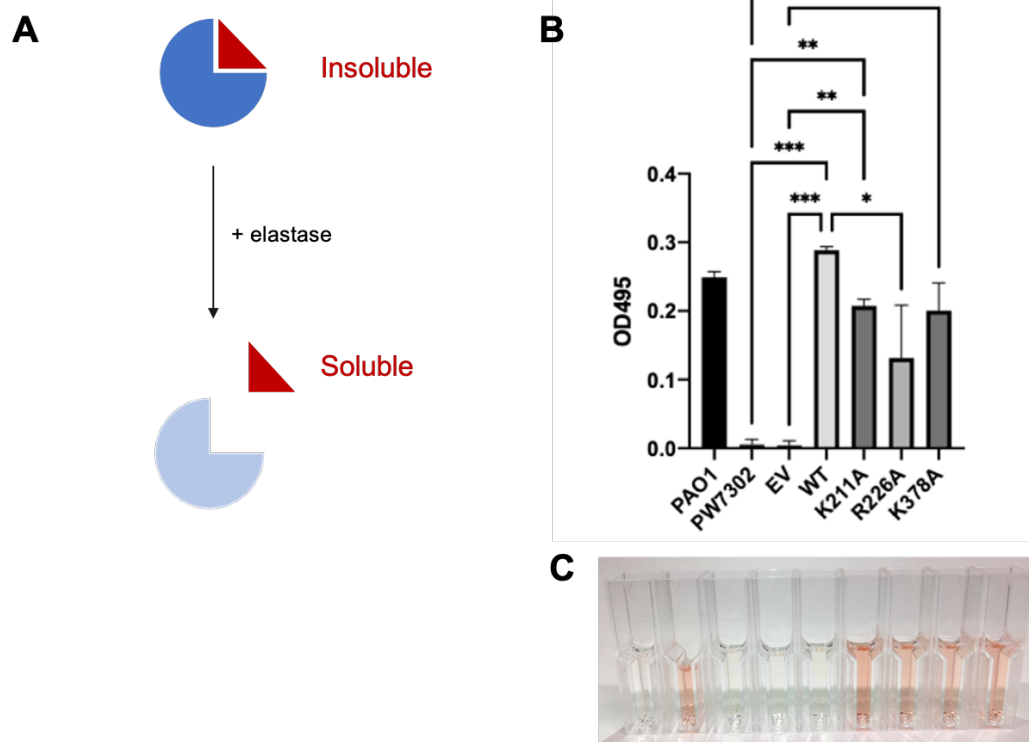
Strain	Zone diameter (mm)
PAO1	2.33 $\pm$ 0.58
PW7302	<1
EV	<1
WT	4
K211A	4
R226A	3
K378A	3.33 $\pm$ 0.29

#### 6.4.1 Effect of LasB SDM on elastinolytic activity

Although there was a clear result in the skimmed milk assays, *P. aeruginosa* also secretes many other proteases, several of which are activated by LasB, which may contribute to the proteolytic activity observed. To more specifically probe differences in elastinolytic activity, the substrate-dye conjugate elastin-Congo red (ECR) was employed. In this assay the ECR substrate starts as an insoluble complex, however upon degradation of the elastin by elastase, the Congo red dye is liberated and can be measured at 495 nm (Figure 6.5A).

In agreement with the previous findings, the results show that PW7302 and the EV control have diminished elastinolytic activity and this is effectively complemented in the WT and SDM strains (Figure 6.5). Interestingly, all three LasB mutants showed marginally lower elastinolytic activity as compared to the wild-type protein. In both biological replicates, the R226A mutation showed the lowest activity of the complemented strains and, despite a relatively large deviation in the absorbancy readings, this was shown to be statistically significant. These results therefore suggest that modification of these residues may contribute to the elastinolytic activity of LasB. Furthermore, the reduction of activity of the R226A strain in all proteolysis assays may be explained by a reduction in activation, or increased degradation, of LasA. Although unconfirmed and inconsistent, the lack of a band at the molecular weight for LasA (21 kDa) in Figure 6.3A could explain the results shown here as active LasA is known to enhance the proteolytic and elastinolytic activity of LasB.





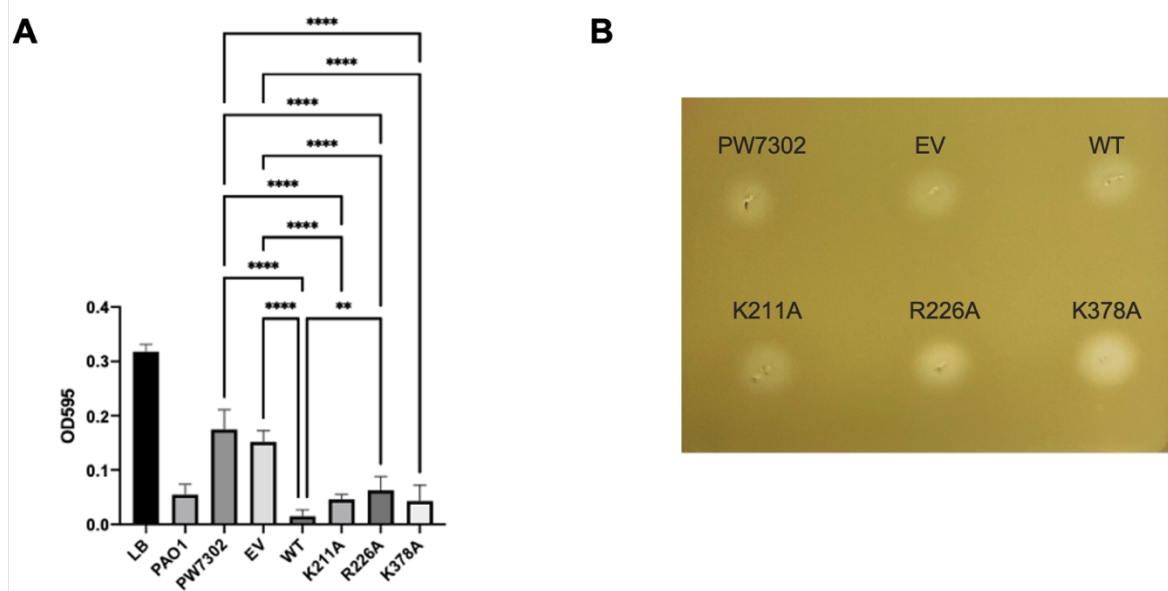
**Figure 6.5 Elastin-Congo red (ECR) assay of LasB mutants**

(A) Assay concept: ECR is an insoluble substrate-dye complex. Following the addition of active elastase, the elastin conjugate is degraded which liberates Congo red as a soluble dye that can be measured at 495 nm. (B) ECR assay of LasB mutant cell-free supernatants. The SDM strains demonstrated slightly reduced elastase activity compared with the WT; this reduction in activity was significant ( $P < 0.05$ ) for the R226A mutated LasB. The data shows the average of two biological replicates and error bars represent the standard deviation. Statistical significance between groups was assessed by one-way ANOVA with Tukey's multiple comparison *post hoc* test (\* =  $P < 0.05$ , \*\* =  $P < 0.01$ , \*\*\* =  $P < 0.001$ ). (C) Image of

## 6.5 Effect of LasB SDM on staphylytic activity

Direct assessment of the staphylytic activity of the strains was subsequently undertaken to substantiate any potential differences in the processing of LasA. Firstly, protein normalised cell-free supernatants were added to heat-killed *S. aureus* in a 96-well microtiter plate. The absorbance at 595 nm was initially measured every 5 minutes for 1.5 hours, however no differences were observed. An end-point reading was therefore taken following 16 hours incubation; a decrease in OD<sub>595</sub> indicates cell lysis.

All strains, including the LasB negative strains, showed some staphylytic capacity compared with the LB control (Figure 6.6A). This is to be expected as the *lasA* gene was not directly disrupted in these strains and the LasA protein can also be activated by other extracellular proteases (Li and Lee, 2019). There were, however, significant differences observed between the elastase null and the complemented strains, with the latter exhibiting greater staphylytic activity. Most noteworthy is the observation that R226A demonstrated significantly reduced staphylytic potential when compared to the wild-type protein ( $p = 0.007$ ). This potentially corroborates the hypothesis that modification of LasB R226 is involved in the activation or extracellular stability of LasA.



**Figure 6.6 Staphylytic activity of LasB and the SDM mutants**

(A) Cell-free supernatants from each strain were added to a heat-killed *S. aureus* suspension and the OD<sub>595</sub> was measured following 16 hours of incubation. A reduction in OD<sub>595</sub> indicates lysis of the *S. aureus* cells. The data shows that the SDM strains demonstrated slightly reduced staphylytic activity compared with the WT, and this reduction in activity was significant ( $P < 0.01$ ) in the R226A complemented strain. Error bars represent the standard deviation of three biological and technical replicates ( $n = 9$ ); statistical significance between groups was assessed by one-way ANOVA with Tukey's multiple comparison *post hoc* test (\*\* =  $P < 0.01$ , \*\*\* =  $P < 0.001$ , \*\*\*\* =  $P < 0.0001$ ). (B) Plate-based assay of staphylytic activity whereby cell free supernatants were added to wells punctured in a *S. aureus* agar overlay; a clear halo represents staphylysin activity

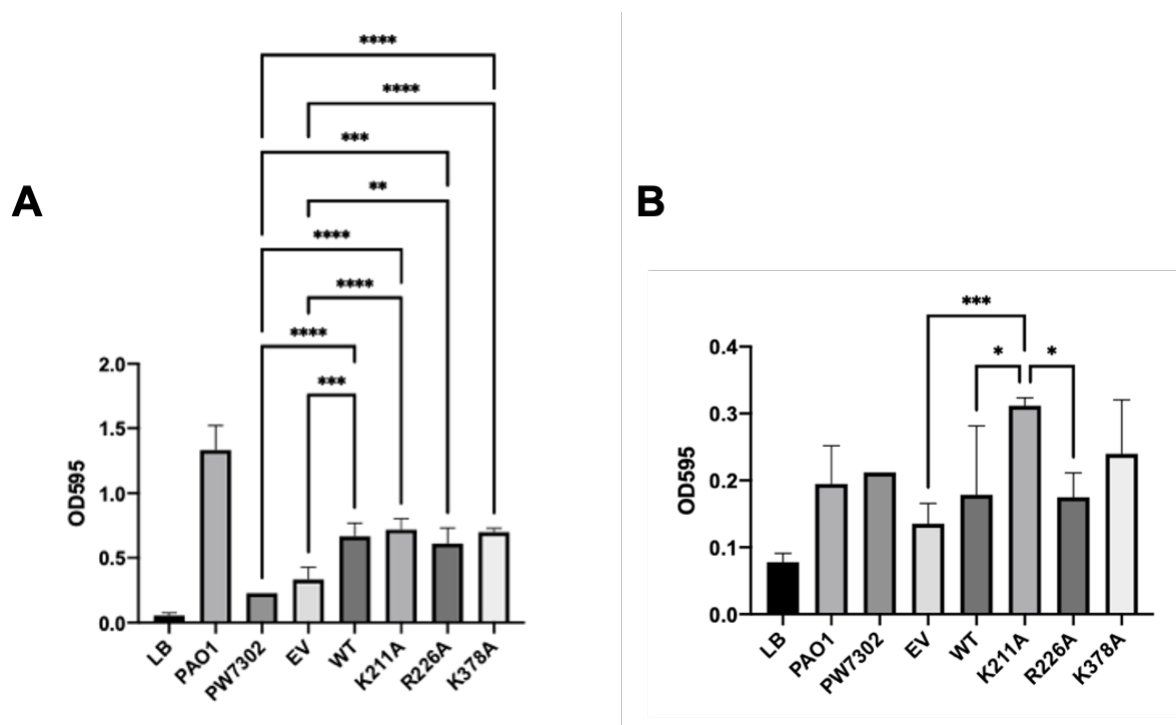
This experiment was repeated in a plate-based assay in which supernatants were added to wells punctured in a *S. aureus* agar overlay (Figure 6.6B). Discrimination of differences in zone size was difficult, however the clearance zones produced by the complemented strains showed slightly less turbidity. Whilst insufficient to complement the quantitative assay, the plate assay does endorse previous findings that staphylolytic activity is not entirely dependent on activation by LasB.

## 6.6 Effect of LasB SDM on biofilm formation

In addition to its direct activity as an extracellular endopeptidase against exogenous host proteins and other secreted proteins, LasB is involved in the complex process of biofilm formation. The ability to form biofilms is an important biological adaptation that increases resistance to antibiotics and aids immune evasion as a means of enhancing community persistence. It involves the transition from planktonic growth to sessile aggregation of bacterial cells on biotic or abiotic surfaces. This requires the production and secretion of exopolysaccharides, such as alginate, that help form the extracellular polymeric substance (EPS) (Mann and Wozniak, 2012; Mulcahy *et al.*, 2014). LasB is thought to promote biofilm formation through the periplasmic activation of nucleoside diphosphate kinase (NDK), generating GTP necessary for the production of alginate (Kamath *et al.*, 1998; Cathcart *et al.*, 2011).

The ability of each strain to form biofilms was therefore measured. Cultures were grown either statically or with rotational agitation for 24 hours and the biomass of adherent cells was quantified through crystal violet staining. In the static incubation condition, an absence of LasB resulted in a significant reduction in biofilm mass, however no differences between the LasB SDM or WT strains were found (Figure 6.7A). In contrast, when the cultures were incubated with shaking, the K211A strain showed a significant increase in biofilm mass when compared with both the WT and R226A strains, although overall cell adherence was lower for all strains (Figure 6.7B). These differences are irrespective of growth rate (Figure 6.3C). The results obtained from the shaking condition display greater variance and therefore more replicates should be generated to verify these results. A lack of concordance between the two conditions precludes robust conclusion about the effect of modifying LasB on biofilm

formation. In spite of this, the differences that have been observed should not be discounted and warrant further experimental validation.



**Figure 6.7 Biofilm formation of the LasB and SDM expressing strains**

Cultures of the indicated strains were grown either statically (A) or shaking (B) for 24 hours, after which the biofilm biomass was measured by crystal violet staining and measurement of absorbance at OD<sub>595</sub>. The LasB-null strains demonstrated reduced biofilm formation capabilities as compared with the complemented strains; R226A indicated slightly reduced biofilm formation compared with the WT and other SDM LasB-complemented strains. The data represents the mean of three biological replicates with two technical replicates each (n=6), except in the case of PW7302 which is representative of one biological replicate. Error bars represent the standard deviation, and statistical significance between groups was assessed by one-way ANOVA with Tukey's multiple comparison *post hoc* test (\* = P < 0.05 \*\* = P < 0.01, \*\*\* = P < 0.001, \*\*\*\* = P < 0.0001)

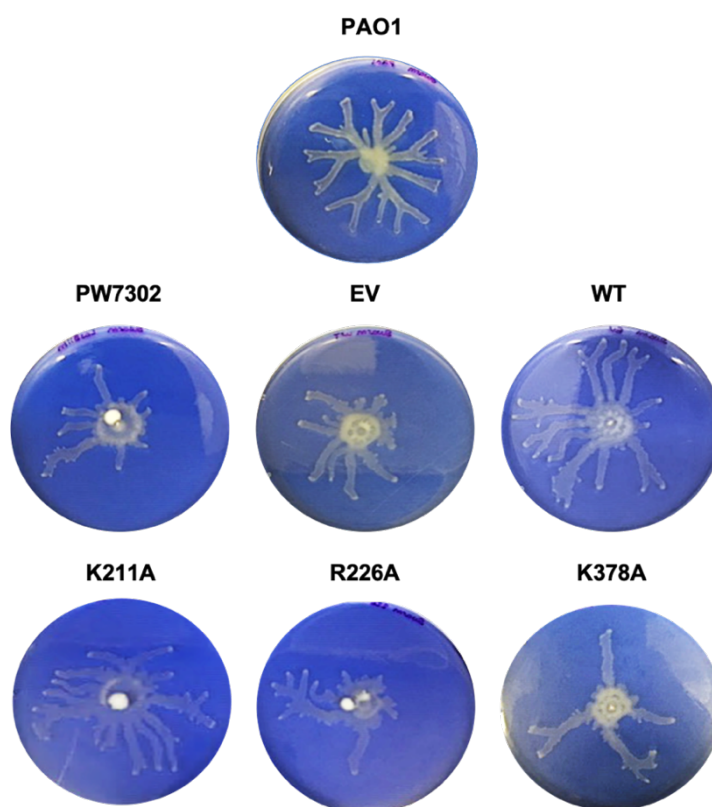
## 6.7 Effect of LasB SDM on motility

### 6.7.1 Swarming motility

Another indirect function of LasB is in the control of cell motility; more specifically, swarming motility. This multicellular phenomenon is highly co-ordinated to allow the movement of a bacterial population across semi-solid surfaces. Swarming is dependent on a functional flagellum as well as the production of biosurfactants, such as rhamnolipids, which lower the surface tension. It has previously been suggested

that LasB is associated with increased extracellular rhamnolipid concentrations which may mediate swarming motility (Tielen *et al.*, 2010; Yu *et al.*, 2014).

To test whether modification of key LasB residues is important to its role in swarming motility, normalised cultures of each strain ( $OD_{600} = 1$ ) were spotted onto swarming agar plates and grown for 24 hours (Figure 6.8). In corroboration with previous findings (Overhage *et al.*, 2008), the LasB deficient strains displayed reduced swarming motility as compared with the wild-type complement, which showed the best swarming ability of all strains. Surprisingly, two of the SDMs (R226A and K378A) demonstrated discernibly weaker swarming motility with reduced tendrill formation. Furthermore, these differences were even more pronounced when magnesium and calcium were removed from the media (see Appendix B). The presence of these elements enhances the production and stability of LasB (Thayer *et al.*, 1991). The results therefore suggest that modification of these residues may contribute to indirect LasB activities, such as control of motility.

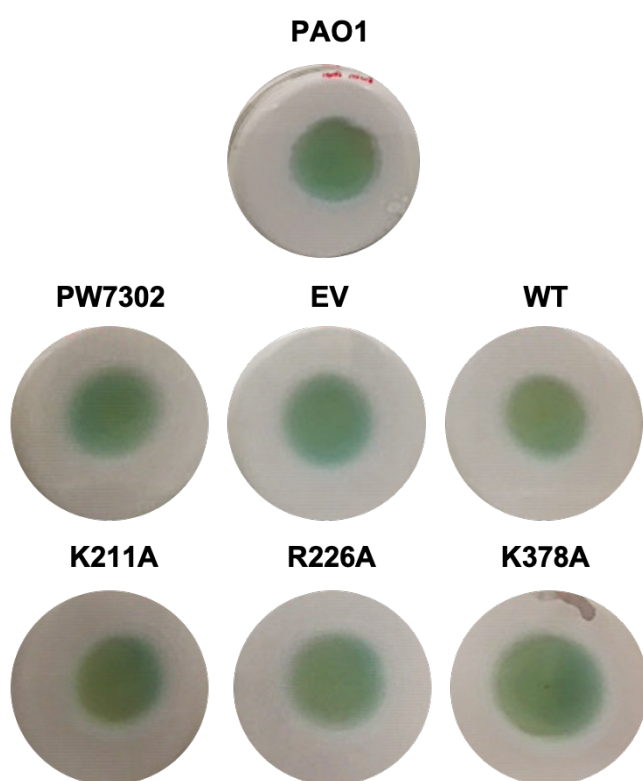


**Figure 6.8 Swarming motility of LasB SDM mutants**

Normalised cultures ( $OD_{600} = 1$ ) of the indicated strains were spotted onto swarming agar plates and motility was measured by branched growth and tendrill formation following growth at 37°C for 24 hours. The LasB-null and the LasB R226A and K378A complemented strains exhibited reduced swarming motility with reduced tendrill formation compared with the WT and K211A strains. Plates are representative of three biological replicates (see Appendix B)

### 6.7.2 Swimming motility

As outlined above, one suggestion as to how LasB may mediate swarming motility is through modulation of rhamnolipid secretion. To determine whether changes in swarming motility were due to the effect of LasB on rhamnolipid production or flagellar function, the swimming motility of the strains was also investigated. This type of motility is solely dependent on rotational propulsion by the polar flagellum to allow individual cell movement in liquid or low-viscosity environments. Differences in swimming motility can be easily determined by measuring the radial migration of cells in low viscosity medium from the point of inoculation.



**Figure 6.9 Swimming motility of LasB SDM mutants**

Normalised cultures ( $OD_{600} = 1$ ) of the indicated strains were inoculated at the bottom of swimming agar plates and motility was measured by total radial growth after 24 hours at 37°C. LasB K378A showed slightly increased swimming motility, however no other significant differences were observed. See Table 6.2 for quantitative measurements.

Strain	Swim diameter (cm)
PAO1	$4.4 \pm 0.4$
PW7302	$4.8 \pm 0.3$
EV	$5.1 \pm 0.1$
WT	$5.1 \pm 0.1$
K211A	$5.2 \pm 0.1$
R226A	$5.05 \pm 0.05$
K378A	$5.95 \pm 0.25$

**Table 6.2 Swimming migration of LasB SDM mutants**

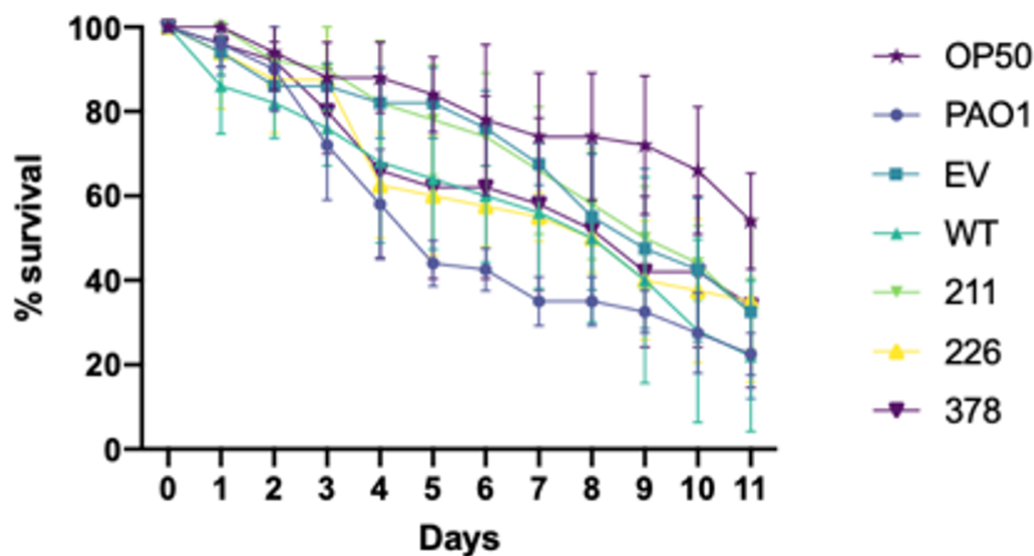
Measurements of the total radial growth of the strains (as shown in Figure 6.9) after 24 hours growth at 37°C in swimming agar. Measurements are representative of three biological replicates  $\pm$  the standard deviation

The results show that there were no major differences in swimming motility between any of the strains tested (Figure 6.9). However, the K378A mutation consistently migrated further from the point of inoculation as compared to all other strains tested (Table 6.2). This is somewhat surprising given the reduced swarming ability shown above, suggesting a potential minor role in flagellar function. In combination, the results of the motility assays provide further gravitas to the notion that LasB can modulate rhamnolipid production or secretion, and further that modification of this protease contributes to its multifunctional activity.

## **6.8 *Caenorhabditis elegans* infection model**

Elastase plays a leading role in colonisation of the host and the establishment of infection. Previous studies have shown that deletion of the *lasB* gene, or challenge of the protein with targeted drugs, causes much less severe and less invasive infection in different *in vivo* models (Tan *et al.*, 1999; Sandri *et al.*, 2018; Galdino, Viganor, *et al.*, 2019). LasB can degrade an extensive number of host proteins, including many components of the extracellular matrix, immune system and surfactant proteins. The nematode *C. elegans* is therefore an ideal model for studying variations in virulence of *P. aeruginosa* strains due to its low-cost, ethical grounding and possession of an innate immune system. This model was utilised to investigate the *in vivo* effects of the SDM strains and resulting pathogenesis within the host.

The SDM and control strains were grown as a small lawn on slow-killing agar for 48 hours and were subsequently seeded with 10 age-matched L4 worms; 5 replicate plates were completed, totalling 50 worms per strain. Every 24 hours worms were scored as either alive or dead. An absence of movement following gentle touch was indicative of death. Alongside fatal infection, there are several other common causes of death when conducting such assays, including dehydration on the sides of the plate and accidental death by researcher handling. Desiccation on the side may be considered an act of suicide as a result of pathogen avoidance. By including this indirect cause of death in the analysis, the behavioural aspects of pathogen avoidance can be appropriately considered (Lewenza *et al.*, 2014).



**Figure 6.10 *Caenorhabditis elegans* infection model survival curve**

A graph showing the percentage survival of *C. elegans* populations following infection with each strain (as detailed in the graph key) and an *E. coli* OP50 negative control. Ten age matched L4 *C. elegans* were added to a lawn of each strain and the number of surviving worms was measured every 24 hours for 11 days. Worms were considered dead if no movement was observed following gentle touch. A total of five replicate plates were conducted for each strain. The number of worms observed to be alive at each time point is shown as a percentage of the total number at time point 0. Error bars represent the standard deviation ( $n = 5$ ) at each time point. Results show that the lethality of the LasB-null and K211A SDM strains is delayed when compared to the WT, K378A and R226A complemented strains.

**Table 6.3 Average lethal time in the *C. elegans* infection assay**

The average time (in days) taken for mortality to occur in 50% of the *C. elegans* population (LT50) after inoculation with the indicated strains

LT50 (days)	
OP50	>11
PAO1	4.5
EV	8.75
WT	8
211	9
226	8
378	8.2



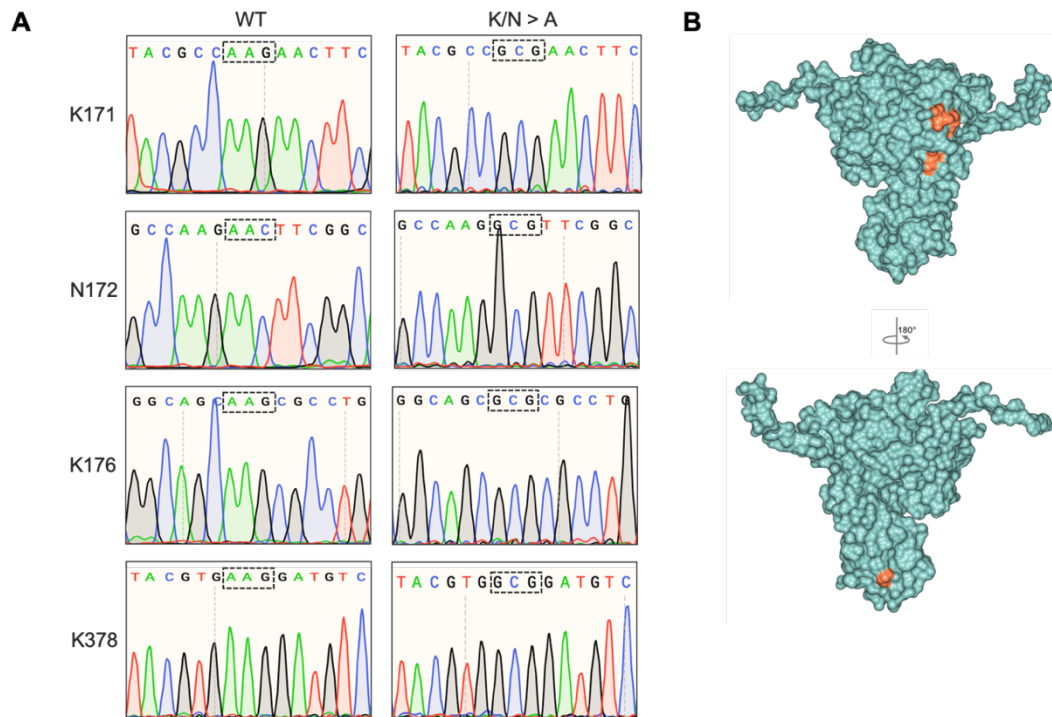
As expected, the PAO1 positive control showed the greatest decline in worm population, and this was matched at the last time point by the wild type LasB expressing strain (Figure 6.10). The SDM strains demonstrated slightly reduced, but overall similar, pathogenicity as compared to the WT. The K211A mutation, however, resulted in a slightly longer delay in fatality consistent with the LasB deficient EV control. Similar findings are also reflected in the median lethal time (LT<sub>50</sub>) for each strain (Table 6.3). This may actually indicate differences in the invasiveness of these strains, as opposed to differences in overall lethality which is due to a multitude of unrelated factors. Upon analysis of survival by two-way ANOVA with Tukey's multiple comparison test, the only statistically relevant differences found were between PAO1 vs K211A and PAO1 vs EV at days 5 to 7 ( $p < 0.008$  and  $p < 0.0085$ , respectively).

## 6.9 Effect of SDM on R2 pyocin activity

In addition to studying the role of modification in LasB, SDM was also used to mutagenise key modified residues of the R2 pyocin sheath protein PA0622. As with LasB, the most commonly modified residues across the PA0622 charge train were selected and substituted for alanine, yielding PA0622-K171A, N172A, K176A and K312A. Mutagenised genes were encoded in pUCP20 and transformed into the PCR-confirmed PA0622 transposon insertion mutant, PA2131. Appropriate empty vector controls were also created, and all constructs were confirmed by Sanger sequencing (Figure 6.11A). The locations of the substituted residues on the subunit structure are highlighted in Figure 6.11B.

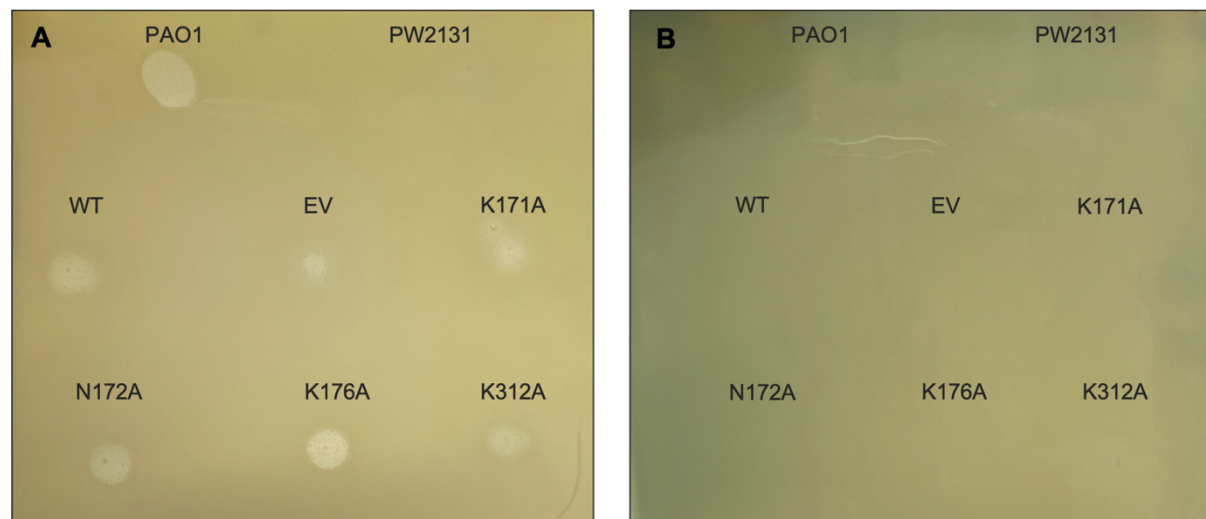
The activity of the complemented strains against several closely related species was investigated by spotting sub-MIC ciprofloxacin induced culture supernatants onto top agar containing the indicator strains. A solid clearance zone is indicative of non-replicative toxicity, and therefore signifies pyocin activity. Of the strains tested, only one (PAK) showed sensitivity to pyocin activity, as previously reported (Köhler *et al.*, 2010). The toxicity profile of the SDM mutants was varied (Figure 6.12). However, although a zone was not identified for the transposon mutant, the empty vector transformed control demonstrated toxicity against PAK. It is unclear whether this is due to technical error, or alternatively an effect of introduction of the vector. This

prohibited the formation of any conclusions and imposes a requirement for further validation.



**Figure 6.11 PA0622 SDM mutants**

Sanger sequencing chromatogram confirmation of the PA0622 WT and alanine substituted site-directed mutant constructs: K171A, N172A, K176A and K312A (A) and the positions of the mutated residues (highlighted in orange) on the PA0622 predicted surface structure (B) (as predicted by I-TASSER)



**Figure 6.12 Pyocin activity of the PA0622 SDM mutants**

Pyocin activity against (A) PAK and (B) PAO1, as a control. Cell-free supernatants from sub-MIC ciprofloxacin (0.5X MIC) induced cultures were harvested at late exponential phase and were spotted onto top agar containing the indicator strains PAK or PAO1. A zone of growth inhibition at the inoculation point indicated pyocin activity in the culture supernatants. As expected, no growth inhibition was observed in PAO1, however all strains, except the PA0622 knock-out (PW2131) showed inhibition of the growth of PAK

## 6.10 Discussion

LasB is one of the most studied virulence factors produced by *P. aeruginosa*. Nearly all investigations highlight novel roles and interactions of the protein without taking into consideration how specific residues, outside of the active site, or PTMs may affect these functions. PTMs can have significant effects on the functioning, stability, location and interactions of proteins. Key residues which I, and others, have shown to be consistently modified were therefore studied. Lysine or arginine was substituted for unreactive alanine to remove the methylation/acetylation potential whilst maintaining the overall protein structure. Phenotypic analyses were then conducted to assess differences in several known functions of LasB using strains expressing these mutant proteins.

**Table 6.4 Overview of the phenotypic assay results**

A table showing an overview of the phenotypic assay results presented in this chapter for each of the SDM and control strains. Each denotation is relative to the other strains.

	Protease	Elastase	Staphylolytic	Biofilm <sup>a</sup>	Swarming	Swimming	In vivo model
PAO1	+++	+++	+++	++++	+++	+	+++
PW7302	+	-	+	+	+	+	
EV	+	-	+	+	+	+	++
WT	+++	+++	++++	+++	+++	+	+++
K211A	+++	++	+++	++++	+++	+	+
R226A	++	+	++	++	++	+	++
K378A	+++	++	+++	+++	++	++	++

a - Average result from both static and rotation conditions

+ Low; ++ Moderate, +++ High, - negative

Inhibition of methylation and/or acetylation was shown to have a significant effect on several key functions of LasB (Table 6.4). The canonical function as an elastase was reduced in all mutants as compared to the wild-type, with a mutation of R226 showing the greatest reduction. This may be a direct effect of the mutation on LasB, or alternatively may be the result of reduced activational processing or increased degradation of other proteases, such as LasA (as was initially suggested by comparison of the mutant secretomes; Figure 6.3A). LasA functions as a staphylolysin,

but also enhances the proteolytic and elastinolytic potential of LasB by cleaving Gly-Gly bonds in substrate proteins (Kessler *et al.*, 1993). A reduction in functional LasA would therefore have a knock-on effect on LasB functioning. To further support the hypothesis that LasA is affected, it was also shown that the R226 mutation resulted in a statistically lower staphylolytic activity. Additionally, these consistencies were not a result of a reduction in the growth rate of this strain.

Preliminary differences identified in mutant secretome profiles were not able to be replicated (Figure 6.3). The results were, however, not discounted as they provide potentially important validation to the ideas presented above; the absence of a band for LasA in the R226A mutant was observed. One potential explanation as to why this result was not replicated is due to differences in the sampling time. It is well known that the secretome can undergo extensive changes during the different phases of growth (Arevalo-Ferro *et al.*, 2003), and so differences in sampling time between experiments may affect enzymatic activity, causing differences in proteolytic processing and thus composition of the samples. If true, this would need to be taken into consideration when conducting future investigations of these strains. What can be taken away from the secretome comparison is that mutation of the residues, and thus removal of PTM, does not appear to affect the presence and stability of LasB in the extracellular milieu. This could be further confirmed with more direct stability assays conducted with purified protein.

The downstream effects which could be attributed to modulation of LasA activity could also be due to LasD. This protein is another extracellular staphylolysin that is activated by LasB and has a similar molecular weight to LasA, therefore making their discrepancy on SDS-PAGE gels difficult without protein sequencing (Park and Galloway, 1995). The residual staphylolytic activity seen in the R226A mutant may therefore be due to compensatory activity by the other protein. LasD itself is a degradation product of chitin binding protein (CbpD) (Folders *et al.*, 2000). Further work could attempt to identify differences in chitinase activity as a proxy measure for LasB proteolytic activation of CbpD to LasD. These results overall suggest that modification of LasB, particularly at R226, potentially has an effect on multi-substrate proteolysis, either directly or indirectly through downstream protein activation.

The complex activation cascade of extracellular proteins, which is thought to be spearheaded by LasB, has recently been updated. Li and Lee found that LasB is activated following secretion by an unidentified extracellular quorum sensing regulated protein, rather than the previously proposed auto-activation (Li and Lee, 2019). Furthermore, activation could not be reversed by the exogenous addition of its pro-peptide, unlike with LasA. It is not fantastical to therefore suggest that the unknown protein may be conferring an activational post-translational modification. LasB has previously been shown to undergo multi-phosphorylation following secretion (Gaviard *et al.*, 2019), and this, or indeed other modifications highlighted here, may contribute to protein maturation and activation. As technology develops and complex protein interactions are put under greater scrutiny, such PTMs should be taken into serious consideration.

In addition to extracellular effects, the cell-associated functions of LasB were also investigated. The results presented here show firstly that PTM appears to have an effect on swarming motility. Interestingly, studies in *Salmonella enterica* have also shown that methylation is important for swarming but not swimming motility (Bogomolnaya *et al.*, 2014). The effect seen in this study is likely to be a result of reduction in rhamnolipid production by the mutants, further developing upon previous findings which showed an increase in rhamnolipid synthesis associated with the production of wild-type LasB (Tielen *et al.*, 2010; Yu *et al.*, 2014).

Previous findings that the expression of LasB also increases biofilm formation are further supported by the results presented here. No major differences between the majority of the complemented strains were identified however, although once again the R226 mutation indicated a potential reduction in LasB function with respect to biofilm development. Swarming ability is thought to affect the type of biofilm that is formed (Shrout *et al.*, 2006). Further investigation looking at the biofilm architecture may therefore provide further insight into how these phenotypes are linked and the role of PTM. The results suggest that whilst the intracellular activity of LasB in activating NDK to generate alginate is generally unlikely to be controlled by PTM, it may have a role in the regulation of rhamnolipid synthesis. The evidence therefore allows speculation that R226 PTM may have extensive effects on the cell physiology.

Clear differences in the functioning of LasB as a result of mutation could not be fully extrapolated to the *in vivo* infection model. Apparent differences in the virulence of the LasB<sup>+</sup> and LasB<sup>-</sup> strains were comparable to those previously published (Zhu *et al.*, 2015). Inter-experimental variation in the LT<sub>50</sub> and survival curves in the study by Zhu *et al* is reflective of the large experimental error seen here. The use of PAO1 in both studies may be counter intuitive as it only exhibits moderate lethality against *C. elegans*. Therefore, should this experiment be repeated, a substitution to the more virulent PA14 may give clearer results. Two of the key LasB modifications were previously confirmed in this strain, supporting this experimental adaptation (Gaviard *et al.*, 2019). Alternatively, a different model, such as *Galleria mellonella*, could be utilised.

Although the ultimate lethality of the strains appeared somewhat similar, upon further inspection of the data, differences in the invasiveness of some strains may be evident. Recently, Biancalani and Gore highlighted a distinct difference between invasiveness and lethality when using this model (Biancalani and Gore, 2019). Differences in invasiveness can be evident in the survival curve until the point of bacterial saturation within the host, at which point the rate of death, which is governed by many different factors, is likely to be near consistent. Taking this into consideration, alongside the role of LasB as an agent of colonisation, a reduction in the invasiveness of the K211A mutant equivalent to the LasB null strain could be proposed.

The promiscuity of LasB provides multiple opportunities for phenotypic analysis of the LasB mutants. In these assays, several phenotypes were shown to be somewhat affected by a lack of specific PTM sites (Table 6.4). There are, however, many other additional experiments that could be performed to provide insight into how PTM may affect other known functions of LasB. This includes the chitinase and rhamnolipid assays, as previously mentioned, as well as assays investigating interactions with other host proteins, such as immunoglobulins and surfactant. In addition, quantifying differences in rhamnolipid production may further support the findings from the motility assays.

Before such assays are undertaken, improvements in the experimental set up are recommended. Firstly, a clean deletion in the *lasB* gene should be constructed to

remove the potential for any transposon-induced effects. Allelic exchange can also be used to introduce plasmid-encoded mutant *lasB* genes into the chromosome via site specific homologous recombination (Hmelo *et al.*, 2015). The advantage here would be a reduction in cell-to-cell variability, an increase in stability and transcriptional consistency, as well as removal of the requirement for antibiotic selection. The choice of amino acid substitution could also be reconsidered. Substitutions that are isobaric for a modified residue can be used as a PTM mimic, i.e. glutamine as a mimic for acetyl-lysine (Carabetta and Cristea, 2017). Furthermore, the use of exogenous modifying and de-modifying enzymes would further home in on the effects of distinct modifications, although as yet these remain uncharacterised. This would also serve to determine whether the effects of mutation to alanine shown here are the result of removal of the modification, or disruption of the core functions and interactions of the residue side chain itself.

An alternative theory as to a potential role of PTM lies not in the control of LasB function, but how the host interacts with the protein. The immune system is primed to challenge any potentially damaging non-host proteins. The secretion of proteins by bacteria is therefore somewhat risky both in terms of extracellular stability and also immune signalling. In fact, *P. aeruginosa* is known to use its extracellular proteases to degrade flagellin as a means of immune evasion (Casilag *et al.*, 2015). Several studies have shown the presence of antibodies against LasB in patient sera, and several K/R containing B-cell epitopes have been identified (Sokol *et al.*, 2000; Wehmhöner *et al.*, 2003). PTM may therefore serve as a means of antigenic variation to generate immune-masking heterogeneity within the protein population. This is also supported by the idea that PTMs are typically sub-stoichiometric, increasing protein variability.

Methylation, in particular, is used for the purpose of immune evasion in the outer membrane proteins of *Leptospira interrogans* and *Rickettsia species* (Eshghi *et al.*, 2012; Abeykoon *et al.*, 2016; Nally *et al.*, 2017). In the same way, methylation also protects against ubiquitin tagging of proteins by the host, although how this relates to minimally invasive *P. aeruginosa* is unclear (Pang *et al.*, 2010). PTM may also mimic host protein epitopes to facilitate invasive interactions between pathogen and host, as in the case of *P. aeruginosa* EF-Tu (Barbier *et al.*, 2013). Acetylation has also been shown to block the immunogenicity of secreted protein Hsp $\alpha$  in *Mycobacterium*

*tuberculosis* (Liu *et al.*, 2014); further investigation of the LasB mutant strains *in vivo* is therefore warranted.

The work conducted here on the function of PTM in pyocin activity is inconclusive. As with LasB, it is recommended to construct a clean deletion mutant and use chromosomal integration of mutant genes instead of plasmid-encoded complementation. This experimental adaptation is potentially more important for investigations into PA0622 PTM as the pyocin genes are encoded in a biosynthetic operon and therefore transposon insertion may have significant downstream effects. The construction of the SDM mutants does however provide a foundation for future experimentation. In addition to functional assays, it would also be interesting to use transmission electron microscopy to also investigate how modification may affect the pyocin structure and stability.

#### **6.10.1 Conclusion**

In conclusion, a distinct and consistent role of methylation and/or acetylation of LasB could not be confirmed. However, the results provide important initial insights into the possible roles of PTM in the proteolytic functioning and interactions of LasB. They also offer a strong platform for future experimentation to encompass more modified residues and phenotypes of LasB that will contribute to a greater understanding of this key virulence factor. This may also help to inform future anti-virulence initiatives.



## Chapter 7

### 7 Final discussion

*P. aeruginosa* poses a significant clinical threat throughout the world where it is the cause of prolific morbidity and mortality. Whilst the regulation of intracellular mechanisms is important to its virulence strategy, it is the proteins which are secreted from the cell that form the front line of attack to ensure successful colonisation of the host (Filloux, 2011). *P. aeruginosa* secretes a wealth of proteins and compounds into the unstable extracellular environment where they disrupt host processes and defences and cause extensive tissue damage. In addition, these virulence factors are important in maintaining the 'upper hand' in the battle against competing organisms. The drug targeting of these factors has been proposed as a means of anti-virulence intervention to reduce the overall pathogenicity of this bacterium. It is therefore important to fully characterise secreted proteins and to understand what underscores their biological activity.

The first aim of this research was to determine how growth conditions may affect *P. aeruginosa* secretome dynamics and, more specifically, charge variation of the secreted proteins. It was found that whilst variations in the growth media, temperature, strain and growth stage can indeed affect the proteins which are secreted, there was paradoxical consistent heterogeneity in the charge of individual proteins leading to charge train formation on 2D gels. In every condition tested, the number of charge variable isoforms for many protein species remained constant. Nonetheless, the findings also supported previous transcriptome studies showing that reduction of the growth temperature to emulate the early stages of infection resulted in modulation of the protein output, including massive upregulation of protease IV secretion (Termine and Michel, 2009; Barbier *et al.*, 2014). Minor alterations in protein composition were observed when *P. aeruginosa* was grown in different media containing different carbon sources, indicating that the extent of modulation of the extracellular proteome in response to carbon source is limited compared to the intracellular proteome (Dolan *et*

*al.*, 2020). It is unclear from the results whether there is temporal regulation of charge variant secretion (and thus PTM) as has been seen in *L. interrogans* (Stewart *et al.*, 2016), however the inclusion of more sampling time points may highlight intricacies in isoform secretion in future studies.

The underlying cause of the observed protein charge heterogeneity is variable post-translational modification. PTM is a dynamic method of control over the proteome and allows a rapid response to both internal and external signals. PTMs can affect protein chemistry, resulting in alterations in protein function, stability, localisation and interactions. Understanding of their implications in bacteria, in particular virulence-associated proteins, is therefore of critical importance. How they are applied within the highly regulated intracellular network is fathomable, but how such stringent control relates to the unstable extracellular environment is unclear.

The second aim of this research was to determine which PTM causes the extensive charge variation of secreted proteins. It quickly became clear that this is an intensely complex and challenging problem that is unlikely to be deduced to modulation of a single PTM at this level of interrogation. Nevertheless, important observations were made with regards to prominent and consistent PTMs of several key secreted proteins. Firstly, extracellular proteases ImpA and LasB were found to undergo extensive methylation and acetylation of mostly surface exposed residues. The location of the modified residues on the protein surface may indicate that PTM occurs post-folding, and this could give insights into where in the cell these modifications may be occurring, i.e., in the periplasm. The high level of PTM identified is likely to reflect the complexity of the protein interactions that take place *in vivo*.

Several residues were modified in every protein spot analysed, indicating that they are potentially more abundant and thus valuable modifications. Several of these residues were also subject to different types of PTM, highlighting that the complexity and dynamism attributed to intracellular regulation also occurs in the extracellular fraction. The subsequent confirmation of several of the LasB PTMs in PA14 by Gaviard *et al* (Gaviard *et al.*, 2019) confers greater significance to the findings and provides an interesting platform for the discovery of the enzymes which carry out these modifications. The majority of the modified residues were also conserved in

homologous proteins expressed by other bacteria and so it would be interesting to determine whether the PTMs are also conserved and serve a similar function in other species. Furthermore, this research provided additional evidence of “atypical” methylation in bacteria, reinforcing the need to incorporate these residues in future bacterial PTM analyses.

Critical residues involved in the binding of zinc ions in ImpA were identified as methylated. This PTM is likely to control co-factor binding and may therefore have an important role in proteolytic functioning. Several components of the pyocin nanomachine were also identified as post-translationally modified. If PTM is indeed involved in pyocin stability or complex formation, it is important to take this into consideration if pyocins are to be developed as targeted antibiotics. In addition, the LasB and ImpA proteases are T2SS substrates, whereas pyocins are released by cell lysis. This further suggests that PTM is likely to be a universal regulatory mechanism for extracellular proteins which are secreted or released by different mechanisms.

An outer membrane porin was also found to undergo PTM, and more importantly methylation at a central substrate binding residue. This modification may well have significant implications for both *P. aeruginosa* survival in low nutritional environments and for antimicrobial resistance. Understanding of the intricate mechanisms of transport across the outer membrane is essential for rational drug design and so investigations into the role of this methylation would be of keen interest for future studies.

An immunoaffinity enrichment of methylated and acetylated proteins in the secretome was planned during the course of this research, however due to COVID-19 related disruptions these experiments were prohibitively delayed. The data generated from these experiments would have provided greater insight into the range of secreted proteins that are post-translationally modified, as well as providing important validation of the PTM sites identified. This would also have helped to identify modified proteins which are less amenable to visualisation by 2-DE, such as those with an extreme pI or low abundance. The data would also be sufficient to conduct more thorough consensus motif searches in this protein subset. Furthermore, this experiment would

have comprised the first methylome analysis in *P. aeruginosa*, contributing to the significantly understudied field of protein methylation in bacteria (André *et al.*, 2017).

Furthermore, such data may have given an indication as to whether specific subsets of proteins are over-represented in the modified population, such as those secreted through the T2SS. Intriguingly, previous research has shown that deletion of the Tat apparatus, which transports folded proteins into the periplasm pre-export, results in an increase in the number of charge isoforms of the main proteins analysed here (Ball *et al.*, 2016). This gives some support to the hypothesis that PTM may occur during the process of secretion from the cell. It is hoped that such valuable experiments can be completed in the future.

The majority of bacterial PTM research relies only on high-throughput enrichment experiments to characterise modified proteins. Whilst such analyses can generate useful and comprehensive lists of proteins that are modified by a specific PTM, these studies tend to lack a depth of understanding and can only focus on a limited range of PTMs. By reducing the scope of analysis and focusing on a subset of proteins as accomplished here, a more thorough perception of PTMs at the single protein level can be achieved. Furthermore, IEF has previously proven more beneficial than antibody enrichment for uncovering a greater diversity in PTM sites (Ouidir *et al.*, 2015). Generally, such global PTM studies also assume the function of a PTM based on the activity of the substrate or from inference based on KEGG and Gene Ontology assessment (M. Zhang *et al.*, 2018). This is because the capacity to identify PTMs currently exceeds the ability to understand their biological functions. Alternatively, when the responsible modifying enzyme has been characterised, the encoding gene is deleted and resulting global disruptions are documented. Once again this generally limits the appreciation of specific PTMs on distinct proteins.

Initial investigations were undertaken into the role of PTM at several key residues in LasB. Site-directed mutagenesis (SDM) was utilised to substitute several consistently modified residues for unreactive alanine. This allowed the preliminary characterisation of phenotypic effects resulting from the disruption of PTM at these sites. One theory was that PTM may mediate protein stability in the extracellular environment, as in the case of *Mycobacterium tuberculosis* (Lanouette *et al.*, 2014). However, all LasB SDM

mutant proteins appeared to be stable and protected from self proteolytic degradation following secretion. The SDM mutants did demonstrate reduced elastinolytic activity compared with the wild type, and the R226A mutant in particular exhibited reduced staphylolytic activity. These results may suggest that PTM of this residue is involved in the activation of, and protein-protein interaction with, LasA. Two of the mutants (R226A and K378A) also displayed reduced swarming motility compared to the wild type, suggesting an alternative role of PTM in rhamnolipid production. The potential multi-functionality of these PTMs further endorses numerous proposals to target these modifications in interventions against *P. aeruginosa* pathogenesis (Lassak *et al.*, 2019; Tiwari, 2019).

Determining the biological role of specific PTMs is commonly like finding a needle in a haystack, particularly with proteins such as LasB that have many functions and a large interactome. In many cases, realisation of an effect comes about serendipitously. The results presented here, however, are a promising starting point for further investigations which should also seek to investigate the role of PTM at more residues, as well as in other secreted proteins. As previously mentioned, identification of the responsible modifying enzymes would be incredibly beneficial for taking these experiments to the next stage and to tease apart the interactions between methylation and acetylation where these sites overlap. This would also help with PTMs where SDM is not a viable option, such as R410 in OprD, which is critical to function.

It is clear that early assumptions of “simpler” bacterial systems being incapable of incorporating such complex PTM coding into their systems is false. As such, future work should also pursue a greater range of PTMs during analysis. The patterns of modification observed across the protein isoforms here did not explain the charge variant phenomenon. This is not only due to the technical difficulties in obtaining sufficient coverage of low abundant PTMs during mass spectrometry analysis, but also the limited scope available during database searches which forms a bottleneck in the depth of novel PTM identification that can be pursued. Additionally, one spot from a 2D gel is commonly not comprised of one distinct proteoform, but actually by several different proteoforms with different PTM site occupancies conferring an overall comparable charge.

Multi-site PTM results in a combinatorial explosion in the number of protein “mod-forms” (Prabakaran *et al.*, 2012). Presently it is technologically unfeasible to distinguish all combinatorial possibilities of PTM, however use of a top-down approach to mass spectrometry may help to characterise the co-occurring PTMs on single proteins (Gault *et al.*, 2014). PTM is clearly an orchestrated process resulting in a paradoxically consistent variable protein population. Moreover, software which aims to unrestrictedly blind-search mass spectrometry data for any PTMs present is currently being developed and improved (M. K. Yang *et al.*, 2014; Kim *et al.*, 2016). The determination of PTM crosstalk is another challenging problem that will require both experimental advancements and mathematical derivation to decipher the PTM code and provide a more holistic understanding of secreted proteins and their charge variation.

Overall, this work demonstrated consistencies in *P. aeruginosa* secreted protein isoform formation in a range of growth conditions and uncovered extensive post-translational modification of several key proteins for the first time. Furthermore, several possible roles have been suggested for these PTMs in the potent protease LasB. This research lays the groundwork for further significant developments to be made in understanding the roles of PTM of *P. aeruginosa* virulence factors and the implications these may have in the survival and pathogenicity of this fascinating bacterium.

## 8 References

- Abeykoon, A. H., Noinaj, N., Choi, B. E., Wise, L., He, Y., Chao, C. C., Wang, G., Gucek, M., Ching, W. M., Chock, P. B., Buchanan, S. K. and Yang, D. C. H. (2016) 'Structural insights into substrate recognition and catalysis in outer membrane protein B (OmpB) by Protein-lysine methyltransferases from rickettsia', *Journal of Biological Chemistry*, 291(38).
- Abeykoon, A., Wang, G., Chao, C. C., Chock, P. B., Gucek, M., Ching, W. M. and Yang, D. C. H. (2014) 'Multimethylation of Rickettsia OmpB catalyzed by lysine methyltransferases', *Journal of Biological Chemistry*, 289(11).
- Ahrné, E., Müller, M. and Lisacek, F. (2010) 'Unrestricted identification of modified proteins using MS/MS', *Proteomics*, 10(4).
- Alleyn, M., Breitzig, M., Lockey, R. and Kolliputi, N. (2018) 'The dawn of succinylation: A posttranslational modification', *American Journal of Physiology - Cell Physiology*, 314(2).
- Aly, K. A., Beebe, E. T., Chan, C. H., Goren, M. A., Sepúlveda, C., Makino, S. I., Fox, B. G. and Forest, K. T. (2013) 'Cell-free production of integral membrane aspartic acid proteases reveals zinc-dependent methyltransferase activity of the *Pseudomonas aeruginosa* prepilin peptidase PilD', *MicrobiologyOpen*, 2(1).
- Ambler, R. P. and Rees, M. W. (1959) 'ε-N-Methyl-lysine in Bacterial Flagellar Protein', *Nature*, 184(4679).
- Anderson, R. D., Roddam, L. F., Bettiol, S., Sanderson, K. and Reid, D. W. (2010) 'Biosignificance of bacterial cyanogenesis in the CF lung', *Journal of Cystic Fibrosis*, 9(3).
- André, P., Bastos, D., Pinto, J. and Vitorino, R. (2017) 'A glimpse into the modulation of post-translational modifications of human-colonizing bacteria', *Journal of Proteomics*, 152.
- Angrill, J., Agustí, C., De Celis, R., Rañó, A., Gonzalez, J., Solé, T., Xaubet, A., Rodriguez-Roisin, R. and Torres, A. (2002) 'Bacterial colonisation in patients with bronchiectasis: Microbiological pattern and risk factors', *Thorax*, 57(1).
- Arevalo-Ferro, C., Buschmann, J., Reil, G., Görg, A., Wiehlmann, L., Tümmler, B., Eberl, L. and Riedel, K. (2004) 'Proteome analysis of intracolonial diversity of two *Pseudomonas aeruginosa* TB clone isolates', *Proteomics*, 4(5).
- Arevalo-Ferro, C., Hentzer, M., Reil, G., Görg, A., Kjelleberg, S., Givskov, M., Riedel, K. and Eberl, L. (2003) 'Identification of quorum-sensing regulated proteins in the opportunistic pathogen *Pseudomonas aeruginosa* by proteomics', *Environmental Microbiology*, 5(12).
- Arora, S. K., Ritchings, B. W., Almira, E. C., Lory, S. and Ramphal, R. (1998) 'The *Pseudomonas aeruginosa* flagellar cap protein, FliD, is responsible for mucin adhesion', *Infection and Immunity*, 66(3).
- Babich, T., Naucle, P., Valik, J. K., Giske, C. G., Benito, N., Cardona, R., Rivera, A., Pulcini, C., Fattah, M. A., Haquin, J., MacGowan, A., Grier, S., Chazan, B., Yanovskay, A., Ami, R. Ben, Landes, M., Nesher, L., Zaidman-Shimshovitz, A., McCarthy, K., *et al.* (2020) 'Risk factors for mortality among patients with *Pseudomonas aeruginosa* bacteraemia: a retrospective multicentre study', *International Journal of Antimicrobial Agents*, 55(2).
- Balasubramanian, D., Kumari, H. and Mathee, K. (2014) '*Pseudomonas aeruginosa* AmpR: An acute-chronic switch regulator', *Pathogens and Disease*, 73(2).

- Ball, G., Antelmann, H., Imbert, P. R. C., Gimenez, M. R., Voulhoux, R. and Ize, B. (2016) 'Contribution of the Twin Arginine Translocation system to the exoproteome of *Pseudomonas aeruginosa*', *Scientific Reports*, 6.
- Ball, G., Durand, É., Lazdunski, A. and Filloux, A. (2002) 'A novel type II secretion system in *Pseudomonas aeruginosa*', *Molecular Microbiology*, 43(2).
- Ballok, A. E. and O'Toole, G. A. (2013) 'Pouring salt on a wound: *Pseudomonas aeruginosa* virulence factors alter Na<sup>+</sup> and Cl<sup>-</sup> flux in the lung', *Journal of Bacteriology*, 195(18).
- Barak, R. and Eisenbach, M. (2001) 'Acetylation of the response regulator, CheY, is involved in bacterial chemotaxis', *Molecular Microbiology*, 40(3).
- Barbier, M., Damron, F. H., Bielecki, P., Suárez-Diez, M., Puchałka, J., Albertí, S., Dos Santos, V. M. and Goldberg, J. B. (2014) 'From the environment to the host: Re-wiring of the transcriptome of *Pseudomonas aeruginosa* from 22°C to 37°C', *PLoS ONE*, 9(2).
- Barbier, M., Owings, J. P., Barbier, M., Owings, J. P., Martínez-ramos, I., Damron, F. H., Gomila, R. and Blázquez, J. (2013) 'Lysine Trimethylation of EF-Tu Mimics Platelet-Activating Factor To Initiate *Pseudomonas aeruginosa* Pneumonia', *American Society for Microbiology*, 4(3).
- Bardoel, B. W., Hartsink, D., Vughs, M. M., de Haas, C. J. C., van Strijp, J. A. G. and van Kessel, K. P. M. (2012) 'Identification of an immunomodulating metalloprotease of *Pseudomonas aeruginosa* (IMPα)', *Cellular Microbiology*, 14(6).
- Barker, A. P., Vasil, A. I., Filloux, A., Ball, G., Wilderman, P. J. and Vasil, M. L. (2004) 'A novel extracellular phospholipase C of *Pseudomonas aeruginosa* is required for phospholipid chemotaxis', *Molecular Microbiology*, 53(4).
- Bassetti, M., Vena, A., Croxatto, A., Righi, E. and Guery, B. (2018) 'How to manage *Pseudomonas aeruginosa* infections', *Drugs in Context*, 7.
- Bastaert, F., Kheir, S., Saint-Criq, V., Villeret, B., Dang, P. M. C., El-Benna, J., Sirard, J. C., Voulhoux, R. and Sallenave, J. M. (2018) '*Pseudomonas aeruginosa* LasB subverts alveolar macrophage activity by interfering with bacterial killing through downregulation of innate immune defense, reactive oxygen species generation, and complement activation', *Frontiers in Immunology*, 9.
- Behrens, H. M., Six, A., Walker, D. and Kleanthous, C. (2017) 'The therapeutic potential of bacteriocins as protein antibiotics', *Emerging Topics in Life Sciences*, 1(1).
- Beltrao, P., Bork, P., Krogan, N. J. and Van Noort, V. (2013) 'Evolution and functional cross-talk of protein post-translational modifications', *Molecular Systems Biology*, 9(1).
- Bendtsen, J. D., Kiemer, L., Fausbøll, A. and Brunak, S. (2005) 'Non-classical protein secretion in bacteria', *BMC Microbiology*, 5.
- Bhagirath, A. Y., Li, Y., Somayajula, D., Dadashi, M., Badr, S. and Duan, K. (2016) 'Cystic fibrosis lung environment and *Pseudomonas aeruginosa* infection', *BMC Pulmonary Medicine*, 16(1).
- Biancalani, T. and Gore, J. (2019) 'Disentangling bacterial invasiveness from lethality in an experimental host-pathogen system', *Molecular Systems Biology*, 15(6).
- Bio-Rad (no date) *2-E Electrophoresis and Analysis*.
- Bleves, S., Viarre, V., Salacha, R., Michel, G. P. F., Filloux, A. and Voulhoux, R. (2010) 'Protein secretion systems in *Pseudomonas aeruginosa*: A wealth of pathogenic weapons',



*International Journal of Medical Microbiology*, 300(8).

Bogomolnaya, L. M., Aldrich, L., Ragoza, Y., Talamantes, M., Andrews, K. D., McClelland, M. and Andrews-Polymenis, H. L. (2014) 'Identification of novel factors involved in modulating motility of salmonella enterica serotype typhimurium', *PLoS ONE*, 9(11).

Bomberger, J. M., Ye, S., MacEachran, D. P., Koeppen, K., Barnaby, R. L., O'Toole, G. A. and Stanton, B. A. (2011) 'A *Pseudomonas aeruginosa* toxin that hijacks the host ubiquitin proteolytic system', *PLoS Pathogens*, 7(3).

Braun, P., De Groot, A., Bitter, W. and Tommassen, J. (1998) 'Secretion of elastinolytic enzymes and their propeptides by *Pseudomonas aeruginosa*', *Journal of Bacteriology*, 180(13).

Breidenstein, E. B. M., de la Fuente-Núñez, C. and Hancock, R. E. W. (2011) '*Pseudomonas aeruginosa*: All roads lead to resistance', *Trends in Microbiology*, 19(8).

Brenner, S. (1974) 'The genetics of *Caenorhabditis elegans*.', *Genetics*, 77(1).

Brown, C. W., Sridhara, V., Boutz, D. R., Person, M. D., Marcotte, E. M., Barrick, J. E. and Wilke, C. O. (2017) 'Large-scale analysis of post-translational modifications in *E. coli* under glucose-limiting conditions', *BMC Genomics*, 18(1).

Bryant, F. R. (1988) 'Construction of a recombinase-deficient mutant *recA* protein that retains single-stranded DNA-dependent ATPase activity', *Journal of Biological Chemistry*, 263(18).

Bucior, I., Pielage, J. F. and Engel, J. N. (2012) '*Pseudomonas aeruginosa* Pili and Flagella mediate distinct binding and signaling events at the apical and basolateral surface of airway epithelium', *PLoS Pathogens*, 8(4).

Burstein, D., Satanower, S., Simovitch, M., Belnik, Y., Zehavi, M., Yerushalmi, G., Ben-Aroya, S., Pupko, T. and Banin, E. (2015) 'Novel type III effectors in *pseudomonas aeruginosa*', *mBio*, 6(2).

Buth, S. A., Shneider, M. M., Scholl, D. and Leiman, P. G. (2018) 'Structure and analysis of R1 and R2 pyocin receptor-binding fibers', *Viruses*, 10(8).

Caiazza, N. C., Shanks, R. M. Q. and O'Toole, G. A. (2005) 'Rhamnolipids Modulate Swarming Motility Patterns of *Pseudomonas aeruginosa*', *Journal of Bacteriology*, 187(21).

Cain, J. A., Solis, N. and Cordwell, S. J. (2014) 'Beyond gene expression: The impact of protein post-translational modifications in bacteria', *Journal of Proteomics*, 97.

Campodónico, V. L., Llosa, N. J., Grout, M., Döring, G., Maira-Litrán, T. and Pier, G. B. (2010) 'Evaluation of flagella and flagellin of *Pseudomonas aeruginosa* as vaccines', *Infection and Immunity*, 78(2).

Cao, X. J., Dai, J., Xu, H., Nie, S., Chang, X., Hu, B. Y., Sheng, Q. H., Wang, L. S., Ning, Z. Bin, Li, Y. X., Guo, X. K., Zhao, G. P. and Zeng, R. (2010) 'High-coverage proteome analysis reveals the first insight of protein modification systems in the pathogenic spirochete *Leptospira interrogans*', *Cell Research*, 20(2).

Carabetta, V. J. and Cristea, I. M. (2017) 'Regulation , Function , and Detection of Protein Acetylation in Bacteria', *Journal of Bacteriology*, 199(16).

Casilag, F., Lorenz, A., Krueger, J., Klawonn, F., Weiss, S. and Häussler, S. (2015) 'LasB elastase of *Pseudomonas aeruginosa* acts in concert with alkaline protease AprA to prevent flagellin-mediated immune recognition', *Infection and Immunity*, 84(1).

- Cathcart, G. R. A., Quinn, D., Greer, B., Harriott, P., Lynas, J. F., Gilmore, B. F. and Walker, B. (2011) 'Novel inhibitors of the *Pseudomonas aeruginosa* virulence factor LasB: A potential therapeutic approach for the attenuation of virulence mechanisms in pseudomonal infection', *Antimicrobial Agents and Chemotherapy*, 55(6).
- Cezairliyan, B., Vinayavekhin, N., Grenfell-Lee, D., Yuen, G. J., Saghatelian, A. and Ausubel, F. M. (2013) 'Identification of *Pseudomonas aeruginosa* Phenazines that Kill *Caenorhabditis elegans*', *PLoS Pathogens*, 9(1).
- Chao, C. C., Zhang, Z., Wang, H., Alkhalil, A. and Ching, W. M. (2008) 'Serological reactivity and biochemical characterization of methylated and unmethylated forms of a recombinant protein fragment derived from outer membrane protein B of *Rickettsia typhi*', *Clinical and Vaccine Immunology*, 15(4).
- Chen, G., Cao, M., Luo, K., Wang, L., Wen, P. and Shi, S. (2018) 'ProAcePred: Prokaryote lysine acetylation sites prediction based on elastic net feature optimization', *Bioinformatics*, 34(23).
- Chen, L., Zou, Y., She, P. and Wu, Y. (2015) 'Composition, function, and regulation of T6SS in *Pseudomonas aeruginosa*', *Microbiological Research*, 172.
- Christensen, D. G., Baumgartner, J. T., Xie, X., Jew, K. M., Basisty, N., Schilling, B., Kuhn, M. L. and Wolfe, A. J. (2019) 'Mechanisms, detection, and relevance of protein acetylation in prokaryotes', *mBio*, 10(2).
- Christensen, David G., Xie, X., Basisty, N., Byrnes, J., McSweeney, S., Schilling, B. and Wolfe, A. J. (2019) 'Post-translational Protein Acetylation: An Elegant Mechanism for Bacteria to Dynamically Regulate Metabolic Functions', *Frontiers in Microbiology*, 10(July).
- Clark, S. T., Diaz Caballero, J., Cheang, M., Coburn, B., Wang, P. W., Donaldson, S. L., Zhang, Y., Liu, M., Keshavjee, S., Yau, Y. C. W., Waters, V. J., Elizabeth Tullis, D., Guttman, D. S. and Hwang, D. M. (2015) 'Phenotypic diversity within a *Pseudomonas aeruginosa* population infecting an adult with cystic fibrosis.', *Scientific reports*, 5.
- Cornelis, P. and Dingemans, J. (2013) '*Pseudomonas aeruginosa* adapts its iron uptake strategies in function of the type of infections', *Frontiers in Cellular and Infection Microbiology*, 4(NOV).
- Costa, T. R. D., Felisberto-Rodrigues, C., Meir, A., Prevost, M. S., Redzej, A., Trokter, M. and Waksman, G. (2015) 'Secretion systems in Gram-negative bacteria: Structural and mechanistic insights', *Nature Reviews Microbiology*, 13(6).
- Cota-Gomez, A., Vasil, A. I., Kadurugamuwa, J., Beveridge, T. J., Schweizer, H. P. and Vasil, M. L. (1997) 'plcR1 and plcR2 are putative calcium-binding proteins required for secretion of the hemolytic phospholipase C of *Pseudomonas aeruginosa*', *Infection and Immunity*, 65(7).
- Crooks, G. E., Hon, G., Chandonia, J. M. and Brenner, S. E. (2004) 'WebLogo: A sequence logo generator', *Genome Research*, 14(6).
- Crosby, H. A., Pelletier, D. A., Hurst, G. B. and Escalante-Semerena, J. C. (2012) 'System-wide studies of N-lysine acetylation in *Rhodopseudomonas palustris* reveal substrate specificity of protein acetyltransferases', *Journal of Biological Chemistry*, 287(19).
- Daury, L., Orange, F., Taveau, J. C., Verchère, A., Monlezun, L., Gounou, C., Marreddy, R. K. R., Picard, M., Broutin, I., Pos, K. M. and Lambert, O. (2016) 'Tripartite assembly of RND multidrug efflux pumps', *Nature Communications*, 7.
- Davey, M. E., Davey, M. E., Caiazza, N. C., Caiazza, N. C., Toole, G. a O. and Toole, G. a O.

(2003) 'Rhamnolipid Surfactant Production Affects Bio Im Architecture in', *Microbiology*, 185(3).

Deng, X., Hahne, T., Schröder, S., Redweik, S., Nebija, D., Schmidt, H., Janssen, O., Lachmann, B. and Wätzig, H. (2012) 'The challenge to quantify proteins with charge trains due to isoforms or conformers', *Electrophoresis*, 33(2).

Dolan, S. K., Pereira, G., Silva-Rocha, R. and Welch, M. (2020) 'Transcriptional regulation of central carbon metabolism in *Pseudomonas aeruginosa*', *Microbial Biotechnology*, 13(1).

Döring, G., Obernesser, H.-J., Botzenbart, K., Flebmig, B., Høiby, N. and Hofmann, A. (1983) 'Proteases of *Pseudomonas aeruginosa* in Patients with Cystic Fibrosis', *The Journal of Infectious Diseases*, 147(4).

Douzi, B., Filloux, A. and Voulhoux, R. (2012) 'On the path to uncover the bacterial type II secretion system', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1592).

Douzi, B., Trinh, N. T. T., Michel-Souzy, S., Desmyter, A., Ball, G., Barbier, P., Kosta, A., Durand, E., Forest, K. T., Cambillau, C., Roussel, A. and Voulhoux, R. (2017) 'Unraveling the self-assembly of the *pseudomonas aeruginosa* XcpQ secretin periplasmic domain provides new molecular insights into type II secretion system secretin architecture and dynamics', *mBio*, 8(5).

Dreier, J. and Ruggerone, P. (2015) 'Interaction of antibacterial compounds with RND efflux pumps in *Pseudomonas aeruginosa*', *Frontiers in Microbiology*, 6(JUL).

Duong, F., Bonnet, E., Géli, V., Lazdunski, A., Murgier, M. and Filloux, A. (2001) 'The AprX protein of *Pseudomonas aeruginosa*: A new substrate for the Apr type I secretion system', *Gene*, 262(1–2).

Durand, E., Alphonse, S., Brochier-Armanet, C., Ball, G., Douzi, B., Filloux, A., Bernard, C. and Voulhoux, R. (2011) 'The assembly mode of the pseudopilus: A hallmark to distinguish a novel secretion system subtype', *Journal of Biological Chemistry*, 286(27).

Duthie, Y. E. S. and Lorenz, L. L. (1952) 'Staphylococcal Coagulase: Mode of Action and Antigenicity', *Journal of General Microbiology*, 6(1–2).

Ellis, T. N., Leiman, S. A. and Kuehn, M. J. (2010) 'Naturally produced outer membrane vesicles from *Pseudomonas aeruginosa* elicit a potent innate immune response via combined sensing of both lipopolysaccharide and protein components', *Infection and Immunity*, 78(9).

Engel, J. and Balachandran, P. (2009) 'Role of *Pseudomonas aeruginosa* type III effectors in disease', *Current Opinion in Microbiology*, 12(1).

Engström, P., Burke, T. P., Iavarone, A. T. and Welch, M. D. (2020) 'Lysine methylation shields an intracellular pathogen from ubiquitylation', *bioRxiv*.

Eren, E., Parkin, J., Adelanwa, A., Cheneke, B., Movileanu, L., Khalid, S. and Van Den Berg, B. (2013) 'Toward understanding the outer membrane uptake of small molecules by *Pseudomonas aeruginosa*', *Journal of Biological Chemistry*, 288(17).

Eren, E., Vijayaraghavan, J., Liu, J., Cheneke, B. R., Touw, D. S., Lepore, B. W., Indic, M., Movileanu, L. and van den Berg, B. (2012) 'Substrate specificity within a family of outer membrane carboxylate channels', *PLoS Biology*, 10(1).

Eshghi, A., Pinne, M., Haake, D. A., Zuerner, R. L., Frank, A. and Cameron, C. E. (2012) 'Methylation and in vivo expression of the surfaceexposed *Leptospira interrogans* outer-

membrane protein OmpL32', *Microbiology*, 158(3).

Farrell, P. M. (2008) 'The prevalence of cystic fibrosis in the European Union', *Journal of Cystic Fibrosis*, 7(5).

Faure, E., Kwong, K. and Nguyen, D. (2018) 'Pseudomonas aeruginosa in Chronic Lung Infections: How to Adapt Within the Host?', *Frontiers in Immunology*, 9(OCT).

Favero, M. S., Carson, L. A., Bond, W. W. and Petersen, N. J. (1971) 'Pseudomonas aeruginosa: Growth in distilled water from hospitals', *Science*, 173(3999).

Feldman, M., Bryan, R., Rajan, S., Scheffler, L., Brunnert, S., Tang, H. and Prince, A. (1998) 'Role of flagella in pathogenesis of Pseudomonas aeruginosa pulmonary infection', *Infection and Immunity*, 66(1).

Fernández, L. and Hancock, R. E. W. (2012) 'Adaptive and mutational resistance: Role of porins and efflux pumps in drug resistance', *Clinical Microbiology Reviews*, 25(4).

Filloux, A. (2011) 'Protein secretion systems in Pseudomonas aeruginosa: An essay on diversity, evolution, and function', *Frontiers in Microbiology*, 2(JULY).

Fleiszig, S. M. J. and Evans, D. J. (2010) 'Pathogenesis of contact lens-associated microbial keratitis', *Optometry and vision science: official publication of the American Academy of Optometry*, 87(4).

Folders, J., Tommassen, J., Van Loon, L. C. and Bitter, W. (2000) 'Identification of a chitin-binding protein secreted by Pseudomonas aeruginosa', *Journal of Bacteriology*, 182(5).

Forrest, S. and Welch, M. (2020) 'Arming the troops: Post-translational modification of extracellular bacterial proteins', *Science Progress*, 103(4).

Fortuin, S., Tomazella, G. G., Nagaraj, N., Sampson, S. L., Gey van Pittius, N. C., Soares, N. C., Wiker, H. G., de Souza, G. A. and Warren, R. M. (2015) 'Phosphoproteomics analysis of a clinical Mycobacterium tuberculosis Beijing isolate: Expanding the mycobacterial phosphoproteome catalog', *Frontiers in Microbiology*, 6(FEB).

Franco, A. A., Buckwold, S. L., Shin, J. W., Ascon, M. and Sears, C. L. (2005) 'Mutation of the zinc-binding metalloprotease motif affects Bacteroides fragilis toxin activity but does not affect propeptide processing', *Infection and Immunity*, 73(8).

Frederiksen, B., Lanng, S., Koch, C. and Høiby, N. (1996) 'Improved survival in the Danish center-treated cystic fibrosis patients: Results of aggressive treatment', *Pediatric Pulmonology*, 21(3).

Frimmersdorf, E., Horatzek, S., Pelnikevich, A., Wiehlmann, L. and Schomburg, D. (2010) 'How Pseudomonas aeruginosa adapts to various environments: A metabolomic approach', *Environmental Microbiology*, 12(6).

Frye, J., Karlinsey, J. E., Felise, H. R., Marzolf, B., Dowidar, N., McClelland, M. and Hughes, K. T. (2006) 'Identification of new flagellar genes of Salmonella enterica serovar Typhimurium.', *Journal of bacteriology*, 188(6).

Fuhrmann, J., Subramanian, V., Kojetin, D. J. and Thompson, P. R. (2016) 'Activity-Based Profiling Reveals a Regulatory Link between Oxidative Stress and Protein Arginine Phosphorylation', *Cell Chemical Biology*, 23(8).

Fulara, A., Vandenberghe, I., Read, R. J., Devreese, B. and Savvides, S. N. (2018) 'Structure and oligomerization of the periplasmic domain of GspL from the type II secretion system of Pseudomonas aeruginosa', *Scientific Reports*, 8(1).

Galdino, A. C. M., de Oliveira, M. P., Ramalho, T. C., de Castro, A. A., Branquinha, M. H. and Santos, A. L. S. (2019) 'Anti-Virulence Strategy against the Multidrug-Resistant Bacterial Pathogen *Pseudomonas aeruginosa*: Pseudolysin (Elastase B) as a Potential Druggable Target', *Current Protein & Peptide Science*, 20(5).

Galdino, A. C. M., Viganor, L., De Castro, A. A., Da Cunha, E. F. F., Mello, T. P., Mattos, L. M., Pereira, M. D., Hunt, M. C., O'Shaughnessy, M., Howe, O., Devereux, M., McCann, M., Ramalho, T. C., Branquinha, M. H. and Santos, A. L. S. (2019) 'Disarming *Pseudomonas aeruginosa* virulence by the inhibitory action of 1,10-phenanthroline-5,6-dione-based compounds: Elastase B (lasB) as a chemotherapeutic target', *Frontiers in Microbiology*, 10(JULY).

Garbom, S., Olofsson, M., Björnfot, A. C., Srivastava, K., Robinson, V. L., Oyston, P. C. F., Titball, R. W. and Wolf-Watz, H. (2007) 'Phenotypic characterization of a virulence-associated protein, VagH, of *Yersinia pseudotuberculosis* reveals a tight link between VagH and the type III secretion system', *Microbiology*, 153(5).

García-Fontana, C., Reyes-Darias, J. A., Muñoz-Martínez, F., Alfonso, C., Morel, B., Ramos, J. L. and Krell, T. (2013) 'High specificity in CheR methyltransferase function: CheR2 of *pseudomonas putida* is essential for chemotaxis, whereas CheR1 is involved in biofilm formation', *Journal of Biological Chemistry*, 288(26).

Garcia, M., Morello, E., Garnier, J., Barrault, C., Garnier, M., Burucoa, C., Lecron, J. C., Si-Tahar, M., Bernard, F. X. and Bodet, C. (2018) '*Pseudomonas aeruginosa* flagellum is critical for invasion, cutaneous persistence and induction of inflammatory response of skin epidermis', *Virulence*, 9(1).

Gault, J., Malosse, C., Machata, S., Millien, C., Podglajen, I., Ploy, M.-C., Costello, C. E., Duménil, G. and Chamot-Rooke, J. (2014) 'Complete posttranslational modification mapping of pathogenic *Neisseria meningitidis* pilins requires top-down mass spectrometry', *Proteomics*. 2014/03/12, 14(10).

Gaviard, C., Broutin, I., Cosette, P., Dé, E., Jouenne, T. and Hardouin, J. (2018) 'Lysine Succinylation and Acetylation in *Pseudomonas aeruginosa*', *Journal of Proteome Research*, 17(7).

Gaviard, C., Cosette, P., Jouenne, T. and Hardouin, J. (2019) 'LasB and CbpD Virulence Factors of *Pseudomonas aeruginosa* Carry Multiple Post-Translational Modifications on Their Lysine Residues', *Journal of Proteome Research*, 18(3).

Gaviard, C., Jouenne, T. and Hardouin, J. (2018) 'Proteomics of *Pseudomonas aeruginosa*: the increasing role of post-translational modifications', *Expert Review of Proteomics*, 15(9).

Ge, P., Scholl, D., Leiman, P. G., Yu, X., Miller, J. F. and Zhou, Z. H. (2015) 'Atomic structures of a bactericidal contractile nanotube in its pre- and postcontraction states.', *Nature structural & molecular biology*, 22(5).

Ge, P., Scholl, D., Prokhorov, N. S., Avaylon, J., Shneider, M. M., Browning, C., Buth, S. A., Plattner, M., Chakraborty, U., Ding, K., Leiman, P. G., Miller, J. F. and Zhou, Z. H. (2020) 'Action of a minimal contractile bactericidal nanomachine', *Nature*, 580(7805).

Ge, R. and Shan, W. (2011) 'Bacterial Phosphoproteomic Analysis Reveals the Correlation Between Protein Phosphorylation and Bacterial Pathogenicity', *Genomics, Proteomics and Bioinformatics*, 9(4–5).

Ge, R., Sun, X., Xiao, C., Yin, X., Shan, W., Chen, Z. and He, Q. Y. (2011) 'Phosphoproteome analysis of the pathogenic bacterium *Helicobacter pylori* reveals over-representation of tyrosine phosphorylation and multiply phosphorylated proteins', *Proteomics*, 11(8).

- Gebhart, D., Lok, S., Clare, S., Tomas, M., Stares, M., Scholl, D., Donskey, C. J., Lawley, T. D. and Govoni, G. R. (2015) 'A modified R-type bacteriocin specifically targeting *Clostridium difficile* prevents colonization of mice without affecting gut microbiota diversity', *mBio*, 6(2).
- Gellatly, S. L. and Hancock, R. E. W. (2013) 'Pseudomonas aeruginosa: New insights into pathogenesis and host defenses', *Pathogens and Disease*, 67(3).
- Gessard, C. (1984) 'On the Blue and Green Coloration that Appears on Bandages', *Reviews of Infectious Diseases*, 6(3).
- Ghequire, M. G. K., Dillen, Y., Lambrichts, I., Proost, P., Wattiez, R. and De Mot, R. (2015) 'Different ancestries of R tailocins in rhizospheric Pseudomonas isolates', *Genome Biology and Evolution*, 7(10).
- Ghequire, M. G. K. and De Mot, R. (2015) 'The Tailocin Tale: Peeling off Phage Tails', *Trends in Microbiology*, 23(10).
- Golovkine, G., Reboud, E. and Huber, P. (2018) 'Pseudomonas aeruginosa takes a multi-target approach to achieve junction breach', *Frontiers in Cellular and Infection Microbiology*, 7(JAN).
- Green, E. R. and Meccas, J. (2016) 'Bacterial Secretion Systems: An Overview', *Virulence Mechanisms of Bacterial Pathogens, Fifth Edition*, 4(1).
- Greer, E. L. and Shi, Y. (2012) 'Histone methylation: a dynamic mark in health, disease and inheritance', *Nature reviews. Genetics*, 13(5).
- Guida, M., Di Onofrio, V., Gallè, F., Gesuele, R., Valeriani, F., Liguori, R., Romano Spica, V. and Liguori, G. (2016) 'Pseudomonas aeruginosa in swimming pool water: Evidences and perspectives for a new control strategy', *International Journal of Environmental Research and Public Health*, 13(9).
- Guillon, A., Brea, D., Morello, E., Tang, A., Jouan, Y., Ramphal, R., Korkmaz, B., Perez-Cruz, M., Trottein, F., O'Callaghan, R. J., Gosset, P. and Si-Tahar, M. (2017) 'Pseudomonas aeruginosa proteolytically alters the interleukin 22-dependent lung mucosal defense', *Virulence*, 8(6).
- Guo, A., Gu, H., Zhou, J., Mulhern, D., Wang, Y., Lee, K. A., Yang, V., Aguiar, M., Kornhauser, J., Jia, X., Ren, J., Beausoleil, S. A., Silva, J. C., Vemulapalli, V., Bedford, M. T. and Comb, M. J. (2014) 'Immunoaffinity enrichment and mass spectrometry analysis of protein methylation', *Molecular and Cellular Proteomics*, 13(1).
- Gupta, K., Sahm, D. F., Mayfield, D. and Stamm, W. E. (2001) 'Antimicrobial Resistance Among Uropathogens that Cause Community-Acquired Urinary Tract Infections in Women: A Nationwide Analysis', *Clinical Infectious Diseases*, 33(1).
- Haba, E., Pinazo, A., Jauregui, O., Espuny, M. J., Infante, M. R. and Manresa, A. (2003) 'Physicochemical characterization and antimicrobial properties of rhamnolipids produced by Pseudomonas aeruginosa 47T2 NCBIM 40044', *Biotechnology and Bioengineering*, 81(3).
- Halder, P. K., Roy, C. and Datta, S. (2019) 'Structural and functional characterization of type three secretion system ATPase PscN and its regulator PscL from Pseudomonas aeruginosa', *Proteins: Structure, Function and Bioinformatics*, 87(4).
- Hall, S., McDermott, C., Anoopkumar-Dukie, S., McFarland, A. J., Forbes, A., Perkins, A. V., Davey, A. K., Chess-Williams, R., Kiefel, M. J., Arora, D. and Grant, G. D. (2016) 'Cellular effects of pyocyanin, a secreted virulence factor of Pseudomonas aeruginosa', *Toxins*, 8(8).

- Han, Y., Wang, T., Chen, G., Pu, Q., Liu, Q., Zhang, Y., Xu, L., Wu, M. and Liang, H. (2019) 'A *Pseudomonas aeruginosa* type VI secretion system regulated by CueR facilitates copper acquisition', *PLoS Pathogens*, 15(12).
- Hancock, R. E. W. and Brinkman, F. S. L. (2002) 'Function of *Pseudomonas* Porins in Uptake and Efflux', *Annual Review of Microbiology*, 56(1).
- Hao, P., Adav, S. S., Gallart-Palau, X. and Sze, S. K. (2017) 'Recent advances in mass spectrometric analysis of protein deamidation', *Mass Spectrometry Reviews*, 36(6).
- Hart-Smith, G., Low, J. K. K., Erce, M. A. and Wilkins, M. R. (2012) 'Enhanced methylarginine characterization by post-translational modification-specific targeted data acquisition and electron-transfer dissociation mass spectrometry', *Journal of the American Society for Mass Spectrometry*, 23(8).
- Hauser, A. R. (2009) 'The type III secretion system of *Pseudomonas aeruginosa*: infection by injection', *Nature reviews. Microbiology*, 7(9).
- Hayakawa, J. (2012) 'Flagellar Glycosylation: Current Advances', in Petrescu, M. I. E.-S. (ed.) *Glycosylation*. Rijeka: IntechOpen.
- Henrichfreise, B., Wiegand, I., Pfister, W. and Wiedemann, B. (2007) 'Resistance mechanisms of multiresistant *Pseudomonas aeruginosa* strains from Germany and correlation with hypermutation', *Antimicrobial Agents and Chemotherapy*, 51(11).
- Henry, R. L., Mellis, C. M. and Petrovic, L. (1992) 'Muroid *Pseudomonas aeruginosa* is a marker of poor survival in cystic fibrosis', *Pediatric Pulmonology*, 12(3).
- Heo, Y. J., Chung, I. Y., Choi, K. B. and Cho, Y. H. (2007) 'R-type pyocin is required for competitive growth advantage between *Pseudomonas aeruginosa* strains', *Journal of Microbiology and Biotechnology*, 17(1).
- Hickey, C., Schaible, B., Nguyen, S., Hurley, D., Srikumar, S., Fanning, S., Brown, E., Crifo, B., Matallanas, D., McClean, S., Taylor, C. T. and Schaffer, K. (2018) 'Increased Virulence of Bloodstream Over Peripheral Isolates of *P. aeruginosa* Identified Through Post-transcriptional Regulation of Virulence Factors', *Frontiers in cellular and infection microbiology*, 8(October).
- Hidron, A. I., Edwards, J. R., Patel, J., Horan, T. C., Sievert, D. M., Pollock, D. A. and Fridkin, S. K. (2008) 'Antimicrobial-Resistant Pathogens Associated With Healthcare-Associated Infections: Annual Summary of Data Reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2006–2007', *Infection Control & Hospital Epidemiology*. 2015/01/02, 29(11).
- Hmelo, L. R., Borlee, B. R., Almblad, H., Love, M. E., Randall, T. E., Tseng, B. S., Lin, C., Irie, Y., Storek, K. M., Yang, J. J., Siehnel, R. J., Howell, P. L., Singh, P. K., Tolker-Nielsen, T., Parsek, M. R., Schweizer, H. P. and Harrison, J. J. (2015) 'Precision-engineering the *Pseudomonas aeruginosa* genome with two-step allelic exchange', *Nature Protocols*, 10(11).
- Ho Sui, S. J., Lo, R., Fernandes, A. R., Caulfield, M. D. G., Lerman, J. A., Xie, L., Bourne, P. E., Baillie, D. L. and Brinkman, F. S. L. (2012) 'Raloxifene attenuates *Pseudomonas aeruginosa* pyocyanin production and virulence', *International journal of antimicrobial agents*. 2012/07/20, 40(3).
- Hoge, R., Pelzer, A., Rosenau, F. and Wilhelm, S. (2010) 'Weapons of a pathogen: proteases and their role in virulence of *Pseudomonas aeruginosa*', *Current Research, Technology and Education Topics in Applied Microbiology and Microbial Biotechnology*, 45.
- Hood, R. D., Singh, P., Hsu, F., Güvener, T., Carl, M. A., Trinidad, R. R. S., Silverman, J. M.,

Ohlson, B. B., Hicks, K. G., Plemel, R. L., Li, M., Schwarz, S., Wang, W. Y., Merz, A. J., Goodlett, D. R. and Mougous, J. D. (2010) 'A type VI secretion system of *Pseudomonas aeruginosa* targets a toxin to bacteria', *Cell host & microbe*, 7(1).

Hooper, N. M. (1994) 'Families of zinc metalloproteases', *FEBS Letters*, 354(1).

Horcajada, J. P., Montero, M., Oliver, A., Sorlí, L., Luque, S., Gómez-Zorrilla, S., Benito, N. and Grau, S. (2019) 'Epidemiology and treatment of multidrug-resistant and extensively drug-resistant *Pseudomonas aeruginosa* infections', *Clinical Microbiology Reviews*, 32(4).

Horsman, S. R., Moore, R. A. and Lewenza, S. (2012) 'Calcium Chelation by Alginate Activates the Type III Secretion System in Mucoid *Pseudomonas aeruginosa* Biofilms', *PLoS ONE*, 7(10).

Horstmann, J. A., Lunelli, M., Cazzola, H., Heidemann, J., Kühne, C., Steffen, P., Szefs, S., Rossi, C., Lokareddy, R. K., Wang, C., Lemaire, L., Hughes, K. T., Uetrecht, C., Schlüter, H., Grassl, G. A., Stradal, T. E. B., Rossez, Y., Kolbe, M. and Erhardt, M. (2020) 'Methylation of *Salmonella Typhimurium* flagella promotes bacterial adhesion and host cell invasion', *Nature Communications*, 11(1).

Huszczynski, S. M., Lam, J. S. and Khursigara, C. M. (2020) 'The role of *Pseudomonas aeruginosa* lipopolysaccharide in bacterial pathogenesis and physiology', *Pathogens*, 9(1).

Hwang, W. and Yoon, S. S. (2019) 'Virulence Characteristics and an Action Mode of Antibiotic Resistance in Multidrug-Resistant *Pseudomonas aeruginosa*', *Scientific Reports*, 9(1).

Isabella, V. M., Campbell, A. J., Manchester, J., Sylvester, M., Nayar, A. S., Ferguson, K. E., Tommasi, R. and Miller, A. A. (2015) 'Toward the rational design of carbapenem uptake in *pseudomonas aeruginosa*', *Chemistry and Biology*, 22(4).

Ittisoponpisan, S., Islam, S. A., Khanna, T., Alhuzimi, E., David, A. and Sternberg, M. J. E. (2019) 'Can Predicted Protein 3D Structures Provide Reliable Insights into whether Missense Variants Are Disease Associated?', *Journal of Molecular Biology*, 431(11).

Jacobs, M. A., Alwood, A., Thaipisuttikul, I., Spencer, D., Haugen, E., Ernst, S., Will, O., Kaul, R., Raymond, C., Levy, R., Chun-Rong, L., Guenther, D., Bovee, D., Olson, M. V. and Manoil, C. (2003) 'Comprehensive transposon mutant library of *Pseudomonas aeruginosa*', *Proceedings of the National Academy of Sciences of the United States of America*, 100.

Jensen, P., Bjarnsholt, T., Phipps, R., Rasmussen, T. B., Calum, H., Christoffersen, L., Moser, C., Williams, P., Pressler, T., Givskov, M. and Høiby, N. (2007) 'Rapid necrotic killing of polymorphonuclear leukocytes is caused by quorum-sensing-controlled production of rhamnolipid by *Pseudomonas aeruginosa*', *Microbiology*, 153(5).

Jiang, F., Waterfield, N. R., Yang, J., Yang, G. and Jin, Q. (2014) 'A *Pseudomonas aeruginosa* type VI secretion phospholipase D effector targets both prokaryotic and eukaryotic cells', *Cell Host and Microbe*, 15(5).

Kabsch, W. and Sander, C. (1983) 'Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features', *Biopolymers*, 22(12).

Kamath, S., Kapatral, V. and Chakrabarty, A. M. (1998) 'Cellular function of elastase in *Pseudomonas aeruginosa*: Role in the cleavage of nucleoside diphosphate kinase and in alginate synthesis', *Molecular Microbiology*, 30(5).

Kang, D., Kirienko, D. R., Webster, P., Fisher, A. L. and Kirienko, N. V. (2018) 'Pyoverdine, a siderophore from *Pseudomonas aeruginosa*, translocates into *C. elegans*, removes iron, and activates a distinct host response', *Virulence*, 9(1).



- Kelly-Wintenberg, K., South, S. L. and Montie, T. C. (1993) 'Tyrosine phosphate in a- and b-type flagellins of *Pseudomonas aeruginosa*', *Journal of Bacteriology*, 175(8).
- Keogh, R. H., Szczesniak, R., Taylor-Robinson, D. and Bilton, D. (2018) 'Up-to-date and projected estimates of survival for people with cystic fibrosis using baseline characteristics: A longitudinal study using UK patient registry data', *Journal of cystic fibrosis : official journal of the European Cystic Fibrosis Society*. 2018/01/06, 17(2).
- Kessler, E., Safrin, M., Olson, J. C. and Ohman, D. E. (1993) 'Secreted LasA of *Pseudomonas aeruginosa* is a staphylolytic protease', *Journal of Biological Chemistry*, 268(10).
- Kessler, E., Safrin, M., Peretz, M. and Burstein, Y. (1992) 'Identification of cleavage sites involved in proteolytic processing of *Pseudomonas aeruginosa* preproelastase', *FEBS Letters*, 299(3).
- Khan, F., Pham, D. T. N., Oloketuyi, S. F. and Kim, Y. M. (2020) 'Regulation and controlling the motility properties of *Pseudomonas aeruginosa*', *Applied Microbiology and Biotechnology*, 104(1).
- Kim, D., Yu, B. J., Kim, J. A., Lee, Y. J., Choi, S. G., Kang, S. and Pan, J. G. (2013) 'The acetylproteome of Gram-positive model bacterium *Bacillus subtilis*', *Proteomics*, 13(10–11).
- Kim, M.-S., Zhong, J. and Pandey, A. (2016) 'Common errors in mass spectrometry-based analysis of post-translational modifications.', *Proteomics*, 16(5).
- Kim, Y. H., Park, K. H., Kim, S. Y., Ji, E. S., Kim, J. Y., Lee, S. K., Yoo, J. S., Kim, H. S. and Park, Y. M. (2011) 'Identification of trimethylation at C-terminal lysine of pilin in the cyanobacterium *Synechocystis* PCC 6803', *Biochemical and Biophysical Research Communications*, 404(2).
- Kim, Y. J., Jun, Y. H., Kim, Y. R., Park, K. G., Park, Y. J., Kang, J. Y. and Kim, S. I. (2014) 'Risk factors for mortality in patients with *Pseudomonas aeruginosa* bacteremia; retrospective study of impact of combination antimicrobial therapy', *BMC Infectious Diseases*, 14(1).
- King, J. D., Kocíncová, D., Westman, E. L. and Lam, J. S. (2009) 'Lipopolysaccharide biosynthesis in *Pseudomonas aeruginosa*', *Innate Immunity*, 15(5).
- Kirienko, N. V, Ausubel, F. M. and Ruvkun, G. (2015) 'Mitophagy confers resistance to siderophore-mediated killing by *Pseudomonas aeruginosa*.', *Proceedings of the National Academy of Sciences of the United States of America*, 112(6).
- Klausen, M., Heydorn, A., Ragas, P., Lambertsen, L., Aaes-Jørgensen, A., Molin, S. and Tolker-Nielsen, T. (2003) 'Biofilm formation by *Pseudomonas aeruginosa* wild type, flagella and type IV pili mutants', *Molecular Microbiology*, 48(6).
- Klockgether, J., Cramer, N., Wiehlmann, L., Davenport, C. F. and Tümmler, B. (2011) '*Pseudomonas aeruginosa* genomic structure and diversity', *Frontiers in Microbiology*, 2(JULY).
- Klockgether, J., Munder, A., Neugebauer, J., Davenport, C. F., Stanke, F., Larbig, K. D., Heeb, S., Schöck, U., Pohl, T. M., Wiehlmann, L. and Tümmler, B. (2010) 'Genome diversity of *Pseudomonas aeruginosa* PAO1 laboratory strains', *Journal of Bacteriology*, 192(4).
- Kloth, C., Schirmer, B., Munder, A., Stelzer, T., Rothsuh, J. and Seifert, R. (2018) 'The role of *Pseudomonas aeruginosa* exoy in an acute mouse lung infection model', *Toxins*, 10(5).
- Kohler, T., Curty, L. K., Barja, F., Van Delden, C. and Pechere, J. C. (2000) 'Swarming of *Pseudomonas aeruginosa* is dependent on cell-to-cell signaling and requires flagella and pili',

*Journal of Bacteriology*, 182(21).

Köhler, T., Donner, V. and Van Delden, C. (2010) 'Lipopolysaccharide as shield and receptor for R-pyocin-mediated killing in *Pseudomonas aeruginosa*', *Journal of Bacteriology*, 192(7).

Kosono, S., Tamura, M., Suzuki, S., Kawamura, Y., Yoshida, A., Nishiyama, M. and Yoshida, M. (2015) 'Changes in the acetylome and succinylome of *Bacillus subtilis* in response to carbon source', *PLoS ONE*, 10(6).

Kropinski, A. M. B., Lewis, V. and Berry, D. (1987) 'Effect of growth temperature on the lipids, outer membrane proteins, and lipopolysaccharides of *Pseudomonas aeruginosa* PAO', *Journal of Bacteriology*, 169(5).

Kuang, Z., Bennett, R. C., Lin, J., Hao, Y., Zhu, L., Akinbi, H. T. and Lau, G. W. (2020) 'Surfactant phospholipids act as molecular switches for premature induction of quorum sensing-dependent virulence in *Pseudomonas aeruginosa*', *Virulence*, 11(1).

Kuang, Z., Hao, Y., Walling, B. E., Jeffries, J. L., Ohman, D. E. and Lau, G. W. (2011) 'Pseudomonas aeruginosa Elastase provides an Escape from phagocytosis by degrading the pulmonary surfactant protein-A', *PLoS ONE*, 6(11).

Kumar, P., Joy, J., Pandey, A. and Gupta, D. (2017) 'PRmePRed: A protein arginine methylation prediction tool', *PLoS ONE*, 12(8).

Laarman, A. J., Bardoel, B. W., Ruyken, M., Fernie, J., Milder, F. J., van Strijp, J. A. G. and Rooijackers, S. H. M. (2012) 'Pseudomonas aeruginosa Alkaline Protease Blocks Complement Activation via the Classical and Lectin Pathways', *The Journal of Immunology*, 188(1).

LaBauve, A. E. and Wargo, M. J. (2012) 'Growth and laboratory maintenance of *Pseudomonas aeruginosa*', *Current Protocols in Microbiology*.

Lai, S. J., Tu, I. F., Wu, W. L., Yang, J. T., Luk, L. Y. P., Lai, M. C., Tsai, Y. H. and Wu, S. H. (2017) 'Site-specific His/Asp phosphoproteomic analysis of prokaryotes reveals putative targets for drug resistance', *BMC Microbiology*, 17(1).

Lange, S., Rosenkrands, I., Stein, R., Andersen, P., Kaufmann, S. H. E. and Jungblut, P. R. (2014) 'Analysis of protein species differentiation among mycobacterial low-Mr-secreted proteins by narrow pH range Immobililine gel 2-DE-MALDI-MS', *Journal of Proteomics*, 97.

Lanouette, S., Mongeon, V., Figeys, D. and Couture, J. F. (2014) 'The functional diversity of protein lysine methylation', *Molecular Systems Biology*, 10(4).

Lassak, J., Koller, F., Krafczyk, R. and Volkwein, W. (2019) 'Exceptionally versatile - Arginine in bacterial post-translational protein modifications', *Biological Chemistry*, 400(11).

Lecoutere, E., Verleyen, P., Haenen, S., Vandersteegen, K., Noben, J. P., Robben, J., Schoofs, L., Ceyssens, P. J., Volckaert, G. and Lavigne, R. (2012) 'A theoretical and experimental proteome map of *Pseudomonas aeruginosa* PAO1', *MicrobiologyOpen*, 1(2).

Lee, J. and Zhang, L. (2015) 'The hierarchy quorum sensing network in *Pseudomonas aeruginosa*', *Protein and Cell*, 6(1).

Lemieux, A. A., Jeukens, J., Kukavica-Ibrulj, I., Fothergill, J. L., Boyle, B., Laroche, J., Tucker, N. P., Winstanley, C. and Levesque, R. C. (2016) 'Genes required for free phage production are essential for pseudomonas aeruginosa chronic lung infections', *Journal of Infectious Diseases*, 213(3).

Lengqvist, J., Eriksson, H., Gry, M., Uhlén, K., Björklund, C., Bjellqvist, B., Jakobsson, P. J.

and Lehtiö, J. (2011) 'Observed peptide pI and retention time shifts as a result of post-translational modifications in multidimensional separations using narrow-range IPG-IEF', *Amino Acids*, 40(2).

Leo, J. C., Grin, I. and Linke, D. (2012) 'Type V secretion: Mechanism(s) of autotransport through the bacterial outer membrane', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1592).

Létoffé, S., Redeker, V. and Wandersman, C. (1998) 'Isolation and characterization of an extracellular haem-binding protein from *Pseudomonas aeruginosa* that shares function and sequence similarities with the *Serratia marcescens* HasA haemophore', *Molecular Microbiology*, 28(6).

Lewenza, S., Charron-Mazenod, L., Giroux, L. and Zamponi, A. D. (2014) 'Feeding behaviour of *Caenorhabditis elegans* is an indicator of *Pseudomonas aeruginosa* PAO1 virulence', *PeerJ*, 2014(1).

Lewenza, S., Gardy, J. L., Brinkman, F. S. L. and Hancock, R. E. W. (2005) 'Genome-wide identification of *Pseudomonas aeruginosa* exported proteins using a consensus computational strategy combined with a laboratory-based PhoA fusion screen', *Genome research*, 15(2).

Li, C. and Clarke, S. (1992) 'A protein methyltransferase specific for altered aspartyl residues is important in *Escherichia coli* stationary-phase survival and heat-shock resistance', *Proceedings of the National Academy of Sciences of the United States of America*, 89(20).

Li, H., Luo, Y. F., Williams, B. J., Blackwell, T. S. and Xie, C. M. (2012) 'Structure and function of OprD protein in *Pseudomonas aeruginosa*: From antibiotic resistance to novel therapies', *International Journal of Medical Microbiology*, 302(2).

Li, J., Ramezanzpour, M., Fong, S. A., Cooksley, C., Murphy, J., Suzuki, M., Psaltis, A. J., Wormald, P. J. and Vreugde, S. (2019) 'Pseudomonas aeruginosa Exoprotein-Induced Barrier Disruption Correlates With Elastase Activity and Marks Chronic Rhinosinusitis Severity', *Frontiers in cellular and infection microbiology*, 9(February).

Li, X. H. and Lee, J. H. (2019) 'Quorum sensing-dependent post-secretional activation of extracellular proteases in *Pseudomonas aeruginosa*', *Journal of Biological Chemistry*, 294(51).

Liang, W., Malhotra, A. and Deutscher, M. P. (2011) 'Acetylation Regulates the Stability of a Bacterial Protein: Growth Stage-Dependent Modification of RNase R', *Molecular Cell*, 44(1).

Liao, G., Xie, L., Li, X., Cheng, Z. and Xie, J. (2014) 'Unexpected extensive lysine acetylation in the trump-card antibiotic producer *Streptomyces roseosporus* revealed by proteome-wide profiling', *Journal of Proteomics*, 106.

Liu, F., Yang, M., Wang, X., Yang, S., Gu, J., Zhou, J., Zhang, X. E., Deng, J. and Ge, F. (2014) 'Acetylome analysis reveals diverse functions of lysine acetylation in *Mycobacterium tuberculosis*', *Molecular and Cellular Proteomics*, 13(12).

Llanes, C., Llanes, C., Hocquet, D., Vagne, C., Benali-baitich, D., Neuwirth, C. and Ple, P. (2004) 'Clinical Strains of *Pseudomonas aeruginosa* Overproducing MexAB-OprM and MexXY Efflux Pumps Simultaneously Clinical Strains of *Pseudomonas aeruginosa* Overproducing MexAB-OprM and MexXY Efflux Pumps Simultaneously', *Antimicrobial Agents and Chemotherapy*, 48(5).

Lorenz, A., Preuß, M., Bruchmann, S., Pawar, V., Grahl, N., Pils, M. C., Nolan, L. M., Filloux, A., Weiss, S. and Häussler, S. (2019) 'Importance of flagella in acute and chronic *Pseudomonas aeruginosa* infections', *Environmental Microbiology*, 21(3).

- Lotfy, W. A., Atalla, R. G., Sabra, W. A. and El-Helow, E. R. (2018) 'Expression of extracellular polysaccharides and proteins by clinical isolates of *Pseudomonas aeruginosa* in response to environmental conditions', *International Microbiology*, 21(3).
- Lothrop, A. P., Torres, M. P. and Fuchs, S. M. (2013) 'Deciphering post-translational modification codes', *FEBS Letters*, 587(8).
- Lu, Z., Cheng, Z., Zhao, Y. and Volchenbom, S. L. (2011) 'Bioinformatic analysis and post-translational modification crosstalk prediction of lysine acetylation', *PLoS ONE*, 6(12).
- Luckett, J. C. A., Darch, O., Watters, C., AbuOun, M., Wright, V., Paredes-Osses, E., Ward, J., Goto, H., Heeb, S., Pommier, S., Rumbaugh, K. P., Cámara, M. and Hardie, K. R. (2012) 'A Novel Virulence Strategy for *Pseudomonas aeruginosa* Mediated by an Autotransporter with Arginine-Specific Aminopeptidase Activity', *PLoS Pathogens*, 8(8).
- Luo, M. (2018) 'Chemical and Biochemical Perspectives of Protein Lysine Methylation', *Chemical reviews*. 2018/06/21, 118(14).
- Luscher, A., Moynié, L., Auguste, P. Saint, Bumann, D., Mazza, L., Pletzer, D., Naismith, J. H. and Köhler, T. (2018) 'TonB-Dependent Receptor Repertoire of *Pseudomonas aeruginosa* for Uptake of Siderophore-Drug Conjugates', *Antimicrobial agents and chemotherapy*, 62(6).
- Lutz, J. K. and Lee, J. (2011) 'Prevalence and antimicrobial-resistance of *pseudomonas aeruginosa* in swimming pools and hot tubs', *International Journal of Environmental Research and Public Health*, 8(2).
- Lyczak, J. B., Cannon, C. L. and Pier, G. B. (2000) 'Establishment of *Pseudomonas aeruginosa* infection: Lessons from a versatile opportunist', *Microbes and Infection*.
- Lyczak, J. B., Cannon, C. L. and Pier, G. B. (2002) 'Lung infections associated with cystic fibrosis', *Clinical microbiology reviews*, 15(2).
- Ma, Q. and Wood, T. K. (2011) 'Protein acetylation in prokaryotes increases stress resistance', *Biochemical and Biophysical Research Communications*, 410(4).
- MacDonald, I. A. and Kuehn, M. J. (2013) 'Stress-induced outer membrane vesicle production by *Pseudomonas aeruginosa*', *Journal of Bacteriology*, 195(13).
- Malhotra, S., Hayes, D. and Wozniak, D. J. (2019a) 'Cystic fibrosis and *pseudomonas aeruginosa*: The host-microbe interface', *Clinical Microbiology Reviews*, 32(3).
- Malhotra, S., Hayes, D. and Wozniak, D. J. (2019b) 'Mucoid *Pseudomonas aeruginosa* and regional inflammation in the cystic fibrosis lung', *Journal of Cystic Fibrosis*, 18(6).
- Mann, E. E. and Wozniak, D. J. (2012) '*Pseudomonas* biofilm matrix composition and niche biology', *FEMS microbiology reviews*. 2012/01/23, 36(4).
- Marquart, M. E., Caballero, A. R., Chomnawang, M., Thibodeaux, B. A., Twining, S. S. and O'Callaghan, R. J. (2005) 'Identification of a novel secreted protease from *Pseudomonas aeruginosa* that causes corneal erosions', *Investigative Ophthalmology and Visual Science*, 46(10).
- Marquart, M. E., Dajcs, J. J., Caballero, A. R., Thibodeaux, B. A. and O'Callaghan, R. J. (2005) 'Calcium and magnesium enhance the production of *Pseudomonas aeruginosa* protease IV, a corneal virulence factor.', *Medical microbiology and immunology*, 194(1–2).
- Masuda, N., Sakagawa, E., Ohya, S., Gotoh, N., Tsujimoto, H. and Nishino, T. (2000) 'Substrate specificities of MexAB-OprM, MexCD-OprJ, and MexXY-OprM efflux pumps in *Pseudomonas aeruginosa*', *Antimicrobial Agents and Chemotherapy*, 44(12).

- Matsui, H., Sano, Y., Ishihara, H. and Shinomiya, T. (1993) 'Regulation of pyocin genes in *Pseudomonas aeruginosa* by positive (prtN) and negative (prtR) regulatory genes', *Journal of Bacteriology*, 175(5).
- Maurice, N. M., Bedi, B. and Sadikot, R. T. (2018) 'Pseudomonas aeruginosa biofilms: Host response and clinical implications in lung infections', *American Journal of Respiratory Cell and Molecular Biology*, 58(4).
- McCarthy, J., Hopwood, F., Oxley, D., Laver, M., Castagna, A., Righetti, P. G., Williams, K. and Herbert, B. (2003) 'Carbamylation of proteins in 2-D electrophoresis - Myth or reality?', *Journal of Proteome Research*, 2(3).
- McCaughey, L. C., Ritchie, N. D., Douce, G. R., Evans, T. J. and Walker, D. (2016) 'Efficacy of species-specific protein antibiotics in a murine model of acute *Pseudomonas aeruginosa* lung infection', *Scientific Reports*, 6(July).
- McDaniel, C. T., Panmanee, W. and Hassett, D. J. (2015) 'An Overview of Infections in Cystic Fibrosis Airways and the Role of Environmental Conditions on *Pseudomonas aeruginosa* Biofilm Formation and Viability', in Panmanee, W. (ed.) *Cystic Fibrosis in the Light of New Research*. Rijeka: IntechOpen.
- McIver, K. S., Kessler, E. and Ohman, D. E. (2004) 'Identification of residues in the *Pseudomonas aeruginosa* elastase propeptide required for chaperone and secretion activities', *Microbiology*, 150(12).
- Mendelson, M. H., Gurtman, A., Szabo, S., Neibart, E., Meyers, B. R., Policar, M., Cheung, T. W., Lillienfeld, D., Hammer, G., Reddy, S., Choi, K. and Hirschman, S. Z. (1994) 'Pseudomonas aeruginosa bacteremia in patients with aids', *Clinical Infectious Diseases*, 18(6).
- Metruccio, M. M. E., Evans, D. J., Gabriel, M. M., Kadurugamuwa, J. L. and Fleiszig, S. M. J. (2016) 'Pseudomonas aeruginosa outer membrane vesicles triggered by human mucosal fluid and lysozyme can prime host tissue surfaces for bacterial adhesion', *Frontiers in Microbiology*, 7(JUN).
- Meuskens, I., Saragliadis, A., Leo, J. C. and Linke, D. (2019) 'Type V secretion systems: An overview of passenger domain functions', *Frontiers in Microbiology*, 10(MAY).
- Michalska, M. and Wolf, P. (2015) 'Pseudomonas Exotoxin A: Optimized by evolution for effective killing', *Frontiers in Microbiology*, 6(SEP).
- Minguez, P., Parca, L., Diella, F., Mende, D. R., Kumar, R., Helmer-Citterich, M., Gavin, A.-C., van Noort, V. and Bork, P. (2012) 'Deciphering a global network of functionally associated post-translational modifications.', *Molecular systems biology*, 8.
- Mittal, R., Aggarwal, S., Sharma, S., Chhibber, S. and Harjai, K. (2009) 'Urinary tract infections caused by *Pseudomonas aeruginosa*: A minireview', *Journal of Infection and Public Health*, 2(3).
- Mizuno, Y., Nagano-Shoji, M., Kubo, S., Kawamura, Y., Yoshida, A., Kawasaki, H., Nishiyama, M., Yoshida, M. and Kosono, S. (2016) 'Altered acetylation and succinylation profiles in *Corynebacterium glutamicum* in response to conditions inducing glutamate overproduction', *MicrobiologyOpen*, 5(1).
- Moradali, M. F., Ghods, S. and Rehm, B. H. A. (2017) 'Pseudomonas aeruginosa lifestyle: A paradigm for adaptation, survival, and persistence', *Frontiers in Cellular and Infection Microbiology*, 7(FEB).

- Morlon-Guyot, J., Méré, J., Bonhoure, A. and Beaumelle, B. (2009) 'Processing of *Pseudomonas aeruginosa* exotoxin A is dispensable for cell intoxication', *Infection and Immunity*, 77(7).
- Mougous, J. D., Cuff, M. E., Raunser, S., Shen, A., Zhou, M., Gifford, C. A., Goodman, A. L., Joachimiak, G., Ordoñez, C. L., Lory, S., Walz, T., Joachimiak, A. and Mekalanos, J. J. (2006) 'A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus', *Science*, 312(5779).
- Mougous, J. D., Gifford, C. A., Ramsdell, T. L. and Mekalanos, J. J. (2007) 'Threonine phosphorylation post-translationally regulates protein secretion in *Pseudomonas aeruginosa*', *Nature Cell Biology*, 9(7).
- Mulani, M. S., Kamble, E. E., Kumkar, S. N., Tawre, M. S. and Pardesi, K. R. (2019) 'Emerging strategies to combat ESKAPE pathogens in the era of antimicrobial resistance: A review', *Frontiers in Microbiology*, 10(APR).
- Mulcahy, L. R., Isabella, V. M. and Lewis, K. (2014) '*Pseudomonas aeruginosa* Biofilms in Disease', *Microbial Ecology*, 68(1).
- Murray, T. S. and Kazmierczak, B. I. (2008) '*Pseudomonas aeruginosa* exhibits sliding motility in the absence of type IV pili and flagella', *Journal of Bacteriology*, 190(8).
- N'Diaye, A. R., Borrel, V., Racine, P. J., Clamens, T., Depayras, S., Maillot, O., Schaack, B., Chevalier, S., Lesouhaitier, O. and Feuilloley, M. G. J. (2019) 'Mechanism of action of the moonlighting protein Eftu as a Substance P sensor in *Bacillus cereus*', *Scientific Reports*, 9(1).
- Nakayama, K., Takashima, K., Ishihara, H., Shinomiya, T., Kageyama, M., Kanaya, S., Ohnishi, M., Murata, T., Mori, H. and Hayashi, T. (2000) 'The R-type pyocin of *Pseudomonas aeruginosa* is related to P2 phage, and the F-type is related to lambda phage', *Molecular Microbiology*, 38(2).
- Nally, J. E., Grassmann, A. A., Planchon, S., Sergeant, K., Renaut, J., Seshu, J., McBride, A. J. and Caimano, M. J. (2017) 'Pathogenic leptospires modulate protein expression and post-translational modifications in response to mammalian host signals', *Frontiers in Cellular and Infection Microbiology*, 7(AUG).
- Nathwani, D., Raman, G., Sulham, K., Gavaghan, M. and Menon, V. (2014) 'Clinical and economic consequences of hospital-acquired resistant and multidrug-resistant *Pseudomonas aeruginosa* infections: A systematic review and meta-analysis', *Antimicrobial Resistance and Infection Control*, 3(1).
- Newman, J. W., Floyd, R. V. and Fothergill, J. L. (2017) 'The contribution of *Pseudomonas aeruginosa* virulence factors and host factors in the establishment of urinary tract infections', *FEMS Microbiology Letters*, 364(15).
- Nguyen, Y., Harvey, H., Sugiman-Marangos, S., Bell, S. D., Buensuceso, R. N. C., Junop, M. S. and Burrows, L. L. (2015) 'Structural and functional studies of the *Pseudomonas aeruginosa* Minor Pilin, Pile', *Journal of Biological Chemistry*, 290(44).
- Noach, I., Ficko-Blean, E., Pluvinau, B., Stuart, C., Jenkins, M. L., Brochu, D., Buenbrazo, N., Wakarchuk, W., Burke, J. E., Gilbert, M. and Boraston, A. B. (2017) 'Recognition of protein-linked glycans as a determinant of peptidase activity', *Proceedings of the National Academy of Sciences of the United States of America*, 114(5).
- Nouwens, A. S., Beatson, S. A., Whitchurch, C. B., Walsh, B. J., Schweizer, H. P., Mattick, J. S. and Cordwell, S. J. (2003) 'Proteome analysis of extracellular proteins regulated by the las

and rhl quorum sensing systems in *Pseudomonas aeruginosa* PAO1', *Microbiology*, 149(5).

Nouwens, A. S., Willcox, M. D. P., Walsh, B. J. and Cordwell, S. J. (2002) 'Proteomic comparison of membrane and extracellular proteins from invasive (PAO1) and cytotoxic (6206) strains of *Pseudomonas aeruginosa*', *Proteomics*, 2(9).

O'callaghan, R., Caballero, A., Tang, A. and Bierdeman, M. (2019) 'Pseudomonas aeruginosa keratitis: Protease iv and pasp as corneal virulence mediators', *Microorganisms*, 7(9).

Oh, J., Li, X. H., Kim, S. K. and Lee, J. H. (2017) 'Post-secretional activation of Protease IV by quorum sensing in *Pseudomonas aeruginosa*', *Scientific Reports*, 7(1).

Olson, J. C. and Ohman, D. E. (1992) 'Efficient production and processing of elastase and LasA by *Pseudomonas aeruginosa* require zinc and calcium ions', *Journal of Bacteriology*, 174(12).

Oluyombo, O., Penfold, C. N. and Diggle, S. P. (2019) 'Competition in biofilms between cystic fibrosis isolates of *pseudomonas aeruginosa* is shaped by R-pyocins', *mBio*, 10(1).

Ouidir, T., Cosette, P., Jouenne, T. and Hardouin, J. (2015) 'Proteomic profiling of lysine acetylation in *Pseudomonas aeruginosa* reveals the diversity of acetylated proteins', *Proteomics*, 15(13).

Ouidir, T., Jarnier, F., Cosette, P., Jouenne, T. and Hardouin, J. (2014) 'Extracellular Ser/Thr/Tyr phosphorylated proteins of *Pseudomonas aeruginosa* PA14 strain', *Proteomics*, 14(17–18).

Ouidir, T., Jouenne, T. and Hardouin, J. (2016) 'Post-translational modifications in *Pseudomonas aeruginosa* revolutionized by proteomic analysis', *Biochimie*, 125.

Overhage, J., Bains, M., Brazas, M. D. and Hancock, R. E. W. (2008) 'Swarming of *Pseudomonas aeruginosa* is a complex adaptation leading to increased production of virulence factors and antibiotic resistance', *Journal of Bacteriology*, 190(8).

Owings, J. P., Kuiper, E. G., Prezioso, S. M., Meisner, J., Varga, J. J., Zelinskaya, N., Dammer, E. B., Duong, D. M., Seyfried, N. T., Albertí, S., Conn, G. L. and Goldberg, J. B. (2016) 'Pseudomonas aeruginosa EftM is a thermoregulated', *Journal of Biological Chemistry*, 291(7).

Oz, T., Guvenek, A., Yildiz, S., Karaboga, E., Tamer, Y. T., Mumcuyan, N., Ozan, V. B., Senturk, G. H., Cokol, M., Yeh, P. and Toprak, E. (2014) 'Strength of selection pressure is an important parameter contributing to the complexity of antibiotic resistance evolution', *Molecular Biology and Evolution*, 31(9).

Pan, J., Chen, R., Li, C., Li, W. and Ye, Z. (2015) 'Global Analysis of Protein Lysine Succinylation Profiles and Their Overlap with Lysine Acetylation in the Marine Bacterium *Vibrio parahaemolyticus*', *Journal of Proteome Research*, 14(10).

Pang, C. N. I., Gasteiger, E. and Wilkins, M. R. (2010) 'Identification of arginine- and lysine-methylation in the proteome of *Saccharomyces cerevisiae* and its functional implications', *BMC Genomics*, 11(1).

Pang, Z., Raudonis, R., Glick, B. R., Lin, T. J. and Cheng, Z. (2019) 'Antibiotic resistance in *Pseudomonas aeruginosa*: mechanisms and alternative therapeutic strategies', *Biotechnology Advances*, 37(1).

Park, S. and Galloway, D. R. (1995) 'Purification and characterization of LasD: a second staphylolytic proteinase produced by *Pseudomonas aeruginosa*.' *Molecular microbiology*, 16(2).

Parkin, J. and Khalid, S. (2014) 'Atomistic molecular-dynamics simulations enable prediction of the arginine permeation pathway through occd1/oprd from pseudomonas aeruginosa', *Biophysical Journal*, 107(8).

Patankar, Y. R., Lovewell, R. R., Poynter, M. E., Jyot, J., Kazmierczak, B. I. and Berwin, B. (2013) 'Flagellar motility is a key determinant of the magnitude of the inflammasome response to *Pseudomonas aeruginosa*', *Infection and Immunity*, 81(6).

Penterman, J., Singh, P. K. and Walker, G. C. (2014) 'Biological cost of pyocin production during the SOS response in *Pseudomonas aeruginosa*', *Journal of Bacteriology*, 196(18).

Pericolini, E., Colombari, B., Ferretti, G., Iseppi, R., Ardizzoni, A., Girardis, M., Sala, A., Peppoloni, S. and Blasi, E. (2018) 'Real-time monitoring of *Pseudomonas aeruginosa* biofilm formation on endotracheal tubes in vitro', *BMC Microbiology*, 18(1).

Persat, A., Inclan, Y. F., Engel, J. N., Stone, H. A. and Gitai, Z. (2015) 'Type IV pili mechanochemically regulate virulence factors in *Pseudomonas aeruginosa*', *Proceedings of the National Academy of Sciences of the United States of America*, 112(24).

Peters, J. E. and Galloway, D. R. (1990) 'Purification and characterization of an active fragment of the LasA protein from *Pseudomonas aeruginosa*: Enhancement of elastase activity', *Journal of Bacteriology*, 172(5).

Pethe, K., Bifani, P., Drobecq, H., Sergheraert, C., Debie, A. S., Loch, C. and Menozzi, F. D. (2002) 'Mycobacterial heparin-binding hemagglutinin and laminin-binding protein share antigenic methyllysines that confer resistance to proteolysis', *Proceedings of the National Academy of Sciences of the United States of America*, 99(16).

Poole, K. (2005) 'Aminoglycoside resistance in *Pseudomonas aeruginosa*', *Antimicrobial Agents and Chemotherapy*, 49(2).

Prabakaran, S., Lippens, G., Steen, H. and Gunawardena, J. (2012) 'Post-translational modification: nature's escape from genetic imprisonment and the basis for dynamic information encoding.', *Wiley interdisciplinary reviews. Systems biology and medicine*, 4(6).

Prezioso, S. M., Duong, D. M., Kuiper, E. G., Deng, Q., Albertí, S., Conn, G. L. and Goldberg, J. B. (2019) 'Trimethylation of Elongation Factor-Tu by the Dual Thermoregulated Methyltransferase EftM Does Not Impact Its Canonical Function in Translation', *Scientific Reports*, 9(1).

Príncipe, A., Fernandez, M., Torasso, M., Godino, A. and Fischer, S. (2018) 'Effectiveness of tailocins produced by *Pseudomonas fluorescens* SF4c in controlling the bacterial-spot disease in tomatoes caused by *Xanthomonas vesicatoria*', *Microbiological Research*, 212–213(April).

Pye, C. (2018) 'Pseudomonas otitis externa in dogs', *The Canadian Veterinary Journal*, 59(November).

Qian, L., Nie, L., Chen, M., Liu, P., Zhu, J., Zhai, L., Tao, S. C., Cheng, Z., Zhao, Y. and Tan, M. (2016) 'Global profiling of protein lysine malonylation in *Escherichia coli* reveals its role in energy metabolism', *Journal of Proteome Research*, 15(6).

Qiu, W. R., Xiao, X., Lin, W. Z. and Chou, K. C. (2014) 'IMethyl-PseAAC: Identification of protein methylation sites via a pseudo amino acid composition approach', *BioMed Research International*, 2014.

Rabilloud, T. and Lelong, C. (2011) 'Two-dimensional gel electrophoresis in proteomics: A tutorial', *Journal of Proteomics*, 74(10).



Ravichandran, A., Sugiyama, N., Tomita, M., Swarup, S. and Ishihama, Y. (2009) 'Ser/Thr/Tyr phosphoproteome analysis of pathogenic and non-pathogenic *Pseudomonas* species', *Proteomics*, 9(10).

Redero, M., Aznar, J. and Prieto, A. I. (2020) 'Antibacterial efficacy of R-type pyocins against *Pseudomonas aeruginosa* on biofilms and in a murine model of acute lung infection', *Journal of Antimicrobial Chemotherapy*, (June).

Redero, M., López-Causapé, C., Aznar, J., Oliver, A., Blázquez, J. and Prieto, A. I. (2018) 'Susceptibility to R-pyocins of *Pseudomonas aeruginosa* clinical isolates from cystic fibrosis patients', *Journal of Antimicrobial Chemotherapy*, 73(10).

Ren, J., Sang, Y., Lu, J. and Yao, Y. F. (2017) 'Protein Acetylation and Its Role in Bacterial Virulence', *Trends in Microbiology*, 25(9).

Ren, J., Sang, Y., Ni, J., Tao, J., Lu, J., Zhao, M. and Yao, Y. F. (2015) 'Acetylation regulates survival of *Salmonella enterica* serovar typhimurium under acid stress', *Applied and Environmental Microbiology*, 81(17).

Righetti, P. G. (2006) 'Real and imaginary artefacts in proteome analysis via two-dimensional maps', *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*, 841(1–2).

Robert, V., Filloux, A. and Michel, G. P. F. (2005) 'Role of XcpP in the functionality of the *Pseudomonas aeruginosa* secretin', *Research in Microbiology*, 156(8).

Ruffin, M. and Brochiero, E. (2019) 'Repair process impairment by *pseudomonas aeruginosa* in epithelial tissues: Major features and potential therapeutic avenues', *Frontiers in Cellular and Infection Microbiology*, 9(MAY).

Russell, A. B., Leroux, M., Hathazi, K., Agnello, D. M., Ishikawa, T., Wiggins, P. A., Wai, S. N. and Mougous, J. D. (2013) 'Diverse type VI secretion phospholipases are functionally plastic antibacterial effectors', *Nature*, 496(7446).

Russell, A. B., Peterson, S. B. and Mougous, J. D. (2014) 'Type VI secretion system effectors: Poisons with a purpose', *Nature Reviews Microbiology*, 12(2).

Sakatos, A., Babunovic, G. H., Chase, M. R., Dills, A., Leszyk, J., Rosebrock, T., Bryson, B. and Fortune, S. M. (2018) 'Posttranslational modification of a histone-like protein regulates phenotypic resistance to isoniazid in mycobacteria', *Science Advances*, 4(5).

Salacha, R., Kovačić, F., Brochier-Armanet, C., Wilhelm, S., Tommassen, J., Filloux, A., Voulhoux, R. and Bleves, S. (2010) 'The *Pseudomonas aeruginosa* patatin-like protein PlpD is the archetype of a novel Type V secretion system', *Environmental Microbiology*, 12(6).

Samsudin, F. and Khalid, S. (2019) 'Movement of Arginine through OprD: The Energetics of Permeation and the Role of Lipopolysaccharide in Directing Arginine to the Protein', *Journal of Physical Chemistry B*, 123(13).

Sandri, A., Ortombina, A., Boschi, F., Cremonini, E., Boaretti, M., Sorio, C., Melotti, P., Bergamini, G. and Lleo, M. (2018) 'Inhibition of *pseudomonas aeruginosa* secreted virulence factors reduces lung inflammation in CF mice', *Virulence*, 9(1).

Sang, Y., Ren, J., Qin, R., Liu, S., Cui, Z., Cheng, S., Liu, X., Lu, J., Tao, J. and Yao, Y. F. (2017) 'Acetylation Regulating Protein Stability and DNA-Binding Ability of HilD, thus Modulating *Salmonella Typhimurium* Virulence', *Journal of Infectious Diseases*, 216(8).

Santajit, S. and Indrawattana, N. (2016) 'Mechanisms of Antimicrobial Resistance in ESKAPE

Pathogens', *BioMed Research International*, 2016.

Sato, H., Hunt, M. L., Weiner, J. J., Hansen, A. T. and Frank, D. W. (2011) 'Modified needle-tip PcrV proteins reveal distinct phenotypes relevant to the control of type III secretion and intoxication by *Pseudomonas aeruginosa*', *PLoS ONE*, 6(3).

Schiessl, K. T., Hu, F., Jo, J., Nazia, S. Z., Wang, B., Price-Whelan, A., Min, W. and Dietrich, L. E. P. (2019) 'Phenazine production promotes antibiotic tolerance and metabolic heterogeneity in *Pseudomonas aeruginosa* biofilms', *Nature Communications*, 10(1).

Schiessl, K. T., Janssen, E. M. L., Kraemer, S. M., McNeill, K. and Ackermann, M. (2017) 'Magnitude and mechanism of siderophore-mediated competition at low iron solubility in the *Pseudomonas aeruginosa* pyochelin system', *Frontiers in Microbiology*, 8(OCT).

Schmidt, A., Trentini, D. B., Spiess, S., Fuhrmann, J., Ammerer, G., Mechtler, K. and Clausen, T. (2014) 'Quantitative phosphoproteomics reveals the role of protein arginine phosphorylation in the bacterial stress response', *Molecular and Cellular Proteomics*, 13(2).

Schmidt, J., Müsken, M., Becker, T., Magnowska, Z., Bertinetti, D., Möller, S., Zimmermann, B., Herberg, F. W., Jänsch, L. and Häussler, S. (2011) 'The *Pseudomonas aeruginosa* chemotaxis methyltransferase CheR1 impacts on bacterial surface sampling', *PLoS ONE*, 6(3).

Scholl, D. and Martin, D. W. (2008) 'Antibacterial efficacy of R-type pyocins towards *Pseudomonas aeruginosa* in a murine peritonitis model', *Antimicrobial Agents and Chemotherapy*, 52(5).

Scott, N. E., Hare, N. J., White, M. Y., Manos, J. and Cordwell, S. J. (2013) 'Secretome of transmissible *Pseudomonas aeruginosa* AES-1R grown in a cystic fibrosis lung-like environment', *Journal of Proteome Research*, 12(12).

Sharma, G., Rao, S., Bansal, A., Dang, S., Gupta, S. and Gabrani, R. (2014) '*Pseudomonas aeruginosa* biofilm: Potential therapeutic targets', *Biologicals*, 42(1).

Shi, Q., Huang, C., Xiao, T., Wu, Z. and Xiao, Y. (2019) 'A retrospective analysis of *Pseudomonas aeruginosa* bloodstream infections: Prevalence, risk factors, and outcome in carbapenem-susceptible and -non-susceptible infections', *Antimicrobial Resistance and Infection Control*, 8(1).

Shien, D.-M., Lee, T.-Y., Chang, W.-C., Hsu, J. B.-K., Horng, J.-T., Hsu, P.-C., Wang, T.-Y. and Huang, H.-D. (2009) 'Incorporating structural characteristics for identification of protein methylation sites', *Journal of Computational Chemistry*, 30(9).

Shrout, J. D., Chopp, D. L., Just, C. L., Hentzer, M., Givskov, M. and Parsek, M. R. (2006) 'The impact of quorum sensing and swarming motility on *Pseudomonas aeruginosa* biofilm formation is nutritionally conditional', *Molecular Microbiology*, 62(5).

Silva, A. M. N., Vitorino, R., Domingues, M. R. M., Spickett, C. M. and Domingues, P. (2013) 'Post-translational modifications and mass spectrometry detection', *Free Radical Biology and Medicine*, 65.

Singhal, A., Arora, G., Virmani, R., Kundu, P., Khanna, T., Sajid, A., Misra, R., Joshi, J., Yadav, V., Samanta, S., Saini, N., Pandey, A. K., Visweswariah, S. S., Hentschker, C., Becher, D., Gerth, U. and Singh, Y. (2015) 'Systematic analysis of mycobacterial acylation reveals first example of acylation-mediated regulation of enzyme activity of a bacterial phosphatase', *Journal of Biological Chemistry*, 290(43).

Skopelja-Gardner, S., Theprungsirikul, J., Lewis, K. A., Hammond, J. H., Carlson, K. M., Hazlett, H. F., Nymon, A., Nguyen, D., Berwin, B. L., Hogan, D. A. and Rigby, W. F. C. (2019)

'Regulation of Pseudomonas aeruginosa-Mediated Neutrophil Extracellular Traps', *Frontiers in Immunology*, 10(JULY).

Smith, E. E., Buckley, D. G., Wu, Z., Saenphimmachak, C., Hoffman, L. R., D'Argenio, D. A., Miller, S. I., Ramsey, B. W., Speert, D. P., Moskowitz, S. M., Burns, J. L., Kaul, R. and Olson, M. V. (2006) 'Genetic adaptation by Pseudomonas aeruginosa to the airways of cystic fibrosis patients', *Proceedings of the National Academy of Sciences of the United States of America*, 103(22).

Soares, N. C., Spät, P., Krug, K. and MacEk, B. (2013) 'Global dynamics of the Escherichia coli proteome and phosphoproteome during growth in minimal medium', *Journal of Proteome Research*, 12(6).

Soares, N. C., Spät, P., Méndez, J. A., Nakedi, K., Aranda, J. and Bou, G. (2014) 'Ser/Thr/Tyr phosphoproteome characterization of Acinetobacter baumannii: Comparison between a reference strain and a highly invasive multidrug-resistant clinical isolate', *Journal of Proteomics*, 102(lidmm).

Sokol, P. A., Kooi, C., Hodges, R. S., Cachia, P. and Woods, D. E. (2000) 'Immunization with a Pseudomonas aeruginosa elastase peptide reduces severity of experimental lung infections due to P. aeruginosa or Burkholderia cepacia', *Journal of Infectious Diseases*, 181(5).

Soong, G., Parker, D., Magargee, M. and Prince, A. S. (2008) 'The type III toxins of Pseudomonas aeruginosa disrupt epithelial barrier function', *Journal of Bacteriology*, 190(8).

Sosnay, P. R., Siklosi, K. R., Van Goor, F., Kaniecki, K., Yu, H., Sharma, N., Ramalho, A. S., Amaral, M. D., Dorfman, R., Zielenski, J., Masica, D. L., Karchin, R., Millen, L., Thomas, P. J., Patrinos, G. P., Corey, M., Lewis, M. H., Rommens, J. M., Castellani, C., *et al.* (2013) 'Defining the disease liability of variants in the cystic fibrosis transmembrane conductance regulator gene', *Nature genetics*. 2013/08/25, 45(10).

Spencer, J., Murphy, L. M., Connors, R., Sessions, R. B. and Gamblin, S. J. (2010) 'Crystal structure of the lasA virulence factor from Pseudomonas aeruginosa: Substrate specificity and mechanism of M23 metallopeptidases', *Journal of Molecular Biology*, 396(4).

Sprung, R., Chen, Y., Zhang, K., Cheng, D., Zhang, T., Peng, J. and Zhao, Y. (2008) 'Identification and validation of eukaryotic aspartate and glutamate methylation in proteins', *Journal of proteome research*, 7(3).

Sriramulu, D. D., Nimtz, M. and Romling, U. (2005) 'Proteome analysis reveals adaptation of Pseudomonas aeruginosa to the cystic fibrosis lung environment', *Proteomics*, 5(14).

Stallcup, M. R. (2001) 'Role of protein methylation in chromatin remodeling and transcriptional regulation', *Oncogene*, 20(24).

Stapleton, F., Dart, J. K. G., Seal, D. V. and Matheson, M. (1995) 'Epidemiology of Pseudomonas aeruginosa keratitis in contact lens wearers', *Epidemiology and Infection*, 114(3).

Stewart, P. E., Carroll, J. A., Rennee Olano, L., Sturdevant, D. E. and Rosa, P. A. (2016) 'Multiple posttranslational modifications of Leptospira biflexa proteins as revealed by proteomic analysis', *Applied and Environmental Microbiology*, 82(4).

Stickland, H. G., Davenport, P. W., Lilley, K. S., Griffin, J. L. and Welch, M. (2010) 'Mutation of nfxB causes global changes in the physiology and metabolism of pseudomonas aeruginosa', *Journal of Proteome Research*, 9(6).

Stover, C. K., Pham, X. Q., Erwin, A. L., Mizoguchi, S. D., Warrenner, P., Hickey, M. J.,

- Brinkman, F. S. L., Hufnagle, W. O., Kowalik, D. J., Lagrou, M., Garber, R. L., Goltry, L., Tolentino, E., Westbrook-Wadman, S., Yuan, Y., Brody, L. L., Coulter, S. N., Folger, K. R., Kas, A., *et al.* (2000) 'Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen', *Nature*, 406(6799).
- Su, M. G., Weng, J. T. Y., Hsu, J. B. K., Huang, K. Y., Chi, Y. H. and Lee, T. Y. (2017) 'Investigation and identification of functional post-translational modification sites associated with drug binding and protein-protein interactions', *BMC Systems Biology*, 11(Suppl 7).
- Sun, J., LaRock, D. L., Skowronski, E. A., Kimmey, J. M., Olson, J., Jiang, Z., O'Donoghue, A. J., Nizet, V. and LaRock, C. N. (2020) 'The *Pseudomonas aeruginosa* protease LasB directly activates IL-1 $\beta$ ', *EBioMedicine*, 60.
- Sundström, J., Jacobson, K., Munck-Wikland, E. and Ringertz, S. (1996) '*Pseudomonas aeruginosa* in Otitis Externa: A Particular Variety of the Bacteria?', *Archives of Otolaryngology–Head & Neck Surgery*, 122(8).
- Suriyanarayanan, T., Periasamy, S., Lin, M. H., Ishihama, Y. and Swarup, S. (2016) 'Flagellin FliC phosphorylation affects type 2 protease secretion and biofilm dispersal in *Pseudomonas aeruginosa* PAO1', *PLoS ONE*, 11(10).
- Swatton, J. E., Davenport, P. W., Maunders, E. A., Griffin, J. L., Lilley, K. S. and Welch, M. (2016) 'Impact of azithromycin on the quorum sensing-controlled proteome of *pseudomonas aeruginosa*', *PLoS ONE*, 11(1).
- Sy, A., Srinivasan, M., Mascarenhas, J., Lalitha, P., Rajaraman, R., Ravindran, M., Oldenburg, C. E., Ray, K. J., Glidden, D., Zegans, M. E., McLeod, S. D., Lietman, T. M. and Acharya, N. R. (2012) '*Pseudomonas aeruginosa* keratitis: Outcomes and response to corticosteroid treatment', *Investigative Ophthalmology and Visual Science*, 53(1).
- Tacconelli, E., Carrara, E., Savoldi, A., Harbarth, S., Mendelson, M., Monnet, D. L., Pulcini, C., Kahlmeter, G., Kluytmans, J., Carmeli, Y., Ouellette, M., Outtersen, K., Patel, J., Cavaleri, M., Cox, E. M., Houchens, C. R., Grayson, M. L., Hansen, P., Singh, N., *et al.* (2018) 'Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis', *The Lancet Infectious Diseases*, 18(3).
- Takeya, K. and Amako, K. (1966) 'A rod-shaped *Pseudomonas* phage.', *Virology*, 28(1).
- Tan, M. W., Mahajan-Miklos, S. and Ausubel, F. M. (1999) 'Killing of *Caenorhabditis elegans* by *Pseudomonas aeruginosa* used to model mammalian bacterial pathogenesis', *Proceedings of the National Academy of Sciences of the United States of America*, 96(2).
- Tang, Y., Romano, F. B., Breña, M. and Heuck, A. P. (2018) 'The *Pseudomonas aeruginosa* type III secretion translocator PopB assists the insertion of the PopD translocator into host cell membranes', *Journal of Biological Chemistry*, 293(23).
- Temmerman, S., Pethe, K., Parra, M., Alonso, S., Rouanet, C., Pickett, T., Drowart, A., Debie, A. S., Delogu, G., Menozzi, F. D., Sergheraert, C., Brennan, M. J., Mascart, F. and Loch, C. (2004) 'Methylation-dependent T cell immunity to *Mycobacterium tuberculosis* heparin-binding hemagglutinin', *Nature Medicine*, 10(9).
- Temporini, C., Calleri, E., Massolini, G. and Caccialanza, G. (2008) 'Integrated analytical strategies for the study of phosphorylation and glycosylation in proteins', *Mass Spectrometry Reviews*, 27(3).
- Terada, L. S., Johansen, K. A., Nowbar, S., Vasil, A. I. and Vasil, M. L. (1999) '*Pseudomonas aeruginosa* hemolytic phospholipase C suppresses neutrophil respiratory burst activity', *Infection and Immunity*, 67(5).

- Termine, E. and Michel, G. P. F. (2009) 'Transcriptome and secretome analyses of the adaptive response of *Pseudomonas aeruginosa* to suboptimal growth temperature', *International Microbiology*, 12(1).
- Thao, S. and Escalante-Semerena, J. C. (2011) 'Control of protein function by reversible N $\epsilon$ -lysine acetylation in bacteria', *Current opinion in microbiology*. 2011/01/14, 14(2).
- Thayer, M. M., Flaherty, K. M. and McKay, D. B. (1991) 'Three-dimensional structure of the elastase of *Pseudomonas aeruginosa* at 1.5-Å resolution', *Journal of Biological Chemistry*, 266(5).
- Tian, Z., Cheng, S., Xia, B., Jin, Y., Bai, F., Cheng, Z., Jin, S., Liu, X. and Wu, W. (2019) 'Pseudomonas aeruginosa ExsA Regulates a Metalloprotease, ImpA, That Inhibits Phagocytosis of Macrophages', *Infection and Immunity*. Edited by A. J. Bäuml, 87(12).
- Tielen, P., Rosenau, F., Wilhelm, S., Jaeger, K. E., Flemming, H. C. and Wingender, J. (2010) 'Extracellular enzymes affect biofilm formation of mucoid *Pseudomonas aeruginosa*', *Microbiology*, 156(7).
- Tingpej, P., Smith, L., Rose, B., Zhu, H., Conibear, T., Al Nassafi, K., Manos, J., Elkins, M., Bye, P., Willcox, M., Bell, S., Wainwright, C. and Harbour, C. (2007) 'Phenotypic characterization of clonal and nonclonal *Pseudomonas aeruginosa* strains isolated from lungs of adults with cystic fibrosis', *Journal of Clinical Microbiology*, 45(6).
- Tiwari, S., Jamal, S. B., Hassan, S. S., Carvalho, P. V. S. D., Almeida, S., Barh, D., Ghosh, P., Silva, A., Castro, T. L. P. and Azevedo, V. (2017) 'Two-component signal transduction systems of pathogenic bacteria as targets for antimicrobial therapy: An overview', *Frontiers in Microbiology*, 8(OCT).
- Tiwari, V. (2019) 'Post-translational modification of ESKAPE pathogens as a potential target in drug discovery', *Drug Discovery Today*, 24(3).
- Touw, W. G., Baakman, C., Black, J., Te Beek, T. A. H., Krieger, E., Joosten, R. P. and Vriend, G. (2015) 'A series of PDB-related databanks for everyday needs', *Nucleic Acids Research*, 43(D1).
- Trentini, D. B., Suskiewicz, M. J., Heuck, A., Kurzbauer, R., Deszcz, L., Mechtler, K. and Clausen, T. (2016) 'Arginine phosphorylation marks proteins for degradation by a Clp protease', *Nature*, 539(7627).
- Turnbull, L., Toyofuku, M., Hynen, A. L., Kurosawa, M., Pessi, G., Petty, N. K., Osvath, S. R., Cárcamo-Oyarce, G., Gloag, E. S., Shimoni, R., Omasits, U., Ito, S., Yap, X., Monahan, L. G., Cavaliere, R., Ahrens, C. H., Charles, I. G., Nomura, N., Eberl, L., *et al.* (2016) 'Explosive cell lysis as a mechanism for the biogenesis of bacterial membrane vesicles and biofilms', *Nature Communications*, 7.
- Turner, K. H., Everett, J., Trivedi, U., Rumbaugh, K. P. and Whiteley, M. (2014) 'Requirements for *Pseudomonas aeruginosa* Acute Burn and Chronic Surgical Wound Infection', *PLoS Genetics*, 10(7).
- Ueda, A. and Wood, T. K. (2009) 'Connecting quorum sensing, c-di-GMP, pel polysaccharide, and biofilm formation in *Pseudomonas aeruginosa* through tyrosine phosphatase TpbA (PA3885)', *PLoS Pathogens*, 5(6).
- Ullah, W., Qasim, M., Rahman, H., Jie, Y. and Muhammad, N. (2017) 'Beta-lactamase-producing *Pseudomonas aeruginosa*: Phenotypic characteristics and molecular identification of virulence genes', *Journal of the Chinese Medical Association*, 80(3).

Upritchard, H. G., Cordwell, S. J. and Lamont, I. L. (2008) 'Immunoproteomics to examine cystic fibrosis host interactions with extracellular *Pseudomonas aeruginosa* proteins', *Infection and Immunity*, 76(10).

Vital-Lopez, F. G., Reifman, J. and Wallqvist, A. (2015) 'Biofilm Formation Mechanisms of *Pseudomonas aeruginosa* Predicted via Genome-Scale Kinetic Models of Bacterial Metabolism', *PLoS Computational Biology*, 11(10).

Waite, R. D. and Curtis, M. A. (2009) '*Pseudomonas aeruginosa* PAO1 pyocin production affects population dynamics within mixed-culture biofilms', *Journal of Bacteriology*, 191(4).

Walker, T. S., Bais, H. P., Déziel, E., Schweizer, H. P., Rahme, L. G., Fall, R. and Vivanco, J. M. (2004) '*Pseudomonas aeruginosa*-Plant Root Interactions. Pathogenicity, Biofilm Formation, and Root Exudation', *Plant Physiology*, 134(1).

Wang, G., Chen, H., Xia, Y., Cui, J., Gu, Z., Song, Y., Chen, Y. Q., Zhang, H. and Chen, W. (2013) 'How are the non-classically secreted bacterial proteins released into the extracellular milieu?', *Current Microbiology*, 67(6).

Wang, Q., Zhang, Y., Yang, C., Lin, Y., Yao, J., Li, H., Xie, L., Zhao, W., Yao, Y., Ning, Z.-B., Zeng, R., Xiong, Y., Guan, K.-L., Zhao, S. and Zhao, G.-P. (2010) 'Acetylation of Metabolic Enzymes Coordinates Carbon Source Utilization and Metabolic Flux', *Science*, 327(5968).

Wargo, M. J., Gross, M. J., Rajamani, S., Allard, J. L., Lundblad, L. K. A., Allen, G. B., Vasil, M. L., Leclair, L. W. and Hogan, D. A. (2011) 'Hemolytic phospholipase C inhibition protects lung function during *Pseudomonas aeruginosa* infection', *American Journal of Respiratory and Critical Care Medicine*, 184(3).

Wehmhöner, D., Häussler, S., Tümmeler, B., Jänsch, L., Bredenbruch, F., Wehland, J. and Steinmetz, I. (2003) 'Inter- and intracolonial diversity of the *Pseudomonas aeruginosa* proteome manifests within the secretome', *Journal of Bacteriology*, 185(19).

Weiner, L. M., Webb, A. K., Limbago, B., Dudeck, M. A., Patel, J., Kallen, A. J., Edwards, J. R. and Sievert, D. M. (2016) 'Antimicrobial-Resistant Pathogens Associated With Healthcare-Associated Infections: Summary of Data Reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2011-2014', *Infection control and hospital epidemiology*. 2016/08/30, 37(11).

Weinert, B. T., Schölz, C., Wagner, S. A., Iesmantavicius, V., Su, D., Daniel, J. A. and Choudhary, C. (2013) 'Lysine succinylation is a frequently occurring modification in prokaryotes and eukaryotes and extensively overlaps with acetylation', *Cell Reports*, 4(4).

Weinstein, R. A., Gaynes, R. and Edwards, J. R. (2005) 'Overview of Nosocomial Infections Caused by Gram-Negative Bacilli', *Clinical Infectious Diseases*, 41(6).

Wettstadt, S., Wood, T. E., Fecht, S. and Filloux, A. (2019) 'Delivery of the *Pseudomonas aeruginosa* phospholipase effectors PldA and PldB in a VgrG- And H2-T6SS-dependent manner', *Frontiers in Microbiology*, 10(JULY).

Widjaja, M., Harvey, K. L., Hagemann, L., Berry, I. J., Jarocki, V. M., Raymond, B. B. A., Tacchi, J. L., Gründel, A., Steele, J. R., Padula, M. P., Charles, I. G., Dumke, R. and Djordjevic, S. P. (2017) 'Elongation factor Tu is a multifunctional and processed moonlighting protein', *Scientific Reports*, 7(1).

Wilhelm, S., Gdynia, A., Tielen, P., Rosenau, F. and Jaeger, K. E. (2007) 'The autotransporter esterase EstA of *Pseudomonas aeruginosa* is required for rhamnolipid production, cell motility, and biofilm formation', *Journal of Bacteriology*, 189(18).

- Williams, B. J., Dehnbostel, J. and Blackwell, T. S. (2010) 'Pseudomonas aeruginosa: Host defence in lung diseases', *Respirology*, 15(7).
- Williams McMackin, E. A., Djapgne, L., Corley, J. M. and Yahr, T. L. (2019) 'Fitting Pieces into the Puzzle of Pseudomonas aeruginosa Type III Secretion System Gene Expression', *Journal of Bacteriology*, 201(13).
- Williams, S. R., Gebhart, D., Martin, D. W. and Scholl, D. (2008) 'Retargeting R-type pyocins to generate novel bactericidal protein complexes', *Applied and Environmental Microbiology*, 74(12).
- Winsor, G. L., Griffiths, E. J., Lo, R., Dhillon, B. K., Shay, J. A. and Brinkman, F. S. L. (2016) 'Enhanced annotations and features for comparing thousands of Pseudomonas genomes in the Pseudomonas genome database', *Nucleic Acids Research*, 44(D1).
- Winter, D. L., Abeygunawardena, D., Hart-Smith, G., Erce, M. A. and Wilkins, M. R. (2015) 'Lysine methylation modulates the protein-protein interactions of yeast cytochrome C Cyc1p.', *Proteomics*, 15(13).
- Witchell, T. D., Eshghi, A., Nally, J. E., Hof, R., Boulanger, M. J., Wunder, E. A., Ko, A. I., Haake, D. A. and Cameron, C. E. (2014) 'Post-translational Modification of LipL32 during Leptospira interrogans Infection', *PLoS Neglected Tropical Diseases*, 8(10).
- Witze, E. S., Old, W. M., Resing, K. A. and Ahn, N. G. (2007) 'Mapping protein post-translational modifications with mass spectrometry', *Nature Methods*, 4(10).
- Wood, T. L., Gong, T., Zhu, L., Miller, J., Miller, D. S., Yin, B. and Wood, T. K. (2018) 'Rhamnolipids from Pseudomonas aeruginosa disperse the biofilms of sulfate-reducing bacteria', *Biofilms and Microbiomes*, 4(1).
- World Health Organisation (2017) *WHO publishes list of bacteria for which new antibiotics are urgently needed*.
- Wozniak, D. J., Tiwari, K. B., Soufan, R. and Jayaswal, R. K. (2012) 'The mcsB gene of the clpC operon is required for stress tolerance and virulence in Staphylococcus aureus', *Microbiology (United Kingdom)*, 158(10).
- Wu, D. Q., Li, Y. and Xu, Y. (2012) 'Comparative analysis of temperature-dependent transcriptome of Pseudomonas aeruginosa strains from rhizosphere and human habitats', *Applied Microbiology and Biotechnology*, 96(4).
- Wu, Y. T., Zhu, L. S., Tam, K. P. C., Evans, D. J. and Fleiszig, S. M. J. (2015) 'Pseudomonas aeruginosa Survival at Posterior Contact Lens Surfaces after Daily Wear', *Optometry and Vision Science*, 92(6).
- Xie, L., Liu, W., Li, Q., Chen, S., Xu, M., Huang, Q., Zeng, J., Zhou, M. and Xie, J. (2015) 'First succinyl-proteome profiling of extensively drug-resistant Mycobacterium tuberculosis revealed involvement of succinylation in cellular physiology', *Journal of Proteome Research*, 14(1).
- Xu, J. Y., Xu, Z., Liu, X. X., Tan, M. and Ye, B. C. (2018) 'Protein Acetylation and Butyrylation Regulate the Phenotype and Metabolic Shifts of the Endospore-forming Clostridium acetobutylicum', *Molecular and Cellular Proteomics*, 17(6).
- Xu, Y., Wang, X. B., Ding, J., Wu, L. Y. and Deng, N. Y. (2010) 'Lysine acetylation sites prediction using an ensemble of support vector machine classifiers', *Journal of Theoretical Biology*, 264(1).
- Yang, D. C. H., Abeykoon, A. H., Choi, B. E., Ching, W. M. and Chock, P. B. (2017) 'Outer

Membrane Protein OmpB Methylation May Mediate Bacterial Virulence', *Trends in Biochemical Sciences*, 42(12).

Yang, H., Chen, J., Yang, G., Zhang, X. H. and Li, Y. (2007) 'Mutational analysis of the zinc metalloprotease EmpA of *Vibrio anguillarum*', *FEMS Microbiology Letters*, 267(1).

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J. and Zhang, Y. (2014) 'The I-TASSER suite: Protein structure and function prediction', *Nature Methods*, 12(1).

Yang, M. K., Yang, Y. H., Chen, Z., Zhang, J., Lin, Y., Wang, Y., Xiong, Q., Li, T., Ge, F., Bryant, D. A. and Zhao, J. D. (2014) 'Proteogenomic analysis and global discovery of posttranslational modifications in prokaryotes', *Proceedings of the National Academy of Sciences of the United States of America*, 111(52).

Yang, M., Wang, Y., Chen, Y., Cheng, Z., Gu, J., Deng, J., Bi, L., Chen, C., Mo, R., Wang, X. and Ge, F. (2015) 'Succinylome analysis reveals the involvement of lysine succinylation in metabolism in pathogenic *Mycobacterium tuberculosis*', *Molecular and Cellular Proteomics*, 14(4).

Yang, Y., Tong, M., Bai, X., Liu, X., Cai, X., Luo, X., Zhang, P., Cai, W., Vallée, I., Zhou, Y. and Liu, M. (2018) 'Comprehensive proteomic analysis of lysine acetylation in the Foodborne pathogen *Trichinella spiralis*', *Frontiers in Microbiology*, 8(JAN).

Yeung, A. T. Y., Torfs, E. C. W., Jamshidi, F., Bains, M., Wiegand, I., Hancock, R. E. W. and Overhage, J. (2009) 'Swarming of *Pseudomonas aeruginosa* is controlled by a broad spectrum of transcriptional regulators, including MetR', *Journal of Bacteriology*, 191(18).

Yu, H., He, X., Xie, W., Xiong, J., Sheng, H., Guo, S., Huang, C., Zhang, D. and Zhang, K. (2014) 'Elastase LasB of *pseudomonas aeruginosa* promotes biofilm formation partly through rhamnolipid-mediated regulation', *Canadian Journal of Microbiology*, 60(4).

Yu, S., Ren, J., Ni, J., Tao, J., Lu, J. and Yao, Y. F. (2016) 'Protein acetylation is involved in salmonella enterica serovar typhimurium virulence', *Journal of Infectious Diseases*, 213(11).

Zhang, J., Sprung, R., Pei, J., Tan, X., Kim, S., Zhu, H., Liu, C. F., Grishin, N. V. and Zhao, Y. (2009) 'Lysine acetylation is a highly abundant and evolutionarily conserved modification in *Escherichia coli*', *Molecular and Cellular Proteomics*, 8(2).

Zhang, M., Xu, J. Y., Hu, H., Ye, B. C. and Tan, M. (2018) 'Systematic Proteomic Analysis of Protein Methylation in Prokaryotes and Eukaryotes Revealed Distinct Substrate Specificity', *Proteomics*, 18(1).

Zhang, Y., Faucher, F., Zhang, W., Wang, S., Neville, N., Poole, K., Zheng, J. and Jia, Z. (2018) 'Structure-guided disruption of the pseudopilus tip complex inhibits the Type II secretion in *Pseudomonas aeruginosa*', *PLoS Pathogens*, 14(10).

Zhang, Z., Tan, M., Xie, Z., Dai, L., Chen, Y. and Zhao, Y. (2011) 'Identification of lysine succinylation as a new post-translational modification', *Nature Chemical Biology*, 7(1).

Zhao, K., Li, W., Li, Jing, Ma, T., Wang, K., Yuan, Y., Li, J. S., Xie, R., Huang, T., Zhang, Y., Zhou, Y., Huang, N., Wu, W., Wang, Z., Zhang, J., Yue, B., Zhou, Z., Li, Jiong, Wei, Y. Q., *et al.* (2019) 'TesG is a type I secretion effector of *Pseudomonas aeruginosa* that suppresses the host immune response during chronic infection', *Nature Microbiology*, 4(3).

Zhao, Y. and Jensen, O. N. (2009) 'Modification-specific proteomics: strategies for characterization of post-translational modifications using enrichment techniques', *Proteomics*, 9(20).



- Zhou, P., Li, W., Wong, D., Xie, J. and Av-Gay, Y. (2015) 'Phosphorylation control of protein tyrosine phosphatase A activity in *Mycobacterium tuberculosis*', *FEBS Letters*, 589(3).
- Zhu, J., Cai, X., Harris, T. L., Gooyit, M., Wood, M., Lardy, M. and Janda, K. D. (2015) 'Disarming *Pseudomonas aeruginosa* virulence factor lasB by leveraging a *Caenorhabditis elegans* infection model', *Chemistry and Biology*, 22(4).
- Zhu, K., Zhao, J., Lubman, D. M., Miller, F. R. and Barder, T. J. (2005) 'Protein pI shifts due to posttranslational modifications in the separation and characterization of proteins', *Analytical Chemistry*, 77(9).
- Zhu, M., Zhao, J., Kang, H., Kong, W. and Liang, H. (2016) 'Modulation of type III secretion system in *Pseudomonas aeruginosa*: Involvement of the PA4857 gene product', *Frontiers in Microbiology*, 7(JAN).
- Zriq, R., Sana, T. G., Vergin, S., Garvis, S., Volfson, I., Bleves, S., Voulhoux, R. and Hegemann, J. H. (2015) 'Genome-wide screen of *Pseudomonas aeruginosa* in *Saccharomyces cerevisiae* identifies new virulence factors', *Frontiers in Cellular and Infection Microbiology*, 5(NOV).
- Zulianello, L., Canard, C., Köhler, T., Caille, D., Lacroix, J.-S. and Meda, P. (2006) 'Rhamnolipids Are Virulence Factors That Promote Early Infiltration of Primary Human Airway Epithelia by *Pseudomonas aeruginosa*', *Infection and Immunity*, 74(6).

## 9 Appendices

### Appendix A

BLAST analysis of *P. aeruginosa* secreted proteins and the full sequence alignments (generated by BoxShade). Black highlight indicates extensive homology, grey highlight indicates some degree of homology.

#### ImpA

BLAST results:

Entry	Entry name	Protein names	Organism	Gene name
<input type="checkbox"/> Q9I5W4	IMPA_PSEAE	<b>Immunomodulating metalloprotease</b>	<i>Pseudomonas aeruginosa</i> (strain ATCC 15692 / DSM 22644 / CIP 104116 / JCM 14847 / LMG 12228 / 1C / PRS 101 / PAO1)	<b>impA</b> PA0572
<input type="checkbox"/> A0A1K2HRD4	A0A1K2HRD4_9BURK	Peptidase M60, enhancin and enhanci...	<i>Chitinimonas taiwanensis</i> DSM 18899	SAMN02745887_03221
<input type="checkbox"/> A0A515ES80	A0A515ES80_9BURK	<b>Peptidase M60 domain-containing pro...</b>	<i>Rhodoferrax sediminis</i>	EXZ61_15890
<input type="checkbox"/> A0A1Y0NBC3	A0A1Y0NBC3_9BURK	<b>Peptidase M60 domain-containing pro...</b>	<i>Curvibacter</i> sp. AEP1-3	AEP_02505
<input type="checkbox"/> A0A6M4A5I5	A0A6M4A5I5_9BURK	<b>Peptidase M60 domain-containing pro...</b>	<i>Undibacterium piscinae</i>	EJG51_012670
<input type="checkbox"/> A0A0C3I3H0	A0A0C3I3H0_9VIBR	<b>Peptidase M60 domain-containing pro...</b>	<i>Vibrio mytili</i>	SU60_17945
<input type="checkbox"/> K2IS43	K2IS43_9GAMM	<b>Peptidase M60 domain-containing pro...</b>	<i>Gallaecimonas xiamenensis</i> 3-C-1	B3C1_10682
<input type="checkbox"/> A0A1S2TSK9	A0A1S2TSK9_9GAMM	<b>Peptidase M60 domain-containing pro...</b>	<i>Shewanella algae</i>	BFS86_02360

Full alignment of PAO1 ImpA with BLAST results:

```

Q9I5W4      1  MSLSTTAFPSLQGENMSRSPIPRHRALLAGFCIA-----CALS
A0A1K2HRD4  1  MH-----ALYRP-----RLLVSTLLL-ALLAACGGGGGEGDGSKPDNGGCTVT
A0A515ES80  1  MR-----LHLAGKTL SARLFV GWATFALV-ALLTACGGGKSGSSPSSVSGSCAVT
A0A1Y0NBC3  1  MR-----PHTANTTTINRHRVGVSTLALA-ALLTACGGGSGDSSSSGASSCTVS
A0A6M4A5I5  1  MR-----PHPKL-----LRFFSIALL-SALLAACGGGGGSTPA--TSGTGTIT
A0A0C3I3H0  1  -----
K2IS43      1  -----MKKTLLLGSLLLAGCGGGGGDPASP-----
A0A1S2TSK9  1  -----MARKKP-----LLLHSILLGSLLLGACGGGGDDESGSG-SSGD-NLQ

```

```

Q9I5W4      39  A-Q-----
A0A1K2HRD4  42  PPV-----
A0A515ES80  50  LVA-----
A0A1Y0NBC3  50  LVA-----
A0A6M4A5I5  40  PPV-----
A0A0C3I3H0  1  -----
K2IS43      26  ---STNQAPSTSPMTLAYQWQTQEASGNWRQFAQDADGDNLNASISQQGQLGRFSLEGDQ
A0A1S2TSK9  40  PPAGVNQAPKGDDIQLAYDSLQSI TVNWQGYVSDPDGDP LQASIAEQGKLGQFSLDGDM

```

```

Q9I5W4      41  -----AATQEEILDAALVSGDSSQLTDSHTV
A0A1K2HRD4  45  -----SN-----AIDQALQSGDASALSDANAI
A0A515ES80  53  -----APQPNPQIVQALATGDASALTDKLV
A0A1Y0NBC3  53  -----APQASSQISQALQSGDASALTDKLV
A0A6M4A5I5  43  -----VTPDSSLDLALRTGDASALTDAAPI
A0A0C3I3H0  1  -----
K2IS43      83  LLYLANSGAQGGDQGGQLTVSDGRGGQATLGITVTQVDGRDGVAVALLAQGDASLTSSPAILL
A0A1S2TSK9  100 LSYKADKDAKGS DQGLLQVSDGRGGSVSLKLA VFGVDGQSPLERALLASGDASGLNPD TILL

```

Q9I5W4 67 ALRLQQOVERIROTRTQLDGLYQNLISQAYDFGAASMWVLPAFNPNTLFFLIGDKGRVLA  
 A0A1K2HRD4 67 AIRARDFTQGLVNRQAGLSAALYQGVSEYOPTOHSQFVLPAFNPNTLFFLIGDKGRVLA  
 A0A515ES80 79 AQHAQYAAQQQISLAQAARTSSLYQGVSTEYDPTNOSHVVOPPLNTATAQPLIVGDQGNAL  
 A0A1Y0NBC3 79 SQHAQYAAQQQISLAQAARTSSLYQGVSTEYDPTNOSHVVOPPLNTATAQPLIVGDQGNAL  
 A0A6M4A5I5 69 SIAARDFVAQLSDQQAARISTLVAGVSSEYAPTONSQFLPLNTELSAPWIIIGDKGOTLA  
 A0A0C3I3H0 1  
 K2IS43 143 ARIKQ-EIQDVQASQQQLVTLQFGNGAIDYAPGNRTQLFNVTEPDRTFSLINANGGOVLA  
 A0A1S2TSK9 160 EAIIV-QITQLRSNEQALRQRFADTALAYAPGNRTQLFNIIEPEIATPLLRANTGOVLA  
  
 Q9I5W4 127 SLSLEAGGRGLAYGTNVLTQLSGT----NAAHAPLLKRAVQWLVNGDPGAATAKDFKVS  
 A0A1K2HRD4 127 SLSIAQGGRSAGYGINVLDQFRKA---NEAHMPAFQRLVSWLVVGPQATLPASINVA  
 A0A515ES80 139 SISVAAGGRSAGFGVOVLERFNGNQ---LIAYRPAFKRLLAWLVRGDAATLPATLNVA  
 A0A1Y0NBC3 139 SISVAGGRSAGFGVOVLSKFNANE---LVAYRPAFKRLLAWLVRGDAATLPATLNVA  
 A0A6M4A5I5 129 SISIASNGRSAGYGANVLEQFRNN---NLNHAPAFKRLLSWLVLGDATLALPASLPAV  
 A0A0C3I3H0 10 TIQGHDCVRAAAGFCNGLIQQDNDNSYSYDAYQKDFKRVLSWLVSQDANAPIPSTLNIA  
 K2IS43 202 IAGEVQGGRYGGIGTHLFARFQAGE---LLAMAPAAEQLLAWLMQD---ADLSQPHTLAL  
 A0A1S2TSK9 219 VAGEQHTGRFAAFGTHLFARFHAGE---LTAMEPANDNLLAWLLNRAQAABEQQPIITVSL  
  
 Q9I5W4 183 VGVVDK-----TAALNGLKSAGLOPADAACNAL-----TDASCSTSKLLVLCNGA--  
 A0A1K2HRD4 184 AGLNA-----TNGLNGLTAAKITAKSLACDPI-----AENSCGQQAADLIIVGGGV--  
 A0A515ES80 196 AGVNA-----NNSAAGMVKADVPVTIACDFI-----ATPACATNAHLVVGDL--  
 A0A1Y0NBC3 196 AGINA-----SSAAGMVKAGVPVTIACDFI-----ATSACATNAHLVVGDL--  
 A0A6M4A5I5 186 AGINA-----SNGLAGLTKIGVTATSLGCDLF-----ADINCNOGAKLLVVGDM--  
 A0A0C3I3H0 70 DLSSGWVCQPAFKDCVEAFPEHIGIDVTQSQCQSVLGGTDEENLACAQQSDVLISAVE--  
 K2IS43 256 SFLAGQE-----EATRWLEAQHPNWTLLGCDVA-----ALDNCLAQADLVISQWRGND  
 A0A1S2TSK9 276 SFLGGQE-----SASRSWLQGRFPNWTIKSCNQVA-----TLEACIADQQLISQWRAAD  
  
 Q9I5W4 228 S-AASLSATVRARLQAGLPILFVHTNGWNQSSSTGQQILAGLGLOEGPYG-GNYWDKDRVP  
 A0A1K2HRD4 229 NADSGLEAKVRNYLSAGKPVLYLHTNGWGDSSASGQOMLSGMGLALGPYA-GNYFAADKVA  
 A0A515ES80 241 APSASLEAGIRALVAAGRPVLYVHTKFSAFSDSGRQVLGAGMGLQFGSYG-GNYWANDKVA  
 A0A1Y0NBC3 241 AASASLETSIRALVAAGRPVLYVHTKSNWTSDSAIOVLGAGMGLQFGPYG-GNYWASDKVA  
 A0A6M4A5I5 231 LANASAEASVRAWVSAGKPVLYLHTKTWGDSESGRQILSGMGLQMGGYA-GNYFSNDLVA  
 A0A0C3I3H0 128 KEVGDLOSQVSALLQSGVPILYLHNHGEFLPEEGQAQILAAAGMEYAVAGASNYFKODQYT  
 K2IS43 306 QDAIAIVASFQGALAQGGKGLFYQHNWYEATSPVADAITELMGASL-PYG-GNYWANASAN  
 A0A1S2TSK9 326 TDAGQVAGVYSRALASGKALYQHNWYEATNAVADAITAGTMGVSL-PYG-GNFWANAAAD  
  
 Q9I5W4 286 SSRTTRTSVE-IGGAYGQDPALVQOIVDGSWRDQYDWSKCTSYVGRITCDDVPGLS-DFS  
 A0A1K2HRD4 288 ANRSSTSNQA-SLNOFGAVTPLRLTLASSNFRSDYDWSKCTNSVGKVSQNVSDLOAEVI  
 A0A515ES80 300 AGRSEATNRT-IGDQFAKTAMLSLSDSFSMPYDWSPTVSAGKTDCSGVASLOSNNL  
 A0A1Y0NBC3 300 AGRSEAGNRT-LSDFAKTTAMLNLLSDSFSMPYDWSPTVSAGKTDCSGVASLOANLL  
 A0A6M4A5I5 290 AGRSEADNLA-KLDQFADSRALINRIANNWRTDYGWSACTTYVGSINCDNVPDLQKDL  
 A0A0C3I3H0 188 GGRSATENIERISNPFHEYLPFFTAINDGSLNTNYDWSQVNSACTFKCPDVPNLLNASLV  
 K2IS43 364 WNQA--QAMLDQAPWLAEQRLTQHFIDQD--FAPDWSGCTSYVGKVSQDQVAGFDQAFI  
 A0A1S2TSK9 384 WSSA--TAMAAFFPLLGSEQRITQHFIDDD--FNFQWSGCTSYVGKVSQDKVNGFESEFL  
  
 Q9I5W4 344 KRVDVLKCALDAYNOKAONLFALEGTSLRLWLLWADAVRONIRYPMDKAADT--ARFOE  
 A0A1K2HRD4 347 TPAEQVREQINSFNAGRNLFSTENTLLRLTLWADTTRKDTYPTDKTKOP--AAFOR  
 A0A515ES80 359 SPVDALRSQIDVFNKAGONLFATADTVLRYLVWADVVRQLRYPMDKASTP--AAFEK  
 A0A1Y0NBC3 359 SPVDALRSQIDVFNKAGONLFATADTVLRYLVWADVVRQLRYPMDKASTP--AAFOK  
 A0A6M4A5I5 349 TPIESLRQQIDNFTSRGRNLFATENTSLRLWLLWADVQRKQIVYVVDKTSKP--VEFOK  
 A0A0C3I3H0 248 EPVEALRDVFDSTYKSGTPLFSAPGNEFYRLANIWADLNRRSMSYPMEKEEDMADYRFOK  
 K2IS43 420 SGARSLKNALNRLDSQGLDLFAEENLTLLKLFVLGLDSYRAQIHYPMDEKSTD-PNOFLA  
 A0A1S2TSK9 440 AGARALKNSLNQLDSRQOMLFQNEGRRLKLFVLGLGLYRADIAFPMDKDTTF-QGRFLA  
  
 Q9I5W4 402 TFVADAIIVGYVREAGAAQKELGSGYAGQRQOSMPVSGSEETLTTLPLSAOGFTAIGRMAAP  
 A0A1K2HRD4 405 ALIADAWVAYVRAGTAQCDLGYMGKTVQGITPSQDETETVTLPLSESGFTAIGRFAAP  
 A0A515ES80 417 ALIADALVAYVRPAANAQADLGSFAGALTGMAVSTDETVDVSVPTSGFTAIGRFAAP  
 A0A1Y0NBC3 417 ALIADALVAYVRPAANAQADLGSFASAMTAGMAVSTDETVDVSVPTSGFTAIGRFAAP  
 A0A6M4A5I5 407 ALVADALVAYVRPLGVAQTDLSSFLSKSAATLPVSTSDETISISLPAQOGFTAIGRFAAP  
 A0A0C3I3H0 308 AYVSDAMVTYLRRTAGEAQPNNMGTFALTDQGLPVSPQOETVTVTLPAESGFTAIGRYIKP  
 K2IS43 479 AYLADHLAYYHRASNPAQADLGNFSDPIATAPAL--VSQTLTFTLNHGSNYRGTCGLYLLP  
 A0A1S2TSK9 499 AYLADHLALYLGRNNAQADLGNFSDPLPQTTLT--ENVSLMAAGKESGYRGSFYLLP  
  
 Q9I5W4 462 GKRLSIRIEDAGQASLAVGLNTORIGSTRLWNTR-----QYDRPRFLKSPDIKLOANOS  
 A0A1K2HRD4 465 QOTLOVEVLNAGTATLALRLNTORTGSTRLWAND-----CFNRPRFLASPSISLSQGSF  
 A0A515ES80 477 GKPVTVELLASGATSLSRLNTORTGSTRLWDPN-----RYNRPRFLASPMVLSTGQA  
 A0A1Y0NBC3 477 GKPVTVELLASGAMVSLRLNTORTGSTRLWDPN-----RYNRPRFLASPMVLSTGQA  
 A0A6M4A5I5 467 KKTLOIELIDAGSATVALRLNTORVSTRYLGAAT-----SYTRPRFLASPMRLAKASF  
 A0A0C3I3H0 368 GRTATIELVDAGNASVSFYINTPNGGTRVWNPKEKSGGIVGYARPRFLKSPNMPNPNSEP  
 K2IS43 537 GORIELERTDQPLALAAFIINTORSGSTREFNAN-----GYLRPRFLKSPQFSLTGGOO  
 A0A1S2TSK9 557 GOSVRLERKDTLPLTVKAFINTORTGSTRFNNO-----GYORPRFLKSPVELTLKPGOP

Q9I5W4 516 VALVSPYGGLLQLVYSGATPGQTVTVKVTGAASQFFLDIOPGEDSSQAIADFIQALDADK  
 A0A1K2HRD4 519 LQLSSPYGGNQLSFSGATLQGVQLRLRGVARQPFLLDLSQG-GD---KASFVAQINAGQ  
 A0A515ES80 531 MOLVSPYGGTLQLVFSNATPQONVOLRLRGVAKHPFLDOSNGAGD---TAAFATTLNAAQ  
 A0A1Y0NBC3 531 MOLVSPYGGTLQLVFSNATPQONVOLRLRGVAKHPFLDOSNGAGD---KAAFVTALNAAQ  
 A0A6M4A5I5 521 LQVNSPYGGTMQLFSGATAGQIVTLIRIGVAQHYPFDLSAG-SD---KAAFI AALNNGQ  
 A0A0C3I3H0 428 VEVNSPYGGTLEVRFNAGAVAGQTLTFNVTNVTGTPVLNLREN-GN---VSSFVDELMTQTP  
 K2IS43 591 QQLTSPYGGPLMVQLPAGS--GQVSLKVSNLSPYPYLSDFAN-----AADYLAQLDDSP  
 A0A1S2TSK9 611 LTLSSPYGGTLMQLPAGE--GVVSVEAQNLLAYPYLKDFNQ-----ASGYLTALLETSP

Q9I5W4 576 ADWLEIRSGSVEVHAKVEKVRGSID-KDYGGDVQRFIRELNEVFIDDAYTLAGFAIPNQA  
 A0A1K2HRD4 575 FDWAEIKMPGVEVHSRVDKIKTVLK-DDYGNDMDRYLAEMRTLFFEDAYQLAGFALSQKT  
 A0A515ES80 588 HEWAEIKLAGVEVHSRADKMRVINGPHYAGDIDKFLNELKTIFFEDLYMLAGYALPGKS  
 A0A1Y0NBC3 588 HEWAEIKLAGIEIHSRADKMRVINGTDYAGDIDKFLNEVKTLFFEDLYMLAGYALPGKS  
 A0A6M4A5I5 577 FDWAEIKMPGAEIHTKVAKMKDVLO-KDYGNDMVRYLDELRLDILLIEGDYRMAGYALKCLD  
 A0A0C3I3H0 484 FEWAEIKLDGVEIHTRVDKMLSUVNNDIYAGDLTRYLNELDYFYGAYHMAAGFAG-TKP  
 K2IS43 643 LAWAGLRDTDFVEINSRKQMMKAFIYADRYQCDVTOALDDVWAYMIKGTYDLAGFQGDGLA  
 A0A1S2TSK9 663 LSWAGLRDTDFVEVNSRKHMMKQFIYAEPYRGDVEQALNDVWRYMIKGTYDLAGFSGECLA

Q9I5W4 635 KTPAIQOECAARGWDCDSETLHKLPGTQHINVDQYACGGGCSGNPYDQTWGLNPRGWGE  
 A0A1K2HRD4 634 LPAAVQSLCNRGWDCSSETVHRVPGTQHINVDAYAECCGAGCSGNPYDQSWGLNPRGWGE  
 A0A515ES80 648 LTAHVQAMCANLGWCTDATHLRVPGTQHINVDNYSQCGSGCAGNPYDQDWGLTPRGWGE  
 A0A1Y0NBC3 648 LTAHVQAMCTSLGNCTDATHLRVPGTQHINVDNYSQCGSGCAGNPYDQDWGLTPRGWGE  
 A0A6M4A5I5 636 LPAAVQTFCNGLAWNCTDAVMHRCGPSTQHINSDDFACGNGCSGNPYDQDWGNPRGWGE  
 A0A0C3I3H0 543 LSAAVLAQCNOFGWDCENAEHLQPNVOHITAGNTAACGAGCSGNPYDQMOGVNPRGFILE  
 K2IS43 703 LAGTVAARCTSLGWDCTDPQIHAKPKVQHINVDEAAECGAGCSGNPYDQSWALSPFGWGE  
 A0A1S2TSK9 723 LAPSVAAARCALLGWDCONEQIHAKPKVQHINVDEAAHCGGCSGNPYDQAVLSPFGWGE

Q9I5W4 695 SHELGHNLQVNRKLVYGGRSGEISNOIFPLHKDWRVLRFCQNLDD-TRVNYRNAYNLIV  
 A0A1K2HRD4 694 SHELGHNLQCGMLNVYGGRSGEVSNNLFPLHKNWRVLRLEDDDRDR-GRINYRSAFDMIK  
 A0A515ES80 708 SHEVGHNOQKGMHKVYDDRSGEVSNNLFPLHKGWRMLSELGYNSGD-NRVAYQSAFNMIM  
 A0A1Y0NBC3 708 SHEVGHNOQKGMHKVYDDRSGEVSNNLFPLHKGWRMLSEMGYNTGD-TRVSYLSAFNMIM  
 A0A6M4A5I5 696 SHELGHNLQOGLLNIYGGRSGEVSNNLFPLRMNWRVLRHEKNT-YVA-NHVGYRSAFDMIN  
 A0A0C3I3H0 603 SHELGHNLQYQEFRVDGGISGEVSNNIFPLQKNGRLYQDFGVDLGS-NKVDYEATFDMLV  
 K2IS43 763 SHEIGHNLQSRSLKIYGGRSGEVSNNIFPLYKGWQRFHDSGERIESCDRQDPALTYGWLO  
 A0A1S2TSK9 783 SHEIGHNLQNRRLKIYGGRSGEVSNNIFPLYKGWQRFKDSGERIASCDRQSPQTTFEWLK

Q9I5W4 754 AGRAEADPLAGVYKRLWEDPGTYALNGERMAFYTQWVHYWADLKN--DPLQGWDIWTLTY  
 A0A1K2HRD4 753 AARAEADPIEGAYQRIWGS DAYAVONGERMAFYMQWVHYWAERNA--DIAKGWDIITLTY  
 A0A515ES80 767 AAKTEADPIEGVYQRIWGS DAYAVONGERMAFYMQWVHYWAQROA--SIATGWDIVTLTY  
 A0A1Y0NBC3 767 AAKLOADPVEAAVQSIWGS DAYAVONGERMAFYMQWVHYWAQROV--SIATGWDIITLTY  
 A0A6M4A5I5 754 AGRTNADPLEARYQSIWGS DAYAVONGERMAFYMQWVHYWHARVN--DEARGWDIVTLTY  
 A0A0C3I3H0 662 AGLKSGNSSDKLNQDLWTDASYSAGQGNKRLAFYTQWALYWADKQNSSDAKTAWEIYTLTY  
 K2IS43 823 QAQTOADPSQAMYDKLWSQGTGYDNAGTRLDLYLQLAFMAD-DIA--GLDNGWQIYTLTY  
 A0A1S2TSK9 843 QAQNSSDPVTSVYEKLWSQGTGYDNAGPRLDLYLQLAALAE-EHA--GLDNGWQIYTLTY

Q9I5W4 812 LHQRQVD-----KSDWDANKAALGYGYAQRPGNSGDASSTDGNDNLLGLSLWL  
 A0A1K2HRD4 811 LHQRQFA-----KLDWAANKDKLGYSQYASRPS-----VDGNDNLLITLSFL  
 A0A515ES80 825 LHQRQFQDAV-----AAADWAANRNKLGYSTYANKPS-----PSGNDNMLITLSWI  
 A0A1Y0NBC3 825 LHQRQFQDAV-----ADADWAANRNKLGYSTYATKPS-----PTGNDNLLITLSWI  
 A0A6M4A5I5 812 LHQRQFA-----RADWAANKDKLGYSQYATRPS-----VDGNDNMLIALSWI  
 A0A0C3I3H0 722 LHAREFLAADVDNSTQETEDDWDNVKGLGFSTYATRPDQDGSRNSQYGNVYLLITLSKL  
 K2IS43 880 LQERLFTQAAA-----DPDRWAAQKGLGMSAFATAPD-----PDGNDFMLMGLSWL  
 A0A1S2TSK9 900 LHERLFSHAIA-----DVARWGMERDALGMSVYIYAPN-----LDGNDFMLLSLSKL

Q9I5W4 861 TQRDQRPFTALWGIRTSAAAQAQVAAYGFAPAEFFYANNRTNE-YSTVKLLDMSQGSFA  
 A0A1K2HRD4 853 TQRDQRPFTDLWGVRYSAATAAAQVTSYGYPAEAAALFYANSSSND-HATVKKVDMLAANPV  
 A0A515ES80 870 TQRDQRPFTDLWGVRYSPAAAAQVAAFGFAAEPAFYANTSTNN-HSTVRKVDMSVANPV  
 A0A1Y0NBC3 870 TQRDQRPFTDLWGVRYSAATAAAQVAAFTFAVEPALFYANTSTNN-HSTVRKVDMSVASPA  
 A0A6M4A5I5 854 TQRDQRPFTDLWGVRYSTAAAQVASYAFAVEPALFYAYKVTDPKVSAYKKIDMSVATPV  
 A0A0C3I3H0 782 TGHQDRNVFDMWGIKYSSAATQQLDTFGLPVEPLFFYTTTPMND-YAAGHRIDMSSDPST  
 K2IS43 927 LORDMRGYFDLWGIYGSADAGAQAQVAAYGFAPAAKVYYQVQSQCA-DMQVFKLPV-DGTSS  
 A0A1S2TSK9 947 LQLDMRPWFDLWGVRYQAQAANQVAAYGYPAAEAVFYRLSSQCA-ELAVPKLTL-DAN--

Q9I5W4 920 WFFP--  
 A0A1K2HRD4 912 WFFV--  
 A0A515ES80 929 WFF--  
 A0A1Y0NBC3 929 WFF--  
 A0A6M4A5I5 914 WFF--  
 A0A0C3I3H0 841 LEYATE  
 K2IS43 985 WBAT--  
 A0A1S2TSK9 003 W----

## LasB

### BLAST results:

Entry	Entry name	Protein names	Organism	Gene name
P14756	ELAS_PSEAE	Elastase	<i>Pseudomonas aeruginosa</i> (strain ATCC 15692 / DSM 22644 / CIP 104116 / JCM 14847 / LMG 12228 / 1C / PRS 101 / PAO1)	<b>lasB</b> PA3724
A0A1G9GYK7	A0A1G9GYK7_9PSED	Neutral metalloproteinase	<i>Pseudomonas indica</i>	SAMN05216186_11447
A0A381F2B8	A0A381F2B8_CHRVL	Neutral metalloproteinase	<i>Chromobacterium violaceum</i>	<b>lasB</b> NCTC8684_03750
A0A254VT81	A0A254VT81_9BURK	Neutral metalloproteinase	<i>Xenophilus</i> sp. AP218F	CEK28_01790
U2ZHC2	U2ZHC2_PSEA4	Neutral metalloproteinase	<i>Pseudomonas alcaligenes</i> (strain ATCC 14909 / DSM 50342 / JCM 20561 / NBRC 14159 / NCIMB 9945 / NCTC 10367 / 1577)	PA6_002_00880
A0A257FTD7	A0A257FTD7_9BURK	Neutral metalloproteinase	<i>Burkholderiales bacterium</i> PB82	CFE41_00260
A0A1K2H955	A0A1K2H955_9BURK	Neutral metalloproteinase	<i>Chitinomonas taiwanensis</i> DSM 18899	SAMN02745887_00792

### Full alignment of PAO1 LasB with BLAST results:

```

P14756      1  MKKVSTL---DLLFVAIMGVSPAFAADLIDVSKLPSKAA---QGAPGPVTLQAAVGAGG
A0A1G9GYK7  1  MKHA---LRIPALFLSALSLPAVSFAADLVDVSKVEARRTVA-----GEAGLHADL-CAS
A0A381F2B8  1  MRK-EQLMLRGLVLSALAVFSSATMAAERIDTEKQGKAQ-----ANSASFT-GVS
A0A254VT81  1  MKR-NSMLQHGLALSVLALACASAOAAQRINTEQQPQLKGI-----DSSASAFSQT-GVS
U2ZHC2      1  MPRK---LIPQSLLLASLAAPATTFAAELVDVASLPQPRAGV-----AAGDIHSTL-GLA
A0A257FTD7  1  MQATFKLLARSALLAAASLACQMAADRVELEGADLA--RVGLVAARGVTGPPAAL-GLA
A0A1K2H955  1  MKPVLFVARPAVL SLLVLGAFGATAAERTINLQERSPDSLLSAALASGQTGVHAML-GLS

```

```

P14756      55  ADELKAIRSTTLPNGKOVTRYEQFHNGVRVVEAITEVKGP GK-----SVAAQRSG
A0A1G9GYK7  52  PDELKVLRSQSAVGKGRVVTRYQQFYNGVRIRGEAITEVSGAQDGERTKRSIDTPAFVRRG
A0A381F2B8  50  QGDLKALRSTQFASGKVTRYQQY YQGVVWGEGVVEEKFAAA--L-----KSVQGLSG
A0A254VT81  54  PRDLKPLRSSQFANGKVTRYQQY YQGVVWGEAVVEEKTGAPGVA-----AKTAGKLSG
U2ZHC2      52  PEELDADSSARLPNGTRVTRYQQFYKGI RVWGEAITEASG-----PRPQRS
A0A257FTD7  58  ADELKPLRSQTYPSGLVVTRYEQHYQGVPIWGEAVVEHLPKG-----SRTPSFSG
A0A1K2H955  60  QSELKAQRSQTYRNGKTVTRYQQY YQGVVWGEAVVAQQTAG-----AVQPALAG

```

```

P14756      106 HFVANIAADLPGSTTAAVSAEQVLAQAQSLKAO---GRKTENDKVELVIRLGENNIAQLV
A0A1G9GYK7  112 HFVTNIAPDLVADTRPTLSAEQALNQAQLKAO---GRRTENERTELVIDLDAONKARLV
A0A381F2B8  103 HYIAGIQADLAS-AKPTLSSAQVLSQAQALKAN---GNPTYNDKAEVLVRLNERNVAQLV
A0A254VT81  109 YYIAGIQSDLAS-VKPALSAQVLSQAQALKAN---SNPTYNEKAELVVQLGKNNVAQLV
U2ZHC2      99  RMLVGIEQDLTRDATPTLTREQALQAQGLVAS---ALPPRNEQSELVVRONEQQAQLA
A0A257FTD7  108 ALVRNINQDLGH-VQPSLSEQAVLAVAKSRAKDASKVGKTSNEQVKLFVKLNERNVAQLV
A0A1K2H955  110 SMLSGIASDLAS-AKPVFSAANMLAQAQSLAR---VNGKTENDQATLYVVKLDKNNVARLV

```

```

P14756      163 YNVSYLI-PGEGLSRPHFVIDAKTGEVLQWEGLAHA-EAGGPGGNQKIGKYTYGSDYGP
A0A1G9GYK7  169 YVVSYFL-PD-VPSRPHFLIDANDGSVLOQWEGLAHRTEASGPGGNEKTGRILYGSYDYG
A0A381F2B8  159 YLVSFVV-DGKEPSRPHFLIDANSQVQLKQWEGLNHA-EAGGPGGNAKTGQYTYGKDYG
A0A254VT81  165 YLVSFVV-EGKEPSRPHFLIDANSQVQLKQWEGLNHA-EAGGPGGNAKTGQYTYGKDYG
U2ZHC2      156 YVVSFFL-GGAKPSRPHFMIEANSGETLQWEGLAHA-EASGPGGNSKTGQYLYGTQYGP
A0A257FTD7  167 YLVSFLLTSHGDEPSRPHLLIDANSQAVLEQWEGLNHA-NATGPGGNAKTGKYEYGTTYGP
A0A1K2H955  166 YLVSFVIEDADHPSRPHFFIDANTGAVLERWEGLAHR-DATGPGGNTKTGQYEYGTDFGP

```

```

P14756      221 LIVNDRCEMDDGNVITVDMNSSDDSKTTPFRFACPTNTYKQVNGAYSPLNDAHFEGGVV
A0A1G9GYK7  227 LVVTNDCHMNSGDVLTVDLHNHSTSNAKRTPFOFACSYNDYKQINGAYSPLNDAHIFGNVV
A0A381F2B8  217 LIVTSDCKMDSGNVATVNLNGGT--SGTTPYKFACPTNTYKQINGAYSPLNDAHIFGNVV
A0A254VT81  223 LIVTSDCKMDSGNVTVNLNGGT--SGSAPYKFACPTNTYKQINGAYSPLNDAHIFGNVV
U2ZHC2      214 LIVSRDCOMNSGDVITVNLNORYDNSSVTPFRFACPYNDYKAINGAFAPLNDAHIFGNLV
A0A257FTD7  226 LVVTADCOMNSGNVITVDLKGGT--SGSTPFKFTCPRNEYKLTNGAYSPLNDAHIFGNVV
A0A1K2H955  225 LVVTDCKMNSGNVITVNLNGGT--SGSTPFQFTCPRNTYKQINGAYSPLNDAHIFGNVI

```



```

P14756      281  FKLYRDWFG-TSPLT-HKLYMKVHYGRSVENALWDGTAMLFGDGATMFYPLVSLDVA AHE
A0A1G9GYK7 287  FRLYNDWFG-LRPLS-QKLYMKVHYSRGYDNAFWDGESMTFGDGARTFFYPLVALDVVAHE
A0A381F2B8 275  FNLYKDWFN-LKPIT-QKLLMKVHYSRNYENAFWDGTAMTFGDGYNTFFYPLVSLDVSAHE
A0A254VT81 281  FGLYKDWFG-KSPLT-SKLYMKVHYSRNYENAFWDGTAMTFGDGYKTFYPLVSLDVSAHE
U2ZHC2      274  FQMYGSWFGGLRPLNERKLYMKVHYGNGYENAFWDGQAMTFGDGRSRFFYPLVSLDVSAHE
A0A257FTD7 284  FNMYQGWFN-IRPIS-OTLYMKVHYSSNYENAFWDGSAMTFGDGASTFFYPLVSLDVSAHE
A0A1K2H955 283  FNMYQSWFG-LRPIS-OTLQMRVHYSTNYENAFWNGSSMTFGDGCSFFYPLVSLDVSAHE

P14756      339  VSHGFTEQNSGLIYRGQSGGMNEAFSDMAGEAAEFYMRCKNDPLICGYDIKKGSGALRYMD
A0A1G9GYK7 345  VSHGFTEQNSNLQYYGQSGGMNEAFSDMAGETAEYIMKCKNDWLVGADIFKKTGALRYMD
A0A381F2B8 333  VSHGFTEQNSGLVYSGQSGGINEAFSDMAGEAAEYIMKCKNDPLVGAEIFKKTGALRYFA
A0A254VT81 339  VSHGFTEQNSGLVYSGQSGGINEAFSDMAGEAAEYIMKCKNDWLVGAEIFKKGSGALRYFA
U2ZHC2      334  VSHGFTEQNSGLVYEGQSGGINEAFSDMAGEAAEYIMKCKNDWLVGADIFKKGSGALRYMD
A0A257FTD7 342  VSHGFTEQNSGLVYSNQSGGMNEAFSDMAGEAAEYIMKCKNDWLVGSEIFKKGSGALRYFA
A0A1K2H955 341  VSHGFTEQNSGLVYSAQSGGMNEAFSDMAGEAAEYIMKCKNDWLVGAEIFKKGSGALRYMD

P14756      399  QPSRDGRSIDNASQYYNGIDVHHSSGVYNRAFYLLANSFGWDTRKAFEVFVDANRYYWTA
A0A1G9GYK7 405  DPTKDGRSIGHAADYRSGMNVHYSSGVYNKAFYLLANTAGWDTRKAFEVFVDANRFFYWTE
A0A381F2B8 393  DPTKDGRSIGNAKDYINGLDVHYSSGVYNKAFYLLATSPNWNTRKAFEVFVDANRLYWNA
A0A254VT81 399  DPTRDGRSIGHAKDYINGLDVHYSSGVYNKAFYLLATSA GWNTRKAFEVFVDANSLYWTA
U2ZHC2      394  QPSRDGASIEHAAQYHDGIDVHHSSGVYNRAFYLLSRTPGWNP RKAFEVFVDANRFFYWTE
A0A257FTD7 402  DPTRDGRSIDNASKYYNGLDVHLSGGVYNKAFYLLATKPGWNTRKAFEVFVDANRLYWTA
A0A1K2H955 401  DPTRDGRSIGHASNYTSCMDVHYSSGVYNKAFYLLANKPGWNTRKAFEVFVDANRLHWTA

P14756      459  TSNYNSGACGVIRSAQNRNYSAADVTRAFSTVGVTCPSAL-----
A0A1G9GYK7 465  TSSFDRGACGVIOSSONRGYGTSAVVOAFROVGVNCP SAL-----
A0A381F2B8 453  NATYNSAACGVVKAADARGYGGADVTKAFAAVGVTCQ-----
A0A254VT81 459  NATYNSAACGVVKAADNRGYN SADVTKAFAAVGVTCQP-----
U2ZHC2      454  TSSFTQGACGVIRSAQNRSYGSAEVVAFAAVGVSCPTLL-----
A0A257FTD7 462  SSTFNQGACGV EKAATARGYATADVTAAFSAVGVS CGTTPPPATIP LTKGVAKTGISLAK
A0A1K2H955 461  NSTFNQGACGVETA AENRGYSKADVTAAFSAVGVS CGTTPPPTGTVLTRGVVPVSGLGATK

```

## PA0622

### BLAST results:

Entry	Entry name	Protein names	Organism	Gene name
<input type="checkbox"/> G3XD39	G3XD39_PSEAE	Probable bacteriophage protein	<i>Pseudomonas aeruginosa</i> (strain ATCC 15692 / DSM 22644 / CIP 104116 / JCM 14847 / LMG 12228 / 1C / PRS 101 / PAO1)	PA0622
<input type="checkbox"/> A0A448BXA3	A0A448BXA3_PSEFL	Phage tail sheath protein	<i>Pseudomonas fluorescens</i>	NCTC10783_05923
<input type="checkbox"/> A0A6I7I634	A0A6I7I634_KLEPN	Phage tail sheath protein	<i>Klebsiella pneumoniae</i>	SAMEA4394740_02551
<input type="checkbox"/> A0A024HAN5	A0A024HAN5_PSEKB	Putative phage tail sheath protein	<i>Pseudomonas knackmussii</i> (strain DSM 6978 / LMG 23759 / B13)	PKB_0574
<input type="checkbox"/> A0A127MLR4	A0A127MLR4_9PSED	Phage tail sheath protein	<i>Pseudomonas citronellolis</i>	PcP3B5_07020

# Full alignment of PAO1 PA0622 with BLAST results:

G3XD39	1	MSFFHGVTVTNVDIGARTIALPASSVIGLCDVFTPGAQASAKPNVPVLLTSKKDAAAAG
AOA448BXA3	1	MSFFHGVTVTNVDIGARTIALPASSVIGLCDVFTPGAQASAKPNVPVLLTSKKDAAAAG
AOA6I7I634	1	MSFFHGVTVTNVDIGARTIALPASSVIGLCDVFTPGAQASAKPNVPVLLTSKKDAAAAG
AOA024HAN5	1	MSFFHGVTVTNVDIGARTIALPSSSIIGLVDTFTPGASVSAQADVPVLLTSLREAAAAG
AOA127MLR4	1	MSFFHGVTVTNVDIGARTIALPSSSIIGLVDTFTPGAPASAEEDVPVLLTSLREAAAAG

G3XD39	61	IGSSIIYLACEAIYNRAQAVIVAVGVEAETPEAQASAVIGGISAAGERTGLQALLDGKSR
AOA448BXA3	61	IGSSIIYLACEAIYNRAQAVIVAVGVEAETPEAQASAVIGGISAAGERTGLQALLDGKSR
AOA6I7I634	61	IGSSIIYLACEAIYNRAQAVIVAVGVEAETPEAQASAVIGGISAAGERTGLQALLDGKSR
AOA024HAN5	61	VDLSAIYKSCVAIYTQSAAVVAVGVKLAETPEMQASAVIGTVTASGORTGLQALLDGKSR
AOA127MLR4	61	TGSAIYKACTAIFTOASAVVAVGVAEVEDPAQQTSAIIGSVTESGORTGLQALLDGKSR

G3XD39	121	FNAQPRLLVAPGHSAAQAVATAMDGLAEKLRAIAILDGNSTDEAAVAYAKNFGSKRLFM
AOA448BXA3	121	FNAQPRLLVAPGHSAAQAVATAMDGLAEKLRAIAILDGNSTDEAAVAYAKNFGSKRLFM
AOA6I7I634	121	FNAQPRLLVAPGHSAAQAVATAMDGLAEKLRAIAILDGNSTDEAAVAYAKNFGSKRLFM
AOA024HAN5	121	FNTQPRLLVAPGHSSTQAVATAMDGLAEKLRAIAILDGNSTDEAAVAYAKNFGSKRLFM
AOA127MLR4	121	FNAQPRLLVAPKHSATEAVATAMDLAGKLRAIAILDGNSTDEAAVAYAKNFGSKRLFM

G3XD39	181	VDPGVQVWDSATNAARNAPASAYAAGLFAWTDAEYGFWSPPSNKEIKGVGTGTSRPVEFLD
AOA448BXA3	181	VDPGVQVWDSATNAARKAPASAYAAGLFAWTDAEYGFWSPPSNKEIKGITGTSRPVEFLD
AOA6I7I634	181	VDPGVQVWDSATNAARKAPASAYAAGLFAWTDAEYGFWSPPSNKEIKGITGTSRPVEFLD
AOA024HAN5	181	VDPGVQVWDSATNAARKAPASAYAAGLFAWTDAEYGFWSPPSNKEIKGITGTSRPVEFLD
AOA127MLR4	181	VDPGVQVWDSATNAARKAPASAYAAGLFAWTDAEYGFWSPPSNKEIKGITGTSRPVEFLD

G3XD39	241	GDETCRANLLNNANIATIIIRDDGYRLWGNRTLSSDSKWAFVTRVRTMDLVMDAILAGHKW
AOA448BXA3	241	GDETCRANLLNNANIATIIIRDDGYRLWGNRTLSSDSKWAFVTRVRTMDLVMDAILAGHKW
AOA6I7I634	241	GDETCRANLLNNANIATIIIRDDGYRLWGNRTLSSDSKWAFVTRVRTMDLVMDAILAGHKW
AOA024HAN5	241	GDETCRANLLNNANIATIIIRDDGYRLWGNRTLSSDSKWAFVTRVRTMDLVMDAILAGHKW
AOA127MLR4	241	GDETCRANLLNNANIATIIIRDDGYRLWGNRTLSSDSKWAFVTRVRTMDLVMDAILAGHKW

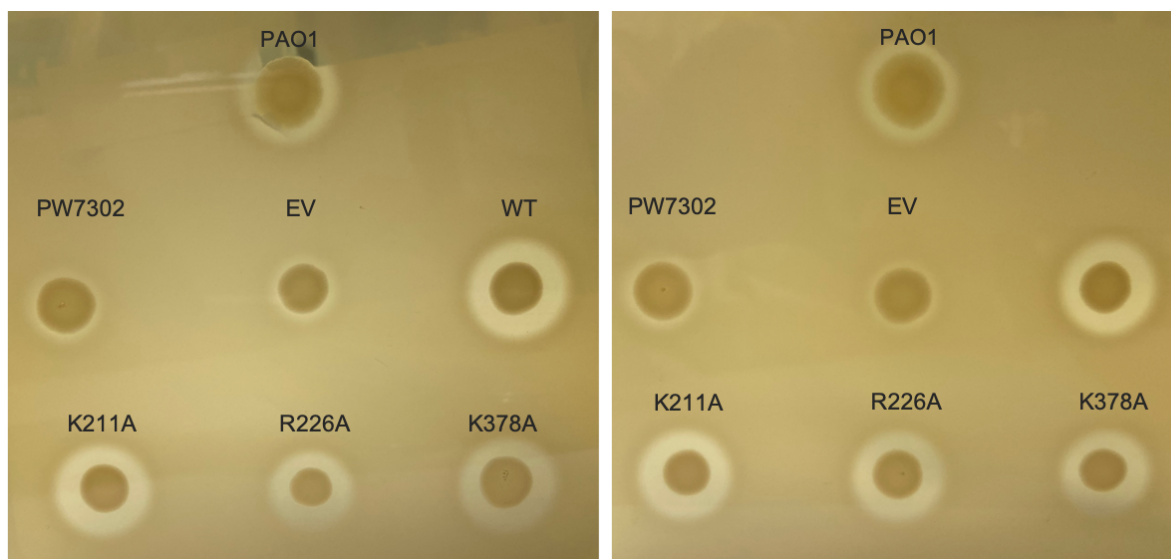
G3XD39	301	AVDRGITKTYVKDVTEGLRAFMRDLKNQGAVINFEVYADPDLNSASQLAQGKVYWNIRFT
AOA448BXA3	301	AVDRGITKTYVKDVTEGLRAFMRDLKNQGAVINFEVYADPDLNSASQLAQGKVYWNIRFT
AOA6I7I634	301	AVDRGITKTYVKDVTEGLRAFMRDLKNQGAVINFEVYADPDLNSASQLAQGKVYWNIRFT
AOA024HAN5	301	AVDRGITKTYVKDVTEGLRAFMRDLKNQGAVINFEVYADPDLNSASQLAQGKVYWNIRFT
AOA127MLR4	301	AVDRGITKTYVKDVTEGLRAFMRDLKNQGAVINFEVYADPDLNSASQLAQGKVYWNIRFT

G3XD39	361	DVPPAENPNFRVEVTDQWLTEVLDA
AOA448BXA3	361	DVPPAENPNFRVEVTDQWLTEVLDA
AOA6I7I634	361	DVPPAENPNFRVEVTDQWLTEVLDA
AOA024HAN5	361	DVPPAENPNFRVEVTDQWLTEVLDA
AOA127MLR4	361	DVPPAENPNFRVEVTDQWLTEVLDA

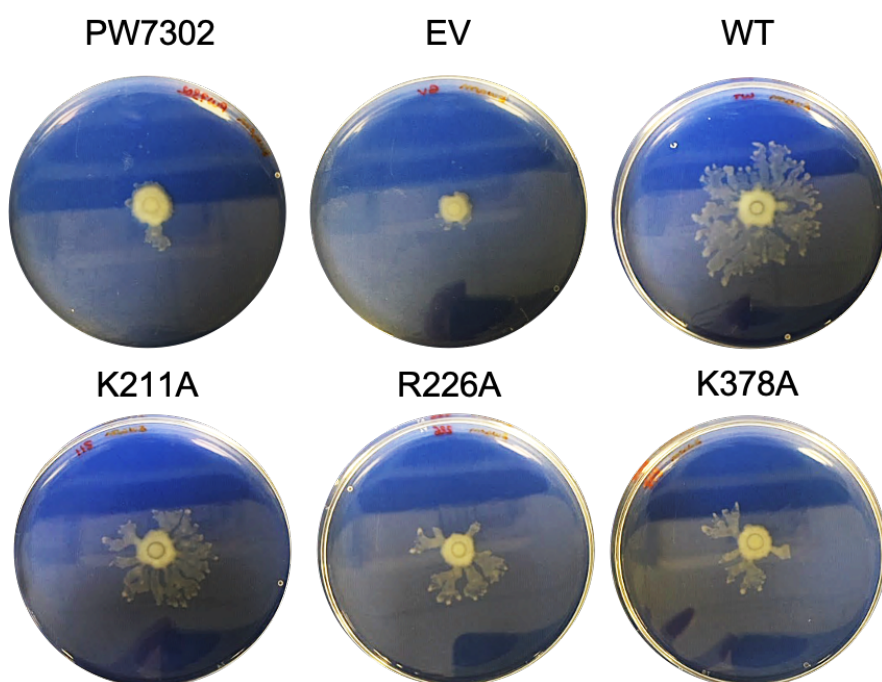
## Appendix B

Images of biological replicate plates from the LasB SDM phenotypic assays (Chapter 6)

Replicate SMA plates from the protease assay (see Section 6.4)

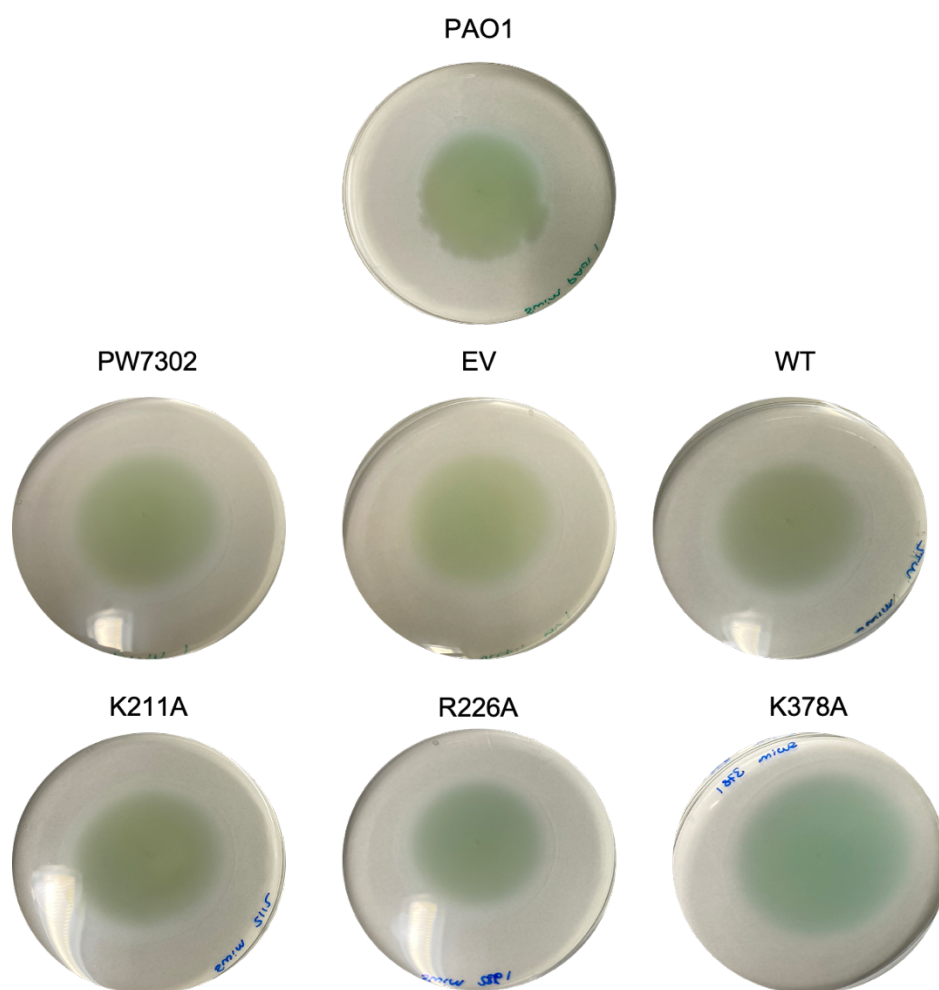


Replicate swarming assay plates, except  $\text{MgCl}_2$  and  $\text{CaCl}_2$  were excluded from the swarming media (see Section 6.7.1)





Replicate swimming assay plates (see Section 6.7.2)



## **Appendix C**

Review paper entitled “Arming the troops: Post-translational modification of extracellular bacterial proteins” published in Science Progress. This review paper was authored by myself and published during the course of this research; it is referred to in Chapter 1.

# Arming the troops: Post-translational modification of extracellular bacterial proteins

Science Progress

2020, Vol. 103(4) 1–22

© The Author(s) 2020

Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/0036850420964317

[journals.sagepub.com/home/sci](https://journals.sagepub.com/home/sci)**Suzanne Forrest and Martin Welch** 

Department of Biochemistry, University of Cambridge, Cambridge, UK

## Abstract

Protein secretion is almost universally employed by bacteria. Some proteins are retained on the cell surface, whereas others are released into the extracellular milieu, often playing a key role in virulence. In this review, we discuss the diverse types and potential functions of post-translational modifications (PTMs) occurring to extracellular bacterial proteins.

## Keywords

Post-translational modification, protein secretion, bacterial pathogens, bacterial virulence factors, proteomics, methylation, acetylation, glycosylation, lipidation

## Introduction

Until relatively recently, the nature and effects of post-translational modifications (PTMs) were principally thought to be restricted to eukaryotic systems. However, PTM in prokaryotes is now appreciated to be just as important and diverse as it is in eukaryotes.<sup>1</sup> The ever-expanding catalogue of bacterial PTMs ranges from methylation and phosphorylation of residues to the addition of complex moieties including lipids and glycans (Figure 1).<sup>1</sup> These modifications can have profound effects on proteins, altering their conformation, activity, stability and localisation, as well as interactions with other molecules.<sup>2</sup> The specific purpose of PTMs is not always clear, although they have been shown to modulate and mediate key biological processes, including central metabolism, signal transduction and virulence.<sup>3,4</sup> Not surprisingly, many reversible PTMs also appear to be involved in mediating rapid responses to environmental stimuli.<sup>5</sup>

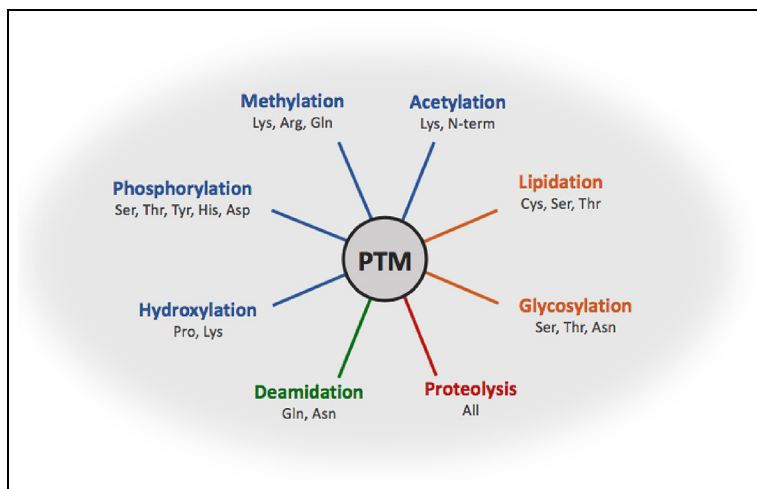
## Corresponding author:

Martin Welch, Department of Biochemistry, University of Cambridge, Hopkins Building, Tennis Court Road, Cambridge, CB2 1QW, UK.

Email: [mw240@cam.ac.uk](mailto:mw240@cam.ac.uk)



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (<https://creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).



**Figure 1.** An overview of the most common post-translational modifications in bacteria, showing the amino acid side chains which are most frequently modified. PTMs are grouped by colour according to their type; small chemical (blue), complex molecule (orange), protein cleavage (red) and amino acid side chain modifications (green).

Historically, the earliest studies focused on the PTMs associated with individual proteins, and often just on the impact of modifications at specific sites. However, advances in proteomic technologies have driven a surge in the number of large-scale global modification studies for a wide range of bacterial species.<sup>5,6</sup> Despite this, there are a number of clear limitations to the technology. High levels of purity and quantity of protein are generally necessary for sufficient sequence coverage and resolution to identify PTMs using mass spectrometry-based approaches.<sup>7</sup> This can be bolstered through selective enrichment of post-translationally modified proteins (e.g. using antibody-based approaches) prior to mass spectrometric analysis, although this is predicated on a high degree of specificity and high binding capacity of the antibodies employed.<sup>8</sup> Perhaps a more pervasive issue is that the greater the number of different types of PTM to be identified, the larger the bioinformatic search space required. This is due to an iterative search mechanism which attempts to identify the presence or absence of each PTM on every modifiable residue. This ultimately increases the rate of false discoveries.<sup>9</sup> There are also limits on the mass shift window authorised for searches, so larger modifications such as glycosylation are frequently excluded.<sup>5</sup> Consequently, independent approaches are often required to validate these “high throughput” technologies.<sup>2,10</sup> These include structural and functional studies, although as always, these too can be challenging.<sup>11</sup>

Although it is now widely accepted that PTMs occur in bacteria, most studies have focused on cell-associated protein modifications, and relatively few have considered the modifications associated with proteins secreted into the extracellular milieu. Indeed, most culture supernatants are usually removed prior to mass spectrometric analysis.<sup>6,12–14</sup> However, recent work has revealed a wealth of unexpected

PTMs associated with extracellular bacterial proteins, including phosphorylation, methylation, acetylation, proteolytic processing, glycosylation and lipidation.<sup>15</sup> In this review, we assess the diversity and likely function(s) of PTMs associated with extracellular bacterial proteins.

## Phosphorylation

Protein phosphorylation is a ubiquitous and abundant PTM, usually associated with intracellular signal transduction.<sup>16,17</sup> The attachment and removal of phosphoryl groups on amino acid side chains is catalysed by kinases and phosphatases, respectively.<sup>2</sup> Serine, threonine and tyrosine are commonly phosphorylated in eukaryotes, whereas histidine and aspartate phosphorylation are thought to be more prevalent in prokaryotes, although this is disputed.<sup>18,19</sup> Microbial phosphoanhydrides (Asp-P) and phosphoramidates (His-P) are more labile than the phosphoesters that form with Ser, Thr and Tyr side chains, and this makes them more difficult to detect, especially if samples are exposed to acidic conditions during preparation.

*Mycobacterium tuberculosis* (Mtb) is a global health burden. The organism is now strongly resurgent, partly due to widespread multi-drug resistance. Effector proteins are secreted directly from the donor Mtb cell cytoplasm to the recipient (host) cell cytoplasm through the multiprotein ESX-1 Type VII secretion system (T7SS). Not only are a sizeable number of these structural proteins phosphorylated; so too are some of the virulence factors which pass through this translocon, such as the immune-triggering proteins EsxB and PtpA.<sup>18</sup> PtpA is a tyrosine phosphatase that is secreted into macrophages following phagocytosis of the bacterium, causing inhibition of phagosome maturation and thus promoting bacterial survival. PtpA is phosphorylated at several S/T/Y residues, and this has been shown to control its activity and secretion.<sup>20</sup> Interestingly, phosphorylation of multiple virulence factors is up-regulated in the hypervirulent Mtb Beijing isolate.<sup>18</sup> Phosphorylation is also used as a regulatory mechanism to temporally control different stages of infection by other bacterial species. The *Helicobacter pylori* cytotoxin CagA is phosphorylated on tyrosine by host membrane-associated Abl1 and/or Src family kinases. Interestingly, CagA phosphorylation is stimulated by another *H. pylori* secreted product, vacuolating toxin (VacA). This effector is itself phosphorylated but also promotes the Src-mediated phosphorylation of CagA.<sup>21,22</sup> CagA and VacA effectors are involved in the early stages of gastric colonisation and are modified following injection into the epithelial cells lining the stomach.<sup>23</sup>

Not all secreted effector proteins are phosphorylated by bacterial kinases. For example, host kinases can phosphorylate some of the effector proteins produced by enteropathogenic *Escherichia coli* and *Citrobacter rodentium*.<sup>24</sup> At least 4 proteins secreted by the T2SS and T3SS of Chlamydial species are similarly modified, including TarP and TeP, which facilitate entry into the mucosal epithelia.<sup>25</sup> Bacteria can also hijack and control host systems through PTM. The Dot/Icm (T4SS) of *Legionella pneumophila* translocates over 300 effectors, including a

kinase, LegK, which can phosphorylate host proteins to interfere with normal cell functioning.<sup>26</sup>

Elastase (LasB) secreted by *Pseudomonas aeruginosa* has also been shown to be phosphorylated; indeed, 19 phospho-residues have been identified in the secreted form, whereas only non-phosphorylated LasB appears intracellularly.<sup>27</sup> The biological role(s) of this phosphorylation have not been elucidated, although it is possible that the modification targets the enzyme for secretion. Indeed, some 28 phospho-exoproteins with a wide range of functions and degree of modification were identified in one study of strain PA14. This suggests that there may be numerous roles for phosphorylation, particularly in *P. aeruginosa* virulence.<sup>28</sup>

Surface-exposed bacterial proteins are also modified in many organisms. For example, flagella proteins from several *P. aeruginosa* strains are S/T/Y phosphorylated and this modification often occurs at a very specific growth stage.<sup>29</sup> Surprisingly, one of the main growth phase-dependent flagella PTMs (phosphorylation of the FliC N-terminus) does not affect swimming motility, but instead increases Type II Secretion System (T2SS) activity whilst concomitantly decreasing biofilm formation.<sup>30</sup> Similarly, phosphorylation of the *Neisseria gonorrhoeae* type IV pilus protein, PilE, at Ser68 also has little apparent effect on the motility-related function of the protein, although it does influence the morphology of the pilus, and consequently, antigenic variation.<sup>31</sup> Outer membrane proteins (OMPs) from *Klebsiella pneumoniae*, *H. pylori* and *Shigella flexneri* are also multi-phosphorylated, although the function of these PTMs remains unclear.<sup>32,33</sup>

## Methylation

Methylation is well characterised in eukaryotes, notably the methylation of histones to control gene transcription,<sup>34</sup> but not well studied in bacteria. It involves the addition of up to two or three methyl groups to the side chain or terminal amine of arginine or lysine, respectively. Glutamine and glutamic acid residues are also modified, but to a lesser extent.<sup>1</sup> S-adenosyl methionine (SAM) acts as a methyl donor, working in concert with methyltransferase enzymes to catalyse these reactions.

The *P. aeruginosa* secreted virulence factors CbpD (chitin binding protein) and elastase are methylated at several lysine residues. However, these lysines are methylated to different degrees, with mono-, di- and tri-methylated forms of the same lysine present. This indicates that methylation can be highly variable, even for the same protein.<sup>27</sup> Variations in side chain methylation are also seen in the Gram-positive organism *Clostridium thermocellum*. This bacterium degrades cellulose by secreting a complex of different proteins, known as the cellulosome. CipA, a cellulosome structural protein, is methylated at Glu1267 and trimethylated at Lys80 and Lys663, although the protein is also secreted in an unmodified form. Contrastingly, the cellulolytic CelK enzyme is consistently di-methylated at Lys652 under different experimental conditions, suggesting that certain residue modifications are invariable. This PTM is postulated to promote protein flexibility, whereas methylation of glutamic acid may aid cellulosome localisation.<sup>10</sup>

Studies closely scrutinising secreted protein methylation are limited. However, outer membrane proteins have been studied in greater detail. OMP 32 of *Leptospira interrogans* undergoes extensive but irregular methylation. Eleven glutamic acid residues are variably modified in response to mammalian host signals.<sup>35,36</sup> This leads to OMP phase variation and reduced recognition by the host immune system. Lysine methylation also alters the antigenicity of *Rickettsia* OMPs.<sup>37</sup> Two different lysine methyltransferases modify OmpB from *Rickettsia prowazekii*, with one enzyme specifically catalysing tri-methylation at consensus sequences.<sup>38,39</sup> Interestingly, the overall number of methylated lysine residues correlates with virulence in this strain.

One of the earliest observations of methylation as a PTM occurred in 1959 during an investigation of flagella structure and function.<sup>40</sup> Over half a century later, investigations have now revealed methylation of flagella proteins in a range of species.<sup>41,42</sup> Flagellin methylation by FliB, a component of the *Salmonella enterica* serovar Typhimurium flagella machinery, is necessary for swarming motility and virulence<sup>43–45</sup> and orthologous methyltransferases produced by other members of the Enterobacteriaceae are encoded in flagellin methylase island loci (FMIs).<sup>42</sup> In addition to flagella, surface-associated pili can also be methylated. The pre-pilin peptidase (PilD) of *P. aeruginosa* acts as a dual modifier, by cleaving the signal peptide and then methylating the N-terminal phenylalanine of mature Type 4 pilus subunits. Methylation happens prior to pilus assembly and is dependent on the binding of zinc as a cofactor.<sup>46</sup> This processing also occurs in *Neisseria meningitidis*, although methylation is not a prerequisite for proper pilus assembly and the true function of this PTM remains unknown.<sup>47,48</sup>

EF-Tu is an elongation factor that delivers charged tRNA to ribosomes in the cytoplasm. It is also transported to the bacterial surface where it “moonlights” as an environmental sensor and aids adherence to epithelial cells in many species.<sup>49</sup> A protein is described to “moonlight” when it has additional function(s) that are not relevant to its primary role within the cell. *P. aeruginosa* EftM exclusively tri-methylates EF-Tu on Lys5. This modification does not impact the canonical function of the protein in translation, but does mediate bacterial attachment to the respiratory epithelia by mimicking phosphorylcholine (a component of the platelet activating factor).<sup>50,51</sup> Methylation is a prerequisite for infection, and deletion of *eftM* decreases *P. aeruginosa* pathogenicity. Interestingly, EftM is thermoregulated, displaying increased stability at 25°C. This may suggest that methylation is critical in the early stages of infection or for survival in non-host environments.<sup>50</sup> This enzyme is well-conserved throughout the *Pseudomonas* and *Vibrio* genera and methylation of EF-Tu has proven indispensable for many pathogens.<sup>51,52</sup>

## Acetylation and succinylation

Acetylation predominantly occurs on the  $\epsilon$ -amine of lysine side chains (N $\epsilon$ -acetylation) and at N-terminal amino acids (N $\alpha$ -acetylation) in a co- or post-translational fashion.<sup>1,53</sup> This type of acylation can also occur on the side chains of serine,

threonine and tyrosine (O-acetylation).<sup>2</sup> Acetylation can occur enzymatically via acetyltransferases (with acetyl groups also removed via the action of deacetylases) or non-enzymatically. Both pathways require an acetyl donor, commonly acetyl-CoA or acetyl-phosphate.<sup>1,54</sup>

Despite being proven to impact upon virulence in several species, the biological significance of acetylation of extracellular proteins is not well defined.<sup>28,55–57</sup> Acetylation of extracellular bacterial proteins increases as cultures approach and enter the stationary phase of growth. This indicates that acetylation may impact upon protein stability, perhaps circumventing the unnecessary use of scarce resources to replenish vital proteins.<sup>5</sup> The protein acetyltransferase (Pat) and deacetylase (CobB) of *S. enterica* serovar Typhimurium are involved in cell survival during growth following acidic stress, invasion of the host and replication within macrophages. *S. enterica* serovar Typhimurium mutants that are unable to acetylate proteins show reduced host inflammation, although how this relates to specific virulence factors is currently unclear.<sup>58</sup>

In a study of the *P. aeruginosa* strain PA14 intracellular lysine acetylome, 522 modified proteins were identified. Additionally, acetylation was enhanced when citrate was used as the sole carbon source.<sup>59</sup> Notably, in addition to identifying many acetylated intracellular proteins, several *P. aeruginosa* virulence factors (some previously identified as methylated, such as CbpD and LasB) were also acetylated, including protease IV, haemolysin, exotoxin A and several components of the T6SS.<sup>27,57,59</sup>

Proteins involved in central metabolism are the main targets of acetylation in Mtb. For example, malate synthase G (GlcB) is a component of the glyoxylate shunt. However, GlcB is also secreted in an acetylated form to the surface of the cell, enhancing bacterial adherence to the lung epithelium. Indeed, a further 45 secreted acetylated proteins from Mtb have been identified. Multi-acetylation of heat shock protein X (HspX) inhibits the host immune response and has been linked to the latency of Mtb infections.<sup>60</sup> ESAT-6 (6 kDa Early Secreted Antigenic Target) is inconsistently acetylated at the N-terminus, preventing protein-protein interactions with its cognate chaperone CFP-10.<sup>61</sup> ESAT-6 is able function alone or in complex with CFP-10 to modulate the host immune response, and therefore the purpose of acetylation-driven segregation of these proteins is unclear.<sup>62</sup> Other members of the ESAT-6 family also undergo N $\alpha$ -acetylation, including EsxN, EsxO, EsxI and EsxA.<sup>63–65</sup> Dephosphorylation of host proteins by PtbB is also critical for Mtb infection. The phosphatase activity of PtbB is controlled by acetylation/succinylation of Lys224, which is found in the lid region that covers the active site. This PTM therefore serves as a negative regulator of enzyme activity.<sup>65</sup>

Succinylation is another form of acylation alongside acetylation, malonylation, propionylation, butyrylation and crotonylation. Although the identification of succinylated proteins is a relatively recent area of research, the overlap between acetylation and succinylation in the secretome of PA14 is significant,<sup>66</sup> with around 41% of succinylation sites also susceptible to acetylation.<sup>59</sup> The functional significance of succinylation remains unclear, although the presence of succinyllysine



within the pro-peptide of LasB indicates a potential role of this PTM in protein maturation or stability. Moreover, the number of acetylated/succinylated lysine residues in LasB increases upon secretion.<sup>27</sup> Global profiling of succinylated proteins in PA14 identified seven sequence motifs that may direct modification, with the same signatures also evident in *Vibrio parahaemolyticus* and Mtb.<sup>59,66,67</sup> This suggests that succinylation of secreted proteins may be a more widespread PTM than previously thought.

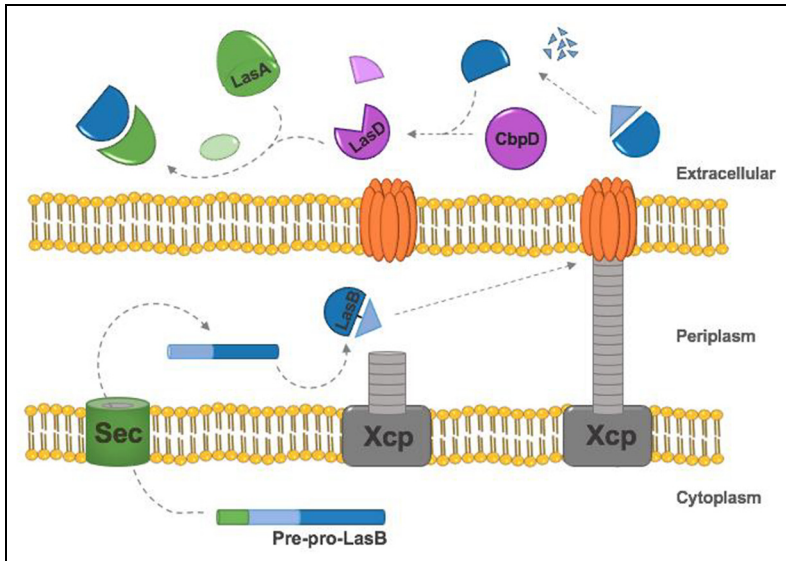
## Proteolysis

Proteolytic cleavage of proteins is a common and irreversible PTM. Endoproteases cleave the polypeptide chain at specific residues within sequence motifs, whereas exoproteases cleave the N- and C- termini.

Many extracytoplasmic proteins are transported across the cytoplasmic membrane (CM) through the sec translocation machinery, guided by an N-terminal signal peptide.<sup>68</sup> In Gram-negative bacteria, additional secretion systems are employed to further transport proteins from the periplasm to outside of the cell.<sup>69</sup> Inhibition of the type 1 signal peptidases that cleave signal peptides results in the accumulation of unprocessed proteins in the cytoplasm and ultimately, cell death.<sup>70,71</sup> Hidden Markov models can accurately predict signal peptide sequences, although different algorithms are necessary for Gram-positive and Gram-negative bacteria.<sup>72</sup> AXA motifs are common at the N-terminal cleavage site,<sup>2,73</sup> but this can vary greatly between species.<sup>74</sup> Somewhat controversially, in a study of Mtb secreted proteins, only 16% of secreted proteins had a cleaved signal peptide; this has also been seen in *Bacillus subtilis*.<sup>72,75</sup> To our knowledge, the reason(s) for these observed discrepancies in the cleavage of signal peptides have not been investigated.

The substrates of the Por (type 9 secretion system, T9SS) of periodontitis-associated *Porphyromonas gingivalis* contain conserved C-terminal domains (CTD) that are essential for secretion. The principal virulence factors translocated through the T9SS are cysteine protease gingipains.<sup>76</sup> Cleavage of the gingipain CTDs by PorU and PorZ following protein folding allows secretion through the outer membrane (OM) and subsequent glycosylation.<sup>77–80</sup> Current data implicate the tertiary structure of the CTD as being the key signal for secretion,<sup>80,81</sup> although some conserved sequence motifs have also been identified.<sup>82</sup> *Bordetella* filamentous haemagglutinin (FHA) also harbours a cleavable CTD. FHA is retained on the cell surface *via* anchoring of the CTD within the FhaC transporter.<sup>83</sup> Cleavage of the CTD releases FHA from the surface, allowing passage of full-length unprocessed FHA through FhaC to act as a transmembrane sensor.<sup>84</sup> This intriguing interplay between the two forms of FHA is tightly-regulated, ensuring successful colonisation and maintenance of infection.

An elegant cleavage cascade which activates several virulence factors has been highlighted in *P. aeruginosa* (Figure 2). Elastase is synthesised as a pre-pro-protein (53 kDa) which is cleaved and transported through the CM and OM by targeting



**Figure 2.** *Pseudomonas aeruginosa* LasB secretion pathway and subsequent proteolytic cleavage cascade. The pre-pro-protein is targeted to the periplasm through the sec translocon via the encoded signal peptide. The protein is then folded, and the pro-domain is cleaved, remaining non-covalently bound to the mature protein. Following secretion via the T2SS Xcp machinery, the pro-domain is degraded and a cleavage cascade begins. Mature LasB cleaves CbpD into LasD, which then activates LasA by proteolysis. Mature LasA enhances the elastolytic activity of LasB.

and sequential cleavage of the pre- and pro-domains, respectively. The pro-domain undergoes autoproteolytic cleavage post-folding in the periplasm but remains non-covalently linked to the mature protein (33 kDa) where it inhibits intracellular protease activity.<sup>85</sup> Both domains are then secreted through the Xcp T2SS machinery and the pro-domain is subsequently degraded. A similar secretion pathway can be seen in the production of subtilisin by *Bacillus subtilis*.<sup>86</sup> Extracellular LasB processes CbpD at the N-terminus, yielding LasD, which in turn, cleaves LasA into its mature staphylytic form.<sup>85,87,88</sup> The processing of these enzymes works as a positive feed-forward mechanism because LasA further enhances the elastolytic activity of LasB.<sup>27,85</sup> LasB also activates leucine aminopeptidase by cleaving the C-terminal pro-sequence, which contains an active site-inhibitory lysine residue. Substitution of this C-terminal lysine with an acidic alternative results in leucine aminopeptidase activation without the need for LasB processing.<sup>89</sup>

Some cleavage events produce multiple functional products. For example, autolysin synthesised by *Staphylococcus aureus* is cleaved at four locations to produce functionally distinct extracellular hydrolases, an amidase and a glucosaminidase.<sup>68</sup> In several diverse pathogens, EF-Tu is also cleaved into fragments which are then expressed on the cell surface. Here they act as adherence factors, binding a range of host substrates and mediating colonisation of different niches.<sup>52</sup>

## Glycosylation

Glycosylation involves the covalent attachment of a carbohydrate to the amide group of asparagine (N-linked) or the hydroxyl group of serine or threonine (O-linked).<sup>90–95</sup> Attachment of glycans is a multi-step process involving different enzymes,<sup>96</sup> which are commonly encoded within gene clusters with their substrates.<sup>42</sup> Promiscuous glucosyltransferases generate extensive variability in the glycosyl moieties of modified proteins.<sup>97,98</sup> Two main glycosylation pathways are used in bacteria, either *via* the *en bloc* transfer of glycan chains (preassembled on lipid carriers) to proteins, or sequential attachment from nucleotide-activated sugars.<sup>99</sup>

There are few examples of bacterial glycoproteins which are fully secreted into the extracellular milieu; most such modified proteins remain attached to the cell.<sup>100–102</sup> The best-studied glycoproteins are flagellins.<sup>41,103–105</sup> *Campylobacter jejuni* flagella are N-glycosylated by PglB with a nine-carbon pseudaminic acid, which transfers glycan moieties *en bloc* to proteins at the sequon D/E-X-N-X-S/T (X ≠ Pro),<sup>106,107</sup> although some exceptions to this rule have been found.<sup>108</sup> Functionally, glycosylation may have multiple roles. For example, adherence of *Clostridium difficile* in the human gut is dependent on N-acetyl-glucosamine glycosylation of flagellar proteins. This glycan-induced switch to a more sessile mode aids biofilm formation.<sup>109</sup> Flagellar glycosylation may also play a protective role. For example, the flagella of *Burkholderia cepacia* are modified at >10 sites, and these modifications are required for auto-agglutination, resistance to acid, and blocking of toll-like receptor 5 recognition.<sup>110–112</sup>

Many Pseudomonads encode a genomic glycosylation island as part of the flagella regulon.<sup>98,103</sup> *P. aeruginosa* produces two types of flagellin proteins which are distinguished by their antigenicity. Both types are glycosylated, although they are modified by different machinery encoded by different gene clusters. O-linked glycosylation occurs at Thr189 and Ser260 found in the central surface-exposed domain of each A type flagellin unit.<sup>98</sup> Interestingly, type B flagellins are also glycosylated twice, at Ser191 and Ser195.<sup>103</sup> Whilst the specific role of O-linked flagella glycosylation is unclear in *P. aeruginosa*, glycosylation of flagella in *Pseudomonas syringae* is involved in bacterial recognition by plants, and can shape host specificity.<sup>113</sup>

O-glycosylation is commonly used to protect surface-associated proteins from degradation by extracellular proteases. *H. pylori* alpha and beta ureases, *Microcystis aeruginosa* microcystin-related protein C, and *Streptococcus mutans* binding protein Cnm all undergo O-glycosylation to increase protein stability.<sup>97,114,115</sup> This is also the case for EmaA, the only *Aggregatibacter actinomycetemcomitans* autotransporter adhesin which is currently known to be glycosylated. The other adhesins, ApiA and AaE, are unmodified.<sup>116</sup> Additional autotransporter proteins and adhesins from unrelated bacteria are also secreted as glycoproteins, including AtaC from *Actinobacillus pleuropneumoniae*, and the self-assembling TibA from *E. coli*.<sup>117,118</sup> Glycosylation of TibA by TibC enhances stability and adhesion to epithelial cells, although the

modification is not known to affect invasion or aggregation of *E. coli*, unlike its other autotransporter glycoproteins.<sup>117</sup>

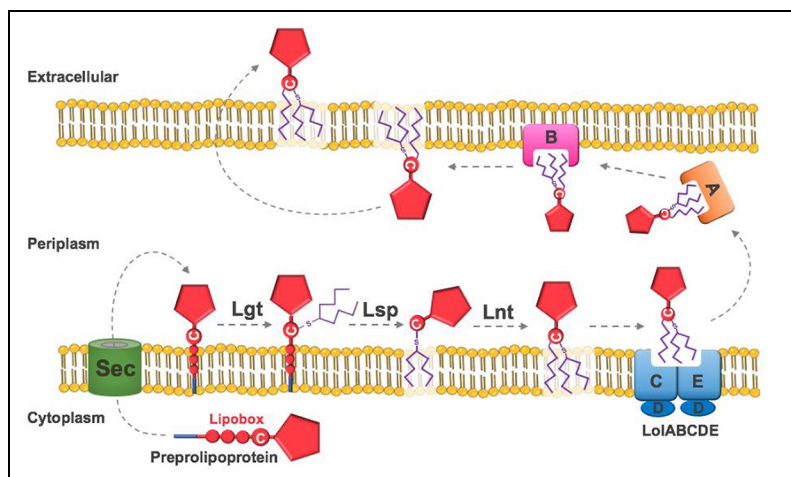
Adherence-promoting surface fimbriae can also be modified by O-linked glycosylation. For example, unmodified fimbria-associated glycoprotein (Fap1) from *Streptococcus parasanguinis* mediates cell adhesion, whereas following glycosylation, the protein becomes essential for biofilm development.<sup>119</sup> Variable glycosylation of Gram-negative pili is also widespread.<sup>48,99</sup> Two of the five type 4 pili produced by *P. aeruginosa* are modified by PilO via the addition of O-antigen or polymers of D-arabinofuranose. The even distribution of glycans on the pili fibrils confers protection against infection by bacteriophage which target pili as receptors for adsorption.<sup>120</sup> In true tit-for-tat competitive style however, phage can mutate their tail proteins and adapt their specificity to re-sensitise against the bacterial host.<sup>121</sup> Overall glycosylation of surface proteins contributes towards enhanced protection and stability; improved adherence, invasion and immune stimulation; and increased uptake of DNA.<sup>107,122</sup>

## Lipidation

Lipidation is another complex PTM, typically involving the addition of two or three lipids to the N-terminal cysteine of lipobox sequence motifs.<sup>123,124</sup> The lipid moieties act as a surface anchor and tend to reflect the fatty acid composition of membrane phospholipids. This contributes to the significant antigenic variability of lipoproteins between- and within-species.<sup>125,126</sup> The lipoprotein biosynthesis pathway involves up to three sequential enzymatic reactions (Figure 3).<sup>127,128</sup> Interestingly, and despite not encoding a homologue of the final N-acyltransferase enzyme (Lnt), many low G + C Gram-positive organisms are still able to produce triacylated lipoproteins.<sup>129–131</sup>

Lipoproteins play a substantial role in bacterial growth and pathogenicity. The role(s) played by lipidation are nothing if not diverse. For example, lipidation enhances Streptococcal adherence to endothelial cells, and loss of a single *S. sanguinis* lipoprotein (the metal ion transporter, SsaB) drastically decreases the ability of the organism to cause endocarditis.<sup>133–135</sup> Lipidation is also known to affect the flagellar-driven motility of *C. jejuni* in the gut<sup>136</sup> and significantly impacts the virulence of *Enterococcus faecalis*. Somewhat unexpectedly, loss of the lipidation enzymes also enhances *E. faecalis* growth under stressed conditions.<sup>137</sup> In contrast, the growth of *S. aureus* during nutrient limitation has been reported to be dependent on lipidation.<sup>138</sup> Interestingly, *S. aureus* triacyl-lipoprotein production is strongly-dependent on environmental parameters and growth phase, with diacyl forms dominating in high-stress conditions.<sup>139</sup> Indeed, the role of the third fatty acid has been questioned in Gram-positive organisms as triacylation is generally thought to target lipoproteins across the outer membrane.

*N. meningitidis* surface antigen Factor H binding protein (FHbp) is normally tri-palmitoylated *in vivo*. In contrast to previous reports, deletion of Lnt (which adds the third palmitoyl group) does not prevent the surface expression of diacylated



**Figure 3.** Biosynthesis pathway of lipoproteins in *E. coli*. Preprolipoproteins are synthesised in the cytoplasm and targeted through the sec translocon to the periplasm via their N-terminal signal peptide. Diacylglycerol transferase (Lgt) transfers diacylglycerol to the thiol group of the last cysteine in the four-residue lipobox motif. After this, lipoprotein signal peptidase (Lsp) cleaves the signal peptide immediately before the derivatised cysteine and apolipoprotein N-acyltransferase (Lnt) transfers another acyl group to the cysteine N $\alpha$  moiety. The mature lipoprotein is recognised by the Lol ABC transporter, which transfers the lipoprotein to the outer membrane.<sup>132</sup>

FHbp.<sup>140,141</sup> However, the outer membrane Lol transporter has a higher affinity for the triacylated form and accumulation of diacylated FHbp results in negative feedback and reduced overall FHbp synthesis. It is possible that production of both the di- and tri-acylated FHbp isoforms may confer a fitness advantage in terms of antigen recognition, or that sole production of the diacylated FHbp isoform leads to a fitness disadvantage (e.g. *via* elevated antibiotic sensitivity due to increased membrane permeability<sup>142</sup>).

Surface dissociation of di- or triacylated lipoproteins during infection activates host TLR2/6 or TLR2/1 heterodimers respectively, which drives inflammation.<sup>132,143,144</sup> In *S. aureus*, lipoprotein release is mediated by quorum sensing-controlled expression of surfactant-like phenol-soluble modulins (PSMs). These small peptides not only promote the release of lipoprotein-containing membrane vesicles during cell turgor in hypotonic conditions, but themselves act as potent virulence factors.<sup>145</sup> Another virulence factor that is regulated by lipidation is the secreted pore-forming toxin, haemolysin, from *E. coli*. This toxin induces apoptosis in target host cells and is activated prior to export by the addition of fatty acids to two internal lysine residues.<sup>146,147</sup> Many other species also use lipidation to regulate haemolysin activity e.g., palmitoylation of *B. pertussis* haemolysin.<sup>148,149</sup>

The localisation of virulence factors within host cells is important for their biological function. The HopZ family of T3SS effectors produced by *P. syringae* are targeted to the plant host plasma membranes *via* myristoylation.<sup>150,151</sup> Proper targeting of these effectors causes programmed cell death through immune modulation.<sup>152</sup> In a remarkable example of molecular subterfuge, some pathogens encode a CAAX motif on their secreted proteins; this motif is lipidated by host acyltransferases, thereby targeting the effectors to distinct organelles. Palmitoylation by host cell enzymes is also essential for the proper localisation, stability and activity of the polyclonal B-cell mitogen, PrpA, produced by *Brucella* species.<sup>153</sup> Indeed, this may be a common feature of many intracellular pathogens, since secreted proteins produced by *L. pneumophila* and *S. enterica* serovar Typhimurium have also been shown to hijack the host cell machinery during targeting, thereby presumably conserving costly resources.<sup>154–156</sup>

## Concluding remarks

Since the start of the 21st century, considerable gains have been made in the field of prokaryotic PTM research. However, until recently, the importance of extracellular protein PTMs has been largely underestimated. Nonetheless, many secreted proteins from diverse bacterial species have been shown to be modified in a variety of ways. Although some of these modifications are proving essential for bacterial physiology and virulence, the purpose of many others remains unresolved. What is clear though, is that PTM is not a singular mechanism of control, and global studies of bacterial proteomes have shown considerable overlap between modifications.<sup>2,157</sup> This notwithstanding, the era of defining the type and extent of such PTMs now seems to be drawing to a close; the challenge for the future generation will be in defining the biology underpinning these modifications.


## Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: SF was supported by a studentship from the BBSRC DTP and by funding from the Evelyn Trust (Cambridge). Work in the MW laboratory is supported by the Cystic Fibrosis Trust, The British Lung Foundation and the BBSRC.

## ORCID iD

Martin Welch  <https://orcid.org/0000-0003-3646-1733>

## References

1. Cain JA, Solis N and Cordwell SJ. Beyond gene expression: the impact of protein post-translational modifications in bacteria. *J Proteomics* 2014; 97: 265–286.
2. André P, Bastos D, Pinto J, et al. A glimpse into the modulation of post-translational modifications of human-colonizing bacteria. *J Proteomics* 2017; 152: 254–275.
3. Grangeasse C, Stülke J and Mijakovic I. Regulatory potential of post-translational modifications in bacteria. *Front Microbiol* 2015; 6: 2014–2015.
4. Michard C and Doublet P. Post-translational modifications are key players of the *Legionella pneumophila* infection strategy. *Front Microbiol* 2015; 6: 1–12.
5. Brown CW, Sridhara V, Boutz DR, et al. Large-scale analysis of post-translational modifications in *E. Coli* under glucose-limiting conditions. *BMC Genomics* 2017; 18: 1–21.
6. Cao XJ, Dai J, Xu H, et al. High-coverage proteome analysis reveals the first insight of protein modification systems in the pathogenic spirochete *Leptospira interrogans*. *Cell Res* 2010; 20: 197–210.
7. Silva AMN, Vitorino R, Domingues MRM, et al. Post-translational modifications and mass spectrometry detection. *Free Radic Biol Med* 2013; 65: 925–941.
8. Zhao Y and Jensen ON. Modification-specific proteomics: strategies for characterization of post-translational modifications using enrichment techniques. *Proteomics* 2009; 9: 4632–4641.
9. Ahrné E, Müller M and Lisacek F. Unrestricted identification of modified proteins using MS/MS. *Proteomics* 2010; 10: 671–686.
10. Dykstra AB, Rodriguez M, Raman B, et al. Characterizing the range of extracellular protein post-translational modifications in a cellulose-degrading bacteria using a multiple proteolytic digestion/peptide fragmentation approach. *Anal Chem* 2013; 85: 3144–3151.
11. Macek B, Forchhammer K, Hardouin J, et al. Protein post-translational modifications in bacteria. *Nat Rev Microbiol*. Epub ahead of print September 2019. DOI: 10.1038/s41579-019-0243-0.
12. Houser JR, Barnhart C, Boutz DR, et al. Controlled measurement and comparative analysis of cellular components in *E. Coli* reveals broad regulatory changes in response to glucose starvation. *PLoS Comput Biol* 2015; 11: 1–27.
13. Kosono S, Tamura M, Suzuki S, et al. Changes in the acetylome and succinylome of *Bacillus subtilis* in response to carbon source. *PLoS One* 2015; 10: 1–24.
14. Soufi B, Krug K, Harst A, et al. Characterization of the *E. Coli* proteome and its modifications during growth and ethanol stress. *Front Microbiol* 2015; 6: 1–11.
15. Maffei B, Francetic O and Subtil A. Tracking proteins secreted by bacteria: what's in the toolbox? *Front Cell Infect Microbiol* 2017; 7: 1–17.
16. Tiwari S, Jamal SB, Hassan SS, et al. Two-component signal transduction systems of pathogenic bacteria as targets for antimicrobial therapy: an overview. *Front Microbiol* 2017; 8: 1–7.
17. Soares NC, Spät P, Méndez JA, et al. Ser/Thr/Tyr phosphoproteome characterization of *Acinetobacter baumannii*: comparison between a reference strain and a highly invasive multidrug-resistant clinical isolate. *J Proteomics* 2014; 102: 113–124.
18. Fortuin S, Tomazella GG, Nagaraj N, et al. Phosphoproteomics analysis of a clinical *Mycobacterium tuberculosis* Beijing isolate: expanding the mycobacterial phosphoproteome catalog. *Front Microbiol* 2015; 6: 1–12.

19. Junker S, Maa S, Otto A, et al. Spectral library based analysis of arginine phosphorylations in *Staphylococcus aureus*. *Mol Cell Proteomics*. Epub ahead of print November 2018. DOI: 10.1074/mcp.RA117.000378.
20. Zhou P, Li W, Wong D, et al. Phosphorylation control of protein tyrosine phosphatase A activity in *Mycobacterium tuberculosis*. *FEBS Lett* 2015; 589: 326–331.
21. Abdullah M, Greenfield LK, Bronte-Tinkew D, et al. VacA promotes CagA accumulation in gastric epithelial cells during *Helicobacter pylori* infection. *Sci Rep* 2019; 9: 1–9.
22. Ge R, Sun X, Xiao C, et al. Phosphoproteome analysis of the pathogenic bacterium *Helicobacter pylori* reveals over-representation of tyrosine phosphorylation and multiply phosphorylated proteins. *Proteomics* 2011; 11: 1449–1461.
23. Salama NR, Otto G, Tompkins L, et al. Vacuolating cytotoxin of *Helicobacter pylori* plays a role during colonization in a mouse model of infection. *Infect Immun* 2001; 69: 730–736.
24. Hansen AM, Chaerkady R, Sharma J, et al. The *Escherichia coli* phosphotyrosine proteome relates to core pathways and virulence. *PLoS Pathog*. Epub ahead of print June 2013. DOI: 10.1371/journal.ppat.1003403.
25. Claywell JE, Matschke LM and Fisher DJ. The impact of protein phosphorylation on chlamydial physiology. *Front Cell Infect Microbiol* 2016; 6: 1–8.
26. Flayhan A, Bergé C, Bailo N, et al. The structure of *Legionella pneumophila* LegK4 type four secretion system (T4SS) effector reveals a novel dimeric eukaryotic-like kinase. *Sci Rep* 2015; 5: 1–13.
27. Gaviard C, Cosette P, Jouenne T, et al. LasB and CbpD virulence factors of *Pseudomonas aeruginosa* carry multiple post-translational modifications on their lysine residues. *J Proteome Res* 2019; 18: 923–933.
28. Ouidir T, Jarnier F, Cosette P, et al. Extracellular Ser/Thr/Tyr phosphorylated proteins of *Pseudomonas aeruginosa* PA14 strain. *Proteomics* 2014; 14: 2017–2030.
29. Kelly-Wintenberg K, South SL and Montie TC. Tyrosine phosphate in a- and b-type flagellins of *Pseudomonas aeruginosa*. *J Bacteriol* 1993; 175: 2458–2461.
30. Suriyanarayanan T, Periasamy S, Lin MH, et al. Flagellin FliC phosphorylation affects type 2 protease secretion and biofilm dispersal in *Pseudomonas aeruginosa* PAO1. *PLoS One* 2016; 11: 1–19.
31. Forest KT, Dunham SA, Koomey M, et al. Crystallographic structure reveals phosphorylated pilin from *Neisseria*: phosphoserine sites modify type IV pilus surface chemistry and fibre morphology. *Mol Microbiol* 1999; 31: 743–752.
32. Lai SJ, Tu IF, Wu WL, et al. Site-specific His/Asp phosphoproteomic analysis of prokaryotes reveals putative targets for drug resistance. *BMC Microbiol* 2017; 17: 1–10.
33. Standish AJ, Teh MY, Tran ENH, et al. Unprecedented abundance of protein tyrosine phosphorylation modulates *Shigella flexneri* virulence. *J Mol Biol* 2016; 428: 4197–4208.
34. Greer EL and Shi Y. Histone methylation: a dynamic mark in health, disease and inheritance. *Nat Rev Genet* 2012; 13: 343–357.
35. Eshghi A, Pinne M, Haake DA, et al. Methylation and in vivo expression of the surface exposed *Leptospira interrogans* outer-membrane protein OmpL32. *Microbiology* 2012; 158: 622–635.
36. Nally JE, Grassmann AA, Planchon S, et al. Pathogenic leptospires modulate protein expression and post-translational modifications in response to mammalian host signals. *Front Cell Infect Microbiol* 2017; 7: 1–15.



37. Abeykoon AH, Noinaj N, Choi BE, et al. Structural insights into substrate recognition and catalysis in outer membrane protein B (OmpB) by protein-lysine methyltransferases from rickettsia. *J Biol Chem* 2016; 291: 19962–19974.
38. Abeykoon AH, Chao CC, Wang G, et al. Two protein lysine methyltransferases methylate outer membrane protein B from Rickettsia. *J Bacteriol* 2012; 194: 6410–6418.
39. Abeykoon A, Wang G, Chao CC, et al. Multimethylation of Rickettsia OmpB catalyzed by lysine methyltransferases. *J Biol Chem* 2014; 289: 7691–7701.
40. Ambler RP and Rees MW.  $\epsilon$ -N-methyl-lysine in bacterial flagellar protein. *Nature* 1959; 184: 56–57.
41. Bubendorfer S, Ishihara M, Dohlich K, et al. Analyzing the modification of the Shewanella oneidensis MR-1 flagellar filament. *PLoS One* 2013; 8: 19–21.
42. De Maayer P and Cowan DA. Flashy flagella: flagellin modification is relatively common and highly versatile among the Enterobacteriaceae. *BMC Genomics* 2016; 17: 1–13.
43. Tronick SR and Martinez RJ. Methylation of the flagellin of Salmonella typhimurium. *J Bacteriol* 1971; 105: 211–219.
44. Bogomolnaya LM, Aldrich L, Ragoza Y, et al. Identification of novel factors involved in modulating motility of salmonella enterica serotype typhimurium. *PLoS One*. Epub ahead of print November 2014. DOI: 10.1371/journal.pone.0111513.
45. Frye J, Karlinsey JE, Felise HR, et al. Identification of new flagellar genes of Salmonella enterica serovar Typhimurium. *J Bacteriol* 2006; 188: 2233–2243.
46. Aly KA, Beebe ET, Chan CH, et al. Cell-free production of integral membrane aspartic acid proteases reveals zinc-dependent methyltransferase activity of the Pseudomonas aeruginosa prepilin peptidase PilD. *Microbiologyopen* 2013; 2: 94–104.
47. Pepe JC and Lory S. Amino acid substitutions in PilD, a bifunctional enzyme of Pseudomonas aeruginosa: effect on leader peptidase and N-methyltransferase activities in vitro and in vivo. *J Biol Chem* 1998; 273: 19120–19129.
48. Gault J, Malosse C, Machata S, et al. Complete posttranslational modification mapping of pathogenic Neisseria meningitidis pilins requires top-down mass spectrometry. *Proteomics* 2014; 14: 1141–1151.
49. N'Diaye AR, Borrel V, Racine PJ, et al. Mechanism of action of the moonlighting protein Eftu as a substance P sensor in Bacillus cereus. *Sci Rep* 2019; 9: 1–14.
50. Prezioso SM, Duong DM, Kuiper EG, et al. Trimethylation of elongation factor-Tu by the dual thermoregulated methyltransferase EftM does not impact its canonical function in translation. *Sci Rep* 2019; 9: 1–12.
51. Barbier M, Owings JP, Barbier M, et al. Lysine trimethylation of EF-Tu mimics platelet-activating factor to initiate Pseudomonas aeruginosa pneumonia. *Am Soc Microbiol* 2013; 4: 3–10.
52. Widjaja M, Harvey KL, Hagemann L, et al. Elongation factor Tu is a multifunctional and processed moonlighting protein. *Sci Rep* 2017; 7: 1–17.
53. Christensen DG, Xie X, Basisty N, et al. Post-translational protein acetylation: an elegant mechanism for bacteria to dynamically regulate metabolic functions. *Front Microbiol* 2019; 10: 1–22.
54. Christensen DG, Baumgartner JT, Xie X, et al. Mechanisms, detection, and relevance of protein acetylation in prokaryotes. *MBio* 2019; 10: 1–20.
55. Ren J, Sang Y, Lu J, et al. Protein acetylation and its role in bacterial virulence. *Trends Microbiol* 2017; 25: 768–779.

56. Zhang J, Sprung R, Pei J, et al. Lysine acetylation is a highly abundant and evolutionarily conserved modification in *Escherichia coli*. *Mol Cell Proteomics* 2009; 8: 215–225.
57. Ouidir T, Cosette P, Jouenne T, et al. Proteomic profiling of lysine acetylation in *Pseudomonas aeruginosa* reveals the diversity of acetylated proteins. *Proteomics* 2015; 15: 2152–2157.
58. Yu S, Ren J, Ni J, et al. Protein acetylation is involved in *Salmonella enterica* serovar typhimurium virulence. *J Infect Dis* 2016; 213: 1836–1845.
59. Gaviard C, Broutin I, Cosette P, et al. Lysine succinylation and acetylation in *Pseudomonas aeruginosa*. *J Proteome Res* 2018; 17: 2449–2459.
60. Liu F, Yang M, Wang X, et al. Acetylome analysis reveals diverse functions of lysine acetylation in *Mycobacterium tuberculosis*. *Mol Cell Proteomics* 2014; 13: 3352–3366.
61. Okkels LM, Müller EC, Schmid M, et al. CFP10 discriminates between nonacetylated and acetylated ESAT-6 of *Mycobacterium tuberculosis* by differential interaction. *Proteomics* 2004; 4: 2954–2960.
62. Sreejit G, Ahmed A, Parveen N, et al. The ESAT-6 protein of *Mycobacterium tuberculosis* interacts with beta-2-microglobulin ( $\beta$ 2M) affecting antigen presentation function of macrophage. *PLoS Pathog*. Epub ahead of print October 2014. DOI: 10.1371/journal.ppat.1004446.
63. Lange S, Rosenkrands I, Stein R, et al. Analysis of protein species differentiation among mycobacterial low-Mr-secreted proteins by narrow pH range Immobililine gel 2-DE-MALDI-MS. *J Proteomics* 2014; 97: 235–244.
64. Berrêdo-Pinho M, Kalume DE, Correa PR, et al. Proteomic profile of culture filtrate from the Brazilian vaccine strain *Mycobacterium bovis* BCG Moreau compared to *M. Bovis* BCG Pasteur. *BMC Microbiol* 2011; 11: 80.
65. Singhal A, Arora G, Virmani R, et al. Systematic analysis of mycobacterial acylation reveals first example of acylation-mediated regulation of enzyme activity of a bacterial phosphatase. *J Biol Chem* 2015; 290: 26218–26234.
66. Xie L, Liu W, Li Q, et al. First succinyl-proteome profiling of extensively drug-resistant *Mycobacterium tuberculosis* revealed involvement of succinylation in cellular physiology. *J Proteome Res* 2015; 14: 107–119.
67. Pan J, Chen R, Li C, et al. Global analysis of protein lysine succinylation profiles and their overlap with lysine acetylation in the marine bacterium *Vibrio parahaemolyticus*. *J Proteome Res* 2015; 14: 4309–4318.
68. Ravipaty S and Reilly JP. Comprehensive characterization of methicillin-resistant *Staphylococcus aureus* subsp. *aureus* COL secretome by two-dimensional liquid chromatography and mass spectrometry. *Mol Cell Proteomics* 2010; 9: 1898–1919.
69. Green ER and Mecsas J. Bacterial secretion systems: an overview. *Virulence Mech Bact Pathog* 2016; 4: 215–239.
70. Van Roosmalen ML, Geukens N, Jongbloed JDH, et al. Type I signal peptidases of Gram-positive bacteria. *Biochim Biophys Acta* 2004; 1694: 279–297.
71. Bonnemain C, Raynaud C, Réglier-Poupet H, et al. Differential roles of multiple signal peptidases in the virulence of *Listeria monocytogenes*. *Mol Microbiol* 2004; 51: 1251–1266.
72. Liversen NA, de Souza GA, Målen H, et al. Evaluation of signal peptide prediction algorithms for identification of mycobacterial signal peptides using sequence data from proteomic methods. *Microbiology* 2009; 155: 2375–2383.

73. Målen H, Berven FS, Fladmark KE, et al. Comprehensive analysis of exported proteins from *Mycobacterium tuberculosis* H37Rv. *Proteomics* 2007; 7: 1702–1718.
74. Payne SH, Bonissone S, Wu S, et al. Unexpected diversity of signal peptides in prokaryotes. *MBio* 2012; 3: 1–6.
75. de Souza GA, Leversen NA, Målen H, et al. Bacterial proteins with cleaved or uncleaved signal peptides of the general secretory pathway. *J Proteomics* 2011; 75: 502–510.
76. Glew MD, Veith PD, Peng B, et al. PG0026 is the C-terminal signal peptidase of a novel secretion system of *Porphyromonas gingivalis*. *J Biol Chem* 2012; 287: 24605–24617.
77. Shoji M and Nakayama K. Glycobiology of the oral pathogen *Porphyromonas gingivalis* and related species. *Microb Pathog* 2015; 94: 35–41.
78. Shoji M, Sato K, Yukitake H, et al. Por secretion system-dependent secretion and glycosylation of *Porphyromonas gingivalis* hemin-binding protein 35. *PLoS One*. Epub ahead of print June 2011. DOI: 10.1371/journal.pone.0021372.
79. Gorasia DG, Veith PD, Chen D, et al. *Porphyromonas gingivalis* type IX secretion substrates are cleaved and modified by a sortase-like mechanism. *PLoS Pathog* 2015; 11: 1–31.
80. Lasica AM, Goulas T, Mizgalska D, et al. Structural and functional probing of PorZ, an essential bacterial surface component of the type-IX secretion system of human oral-microbiomic *Porphyromonas gingivalis*. *Sci Rep* 2016; 6: 1–22.
81. Lasica AM, Ksiazek M, Madej M, et al. The type IX secretion system (T9SS): highlights and recent insights into its structure and function. *Front Cell Infect Microbiol*. Epub ahead of print May 2017. DOI: 10.3389/fcimb.2017.00215.
82. Veith PD, Nor Muhammad NA, Dashper SG, et al. Protein substrates of a novel secretion system are numerous in the Bacteroidetes phylum and have in common a cleavable C-terminal secretion signal, extensive post-translational modification, and cell-surface attachment. *J Proteome Res* 2013; 12: 4449–4461.
83. Inatsuka CS, Julio SM and Cotter PA. *Bordetella* filamentous hemagglutinin plays a critical role in immunomodulation, suggesting a mechanism for host specificity. *Proc Natl Acad Sci USA* 2005; 102: 18578–18583.
84. Melvin JA, Scheller EV, Noël CR, et al. New insight into filamentous hemagglutinin secretion reveals a role for full-length FhaB in *Bordetella* virulence. *MBio* 2015; 6: 12–15.
85. Braun P, De Groot A, Bitter W, et al. Secretion of elastinolytic enzymes and their propeptides by *Pseudomonas aeruginosa*. *J Bacteriol* 1998; 180: 3467–3469.
86. Forster BM, Zemansky J, Portnoy DA, et al. Posttranslocation chaperone PrsA2 regulates the maturation and secretion of *listeria monocytogenes* proprotein virulence factors. *J Bacteriol* 2011; 193: 5961–5970.
87. Folders J, Tommassen J, Van Loon LC, et al. Identification of a chitin-binding protein secreted by *Pseudomonas aeruginosa*. *J Bacteriol* 2000; 182: 1257–1263.
88. Park S and Galloway DR. *Pseudomonas aeruginosa* LasD processes the inactive LasA precursor to the active protease form. *Arch Biochem Biophys* 1998; 357: 8–12.
89. Sarnovsky R, Rea J, Makowski M, et al. Proteolytic cleavage of a C-terminal prosequence, leading to autoprocessing at the N terminus, activates leucine aminopeptidase from *Pseudomonas aeruginosa*. *J Biol Chem* 2009; 284: 10243–10253.
90. Schäffer C and Messner P. Emerging facets of prokaryotic glycosylation. *FEMS Microbiol Rev* 2017; 41: 49–91.

91. Tytgat HLP and Lebeer S. The sweet tooth of bacteria: common themes in bacterial glycoconjugates. *Microbiol Mol Biol Rev* 2014; 78: 372–417.
92. Nothaft H and Szymanski CM. Bacterial protein n-glycosylation: new perspectives and applications. *J Biol Chem* 2013; 288: 6912–6920.
93. Lu Q, Li S and Shao F. Sweet talk: protein glycosylation in bacterial interaction with the host. *Trends Microbiol* 2015; 23: 630–641.
94. Latousakis D and Juge N. How sweet are our gut beneficial bacteria? A focus on protein glycosylation in *Lactobacillus*. *Int J Mol Sci* 2018; 19: 1–18.
95. Eichler J and Koomey M. Sweet new roles for protein glycosylation in prokaryotes. *Trends Microbiol* 2017; 25: 662–672.
96. Zhou M, Zhu F, Li Y, et al. GapI functions as a molecular chaperone to stabilize its interactive partner Gap3 during biogenesis of serine-rich repeat bacterial adhesin. *Mol Microbiol* 2012; 83: 866–878.
97. Champasa K, Longwell SA, Eldridge AM, et al. Targeted identification of glycosylated proteins in the gastric pathogen *Helicobacter pylori* (Hp). *Mol Cell Proteomics* 2013; 12: 2568–2586.
98. Schirm M, Arora SK, Verma A, et al. Structural and genetic characterization of glycosylation of type a flagellin in *Pseudomonas aeruginosa*. *J Bacteriol* 2004; 186: 2523–2531.
99. Faridmoayer A, Fentabil MA, Mills DC, et al. Functional characterization of bacterial oligosaccharyltransferases involved in O-linked protein glycosylation. *J Bacteriol* 2007; 189: 8088–8098.
100. Paramonov N, Rangarajan M, Hashim A, et al. Structural analysis of a novel anionic polysaccharide from *Porphyromonas gingivalis* strain W50 related to Arg-gingipain glycans. *Mol Microbiol* 2005; 58: 847–863.
101. Nothaft H, Scott NE, Vinogradov E, et al. Diversity in the protein N-glycosylation pathways within the *Campylobacter* genus. *Mol Cell Proteomics* 2012; 11: 1203–1219.
102. Young NM, Brisson JR, Kelly J, et al. Structure of the N-linked glycan present on multiple glycoproteins in the Gram-negative bacterium, *Campylobacter jejuni*. *J Biol Chem* 2002; 277: 42530–42539.
103. Verma A, Schirm M, Arora SK, et al. Glycosylation of b-type flagellin of *Pseudomonas aeruginosa*: structural and genetic basis. *J Bacteriol* 2006; 188: 4395–4403.
104. Merino S and Tomás JM. Gram-negative flagella glycosylation. *Int J Mol Sci* 2014; 15: 2840–2857.
105. Sun L, Jin M, Ding W, et al. Posttranslational modification of flagellin FlaB in *Shewanella oneidensis*. *J Bacteriol* 2013; 195: 2550–2561.
106. Thibault P, Logan SM, Kelly JF, et al. Identification of the carbohydrate moieties and glycosylation motifs in *Campylobacter jejuni* flagellin. *J Biol Chem* 2001; 276: 34862–34870.
107. Alemka A, Nothaft H, Zheng J, et al. N-glycosylation of *Campylobacter jejuni* surface proteins promotes bacterial fitness. *Infect Immun* 2013; 81: 1674–1682.
108. Iwashkiw JA, Fentabil MA, Faridmoayer A, et al. Exploiting the *Campylobacter jejuni* protein glycosylation system for glycoengineering vaccines and diagnostic tools directed against brucellosis. *Microb Cell Fact* 2012; 11: 13.
109. Valiente E, Bouché L, Hitchen P, et al. Role of glycosyltransferases modifying type B flagellin of emerging hypervirulent *Clostridium difficile* lineages and their impact on motility and biofilm formation. *J Biol Chem* 2016; 291: 25450–25461.

110. Guerry P, Ewing CP, Schirm M, et al. Changes in flagellin glycosylation affect *Campylobacter* autoagglutination and virulence. *Mol Microbiol* 2006; 60: 299–311.
111. Hanuszkiewicz A, Pittock P, Humphries F, et al. Identification of the flagellin glycosylation system in *Burkholderia cenocepacia* and the contribution of glycosylated flagellin to evasion of human innate immune responses. *J Biol Chem* 2014; 289: 19231–19244.
112. Khodai-Kalaki M, Andrade A, Mohamed YF, et al. *Burkholderia cenocepacia* lipopolysaccharide modification and flagellin glycosylation affect virulence but not innate immune recognition in plants. *MBio* 2015; 6: 1–11.
113. Takeuchi K, Taguchi F, Inagaki Y, et al. Flagellin glycosylation island in *Pseudomonas syringae* pv. *Glycinea* and its role in host specificity. *J Bacteriol* 2003; 185: 6658–6665.
114. Zilliges Y, Kehr JC, Mikkat S, et al. An extracellular glycoprotein is implicated in cell-cell contacts in the toxic cyanobacterium *Microcystis aeruginosa* PCC 7806. *J Bacteriol* 2008; 190: 2871–2879.
115. Avilés-Reyes A, Miller JH, Simpson-Haidaris PJ, et al. Modification of *Streptococcus mutans* Cnm by PgfS contributes to adhesion, endothelial cell invasion, and virulence. *J Bacteriol* 2014; 196: 2789–2797.
116. Tang G, Ruiz T and Mintz KP. O-polysaccharide glycosylation is required for stability and function of the collagen adhesin EmaA of *Aggregatibacter actinomycetemcomitans*. *Infect Immun* 2012; 80: 2868–2877.
117. Côté JP, Charbonneau MÉ and Mourez M. Glycosylation of the *Escherichia coli* TibA self-associating autotransporter influences the conformation and the functionality of the protein. *PLoS One* 2013; 8: 1–9.
118. Cuccui J, Terra VS, Bossé JT, et al. The N-linking glycosylation system from *Actinobacillus pleuropneumoniae* is required for adhesion and has potential use in glycoengineering. *Open Biol.* Epub ahead of print January 2017. DOI: 10.1098/rsob.160212.
119. Wu H, Zeng M and Fives-Taylor P. The glycan moieties and the N-terminal polypeptide backbone of a fimbria-associated adhesin, Fap1, play distinct roles in the biofilm development of *Streptococcus parasanguinis*. *Infect Immun* 2007; 75: 2181–2188.
120. Smedley JG, Jewell E, Roguskie J, et al. Influence of pilin glycosylation on *Pseudomonas aeruginosa* 1244 pilus function. *Infect Immun* 2005; 73: 7922–7931.
121. Harvey H, Bondy-Denomy J, Marquis H, et al. *Pseudomonas aeruginosa* defends against phages through type IV pilus glycosylation. *Nat Microbiol* 2018; 3: 47–52.
122. Karlyshev AV, Everest P, Linton D, et al. The *Campylobacter jejuni* general glycosylation system is important for attachment to human epithelial cells and in the colonization of chicks. *Microbiology* 2004; 150: 1957–1964.
123. Sobocinska J, Roszczenko-Jasinska P, Ciesielska A, et al. Protein palmitoylation and its role in bacterial and viral infections. *Front Immunol* 2018; 8: 1–19.
124. Kovacs-Simon A, Titball RW and Michell SL. Lipoproteins of bacterial pathogens. *Infect Immun* 2011; 79: 548–561.
125. Buddelmeijer N. The molecular mechanism of bacterial lipoprotein modification—how, when and why? *FEMS Microbiol Rev* 2015; 39: 246–261.
126. Bürki S, Frey J and Pilo P. Virulence, persistence and dissemination of *Mycoplasma bovis*. *Vet Microbiol* 2015; 179: 15–22.

127. Sankaran K and Wus HC. Lipid modification of bacterial prolipoprotein. Transfer of diacylglycerol moiety from phosphatidylglycerol. *J Biol Chem* 1994; 269: 19701–19706.
128. Hillmann F, Argentini M and Buddelmeijer N. Kinetics and phospholipid specificity of apolipoprotein N-acyltransferase. *J Biol Chem* 2011; 286: 27936–27946.
129. Asanuma M, Kurokawa K, Ichikawa R, et al. Structural evidence of  $\alpha$ -aminoacylated lipoproteins of *Staphylococcus aureus*. *FEBS J* 2011; 278: 716–728.
130. Kurokawa K, Ryu KH, Ichikawa R, et al. Novel bacterial lipoprotein structures conserved in low-GC content gram-positive bacteria are recognized by Toll-like receptor 2. *J Biol Chem* 2012; 287: 13170–13181.
131. Serebryakova MV, Demina IA, Galyamina MA, et al. The acylation state of surface lipoproteins of mollicute *Acholeplasma laidlawii*. *J Biol Chem* 2011; 286: 22769–22776.
132. Kang JY, Nan X, Jin MS, et al. Recognition of lipopeptide patterns by Toll-like receptor 2-Toll-like receptor 6 heterodimer. *Immunity* 2009; 31: 873–884.
133. Das S, Kanamoto T, Ge X, et al. Contribution of lipoproteins and lipoprotein processing to endocarditis virulence in *Streptococcus sanguinis*. *J Bacteriol* 2009; 191: 4166–4179.
134. Bray BA, Sutcliffe IC and Harrington DJ. Impact of lgt mutation on lipoprotein biosynthesis and in vitro phenotypes of *Streptococcus agalactiae*. *Microbiology* 2009; 155: 1451–1458.
135. Petit CM, Brown JR, Ingraham K, et al. Lipid modification of prelipoproteins is dispensable for growth in vitro but essential for virulence in *Streptococcus pneumoniae*. *FEMS Microbiol Lett* 2001; 200: 229–233.
136. Cullen TW and Trent MS. A link between the assembly of flagella and lipooligosaccharide of the Gram-negative bacterium *Campylobacter jejuni*. *Proc Natl Acad Sci USA* 2010; 107: 5160–5165.
137. Reffuveille F, Serron P, Chevalier S, et al. The prolipoprotein diacylglycerol transferase (Lgt) of *Enterococcus faecalis* contributes to virulence. *Microbiology* 2012; 158: 816–825.
138. Stoll H, Dengjel J, Nerz C, et al. *Staphylococcus aureus* deficient in lipidation of prelipoproteins is attenuated in growth and immune activation. *Infect Immun* 2005; 73: 2411–2423.
139. Kurokawa K, Kim MS, Ichikawa R, et al. Environment-mediated accumulation of diacyl lipoproteins over their triacyl counterparts in *Staphylococcus aureus*. *J Bacteriol* 2012; 194: 3299–3306.
140. Kumar S, Balamurali MM and Sankaran K. Bacterial lipid modification of proteins requires appropriate secretory signals even for expression—implications for biogenesis and protein engineering. *Mol Membr Biol* 2014; 31: 183–194.
141. LoVullo ED, Wright LF, Isabella V, et al. Revisiting the Gram-negative lipoprotein paradigm. *J Bacteriol* 2015; 197: 1705–1715.
142. da Silva RAG, Churchward CP, Karlyshev AV, et al. The role of apolipoprotein N-acyl transferase, Lnt, in the lipidation of factor H binding protein of *Neisseria meningitidis* strain MC58 and its potential as a drug target. *Br J Pharmacol* 2017; 174: 2247–2260.
143. Luo Y, Friese OV, Runnels HA, et al. The dual role of lipids of the lipoproteins in trumenba, a self-adjuvanting vaccine against meningococcal meningitis B disease. *AAPS J* 2016; 18: 1562–1575.

144. Machata S, Tchatalbachev S, Mohamed W, et al. Lipoproteins of *Listeria monocytogenes* are critical for virulence and TLR2-mediated immune activation. *J Immunol* 2008; 181: 2028–2035.
145. Schlatterer K, Beck C, Hanzelmann D, et al. The mechanism behind bacterial lipoprotein release: phenol-soluble modulins mediate Toll-like receptor 2 activation via extracellular vesicle release from *Staphylococcus aureus*. *MBio* 2018; 9: 1–13.
146. Issartel J-P, Koronakis V and Hughes C. Activation of *Escherichia coli* prohaemolysin to the mature toxin by acyl carrier protein-dependent fatty acylation. *Nature* 1991; 351: 759–761.
147. Greene NP, Crow A, Hughes C, et al. Structure of a bacterial toxin-activating acyltransferase. *Proc Natl Acad Sci USA* 2015; 112: E3058–E3066.
148. Lee VT and Schneewind O. Protein secretion and the pathogenesis of bacterial infections. *Genes Dev* 2001; 15: 1725–1752.
149. Bouchez V, Douché T, Dazas M, et al. Characterization of post-translational modifications and cytotoxic properties of the adenylate-cyclase hemolysin produced by various *Bordetella pertussis* and *Bordetella parapertussis* isolates. *Toxins (Basel)*. Epub ahead of print September 2017. DOI: 10.3390/toxins9100304.
150. Lewis JD, Wilton M, Mott GA, et al. Immunomodulation by the *Pseudomonas syringae* HopZ type III effector family in *Arabidopsis*. *PLoS One* 2014; 9: 1–20.
151. Thieme F, Szczesny R, Urban A, et al. New type III effectors from *Xanthomonas campestris* pv. *Vesicatoria* trigger plant reactions dependent on a conserved N-myristoylation motif. *Mol Plant Microbe Interact* 2007; 20: 1250–1261.
152. Lewis JD, Abada W, Ma W, et al. The HopZ family of *Pseudomonas syringae* type III effectors require myristoylation for virulence and avirulence functions in *Arabidopsis thaliana*. *J Bacteriol* 2008; 190: 2880–2891.
153. Spera JM, Guaimas F, Corvi MM, et al. *Brucella* hijacks host-mediated palmitoylation to stabilize and localize PrpA to the plasma membrane. *Infect Immun* 2018; 86: 1–8.
154. Ivanov SS, Charron G, Hang HC, et al. Lipidation by the host prenyltransferase machinery facilitates membrane localization of *Legionella pneumophila* effector proteins. *J Biol Chem* 2010; 285: 34686–34698.
155. Reinicke AT, Hutchinson JL, Magee AI, et al. A *Salmonella typhimurium* effector protein SifA is modified by host cell prenylation and S-acylation machinery. *J Biol Chem* 2005; 280: 14620–14627.
156. Schroeder GN, Aurass P, Oates CV, et al. *Legionella pneumophila* effector LpdA is a palmitoylated phospholipase D virulence factor. *Infect Immun* 2015; 83: 3989–4002.
157. Soufi B, Soares NC, Ravikumar V, et al. Proteomics reveals evidence of cross-talk between protein modifications in bacteria: focus on acetylation and phosphorylation. *Curr Opin Microbiol* 2012; 15: 357–363.

## Author biographies

Suzanne Forrest is currently a BBSRC-funded PhD student looking at the diversity and function(s) of post-translational modifications associated with the proteins secreted by *Pseudomonas aeruginosa*. The *P. aeruginosa* secretome plays a key role in pathogenicity, yet little is currently understood about how these post-translational modifications are made, or what their function is.

**Martin Welch** is Reader in Microbial Physiology and Metabolism at the University of Cambridge, and has a long-standing interest in *Pseudomonas aeruginosa* virulence and bio-film formation. He is particularly interested in understanding how metabolism affects *P. aeruginosa* pathogenicity, and how inter-species signalling impinges upon the spectrum of secreted virulence factors.



Other academic contributions during the course of study:

Synthesis and biological evaluation of 1,2-disubstituted 4-quinolone analogues of *Pseudonocardia* sp. natural products

Geddis, S. et al, *Beilstein J. Org. Chem.* **2018**, *14*, 2680–2688

Divergent Synthesis of Novel Cylindrocyclophanes that Inhibit Methicillin-Resistant *Staphylococcus aureus* (MRSA)

Freudenreich, J. et al, *ChemMedChem* **2020**, *15*, 1289



# Synthesis and biological evaluation of 1,2-disubstituted 4-quinolone analogues of *Pseudonocardia* sp. natural products

Stephen M. Geddis<sup>1</sup>, Teodora Coroama<sup>1</sup>, Suzanne Forrest<sup>2</sup>, James T. Hodgkinson<sup>\*3</sup>, Martin Welch<sup>\*2</sup> and David R. Spring<sup>\*1</sup>

## Letter

[Open Access](#)

### Address:

<sup>1</sup>Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, UK, <sup>2</sup>Department of Biochemistry, University of Cambridge, 80 Tennis Road, Cambridge, CB2 1GA, UK and <sup>3</sup>Leicester Institute of Structural and Chemical Biology, and Department of Chemistry, University of Leicester, George Porter Building, University Road, Leicester, LE1 7RH, UK

### Email:

James T. Hodgkinson\* - jthodgkinson@leicester.ac.uk;  
Martin Welch\* - mw240@cam.ac.uk; David R. Spring\* - spring@ch.cam.ac.uk

\* Corresponding author

### Keywords:

antibiotics; cross-coupling; heterocycles; quorum-sensing; structure–activity relationships

*Beilstein J. Org. Chem.* **2018**, *14*, 2680–2688.  
doi:10.3762/bjoc.14.245

Received: 31 July 2018  
Accepted: 10 October 2018  
Published: 19 October 2018

This article is part of the thematic issue "Antibacterials, bacterial small molecule interactions and quorum sensing".

Associate Editor: I. R. Baxendale

© 2018 Geddis et al.; licensee Beilstein-Institut.  
License and terms: see end of document.

## Abstract

A series of analogues of *Pseudonocardia* sp. natural products were synthesized, which have been reported to possess potent anti-bacterial activity against *Helicobacter pylori* and induce growth defects in *Escherichia coli* and *Staphylococcus aureus*. Taking inspiration from a methodology used in our total synthesis of natural products, we applied this methodology to access analogues possessing bulky N-substituents, traditionally considered to be challenging scaffolds. Screening of the library provided valuable insights into the structure–activity relationship of the bacterial growth defects, and suggested that selectivity between bacterial species should be attainable. Furthermore, a structurally related series of analogues was observed to inhibit production of the virulence factor pyocyanin in the human pathogen *Pseudomonas aeruginosa*, which may be a result of their similarity to the *Pseudomonas* quinolone signal (PQS) quorum sensing autoinducer. This provided new insights regarding the effect of N-substitution in PQS analogues, which has been hitherto underexplored.

## Introduction

The quinolone core has long been implemented in structures possessing formidable activity in a broad range of fields, including antibiotics, bacterial signalling and iron metabolism [1].

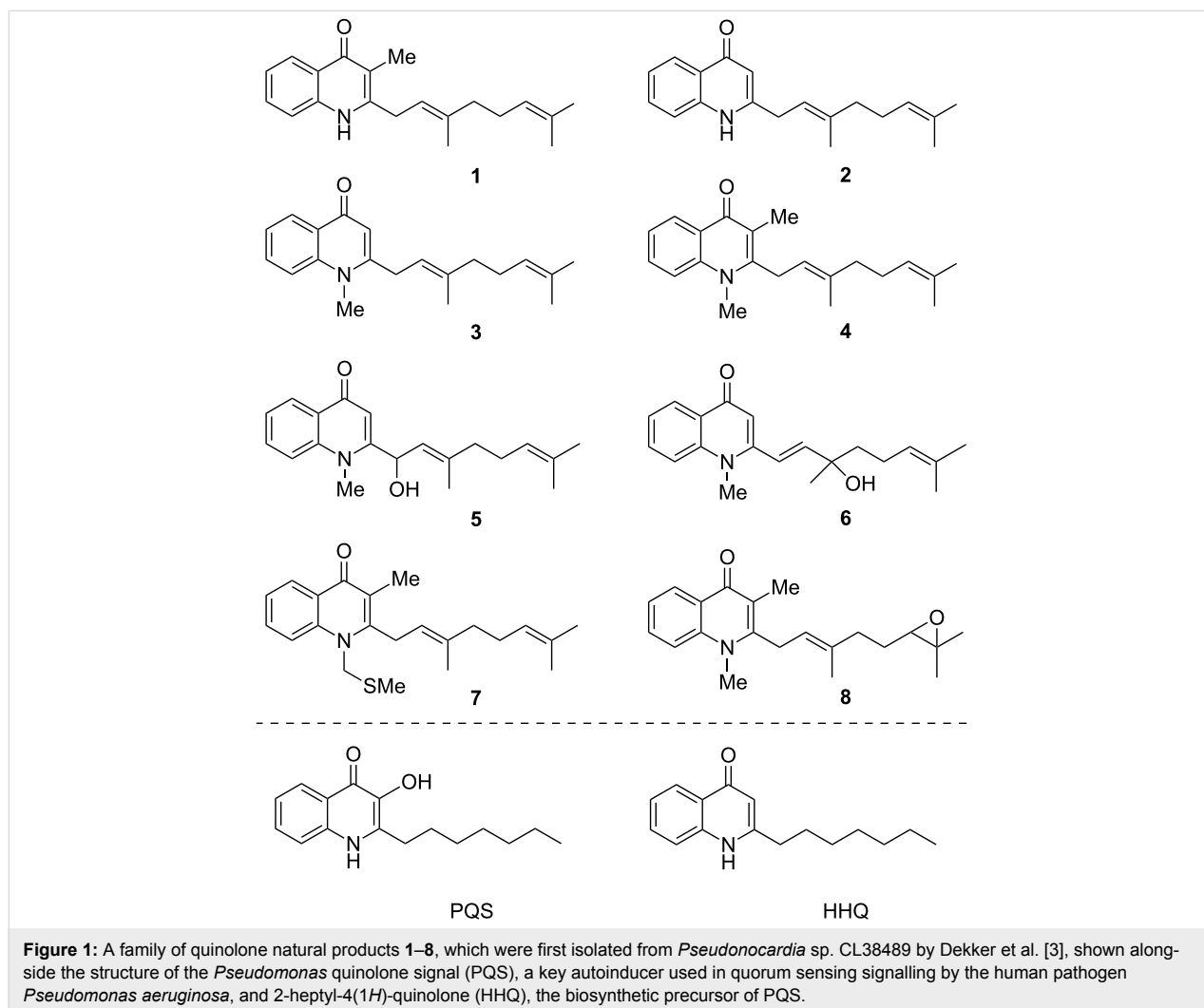
Structural optimisation of quinolones possessing intriguing properties can lead to the discovery of important drug classes, as demonstrated by the fluoroquinolone antibiotics, which were

inspired by the observation of an antibacterial quinolone side-product generated during the synthesis of the antimalarial chloroquine [2].

Given this high potential for the discovery of useful chemical entities, we have recently been engaged in research regarding a family of quinolone natural products which are produced by the actinomycete *Pseudonocardia* sp. CL38489, and were first isolated by Dekker et al. (**1–8**, Figure 1) [3]. The authors noted the potent antibacterial activity of these compounds against *Helicobacter pylori*, which is responsible for the generation of numerous digestive disorders [4]. Furthermore, with the presence of a lipophilic chain in the 2-position, there is a structural resemblance to the *Pseudomonas* quinolone signal (PQS), and its biosynthetic precursor 2-heptyl-4(1*H*)-quinolone (HHQ), which are vital to the cooperative behaviour of the human pathogen *Pseudomonas aeruginosa* via quorum sensing (QS). This is a means by which bacteria alter their phenotype in response to changes in population density, regulating virulence

and biofilm formation when most impactful to the host organism [5]. This process is mediated by signalling molecules such as PQS, and natural product structures **1–8** analogous to PQS may provide interspecies QS-modulator chemical probes. It has been proposed that such a strategy may perturb bacterial virulence and pathogenicity associated with QS, thus conferring a therapeutic benefit, without applying a selection pressure for resistance [6]. Whilst recent experiments suggest that resistance may still emerge, it has been proposed that this development should be limited under certain conditions [7].

We wished to investigate the potential of **1–8** to modulate QS in *P. aeruginosa*, however, the compounds are available in only trace amounts from natural sources [3], and so we embarked on the total synthesis of the compounds. We first developed a strategy which constructed natural products **1–4** by uniting the quinolone cores with the side chain by means of an  $sp^2$ – $sp^3$  Suzuki–Miyaura coupling reaction [8]. Whilst these compounds unfortunately provided no modulation of PQS quorum



screening (as determined using a heterologous *Escherichia coli* reporter system [9]), an intriguing effect upon the growth of *E. coli* and *Staphylococcus aureus* was noted, which showed an extended lag phase in response to the compounds (except **4**, which was inactive towards *E. coli*). It should be noted that in this previous publication, the graphical data for compounds **3** and **4** was erroneously switched). It is tentatively proposed that this is as a result of disruption of electron transport, as the compounds bear resemblance to the menaquinones which are used by bacteria for this purpose [10]. Following on from this, we recently reported a divergent strategy which granted access to remainder of the natural products **5–8**, alongside offering more efficient synthesis of **1** and **4** [11]. Allylic alcohols **5** and **6** were accessed from a mutual precursor (constructed using methodology adapted from that reported by Bernini et al. [12]) using an acid-catalysed transposition, whilst **4**, **7** and **8** were derivatised from **1**.

In this current work, we turn to the further elucidation of the biological activity of this class of compounds. In order to gain additional insight into the associated structure–activity relationships (SAR), it was desired to generate analogues of the natural products. The chemistry developed towards the allylic alcohols **5** and **6**, outlined in Scheme 1, seemed ideal to this end. A range of alkynes **10** could undergo Sonogashira coupling with the commercially available acid chloride **9**. The resultant ynones **11** could then undergo conjugate addition with primary amines **12**, which following metal-catalysed cyclisation would give 1,2-disubstituted quinolones **14**.

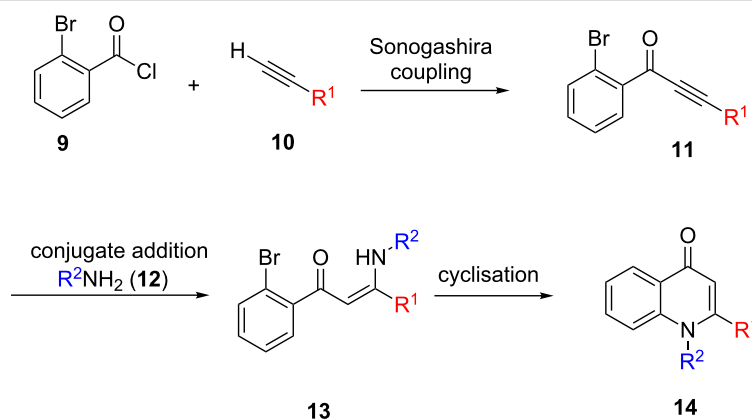
Upon the successful synthesis of analogues of the form **14**, biological evaluation of these and the natural products **1**, **4** and **5–7** (synthesised during our previous study [11]) would then be possible. In particular, it was desired to further probe the intriguing growth defects which had been observed for natural

products **1–4** in *E. coli* and *S. aureus*. Furthermore, exploration of any effect on QS of the analogues **14** would be valuable, as to our knowledge studies on the SAR of PQS analogues have not yet thoroughly assessed substitution at the 1-position of the quinolone system [13]. This is perhaps as a result of direct alkylation at this position being very challenging, with low yields and poor O- vs N-selectivity being encountered, particularly with a sterically demanding substituent present in the 2-position [14,15].

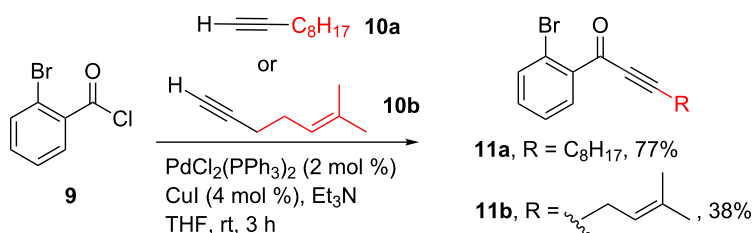
As a measure of modulation of QS in *P. aeruginosa*, it was desired to measure the amount of pyocyanin produced by bacterial cultures after treatment with the compounds. This virulence factor is known to be under the regulation of PQS signalling system, and is capable of disrupting many important biochemical processes [16]. This leads to numerous deleterious effects on human cells, including inhibited respiration and ciliary action [17]. These effects allow pyocyanin to play a critical role in infection; indeed, mutant *P. aeruginosa* strains which are unable to produce pyocyanin have been shown to be unsuccessful in infecting the lungs of mice [18]. Being able to prevent the production of pyocyanin could therefore be of great therapeutic benefit.

## Results and Discussion

In the implementation of the strategy outlined in Scheme 1, alkynes **10a** and **10b** were first subjected to Sonogashira coupling with commercially available acid chloride **9** according to the previously reported conditions (Scheme 2) [12]. These alkynes were chosen so as to allow access to valuable SAR data regarding the side chain of the natural products **1–8**. Commercially available alkyne **10a** would ultimately lead to a simple saturated side chain of the same length as that observed naturally, whilst **10b** (itself synthesised according to a literature procedure [19]) would provide analogues possessing a truncated



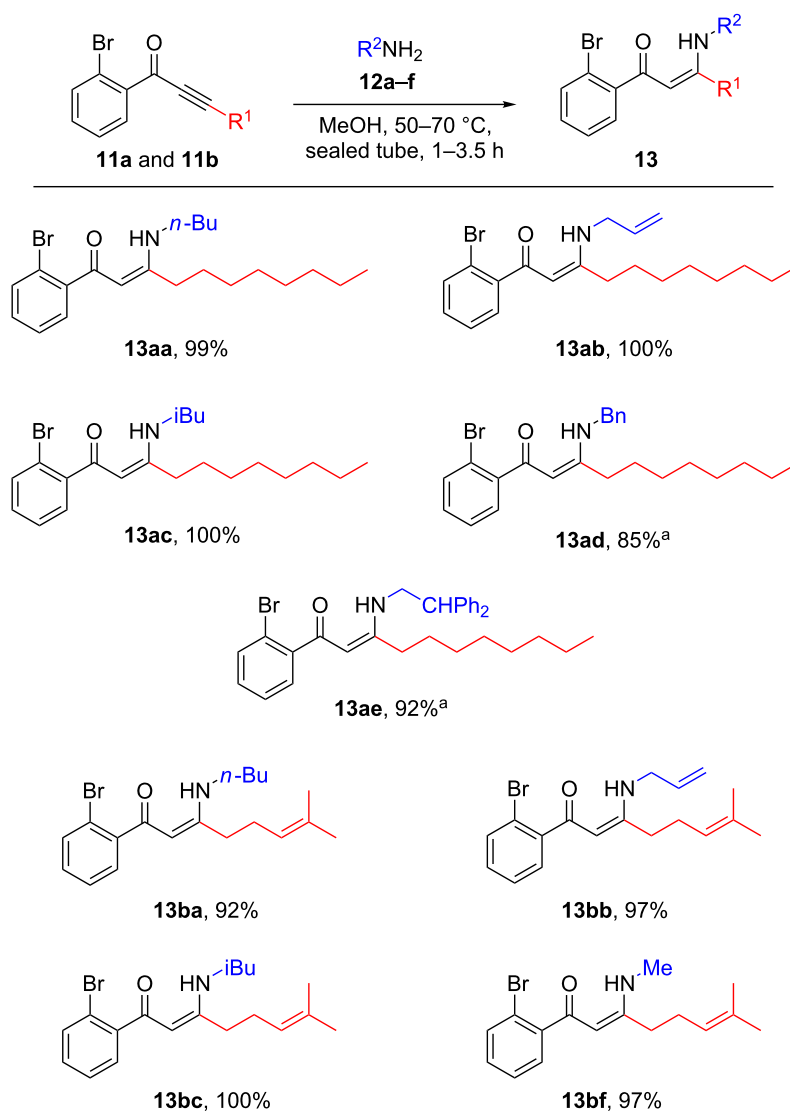
**Scheme 1:** Proposed use of the chemistry developed towards the total synthesis of **5** and **6** for generation of natural product analogues. Modular coupling of alkynes **10** and amines **12** with commercially available acid chloride **9** was proposed to give 1,2-disubstituted quinolones **14**.



**Scheme 2:** Sonagashira coupling of alkynes **10a** and **10b** with commercially available acid chloride **9** to give ynones **11a** and **11b**.

prenyl-type substituent. In the event, ynone **11a** was obtained with good yield, however, a poorer yield resulted for **11b**, which was attributed to difficulties in obtaining its precursor **10b** with high purity which stemmed from its volatility.

These ynones were then subjected to a conjugate addition with an assortment of primary amines **12a–f** (Scheme 3). The reactions proceeded with excellent yield in all cases, with aliphatic and aromatic moieties well tolerated. Given the high volatility



**Scheme 3:** Conjugate addition of primary amines **12a–f** with ynones **11a** and **11b**. <sup>a</sup>Following concentration in vacuo, further purification using silica gel flash chromatography was required.

of most of the amine starting materials, the products **13** were in general analytically pure following concentration in vacuo of the reaction mixture. However, use of higher boiling-point amines necessitated purification by flash column chromatography, which may account for the slightly lower yields in these cases (**13ad** and **13ae**).

With the compounds **13** now in hand, their cyclisation to the desired analogues **14** was explored. However, whilst the conditions which had proved successful in the total synthesis of natural products **5** and **6** proved satisfactory in most cases, some optimisation was required for substrates possessing unsaturated functionality attached to the amine (Table 1). When the palladium-catalysed conditions were employed [20], a complex mixture resulted, from which no product could be obtained (Table 1, entry 1). Meanwhile, use of base-induced  $S_NAr$ -type conditions allowed a small amount of product to be isolated (Table 1, entry 2) [21], but copper-catalysed conditions offered a higher yield (Table 1, entry 3) [12]. This behaviour stands in contrast to that noted for substrates bearing an alkyl substituent in our previous study, for which these copper-catalysed conditions resulted in dimerization [11].

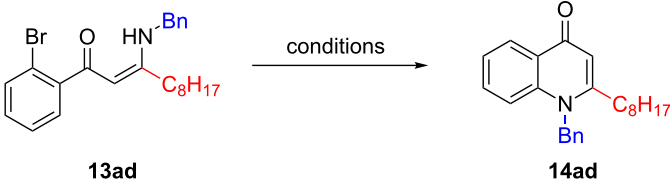
Following the discovery of this substrate-dependent dichotomy with respect to optimal reaction conditions, the entire library of compounds was successfully cyclised. Whilst the yields ranged from low to moderate (Scheme 4), sufficient quantities were obtained to facilitate biological screening. It appeared that bulkier N-substituents (e.g., **14ae**) resulted in lower yields than less bulky derivatives (e.g., **14bf**), underlining the importance of steric factors during cyclisation. Interestingly, the allyl-substituted substrates **13ab** and **13bb** underwent an isomerisation under the reaction conditions, with the double bond moving into conjugation with the amine to give inseparable mixtures of en-

amine-type products **14ab** and **14bb**. Given the likely hydrolytic instability of synthetic precursors possessing an enamine moiety, these compounds would likely be highly challenging to synthesise by other means. However, it is proposed that the involvement of the nitrogen lone-pair in the aromaticity of the quinolone system attenuates the susceptibility of **14ab** and **14bb** towards hydrolysis.

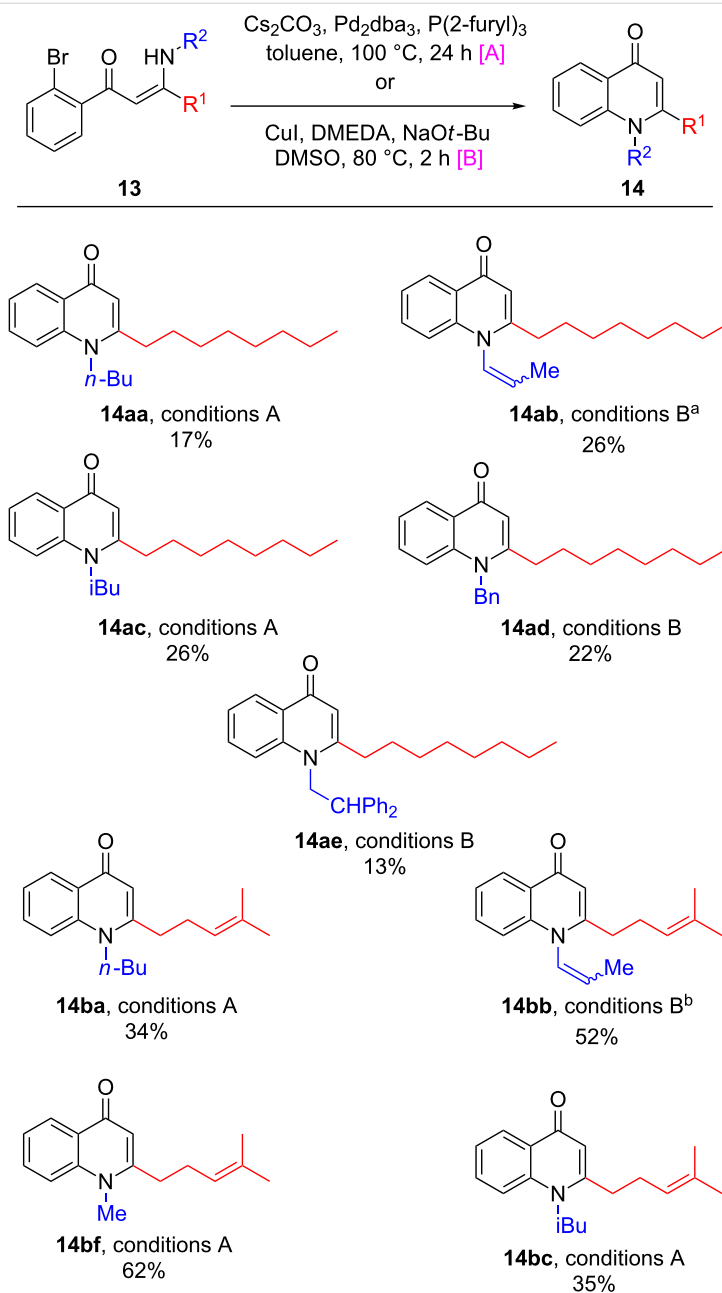
Intriguingly, the employment of an excess of DMEDA in the Cu-catalysed cyclisation of **13bb** generated **14bg** as a side-product, which represents another interesting analogue for biological study (Scheme 5). This may putatively result from the displacement of the allylamine in **13bb** by the DMEDA ligand, followed by heterocyclisation with concomitant N→N' methyl transfer.

With the library of natural products and analogues in hand, our attention turned to their biological activity. It was desired to further explore the growth defects which had been previously noted for natural products **1–4** against *E. coli* and *S. aureus*, and so these species were grown in the presence of the compounds. The results for *E. coli* ESS are shown in Figure 2, split into the natural product series, the series of analogues with a saturated side chain, and the truncated series of analogues. As can be seen in Figure 2A, natural product **1** resulted in slowed bacterial growth whilst **4** elicited no such effect, consistent with our previous observations (although the later recovery in population in the presence of **4** was less pronounced in the present case) [8]. Meanwhile, **5** appeared to show a moderate growth-slowing effect, which when compared to the stronger effect previously observed for **3**, demonstrates that oxidation of the geranyl side chain is deleterious to the biological effect under investigation. The regioisomeric **6** showed a very small effect, further showing the lack of tolerance of the effect towards side-

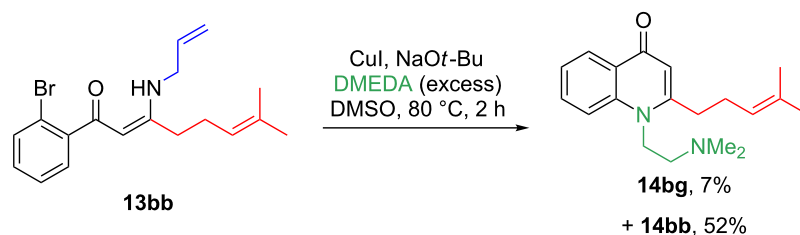
**Table 1:** Optimisation of conditions for the cyclisation of **13ad** to natural product analogue **14ad**.

		
entry	conditions	result
1	$\text{Cs}_2\text{CO}_3$ , $\text{Pd}_2\text{dba}_3$ , $\text{P}(\text{2-furyl})_3$ , toluene, 100 °C, 24 h [20]	complex mixture <sup>a</sup>
2	$\text{KOt-Bu}$ , dioxane, 90 °C, 24 h [21]	13% <sup>b</sup>
3	$\text{CuI}$ , DMEDA, $\text{NaOt-Bu}$ , DMSO, 80 °C, 2 h [12]	22% <sup>b</sup>

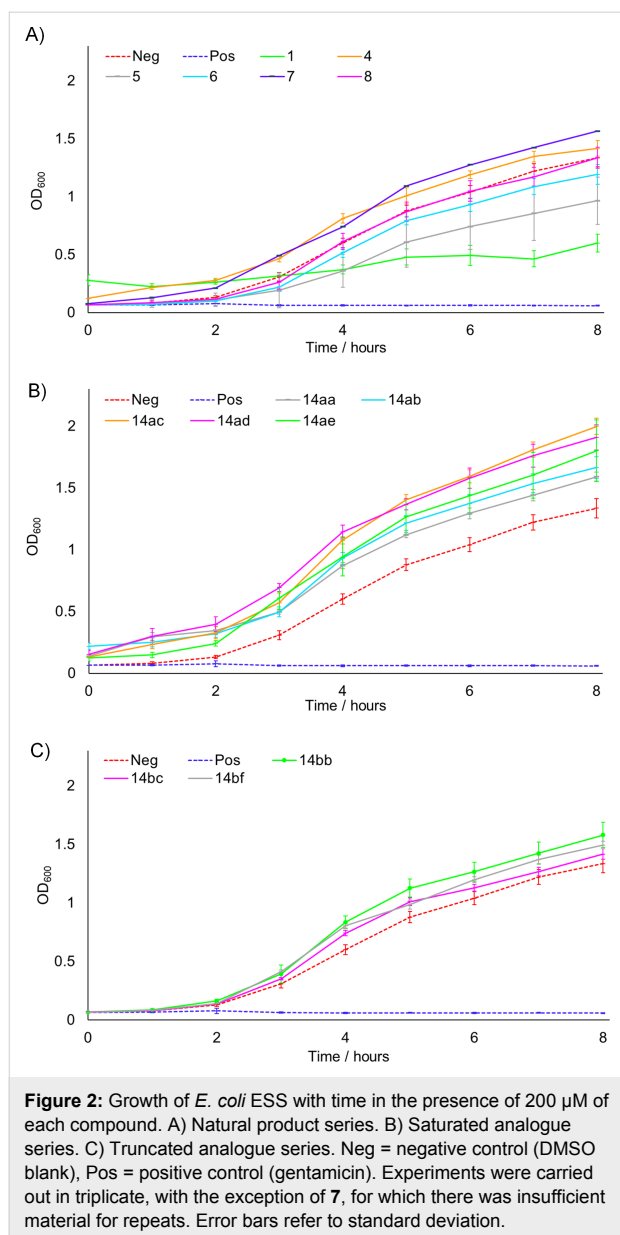
<sup>a</sup>As determined by LCMS and <sup>1</sup>H NMR analysis of the crude reaction product. <sup>b</sup>Isolated yield.



**Scheme 4:** Cyclisation of precursors **13** to natural product analogues **14** using palladium- or copper-catalysed conditions. Yields quoted are isolated. <sup>a</sup>**13ab** used as starting material, *E/Z* ratio = 29:71 based on <sup>1</sup>H NMR data. <sup>b</sup>**13bb** used as starting material, *E/Z* ratio = 15:85 based on <sup>1</sup>H NMR data.

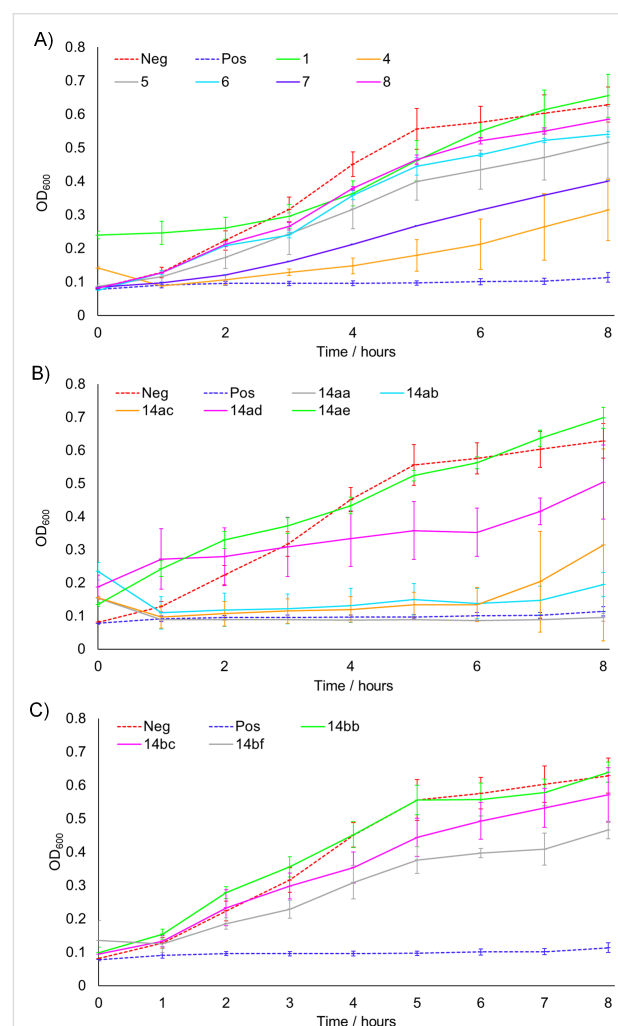


**Scheme 5:** Use of an excess of  $\text{DMEDA}$  in the Cu-catalysed cyclisation of **13bb** resulted in the generation of **14bg** as a side-product, alongside **14bb**.



chain oxidation. Finally, neither replacement of the N-Me of **4** with the methylthiomethylene substituent of **7**, nor epoxidation of **4**'s side chain to give **8**, offered any improvement in the biological activity. The result for **8** is particularly intriguing, as this natural product was noted to have the strongest effect upon the growth of *H. pylori* in the study by Dekker et al., which may imply that these compounds are acting through different mechanisms upon each bacterial species. Next, considering Figure 2B and C, we see that none of the analogues were capable of affecting the growth of *E. coli*, further demonstrating the importance of the geranyl side for biological activity against this species. This observation is particularly striking for analogue **14bf**, which possesses an identical quinolone core structure to natural product **3**, which was previously shown to be active.

Meanwhile, Figure 3 shows the data for *S. aureus* 25923, which is split into the same three series as before. Considering the natural products (Figure 3A), we see that both **1** and **4** resulted in a slowing of growth consistent with that reported previously, although the effect for **1** was less pronounced in the present case, operating for only the first three hours [8]. Whilst most of the other natural products appeared to show only slight effects, moderate activity was observed for **7**, which stands in an interesting contrast to the inactivity of this compound against *E. coli*. This implies that the structural requirements for optimal activity differ between the species, a conclusion which is further bolstered by the results for the saturated series of analogues shown in Figure 3B. Whilst these compounds were completely inactive against *E. coli*, in this case a strong effect was observed, which for **14aa**, **14ab** and **14ac** was comparable to the





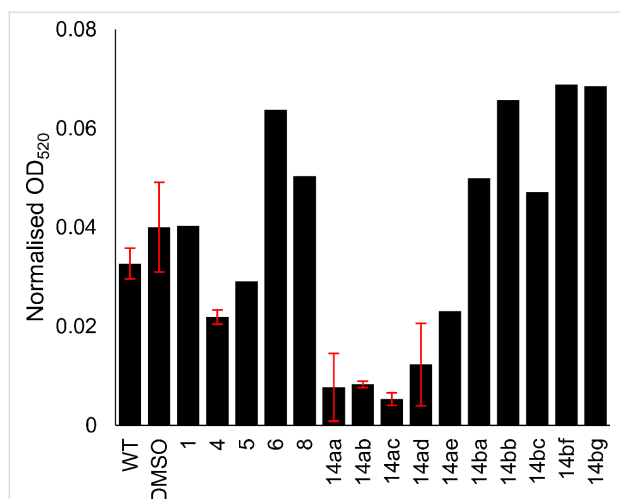
positive gentamicin control over the timescale of concern. These results show the high efficacy of the saturated side chain against *S. aureus*, however, considering the data for **14ad** and **14ae**, we can observe that adding bulky aromatic moieties to the N-substituent results in reduced activity, with smaller alkyl groups in this position instead being optimal. Finally, the truncated series of analogues appeared to show only small effects upon the growth of *S. aureus* (Figure 3C).

Attention then turned to the ability of the compounds to modulate *P. aeruginosa* PA01 QS, as measured by the production of the virulence factor pyocyanin. Bacterial cultures were grown for eight hours in the presence of each compound, followed by extraction of the pyocyanin under acidic conditions [22]. This was then quantified by measurement of the OD<sub>520</sub>, which corresponds to absorption by the toxin. The results are shown in Figure 4, normalised by the bacterial density as measured by OD<sub>600</sub> (no significant effect on the overall growth of the bacteria was observed for any of the compounds, see Supporting Information File 1 for details). The most promising results were replicated to ensure validity (due to the large amount of chemicals required for the assay, it was not practical to perform this for every compound). Of the natural products, only **4** seemed to show attenuation of pyocyanin production relative to the negative control. Meanwhile, whilst the truncated series of analogues appeared to lack activity, a marked reduction in pyocyanin production was noted for the compounds possessing a linear octan-1-yl side chain. We speculate that this is due to the similarity to the heptan-1-yl chain present in the native

PQS ligand. This observation is highly intriguing, as HHQ (Figure 1), which differs from these analogues only by the lack of an N-substituent and a slightly shortened side chain, is known to activate *P. aeruginosa* QS (although it is incapable of inducing full pyocyanin expression) [23]. It would therefore appear that this inhibitory activity is likely due to the N-substitution, an avenue which has been to our knowledge underexplored in the SAR analysis of PQS and HHQ. In particular, analogues possessing smaller N-substituents (**14aa**, **14ab** and **14ac**) appeared to elicit a stronger effect than those possessing larger aromatic moieties (**14ad** and **14ae**).

## Conclusion

We have reported the synthesis of structural analogues of a family of 4-quinolone *Pseudonocardia* sp. natural products, which encompassed variation of both the side chain and N-substituent. This represented an extension of the chemistry which we employed towards the natural products, utilising sequential Sonogashira coupling, high-yielding conjugate addition, and metal-catalysed cyclisation. A dichotomy in the optimal conditions for the final step was discovered, depending on the nature of the N-substituent. These analogues were then tested, together with a number of previously synthesised natural products, for their ability to bring about an intriguing “growth-slowing” effect towards certain species of bacteria. Whilst the presence of a geranyl side chain was shown to be vital for strong activity in *E. coli*, analogues possessing a saturated side chain exhibited marked inhibition of *S. aureus* growth. This intriguing result implies that slightly different mechanisms may be at work in each case, and suggests that it may be possible to attain selective therapeutic treatment of a specific species. Furthermore, the saturated series of analogues were demonstrated to inhibit the production of pyocyanin by *P. aeruginosa*, a virulence factor known to be under QS regulation, providing valuable new SAR insights regarding N-substitution of PQS and HHQ analogues.



**Figure 4:** OD<sub>520</sub> (absorption corresponding to pyocyanin) normalised by the culture population (measured by OD<sub>600</sub>) for cultures of *P. aeruginosa* PA01 grown in the presence of concentrations of 200 μM of natural products and analogues after 8 hours. WT = wild type, no treatment added. DMSO = treated with DMSO blank. The experiment for **7** was not performed due to lack of material. The most promising results were replicated once to ensure validity, as shown by error bars.

## Supporting Information

### Supporting Information File 1

Experimental procedures and analytical data.

[<https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-14-245-S1.pdf>]

## Acknowledgements

SF was supported by a BBSRC studentship. DRS acknowledges support from the Engineering and Physical Sciences Research Council (EP/P020291/1) and Royal Society (Wolfson Research Merit Award). Data accessibility: all data supporting this study are provided as Supporting Information accompanying this paper.

## ORCID® iDs

Suzanne Forrest - <http://orcid.org/0000-0002-9605-7962>

Martin Welch - <http://orcid.org/0000-0003-3646-1733>

David R. Spring - <http://orcid.org/0000-0001-7355-2824>

## References

- Heeb, S.; Fletcher, M. P.; Chhabra, S. R.; Diggle, S. P.; Williams, P.; Cámara, M. *FEMS Microbiol. Rev.* **2011**, *35*, 247–274. doi:10.1111/j.1574-6976.2010.00247.x
- Bisacchi, G. S. *J. Med. Chem.* **2015**, *58*, 4874–4882. doi:10.1021/jm501881c
- Dekker, K. A.; Inagaki, T.; Gootz, T. D.; Huang, L. H.; Kojima, Y.; Kohlbrenner, W. E.; Matsunaga, Y.; McGuirk, P. R.; Nomura, E.; Sakakibara, T.; Sakemi, S.; Suzuki, Y.; Yamauchi, Y.; Kojima, N. *J. Antibiot.* **1998**, *51*, 145–152. doi:10.7164/antibiotics.51.145
- Rho, T. C.; Bae, E.-A.; Kim, D.-H.; Oh, W. K.; Kim, B. Y.; Ahn, J. S.; Lee, H. S. *Biol. Pharm. Bull.* **1999**, *22*, 1141–1143. doi:10.1248/bpb.22.1141
- Whiteley, M.; Diggle, S. P.; Greenberg, E. P. *Nature* **2017**, *551*, 313–320. doi:10.1038/nature24624
- Clatworthy, A. E.; Pierson, E.; Hung, D. T. *Nat. Chem. Biol.* **2007**, *3*, 541–548. doi:10.1038/nchembio.2007.24
- Allen, R. C.; Popat, R.; Diggle, S. P.; Brown, S. P. *Nat. Rev. Microbiol.* **2014**, *12*, 300–308. doi:10.1038/nrmicro3232
- Salvaggio, F.; Hodgkinson, J. T.; Carro, L.; Geddis, S. M.; Galloway, W. R. J. D.; Welch, M.; Spring, D. R. *Eur. J. Org. Chem.* **2016**, 434–437. doi:10.1002/ejoc.201501400
- Cugini, C.; Calfee, M. W.; Farrow, J. M., III; Morales, D. K.; Pesci, E. C.; Hogan, D. A. *Mol. Microbiol.* **2007**, *65*, 896–906. doi:10.1111/j.1365-2958.2007.05840.x
- Kurosu, M.; Begari, E. *Molecules* **2010**, *15*, 1531–1553. doi:10.3390/molecules15031531
- Geddis, S. M.; Carro, L.; Hodgkinson, J. T.; Spring, D. R. *Eur. J. Org. Chem.* **2016**, 5799–5802. doi:10.1002/ejoc.201601195
- Bernini, R.; Cacchi, S.; Fabrizi, G.; Sferrazza, A. *Synthesis* **2009**, 1209–1219. doi:10.1055/s-0028-1087990
- Ó Muimhneacháin, E.; Reen, F. J.; O’Gara, F.; McGlacken, G. P. *Org. Biomol. Chem.* **2018**, *16*, 169–179. doi:10.1039/C7OB02395B
- Abe, H.; Kawada, M.; Inoue, H.; Ohba, S.-i.; Nomoto, A.; Watanabe, T.; Shibasaki, M. *Org. Lett.* **2013**, *15*, 2124–2127. doi:10.1021/ol400587a
- Mehra, M. K.; Tantak, M. P.; Arun, V.; Kumar, I.; Kumar, D. *Org. Biomol. Chem.* **2017**, *15*, 4956–4961. doi:10.1039/C7OB00940B
- Lau, G. W.; Hassett, D. J.; Ran, H.; Kong, F. *Trends Mol. Med.* **2004**, *10*, 599–606. doi:10.1016/j.molmed.2004.10.002
- Sorensen, R. U.; Klinger, J. D. *Basic Research and Clinical Aspects of Pseudomonas Aeruginosa. International Symposium on Pseudomonas Aeruginosa*; 1987; Vol. 39, pp 113–124.
- Lau, G. W.; Ran, H.; Kong, F.; Hassett, D. J.; Mavrodi, D. *Infect. Immun.* **2004**, *72*, 4275–4278. doi:10.1128/IAI.72.7.4275-4278.2004
- Smith, W. N.; Beumel, O. F., Jr. *Synthesis* **1974**, 441–443. doi:10.1055/s-1974-23341
- Wolfe, J. P.; Rennels, R. A.; Buchwald, S. L. *Tetrahedron* **1996**, *52*, 7525–7546. doi:10.1016/0040-4020(96)00266-9
- Janni, M.; Arora, S.; Peruncheralathan, S. *Org. Biomol. Chem.* **2016**, *14*, 8781–8788. doi:10.1039/C6OB01568A
- Frank, L. H.; DeMoss, R. D. *J. Bacteriol.* **1959**, *77*, 776–782.
- Xiao, G.; Déziel, E.; He, J.; Lépine, F.; Lesic, B.; Castonguay, M.-H.; Milot, S.; Tampakaki, A. P.; Stachel, S. E.; Rahme, L. G. *Mol. Microbiol.* **2006**, *62*, 1689–1699. doi:10.1111/j.1365-2958.2006.05462.x

## License and Terms

This is an Open Access article under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>). Please note that the reuse, redistribution and reproduction in particular requires that the authors and source are credited.

The license is subject to the *Beilstein Journal of Organic Chemistry* terms and conditions: (<https://www.beilstein-journals.org/bjoc>)

The definitive version of this article is the electronic one which can be found at: [doi:10.3762/bjoc.14.245](https://doi.org/10.3762/bjoc.14.245)

# Divergent Synthesis of Novel Cyliindrocyclophanes that Inhibit Methicillin-Resistant *Staphylococcus aureus* (MRSA)

Julien J. Freudenreich,<sup>[a]</sup> Sean Bartlett,<sup>[a]</sup> Naomi S. Robertson,<sup>[a]</sup> Sarah L. Kidd,<sup>[a]</sup> Suzie Forrest,<sup>[b]</sup> Hannah F. Sore,<sup>[a]</sup> Warren R. J. D. Galloway,<sup>[a]</sup> Martin Welch,<sup>[b]</sup> and David R. Spring<sup>\*[a]</sup>

The cyliindrocyclophanes are a family of macrocyclic natural products reported to exhibit antibacterial activity. Little is known about the structural basis of this activity due to the challenges associated with their synthesis or isolation. We hypothesised that structural modification of the cyliindrocyclophane scaffold could streamline their synthesis without significant loss of activity. Herein, we report a divergent synthesis of the cyliindrocyclophane core enabling access to symmetrical

macrocycles by means of a catalytic, domino cross-metathesis-ring-closing metathesis cascade, followed by late-stage diversification. Phenotypic screening identified several novel inhibitors of methicillin-resistant *Staphylococcus aureus*. The most potent inhibitor has a unique tetrabrominated [7,7]paracyclophane core with no known counterpart in nature. Together these illustrate the potential of divergent synthesis using catalysis and unbiased screening methods in modern antibacterial discovery.

## Introduction

*Staphylococcus aureus* is a serious cause of community- and healthcare-associated infection worldwide.<sup>[1]</sup> A particular health burden is the treatment of methicillin-resistant *S. aureus* (MRSA) infection, which is associated with a significant increase in mortality and long-term patient care.<sup>[2]</sup> As such, the World Health Organization has recently designated MRSA as a high-priority pathogen for focused antibacterial research and development.<sup>[3]</sup>

New antibiotics are needed just to keep up with the spread of resistance, but this need is not being met by the development pipeline.<sup>[4]</sup> For decades, pharmaceutical companies have struggled with the complexities of bringing novel antibiotics to market.<sup>[5,6]</sup> Accordingly, most antibiotics available today are derivatives of older antibiotics that have since been phased out. This commonality limits the lifespan of new treatments before cross-resistance renders them ineffective.<sup>[7]</sup>

In an attempt to break this deadlock, recent years have seen growing interest in the exploration of new antibacterial scaffolds and targets in screening.<sup>[8]</sup> In particular, we and others have sought to make use of divergent synthesis to identify


novel antibacterial leads for drug development.<sup>[9–11]</sup> The cyliindrocyclophanes are a family of macrocyclic natural products isolated from marine and terrestrial cyanobacteria.<sup>[12–14]</sup> They are structurally related to the corresponding carbamido-, nosto- and merocyclophanes, which share a common [7,7]paracyclophane backbone but vary in  $\alpha$ -,  $\beta$ - and peripheral substitution patterns and oxidation level (Figure 1).<sup>[15–20]</sup> For an excellent review on alkylresorcinols such as cyliindrocyclophanes, see Martins et al.<sup>[21]</sup>


The biochemical and chemical synthesis of cyclophane natural products has interested and occupied chemists for decades.<sup>[22–31]</sup> Several reports describe the antibacterial activities of related carbamidocyclophane natural products; however, the cyliindrocyclophanes have been subject to rather less attention in this regard. To our knowledge, all studies to date describing the antibacterial evaluation of the cyliindrocyclophane family are restricted to naturally occurring [7,7]paracyclophanes of which 16 members have been identified.<sup>[32,33]</sup> This limits the chemical diversity and hence scope of any such investigation, meaning that little is known about the structure–activity relationships of these compounds or their derivatives. The cyliindrofridins (linear congeners of the cyliindrocyclophanes) display reduced activity against MRSA and *Streptococcus pneumoniae*, thus suggesting that cyclisation augments the antibacterial activity of this scaffold.<sup>[32]</sup> The cyliindrocyclophane  $\alpha$ -OH motif is not required for activity against *Mycobacterium tuberculosis*, although  $\alpha$ -acetylated cyliindrocyclophanes display reduced activity against MRSA.<sup>[33]</sup>

These observations prompted us to question which structural motifs might be responsible for the antibacterial activity of the cyliindrocyclophanes. We thought it possible that we could design cyliindrocyclophane analogues with streamlined syntheses that retain the antibacterial activity of the parent scaffold. If so, this would significantly reduce the effort needed to synthesise and study this family of compounds. As such, we sought to develop of a chemical synthesis of cyliindrocyclo-

[a] Dr. J. J. Freudenreich, Dr. S. Bartlett, Dr. N. S. Robertson, Dr. S. L. Kidd, Dr. H. F. Sore, Dr. W. R. J. D. Galloway, Prof. D. R. Spring  
Department of Chemistry  
University of Cambridge  
Lensfield Road, Cambridge CB2 1EW (UK)  
E-mail: spring@ch.cam.ac.uk

[b] Dr. S. Forrest, Dr. M. Welch  
Department of Biochemistry  
University of Cambridge  
Downing Site, Cambridge CB2 1QW (UK)

 Supporting information for this article is available on the WWW under <https://doi.org/10.1002/cmdc.202000179>

 © 2020 The Authors. Published by Wiley-VCH GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

phane scaffolds to enable the exploration of the cylindrocyclophanes as novel antibacterials. Herein, we report the divergent synthesis of a collection of novel cylindrocyclophanes varying in both ring architecture and functionalisation around the core. Phenotypic screening was performed using the library of novel cylindrocyclophane analogues against a panel of common clinical pathogens. The results of the phenotypic screening and implications of these findings for the discovery of novel antibacterials using divergent synthesis are discussed.

## Results

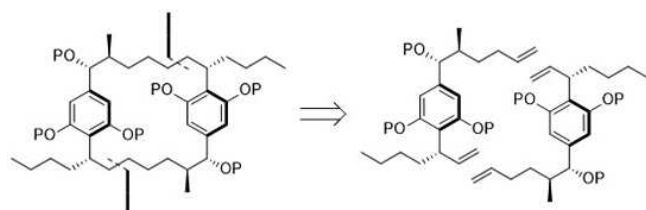
We sought an efficient route towards a simplified cylindrocyclophane core. We envisaged a key disconnection of the cylindrocyclophane scaffold into symmetrical monomers. We hoped to make use of a domino cross-metathesis-ring-closing metathesis cascade as the pivotal ring-forming step based on recent syntheses that have employed similar head-to-tail cyclodimerisations to good effect.<sup>[26–30,34]</sup> From here, we felt that this common intermediate would be accessible from commercial materials in just a few steps (Scheme 1). This head-to-tail cyclodimerisation would enable the study of a homologous series of symmetrical  $[m,n]$ cylindrocyclophanes through variation of the chain length installed during synthesis. We hoped to use this route to investigate the role of the resorcinol core by testing protected derivatives, as well as the role of the substituent in the  $\alpha$ -position, further unsaturation and other late-stage modifications of the parent scaffold.

Gratifyingly, the synthesis of cylindrocyclophanes **1a–c** could be achieved effectively in this way (Scheme 2). Manipulation of acid **2** provided the Suzuki substrate **3**, which was

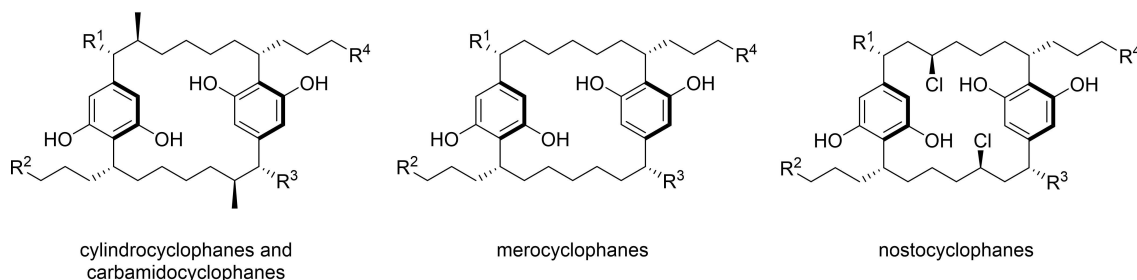
coupled with the allylic boronate ester to provide Weinreb amide **4** in good yield to serve as the branching point in our synthesis. Reaction with the  $n$ -alkenyl Grignards yielded compounds **5a–c**, which constitute a homologous series of acyclic precursors varying in alkenyl chain length. Subsequent deprotection and acetylation of **5a–c** yielded the requisite precursors **6a–c** for our domino cross-metathesis-ring-closing metathesis end-game (Scheme 3).

Treatment of **6a–c** with Grubbs' catalyst, followed by hydrolysis of the acetyl protecting group promoted the desired regioselective cyclodimerisation to construct the  $[m,n]$ paracyclophane scaffolds **7a–c**. Predominant head-to-tail cyclodimerisation was confirmed by single-crystal X-ray diffraction of **7b**, and for all other analogues by analogy. During reaction we saw trace trimerisation of **6a** and **6c**. In both cases regioselectivity was poor; symmetric and asymmetric trimers **11a** and **11c** were formed and purified in almost equal amounts. Finally, olefin reduction was achieved using hydrogen over palladium on barium sulfate to conclude our synthesis of the desired cylindrocyclophane analogues **1a–c**. Using this route we were able to prepare more than one gram of [7.7]cylindrocyclophane **1b** for study in eight steps and 25% overall yield (cf 11–16 steps, 8–22% yield for the natural products).

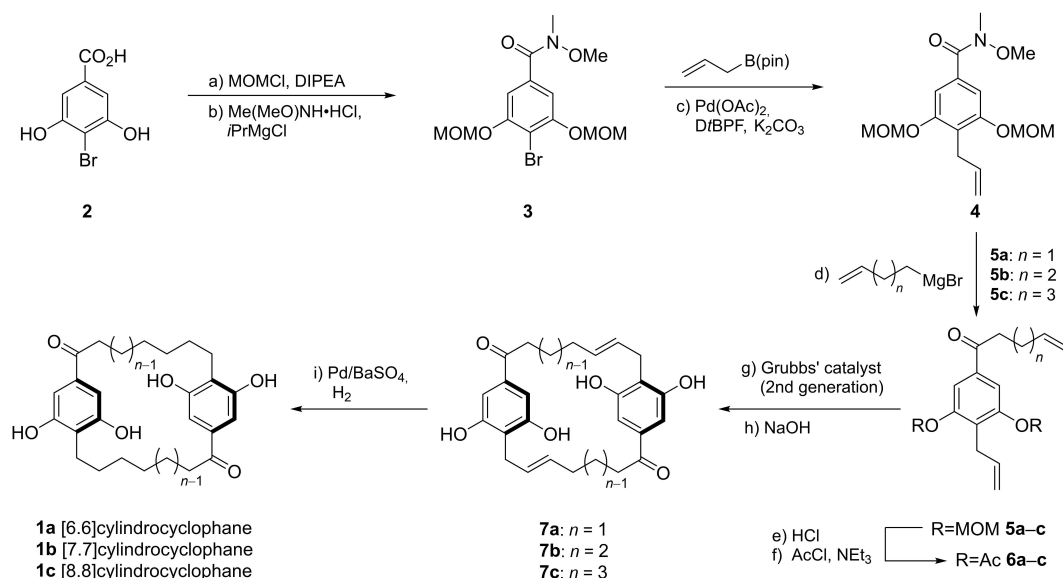
We were able to build on this synthesis by diversifying the unsaturated [7.7]paracyclophane intermediate **7b** at this stage. As mentioned, conversion to **1b** was achieved using hydrogen and palladium over barium sulfate. In addition, reduction of **7b** to **8** and doubly reduced cylindrocyclophane **9** completed this set of structural analogues, enabling us to investigate the effect of sequential reductions upon the antibacterial activity of this scaffold. Finally, based on a series of isolated natural cyclophanes with some uncommon substituents we also sought to investigate a particularly unusual modification to the cylindrocyclophane core to see if bromination had an effect of antibacterial activity. Previous work has identified a family of brominated cylindrocyclophanes isolated from *Nostoc* when cultured under particular conditions.<sup>[14]</sup> Amongst these, a tetrabrominated cylindrocyclophane (cylindrocyclophane  $A_{B4}$ ), where the bromination is on the alkylresorcinol motif, has been shown to be 40 times more potent against *M. tuberculosis* than its tetrachlorinated analogue (cylindrocyclophane  $A_4$ ).<sup>[33]</sup> Interested by the unique effect of this modification, we aimed to investigate the effect a similar transformation upon the



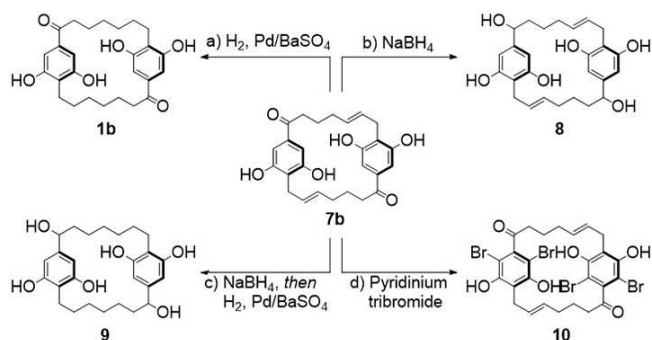
**Scheme 1.** Disconnection of the cylindrocyclophane core into symmetrical monomers. Dashed lines indicate the intended location of bond disconnection. P = protecting group.



**Figure 1.** Structural features of [7.7]paracyclophane natural products. All share a dimeric alkylresorcinol motif but differ in substitution pattern.  $R^1$ – $R^4$  represent side chain substituents.



**Scheme 2.** Synthesis of cylindrocyclophanes **1a–1c**. a) chloromethyl methyl ether (3.3 equiv), *N,N*-diisopropylethylamine/ $\text{CH}_2\text{Cl}_2$  (1:1), 0 °C to rt, 16 h, quant.; b)  $\text{Me}(\text{MeO})\text{NH}\cdot\text{HCl}$  (1.6 equiv), *i*PrMgCl (3.2 equiv), THF, –10 °C, 30 min, 75%; c)  $\text{Pd}(\text{OAc})_2$  (3 mol %), 1,1'-bis(di-*tert*-butylphosphino)ferrocene (3.6 mol %),  $\text{K}_2\text{CO}_3$  (3 equiv), allylboronic acid pinacol ester (2.5 equiv), THF, reflux, overnight, 78%; d) alkenyl magnesium bromide (2 equiv), THF, 0 °C to rt, 2 h, 72–89%; e)  $\text{HCl}/\text{MeOH}$  (1:2), 60 °C, 1–3 h, quant.; f)  $\text{NEt}_3$  (4.4 equiv),  $\text{AcCl}$  (4.4 equiv),  $\text{CH}_2\text{Cl}_2$ , 0 °C to rt, overnight, 66–87%; g) Grubbs' 2nd-generation catalyst (5 mol %),  $\text{CH}_2\text{Cl}_2$ , reflux, 20 h, 4–61%; h)  $\text{NaOH}$  (12 equiv),  $\text{MeOH}/\text{CH}_2\text{Cl}_2/\text{H}_2\text{O}$  (4:1:1), rt, 1 h, 84–89%; i)  $\text{H}_2$  (1 atm),  $\text{Pd}/\text{BaSO}_4$  (10 wt %), acetone, rt, overnight, 35–60%.



**Scheme 3.** Late-stage diversification of [7.7]cylindrocyclophane **1b**. a)  $\text{H}_2$  (1 atm),  $\text{Pd}/\text{BaSO}_4$  (10 wt %), acetone, rt, overnight, 35%; b)  $\text{NaBH}_4$  (4.8 equiv),  $\text{MeOH}$ , rt, 30 min, 55%; c)  $\text{NaBH}_4$  (4.8 equiv),  $\text{MeOH}$ , rt, 30 min, then  $\text{Pd}/\text{BaSO}_4$  (10 wt %), acetone, rt, overnight, 68% for second step; d) pyridinium tribromide (2.4 equiv),  $\text{EtOH}$ , rt, overnight, 23%.

resorcinol core of **7b**. We were able to effect a selective late-stage bromination of **7b** using pyridinium tribromide, which yielded tetrabrominated cylindrocyclophane **10** to complete the synthesis for this study.

We screened compounds **1a–c**, **6a–c**, **7a–c** and **8–10** for activity against a range of clinical pathogens using an adapted broth microdilution method.<sup>[35]</sup> Compounds were tested by using a twofold dilution series in biological duplicate and technical triplicate against *S. aureus* (Newman), epidemic MRSA type 15 (EMRSA-15), *Serratia marcescens* (Sma12), *Escherichia coli* (Beecham's) and *Pseudomonas aeruginosa* (PA01).

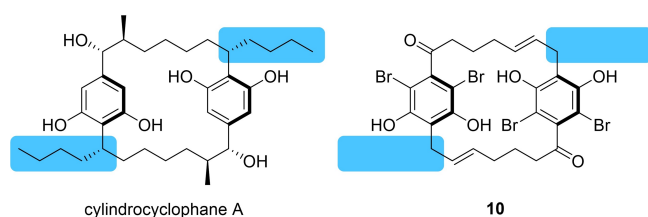
The cylindrocyclophanes in this work inhibited the growth of *S. aureus* and MRSA (Table 1) selectively, which corroborates the antibacterial activity of cylindrocyclophane natural products

reported elsewhere.<sup>[18]</sup> Gram-negative bacteria *S. marcescens*, *E. coli* and *P. aeruginosa* were not susceptible to any of the compounds in this work (minimum inhibitory concentration (MIC) > 200  $\mu\text{M}$ ). In addition, acetate-protected monomers **6a–c** and their metathesis products **12a–c** were inactive in all assays, corroborating a previous observation that the resorcinol core is required for biological activity of the cylindrocyclophanes.<sup>[30]</sup>

The [6.6]cylindrocyclophanes **1a** and **7a** exhibited little activity when tested, whereas the [7.7]- (**1b**, **7b**, **8**, **9** and **10**) and [8.8]cylindrocyclophane (**1c** and **7c**) series were more effective in this regard. Doubly oxidized compound **7b** was the only member of the [7.7]cylindrocyclophanes unable to arrest growth of *S. aureus*. Both [8.8]cylindrocyclophanes **1c** and **7c** were effective inhibitors, suggesting that expansion, but not contraction, of the 22-membered ring may be tolerated by members of this family. An authentic sample of a natural cylindrocyclophane was not available to us at this time, but not one of the analogues in this work exhibited MICs as potent as those reported for cylindrocyclophane A (0.45  $\mu\text{M}$ ).<sup>[18]</sup> This suggests that the alkylresorcinol motif absent in these analogues imparts activity upon the [7.7]cylindrocyclophane core, although this remains unconfirmed in the absence of a direct comparison (Figure 2).

As such, we focused our study on the natural [7.7] architecture, and in particular the remarkable effect of tetrabrominated compound **10** (MIC 12.5  $\mu\text{M}$ ) relative to its unsubstituted congener **7b** (MIC > 200  $\mu\text{M}$ ). This result suggests that substitution beyond naturally occurring paracyclophanes might not only be tolerated, but also perhaps a fruitful endeavour in the search for new inhibitors of *S. aureus*. More generally, it may be that bromination, although rarely explored as part of





**Figure 2.** Comparison of cylindrocyclophane A and **10**, highlighting the alkylresorcinol motif in cylindrocyclophane A and the lack thereof in **10**.

systematic SAR, can improve the activity of related inhibitors of *S. aureus* or other pathogens. We evaluated tetrabrominated macrocycle **10** further against *S. aureus* and determined its minimum bactericidal concentration (MBC) as 25  $\mu\text{M}$ , suggesting a bactericidal mechanism of action for **10**. Cell viability was unaffected by **10** below its MBC but some bacteriostatism was observed at concentrations as low as 6.25  $\mu\text{M}$ .

Many respiratory inhibitors are uncouplers, which dissipate the transmembrane proton gradient to uncouple electron transport from adenosine triphosphate (ATP) synthesis. Uncouplers are typically large, amphiphilic weak acids that can permeate into the cell, ionise once inside, and traverse back out to the cell exterior where it reprotonates and this process repeats. This “short-circuits” ATP synthesis as a means for proton translocation, dissipating the proton motive force (PMF) and rendering the cell unable to generate energy in the form of ATP. At first glance the structure of **10** lends itself to uncoupling. We sought to characterise its uncoupling ability by measuring oxygen consumption in *S. aureus* using a Clark-type oxygen microsensor (oxygraph). Treatment of *S. aureus* with **10** (25  $\mu\text{M}$ ) was followed by an immediate decrease in oxygen consumption, which suggests that **10** is not an uncoupler – rather that it inhibits some part of the respiratory chain.

We measured susceptibility data for **10** across the pH range 5.0–9.0. Sensitivity to **10** decreased with increasing pH; the MIC increased from 6.25  $\mu\text{M}$  at pH 5.0 to > 100  $\mu\text{M}$  at pH 9.0, which supports an mechanism of action involving the transmembrane proton gradient,  $\Delta\text{pH}$ . To corroborate these findings we looked at the ability of sublethal concentrations of **10** to modulate the activity of clinical antibiotics kanamycin and tetracycline. Kanamycin and tetracycline uptake are driven by the electrical potential ( $\Delta\psi$ ) and  $\Delta\text{pH}$ , respectively. As such, dissipation of  $\Delta\text{pH}$  increases sensitivity to kanamycin and decreases sensitivity to tetracycline. In line with this, co-administration of *S. aureus* with **10** (6.25  $\mu\text{M}$ ) and the corresponding antibiotic resulted in a modest changes to kanamycin (sensitivity increased ca. twofold) and tetracycline (sensitivity decreased ca. four- to eightfold) relative to their untreated controls. These observations suggest that dissipation of  $\Delta\text{pH}$  contributes to the antibacterial activity of **10**, although more work is needed to build on the weak cooperativity seen in these experiments before firm conclusions are drawn.

## Discussion

Both tetrahalogenated cylindrocyclophane analogues (**A<sub>4</sub>** and **A<sub>B4</sub>**) exhibit similar cytotoxicity,<sup>[33]</sup> this is an intriguing prospect, as it suggests that cytotoxicity of [7.7]paracyclophanes might not be related to increasing lipophilicity alone, but is still primarily based in the core resorcinol structure.

Based on our findings, we hypothesised that **10** disrupts the *S. aureus* cytoplasmic membrane or cell wall to compromise structure or function. These two targets are more easily accessible than intracellular targets and play crucial roles in cell structure and function (including cellular processes such as resistance, substrate transport, respiration, quorum sensing and energetics), and are conserved across bacteria.<sup>[36]</sup> Although the cell wall is an established target inhibited by antibiotics such as  $\beta$ -lactams and glycopeptides, the cell membrane is relatively unexplored due to concerns around mammalian toxicity.<sup>[37]</sup> In line with previous studies, which suggest brominating rigidifies the ore scaffold of membrane-active macrocycles and increases potency against MRSA,<sup>[38]</sup> we thought the chemoselective bromination of cylindrocyclophane **7b** to afford its halogenated congener **10** would also increase its rigidity and so support a similar mode of action involving membrane disruption. The chemoselective late-stage bromination using pyridinium tribromide in this study and the scaffolds explored herein may find use in future studies of the cylindrocyclophanes and supramolecular chemistry.<sup>[39,40]</sup>

## Conclusions

We have reported the development of a divergent synthetic strategy for the study of novel cylindrocyclophane scaffolds. Application of this method enabled us to generate a range of novel macrocycles varying in ring size, oxidation level and functionalisation around the cyclophane core. Antibacterial evaluation of these compounds demonstrated that modification of the cylindrocyclophane natural products can be achieved without total loss of activity, and from these we identified several novel inhibitors of *S. aureus* and methicillin-resistant *S. aureus*. We have described preliminary structure–activity requirements of these scaffolds, including the requirement for unprotected resorcinols and superiority of the natural [7.7] paracyclophane motif and larger ring sizes. In general, structural simplification of the cylindrocyclophanes was associated with decreased antibacterial activity. Nonetheless, in line with other studies,<sup>[14]</sup> we found that bromination increased activity at least eightfold relative to its non-halogenated congener. This compound (**10**) was less active than has been reported for the cylindrocyclophane natural products,<sup>[18]</sup> however bromination of the natural products may yet identify more potent inhibitors than those already known. Detailed profiling of **10** and this family is underway, and developments will be reported in due course.

## Acknowledgements

Research in our laboratory is supported by the European Research Council (FP7/2007-2013 and 279337/DOS), Engineering and Physical Sciences Research Council, Biotechnology and Biological Sciences Research Council (BB/M019411), European Commission, Medical Research Council, Royal Society and Wellcome Trust. We have also received support from the Swiss National Science Foundation and the Herchel Smith Fund. These funding agencies had no role in the design of, conduct of or decision to publish this study.

## Conflict of Interest

The authors declare no conflict of interest.

**Keywords:** cross metathesis • cylindrocyclophane • macrocycles • ring-closing metathesis

- [1] S. Y. C. Tong, J. S. Davis, E. Eichenberger, T. L. Holland, V. G. Fowler, *Clin. Microbiol. Rev.* **2015**, *28*, 603–61.
- [2] World Health Organization, *Antimicrobial Resistance: Global Report on Surveillance*, **2014**.
- [3] E. Tacconelli, E. Carrara, A. Savoldi, S. Harbarth, M. Mendelson, D. L. Monnet, C. Pulcini, G. Kahlmeter, J. Kluytmans, Y. Carmeli, et al., *Lancet Infect. Dis.* **2018**, *18*, 318–327.
- [4] M. S. Butler, M. A. Blaskovich, M. A. Cooper, *J. Antibiot. (Tokyo)*. **2017**, *70*, 3–24.
- [5] D. J. Payne, M. N. Gwynn, D. J. Holmes, D. L. Pompliano, *Nat. Rev. Drug Discovery* **2007**, *6*, 29–40.
- [6] R. Tommasi, D. G. Brown, G. K. Walkup, J. I. Manchester, A. A. Miller, *Nat. Rev. Drug Discovery* **2015**, *14*, 529–542.
- [7] E. D. Brown, G. D. Wright, *Nature* **2016**, *529*, 336–343.
- [8] K. M. G. O'Connell, J. T. Hodgkinson, H. F. Sore, M. Welch, G. P. C. Salmond, D. R. Spring, *Angew. Chem. Int. Ed.* **2013**, *52*, 10706–10733.
- [9] M. Dow, F. Marchetti, K. A. Abrahams, L. Vaz, G. S. Besra, S. Warriner, A. Nelson, *Chem. Eur. J.* **2017**, *23*, 7207–7211.
- [10] I. B. Seiple, Z. Zhang, P. Jakubec, A. Langlois-Mercier, P. M. Wright, D. T. Hog, K. Yabu, S. R. Allu, T. Fukuzaki, P. N. Carlsen, et al., *Nature* **2016**, *533*, 338–345.
- [11] W. R. J. D. Galloway, A. Bender, M. Welch, D. R. Spring, *Chem. Commun.* **2009**, 2446.
- [12] B. S. Moore, J. L. Chen, G. M. L. Patterson, R. E. Moore, L. S. Brinen, Y. Kato, J. Clardy, *J. Am. Chem. Soc.* **1990**, *112*, 4061–4063.
- [13] B. S. Moore, J.-L. Chen, G. M. L. Patterson, R. E. Moore, *Tetrahedron* **1992**, *48*, 3001–3006.
- [14] G. E. Chlipala, M. Sturdy, A. Kronic, D. D. Lantvit, Q. Shen, K. Porter, S. M. Swanson, J. Orjala, *J. Nat. Prod.* **2010**, *73*, 1529–1537.
- [15] J. L. Chen, R. E. Moore, G. M. L. Patterson, *J. Org. Chem.* **1991**, *56*, 4360–4364.
- [16] H. T. N. Bui, R. Jansen, H. T. L. Pham, S. Mundt, *J. Nat. Prod.* **2007**, *70*, 499–503.
- [17] S. Luo, H.-S. Kang, A. Kronic, G. E. Chlipala, G. Cai, W.-L. Chen, S. G. Franzblau, S. M. Swanson, J. Orjala, *Tetrahedron Lett.* **2014**, *55*, 686–689.
- [18] M. Preisitsch, K. Harmrolfs, H. T. Pham, S. E. Heiden, A. Füssel, C. Wiesner, A. Pretsch, M. Swiatecka-Hagenbruch, T. H. Niedermeyer, R. Müller, et al., *J. Antibiot. (Tokyo)*. **2015**, *68*, 165–177.
- [19] H.-S. Kang, B. D. Santarsiero, H. Kim, A. Kronic, Q. Shen, S. M. Swanson, H. Chai, A. D. Kinghorn, J. Orjala, *Phytochemistry* **2012**, *79*, 109–115.
- [20] D. S. May, W.-L. Chen, D. D. Lantvit, X. Zhang, A. Kronic, J. E. Burdette, A. Eustaquio, J. Orjala, *J. Nat. Prod.* **2017**, *80*, 1073–1080.
- [21] T. P. Martins, C. Rouger, N. R. Glasser, S. Freitas, N. B. de Fraissinette, E. P. Balskus, D. Tasdemir, P. N. Leão, *Nat. Prod. Rep.* **2019**, *36*, 1437–1461.
- [22] S. C. Bobzin, R. E. Moore, *Tetrahedron* **1993**, *49*, 7615–7626.
- [23] H. Nakamura, H. A. Hamer, G. Sirasani, E. P. Balskus, *J. Am. Chem. Soc.* **2012**, *134*, 18518–18521.
- [24] T. Gulder, P. S. Baran, *Nat. Prod. Rep.* **2012**, *29*, 899.
- [25] A. B. Smith, S. A. Kozmin, D. V. Paone, *J. Am. Chem. Soc.* **1999**, *121*, 7423–7424.
- [26] A. B. Smith, S. A. Kozmin, C. M. Adams, D. V. Paone, *J. Am. Chem. Soc.* **2000**, *122*, 4984–4985.
- [27] T. R. Hoye, P. E. Humpai, B. Moon, *J. Am. Chem. Soc.* **2000**, *122*, 4982–4983.
- [28] A. B. Smith, C. M. Adams, S. A. Kozmin, *J. Am. Chem. Soc.* **2001**, *123*, 990–991.
- [29] A. B. Smith, C. M. Adams, S. A. Kozmin, D. V. Paone, *J. Am. Chem. Soc.* **2001**, *123*, 5925–5937.
- [30] H. Yamakoshi, F. Ikarashi, M. Minami, M. Shibuya, T. Sugahara, N. Kanoh, H. Otori, H. Shibata, Y. Iwabuchi, *Org. Biomol. Chem.* **2009**, *7*, 3772.
- [31] K. C. Nicolaou, Y. P. Sun, H. Korman, D. Sarlah, *Angew. Chem. Int. Ed.* **2010**, *49*, 5875–5878.
- [32] M. Preisitsch, T. H. J. Niedermeyer, S. E. Heiden, I. Neidhardt, J. Kumpfmüller, M. Wurster, K. Harmrolfs, C. Wiesner, H. Enke, R. Müller, et al., *J. Nat. Prod.* **2016**, *79*, 106–115.
- [33] M. Preisitsch, S. E. Heiden, M. Beerbaum, T. H. J. Niedermeyer, M. Schneefeld, J. Herrmann, J. Kumpfmüller, A. Thürmer, I. Neidhardt, C. Wiesner, et al., *Mar. Drugs* **2016**, *14*, 21.
- [34] D. Berthold, B. Breit, *Chem. Eur. J.* **2018**, *24*, 16770–16773.
- [35] I. Wiegand, K. Hilpert, R. E. W. Hancock, *Nat. Protoc.* **2008**, *3*, 163–175.
- [36] M. T. Cabeen, C. Jacobs-Wagner, *Nat. Rev. Microbiol.* **2005**, *3*, 601–610.
- [37] L. L. Silver, *Clin. Microbiol. Rev.* **2011**, *24*, 71–109.
- [38] B. Sun, M. Zhang, Y. Li, Q. Hu, H. Zheng, W. Chang, H. Lou, *Bioorg. Med. Chem. Lett.* **2016**, *26*, 3617–3620.
- [39] N. K. Mitra, R. Meudom, J. D. Gorden, B. L. Merner, *Org. Lett.* **2015**, *17*, 2700–2703.
- [40] K. S. Unikela, T. L. Roemmele, V. Houska, K. E. McGrath, D. M. Tobin, L. N. Dawe, R. T. Boeré, G. J. Bodwell, *Angew. Chem. Int. Ed.* **2018**, *57*, 1707–1711.

Manuscript received: March 23, 2020

Accepted manuscript online: May 18, 2020

Version of record online: June 12, 2020