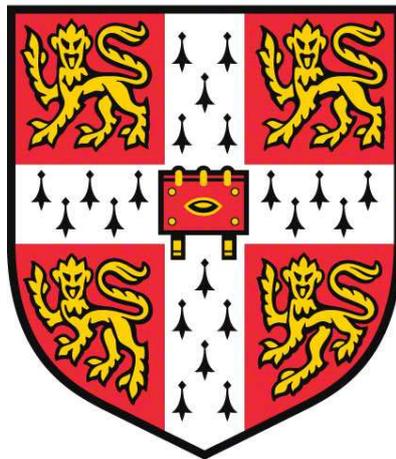


**Rational development of novel small molecule leads
against the transcriptional activator protein ExsA of
*Pseudomonas aeruginosa***



Jack Calum Greenhalgh

University of Cambridge

Departments of Pharmacology and Biochemistry

Corpus Christi College

June 2022

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

Declaration

The work presented in this dissertation was conducted at the Departments of Pharmacology and Biochemistry, University of Cambridge between October 2017 and July 2022 under the supervision of Dr. Taufiq Rahman and Dr. Martin Welch. This thesis is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. It is not substantially the same as any that I have submitted or is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution. I further state that no substantial part of my thesis has already been submitted or is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution. The length of this thesis does not exceed the prescribed word limit stated by the Biology Degree Committee.

Rational development of novel small molecule leads against the transcriptional activator protein ExsA of *Pseudomonas aeruginosa*

Jack Calum Greenhalgh

Pseudomonas aeruginosa is a Gram-negative opportunistic pathogen. It is present in aquatic, marine, and soil environments and can also be found in a variety of anthropological environments, including water distribution systems and hospitals. This prevalence is concerning due to *P. aeruginosa*'s status as an opportunistic pathogen, and the latter particularly so because it is a common cause of nosocomial infections. Burn victims and cystic fibrosis patients are particularly vulnerable to *P. aeruginosa* infections. The pathogen is highly virulent, and consequently it is a leading cause of death in intensive care units, especially among cases of ventilator associated pneumonia. Treating *P. aeruginosa* infections is complicated by its resistance to multiple drugs. Besides acquiring resistance mutations to antibiotic targets, the pathogen also possesses a suite of efflux pumps capable of exporting a wide range of antibiotics. *P. aeruginosa* has two distinct virulent lifestyles which correspond to acute and chronic infections. *P. aeruginosa*'s type three secretion system (T3SS) dominate in acute infection, and the expression of this system is regulated by the transcription factor ExsA. The latter therefore represents an attractive target for developing anti-virulence drugs against this pathogen.

I have further characterised ExsA and its regulon with a proteomics experiment utilising deletion mutants which have an "ExsA always on" phenotype as well as an *exsA* deletion mutant. An extended ExsA regulon was revealed, including well known virulence factors such as HCN, and potential novel factors such as an uncharacterised non-ribosomal peptide synthase. Effectors of, and components for, the type six secretion system (associated with chronic infections and generally inversely regulated compared to the type three secretion system) were also identified as overexpressed in the "ExsA always on" mutants. Potential connections to other signalling systems are also examined. This work strengthened the case for ExsA as a therapeutic target, expanding its virulence inducing role beyond the T3SS. A number of other phenotypes, such as the downregulation of denitrification proteins, are also identified and validation is sought through phenotypic assays.

ExsA is subsequently examined bioinformatically, and the inhibitors and ligands of related proteins (i.e. members of the AraC family of transcription factors) are examined for potential inhibitory effects. Due to dearth of potent inhibitors with well-characterised mechanism of action, some potential small molecule binding sites were predicted and subsequently utilised for *in silico* screening of commercial lead like libraries. Parallel to this, attempts were made to obtain the full-length crystal structure of ExsA which were ultimately unsuccessful.

Several iterations of *in silico* docking experiments were performed, utilising a combination of published and modelled structures of the ExsA. This led to the identification of novel chemical scaffolds as potential binders against the chosen pocket of ExsA. Best hits from the *in silico* screening were then subjected to *in vivo* and biophysical analysis with mixed results. Due to the impact on Covid-19, the complete characterisation of those compounds was not feasible, though several of them appeared to be promising leads. Finally, a comprehensive effort to obtain an optimal structure of ExsA was undertaken. Whilst an experimental structure remained elusive, state of the art structural modelling was undertaken alongside all-atoms molecular dynamics simulations.

Together the research presented in the thesis offers a firm foundation and several leads for further inhibitor discovery efforts against ExsA, as well as findings of biological significance concerning the full regulator effects of ExsA.

Acknowledgements

I would like to thank Dr Taufiq Rahman and Dr Martin Welch for their help and guidance throughout my PhD. There is nothing more I could have asked for from either of them and I deeply appreciate all their support and advice. My thanks go to the BBSRC for funding my PhD, and also to Shionogi for funding the proteomics project. I would also like to thank Corpus Christi for their support throughout.

I cannot thank my Welch lab colleague Dr Meng Wang enough for making the final preparations for the proteomics experiments, preparing and submitting the samples to the Cambridge Centre for Proteomics on my behalf. I would also like to thank her profusely for her help with the data analysis, and for generally being a fantastic post doc around the lab.

I would like to thank the rest of the Rahman and Welch labs. In the Rahman lab Maria Eznarriaga Gutierrez deserves a special thanks for keeping the place ship shape, as do Xavier Chee and Rory Triniman for showing me the ropes. Dr Shunsuke Numata was fantastically helpful, both during his time in the Welch lab and once he had returned to Shionogi. Dr Stephen Dolan provided numerous strains which I made use of, as well as establishing several useful protocols. Dr Stephen Trigg provide excellent advice regarding denitrification and was one of many that made the Welch lab a great place to be. The full list of colleagues from the Welch lab which deserve thanks is too long to list here; but thank you all!

Dr Thales Kronenberger of the Dept. of Pharmacy, University of Tübingen conducted the molecular dynamics simulations presented here, and I would like to thank him for being a fantastic collaborator. I would also like to thank Professor Antti Poso of the University of Eastern Finland for the opportunity to visit Finland (even though Covid-19 prevented me from doing so), and for putting me in touch with Thales.

Finally, all of my family deserve my thanks for supporting and putting up with me, especially my wife and children for letting me drag them around the country and everyone who supported and made space for me whilst writing up.

Contents

Contents	ii
List of Figures	ix
List of Tables	xii
List of Appendixes.....	xiii
Abbreviations	xiv
1. Introduction	1
1.1 Antimicrobial resistance – a formidable biomedical challenge for the 21 st century.....	1
1.2 Anti-virulence – an effective therapeutic strategy	4
1.3 <i>Pseudomonas aeruginosa</i> - a priority pathogen.....	7
1.4 The virulence strategies of <i>P. aeruginosa</i>	8
1.5 The T3SS of <i>P. aeruginosa</i> - a potent anti-virulence target.....	12
1.5 ExsA – a target for the upstream regulation of the T3SS	13
1.6 <i>In silico</i> drug discovery.....	17
1.6.1 Ligand-guided virtual screening	18
1.6.2 Structure-guided virtual screening	18
1.6.3 Molecular dynamics simulations	19
1.7 Aims and objectives.....	20
2. General methods	21
2.1 General <i>in silico</i> methods.....	21
2.1.1 Preparation and analysis software (Ftsite, MMV, ICM)	21
2.1.2 Focused docking using GOLD suite	21
2.1.3 Blind docking Vina.....	21
2.1.4 Ligand oriented screening with Rocs	22
2.2 Molecular microbiology	22
2.2.1 Bacterial strains and growth conditions.....	22
2.2.2 Storage of bacterial cultures	23
2.2.3 Molecular cloning	23
2.2.3.1 Polymerase chain reactions.....	23
2.2.3.2 Agarose gel electrophoresis for DNA.....	23

2.2.3.3 Gel extraction	24
2.2.3.4 Purification and isolation of DNA	24
2.2.3.5 Restriction endonuclease digestion	24
2.2.3.6 Ligation.....	24
2.2.3.7 Transformation by electroporation and chemically competent cells.	25
2.2.3.7 DNA sequencing	25
2.3 Protein purification	25
2.3.1 Expression testing of novel constructs.....	25
2.3.2 Protein expression.....	26
2.3.3 Nickel affinity chromatography.....	27
2.3.4 Heparin affinity chromatography	27
2.3.5 Amylose affinity chromatography.....	28
2.3.6 Preparatory size exclusion chromatography	28
2.3.7 Analytical size exclusion chromatography	28
2.4 Protein analysis	28
2.4.1 Quantification of protein.....	28
2.4.2 Bradford assay	29
2.4.3 Acrylamide electrophoresis for protein.....	29
2.4.4 Coomassie staining	29
2.4.5 Preparation of cell culture for Western blotting	30
2.4.6 Pre-absorption of antibodies.....	30
2.4.7 Western blotting	30
2.4.8 Protein sequencing.....	31
2.5 <i>in vivo</i> assays	31
2.5.1 Luciferase reporter assays and growth curves.....	31
2.5.2 Growth curves	31
2.5.3 Preparation of samples for ExsA inhibition Western blotting	32
2.6 Thermal shift assay.....	32
3. Proteomic Analysis of the ExsA regulatory cascade.....	33

3.1 Aims and Approach	33
3.2 Methods.....	33
3.2.1 Unmarked mutant creation	33
3.2.1.1 Splicing by overlap extension	34
3.2.1.2 Validation of mutants	34
3.2.2 Growth in 50 mL shaking flasks	34
3.2.3 Concentration of secreted proteins	34
3.2.4 Artificial sputum medium continuous flow model.....	35
3.2.5 Proteomic analysis	36
3.2.6 STRING map construction.....	36
3.2.7 Pyocyanin Quantification	36
3.2.8 Siderophore and nitrite quantifications.....	37
3.3 A slight growth defect is apparent in the “always on” mutants.....	37
3.4 Preliminary characterisation of ExsA activity by western blot	39
3.4.1 Mutant expression of PcrV correspond to expected phenotypes	40
3.4.2 An ASM continuous flow model induces little PcrV expression	40
3.4.3 AGSY supports more robust PcrV expression than LB	40
3.4.4 PcrV and Hcp1 expression increase in late growth phase	44
3.5 Proteomics.....	45
3.5.1 The proteomes of $\Delta exsA$ and WT are similar whilst $\Delta exsD$ and $\Delta exsE$ cluster separately	45
3.5.2 The $\Delta exsA$ mutant diverges from wild type on a small number of proteins.....	46
3.5.3 $\Delta exsD$ and $\Delta exsE$ mutants present similar proteomes.....	47
3.5.4 $\Delta exsD$ compared to $\Delta exsA$ offers a sharper contrast than compared to wild type	57
.....	59
3.5.5 Pyocyanin synthesis enzymes are upregulated in “always on” mutants.....	60
3.5.7 The ExsA cascade has a regulatory role in denitrification.....	62
3.5.9 The ExsA cascade has a limited impact on broader virulence signalling.....	66

3.5.10 Not all reported members of the ExsA regulon are differentially regulated ...	68
3.6 Discussion	70
3.6.1 Interpretation of preliminary results.....	70
3.6.2 Virulence and signalling implications of ExsA activity	71
3.6.3 Downregulation in “always on” mutants	75
3.7 Conclusions	77
4. Bioinformatic and structural investigation of approaches for ExsA inhibition.....	78
4.1 Aims and approaches	78
4.2 Methods.....	78
4.2.1 Comparison of ligand-binding pockets between ToxT and ExsA	78
4.2.2 Blind docking of ToxT ligands to ExsA with Autodock 4.....	78
4.2.3 Phylogeny generation.....	79
4.2.4 SNP analysis.....	79
4.2.5 Development of methods to copurify ExsA and ExsD	80
4.2.6 Crystallisation screening.....	80
4.3 ToxT inhibitors do not inhibit ExsA.....	80
4.4 The close phylogenetic relations of ExsA offer no probable inhibitors	83
4.3 Analysis of single nucleotide polymorphisms (SNPs) identify inconsistent regions across ExsA.....	84
4.5 ExsD and ExsA and co-purifiable as a complex but evaded co-crystallisation under any tested condition.....	86
4.5.1 ExsA-ExsD co-purification requires different conditions to ExsA alone	86
4.5.2 ExsA-ExsD has extremely low salt tolerance	88
4.5.3 Size exclusion chromatography was necessary for purity ExsA-ExsD, which were confirmed to co-express	90
4.6 Discussion	94
4.7 Conclusion.....	96
5. Preliminary identification of ExsA inhibitors targeting the NTD	97
5.1 Aims and approaches	97
5.2 Methods.....	97

5.2.1 Identification of potential inhibitors of the ExsA ligand pocket	97
5.2.2 Identification of small molecule ligands that could potentially bind to the ExsA dimer interface	98
5.2.3 Development of a data handling R script	100
5.2.4 Identification and docking of ROCS hits for L198.....	100
5.2.5 Electromobility shift assay	100
5.3 Computational identification of potential ligands against the NTD of ExsA.....	101
5.4 Identification of two small molecules that reproducibly inhibited luminescence of a ExsA dependant Lux reporter	104
5.5 Compound L198 was lethal to several bacterial species during stationary phase .	107
5.6 Compounds L980 and L198 did not inhibit PlacZ-Lux reporter luminescence	111
5.7 Strain YM64 is not suitable for modification as a luciferase reporter	111
5.8 No FDA approved drugs are an appropriate replacement for L198.....	111
5.9 Optimization of the ExsA NTD purification	115
5.10 Thermal Shift Assay for putative ExsA NTD binding compounds	115
5.11 Attempt to establish an electromobility shift assay	117
5.12 Discussion	119
5.13 Conclusions	121
6. Identification of small molecule inhibitors of ExsA - an extended investigation	122
6.1 Aims and approaches.....	122
6.2 Methods.....	122
6.2.1 Generation of protein models	122
6.2.4 Overall docking approach.....	123
6.2.5 Docking with FRED	123
6.2.6 Docking with GLIDE	123
6.3 A composite structure of 4ZUA and a CTD model provides the most suitable docking receptor	123
6.4 Varied docking algorithms show little correlation in score	129
6.5 Structure-guided screening reveals potential hits.....	129

6.6 Virtual screening and subsequent experimental testing of the in-house putative PARP inhibitors	136
6.7 Discussion	143
6.7.1 A full length model of ExsA.....	143
6.7.2 Small molecule inhibitors of ExsA.....	145
6.8 Conclusion.....	147
7. Modelling of full length ExsA using AlphaFold	148
7.1 Aims and approaches	148
7.2 Methods.....	148
7.2.1 AlphaFold for protein modelling	148
7.2.2 Prediction of disorder with IUPred2A	148
7.2.3 Prediction of conservation with ConSurf	148
7.2.4 Virtual alanine interface scanning	149
7.2.5 Model generation with ClusPro.....	149
7.2.6 Molecular dynamics simulations	149
7.3 AlphaFold predicts an open model of monomeric ExsA	150
7.4 AlphaFold successfully recreates published dimeric AraC family structures.....	154
7.5 AlphaFold predicts two different conformations of ExsA dimer.....	157
7.6 AlphaFold model 2 and ClusPro model 1 of the ExsA homodimer are highly similar	160
7.7 The dimerization interface is broadly conserved between models 1 and 2.....	160
7.8 Alanine interface scanning highlights discrepancies surrounding Met136 and Glu144 between AlphaFold models 1 and 2.....	162
7.9 Together dimeric models 1 and 2 offer a glimpse of ExsA flexibility	163
7.10 Molecular dynamics simulations support dimer model 2 and find correlation between dimeric and DNA bound structures	166
7.11 AlphaFold predicts novel ExsA-ExsD binding conformation	171
7.13 AlphaFold models of ExsA mutations and an ExsD fragment conflict with published <i>in vitro</i> data.....	171
7.14 Discussion	174

7.14.1 Analysis of the AlphaFold-derived model of an ExsA monomer	174
7.14.2 Analysis of the AlphaFold-derived models of ExsA dimers	175
7.14.3 Analysis of models of the ExsA-ExsD heterodimer	177
7. 15 Conclusions	177
8. Final Conclusions.....	179
References	183
Appendixes	210

List of Figures

1. Introduction

Figure 1.1 Schematic of future combined therapeutic approaches to bacterial infection.....	3
Figure 1.2 The structure of a T3SS.....	12
Figure 1.3 The function of the ExsA cascade.....	14
Figure 1.4 ExsA's dimerization and DNA binding.....	15
Figure 3.1 Growth curves of PAO1 and novel mutants.....	38
Figure 3.2 Growth curves from microplate shaking incubator are comparable.....	39
Figure 3.3 Preliminary phenotypic Western blots for deletion mutants.....	41
Figure 3.4 Western blots for samples from an ASM continuous flow system.....	42
Figure 3.5 Western blots for AGSY and LB.....	43
Figure 3.6 Western blot of PcrV in AGSY cultures.....	44
Figure 3.7 Western blot of samples across a growth curve.....	45
Figure 3.8 Scatter of proteomic samples after PLS-DS analysis.....	46
Figure 3.9 Volcano plot comparison of Δ exsA and WT.....	47
Figure 3.10 Volcano plot comparisons of always on mutants and WT.....	48
Figure 3.11 Annotated STRING map of proteins upregulated in Δ exsD mutants.....	50
Figure 3.12 Annotated STRING map of proteins upregulated in Δ exsE mutants.....	51
Figure 3.13 Volcano plot comparisons of always on mutants and WT for the NRPS operon PA1216-PA1221.....	52
Figure 3.14 Annotated STRING map of proteins down regulated in Δ exsD mutants.....	53
Figure 3.15 Annotated STRING map of proteins downregulated in Δ exsE mutants.....	54
Figure 3.16 Volcano plot comparisons of Δ exsD and Δ exsE.....	55
Figure 3.17 Annotated STRING map of proteins differentially expressed between Δ exsD and Δ exsE.....	56
Figure 3.18 Volcano plot comparisons of Δ exsD and Δ exsA.....	57
Figure 3.19 String map of upregulated proteins in Δ exsD compared Δ exsA.....	58
Figure 3.20 String map of downregulated proteins in Δ exsD compared Δ exsA.....	59
Figure 3.21 Pyocyanin synthesis volcano plot and pyocyanin extraction variation between mutants.....	61
Figure 3.22 Siderophore quantification assay.....	62
Figure 3.23 NirS Luciferase reporter activity in MOPS with acetate.....	63
Figure 3.24 NirS Luciferase reporter activity in MOPS with acetate and glucose.....	64
Figure 3.25 NirS Luciferase reporter activity in AGSY.....	65
Figure 3.26 Nitrite detection assay.....	66
Figure 3.27 Proteomic data for quorum sensing proteins.....	67

Figure 3.28 GacA network protein abundance comparison for Δ exsD and Δ exsA mutants.	68
Figure 3.29 chIP-seq hits compared to proteomic data.	69
Figure 4.1 Computational and <i>in vivo</i> analysis of virstatin and ExsA.	82
Figure 4.2 The phylogenetic branch of the AraC family containing ExsA.	84
Figure 4.3 Residues at which SNPs have been identified.	86
Figure 4.4 Co-purified ExsA and ExsD Coomassie stain.	87
Figure 4.5 ExsA and ExsD copurification buffer screen by thermal shift.	89
Figure 4.6 Crude estimation of ExsA and ExsD solubility.	90
Figure 4.7 Size exclusion chromatography of ExsA-ExsD copurification.	92
Figure 4.8 Analytical size exclusion chromatography of ExsA-ExsD copurification.	93
Figure 5.1 Potential ligand binding pocket within the N terminus of ExsA.	98
Figure 5.2 Docking workflows.	99
Figure 5.3 The ExsA dimerization interface.	103
Figure 5.4 Maximum luminescence of <i>PpcrV</i> luciferase reporter with putative inhibitors of ExsA.	105
Figure 5.5 Luciferase reporter for selected compounds.	107
Figure 5.6 Growth curves for candidate compounds with <i>PpcrV</i> luciferase reporter strain.	108
Figure 5.7 Growth of <i>E. coli</i> , <i>S. aureus</i> , and wild type PAO1 with L198.	109
Figure 5.8 Growth and luminescence of <i>PpcrV</i> luciferase reporter strain with sub minimal inhibitory concentrations of carbenicillin.	110
Figure 5.9 The lead compounds do not effect <i>PlacZ</i> luciferase reporter luminescence.	112
Figure 5.10 YM64 strain is no suitable for reporter strain.	113
Figure 5.11 Analysis of two FDA approved drugs with potential to replace L198.	114
Figure 5.12 Electrophoresis gels for ExsA NTD expression.	116
Figure 5.13 Thermal shift assay for lead compounds.	117
Figure 5.14 Preliminary EMSA gels.	118
Figure 5.15 EMSA gel for loading volume.	119
Figure 6.1 Alignment of selected sequences to ExsA.	125
Figure 6.2 Visual Comparison of ExsA Models and 4ZUA.	126
Figure 6.3 Visual comparison of final ExsA models with a focus on the target pocket.	128
Figure 6.4 SNPs identified on the ExsA CTD.	129
Figure 6.5 Comparison of the top 500 FRED compounds with scoring in GOLD and Glide.	131
Figure 6.6 An illustration of 4-hydroxyquinazoline ligands.	133
Figure 6.7 Docked poses of Z20230049 and 4-hydroxyquinazoline.	134
Figure 6.8 A visual comparison of docked poses for Z195013052.	135
Figure 6.9 Thermal shift assay of compounds from <i>in silico</i> docking.	137

Figure 6.10 Thermal shift assays for compounds from putative PARP inhibitor library.	138
Figure 6.11 Luciferase reporter assays for novel candidate molecules.	140
Figure 6.12 Luciferase reporter assays for other screening molecules.	141
Figure 6.13 Luciferase control assays for hit compounds.	142
Figure 6.14 PcrV Western blots with lead compounds.	143
Figure 7.1 AlphaFold models of ExsA and ToxT compared to crystal structures and prior homology modelling.	151
Figure 7.2 IDDT and predicted disorder in the AlphaFold model of ExsA.	152
Figure 7.3 Residue conservation in ExsA.	153
Figure 7.4 AlphaFold recreations of 1XJA and 6NX3 dimeric structures.	155
Figure 7.5 Attempts at AlphaFold recreations of 5U9E.	156
Figure 7.6 Monomeric AlphaFold variants of 5U9E.	157
Figure 7.7 The varied AlphaFold models of the ExsA dimer and their IDDT scores.	158
Figure 7.8 AlphaFold ExsA dimer models 1 and 2 compared by IDDT.	159
Figure 7.9 ClusPro recreations of AlphaFold models.	161
Figure 7.10 ExsA dimeric interface for model 2.	162
Figure 7.11 AlphaFold models of the ExsA monomer and dimer compared.	164
Figure 7.12 AlphaFold ExsA dimer models 1 and 2 with MarA.	165
Figure 7.13 AlphaFold ExsA dimer models with two DNA molecules.	166
Figure 7.14 MD simulation protein backbone motion.	167
Figure 7.15 Overview of the ExsA model structure MD simulations.	169
Figure 7.16 CTD and dimerization domain MD simulation overview.	170
Figure 7.17 An AlphaFold model of ExsA-ExsD interaction.	172
Figure 7.19 Modelling the ExsD 40 N terminal amino acids with ExsA.	174

List of Tables

Table 2.1 Frequently used bacterial strains.....	22
Table 2.2 Protein expression plasmids.....	26
Table 2.3 Salt concentration in protein purification buffers	27
Table 2.4 Antibodies used for western blotting	31
Table 3.1 Plasmids used for the generation of unmarked deletion mutants.....	33
Table 3.2 Primers utilised for unmarked deletion creation	35
Table 3.3 Expression change of Δ exsA mutant hits in Δ exsD and Δ exsR mutants.....	59
Table 3.4 Proteomic results for ImpA	70
Table 4.1 Results of Gold and Vina docking of chitin metabolites and cellobiose to the ExsA NTD.	83
Table 4.2 SNPs detected in ExsA	85
Table 5.1 Top scoring compounds docked to the ExsA ligand binding pocket.....	102
Table 5.2 Top scoring compounds docked to the ExsA dimerization interface	104
Table 5.3 Docking scores for L198 ROCS hits	113
Table 5.4 Mascot search results for ExsA NTD purification products	115
Table 6.1 Candidate template structures for ExsA	124
Table 6.3 A Selected compounds docking scores	132
Table 6.4 B Select Vina analysis.....	132
Table 6.5 Parp Library compounds docking scores.....	136
Table 7.1 Comparison of virtual and experimental mutagenesis for ExsA-ExsD heterodimer	173

List of Appendixes

Appendix 1: R script for the handling and analysis of Autodock Vina outputs.....	210
Appendix 2: A Phylogeny for ExsA.....	214

Abbreviations

AGSY	Alanine glycerol salt yeast extract media
AMR	Antimicrobial resistance
ASM	Artificial sputum media
chIP-seq	Chromatin immunoprecipitation sequencing
c-di-GMP	Cyclic diguanylate monophosphate
CTD	C terminal domain
DDM	N-dodecyl- β -D-maltoside
DNA	deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethylene glycol-bis(β -aminoethyl ether)-N,N,N',N'-tetraacetic acid
EMSA	Electromobility shift assay
HCN	Hydrogen cyanide
HTH	Helix-turn-helix
IPTG	Isopropyl β -D-1-thiogalactopyranoside
LB	Luria-Bertani (broth or agar)
LC-MS/MS	Liquid chromatography mass spectrometry – mass spectrometry
M	The full length AlphaFold model of ExsA when subjected to MD simulation
M _{DNA}	The full length AlphaFold model of ExsA when subjected to MD simulation with DNA present
MD	Molecular dynamics
NRPS	Non-ribosomal peptide synthase
NTD	N terminal domain
OD ₆₀₀	Optical density at $\lambda = 600$ nm
PBS	Phosphate buffered saline
PDB	Protein data base code
PLS-DA	Partial least squares discriminant analysis
QS	Quorum sensing
RMSD	Root mean squared deviation
RMSF	Root mean squared fluctuation
rpm	Revolutions per minute
SD	Standard Deviation
SDS	Sodium dodecyl sulphate
SNP	Non-synonymous single nucleotide polymorphism
T1SS	Type I secretion system
T3SS	Type III secretion system

T4SS	Type IV secretion system
T6SS	Type VI secretion system
TF	Transcription factor
VS	Virtual screening
WHO	World Health Organisation
WT	Wild type

1. Introduction

1.1 Antimicrobial resistance – a formidable biomedical challenge for the 21st century

Despite only featuring in apocalyptic newspaper headlines and the public's conscience in recent years, antimicrobial resistance (AMR) is not a new issue or recent discovery (Gallagher, 2015; Sample, 2018). Alexander Fleming warned of its emergence and consequences in his Nobel Prize lecture for the discovery of the first antibiotic to enter clinical use, penicillin (Fleming, 1945), and five years prior to the lecture a resistance mechanism had been characterised (Abraham and Chain, 1940). Similarly, the lack of progress toward truly novel antibiotics, rather than the modifications of existing compounds, and pharmaceutical companies' failure to engage with the challenge due to the disassociation of the profit motive and social good, has been bemoaned for some time (Projan, 2003; Shlaes, 2003). There were at least 700,000 deaths caused by antimicrobial resistant pathogens in 2017, with a projected rise to 10 million by 2050 (Neill, 2016). A statistical analysis conducted in 2022 places annual deaths due to bacterial AMR at 4.95 million, indicating that the lethality of AMR is increasing rapidly (Murray *et al.*, 2022). The large increase between confirmed deaths in 2017 and statistically probable deaths in 2022 likely reflects a real increase in part, however due to the different methods by which the numbers were generated they are not directly comparable and therefore not indicative of an increase of 4.25 million deaths in 5 years. Besides this grim projection of morbidity, the rising threat of AMR is projected to cost the global economy \$100 trillion in lost production by 2050 unless further action is taken, pushing 24 million people into extreme poverty by 2030 (Bryan-Wilson, 2016). With the "golden age" of antibiotic discovery behind us, little interest in discovery of novel agents from the pharmaceutical industry, and the rising clinical and economic burden of AMR, it is quite apparent that urgent research is required to combat resistant pathogens.

“A post-antibiotic era-in which common infections and minor injuries can kill, far from being an apocalyptic fantasy, is instead a very real possibility for the 21st century.”

WHO, 2014 (World Health Organization, 2014).

Very recently there have been reassuring advances in the discovery of entirely novel antimicrobials. One area of research is the identification of natural broad-spectrum antimicrobials, which could be readily added to our existing clinical arsenal. Examples of progress in this field include the ongoing characterisation of monotreme lactation protein (Enjapoori *et al.*, 2014; Newman *et al.*, 2018), a variety of candidates from frog skin secretions

including a particularly promising compound from *Pelophylax nigromaculatus* (Lu *et al.*, 2022), and the identification of a novel non-ribosomal peptide synthase (NRPS) products with strong antibiotic potential identified from uncultured bacteria (Ling *et al.*, 2015; Lam and Crawford, 2018). It is also probable exploration of other secondary metabolites could yield viable compounds (Onaka, 2017). Repurposing of existing drugs offers a faster route from the identification of novel antimicrobials to deployment in the clinic as they have already undergone clinical trials and are known to be safe. Progress is also being made in this field, for instance the identification of a psychoactive compound with an antimicrobial effect against *Salmonella enterica* serovar Typhimurium (Ellis *et al.*, 2019). The generation of novel antimicrobials by synthesis is also seeing progress, for instance the generation and screening of novel pyrazolidinone derivatives, some of which showed promise as lead compounds (Mokbel *et al.*, 2020).

However, utilisation of novel antimicrobial agents which are either bactericidal or bacteriostatic, will generate the same strong selective pressures toward resistance which drove the historic evolution and spread of AMR. Therefore, whilst the developments discussed prior are welcome steps forward, they are very much steps forward within the context of an ongoing arms race which we have been losing for over a decade. A number of alternative approaches have been suggested including immunotherapy, bacteriophage therapy, and anti-virulence strategies (Nicolle, 1952; Hotchkiss and Opal, 2010; Propst *et al.*, 2010; Rasko and Sperandio, 2010; Cisek *et al.*, 2017). A key limitation of all these approaches is their narrow spectrum of activity, with each novel agent affecting only one or a small subset of pathogens. This represents a clear clinical limitation, requiring accurate diagnosis prior to prescription and therefore perhaps leaving them as a secondary option after broad spectrum antimicrobials have failed. However, these alternative strategies do also offer several advantages. Firstly, they are thought to be less vulnerable to the development of resistance because, compared to bacteriostatic or bactericidal treatments, the selective pressure towards resistance is less acute. Secondly, they offer a great diversification of therapeutic options, so alternating and combinatorial therapeutic choices can be explored (Figure 1.1). They also fit well into the ongoing World Health Organisation's (WHO's) focus on a One Health approach, not least because of their reduced pressure toward resistance *ex vivo* (Pollock *et al.*, 2020).

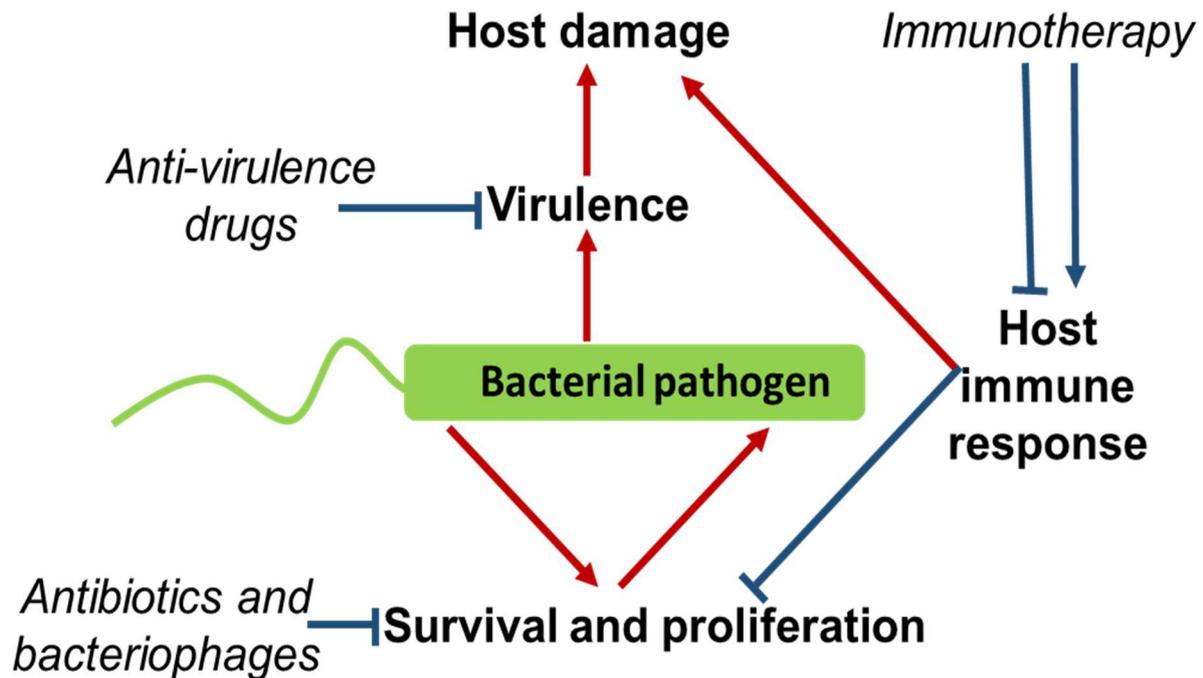


Figure 1.1 Schematic of future combined therapeutic approaches to bacterial infection. A simplified schematic representation of the cooperative and complementary nature of the novel and existing therapeutic strategies to overcome bacterial infection. Positive effects are depicted in blue, whilst negative effects are depicted in red. Arrows indicate causation and promotion, whilst blocking lines indicate inhibition, both together indicate modulation as required. Interventions are in italics.

Immunotherapy aims to beneficially modulate the patient's immune system, which can be either a promotion of the immune response or restricting it to prevent deleterious immune (over) responses such as cytokine storms. Examples include the addition of natural killer cells (Schmidt *et al.*, 2018) or the use of cytokines to modulate the immune response (Propst *et al.*, 2010). This field has been accelerated by Covid-19, with a plethora of new clinical options for immunotherapy being explored prior to the discovery of effective anti-viral drugs. This research however highlighted the complexities of immune intervention with patient variability and disease severity playing an important role in therapeutic effectiveness (Angriman *et al.*, 2021; van de Veerdonk *et al.*, 2022).

The use of bacteriophages as therapeutic agents is a historic practice (Lipska, 1951; Nicolle, 1952; Sibirtsev, Bel'skaia and Lavrova, 1953), which has fallen out of usage in much of the world. Strongholds of scientific knowledge on bacteriophage therapy gathered over the proceeding century exist in Poland and Georgia (Parfitt, 2005; Żaczek *et al.*, 2020), and there is renewed global interest in phage as an alternative or compliment to antibiotics (Cisek *et al.*,

2017; Gordillo Altamirano and Barr, 2019). Bacteriophage therapy shows promise, especially in conjunction with antibiotics (Altamirano *et al.*, 2022). However, a recent clinical review concluded that “high-quality trials are urgently required” for this promise to be actualised (Uyttebroek *et al.*, 2022), reflecting the lack of double blinding and placebo controls in many of the historic trials.

1.2 Anti-virulence – an effective therapeutic strategy

The goal of anti-virulence strategies is to disarm the pathogen, providing an opportunity for the host’s immune system to clear the infection (Rasko and Sperandio, 2010). Resistance to anti-virulence compounds has been observed both *in vitro* and *in vivo* (Smith *et al.*, 2012; García-Contreras *et al.*, 2013). However, the evolutionary pressures selecting for such traits are thought to be reduced because the prevention of proliferation, by the host immune system, is an indirect consequence of the compound (Rasko and Sperandio, 2010). The status of many virulence factors as “public goods” could also contribute to this effect (Davis and Isberg, 2019), as a newly resistant mutant would not receive the full reproductive benefit of resistance; it would be shared with the non-resistant members of the population. The presence of “cheater” cells which benefit from the populations virulence factors without producing them, including the type three secretion system of *Pseudomonas aeruginosa*, underlines this possibility (Czechowska *et al.*, 2014). Further, mutations conveying resistance to anti-virulence can impair the function of the virulence factor targeted and therefore are not as deleterious in terms of prognosis as might first be imagined (Tkaczyk *et al.*, 2018). Anti-virulence strategies can also eliminate the current issues of *ex vivo* selection for antibiotic resistance, caused by the environmental prevalence of antibiotics, particularly associated with agriculture and sewage (Zhang *et al.*, 2014; Xu *et al.*, 2015).

There are instances of virulence associated traits expressed and important for survival outside of an infection, implying the possibility of environmental pressure towards resistance to anti-virulence medications could be an issue. An interesting example of this is *Mycobacterium* species, in which environmental predation by Amoebae provided evolutionary pressure which drove adaptations later applicable in human pathogenesis (Salah, Ghigo and Drancourt, 2009). *Legionella pneumophila* also appears to have conserved traits for survival within Amoebae and human cells (Escoll *et al.*, 2013; Price *et al.*, 2014). With that said, virulence traits are generally more species specific than the core functions targeted by bactericidal antibiotics, and not all virulence traits are advantageous outside of infection. The issue of environmental selection is therefore reduced in all instances, and in some instances is entirely irrelevant.

An anti-virulence approach also leaves the host’s commensal bacteria unharmed, thus

minimising disruption to the host microbiome. This is vital because commensal bacteria, which are killed by conventional antibiotic treatment, can prevent colonization by nosocomial pathogens, notably Methicillin-Resistant *Staphylococcus aureus* MRSA (Yang *et al.*, 2018). Antibiotic treatment removing commensal bacteria has also been shown to leave patients more vulnerable to the causative agent of whooping cough, *Bordetella pertussis* (Zhang *et al.*, 2019). These are not isolated examples; an extensive literature on the phenomenon is now available, as previously reviewed (Zhang and He, 2015; Shi *et al.*, 2017).

The identification and targeting of proteins or structures either important or vital for virulence is therefore an important goal in current scientific research, with clear clinical relevance. Several types of target have been proposed for anti-virulence drugs. Firstly, toxins produced by pathogens can be targeted directly and inhibited, alternatively their delivery mechanism can be disrupted, finally signalling for and transcription of virulence factors can be suppressed (Rasko *et al.*, 2008; Sharifahmadian and Baron, 2017; Johnson and Gerding, 2019).

Human monoclonal antibodies targeting toxins have experienced some early success, for instance Bezlotoxumab, which is in clinical use to suppress recurrent *Clostridium difficile* infections (Johnson and Gerding, 2019). Similarly, a human monoclonal antibody targeting a *Bacillus anthracis* toxin has been approved for use in both the treatment and prevention of anthrax (Greig, 2016). This approach is therefore of proven validity; however, in instances where a greater range of less individually vital toxins are utilised by the pathogen a large number of antitoxins would be required for a significant effect. One pathogen for which this is true is *Pseudomonas aeruginosa*, as discussed at length below.

Where there is no single toxin or effector that can be selected to quench virulence, the mechanism by which they are delivered to the host can be a potential target. In Gram-negative bacteria secretion systems are generally employed for this function. Type IV Secretion Systems (T4SS), first characterised in plant pathogen *Agrobacterium tumefaciens*, are one potential target (Sharifahmadian and Baron, 2017). Attempts to develop inhibitors for human pathogens which depend on a T4SS have been made, including the discovery of some promising fragments which inhibit the T4SS of *Helicobacter pylori* (Arya *et al.*, 2019). Type VI secretion systems (T6SS) can target toxins directly into both prokaryotes and eukaryotes, and where the latter is the case can contribute to virulence. As such several T6SSs have been identified as targets, and both immunological and small molecule lead compounds have been found for further investigation (Sun *et al.*, 2014; Nguyen *et al.*, 2015). Given the dual eukaryotic-prokaryotic toxicity of some T6SS effectors, bacteria utilising these systems often express anti-toxins to prevent damage to their own cell. This has encouraged attempts to

interfere with the toxin-antitoxin interface and thus create species specific bactericidal agents which would apply a selective pressure toward disabling a virulence feature (Lu *et al.*, 2014; Gao *et al.*, 2017). Type three secretion systems (T3SS) also offer an excellent anti-virulence target in many organisms, as will be detailed at length below.

Virulence factors are often metabolically expensive; thus their production is tightly regulated. There are often common systems for the regulation of multiple virulence factors, therefore one target can allow the suppression of multiple virulence traits. A key example of this is quorum sensing (QS), which allows bacteria to gauge the population's cell density and commence the expression of virulence factors only once a sufficient population density has been reached (Rutherford and Bassler, 2012; Majerczyk *et al.*, 2014). The system works by the production of a signalling molecule which is released into the environment, and its subsequent detection. In the system first characterised, the LuxIR system on *Vibrio fischeri*, detection of the signalling molecule is performed directly by a deoxyribonucleic acid (DNA) binding transcription factor (Fuqua, Winans and Greenberg, 1994), however in some systems this occurs *via* a two component system. Two component systems consist of a histidine kinase and a response regulator, which when activated by the histidine kinase generally functions as a transcription factor (TF). QS systems have been the target of an extensive search for inhibitors, including the design and synthesis of signalling molecule analogues (Hodgkinson *et al.*, 2012; Morkunas *et al.*, 2012; Almohaywi *et al.*, 2018), the ongoing search for natural inhibitors (Branco-Vanegas *et al.*, 2014; Zhao *et al.*, 2019), and screening efforts both high throughput and virtual (D'Angelo *et al.*, 2018; Gatta *et al.*, 2019; Zhong *et al.*, 2020). A particularly notable example is LED209, an inhibitor of the QS two component system QseBC (Curtis *et al.*, 2014), which has been reviewed as an important target for some time (Rasko *et al.*, 2008; Rooks *et al.*, 2017).

Targeting QS, or two component systems involved in other areas of signalling, is not always distinct from the inhibition of TFs directly. When screens are conducted it is often unclear whether the stimuli detecting histidine kinase or the TF response regulator is targeted. For instance, a screening study conducted for inhibitors of the SaeR/S two component system, which found two compounds of significant interest, made no attempt to distinguish which protein was inhibited (Yeo *et al.*, 2018). Further, it was found several other TFs of relevance to virulence are inhibited by the compounds discovered (Yeo *et al.*, 2018). This is not the case during the discovery of inhibitors of AraC family TFs, which are frequently found to be virulence determinants, and are not directly connected to QS or two component systems.

The AraC family protein VirF in *Shigella* species is a key TF for virulence regulation, several

inhibitors which appear to bind to its DNA binding domain were identified in a high throughput screen and have been characterised, determining both their mode of action and demonstrating efficacy in cell culture models (Hurt *et al.*, 2010; Emanuele *et al.*, 2014; Emanuele and Garcia, 2015). Similarly, an AraC family TF responsible for the regulation of the toxin-coregulated pilus and cholera toxin in *Vibrio cholerae* was found was also discovered through high throughput screening (Hung *et al.*, 2005), and has since been shown to bind to a regulatory pocket in the protein, preventing dimerization (Shakhnovich *et al.*, 2007). Rational design has since provided more inhibitors for the same target and site (Woodbrey *et al.*, 2017). Further, DNA binding domain inhibitors for the related *Escherichia coli* MarA AraC family protein have been developed by *in silico* led structure-based drug design (Bowser *et al.*, 2007). The compound has since been modified and validated against the related LcrF, and important regulator of the T3SS in *Yersinia* species (Garrity-Ryan *et al.*, 2010). The same class of compounds, N-hydroxybenzimidazoles, has also been assessed for activity against *P. aeruginosa* master T3SS regulator ExsA, yet another AraC family TF important for virulence. They were found to have activity, although with different specificity from LcrF in a dual computational and mutagenesis assessment (Marsden, King, *et al.*, 2016).

1.3 *Pseudomonas aeruginosa* - a priority pathogen

Pseudomonas aeruginosa is a frequent target for anti-virulence drug development (Wagner *et al.*, 2016). It is an environmentally ubiquitous, Gram-negative, rod shaped γ -proteobacterium. The organism is present in aquatic, marine, and soil environments (Pirnay *et al.*, 2005; Deredjian *et al.*, 2014; Nair *et al.*, 2015). *P. aeruginosa* has a large and flexible genome (Olson *et al.*, 2000), with significant genetic diversity observed within limited geographic areas (Pirnay *et al.*, 2005). It can also be found in a variety of anthropological environments, including water distribution systems (Lu *et al.*, 2016), and hospitals (Anaissie, Penzak and Dignani, 2002).

This prevalence is concerning due to *P. aeruginosa*'s status as an opportunistic pathogen, particularly so because it is a common nosocomial infection. The bacterium was first observed in a clinical setting in the 19th century (Gessard, 1882), and the groups predisposed to infection include burns victims (Turner *et al.*, 2014), and cystic fibrosis patients (Murray, Egan and Kazmierczak, 2007). The pathogen is highly virulent, and consequently it is a leading cause of death in intensive care, especially among cases of ventilator associated pneumonia (Bassetti, Villa and Pecori, 2014). *P. aeruginosa* bacteraemia has reported mortality rates as high as 48% if treated ineffectively, and 14.3% with appropriate treatment (Ferreira *et al.*, 2014). Further, the pathogen is the leading cause of death among cystic fibrosis patients (Murray, Egan and Kazmierczak, 2007), and has been identified as a common coinfection in severe Covid-19 cases (Qu *et al.*, 2021). Its virulence combined with the difficulties associated

with treatment have seen it placed second on the WHO's list of priority pathogens (World Health Organisation, 2017).

Treating *P. aeruginosa* infections is complicated by its multiple drug resistances. As well as acquiring target mutations (Nakano *et al.*, 1997), the pathogen also possesses a suite of efflux pumps capable of exporting a wide range of antibiotics (Poole *et al.*, 1993, 1996; Köhler *et al.*, 1997; Aires *et al.*, 1999). Within the global *P. aeruginosa* population extended spectrum β -lactamases are present and proliferating, providing resistance to clinicians' previous drug of choice, carbapenem (Hopkins *et al.*, 2016; Mohanam and Menon, 2017). This potent arsenal of antibiotic resistance is further compounded by antibiotic tolerance, particularly that conferred by biofilms (Babin *et al.*, 2017), although "dormant" persister cells also contribute (Lewis, 2010). Given the severity of the threat posed by *P. aeruginosa* and the difficulty treating it with current antibiotics, investigation of the pathogen's virulence factors and subsequent development of anti-virulence drugs is an urgent priority.

1.4 The virulence strategies of *P. aeruginosa*

P. aeruginosa has two distinct virulent lifestyles which are causative of acute and chronic infections (Valentini *et al.*, 2018). Chronic infections can persist for multiple years, with the infection of cystic fibrosis patients' lungs providing the archetypal and most clinically important example. These infections are characterised by biofilm formation (Mulcahy, Isabella and Lewis, 2014), adopting a mucoid phenotype (Pritt, O'Brien and Winn, 2007), and expression of a T6SS (Chen *et al.*, 2015). The sites at which chronic infections are established are frequently polymicrobial, giving rise to a complex interplay of cross species signalling, cooperation, and competition (Bisht, Baishya and Wakeman, 2020). Over the long duration of a chronic infection, especially that of the cystic fibrosis lung, mutations accumulate within the *P. aeruginosa* population, tailoring its genome to the chronic infection lifestyle. During this process adaptations observed frequently include: acquisition of amino acid autotrophies, loss of acute virulence traits such as motility and the T3SS, inactivation of QS, and resistance to antibiotics (Winstanley, O'Brien and Brockhurst, 2016).

The two virulence lifestyle are reciprocally regulated by the GacS network, a complex system of TCSs, interacting orphan histidine kinases, and regulatory RNAs (Francis, Stevenson and Porter, 2017). Activation of the network triggers a chronic virulence phenotype and suppresses acute virulence traits, in part through downstream signalling by cyclic diguanylate monophosphate (c-di-GMP) (Francis, Stevenson and Porter, 2017). The network is perhaps most notable for the unorthodox interaction between histidine kinases which allows the integration of multiple signals and fine control (Francis *et al.*, 2018). Given the networks role

in activating chronic virulence traits and suppression of acute virulence it is perhaps unsurprising that a strain adapted to acute infections and isolated from a burn wound, PA14, has a frameshift mutation in LadS, a member of the GacS network (Mikkelsen, McMullan and Filloux, 2011).

Whilst virulence traits such as biofilm formation and T6SS are specific to chronic virulence, and others such as the T3SS are more implicated in acute virulence, there are a number of virulence factors utilised in both infection forms. These virulence factors are frequently under the control of QS, for which *P. aeruginosa* has four interconnected systems. The four systems are hierarchical, with a classical QS system LasIR upregulating the transcription of the other three systems (Lee and Zhang, 2015). Like other QS systems they allow *P. aeruginosa* to determine the population level and activate the expression of virulence traits only when the population is sufficiently dense for virulence strategies to be viable. As well as the universal virulence traits outlined below QS protein PqsR (also known as MvfR) has been found to upregulate T6SS expression (Maura *et al.*, 2016). It has also been reported that QS in *P. aeruginosa* downregulates the transcription of the T3SS, however there are a number of studies contradicting this finding with either an absence of QS regulation of the T3SS or even an apparent upregulation (Schuster *et al.*, 2003; Liang *et al.*, 2008; Zhang *et al.*, 2018). It has been suggested that the Ca²⁺ concentration in the media is causative of these discrepancies (Pena *et al.*, 2019). These discrepancies are significant because the Ca²⁺ concentration has a significant impact on the expression of the T3SS (Frank, 1997), and the different concentrations used would therefore have an impact on the result. It is clear that further study is required to elucidate the connection between QS and T3SS.

An example of a virulence factor not specific to one infection type is hydrogen cyanide (HCN) production. HCN been reported in both the chronic lung infections and in a model of acute infection in nematodes, where it appeared to be the main lethal virulence factor (Gallagher and Manoil, 2001; B *et al.*, 2008; Sanderson *et al.*, 2008). HCN production is regulated by QS in *P. aeruginosa*, as is the expression of a modified respiratory electron transport chain component conveying resistances to cyanide poisoning (Brint and Ohman, 1995; Yan *et al.*, 2019).

Similarly, pyocyanin is known to be expressed in both acute and chronic infections such as mouse models of pneumonia and in the cystic fibrosis lung (Lau *et al.*, 2004; Fothergill *et al.*, 2007). Production of pyocyanin is also regulated by QS (O'Loughlin *et al.*, 2013; Dong *et al.*, 2021), alongside other factors such as the metallothionein protein PmtA (Thees *et al.*, 2021). Pyocyanin is an exported secondary metabolite with severe cellular toxicity driven by the

generation of reactive oxygen species and the subsequent oxidative stress (Hall *et al.*, 2016). Pyocyanin can also facilitate the reduction of Fe³⁺ into soluble Fe²⁺ which *P. aeruginosa* can more readily uptake, assisting in iron acquisition (Cox, 1986; Sass *et al.*, 2021).

Another mechanism for iron uptake in *P. aeruginosa* which is considered a virulence factor is the production of siderophores, iron chelating molecules which are exported from the cell and actively transported back into the cytoplasm once bound to Fe³⁺. *P. aeruginosa* can produce two siderophores, pyoverdine and pyochelin, with the former having a higher affinity for Fe³⁺ than the latter (Albrecht-Gary *et al.*, 1994; Brandel *et al.*, 2012). Iron is an essential element for the bacterial growth and is scarce in many infection contexts, and that which is present is often bound within ferritin, haem, or similar molecules. It is therefore unsurprising that pyoverdine has been found to be essential in mouse burn infection models, and both pyoverdine and pyochelin contribute to virulence in other animal models of infection, though pyoverdine has the more significant impact (Meyer *et al.*, 1996; Takase *et al.*, 2000). *P. aeruginosa* also produces pyoverdine in the relatively iron rich environment of the cystic fibrosis lung, however there are examples of cystic fibrosis adapted strains which do not produce pyoverdine suggesting that it is not essential (Lamont, Konings and Reid, 2009). It is possible that in the cystic fibrosis lung pyoverdine is more important as a counter measure to fungal competitors than for iron acquisition *per se* (Sass *et al.*, 2017), although comparative data on the entire lung microbiome of *P. aeruginosa* infected patients with and without pyoverdine expression is not sufficiently available to demonstrate this. Siderophore synthesis in *P. aeruginosa* is, rather intuitively, regulated according to iron availability with circumstantial optimisation of the production of either pyoverdine or pyochelin (Cunrath *et al.*, 2020). QS also contributes to the regulation of pyoverdine synthesis, increasing pyoverdine production when QS signalling is active (Stintzi *et al.*, 1998). Whilst it is unclear how it fits into the broader regulatory framework of both virulence factors there is also at least one TCS which reciprocally regulates pyocyanin and pyoverdine metabolism (Little *et al.*, 2018).

Siderophores in *P. aeruginosa* are in part synthesised by NRPS proteins. NRPSs are ribosomal proteins which can create a specific peptide product in a production line like manner, often incorporating amino acids or chemical alterations not available to peptides produced by the ribosome. As well as siderophores NRPSs can produce other virulence factors with more direct host damaging effects, such as the *E. coli* NRPS product colibactin (Nougayrède *et al.*, 2006). It is therefore intriguing that *P. aeruginosa* has multiple NRPS operons which are either uncharacterised or have been subjected to only a very limited investigation (Gulick, 2017). One such NRPS has been shown to be within the QS regulon, which implies a link to virulence, however mutants deficient in this NRPS has increased virulence in a *Galleria mellonella*

infection model (Hong *et al.*, 2019).

The intensive energy demands of *P. aeruginosa* during infection are underwritten by metabolic diversity, allowing the pathogen to adapt to the divergent environmental and infection niches which it can occupy. For instance, *P. aeruginosa* can grow on a variety of carbon sources, utilising several divergent pathways around the conserved core of carbon metabolism to facilitate this flexibility (Dolan and Welch, 2018; Dolan *et al.*, 2020). Similarly, the *P. aeruginosa* can utilise alternative terminal electron acceptors. Rather than oxygen it can switch to utilising nitrite and/or nitrate, a process known as denitrification which allows respiration in anaerobic environments (Carlson and Ingraham, 1983).

The relevance of denitrification to virulence is highlighted by its regulatory association with QS, and the coregulation of denitrification proteins and virulence factors by the oxygen sensing TF Anr (Hammond *et al.*, 2015). Whilst denitrification is often thought of as a “backup” system for anaerobic conditions *P. aeruginosa* can utilise nitrite and nitrate as terminal electron acceptors alongside oxygen in lower oxygen conditions (Chen, Xia and Ju, 2006). A proteomic investigation of the changes in expression of PAO1 and cystic fibrosis clinical isolates between aerobic and hypoxic cultures found that proteins required to utilise oxygen were upregulated alongside denitrification proteins and a fermentation pathway in hypoxic cultures (Kamath *et al.*, 2017). The production of NO during denitrification is also a signal for acute virulence expression. Mutants unable to perform denitrification express T3SS proteins at a lower level and have reduced virulence against *Caenorhabditis elegans* and human cells – a phenotype which can be restored by the introduction of exogenous NO (van Alst *et al.*, 2007).

In contrast to chronic infections, acute infections are characterised by a motile, planktonic lifestyle with the expression of a range of toxins – especially those associated with the T3SS (Arora *et al.*, 2005). A minority of acute infections are caused by strains of *P. aeruginosa* which lack a T3SS, and instead express the toxin exolysin A. However, these strains are both less virulent, and only cause a small fraction (1.5%) of cases (Medina-Rojas *et al.*, 2020). Other, non-T3SS dependent toxins and effectors, such as protease IV (Engel *et al.*, 1998), the type 1 secretion system (T1SS) metalloprotease ImpA (Tian *et al.*, 2019), and PlcH and PlcN phospholipases (Voulhoux *et al.*, 2001), are known to impact virulence. However, the pivotal role of the T3SS well documented for burn infections (Holder, Neely and Frank, 2001), bacteraemia and systemic infection (Vance, Rietsch and Mekalanos, 2005), and acute pneumonia (Lee *et al.*, 2005).

1.5 The T3SS of *P. aeruginosa* - a potent anti-virulence target

The T3SS is a needle like multi-protein complex which allows the delivery of toxins, commonly referred to as effectors, into the host's cells (Figure 1.2). The T3SS spans both *P. aeruginosa*'s inner and outer membranes, forms a translocon pore in the host membrane, and proceeds to deliver effector. *P. aeruginosa*'s T3SS effectors ExoT, ExoY, ExoS, and ExoU are well characterised (Shaver and Hauser, 2004), and others (PemA and PemB) have been identified (Burstein *et al.*, 2015). It has also been reported that the translocon pore itself causes damage to the host cells (Dortet *et al.*, 2018), although the possibility of the observed host damage being caused by novel effectors was not conclusively disproven. The importance of individual T3SS effectors for virulence has led to ExoU's identification as a prospective drug target (Foulkes *et al.*, 2019), however it is apparent that inhibiting the entire apparatus is preferable to inhibiting a minority of the effectors it translocates.

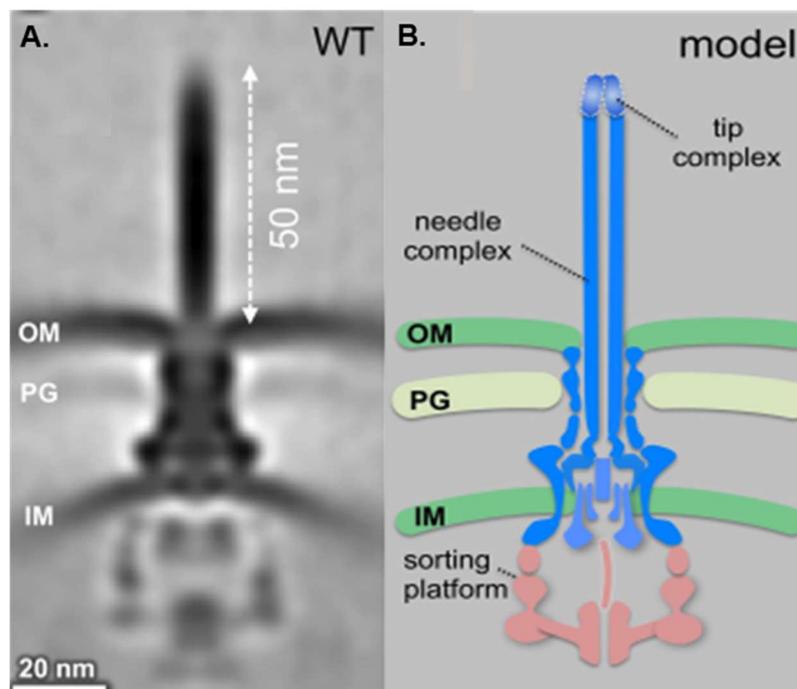


Figure 1.2 The structure of a T3SS.

A. A electron microscopy image of the *Salmonella enterica* serovar Typhimurium T3SS, which has a highly similar structure to that of *P. aeruginosa*.

B. A cartoon representation of T3SS to illustrate its components.

OM, PG, and IM indicate the outer membrane, peptidoglycan, and inner membrane respectively. Both images are modified from (Park *et al.*, 2018).

Multiple mechanisms for T3SS inhibition are currently under development. The protein PcrV, which is part of the T3SS's tip for insertion into the host membrane, is a common target due to its relative exposure and proven importance. Humanised monoclonal antibodies targeting

PcrV have shown promising results in mouse and rabbit infection models (Sawa *et al.*, 2014), and has been demonstrated to be safe for use in humans (Jain *et al.*, 2017). The clinical utility remains unproven, possibly due to the choice of cystic fibrosis patients in clinical trials, rather than patients with acute *P. aeruginosa* infections. Other attempts to target PcrV with antibodies are also underway (Ali *et al.*, 2018; Tabor *et al.*, 2018).

Chemical inhibitors of the T3SS have also been reported, including MBX 1641, which targets T3SS protein PscF (Williams *et al.*, 2015). However, a single codon mutation was sufficient to cause resistance to this inhibitor (Bowlin *et al.*, 2014), indicating the possibility of resistance in a clinical setting. In a more recent study, the compound was observed to reduce T3SS protein ExoT secretion in culture but had little effect murine infection models (Aburto-Rodríguez *et al.*, 2021). A novel small molecule inhibitor of the *P. aeruginosa* T3SS has also been reported, with convincing *in vivo* data to support the suppression of virulence. However, the only direct evidence of reduced T3SS activity was a somewhat messy western blot with no loading control, and no mechanism of action was reported (Sheremet *et al.*, 2018). More generally the natural product cinnamaldehyde has been identified as a T3SS inhibitor in *S. typhimurium* (Liu *et al.*, 2019), however the compound's broader antimicrobial function and status as an allergen make it a poor choice as an anti-virulence therapeutic (Doyle and Stephens, 2019), as does its tendency to increase the expression of efflux systems (Tetard *et al.*, 2019). Similarly, the T3SS of a variety of plant pathogens, including *Pseudomonas syringae*, has been shown to be inhibited by benzyloxy carbonimidoyl dicyanide derivatives very convincingly. However, no mode of action was demonstrated, and there has yet to be any further investigation exploring the compounds applicability to human pathogens (Ma *et al.*, 2019). In short despite a plethora of attempts to target T3SSs, of which only a selection are presented here, no inhibitors have yet been approved for use, and few have undergone clinical trials.

An alternative strategy to disable the T3SS is targeting its regulation, rather than to directly impairing its function. If both strategies came to fruition they could be used synergistically, thus rendering the T3SS either absent or disabled if the regulatory inhibition was bypassed – thereby further diminishing the already reduced risk of resistance. In *P. aeruginosa*, expression of the T3SS and its effectors are controlled by the master regulator ExsA (Brutinel *et al.*, 2008). Inhibition of ExsA therefore offers a direct mechanism for inhibiting the expression of the T3SS and its effectors, making it a compelling anti-virulence target.

1.5 ExsA – a target for the upstream regulation of the T3SS

ExsA is a member of the AraC family of transcription factors and consists of two domains. The

C terminal domain (CTD) is the functional domain; it binds to a specific promoter region of DNA. The N terminal domain (NTD) is responsible for dimerization and the regulatory interaction with the anti-activator ExsD (Brutinel *et al.*, 2008). The NTD also contains an ancestral ligand binding pocket without a characterised function within ExsA. ExsA is regulated by a non-canonical partner switching cascade (Figure 1.3). This cascade operates as follows: in a “T3SS off” state ExsE is bound to ExsC and ExsD is bound to ExsA, contact between a host cell and basally expressed T3SSs triggers the export of ExsE (Urbanowski, Brutinel and Yahr, 2007); this releases ExsC. The liberated ExsC subsequently binds ExsD, the cognate anti-activator of ExsA (Brutinel, Vakulskas and Yahr, 2010). Thus liberated, ExsA binds its cognate promoter regions, activating T3SS transcription and rapidly increasing the number of these structures in the cell (Zheng *et al.*, 2007).

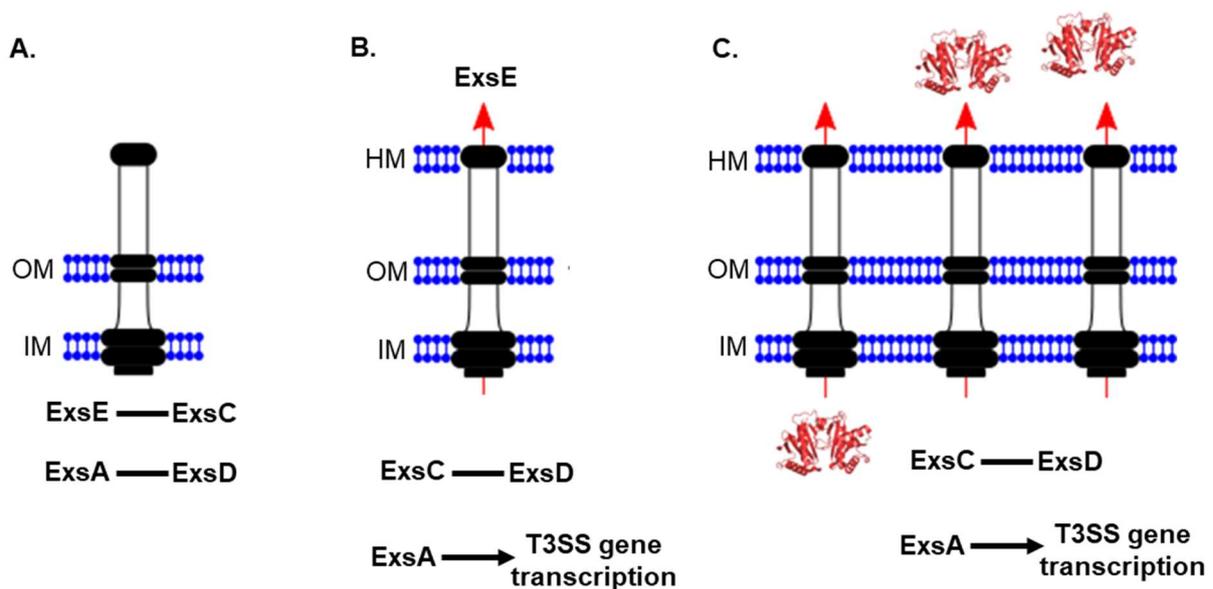


Figure 1.3 The function of the ExsA cascade.

Three cartoon images depicting the sequential stages of ExsA activation. The bilayers labelled IM, OM, and HM indicate the bacterial inner membrane, outer membrane, and host membrane respectively. In **A.** the T3SS is ‘idle’ and the ExsA cascade is in a stable inactivated state. In **B.** the red arrow indicates the translocation of ExsE, resulting in partner switching which allows ExsA to activate T3SS transcription. Panel **C.** shows the proliferation of T3SSs and the export of the effector ExoT (shown as PDB structure 4JMF).

A wider view, both in terms of function and regulation, of ExsA beyond this classical depiction is emerging. For instance, the TF Vfr has been demonstrated to directly promote the transcription of *exsA* (Marsden, Intile, *et al.*, 2016), and RplI regulates the translation of ExsA (Wang *et al.*, 2022). There is also an unknown mechanism by which ExsA expression is

heterogeneous within a *P. aeruginosa* population, with some cells being primed for T3SS activation due to high expression levels of ExsA and others taking longer to respond to activating signals (Lin *et al.*, 2021). The known regulon of ExsA has also expanded beyond the transcription of T3SS genes. ExsA has previously been identified as the transcription factor for at least one virulence factor that is independent of the T3SS, the T1SS metalloprotease ImpA (Tian *et al.*, 2019). It has also been implicated in both positive and negative regulation of a variety non-T3SS genes *via* a chIP-seq screen, notably *phrS*, which encodes a small RNA linked to oxygen availability and QS (Huang *et al.*, 2019). The fact that negative regulation by ExsA has been discovered is yet to be explored to its full potential and does not feature in much of the literature on the protein.

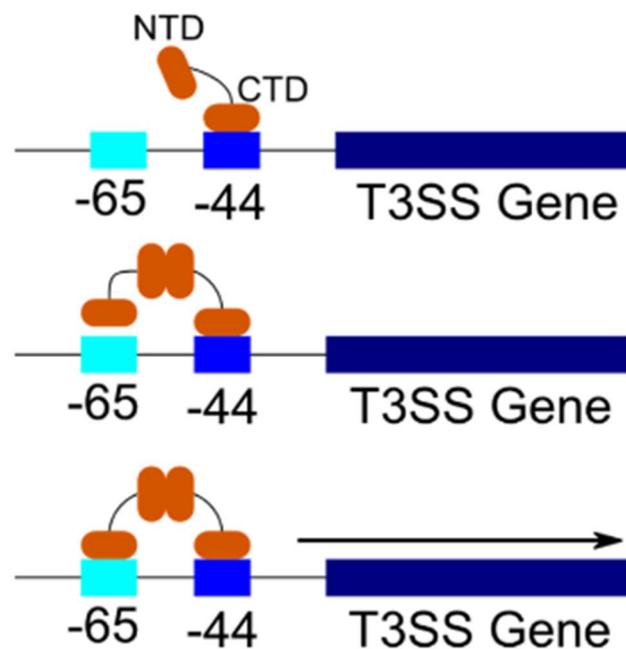


Figure 1.4 ExsA's dimerization and DNA binding.

A cartoon representation of the ExsA's DNA binding mechanism. Each row represents a sequential stage, with boxes showing specific regions of DNA as indicated. The -44 and -65 boxes indicate ExsA binding sites and their approximate position to the start of the T3SS gene.

Within the cytoplasm ExsA exists as a monomer (when not bound to ExsD), it binds to DNA as follows: a helix turn helix motif in the CTD binds to a promoter region which overlaps the -35 RNA polymerase site upstream of T3SS genes (Figure 1.4) Then the NTD facilitates dimerization, and a second ExsA protein then binds another, less well conserved site further upstream (Brutinel *et al.*, 2008). Whether the dimerization facilitates binding to this lower

affinity site, or lack of dimerization sterically hinders binding is uncertain, and seems to vary between promoter regions (Marsden, Schubot and Yahr, 2014). The interruption of any stage in this process would prevent the transcription of the T3SS genes, and thereby significantly impair *P. aeruginosa*'s virulence. The viability of this strategy was demonstrated in a murine infection model in which *P. aeruginosa* was subjected to an ExsA knockdown via Mobile-CRISPRi resulting in a marked increase in recovery rate (Qu *et al.*, 2019).

As mentioned prior there is a characterised class of compounds, N-hydroxybenzimidazoles, which disrupt the CTD-DNA interface of ExsA (Marsden, King, *et al.*, 2016), however due to the conserved nature of the CTD, the compounds also affect other AraC family members (Bowser *et al.*, 2007). This has clinical implications, not just because of the enormous variety of possible effects on other organisms, but also because inhibition of two other *P. aeruginosa* AraC family proteins, VqsM and SphR, is likely to increase virulence (LaBauve and Wargo, 2014; Liang *et al.*, 2014; Okino and Ito, 2016). To nullify this issue, molecules with greater specificity should be sought. This line of research is already being undertaken within the laboratory.

The NTD could be targeted with three inhibitory strategies: mimicking ExsA's antiactivator ExsD, blocking the dimer interface, and targeting the ancestral ligand binding pocket. Mimicking the activity of ExsD is theoretically possible, a molecule which bound to the interface on ExsA, or which stimulated a similar effect through another mechanism, would facilitate inhibition. However, at present there is no structure available for the ExsD-ExsA complex, and the only ExsD structure published is in the form of a homotrimer (Bernhards *et al.*, 2009), and thus there is a possibility of significant conformational change. Without such a structure this approach is not viable, an issue which will be engaged with within this research.

Targeting the dimer interface is also complicated by the lack of a dimeric structure, however a symmetry mate of the NTD crystal structure offers a reasonable approximation and is supported by comparison with other AraC family dimers (Shrestha *et al.*, 2015). Further, key residues within this interface have been identified previously via mutagenesis (Marsden, Schubot and Yahr, 2014; Shrestha *et al.*, 2015). It is therefore possible to computationally predict compounds which will disrupt this interface.

Finally, the NTD's ligand binding pocket offers a possible route for the inhibition of ExsA dimerization. In *Vibrio cholerae*, the ExsA homolog ToxT has several known inhibitors, all of which bind within the homologous pocket (Shakhnovich *et al.*, 2007; Lowden *et al.*, 2010; Anthouard and DiRita, 2013). The possibility of the same class of inhibitors functioning for

ExsA, and of targeting the homologues ExsA pocket with novel compounds, have been explored within the present study.

1.6 *In silico* drug discovery

The discovery of novel therapeutic agents *in silico* is often subject to scepticism by more traditional scientists. The reader will have noted the numerous success stories cited above, as such I will not present a detailed argument for their utility here. In short computational methods allow for the testing of far more compounds at a fraction of the cost, time, and labour *vis a vis* traditional screening systems. Computational methods can facilitate both rational design and virtual screening (VS) of extensive libraries for the discovery and improvement of lead compounds and therapeutic agents. When considering modification or rational design, modelling the potential ligand-receptor interactions, and comparing potential variants predicted binding affinities can inform the choices made. These two approaches are often overlapping, with screening often employed to find commercially available compounds which fulfil rationally designed criteria (Matsoukas *et al.*, 2015).

VS is somewhat comparable to assay-based high-throughput screening methods, where a library of molecules is screened for activity either *in vitro* or *in vivo*. Both methods have comparative advantages and disadvantages. High-throughput screening has greater accuracy; a hit indicates a biological effect rather than a prediction as in VS. However, VS allows a greater number of molecules to be screened in a fraction of the time and at a negligible cost. Whilst VS results contain more false positives and negatives, and require experimental validation, the efficiency of the method can allow a greater chance of success for the time and funding invested.

There are two main VS methodologies: ligand-guided or structure-guided. Ligand-guided methods are utilised to find ligands similar to a known active compound, whilst structure-guided methods are used to dock ligands to potential receptors *in silico*, referred to from here on as docking. The methods are often used sequentially. For instance, novel modulators of the human acid-sensing ion channel 3 were discovered by employing a ligand-oriented approach, after experimental verification the ligands were further characterised in docking experiments (Callejo *et al.*, 2020). Similarly, pharmacophore hypotheses can be generated from known active compounds, and used to filter compound libraries prior to VS employing docking methods (Upadhyay, Gajjar and Suhagia, 2018). Conversely, a study which reported novel inhibitors of IdeR, a *Mycobacterium tuberculosis* TF, first employed a docking method for VS, then used a ligand-oriented method on experimentally validated hits to find additional active compounds (Rohilla, Khare and Tyagi, 2017).

1.6.1 Ligand-guided virtual screening

Ligand-guided VS takes a known ligand and finds similar compounds either in terms of shape, electrochemical properties, or preferably both. Overlaying the volume of two compounds allows for a direct comparison of 3D shape; whilst this is theoretically simple, in practice the exact conformation of both compounds and inclusion of other chemical information thought relevant to the pharmacophore can generate significant complexity (Kirchmair *et al.*, 2009). The complexities of the non-covalent interactions between a receptor and ligand mean the chemical groups on the query or bait compounds create additional issues. Two of the mechanisms to solve this problem are considering pharmacophore features, e.g. hydrogen bond donors, as a sphere and accepting any equivalent groups which is positioned near enough to the original. This is the approach taken by the ROCs software in its colour score methods (OpenEye Scientific Software). An alternative approach is to examine projected electrostatic fields, as with the EON software (OpenEye Scientific Software). This approach has been found to perform better in some systems (Mishra and Basu, 2013), but it is worth noting that software's accuracy can vary between ligand-receptor systems. Ligand-guided approaches are often considered more reliable than structure-guided approaches, which is supported by a comparative study (Hawkins, Skillman and Nicholls, 2007). Similarly, a large comparative study found a joint ligand-structure approach (HYBRID, OpenEye Scientific Software) to be more accurate than its purely structure-guided equivalent (FRED, OpenEye Scientific Software) (Li *et al.*, 2018). However, the obvious drawback is that to employ a ligand-oriented approach you must first have a known ligand and, ideally, structural information regarding both the conformation of the bound ligand and the relevant ligand-receptor interactions.

1.6.2 Structure-guided virtual screening

Structure-guided VS or large-scale docking uses a receptor structure, either generated by homology modelling or an experimentally determined, real structure, and the software predicts the likely binding pose of a library of ligands as well as providing a score for each. Most VS cited thus far relies on focus docking, where a target pocket is selected and ligands are assessed for compatibility. Blind docking is an alternative approach in which the entire receptor structure is considered, this is useful to either find the binding location of a compound known to bind that target when structural information of their interaction is limited or absent (Greenhalgh *et al.*, 2020; Zarkan *et al.*, 2020), and to analyse the results of focused docking to remove "sticky" compounds which bind with an apparent high affinity but lack specificity to the target pocket. As more receptor surface area is considered, these approaches are often significantly slower. The flexibility of both the ligand and receptor are also considered to

varying degrees, with greater flexibility allowing for greater accuracy at the expense of additional computing time. Generally fast methods are used to create a shortlist, and more accurate methods are then used to select the final compounds to be tested.

Benchmarking exercises provide a useful guide to the efficacy of different software and algorithms. In these studies different software is used to recreate known receptor-ligand complexes, and the ability of each method to do so is assessed and compared. The main limitation of benchmarking studies is the divergence in data sets and methodologies employed, which limits the ease of comparison between studies investigating different sets of programs or providing with divergent results.

Multiple factors contribute to the quality of a docking software, which are reported separately upon to varied degrees in the literature. The main two functions are pose prediction of ligand pose and its scoring. They are generally assessed by the accuracy, assessed by root mean squared deviation (RMSD) of the docked pose of the ligand with respect to its real pose, with which a crystallographic pose is recreated, and how well the correct pose is scored compared to other binding modes respectively. GOLD (Jones *et al.*, 1997) and Autodock Vina (Trott and Olson, 2009) both performed very well in a thorough and repeatable benchmarking exercise (Li *et al.*, 2014), with two GOLD scoring methods, namely ChemPLP and ChemScore performing best in reproducing the crystallographic pose in the original study, whilst Autodock Vina scored even higher when assessed by the same criteria (Gaillard, 2018). The same software, and another named LeDock (Zhao and Caflich, 2013), also performed exceptionally well in another benchmarking test (Wang *et al.*, 2016). It has been demonstrated that using multiple approaches yields better results (Li *et al.*, 2014), therefore either constructing a pipeline or simply utilising multiple software is desirable.

1.6.3 Molecular dynamics simulations

Molecular dynamics (MD) simulations, which are *in silico* simulations encompassing every atom within a protein, are becoming increasingly accessible, both in terms of cost and ease of use (Hollingsworth and Dror, 2018). They allow explicit consideration of the flexibility of the target protein structure into consideration whether that structure is determined by crystallography or modelling. This dynamic picture of a proteins structure can be used to inform inhibitory strategies, or simply to improve our biological understanding of a protein (Medarametla *et al.*, 2021; Yang *et al.*, 2022). Used in conjunction with either docking methods as outlined above, or experimental data, they can also greatly enhance our understanding of ligand-protein interaction (Adla *et al.*, 2021).

1.7 Aims and objectives

ExsA appears to be a prime target for the development of much needed novel anti-virulence therapeutics against *P. aeruginosa*. However, questions remain as to the full extent of the TF regulon, and which target sites are most likely to be efficacious. It is also apparent that targeting the CTD is likely to lead to non-specific compounds liable to have an affect against the wider AraC family.

The overarching aim of this investigation was therefore to further investigate the regulon of ExsA and begin the discover of inhibitory compounds for ExsA, which do not target the highly conserved DNA binding domain, and thus are highly specific. To this end the following had been undertaken:

- The creation of deletion mutants of ExsA cascade proteins and phenotypic characterisation of the mutants.
- A proteomic investigation of said mutants to explore the wider ExsA regulon.
- Investigation of an inhibitor of ToxT, a homologous protein.
- Identification of closely related homologues, and their known ligands.
- Structure-guided VS screening to identify novel ExsA inhibitors.
- Both *in vivo* and biophysical testing of identified putative inhibitors.
- The development of assays and methods for functional analysis of ExsA inhibitors.
- Modelling and molecular dynamics simulation of ExsA, providing an optimal structure of ExsA for future use.

2. General methods

2.1 General *in silico* methods

2.1.1 Preparation and analysis software (Ftsite, MMV, ICM)

FTSite (Ngan *et al.*, 2012) was accessed through the dedicated web portal: <https://ftsitesite.bu.edu/>. All protein structures were prepared in PyMOL, including the removal of any unnecessary molecules from the file and alignment to other structures for the purpose of comparison, both of which were conducted prior to submission.

The Molegro Molecular Viewer (MMV) (<http://molexus.io/molegro-molecular-viewer/>) was used to prepare all ligands for docking. This entailed the addition of H atoms and energy minimisation.

ICM Pro version 3.8 (Molsoft Inc.) was used to prepare receptors prior to docking, which involved energy minimisation to remove steric clashes and addition of H atoms that are not usually assigned in experimental structures

2.1.2 Focused docking using GOLD suite

The GOLD version 5.3 (CCDC, Cambridge) suite was used as the predominant focus docking software (Jones *et al.*, 1997). The target site was selected via coordinates extracted from the receptor file, and the software's protein preparation wizard was used as per the vendor's guidance. A large pocket was allowed, 15 Å from the given coordinate to allow enough search space for the ligand in the given poses. GoldScore, ChemPLP, and ChemScore were used to sort the results, and where indicated were combined into a Z score for each compound by taking the mean of the Function Z scores for each scoring algorithm, which were in turn calculated by the following equation: $\text{Function Zscore} = (\text{score} - \text{mean}) / \text{standard deviation}$. The pose corresponding to the highest score was inspected in PyMOL with reference to the receptor and any relevant poses from other software such as AutoDock Vina or ICM Pro.

2.1.3 Blind docking Vina

Autodock Vina version 1.1.2 was employed for the majority of blind docking (Trott and Olson, 2009). Docking in Autodock Vina was conducted with 100 replicates per compound unless otherwise specified. Energy minimised ligands in an .SDF written by MMV were then prepared for Vina using OpenBabel to convert inputs to pdbqt files. The latter format is required for docking in AutoDock Vina. Blind docking was undertaken at exhaustiveness 16 with a maximised search grid covering the entire protein structure and a randomly generated seed. The top 100 scoring results were extracted and analysed using the R script available in

Appendix 1. The relevant pose or poses were then extracted and visualised in PyMOL with reference to the receptor and any relevant poses from other software.

2.1.4 Ligand oriented screening with Rocs

OpenEye's ROCS version 3.2.1.4 (Lantz, 2005) was employed for ligand-guided screening. Any adjustments made to raw inputs were conducted in the program's query adjustment suite (vROC) and the top 500 hits as output from ROCS-based VS were saved. These hits were ranked by descending order of their Tanimoto Combo score were saved (Naylor *et al.*, 2009).

2.2 Molecular microbiology

2.2.1 Bacterial strains and growth conditions

Unless otherwise specified all bacterial cultures were incubated at 37°C, with 180 revolutions per minute (rpm) for liquid cultures. Either Luria-Bertani media (LB) (Thermofisher Scientific, USA), with 1.5% (w/v) agar where appropriate, or Alanine glycerol salt yeast extract (AGSY) media (56 mM L-alanine, 17 mM K₂HPO₄, 86 mM NaCl, 3 g Oxoid yeast extract (Thermofisher Scientific, USA), 100 µM CaCl₂, 10 mM MgSO₄, 5 µM FeCl₂, 7.5 µM ZnSO₄, and 0.5% (v/v) glycerol), were used as indicated. Antibiotics gentamycin (*E. coli* = 20 µg/ml, *P. aeruginosa* = 50 µg/ml), carbenicillin (50 µg/ml), and chloramphenicol (34 µg/ml) were used were noted. All antibiotics were dissolved in deionised water, except chloramphenicol for which ethanol was use, sterilised by 0.22 µm syringe filter units (Millipore), and stored at -20°C prior to use. General strains used are recorded in Table 2.1.

Table 2.1 Frequently used bacterial strains

Organism	Strain	Source
<i>Escherichia coli</i>	DH5α	Welch lab stocks
<i>Escherichia coli</i>	Rosetta (Cm ^R)	Novagen
<i>Staphylococcus aureus</i>		Welch lab stocks
<i>Pseudomonas aeruginosa</i>	PAO1	Welch lab stocks
<i>Pseudomonas aeruginosa</i>	PAO1 P _{pcrV} -Lux (Gent) ^{**}	Dr Stephen Dolan
<i>Pseudomonas aeruginosa</i>	PAO1 P _{lacZ} -Lux (Gent ^R) [*]	Dr Stephen Dolan
<i>Pseudomonas aeruginosa</i>	PAO1 P _{nirS} -Lux (Gent ^R) [*]	Dr Stephen Trigg
<i>Pseudomonas aeruginosa</i>	YM64	(3)

* Luciferase reporter strains were created by the chromosomal integration of pUC18-mini-Tn7T-Gm-lux modified to include the indicated promoter.

+ The *popN* promoter, which is at the start of the operon containing *pcrV*, was used. *PpcrV* is referred to as for consistency with other assays within the study.

2.2.2 Storage of bacterial cultures

Short term storage of bacteria was conducted by streak plating onto LB + 2% w/v agar with selection as necessary. Plates were incubated at 37°C overnight, then sealed with parafilm and stored at 4°C for all bacteria except *P. aeruginosa* which was kept at room temperature. No plates were used after more than 5 days of storage.

For long term storage glycerol stocks were utilised. A 1:1 ratio of sterile 50% glycerol and overnight culture were mixed in a cryovault tube and stored at -80°C. Revival of samples was undertaken by streaking onto an LB agar plate with appropriate selection followed by overnight incubation.

2.2.3 Molecular cloning

2.2.3.1 Polymerase chain reactions

Polymerase chain reactions (PCR) for molecular cloning were undertaken with the following reaction mixtures at a total volume of 50 µl: 1 µl Q5[®] High-Fidelity DNA Polymerase (0.02 Units/µl, New England Biolabs, USA), 10 µl 5X Q5 Reaction Buffer (New England Biolabs), 10 mM dNTPs (Invitrogen, USA), 10 µM of both the forward and reverse primers, and 60 ng of genomic template DNA unless otherwise noted. Unless otherwise specified reactions were conducted in a thermal cycler as follows: 98°C for 30 seconds, followed by 30 replicates of 98°C for 10 seconds and 72°C for 35 seconds with an additional 15 seconds for every 500 base pairs in length over 500. Finally, the reactions were held at 72°C for 120 seconds before being cooled to and stored at 10°C until removal from the thermal cycler and storage at 4°C. All reactions took place in a Veriti Thermal cycler. It is noted that a two step PCR was employed due to the high T_m of the primers caused by *P. aeruginosa*'s GC rich genome.

Samples for colony PCR were prepared by resuspending a colony in 10 µl of deionised H₂O. The reactions were undertaken with the following reaction mixtures to a total volume of 20 µl: 0.1 µl *Taq* DNA Polymerase (1.25 units, New England Biolabs), 4 µl 5X Standard *Taq* Reaction Buffer (New England Biolabs), 200 µM dNTPs (Invitrogen, USA), 0.2 µM of both the forward and reverse primers, and 1 µl of resuspended colony. Reactions were conducted in a thermal cycler as follows: 95°C for 5 minutes, followed by 30 replicates of 95°C for 20 seconds and 68°C for 70 seconds with an additional 30 seconds for every 500 base pairs in length over 500. Finally, the reactions were held at 68°C for 5 minutes before being cooled to and stored at 10°C

2.2.3.2 Agarose gel electrophoresis for DNA

PCR reaction products were analysed separation according to length *via* gel electrophoresis.

Gels were made from 1% weight/volume agarose (Melford, UK) dissolved in TAE buffer (40 mM Trizma[®] base (Sigma), 20 mM acetic acid (Honeywell, Germany), 1 mM Ethylenediaminetetraacetic acid (EDTA) pH 8) with gentle heating and manual agitation. Once the solution had cooled sufficiently to cease emitting vapour ethidium bromide was added to a final concentration of 0.5 µg/ml. Gels were cast and run submerged in TAE buffer in a Wide Mini-Sub Cell GT Horizontal Electrophoresis System (Bio-Rad). DNA hyperLadder (10 kb, Bionline) was loaded alongside samples to provide reference product lengths. Gels were run at 100 V until the dye front approached the end of the gel, taking approximately 90 minutes, and visualised by UV transillumination.

2.2.3.3 Gel extraction

To isolate a specific PCR or endonuclease restriction product agarose gels were performed as above. The desired product was identified by its size with reference to the ladder and excised with a scalpel on a UV transilluminator; eye protection was used throughout. The excised section of gel was purified with a GeneJET Gel Extraction Kit (Thermo Scientific) in accordance with the manufacturer's instructions.

2.2.3.4 Purification and isolation of DNA

Genomic DNA isolation, PCR clean ups, and plasmid purifications were performed using the GeneJET Genomic DNA Purification Kit (thermo scientific), GeneJET PCR Purification Kit (Thermo Scientific), and GeneJET Plasmid Miniprep Kit (Thermo Scientific) respectively, in accordance with the manufacturer's instructions. All DNA concentrations were determined NanoDrop ND-100 spectrophotometer.

2.2.3.5 Restriction endonuclease digestion

Restriction digests were performed in 50 µl reactions with the High-Fidelity (HF[®]) Restriction Endonuclease (New-England Biolabs) with 1X CutSmart[®] Buffer (New-England Biolabs) for 90 minutes at 37°C. Two high-fidelity restriction endonucleases were used per reaction, both with a 2 µl per 50 µl reaction. Unless the product was immediately isolated the reaction was stopped by heating to 80°C for 5 minutes. Where required digestion was confirmed by agarose gel electrophoresis.

2.2.3.6 Ligation

Ligation reactions were undertaken for 17 hours at room temperature, with T4 DNA Ligase (New-England Biolabs) with 1X T4 DNA Ligase Buffer (New-England Biolabs). Insert DNA was provided at a 5:1 molar ratio to the digested plasmid DNA. Ligations were stored at -20°C prior to transformation.

2.2.3.7 Transformation by electroporation and chemically competent cells.

Two types of *E. coli* transformation were undertaken, transformation of ligations and transformation of confirmed plasmids. Transformation of ligation products was undertaken with chemically competent dh5 α *E. coli* cells, as previously described (Green and Rogers, 2013), except in the case of the novel deletion constructs described in Chapter 3.

Transformation of plasmid DNA into *E. coli* by electroporation was initiated by growing *E. coli* DH5 α overnight in LB media. The culture was diluted 1:100 in LB media and grown to an Optical density at $\lambda = 600$ nm (OD₆₀₀) of 0.3, after which it was chilled on ice for 30 minutes, then pelleted by centrifugation at 3200 x *g* at 4°C for ten minutes. The supernatant was discarded, and the cell pellet was washed three times by resuspension in 10 ml of ice cold 10% glycerol and subsequent pelleting. Finally the cells were resuspended in 1 ml of ice cold 10% glycerol, and 100 μ l of the resulting suspension was added to 4 μ l of ligation mixture in an ice-cold electroporation cuvette. Electroporation was then conducted in an Eppendorf Electroporator at 2.5 kV (25 μ F, 200 Ω), and 1 ml of 37°C LB was added immediately, followed by a 1 hour recovery at 37°C in a rotating drum. The recovered culture was pelleted at 3200 x *g* for 5 minutes at room temperature, 900 μ l of the supernatant was discarded, and the pellet was resuspended in the remaining liquid. The resuspended culture was plated on LBA with appropriate antibiotic selection at a maximum volume of 50 μ l and incubated overnight at 37°C

Transformation of *P. aeruginosa* was undertaken *via* the same method except that all steps were undertaken at 20°C with no chilling utilized, and the overnight cultures were grown with any agitation.

2.2.3.7 DNA sequencing

DNA sequencing was utilised for plasmid confirmation, it was conducted using GATC Biotech's sequencing service and isolated DNA samples were prepared in accordance with their instructions.

2.3 Protein purification

2.3.1 Expression testing of novel constructs

Overnight cultures of expression strains were diluted 1 in 100 in 50 mL of LB media supplemented with chloramphenicol and carbenicillin and grown until an OD₆₀₀ of 0.6. Isopropyl β -D-1-thiogalactopyranoside (IPTG) was then added to a concentration of 1 mM. Expression strains were constructed by the transformation of *E. coli* strain Rossetta with the relevant plasmid as listed in Table 2.2 The induced cultures were then incubated at 18°C for 18 hours, then centrifuged at 3220 *g* for 30 minutes. The supernatant was discarded, and the

pellet stored at -20°C. The pellets were resuspended in 1 mL lysis buffer (500 mM NaCl, 25 mM imidazole, 50 mM Tris-HCl pH 7.4), and lysed by sonication (30 second bursts with equal rest, 6 repeats, 13 amps). The lysates were added to 4x Sodium dodecyl sulphate (SDS) loading buffer (240 mM Tris-HCl pH 6.8, 8% (w/v) SDS, 40% (v/v) glycerol, 0.043 M EDTA, 0.04% (w/v) bromophenol blue, 0.1 M DTT) and heated to 95°C for 20 minutes. The samples were then analysed as details in 2.4.3 Acrylamide electrophoresis for protein.

Table 2.2 Protein expression plasmids

Plasmid	Description	Source
pMAL-c2x-NTD	MBP-tagged NTD expression plasmid, used predominantly for thermal shift assays; Cb ^R	Present study
pET-19m-ExsA-ExsD	His-tagged ExsA expression plasmid with untagged ExsD co-expressed, used in crystal trials; Cb ^R	Present study
pET16b-ExsA	His-tagged ExsA expression plasmid, used for electromobility shift assays; Cb ^R	(Brutinel <i>et al.</i> , 2008)
pET16b-ExsA-NTD	His-tagged NTD expression plasmid; Cb ^R	(Brutinel <i>et al.</i> , 2008)

2.3.2 Protein expression

Protein expression plasmids were transformed into *E. coli* Rosetta cells and either transformants or single colonies revived from glycerol stocks thereof, were used to inoculate 10 mL LB overnight cultures supplemented with carbenicillin and chloramphenicol. A 1 L of LB supplemented with carbenicillin and chloramphenicol was inoculated with the overnight culture and grown to an OD₆₀₀ of 0.6 at 37°C with 180 rpm shaking. IPTG was added to a final concentration of 1 mM and the temperature was lowered to 16°C. The cultures were incubated overnight, then pelleted and stored at -20°C if stored for under a week and -80°C if stored for over a week. They were thawed on ice, resuspended in 10 mL of lysis buffer (100 mM NaCl, 25 mM imidazole, 50 mM Tris-HCl, 1 cOmplete Mini protease inhibitor cocktail tablet (Roch diagnostics, Germany)), lysed by sonication (30 second bursts with equal rest, 10 repeats, 13 amps, all on ice), and spun down at 20,000 g for 20 minutes at 4°C. The supernatant was then

filtered with a 0.22 µm syringe filter (Qiagen) and the protein was purified with affinity chromatography as described below.

2.3.3 Nickel affinity chromatography

Varied salt concentrations were employed for the purification of different products; unless otherwise noted all buffers contained the NaCl concentrations listed in Table 2.3. Nickel affinity chromatography was used to purify all his tagged constructs, with a 5 mL Ni-NTA Superflow Cartridge (His-trap) (Qiagen) in an ÄKTA pure machine. The basic protocol was as follows: loading of the sample at 0.5 mL/minute until the entire sample is loaded, a 20 column volume wash with lysis buffer at 5 mL/minute, followed by elution with 7 column volumes of elution buffer (NaCl as specified in Table 2.3, 1 M imidazole, 50 mM Tris-HCl) at 1 mL/minute. The elution was collected in 5 mL aliquots, and the UV₂₃₀ reading was used to select aliquots which contained protein. For ExsA-ExsD copurification's two additional wash steps were employed before elution, the first consisted of a 20 column volumes of lysis buffer containing 500 mM NaCl, and the second of 7 column volumes of a lysis buffer containing no NaCl.

Where subsequent TEV protease cleavage of His tag was desired the dialysis step below was conducted with 1 mg of TEV per 50 mg of protein, and the dialysed protein was reapplied to the His-trap and the initial run through which did not bind to the nickel was collected.

The elute was dialysed overnight at 4°C in dialysis buffer (0.025 M Tris-HCl, NaCl as specified in Table 2.3, 1 mM DTT), and concentrated by ultracentrifugation in a Vivaspinn column (Sartorius) with the largest available pore size smaller than the protein product. Protein concentration was then approximated by absorbance at a wavelength of 280 nm, before snap freezing in liquid N₂ and storage at -80°C

Table 2.3 Salt concentration in protein purification buffers

NaCl Concentration (mM)	Expression Plasmid	Product
0	pET-19m-ExsA-ExsD	ExsA and ExsD
100	pMAL-c2x-NTD	MBT-ExsA NTD
500	pET16b-ExsA-NTD	ExsA NTD
500	pET16b-ExsA	ExsA

2.3.4 Heparin affinity chromatography

Heparin affinity chromatography was employed to purify full length ExsA. The input sample was the product of Nickel affinity chromatography, after dialysis but prior to concentration. The process was much the same as with the His trap, however a HiTrap Heparin HP (Cytiva,

Thermo Fisher Scientific, USA) was used rather than the HisTrap. A 7-column volume wash was undertaken with dialysis buffer, followed by elution with salt elution buffer (0.025 M Tris-HCl, 2 M NaCl, 1 mM DTT). A further dialysis step was then undertaken, followed by concentration and storage as in 2.3.

2.3.5 Amylose affinity chromatography

Amylose affinity chromatography was used to purify MBP tagged proteins. The input sample was the product of Nickle affinity chromatography, after dialysis but prior to concentration. The process was much the same as with the His trap, however a MBPTrap HP (Cytiva) was used. A 7-column volume wash was undertaken with maltose wash buffer (0.025 M Tris-HCl, 0.1 M NaCl, 0.01 M maltose) followed by elution with maltose elution buffer (0.025 M Tris-HCl, 0.1 M NaCl, 1 mM DTT, 0.5 M maltose). A further dialysis step was then undertaken, followed by concentration and storage as in 2.3.

2.3.6 Preparatory size exclusion chromatography

Size exclusion chromatography for the increased purity of purified proteins was undertaken using a HiLoad Superdex 200 pg preparative SEC column (Cytiva), with samples from nickel affinity chromatography once dialysed and concentrated. The sample was filtered with a 0.22 µm syringe filter (Qiagen). The column was equilibrated with dialysis buffer, the sample was then loaded via loop injection, and the column was then run for 1.5 column volumes with dialysis buffer at a flow rate of 0.5 ml/minute. The elute was collected in 1.5 ml aliquots and concentration and storage were undertaken as in 2.3.

2.3.7 Analytical size exclusion chromatography

Analytical size exclusion chromatography was undertaken with a Superdex 200 5/150 GL (Cytiva) on an ÄKTA pure machine. The column was equilibrated ExsA-ExsD dialysis buffer or a standard buffer (0.1 M NaCl, 50 mM Tris-HCl pH 7.4). Protein samples were manually injected, and 1.5 column volumes of the buffer used to equilibrate were run through at 0.3 ml/minute. Gel Filtration Standard (Bio-Rad) was used to generated standards.

2.4 Protein analysis

2.4.1 Quantification of protein

Protein quantification was done with a NanoDrop ND-100 spectrophotometer A_{280} reading, adjusted by the protein's extinction coefficients with the Beer-Lambert law. Specifically, the extinction coefficients were obtained using the ExPasy ProtParam tool (Gasteiger et al., ExPASy - ProtParam tool), and the concentration was calculated with the following equation:
 $A_{280}/\text{extinction coefficient} \times \text{pathlength} = \text{concentration}$

2.4.2 Bradford assay

When western blotting was utilised for total protein quantification an approximation of protein concentration was first made with a NanoDrop ND-100 spectrophotometer A_{280} reading, and a 50 μ l sample of the protein was adjusted to an approximate protein concentration of 1 mg/ml. A set of samples of known protein concentration were also made by dissolving bovine serum albumin in the same buffer as the samples to final concentrations of 0 mg/ml, 0.25 mg/ml, 0.5mg/ml, 1mg/ml, and 1.4 mg/ml. The final readings from these samples were used to construct a standard curve. Each sample had added to it 1.5 ml Bradford reagent (Sigma) and was mixed thoroughly by vortex. The samples were allowed to rest for 30 minutes at room temperature, and the absorbance at 596 nm was measured with a spectrophotometer. The readings from the samples of know concentration were used to plot a standard curve from which the concentrations of the unknown samples were calculated. All samples were tested in triplicate and mean values were used.

2.4.3 Acrylamide electrophoresis for protein

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) gels were cast as follows: a running gel was poured (15% w/v acrylamide (Severn Biotech, UK), 0.325 M TrisHCl pH 8.3, 0.1% SDS, 0.8% w/v Ammonium persulfate, 5 μ L Tetramethylethylenediamine (VWR, UK) between clamped glass plates from the Mini-PROTEAN System (Bio-Rad). The concentration of acrylamide was lowered for proteins of interest with larger molecular masses. A thin layer of deionised water was then gently pipetted above the gel mixture. Oncethe running gel had set completely the water was poured off, and a stacking gel (6% (w/v) acrylamide, 1.25 M Tris, 0.1% SDS, 0.8% w/v, APS, 5 μ LTEMED) was cast on top with a comb inserted.

Protein samples were prepared by adding SDS loading buffer and heating to 95°C for 15 minutes, then they were left to cool and briefly centrifuged to return condensation to the sample. The samples, along with Precision Plus Protein™ All Blue Standards (Bio-Rad), were subsequently loaded onto the gels and were run at 150 volts for approximately 1 hour, with variation allowed for the molecular mass of the protein, again in a Mini-PROTEAN System (Bio-Rad).

Where loading a known mass of protein was desired a Bradford assay was conducted on the samples to calculate the concentration, from which the volume to be loaded for the desired mass was derived.

2.4.4 Coomassie staining

Coomassie staining was performed as follows; polyacrylamide gels were stained for 40 minutes with Coomassie, then rinsed with de-ionised water, washed with destain 1 (50% methanol, 7% acetic acid) overnight, followed by destain 2 (10% methanol, 7% acetic acid) for 2 hours. The de-stained gels were then imaged with either with photography or scanning in a LI-COR Odyssey at a wavelength of 680 nm.

2.4.5 Preparation of cell culture for Western blotting

When performing Western blots with samples from *P. aeruginosa* cell pellets were resuspended in phosphate buffered saline (PBS) and either sonicated if the volume was sufficient, or subject to at least five freeze thaw cycles. The lysed cells were then centrifuged at 8050 g for 5 minutes to remove cell debris. Finally, a Bradford assay was conducted as above and acrylamide electrophoresis and Western blotting were conducted as below.

2.4.6 Pre-absorption of antibodies

In order to pre-absorb anti-PcrV antibodies a 1 L culture of $\Delta pcrV$ PAO1 was grown for 15 hours. The culture was then harvested by centrifugation at 8000 g for 20 minutes. The cell pellet was stored on ice, whilst the supernant was chilled on ice prior to the addition of ammonium sulphate to a final concentration of 65% W/V. The mixture was then left with gentle stirring for 2 hours at 4°C before the collection of precipitated proteins *via* centrifugation at 10,000 g for 30 minutes at 4°C. The protein pellet was resuspended in 10 ml of PBS and combined with the cell pellet. An additional 40 ml of PBS was added, the mixture was thoroughly resuspended, then lysed by sonication (30 second bursts with equal rest, 10 repeats, 13 amps, all on ice). The lysed cells were then treated with 200 ml of -20°C acetone. After 20 minutes a protein precipitate had formed, the excess acetone was removed and a further 200 ml of -20°C acetone was added, and the solution was mixed. After a further 30 minutes the acetone was again removed, the protein precipitate was then left to dry at room temperature for 3 days. Once dry the protein powder was ground in a pestle and mortar and 10 mg of the powder was mixed with 100 μ l of antisera (anti-PcrV antiserum from rabbit, BioGenes GMBH, Germany) and mixed end over end a room temperature for 1 hour. The mixture was then separated by centrifugation at 15000 rpm for 1 hour and the supernatant was collected and stored at -20°C.

2.4.7 Western blotting

Acrylamide gels were transferred to a membrane using a Trans-Blot Turbo RTA Kit (Bio-Rad, USA). The nitrocellulose and LF PVDF variants were used, both in accordance with the manufactures instructions in a Trans-Blot Turbo Transfer System (Bio-Rad). Once transferred the blots were blocked with 10 ml of 10% (w/v) dried skimmed milk powder (MARVEL, UK) dissolved in PBS (notated as PBS milk from hereon) for at least one hour at room temperature

with shaking. Subsequently the blots were transferred to 10 ml 5% PBS milk and primary antibodies were added and left shaking for one hour. Ratios of all primary antibodies are listed in Table 2.5. Blots were then rinsed with PBS, and subsequently washed three times with 10 ml PBS for 5 minutes with shaking. Secondary antibodies were then used

Gels for western blots were transferred to nitrocellulose membranes using the Trans-Blot® Turbo™ Transfer System (Bio-Rad) in accordance with the manufacturer's instructions. The membranes were visualised with anti His tag primary antibodies (Bio-Rad) and fluorescent secondary antibodies (Bio-Rad) in accordance with the manufacturer's instructions.

Table 2.4 Antibodies used for western blotting

Antibody	Ratio	Source
Anti-PcrV	1:5000	Present study
Anti-ICD	1:5000	Welch lab stocks
Anti-Hcp1	1:2000	Shunsuke Numata, Welch lab

2.4.8 Protein sequencing

Protein sequencing was conducted by the Cambridge Centre for Proteomics via their liquid chromatography – mass spectroscopy service.

2.5 *in vivo* assays

2.5.1 Luciferase reporter assays and growth curves

To perform the *PnirS*-Lux, *PpcrV*-Lux, and *PlacZ*-Lux reporter assays, overnight cultures of gentamycin supplemented LB with the appropriate strain were diluted 1:500 with AGSY media, or MOPS media with the specified carbon source. MOPS media was assembled as previously published (LaBauve and Wargo, 2012). ExsA activity was induced with 5 mM ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid (EGTA) immediately prior to the assays where indicated. Compounds were added to a final concentration of 200 μM unless otherwise noted, with an equal volume of solvent (DMSO) added to controls. The assays were performed in μclear 96 well plate (Greiner Bio-One, Austria) sealed with breathable membranes, with a total volume of 100 μL, with at least three biological replicates. The plates were then incubated at 37°C with shaking, for 13 hours. During this time course both the OD₆₀₀ and the luminescence were recorded every 30 minutes, using a FLUOstar Omega (BMG Labtech, UK). Adjusted Lux values were calculated by dividing Lux values by OD₆₀₀ x 10⁴.

2.5.2 Growth curves

Automated bacterial growth curves were performed by diluting an overnight culture of the relevant bacterial strain 1:500 in LB or a different media as specified, followed by growth with 37°C with shaking, for 13 hours. Assays were performed in 96 well plates sealed with breathable membranes with a total volume of 100 μ L, with four biological replicates. OD_{600} measurements were taken every 10 minutes using a FLUOstar Omega.

2.5.3 Preparation of samples for ExsA inhibition Western blotting

Overnight cultures of PAO1 or PAO1: Δ exsD were diluted 1:100 in AGSY, with or with 5 mM EGTA. Test compounds were added to a final concentration of 200 μ M, or an equivalent volume of DMSO was added. The cultures were grown to an OD_{600} of 1, then prepared for Western blotting as in 2.4.7.

2.6 Thermal shift assay

Hard-Shell® 96-Well PCR Plates (BIO-RAD) were prepared with 40x sypro orange dye (ThermoFisher) and 10 μ M of protein to a final volume of 12.5 μ L per well. For buffer screen the volume was obtained by dilution in deionised water, for the ligand binding assays the buffer in which the protein was stored was for dilution instead. Test solutions were prepared with 2x dialysis buffer and 200 μ M of test compound or equivalent volume of DMSO for ligand binding experiments, or 2x the relevant test buffer during the buffer screen. The test solutions were then added to the plates at a volume of 12.5 μ L per well and the assay was conducted on a Bio-Rad CFX Connect with the following parameters: temperature range: 4-95°C, incremental increase of 1°C every 45 seconds, fluorescence signal was measured every 15 seconds with a λ excitation of 470 nm and λ emission of 570 nm.

For the buffer screen 0.025 M Tris, 0.025 M HEPES, and 1x PBS buffers were tested at pH 6, 7, and 8. Each buffer condition and pH were tested with NaCl concentrations: 0 M, 0.075M, 0.2M, and 0.5 M. and every condition was tested with and without 10% glycerol.

3. Proteomic Analysis of the ExsA regulatory cascade

3.1 Aims and Approach

The opportunistic pathogen *P. aeruginosa* is dependant, in the vast majority of strains, on the T3SS to cause acute infection or establish a chronic infection. The T3SS is transcriptionally controlled by the AraC family TF ExsA, thus making ExsA an interesting drug target. It has also been reported that ExsA controls the expression of a virulence factor which is exported independently of the T3SS; ImpA. To establish a more complete picture of the ExsA regulon, especially novel virulence factors within the ExsA regulon and any downstream signalling effects, unmarked deletions of the four genes within the ExsA partner switching cascade were created and three were the subjects of proteomic analysis. The mutants provide “always off” and “always on” phenotypes for ExsA activity, allowing a full exploration of the regulon. Prior to the commencement of proteomic analysis, preliminary experiments were conducted to determine conditions in which the mutants would give varied phenotypes, using the T3SS protein PcrV as an indicator of ExsA activity. Phenotypic confirmation of changes indicated by the proteomic data was then sought.

3.2 Methods

Table 3.1 Plasmids used for the generation of unmarked deletion mutants

Name	Notes	Source
pEX19-Gm	Suicide vector in <i>P. aeruginosa</i> , <i>aacC1</i> mediated (Hmelo et al.) gentamicin resistance, <i>sacB</i> counter selectable marker.	
pEX18-Gm	Suicide vector in <i>P. aeruginosa</i> , <i>aacC1</i> mediated (Hmelo et al.) gentamicin resistance, <i>sacB</i> counter selectable marker.	
pEX18-Gm- Δ exsA	Derived from pEX18-Gm	Present study
pEX18-Gm- Δ exsD	Derived from pEX18-Gm	Present study
pEX18-Gm- Δ exsE	Derived from pEX18-Gm	Present study
pEX19-Gm- Δ exsC	Derived from pEX19-Gm	Present study

3.2.1 Unmarked mutant creation

Primers were designed in accordance with an established protocol for generating unmarked deletion mutants in *P. aeruginosa* (Hmelo *et al.*, 2015), care was taken to ensure the mutants remained in frame and downstream codons were not affected. In brief the strategy entails the PCR amplification of two flanks of the genes of interest between 400 and 500 nucleotides, which can then be joined via overlap extension PCR and cloned into the suicide vectors pEX19-Gm or pEX18-GM. Both vectors were used in the present study, and the plasmids

created are listed in Table 3.1. The primers were designed to introduce a site to facilitate overlap extension PCR and EcoRI and BamHI restriction sites to the 5' and 3' ends of the combined flanks respectively and are listed in Table 3.2.

3.2.1.1 Splicing by overlap extension

PCR amplification of flanking regions was conducted as detailed in general methods, with the addition of 5% DMSO for all PCRs concerning *exsC*. The primer pairs Up-F and Up-R or Down-F and Down-R were used for each gene, e.g. *exsA*-Up-F and *exsA*-Up-R, to generate flanking regions for each gene. The two flanking regions for each gene were then spliced together in an overlap extension PCR, which followed the same methodology except with the template DNA consisting of 0.1 ng of each amplified flank, using the primers Up-F and Down-R for each respective gene. The products of each step were isolated via gel extraction where the PCRs generated non-specific bands.

3.2.1.2 Validation of mutants

Once mutants were recovered from selection and counter selection as in the protocol followed (Hmelo *et al.*, 2015), they were validated via colony PCR using UP-F and Down-R primers for each gene. Wild type PAO1 was included as a negative control in which a band corresponding to the deleted region would be seen, whilst the mutants would produce a shorted band.

3.2.2 Growth in 50 mL shaking flasks

Overnight cultures of the relevant strains in LB were inoculated into 50 mL of fresh media to a final OD₆₀₀ of 0.05 and sealed with bung to allow aeration. The cultures were then grown in a shaking water bath at 37°C with 200 rpm shaking. The OD₆₀₀ was measured every hour until completion by withdrawing 1 mL samples and reading on an Eppendorf Biospectrometer.

3.2.3 Concentration of secreted proteins

In order to concentrate secreted proteins sufficiently for quantitative western blot, cultures were centrifuged for 10 minutes at 4000 rpm and the supernatant was extracted. The following methods were employed to concentrate the protein from the supernatant:

Acetone extraction began by the addition -20°C acetone to the supernatant to a total of 5x the original volume. The mixture was then vortexed thoroughly and stored at -20°C for 24 hours. The samples were then spun down for 30 minutes at 4000 rpm, 4°C, and the supernatant was discarded. The protein pellet was then air dried, and resuspended in the following buffer: 50 mM tris (pH 7), 100 mM NaCl, 8% (w/v) SDS. After vortexing and shaking at room temperature for 30 minutes any remaining insoluble fraction were spun down and the supernatant removed for Bradford assay quantification and Western blotting.

Ultracentrifugation conducted by placing the culture supernatants in a 10 kDa Vivaspin column (Sartorius), followed by ultracentrifugation at 4000 rpm at 4°C until a 10 fold reduction in volume was achieved.

Table 3.2 Primers utilised for unmarked deletion creation

Name	Sequence	Feature
ExsA-Up-F	cgcgaaattcTTGGAGGGCGAGGCATTG	EcoR1
ExsA-Up-R	GCCCGGCATTTCGTCCTTCCCTTGCATATTATAAGAACCCC	SOE overlap
ExsA-Down-F	GGGAAGGACGAATGCCGGGCT	
ExsA-Down-R	atggatccGGCAGCAGGACCCAGTCGATC	BamH1
ExsA-SeqF	CCTTCGATCTGGAGGTCGAC	
ExsA-SeqR	GCAGGCTTTCCCACCAGCCA	
ExsD-Up-F	cgcgaaattcAGGGGCCGCTGCTGATGTCG	EcoR1
ExsD-Up-R	GCTCAGCTCTGCCAGTAGAAATCGTCTTCCTGCTCCATTC	SOE overlap
ExsD-Down-F	TTCTACTGGCAGAGCTGAGCGGCCG	
ExsD-Down-R	atggatccTCAGGGACGCCACACCGG	BamH1
ExsD SeqF	GCTGATGCTCTTCGCGTTCA	
ExsD SeqR	GCCAGGCAGCAGCGCCAGCA	
ExsE-Up-F	cgcgaaattcCGAGTTCGCAGGCCGTAT	EcoR1
ExsE-Up-R	ACAGCATCCAGCACCTCATTTCGATTTTCATGGGGCC	SOE overlap
ExsE-Down-F	TGAGGTGCTGGATGCTGTTGCCG	
ExsE-Down-R	atggatccTGTATCAATCGTTGCCAGATC	BamH1
ExsE SeqF	AAGCGATGCATAGCCCGGTG	
ExsE SeqR	CGAGGCCTTCCCTGAAAACG	
ExsC-Up-F	cgcgaaattcTGCCTTCGAGAGCCGCAA	EcoR1
ExsC-Up-R	TTTCATGGGGCCGCTCAAACCTTGCTCGTTAAATCCAT	SOE overlap
ExsC-Down-F	GTTTGAGCGGCCCATGAAAATCGA	
ExsC-Down-R	atggatccAGGCGCAGCCTGAGTTGC	BamH1
ExsC SeqF	GCGCTTGCAAGACCTCCGA	
ExsC SeqR	CCAACCTGTGACCTCCAGA	

3.2.4 Artificial sputum medium continuous flow model

A steady state culture utilising artificial sputum media has previously been developed within the lab (O'Brien and Welch, 2019). This model was employed as previously published for the initial 96 hour run. In brief an artificial sputum medium (ASM, a highly complex media produced

as referenced (O'Brien and Welch, 2019)) culture was inoculated with PAO1, with new media added and spent culture removed at a rate proven to maintain a steady state culture, with stirring at 37 °C. For the EGTA induced samples the system was run as prior for 24 hours, at which point a 1.5 mL sample was collected and EGTA was injected into each sample to a final concentration of 5 mM. The system was incubated for a further hour before 1.5 mL samples were collected and cells were harvested by centrifugation for 10 minutes at 4000 rpm.

3.2.5 Proteomic analysis

The preparation of final proteomic samples and submission to the Cambridge Centre for Proteomics was conducted by Dr Meng Wang. Preparation of samples for proteomics began as in 3.2.2 Growth in 50 mL shaking flasks, with 50 ml cultures of each sample being grown from an OD₆₀₀ of 0.05 for 8 hours. A reference flask was used for each genotype to construct a growth curve without disturbing the samples. Cultures were then spun down at 4 °C, the supernatant was discarded and the pellet was resuspended in 10 mL of sterile PBS. The resuspension was again spun down, the supernatant discarded, and the pellets were snap frozen in liquid nitrogen and stored at -80 °C. The samples were then thawed on ice and resuspended in 1.25 mL of lysis buffer (100mM Tris/HCl pH7.5, 50mM NaCl, 10% v/v glycerol, 1mM tris(2-carboxyethyl)phosphine), and sonicated for 5 second burst with 15 seconds rest between for 15 cycles at 10 amplitude. Protein concentration was determined via a Bradford assay and the concentration was adjusted by dilution with lysis buffer to provide 100 µg samples at approximately 50 µL. Protein samples were snap frozen in liquid nitrogen and stored at -80 °C.

A TMTpro™ 16plex Label Reagent Set (Thermo Scientific) was used to label the protein samples in accordance with the manufacturer's instructions, and the labelled samples were submitted to the Cambridge Centre for Proteomics for LC-MS/MS analysis. Fragmentation data was analysed by searching against the National Centre for Biotechnology Information (NCBI) database with the MASCOT (Matrix Science) search engine. Statistical analysis was conducted with the empirical Bayes moderated T-test via the limma package for R. P-values were adjusted with the Benjamini-Hochberg method (FDR ≤ 0.05), and fold change in expression was calculated from normalized log₂ ratios.

3.2.6 STRING map construction

STRING maps were constructed using the STRING database web portal [<https://string-db.org/>] with the "Multiple Proteins" function and the organism specified as *Pseudomonas aeruginosa* (PAO1) (von Mering *et al.*, 2005; Szklarczyk *et al.*, 2021).

3.2.7 Pyocyanin Quantification

AGSY cultures were prepared and grown as in 3.2.2 for 8 hours. The cultures were then pelleted by centrifugation at 4000 rpm for 15 minutes, and 15 ml of the supernatants was extracted and filter sterilised. To this 9 ml of chloroform was added and mixed thoroughly. The chloroform extractions were then separated by centrifugation at 4000 rpm at 4°C for 10 minutes and 6 ml of the lower layer was extracted. To this 3 ml of 0.2 M HCl was added, mixed thoroughly, and then separated via centrifugation as prior. The top layer was then extracted and the OD₅₂₀ was measured. Pyocyanin concentration was calculated as follows: OD₅₂₀ x 17.072 x 1.5 = pyocyanin (µg/ml)

3.2.8 Siderophore and nitrite quantifications

The samples were prepared by growth of the strains as for the proteomic sample preparation, followed by spinning down at 3200 x g at 4°C for ten minutes. The supernatant was then removed and filter sterilised for use in the assays. Siderophore quantification was undertaken longstanding protocol for the “Chrome Azurol S Liquid Assay”, with the total assay volume reduced to 100 µL for compatibility with 96 well plates and each component reduced proportionally (Payne, 1994). A Griess Reagent System (Promega) was used in accordance with the manufacturer’s instructions to quantify nitrite.

3.3 A slight growth defect is apparent in the “always on” mutants

The mutants were initially examined for any effects on growth by conducting growth curves in both manual and automated fashion. The “always on” mutant strains $\Delta exsD$ and $\Delta exsE$ exhibited a slight growth defect in both LB and AGSY media in both methodologies. The growth defect was observed both in the growth of 50 mL cultures (Figure 3.1 A and B) and in 200 µL cultures grown in a 96 well plate (Figure 3.2 A and B). In the 50 mL cultures a greater divergence in OD₆₀₀ was observed in LB than AGSY, from time points 7 till 11 hours in LB an average difference of 12.2% for $\Delta exsD$ and 5.9% for $\Delta exsE$ were recorded, whilst in AGSY the respective values are 8.3% and 2.9%. All data was analysed with a 2-way ANOVA, which demonstrated statistically significant difference within each data set, and Tukey’s test which identified time points for $\Delta exsD$ and $\Delta exsE$ which were significantly different from PAO1.

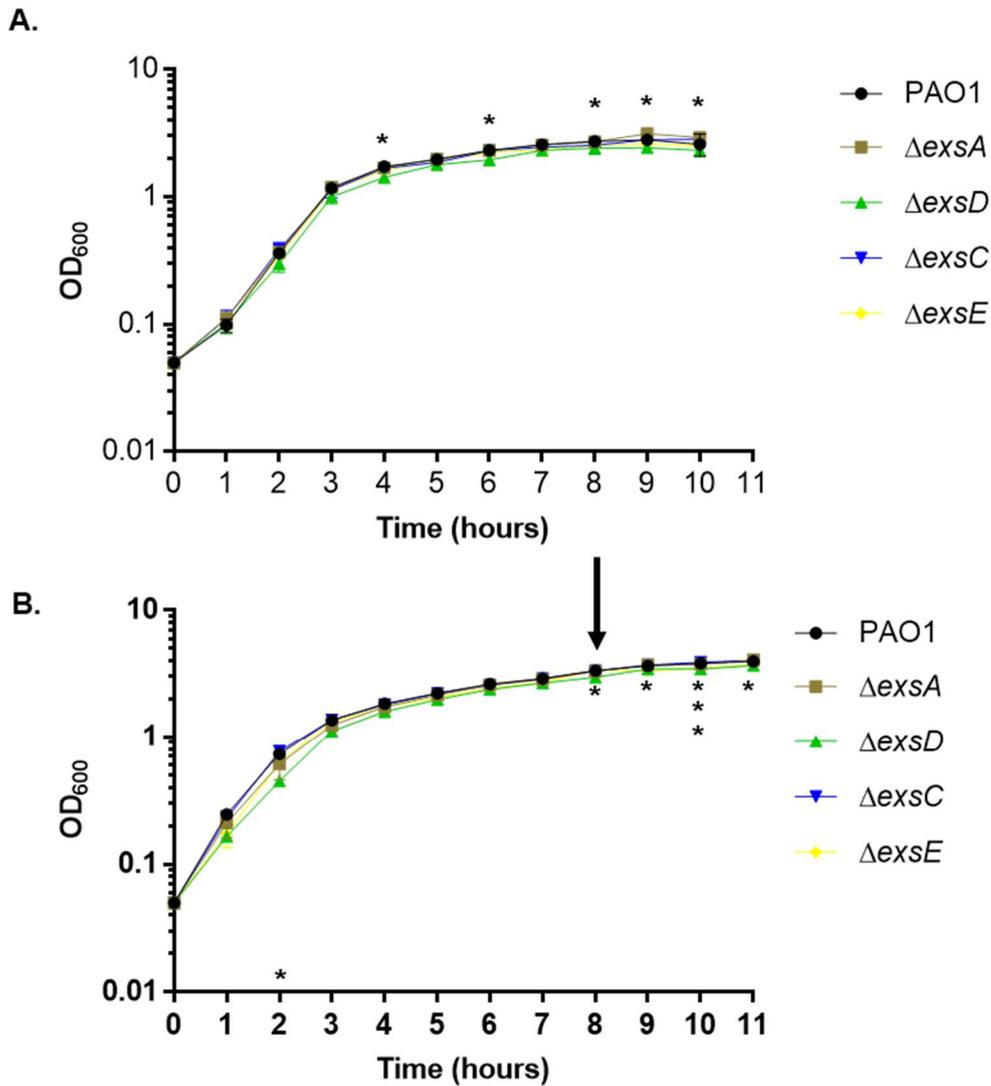


Figure 3.1 Growth curves of PAO1 and novel mutants.

Growth curves of mutant stains and PAO1 *P. aeruginosa* in 50 mL shaking water bath cultures. Y axis utilises a Log_{10} scale. **A** and **B** represent data from experiments conducted using LB medium and AGSY medium, respectively. The arrow indicates the timepoint at which proteomic and other samples were taken, unless otherwise noted.

* indicates time points at which $\Delta exsD$ is significantly different ($P < 0.01$) from PAO1 and ** indicates time points at which $\Delta exsE$ is significantly different ($p < 0.01$) from PAO1, when assessed with a Tukey post hoc test following a 2-way ANOVA. Each data point represents the mean of three biological replicates, and all data is representative of two independently conducted experiments. Error bars indicate standard deviation (SD) and were omitted in the plot when too small to be visible.

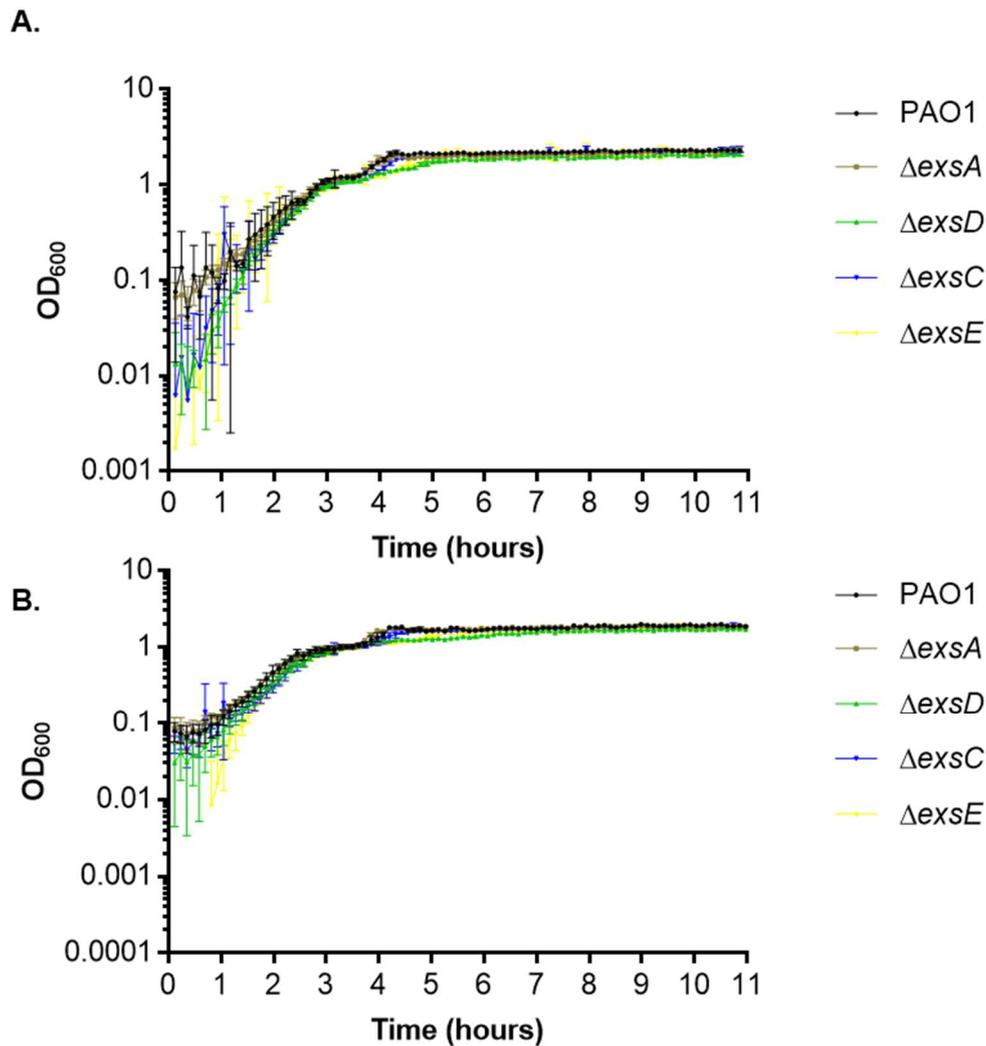


Figure 3.2 Growth curves from microplate shaking incubator are comparable.

Growth curves of mutant stains and PAO1 in 200 μ L cultures conducted with a Flurostar. Y axis utilises a Log_{10} scale. **A** and **B** represent data from experiments conducted using LB medium and AGSY medium, respectively.

Error bars indicate SD and are omitted when too small to be visible. Each data point represents the mean of five biological replicates \pm SD.

3.4 Preliminary characterisation of ExsA activity by western blot

The mutants' phenotypes with respect to T3SS were examined, and various conditions trialled to find a suitable condition for proteomic investigation of the mutants. This was achieved with Western blots for PcrV to detect T3SS expression, Hcp1 as an indicator of T6SS expression, and the metabolic protein ICD (isocitrate dehydrogenase) as a loading control due to its consistent expression (Crousilles *et al.*, 2018).

3.4.1 Mutant expression of PcrV correspond to expected phenotypes

To test the phenotypes a basic culture method was employed with a 1:100 dilution of an overnight culture into AGSY grown to an OD₆₀₀ of 1 with or without EGTA induction. As expected, the “always on” mutants ($\Delta exsD$ and $\Delta exsE$) displayed abundant expression of PcrV, in both extracellular and cellular culture fractions, even whilst not induced with EGTA (Figure 3.3 A&C). The opposite phenotype was observed in the “always off” mutants ($\Delta exsA$ and $\Delta exsC$) which showed no PcrV visible when induced despite PAO1 demonstrating robust expression in this condition (Figure 3.3 A&B). Hcp1 expression appeared unchanged in the cellular fraction, however in the extracellular fraction it appears to be higher in the “always on” mutants compared to PAO1 (Figure 3.3 A).

3.4.2 An ASM continuous flow model induces little PcrV expression

The most physiologically relevant model available in the lab was artificial sputum media in a continuous flow system developed to mimic chronic infection conditions. It was therefore explored first as a potential system for further examination of the mutants. It was observed that over the course of a 96-hour continuous flow culture that PcrV was not detected at an observable level in PAO1 (Figure 3.4 A). It was further demonstrated that the $\Delta exsD$ mutant has robust PcrV expression in the system and EGTA induction does not stimulate PcrV production in PAO1 in this system (Figure 3.4 B). Hcp1 expression was constant whether or not EGTA expression takes place and appears to be stronger in the $\Delta exsD$ mutant (Figure 3.4 B&C).

3.4.3 AGSY supports more robust PcrV expression than LB

Given the muted PcrV expression of wild type (WT) PAO1 in the ASM continuous flow system a different culture method was sought which would allow greater differentiation of the WT and $\Delta exsA$ phenotypes. To determine which media would be preferable for the proteomics experiment both LB and AGSY were examined. It was generally observed that the degree of PcrV expression was higher in AGSY. On direct comparison gels between LB and AGSY for wild type induced and non-induced and the $\Delta exsD$ mutant, PcrV was detectable in both the $\Delta exsD$ mutant samples and the induced sample for AGSY but not in LB (Figure 3.5). A further examination of AGSY without induction was undertaken, showing both always on mutants had strong PcrV bands, with $\Delta exsD$ appearing to have more than $\Delta exsE$, PAO1 has a dimmer band, $\Delta exsC$ a very faint band, and $\Delta exsA$ has no visible band at all (Figure 3.6).

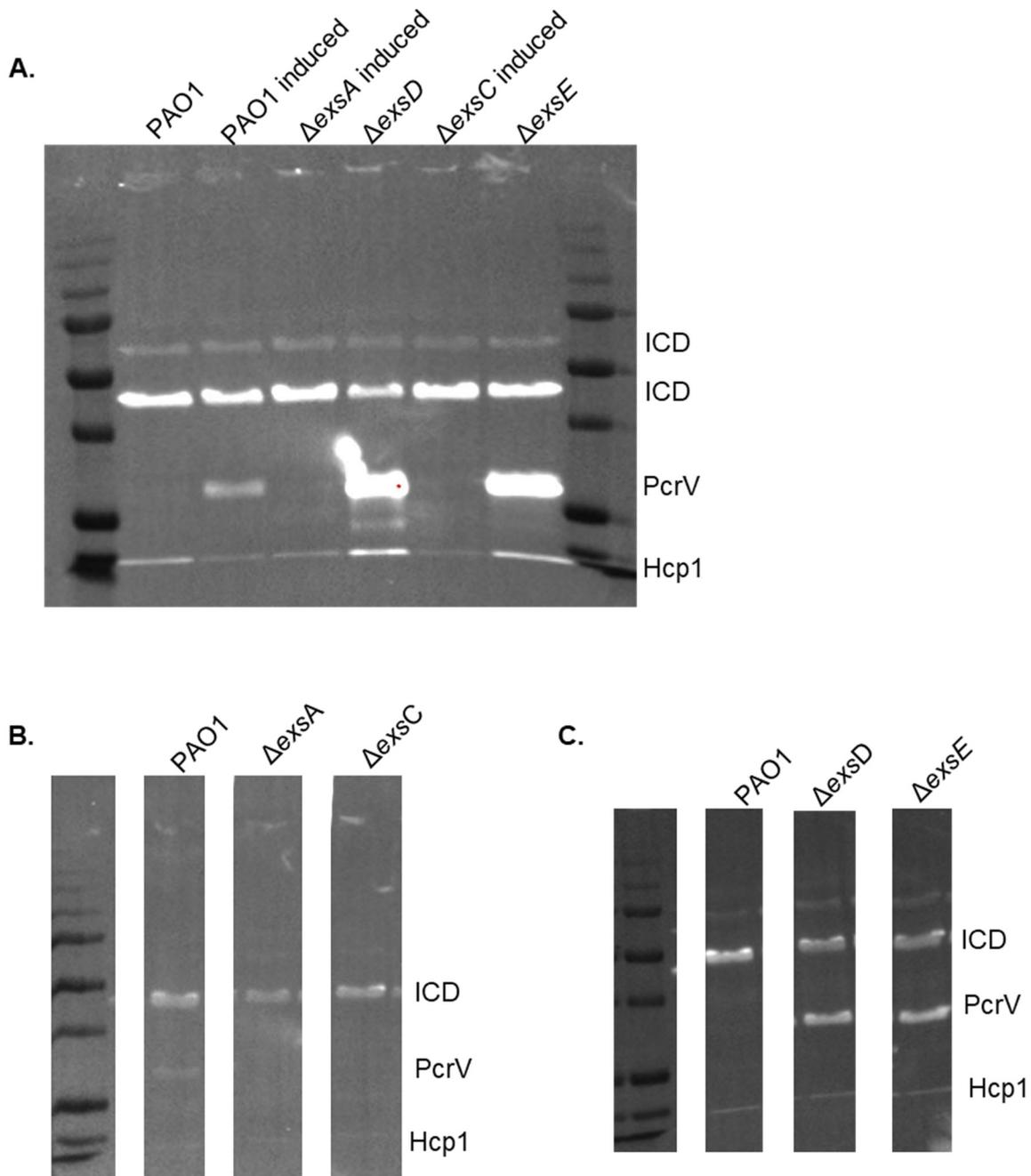


Figure 3.3 Preliminary phenotypic Western blots for deletion mutants.

Western blots of AGSY cultures grown to an $OD_{600} = 1$ using anti-PcrV, anti-Hcp1, and anti-ICD antibodies. A total of 5 μ g of protein was loaded into each sample well. Images **B** and **C** have been cropped to remove technical replicates for the sake of clarity. All images are representative of three independent biological replicates.

A. Extracellular culture fraction concentrated by ultracentrifugation from induced or non-induced cultures as indicated.

B. Intracellular culture fractions from samples induced with EGTA.

C. Intracellular culture fractions from samples without induction.

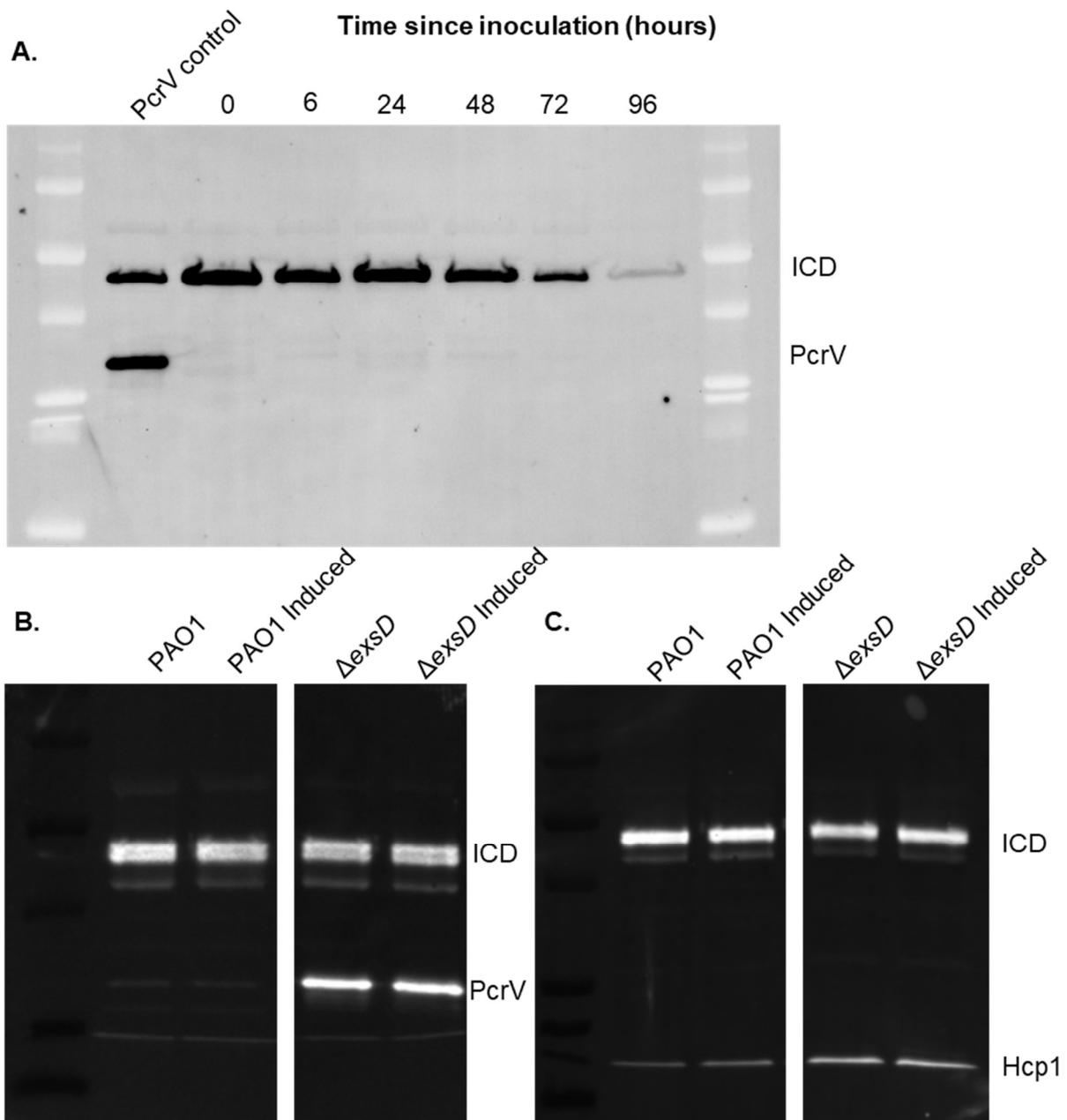


Figure 3.4 Western blots for samples from an ASM continuous flow system.

Samples were extracted from a continuous flow system and prepared for blotting by cell lysis. All sample wells were loaded with 5 μ g of total protein.

A. Typical Western blot showing samples of PAO1 extracted from the same flow system at multiple time points. Due to a lack of detectable PcrV in earlier blots a PcrV positive control was included in the form of a Δ exsD sample grown in AGSY. Anti-ICD and anti-PcrV antibodies were used.

B. Samples of PAO1 and Δ exsD from the continuous flow system pre and post EGTA induction. Anti-ICD and anti-PcrV antibodies were used. Image cropped to remove technical replicates.

C. Samples of PAO1 and Δ exsD from the continuous flow system pre and post EGTA induction. Anti-ICD and anti-Hcp1 antibodies were used. Image cropped to remove technical replicates.

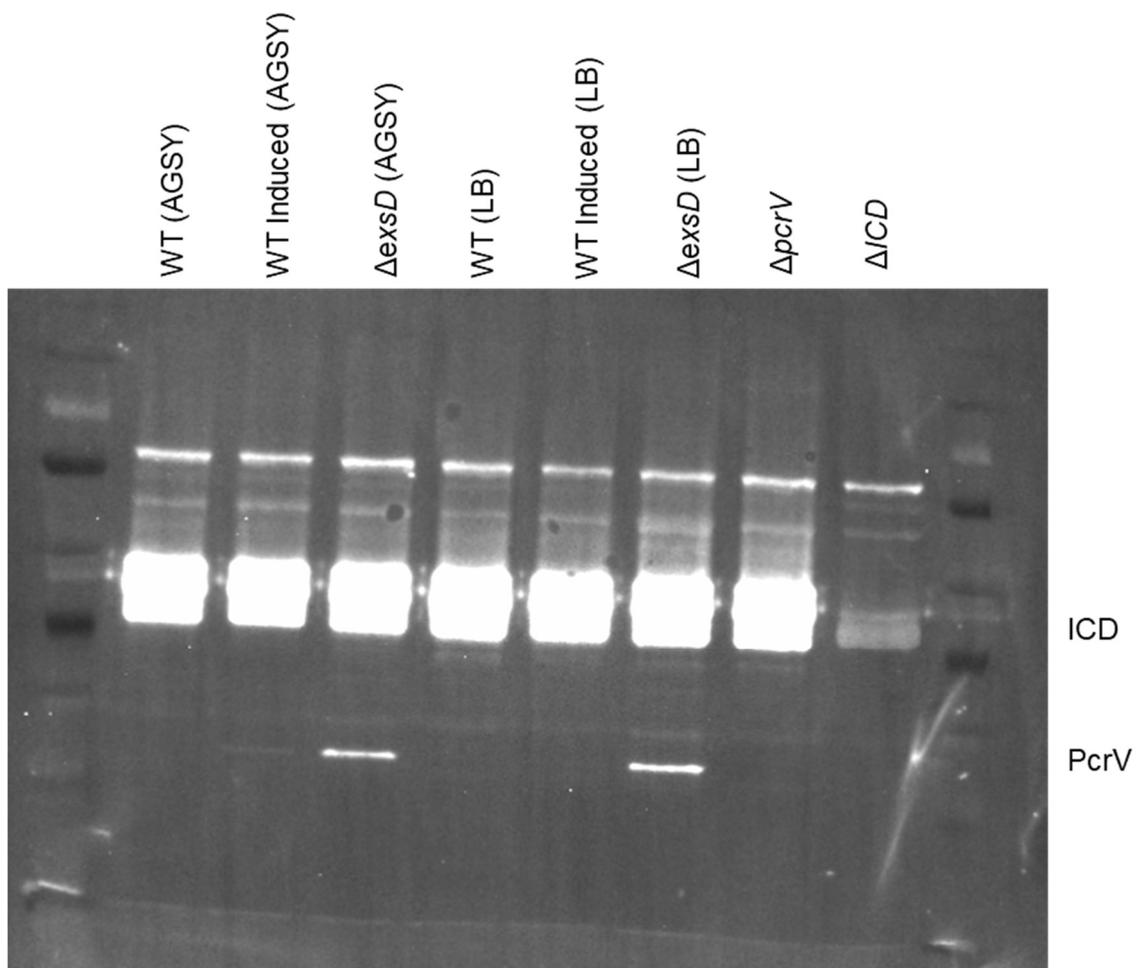


Figure 3.5 Western blots for AGSY and LB.

AGSY and LB cultures grown to an OD_{600} of 1 with and without induction. All sample wells were loaded with 10 μ g of total protein and anti-PcrV and anti-ICD antibodies were used. Significant smearing is seen above the ICD band, however the PcrV band is clear.

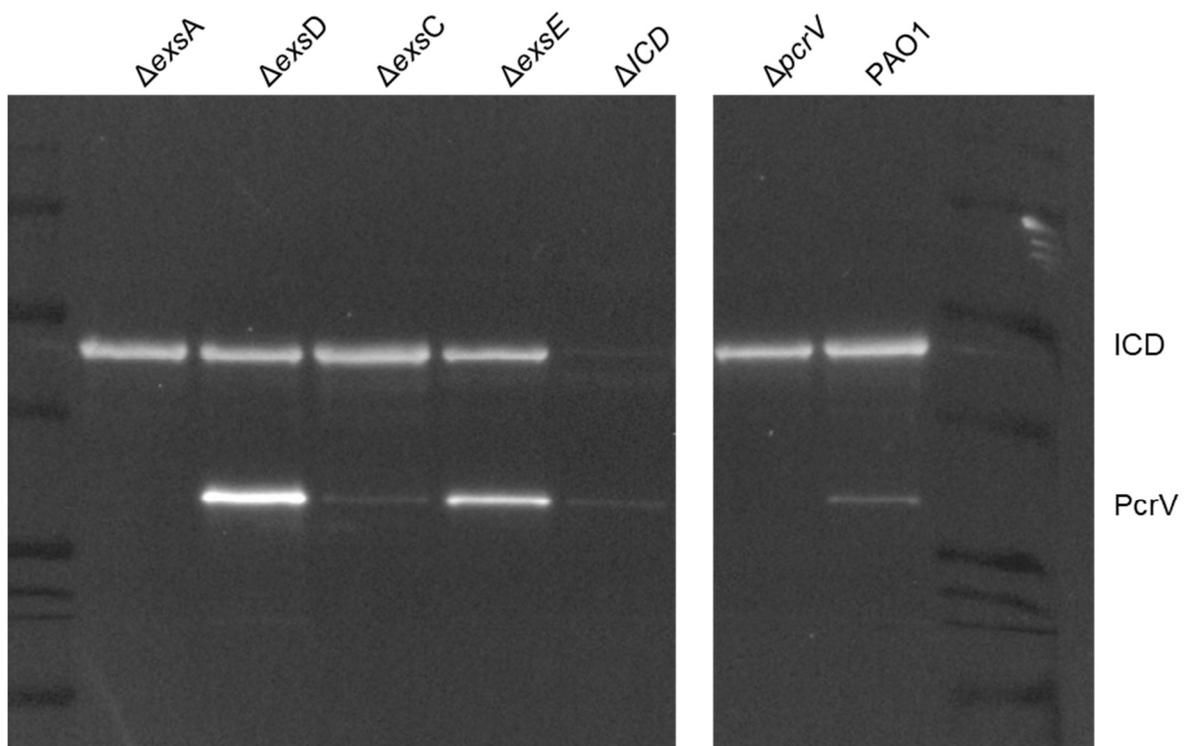


Figure 3.6 Western blot of PcrV in AGSY cultures.

PcrV in AGSY cultures grown without induction to $OD_{600} = 1$, with 10 μ g of total protein loaded per well. Blotted with anti-ICD and anti-PcrV antibodies. This image is representative of biological triplicate and has been cropped to remove a lane containing a mutant irrelevant to the current investigation.

3.4.4 PcrV and Hcp1 expression increase in late growth phase

To determine a time point at which a 50 mL shaking water bath culture in AGSY would provide an interesting comparison for proteomics, samples from across the growth curve were taken and blotted for PcrV and Hcp1. Hcp1 was detectable throughout but appeared to be induced more strongly after 7 hours, whilst PcrV expression became detectable at 7 hours and remained similar from then on (Figure 3.7).

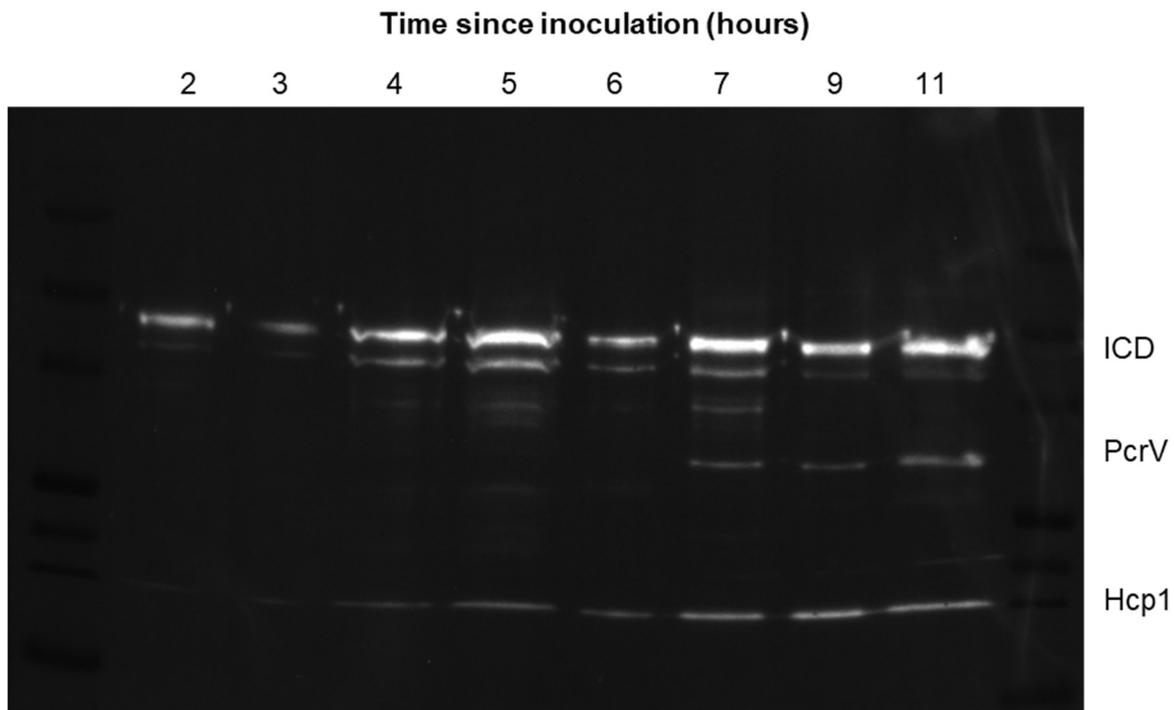


Figure 3.7 Western blot of samples across a growth curve.

A time course of WT PAO1 intracellular samples taken from a 50 mL shaking water bath cultures in AGSY. Blotted with anti-ICD, anti-PcrV, and anti-Hcp1 antibodies. 5 μ g of total protein loaded per well. This image is representative of biological triplicate.

3.5 Proteomics

From the above data it was decided to grow the strains Δ exsA, Δ exsD, and Δ exsE, along with wild type, in 50 mL shaking flasks for 8 hours and then harvest the cells for proteomic investigation. Given the 16 plex probe sets and desire to run 4 replicates of each sample Δ exsC was excluded.

3.5.1 The proteomes of Δ exsA and WT are similar whilst Δ exsD and Δ exsE cluster separately

Partial Least Squares Discriminant Analysis (PLS-DA) clustered the WT and the Δ exsA mutant proteomes closely, though they are somewhat separated. Meanwhile the Δ exsD, and Δ exsE mutants are again clustered near to one another though noticeably separate. It is noteworthy that Δ exsD and Δ exsE are separated on the vertical axis, which accounts for almost 10x less variation than the horizontal axis on which they are much closer. The two pairs of clusters are separated very strongly within the analysis (Figure 3.8). A total of 3312 proteins were detected out of 5587 hypothetical proteins annotated to the PAO1 genome. Throughout the following section proteins will be referred to by both their UniProtKB accession and subsequently protein name where available and the corresponding genetic locus tag where unnamed.

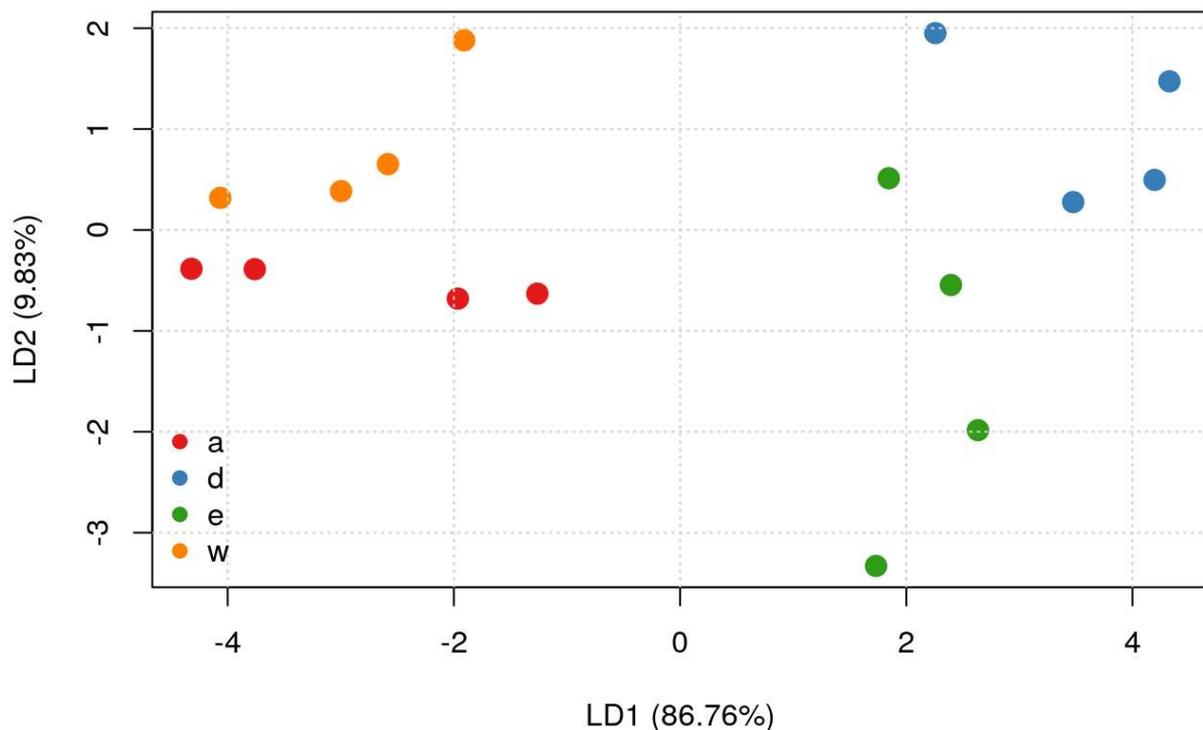


Figure 3.8 Scatter of proteomic samples after PLS-DS analysis.

A scatter graph of PLS-DA data analysis of the groupings and distances for the proteomes of each replicate sample for $\Delta exsA$, $\Delta exsD$, $\Delta exsE$ mutants and the wild type PAO1 background strain after normalisation.

3.5.2 The $\Delta exsA$ mutant diverges from wild type on a small number of proteins

The $\Delta exsA$ mutant diverged from the WT strain in a statistically significant manner for five proteins (Figure 3.9 A). However only two of these were beyond a fold change limit of 1, P26993/ExsA which was unsurprisingly very reduced in the mutant, and Q9HVZ0/GmhA, a lipopolysaccharide synthesis gene, which was up-regulated in the mutant (Figure 3.9 B). Three other proteins were statistically different with a modest change in expression: Q9I481/PA1263 is slightly downregulated, whilst Q9I318/PscD and O68822/PepA and upregulated (Figure 3.9 B).

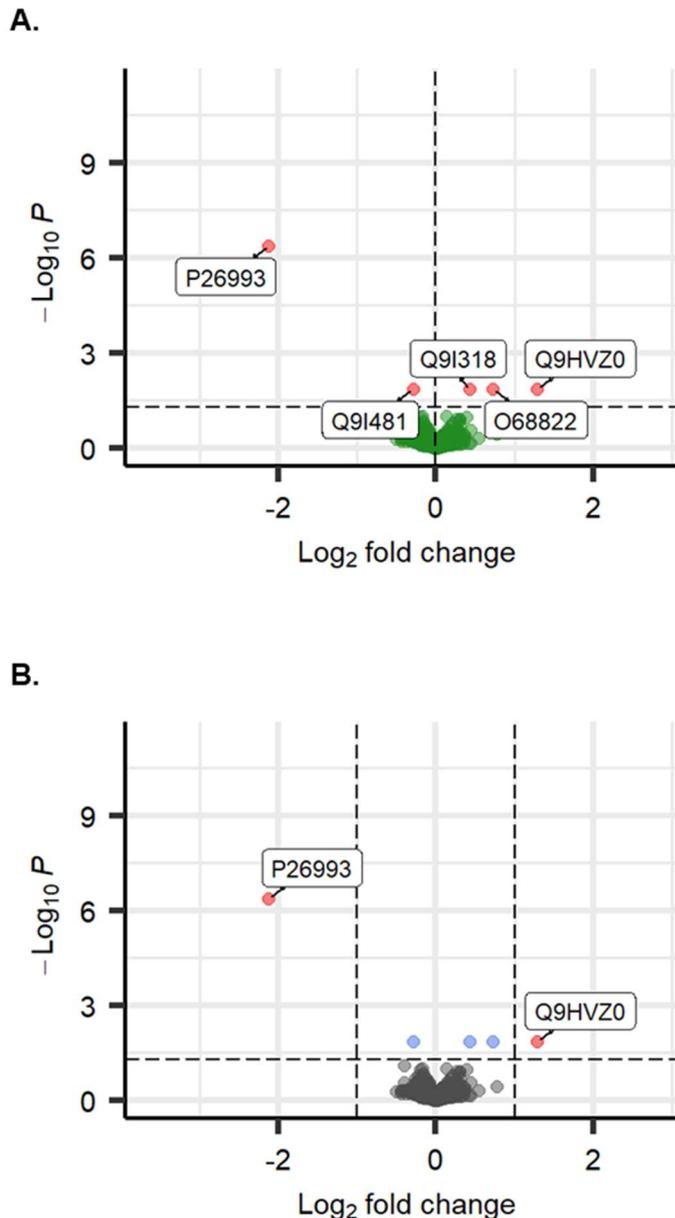


Figure 3.9 Volcano plot comparison of $\Delta exsA$ and WT.

Two volcano plots depicting a comparison between $\Delta exsA$ and WT, with change in $\Delta exsA$ with respect to WT for every detect protein shown. **A.** shows all statistically significant data points ($p < 0.05$) in red, whilst **B.** imposes a \log_2 (fold change) cut off of 1 and proteins failing to meet that threshold are blue.

3.5.3 $\Delta exsD$ and $\Delta exsE$ mutants present similar proteomes

The $\Delta exsD$ mutant and $\Delta exsE$ mutant presented very similar proteomes, with similar volcano plots (Figure 3.10). The $\Delta exsD$ mutants had more proteins divergently regulated within the criteria of statistical significance and a \log_2 fold change greater than one (36 downregulated, 49 upregulated) compared to $\Delta exsE$ (15 downregulated, 40 upregulated).

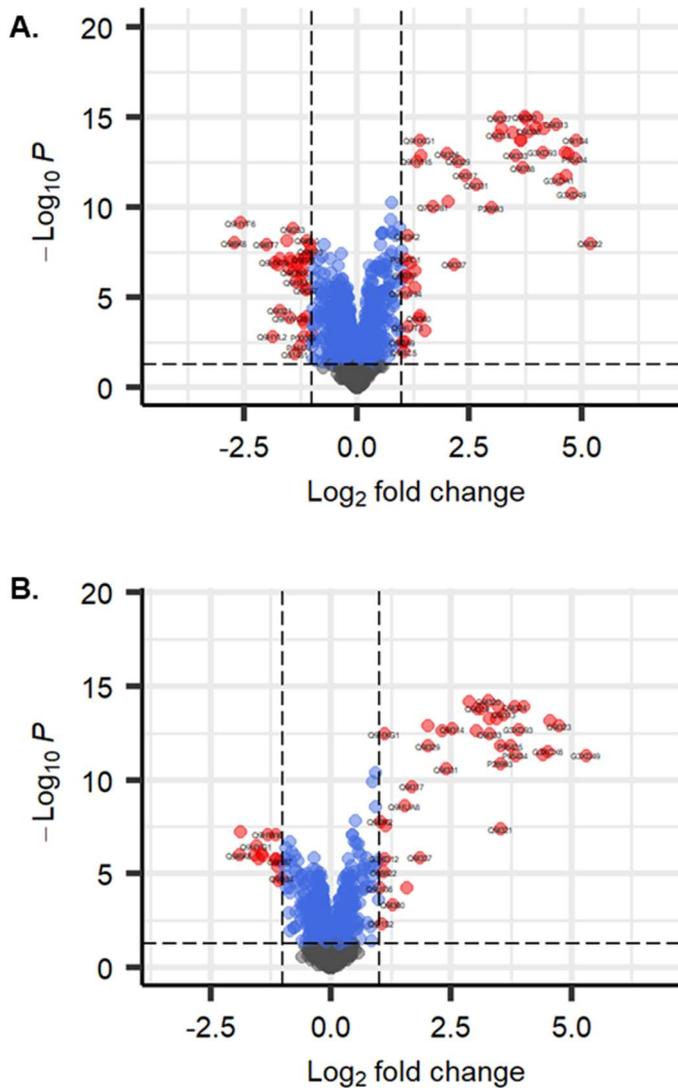


Figure 3.10 Volcano plot comparisons of always on mutants and WT.

Volcano plots depicting a comparison between ΔexsD and ΔexsE mutants with the WT, with change in the mutants with respect to WT for every detect protein shown. Blue and red dots represent statistically significant data points ($p < 0.05$), and red dots are those which meet a Log_2 fold change minimum of ± 1 .

A. Comparison between ΔexsD and WT, with 49 up regulated proteins and 36 down regulated proteins which meet the statistical and fold change thresholds.

B. ΔexsE compared to WT, with 40 up regulated proteins and 15 down regulated proteins.

To aid analysis STRING maps were constructed, identifying clusters of hits which are share connections. The main upregulated grouping for both $\Delta exsD$ and $\Delta exsE$ mutants were T3SS proteins (Figures 3.11 & 3.12). T6SS proteins were also identified, as was the T6SS effector Q9I3K2/TpIE. Two uncharacterised proteins predicted to be transcription factors (TFs), Q9HXG1/PA3845 and Q9I787/PA0045, were detected as upregulated in the $\Delta exsD$ mutants, one of which (Q9HXG1/PA3845) was similarly detected in the $\Delta exsE$ mutants. HCN synthesis proteins were also identified, as was a pyocyanin synthesis protein. Two members of an NRPS operon, PA1212 and PA1218, are upregulated in the $\Delta exsD$ mutants. Upon further inspection it was apparent that the rest of the operon was also upregulated, including PA1217, but fell short of the fold change threshold set in the main analysis, Q9I4B9/PA1219 was undetected (Figure 3.13).

Down regulation revealed two shared clusters, corresponding to pyoverdine synthesis and a set of metabolic proteins used in sulphur and taurine metabolism (Figures 3.14 & 3.15). An additional cluster identified in the $\Delta exsD$ mutant was composed of 4 denitrification proteins.

When directly compared the two always on mutants exhibited only one protein other than Q9I322/ExsE and Q9I321/ExsD that was differentially expressed between the two with a Log_2 fold change of greater than 1, P95434/PscF – a T3SS protein (Figure 3.16 A). When the Log_2 fold change limit was reduced to 0.5 a larger number of proteins were detected allowing string maps to be constructed. $\Delta exsD$ contained 15 upregulated and 12 downregulated proteins compared to $\Delta exsE$ within these criteria, which is more in line with the differences between the two in comparison to WT (Figure 3.16 B). When grouped into STRING maps it was observed that the differences were predominantly in groups observed similarly between the two mutants in comparison to WT but larger in $\Delta exsD$ (Figure 3.17).

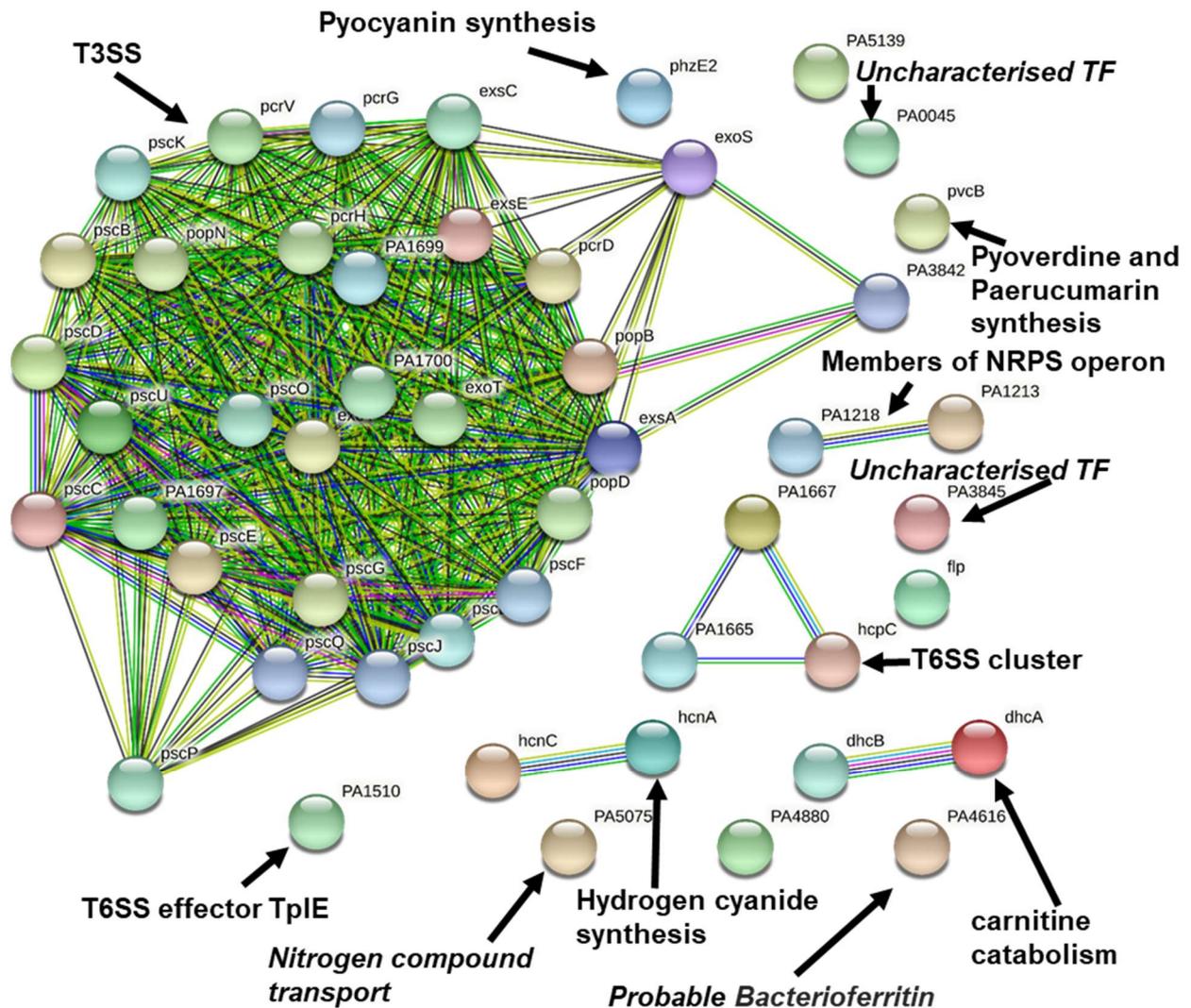


Figure 3.11 Annotated STRING map of proteins upregulated in $\Delta exsD$ mutants.

All proteins which were of increased abundance in $\Delta exsD$ compared to WT, where that increase was both statistically significant and greater than Log_2 fold change = 1. The connections are automatically generated by the STRING webservice and are indicative of shared function or regulation. Point annotations contain the relevant gene name, whilst manual annotations added with arrows indicate known or predicted biological functions (with predicted functions in italics) for the relevant data point or cluster.

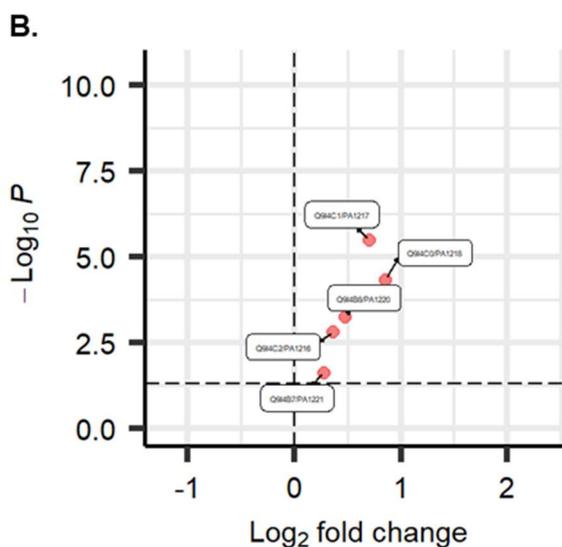
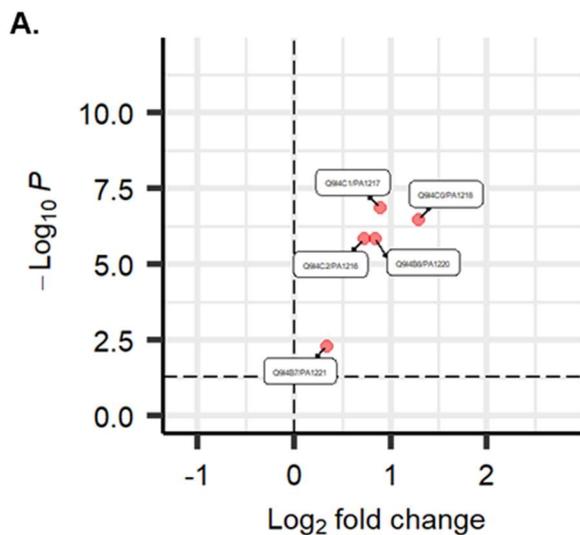


Figure 3.13 Volcano plot comparisons of always on mutants and WT for the NRPS operon PA1216-PA1221.

Volcano plots depicting a comparison between $\Delta exsD$ and $\Delta exsE$ mutants with the WT, with change in the mutants with respect to WT for every detected member of the PA1216-PA1221 NRPS operon shown. Red dots represent statistically significant data points ($p < 0.05$). Comparison between $\Delta exsD$ and WT, and $\Delta exsE$ and WT are shown in **A.** and **B.** respectively.

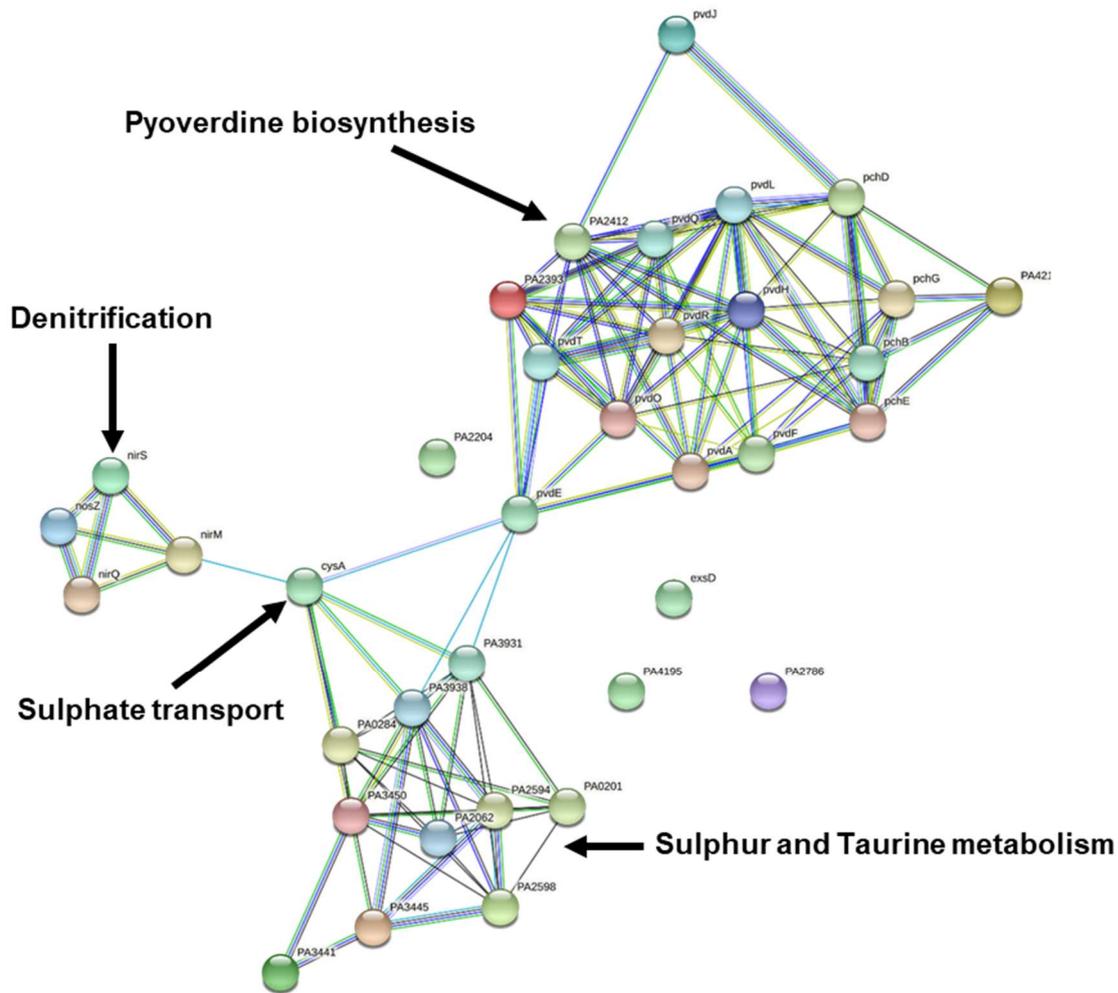


Figure 3.14 Annotated STRING map of proteins down regulated in $\Delta exsD$ mutants.

All proteins which were decreased in abundance in $\Delta exsD$ compared to WT, where that decrease was both statistically significant and greater than \log_2 fold change = 1. The connections are automatically generated by the STRING webservice, and are indicative of shared function or regulation. Point annotations contain the relevant gene name, whilst manual annotations added with arrows indicate known or predicted biological functions (with predicted functions in *italics*) for the relevant data point or cluster.

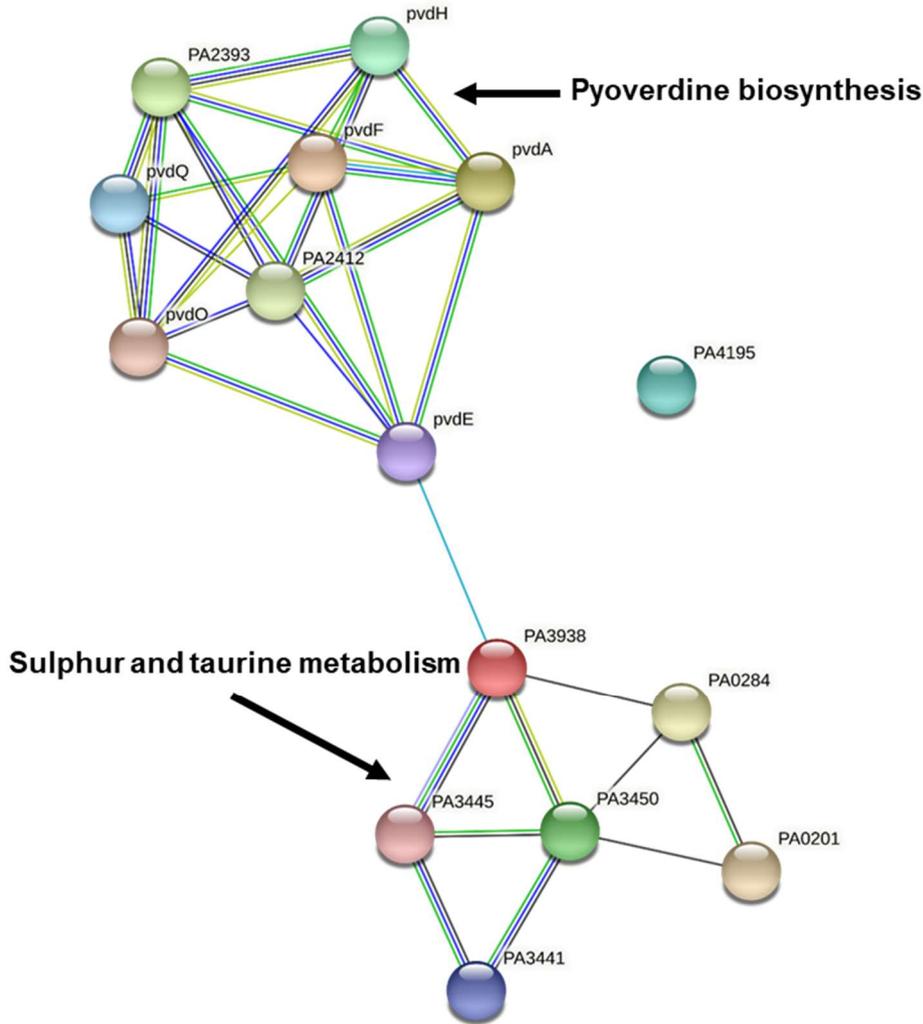


Figure 3.15 Annotated STRING map of proteins downregulated in Δ exsE mutants.

All proteins which were of decreased abundance in Δ exsE compared to WT, where that increase was both statistically significant and greater than Log_2 fold change = 1. The connections are automatically generated by the STRING webservice, and are indicative of shared function or regulation. Point annotations contain the relevant gene name, whilst manual annotations added with arrows indicate known or predicted biological functions (with predicted functions in *italics*) for the relevant data point or cluster.

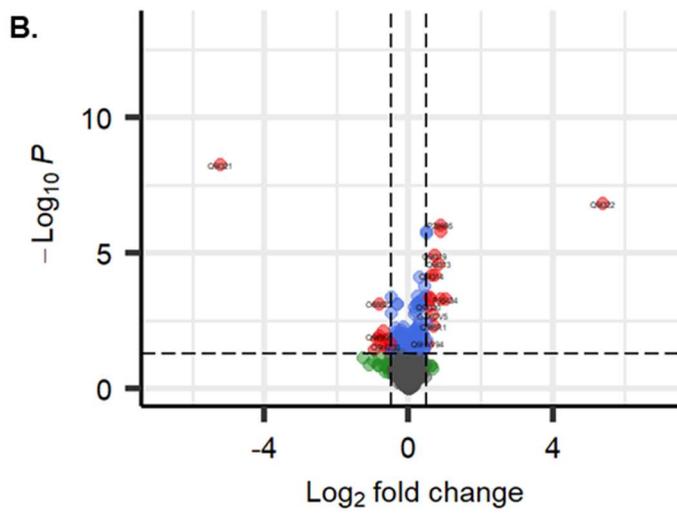
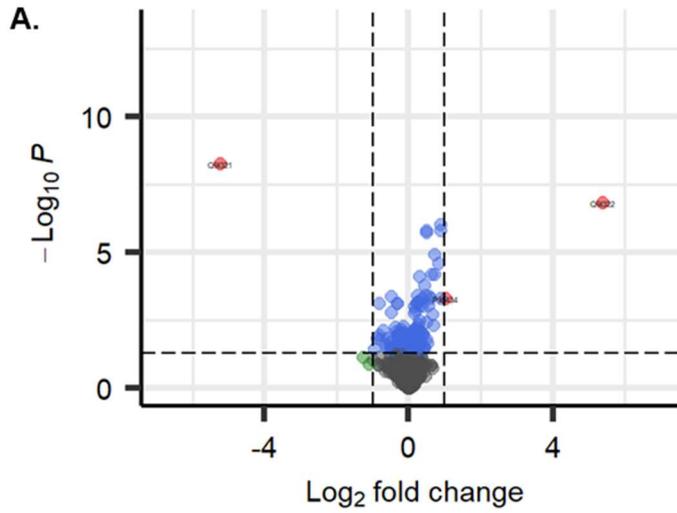


Figure 3.16 Volcano plot comparisons of ΔexsD and ΔexsE .

Volcano plots depicting a comparison between ΔexsD and ΔexsE mutants. Blue and red dots represent statistically significant data points ($p < 0.05$), and red dots are those which meet a Log_2 fold change minimum of ± 1 in **A.** or ± 0.5 in **B.**

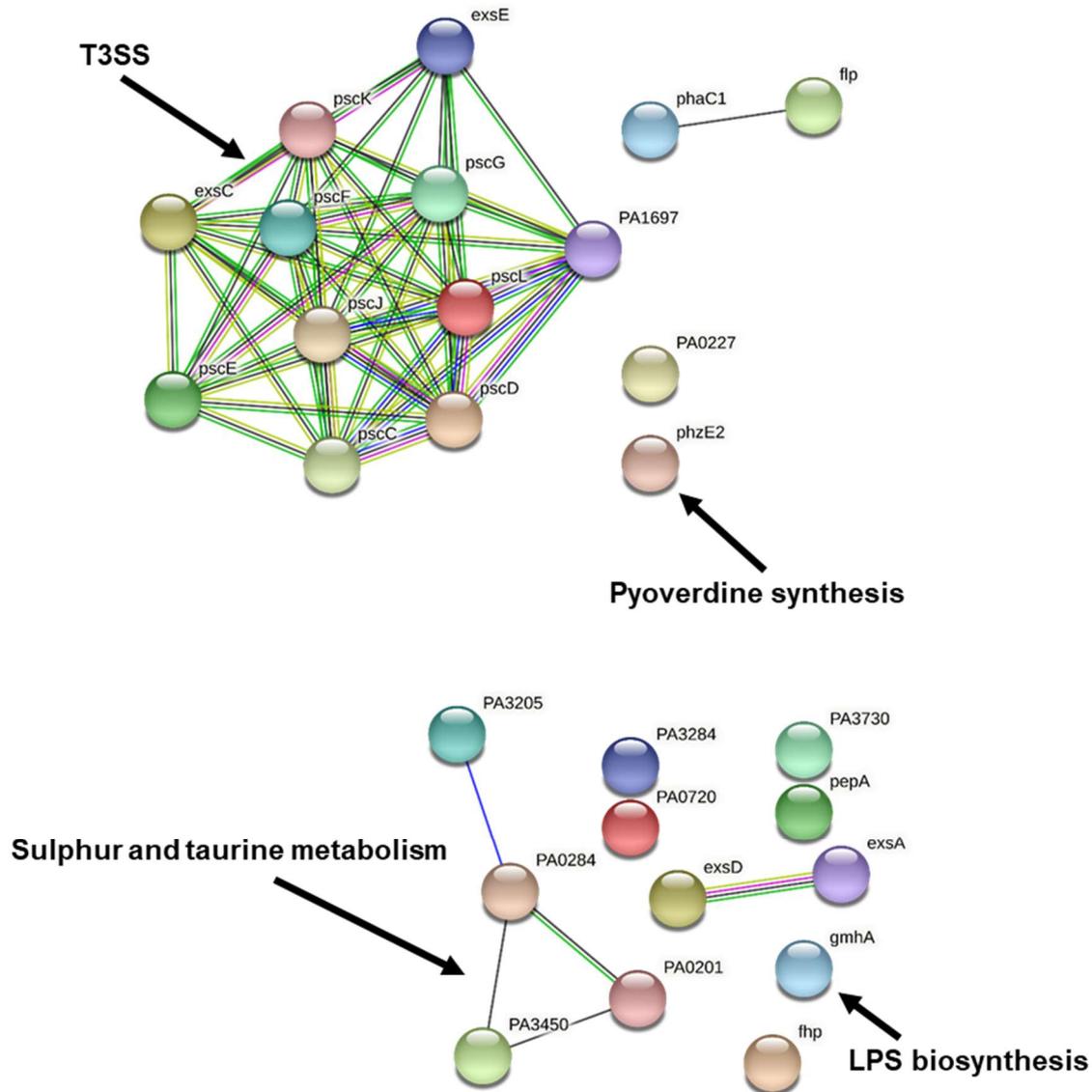


Figure 3.17 Annotated STRING map of proteins differentially expressed between $\Delta exsD$ and $\Delta exsE$.

All proteins which were of increased or decreased abundance in $\Delta exsD$ compared to $\Delta exsE$ (**A.** and **B.** respectively), where that increase was both statistically significant and greater than Log_2 fold change = 0.5. The connections are automatically generated by the STRING webservice, and are indicative of shared function or regulation. Point annotations contain the relevant gene name, whilst manual annotations added with arrows indicate known or predicted biological functions (with predicted functions in *italics*) for the relevant data point or cluster. It is noted that the link indicated between Flp and phaC1 link is tenuous, it is generated due to co-expression of homologs in a different organism.

3.5.4 Δ exsD compared to Δ exsA offers a sharper contrast than compared to wild type

Direct comparison between Δ exsD and Δ exsA revealed an increased number of divergently regulated proteins than between Δ exsD and WT. The total number of proteins to meet the criteria of statistical significance and a Log_2 fold change of greater than one was 52 downregulated and 51 upregulated (Figure 3.18). When mapped to string plots several of the novel upregulated proteins were members of established groups. The small groupings for HCN synthesis, carnitine catabolism, T6SS, and an NRPS operon, all gained an additional member, and the T3SS cluster also grew (Figure 3.19). The full HCN synthesis operon (*hcnABC*) is present. Downregulation saw a more dramatic increase in the number of hits in comparison to Δ exsD compared to WT, with the denitrification, pyoverdine synthesis, and sulphur/taurine metabolism clusters all growing significantly (Figure 3.20). A number of unconnected proteins were also identified, most were either uncharacterised enzymes of no clear function or putative components of ABC transporters.

Proteins with differential regulation between Δ exsA and WT did not necessarily have the inverse observed in the always on mutants (Table 3.3). Q9HVZ0/GmhA and O68822/PepA which are upregulated in Δ exsA mutants have no change in regulation in Δ exsD mutants, and an increase in the Δ exsE mutants. Q9I481/PA1263 has a statistically significant but very minor change in both always on mutants, whilst T3SS protein Q9I318/PscD is substantially upregulated in both.

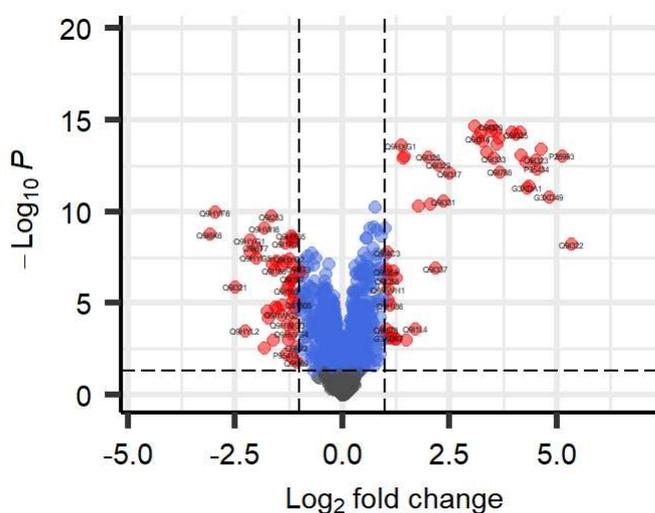


Figure 3.18 Volcano plot comparisons of Δ exsD and Δ exsA.

Volcano plots depicting a comparison between Δ exsD and Δ exsA mutants. Blue and red dots represent statistically significant data points ($p < 0.05$), and red dots are those which meet a Log_2 fold change minimum of ± 1 .

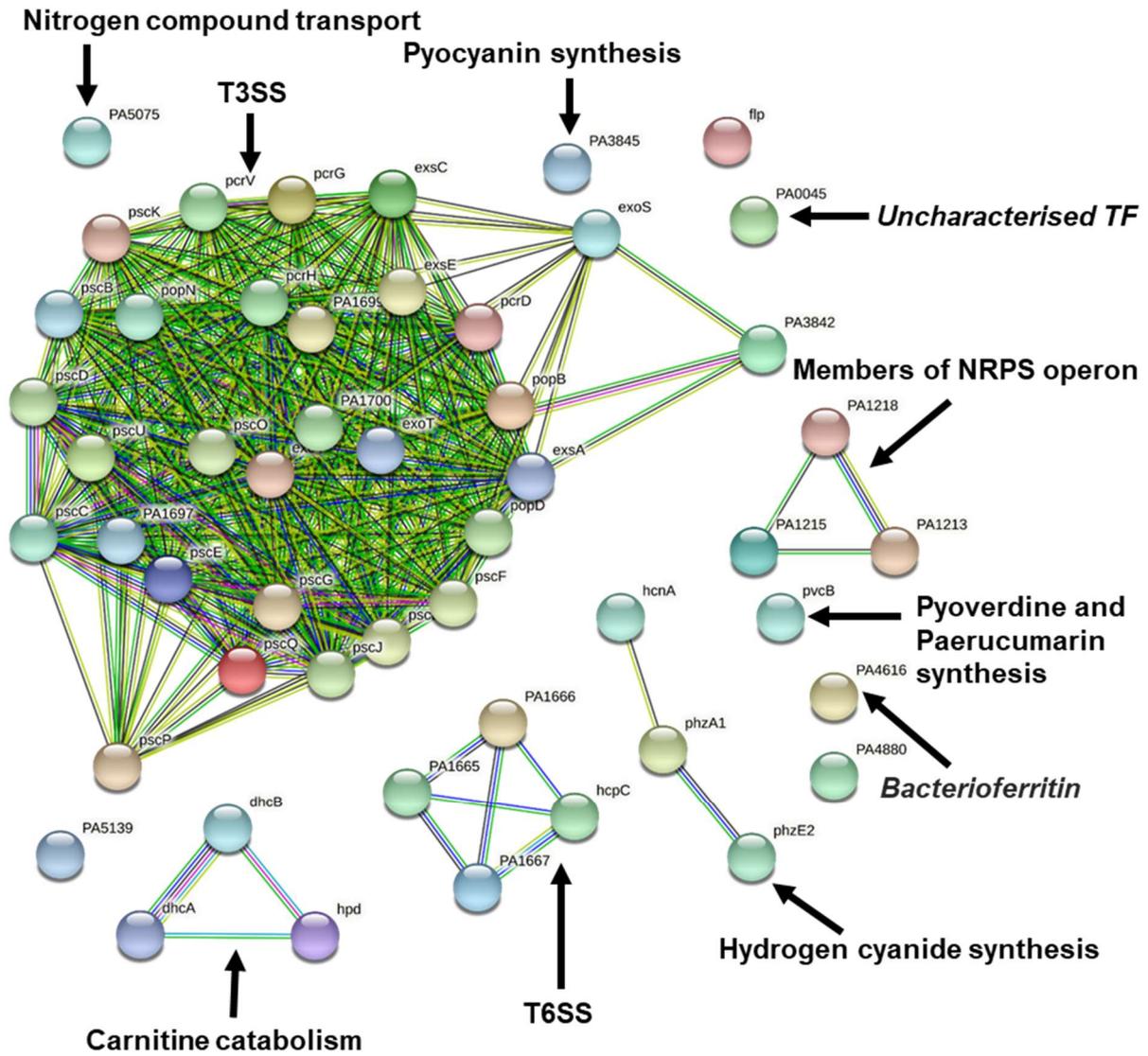


Figure 3.19 String map of upregulated proteins in $\Delta exsD$ compared $\Delta exsA$.

All proteins which were of increased abundance in $\Delta exsD$ compared to $\Delta exsA$, where that increase was both statistically significant and greater than Log_2 fold change = 1. The connections are automatically generated by the STRING webservice and are indicative of shared function or regulation. Point annotations contain the relevant gene name, whilst manual annotations added with arrows indicate known or predicted biological functions (with predicted functions in *italics*) for the relevant data point or cluster.

Table 3.3 Expression change of Δ exsA mutant hits in Δ exsD and Δ exsR mutants

Protein	Statistically significant Log_2 fold change Δ exsD compared to WT	Statistically significant Log_2 fold change Δ exsE compared to WT
Q9HVZ0/GmhA	NA	0.603078
Q9I481/PA1263	-0.1869	-0.29263
Q9I318/PscD	4.005022	3.120777
O68822/PepA	NA	1.08069

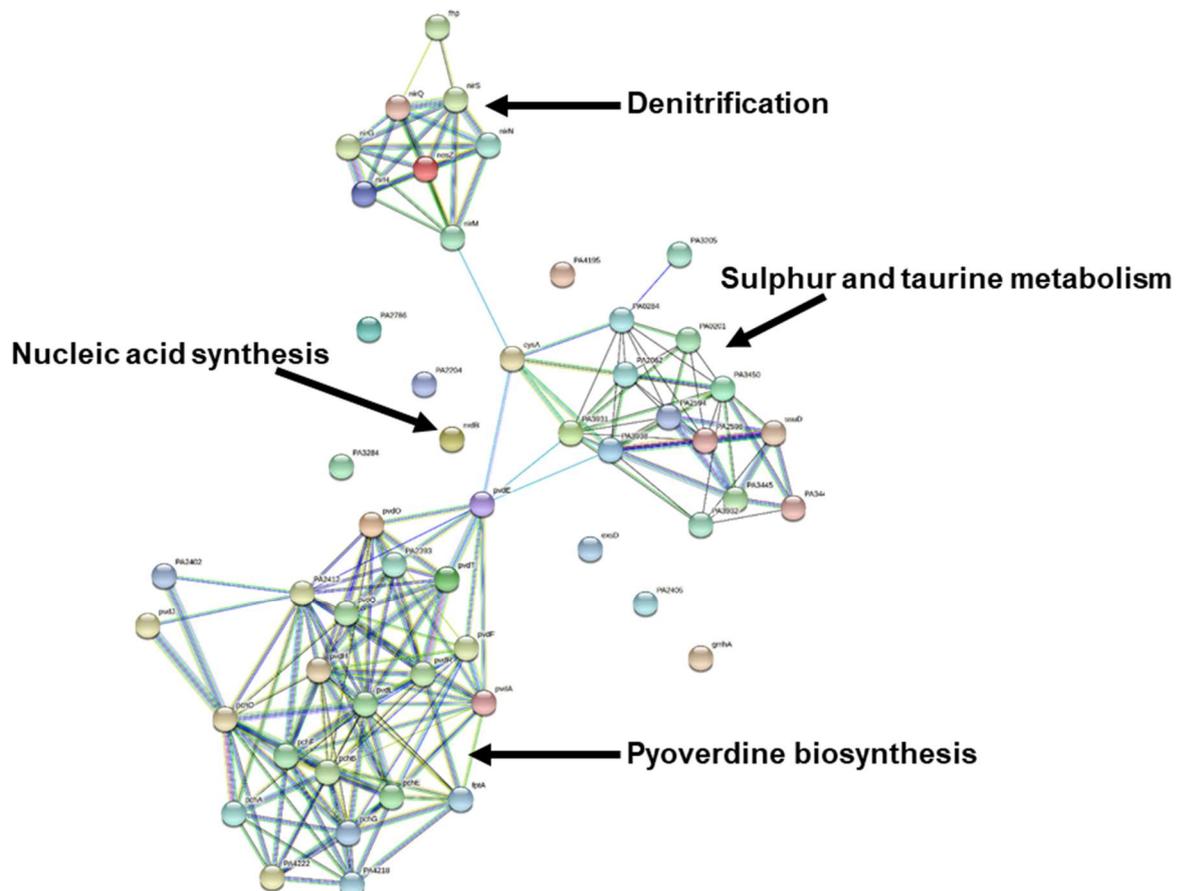


Figure 3.20 String map of downregulated proteins in Δ exsD compared Δ exsA.

All proteins which were of decreased abundance in Δ exsD compared to Δ exsA, where that decrease was both statistically significant and greater than Log_2 fold change = 1. The connections are automatically generated by the STRING webservice and are indicative of shared function or regulation. Point annotations contain the relevant gene name, whilst manual annotations added with arrows indicate known or predicted biological functions (with predicted functions in italics) for the relevant data point or cluster.

3.5.5 Pyocyanin synthesis enzymes are upregulated in “always on” mutants

As noted prior Q7DC81/PhzE1 is upregulated in the “always on mutants”, and Q9HWH2/PhzM is upregulated in the ΔexsA mutant. Both these proteins are part of the pyocyanin synthesis pathway. Further analysis of the pathway was undertaken by comparing the ΔexsD mutant to the ΔexsA mutant. There is a statistically significant difference in protein abundance between the two mutants for all the examined proteins except Q9HWH2/PhzM (Figure 3.21 A). The only protein to be comparatively downregulated in the ΔexsD mutant was Q06553/PtrR, a transcriptional repressor of the pyocyanin synthesis proteins. *Pseudomonas aeruginosa* contains 2 parallel pyocyanin synthesis operons PhzA1-G1 and PhzA2-G2, however barring the respective A and B members the two operons share accession codes and only one could therefore be displayed (and which operon is uncertain).

A pyocyanin extraction was undertaken to confirm the phenotypic consequences of the upregulation of these synthesis proteins, with both ΔexsD and ΔexsE showing statistically significant increases in pyocyanin present compared to wild type PAO1 (Figure 3.21 B). Whilst both the “always off” mutants had lower mean values than wild type this difference was not found to be statistically significant.

3.5.6 The ExsA cascade has a regulatory role in siderophore synthesis

Given the large cluster of down regulated pyoverdine siderophore synthesis proteins in the always on mutants it was decided to determine if the siderophore concentration in the cultural supernatant varied between the mutants accordingly. A stark drop in siderophore activity was observed in the ΔexsD mutant, but not in the ΔexsE mutant which retained equivalent iron chelation to WT and the always off mutants (Figure 3.22).

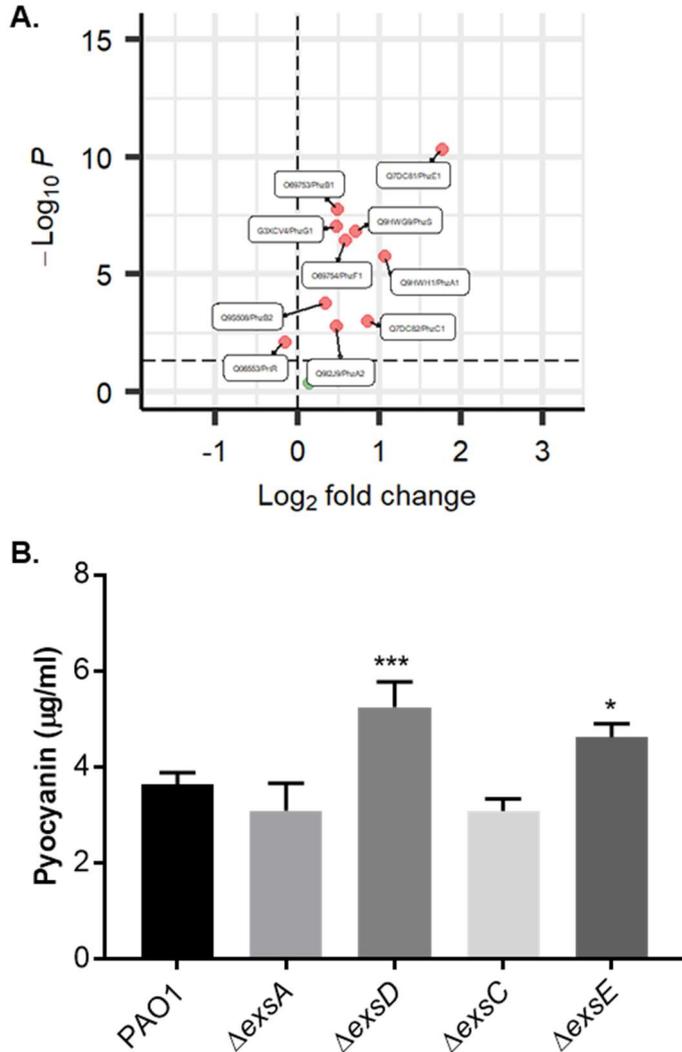


Figure 3.21 Pyocyanin synthesis volcano plot and pyocyanin extraction variation between mutants.

A. Volcano plot comparing ΔexsD and ΔexsA mutants for every pyocyanin synthesis proteins and one regulatory protein. Red dots represent statistically significant data points ($p < 0.05$), the green dot represents Q9HWH2/PhzM, the only protein examined which does not diverge with statistical significance. Q06553/PrtR is a negative regulator of pyocyanin synthesis, and the only protein to be downregulated in ΔexsD compared to ΔexsA in this data set.

B. Pyocyanin concentrations from cultural supernatant in the same conditions as the proteomic experiments. Statistical analysis was conducted with a one way ANOVA, followed by individual comparisons of each mutant to PAO1. * indicates a P value of < 0.05 and *** indicates a p value of < 0.0005 . Each bar represents a mean of three biological replicates, error bars indicate SD.

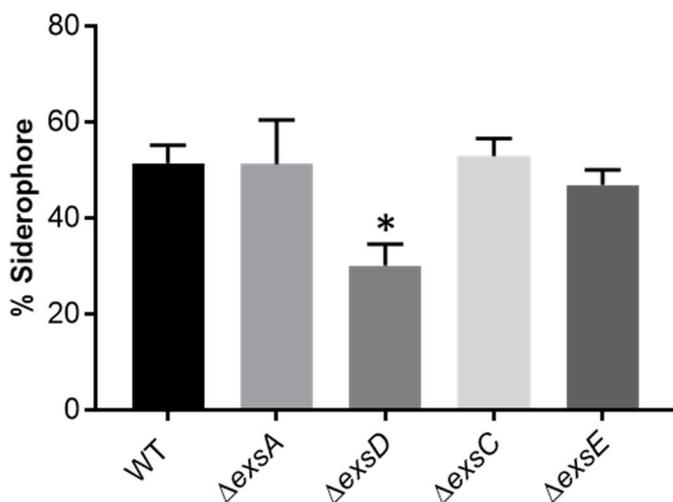


Figure 3.22 Siderophore quantification assay.

Relative siderophore concentrations from cultural supernatant in the same conditions as the proteomic experiments. Statistical analysis was conducted with a one way ANOVA, followed by individual comparisons of each mutant to PAO1. The units % siderophore are calculated using the sample as a percentage of a siderophore free reference ($\frac{\text{absorbance of reference} - \text{absorbance of sample}}{\text{absorbance of reference}} \times 100$). * indicates a P value of < 0.05. Each bar represents a mean of three biological replicates, error bars indicate standard deviation.

3.5.7 The ExsA cascade has a regulatory role in denitrification

As observed above a cluster of denitrification proteins were present at reduced levels in the ΔexsD mutant in compared to the WT and this was larger in comparison to the ΔexsA mutant. However, no cluster was observed in ΔexsE compared to wild type, yet there was also no significant dysregulation of denitrification proteins apparent between ΔexsD and ΔexsE proteins (Figure 3.16). Indeed, there is no significant change compared to WT for P24474/NirS (p=0.36).

An initial attempt to explore the link between the expression of denitrification proteins and was undertaken using a luciferase reporter under the control of the *nirS* promoter, with EGTA induction utilised to stimulated ExsA activity. Three reported strains available within the lab were used, *PnirS*-Lux to explore possible effects of induction on NirS expression, and *PpcrV*-Lux and *PlacZ*-Lux as a positive and negative control for ExsA mediated expression respectively. Initially a media known to stimulate NirS expression (MOPS + 40 mM acetate) was employed, however no increase in luciferase activity was observed in the positive control during induction (Figure 3.23). Glucose was subsequently added to the media, however the

same issue persisted (Figure 3.24). Finally, AGSY media was utilised to support robust activity in the control. Here the positive control behaved as expected, however the negative control and both test runs showed a reduced luciferase activity (Figure 3.25).

An alternative validation was undertaken utilising the same conditions as the proteomics experiment and measuring nitrite concentration from the supernatant at 8 hours. The results showed a reduced concentration of nitrite in the always off mutants, whilst the always on mutants had levels consistent with that of the WT (Figure 3.26).

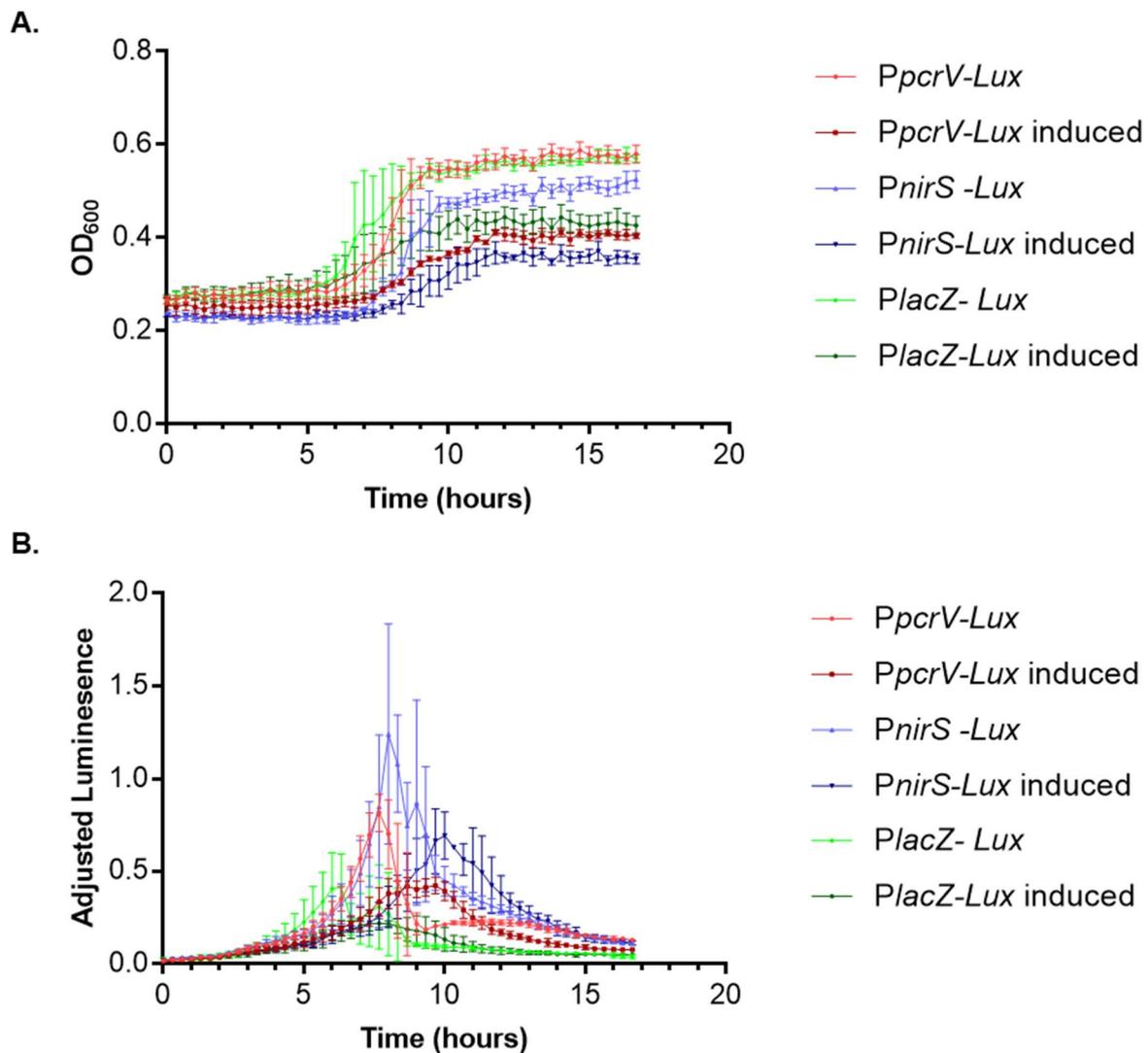


Figure 3.23 NirS Luciferase reporter activity in MOPS with acetate.

Growth curves (**A.**) and OD₆₀₀ adjusted luminescence data (**B.**) for a reporter strains in MOPS + 40 mM acetate at 37 °C with shaking. Induced samples contain 5 mM EGTA. Data points represent the mean of three biological replicates and error bars depict standard deviation.

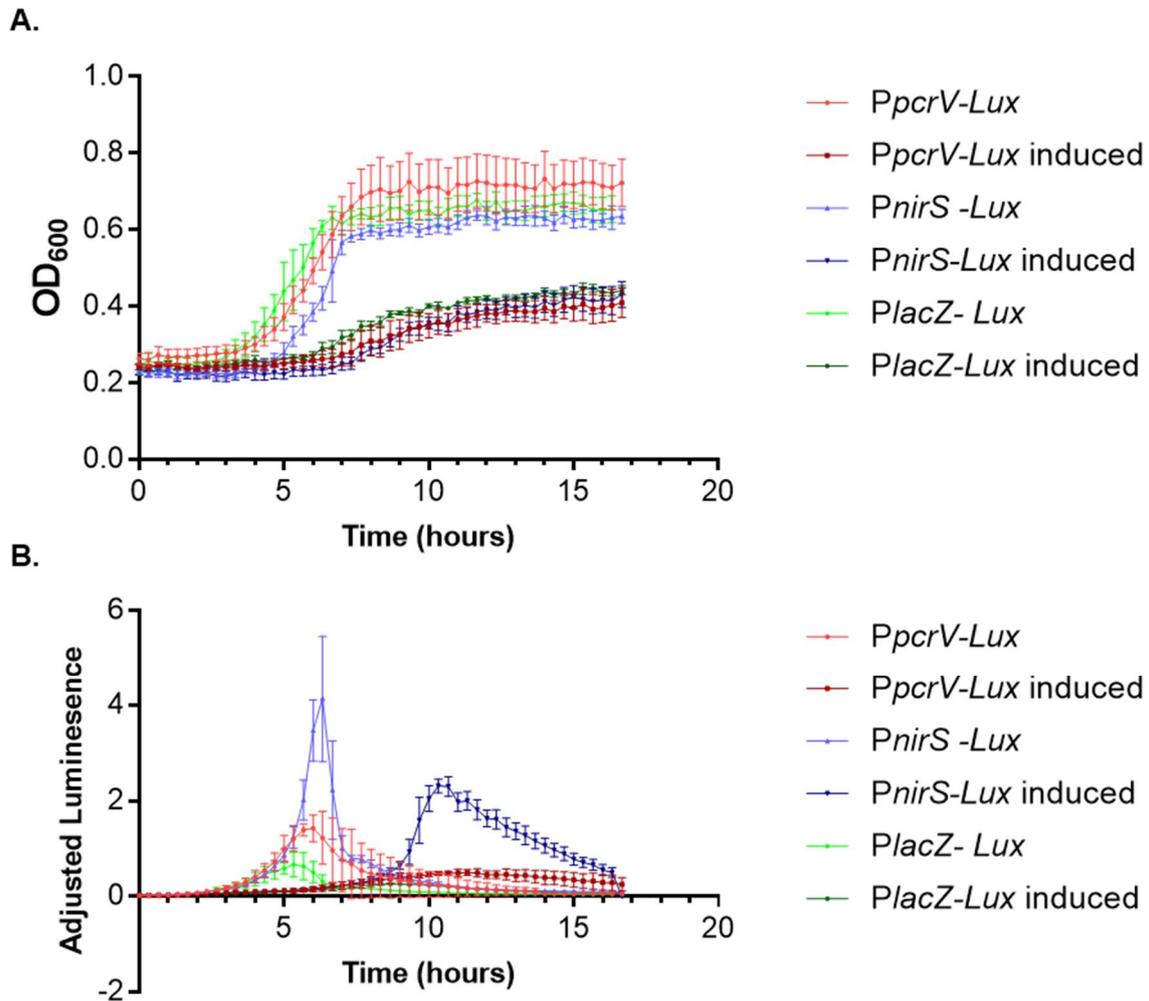


Figure 3.24 NirS Luciferase reporter activity in MOPS with acetate and glucose.

Growth curves (**A.**) and OD₆₀₀ adjusted luminescence data (**B.**) for a reporter strains in MOPS + 40 mM acetate + 15 mM glucose, at 37 °C with shaking. Induced samples contain 5 mM EGTA. Data points represent the mean of three biological replicates and error bars depict standard deviation.

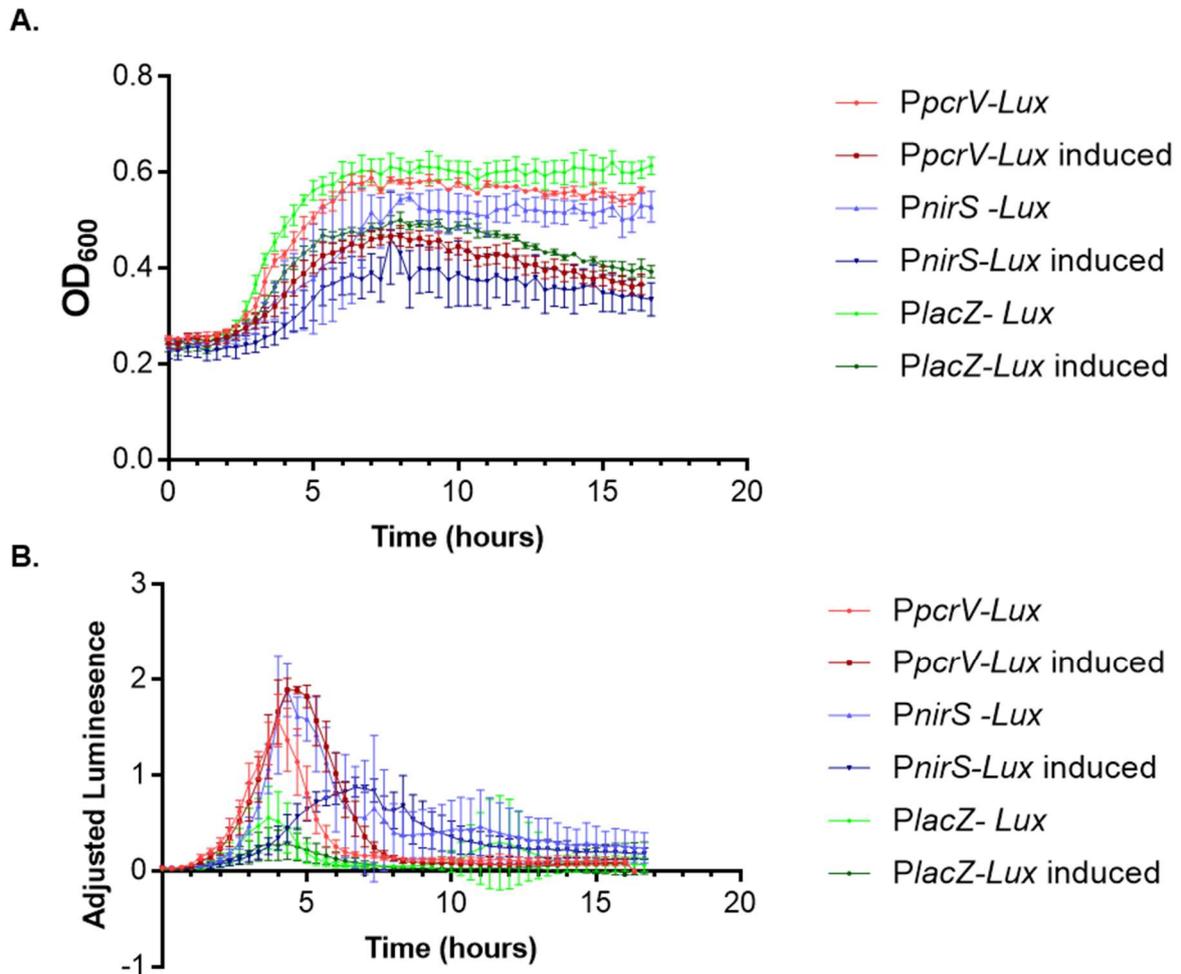


Figure 3.25 NirS Luciferase reporter activity in AGSY.

Growth curves (**A.**) and OD₆₀₀ adjusted luminescence data (**B.**) for a reporter strains in AGSY at 37 °C with shaking. Induced samples contain 5 mM EGTA. Data points represent the mean of three biological replicates and error bars depict standard deviation.

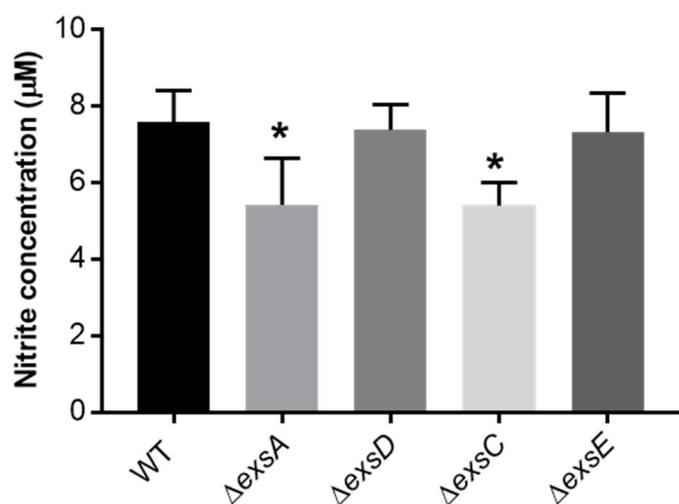


Figure 3.26 Nitrite detection assay.

Nitrite concentrations from cultural supernatant in the same conditions as the proteomic experiments. Statistical analysis was conducted with a one-way ANOVA, followed by individual comparisons of each mutant to PAO1. * indicates a P value of < 0.05. Each bar represents a mean of three biological replicates, error bars indicate standard deviation.

3.5.9 The ExsA cascade has a limited impact on broader virulence signalling

Within the always on mutants there is a small but statistically significant increase in the abundance of several quorum sensing (QS) proteins, especially the members of the PQS system (Figure 3.27). Comparison between $\Delta exsD$ and $\Delta exsA$ mutants and the “always on” mutants compared to WT are similar, with a more distinct change being observed for the $\Delta exsD$ mutant than $\Delta exsE$ mutant. Only P33883/LasI and P54292/RhIR are do not diverge in a statistically significant manner between the $\Delta exsD$ and $\Delta exsA$ mutants (Figure 3.27).

Some proteins within the Gac MKN have a small but statistically significant decrease in abundance in $\Delta exsD$ mutant compared to $\Delta exsA$ mutants (Figure 3.28), however it is not apparent for all of them, nor is there a drive towards pro-acute or pro-chronic proteins being under expressed.

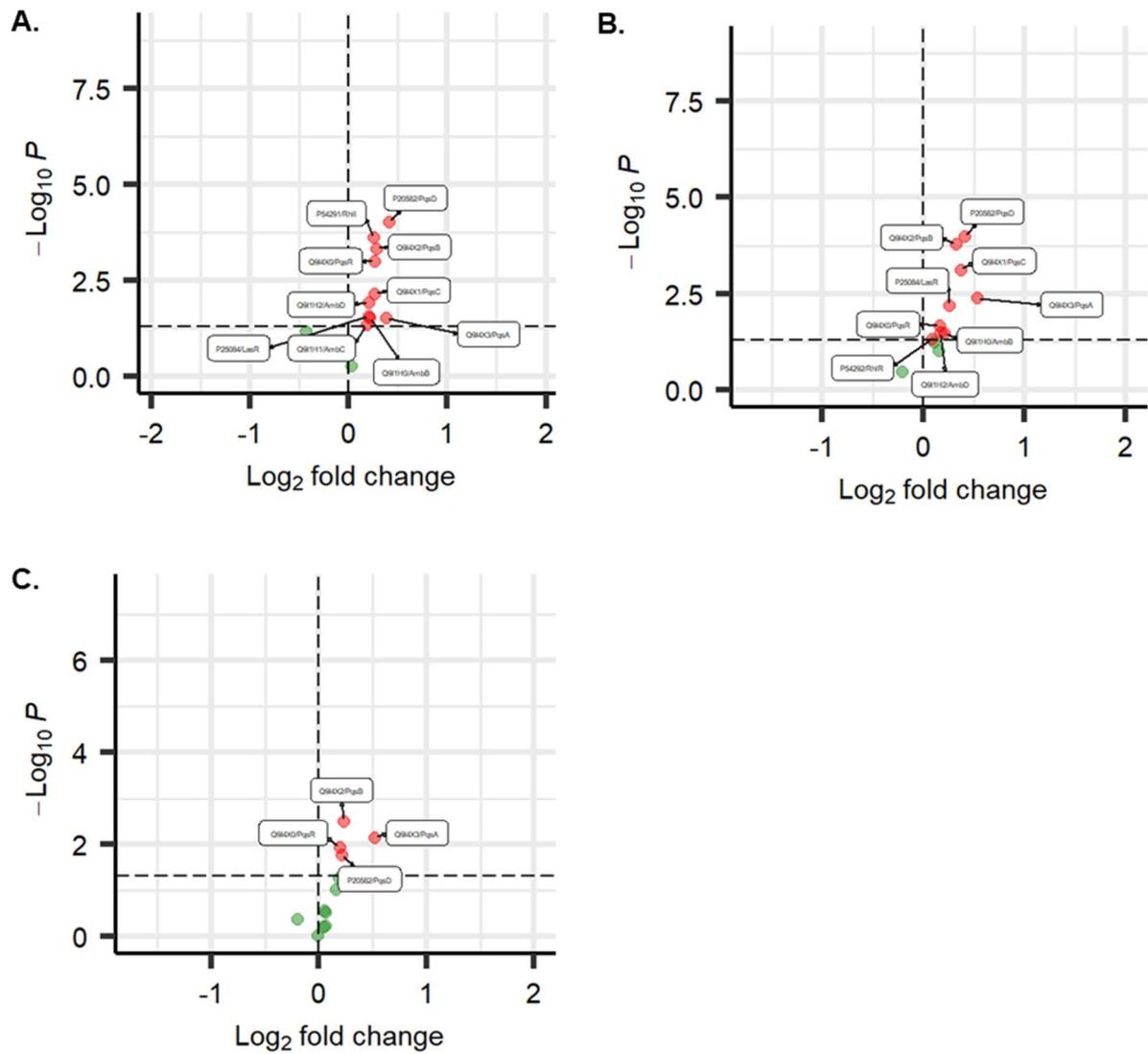


Figure 3.27 Proteomic data for quorum sensing proteins.

Volcano plot comparisons for quorum sensing proteins. Red dots represent statistically significant data points ($p < 0.05$).

A. Δ exsD and Δ exsA compared.

B. Δ exsD and WT compared.

C. Δ exsE and WT compared.

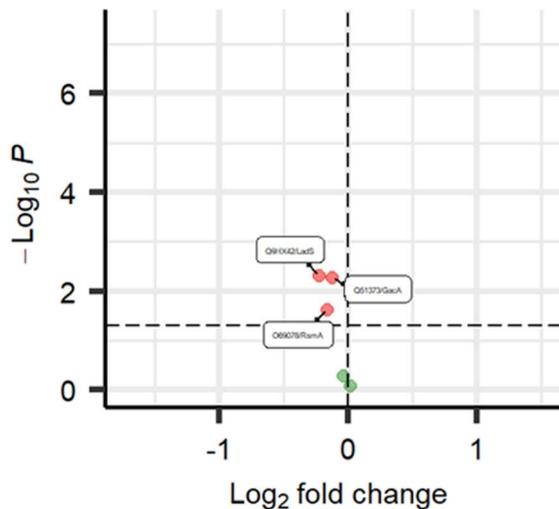


Figure 3.28 GacA network protein abundance comparison for Δ exsD and Δ exsA mutants.

Volcano plots depicting a comparison between Δ exsD and Δ exsA mutants for proteins selected as members of the chronic/acute switch network surrounding GacA. Red dots represent statistically significant data points ($p < 0.05$).

3.5.10 Not all reported members of the ExsA regulon are differentially regulated

The reported regulation of T1SS effector Q9I5W4/ImpA by ExsA was part of the rationale for this project (Tian *et al.*, 2019), however the protein had no divergence in its abundance between mutants (Table 3.4).

The most extensive prior examination of the ExsA regulon was a chromatin immunoprecipitation sequencing (chIP-seq) investigation, which found 17 promoter regions bound by ExsA (Huang *et al.*, 2019). The non-T3SS promoter regions detected which are prior to a proteins coding gene were retrieved from the proteomic data comparing the Δ exsD and Δ exsA mutants. Q9HI36/HcpA was used as a negative control in the prior study, with no record of binding, however in the present study it was found to be statistically significantly upregulated by above a Log_2 fold change threshold of 1. None of the other proteins met these criteria, however Q9I3G0/ccON2, Q9HTV1/Rho, and G3XD85/WbpH were all statistically significantly downregulated in the Δ exsD mutant compared to the Δ exsA to a minor extent. When compared to WT the Q9I3N8/PA1474 and O54438/FabG were also statistically significantly downregulated in the Δ exsD mutant to a minor extent.

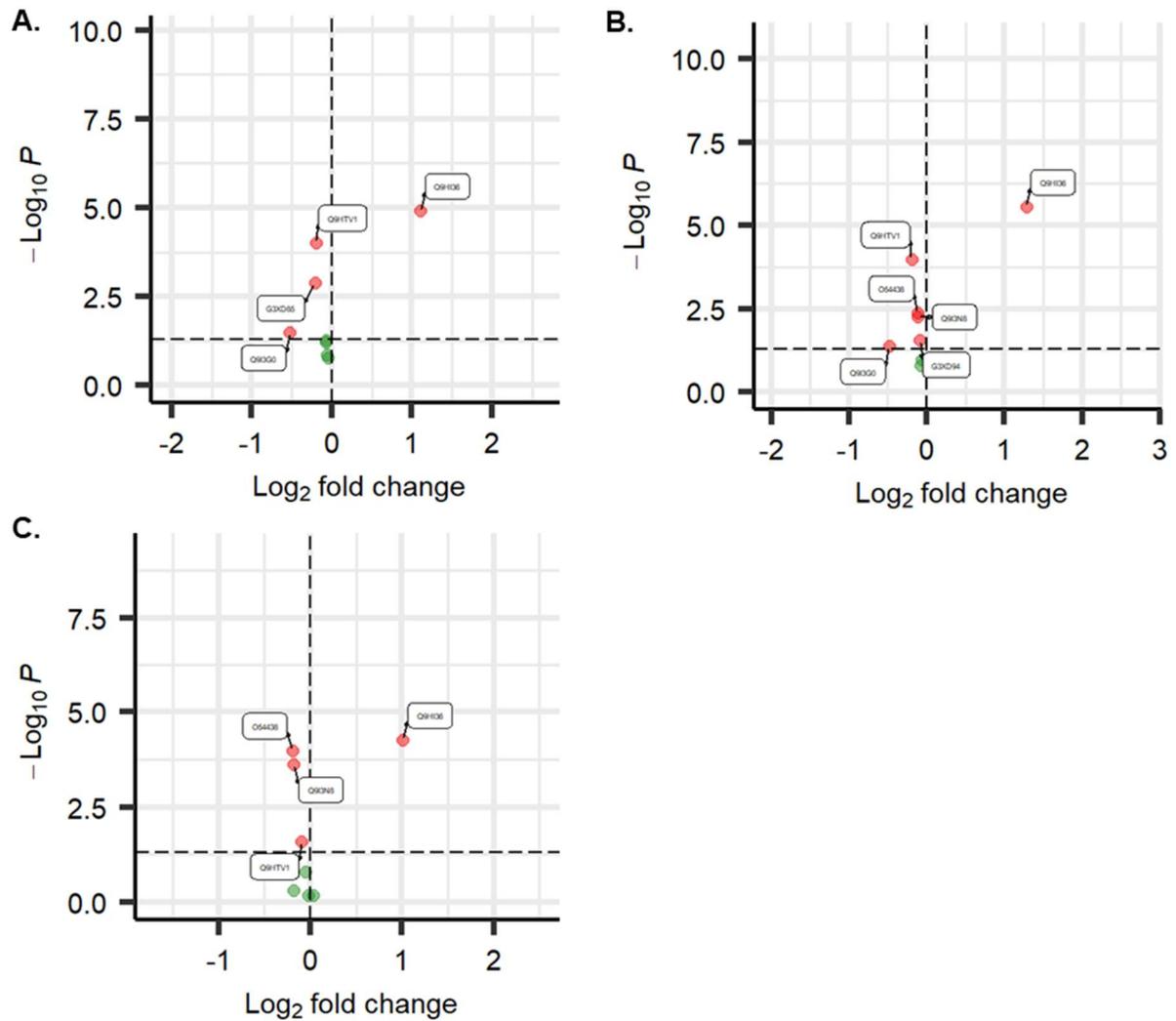


Figure 3.29 chIP-seq hits compared to proteomic data.

Volcano plots showing the comparison of the Δ exsD and Δ exsA data sets for the proteins corresponding to the chIP-seq hits reported for ExsA (Huang et al., 2019). Labelled red dots correspond to results which met a statistical significance threshold of $p < 0.05$.

A. Δ exsD and Δ exsA compared.

B. Δ exsD and WT compared.

C. Δ exsE and WT compared.

Table 3.4 Proteomic results for ImpA

Comparison	Log ₂ fold change	Adjusted P value
$\Delta exsD$ to WT	-0.16989	0.045669
$\Delta exsA$ to WT	-0.15533	0.450685
$\Delta exsE$ to WT	0.001793	0.99076
$\Delta exsD$ to $\Delta exsA$	-0.01456	0.880903

3.6 Discussion

3.6.1 Interpretation of preliminary results

The role of the ExsA cascade in regulating the T3SS in *P. aeruginosa* is well characterised, the $\Delta exsD$ and $\Delta exsE$ mutants were therefore predicted to have a T3SS “always on” overexpression phenotype, whilst the $\Delta exsA$ and $\Delta exsC$ mutants were predicted to have a T3SS “always off” phenotype. The initial mutant phenotypes conform to this existing knowledge, with the “always on” mutants $\Delta exsD$ and $\Delta exsE$ were shown in western blot to produce more PcrV, a component of the T3SS, than WT. Conversely the “always off” mutants $\Delta exsA$ and $\Delta exsC$ were deficient for PcrV production. The slight growth defect observed in the always on mutants is explicable; T3SS production is an energy and resource intensive process (hence the tight regulation in WT), by causing over expression of the T3SS the always on mutations deprive the cells of energy that could be used to grow and replicate.

The phenotypic distinctions between $\Delta exsA$ and $\Delta exsC$, and between $\Delta exsD$ and $\Delta exsE$, shows stronger phenotypes in $\Delta exsA$ and $\Delta exsD$, which are further down the cascade than $\Delta exsE$ and $\Delta exsC$. The Western blots demonstrating this are further supported by the general distinction in the proteomic data for $\Delta exsD$ and $\Delta exsE$, whereby a more extreme phenotype is observed in the $\Delta exsD$ mutant. This provides material evidence for targeting ExsA rather than attempting an intervention higher up the cascade, for instance by targeting $\Delta exsC$ or attempting to prevent the release of $\Delta exsE$.

With the preliminary characterisation of the mutants complete, conditions for the proteomic comparison were sought. The ASM continuous flow model proved unsuitable for further investigation of the ExsA cascade, as it is unlikely to show a distinction between $\Delta exsA$ and PAO1 and seemed to favour the expression of the T6SS over the T3SS. This is unsurprising

as it is a model of chronic virulence and the T3SS is associated with acute virulence. Whilst other available culturing methods are less physiologically relevant, it was decided upon to test them with a view towards highlighting phenotypic distinctions more clearly.

Comparison of LB and the richer AGSY showed a clear preference in favour of AGSY as it supported both a clearly distinct phenotype between each genotype, higher T3SS expression than LB, and the apparent growth defect in the always on mutants was less pronounced. Across the time course of the culture, it was clear that a time after 6 hours would be needed for the proteomic investigation as that is when the T3SS became detectable. The 8th hour was selected to ensure T3SS expression was established in the cultures.

3.6.2 Virulence and signalling implications of ExsA activity

In a similar manner to the preliminary investigations, one of the proteomic results was highly predictable. In both “always on” mutants the T3SS was expressed to a far greater degree than in the WT. However, the Δ exsA mutant did not show a downregulatory effect on the T3SS genes in comparison to the WT, though when compared to the Δ exsD mutant directly a more substantial dysregulation was observed than when the Δ exsD mutant is compared to WT. This indicates that much of the change in the Δ exsA mutant was beneath the detection threshold of the experiment in comparison to WT, but that a biological difference was present.

Of the few proteins to be dysregulated in the Δ exsA mutant, and the only one to have a Log₂ fold change of >1 was Q9HVZ0/GmhA, an enzyme involved in the production of virulence factor lipopolysaccharide (LPS) (Taylor *et al.*, 2008). Whilst LPS is a virulence determinant with a role in both chronic and acute infection (Pier, 2007), the fact that it is also upregulated in Δ exsD, and the lack of a more general trend in the proteomic detection of the wider LPS synthesis pathway it is hard to infer anything from this isolated change. The modest upregulation of Q9I318/PscD and O68822/PepA in the Δ exsA mutants is overshadowed by their much larger upregulation in the Δ exsD mutants. That said, the upregulation of T3SS protein O68822/PepA is surprising, and possibly due to a secondary effect. One potential source of secondary effects in ExsA's absence is the downregulation of the uncharacterized TF Q9I481/PA1263, although as it is also modestly downregulated in the absence of ExsD it again doesn't offer a clear insight. In short there is little to be discovered from the direct comparison of Δ exsA mutants and WT in the conditions utilised, despite promising preliminary results via Western blot. Conditions more conducive to T3SS expression, such as AGSY with EGTA added to cause Ca²⁺ depletion, would create greater distinction between the Δ exsA mutant in the two strain's proteomes, however the stark contrast with the Δ exsD mutants would have then been lost.

The “always on” mutants had higher Hcp1 expression than WT, as observed both in the Western blots and proteomics, and a number of other T6SS proteins and effectors including Q9HI36/HcpA were found to be upregulated in the “always on” mutants in the proteomic experiments. The T3SS and T6SS are generally thought of as inversely regulated, particularly through the GacS network which controls the acute-chronic lifestyle switch (Chambonnier *et al.*, 2016; Francis *et al.*, 2018). That the “always on” mutants showed increased T6SS expression indicates that the ExsA cascade does not cross talk with this network specifically and possibly that the ExsA cascade is downstream of this switch, which is further supported by the proteomic data on members of the GacS network, where no clear distinction which would affect the pathway’s signalling was seen between the Δ exsA and Δ exsD mutants. This does not however indicate whether the pathway is actively signalling or not. Phenotypically it appears unlikely that the GacS network is the cause of increased expression of some T6SS proteins because an increase across a wider array of T6SS proteins (like that observed for T3SS proteins) would be expected in that instance.

Hydrogen cyanide production proteins were also upregulated in the “always on” mutants. HCN is a *P. aeruginosa* virulence factor, the regulation of which not previously been attributed to the ExsA cascade. Unfortunately, facilities for the detection of HCN production were not available, leaving further investigation of the cascade’s role in its regulation an open question.

To my knowledge, the link between the ExsA cascade and pyocyanin has also not been previously reported. In this study we have demonstrated that a number of pyocyanin synthesis proteins are upregulated in “always on” mutants, and that there is an observable increase in pyocyanin concentration in the mutant cultures as a result. Pyocyanin is a long-established virulence factor (Allen *et al.*, 2005; Hall *et al.*, 2016), with a complex web of regulation (Little *et al.*, 2018; Thees *et al.*, 2021). Whilst it is not clear how the ExsA cascade fits into the wider regulation of pyocyanin, it is evident that the rampant ExsA activity induced in an ExsD deficient mutant increases the virulence factor’s production. This suggests that, alongside the other non-T3SS virulence factors increasingly ascribed to the ExsA regulon, ExsA offers a generalist anti-virulence drug target as well as specific target for the repression of the T3SS.

Other upregulated proteins included the NRPS operon PA1221-PA1211. Whilst the cluster remains broadly uncharacterized (Gulick, 2017), PA1221 has been crystallised and contains adenylation and peptidyl carrier protein domains, though the final product remains to be determined (Mitchell *et al.*, 2012). Finally, regulation via QS has previously been reported for the operon (Hendrix *et al.*, 2022). The combination of ExsA and QS regulation strongly imply

that the NRPS creates peptides that are useful in infection or are directly virulence factors.

Two uncharacterized TFs were also detected in increased abundance in the Δ *exsD* mutant proteome. Without a focused investigation it is difficult to speculate on their role, but they certainly imply downstream regulatory effects of ExsA activity. Conceivably the transcription of these TFs is directly regulated by ExsA. Whether their increased abundance in Δ *exsD* mutant were a direct function of increased ExsA activity or an indirect effect itself they offer a possible route for the expansion of ExsA's regulatory influence beyond the protein's direct transcriptional regulon.

These TFs could explain a discrepancy between the proteomic data presented here, in which the Δ *exsD* mutant shows increased Q9HI36/HcpA, and chIP-seq data previously published which found that ExsA does not bind to the Q9HI36/HcpA promoter (Huang *et al.*, 2019). Given this discrepancy it is strongly implied that the increased presence of T6SS proteins in the "always on" mutants is via a secondary mechanism downstream of ExsA. It is also possible that other virulence factors which are upregulated in the "always on" mutants are similarly regulated by a factor downstream of ExsA.

The same chIP-seq study identified several more non-T3SS factors regulated by ExsA (Huang *et al.*, 2019). The results presented here partially agree, with O54438/FabG and Q9I3G0/ccoN2 being negatively regulated in both data sets. Several other factors which were only reported as bound by ExsA were found to be negatively regulated in the proteomic data. Other proteins reported to be positively regulated by ExsA, such as Q9HTV1/rho were negatively regulated in "always on" mutants in the proteome comparisons. There were also proteins reported to be regulated by ExsA which had no statistically significant change. Whether statistically significant or not, with the exception of Q9HI36/HcpA, any change was small, lower than a Log_2 fold change of 0.5.

The gap between this small change and the very large change observed in other proteins which were not reported in the chIP-seq study is intriguing. Whilst the canonical members of the ExsA regulon were identified in the chIP-seq study (Diaz, King and Yahr, 2011; Huang *et al.*, 2019), it did not report the positive regulation of Q9I5W4/ImpA which has been previously reported as ExsA regulated (Tian *et al.*, 2019). It therefore appears that Huang *et al.*'s study does not provide a comprehensive list of promoters to which ExsA binds. The culture conditions of the present study and Huang *et al.*'s chIP-seq work vary, with $\text{OD}_{600} = 0.8$ in AGSY media and $\text{OD}_{600} = 0.6$ in LB media being used respectively (Huang *et al.*, 2019). The discovery of ExsA's regulation of ImpA was independent of cultural methods (Tian *et al.*,

2019). Whilst the variation of methods cannot be ruled out as causative of the discrepancies given that to the best of our knowledge ExsA activity is only regulated by the inhibitory binding of ExsD it seems unlikely that different promoters would be bound by the TF in different cultural conditions.

Huang *et al.*'s ChIP-seq investigation also identified the regulatory RNA PhrS as negatively regulated by ExsA (Huang *et al.*, 2019). In turn PhrS (which as a regulatory RNA is not reported on in the present data set) has been shown to positively regulate the transcription of the QS system in which PqsR features in an ANR dependant manner (Sonnleitner *et al.*, 2011). The heightened levels of expression of PqsR in “always on” mutants demonstrated in the proteomics data set do not conflict with this result as ANR activity is only prevalent in anaerobic conditions (Zimmermann *et al.*, 1991), and thus neither ANR nor PhrS are likely to be active in the highly aerobic conditions employed. ANR transcriptionally promotes HCN activity and denitrification (Zimmermann *et al.*, 1991), which are differentially regulated by “always on” mutants in the present study. There is also a growing appreciation of the relationship between ANR, QS, and virulence phenotypes (Hammond *et al.*, 2015b; Clay *et al.*, 2020). The regulatory link between ExsA, ANR, and QS mediated by PhrS is therefore of added interest given the overlap between the ExsA and ANR regulons highlighted above.

Q9I5W4/ImpA, a T1SS metalloprotease, was absent from the upregulated proteomes of “always on” mutants despite previous *in vivo* evidence of ExsA regulation (Tian *et al.*, 2019). Two possible explanations for this are: a post transcriptional regulation of Q9I5W4/ImpA or ImpA's rapid export from the “always on” mutants and thus the exclusion of the increased abundance from the present data set. Whilst both T3SS and T6SS secreted effectors were observed in increased abundance within the “always on” mutants these systems are designed for cell-to-cell effector delivery, for which there were no targets. Conversely T1SS secreted proteins are for extracellular deployment, and thus could have been exported from the cells into the cultural supernatant. In either case the implication that more virulence factors directly regulated by ExsA at the transcriptional level remain undetected in this study is clear.

The influence of the ExsA cascade appears to extend to an endorsement of QS, with almost all QS proteins being upregulated in the Δ exsD mutant. In theory the increased abundance of both signal production and detection proteins should lower the cell density threshold for the activation of the QS regulon, which is comprised of a plethora of virulence factors. It does not necessarily indicate the causation of QS associated phenotypes such as pyocyanin production given that the cell density is high at the relevant time point, so QS is likely to be active in each genotype.

To summarise, the diversity of upregulated proteins in the Δ exsD mutants demonstrates that ExsA upregulates the expression of a variety of non-T3SS virulence factors on top of its canonical T3SS regulon. ExsA therefore offers a generalist anti-virulence target as well as a specific target for the repression of the T3SS.

3.6.3 Downregulation in “always on” mutants

A large cluster of proteins linked to sulphur and taurine metabolism were downregulated in the always on mutants. The causation and impact of this change is very much unclear, and further investigation of such a metabolic shift is beyond the scope of the present study. There is no prior indication of ExsA directly regulating the transcription of any similar metabolic genes, however it has been demonstrated that ExsA can downregulate transcription at certain promoters (Huang *et al.*, 2019), so direct regulation is a possibility for all downregulated proteins. The protein Q9I6L0/CysA, which is connected to this cluster as well as P00099/NirM (denitrification) and Q9I183/PvdE (siderophore synthesis), is a sulphur transported and should be considered part of the sulphur metabolism cluster. Due to the links from Q9I6L0/CysA to all three downregulated clusters it would be easy to ascribe false importance to marginal connections by focusing on this protein. The links between Q9I6L0/CysA, P00099/NirM, and Q9I183/PvdE are drawn due both being present in KEGG maps of other organism's sulphur metabolisms, and there is no specific association known in *P. aeruginosa*.

The “always on” mutants had large cluster of downregulated proteins for siderophore synthesis. Siderophores are used to sequester Fe^{3+} and facilitate its uptake by *P. aeruginosa*, both depriving other organisms of the mineral and increasing its availability to *P. aeruginosa* (Saha *et al.*, 2013). In aerobic environments, such as the shaking flasks used in the present study, Fe^{3+} is the main ionic state of iron, whilst the more bioavailable Fe^{2+} is predominant in deoxygenated environments. Siderophore production involves a large number of enzymes and has an extensive regulatory network (Cornelis, Matthijs and van Oeffelen, 2009; Cornelis *et al.*, 2022). Interestingly pyocyanin and other phenazines are known to oxidise Fe^{3+} to Fe^{2+} , which can be readily up taken by *P. aeruginosa* without the aid of siderophores (Wang *et al.*, 2011). It is therefore possible that in the iron rich and aerobic conditions examined the increase in pyocyanin production causes an increased bioavailability of iron and therefore a reduction in siderophore synthesis.

Siderophores are an expensive investment for bacteria, as attested by “anti-social” members of the *P. aeruginosa* population which cheat, ceasing the production of siderophores whilst still up taking those bound to iron which they encounter (O'Brien *et al.*, 2017; Stilwell, Lowe

and Buckling, 2018). It could also therefore be the case that with the increased energetic expense of T3SS and other virulence system expression in the always on mutants, the marginal gains in iron availability gained in the iron rich medium by siderophore synthesis is outweighed by the increased anabolic burden. This indirect explanation is entirely compatible, even cooperative with, the increased availability of Fe²⁺ due to raised pyocyanin levels.

It is also possible that there is a more direct mechanism, which could also be implied by the upregulation of an uncharacterised bacterioferritin like protein which likely has a role in the cytoplasmic storage of iron. A further interesting observation on the mutant's siderophores is that the Δ *exsD* mutants have a larger downregulated siderophore synthesis cluster than Δ *exsE* mutants, and only the Δ *exsD* mutants have decreased siderophore activity compared to WT. There is however no apparent dysregulation apparent for siderophore synthesis proteins between the two "always on" mutants in the direct comparison of their proteomes, which indicates that, much like with the T3SS western blots earlier in the investigation, the Δ *exsD* phenotype is stronger than the Δ *exsE* phenotype. It could be speculated that a threshold is not met in the Δ *exsE* mutants for the measurable decrease in siderophore production, despite the downregulation of many involved proteins.

The starkest difference between the two always on mutants is the presence of a downregulated cluster of denitrification proteins in the Δ *exsD* mutants, which is entirely absent in the STRING map for the Δ *exsE* mutants. There was not, however, a corresponding difference between them in the direct comparison of the two "always on" mutant proteomes. Further examination of denitrification within the mutants failed to clarify this quandary, with a luciferase reporter of *nirS* transcription offering data supportive of induced ExsA activity reducing *nirS* transcription which could not be validated due to the control reporters showing unexpected changes in the same conditions. A nitrite assay was therefore employed to quantify any denitrification that took place. Contrary to expectations from the proteomic data nitrite was reduced in the always off mutants, and unchanged between always on mutants and WT. Together the data indicates that the ExsA cascade has a role in the regulation of denitrification in which ExsA limits denitrification, though how and why remain to be elucidated.

Given that both "always off" mutants have the same measurable decrease in nitrite concentration it appears that once again the proteomic difference between the Δ *exsD* and Δ *exsE* strains is a matter of phenotypic strength, with the impact of Δ *exsE* on the expression of denitrification proteins falling below the required level for detection. An alternative explanation for the distinctions between the two always on mutants is an independent DNA binding regulatory role for ExsD. This has previously been suggested from the appearance of

the crystal structure (Bernhards *et al.*, 2009). However, no data indicating that ExsD can bind to DNA has been published, and the same group subsequently claimed that the homotrimerization of ExsD is a regulatory feature of the ExsA cascade which prevents ExsD-ExsA interaction in some circumstances in a paper which makes no reference to putative ExsD-DNA interaction (Bernhards *et al.*, 2013). In the same paper they demonstrate that this regulatory effect is eliminated at 37°C, rendering it irrelevant to the present study and many infection scenarios. There is therefore no evidence of an independent regulatory role for ExsD and, given the lack of clear distinction between the two “always on” mutants, the present data does not support this theory either.

The data to hand therefore indicates that the ExsA cascade is watertight, with a stronger phenotypic impact caused by mutations lower down the cascade, but ultimately one signal transmitted from ExsE export, amplified through the partner switching cascade, and finally causing the initiation of ExsA transcriptional activity.

3.7 Conclusions

Despite promising preliminary results, the distinction between the Δ exsA mutant and WT proteomes was minimal. However, a plethora of data was generated in the Δ exsD mutant, and the mutant's proteome contrasted more strongly with Δ exsA than WT. A raft of novel additions to the ExsA regulon of virulence factors have been suggested by this data, notably pyocyanin, HCN, a subset of T6SS proteins, and the NRPS operon PA1221-PA1211. Complementation of the mutants with a plasmid borne copy of the gene demonstrating restoration of a WT phenotype would provide stronger evidence, however time constraints prevented this in the present study. This confirms the utility of ExsA inhibitors for an anti-virulence strategy, providing a both a specific T3SS inhibition and a broader inhibition of virulence factors, with no indication that it could push the pathogen into a chronic virulence lifestyle or have other deleterious consequences.

4. Bioinformatic and structural investigation of approaches for ExsA inhibition

4.1 Aims and approaches

With the validity of ExsA as a target for an anti-virulence therapeutic strategy for acute *P. aeruginosa* firmly established both in the published literature as described in introduction section 1.5 and the prior results chapter, further investigation was undertaken towards exploring various approaches to inhibit this protein. In this chapter ExsA has been subjected to bioinformatic analyses focused on identification and investigation of homologous proteins, *in silico* screening and *in vivo* testing of chosen small molecules as potential inhibitors of ExsA. For *in silico* screening the only published ExsA structure was used (Protein data base code (PDB): 4ZUA), which contains only the NTD, this truncated structure was deemed acceptable as it contains the ligand binding pocket of interest.

4.2 Methods

4.2.1 Comparison of ligand-binding pockets between ToxT and ExsA

The crystal structure of the ToxT from *V. cholerae* (PDB: 5SUX) excluding the bound ligand was aligned with the ExsA monomer (PDB: 4ZUA) (Shrestha et al., 2015) utilising the structural protein alignment function of Molegro molecule viewer 2.5 (MMV). Both the structures were then subjected to FTSite server (give the url) (Ngan et al., 2012) to identify potential small molecule ligand binding pockets. These latter were then viewed in PyMOL for comparative visual analysis

4.2.2 Blind docking of ToxT ligands to ExsA with Autodock 4

The following X-ray crystal structures of ToxT were obtained from the RCSB protein data bank: 3GBG (Lowden et al., 2010), 4MLO (Li et al., 2016), 5SUX, and 5SUW. The respective ligands namely *cis*-palmitoleate (TL1), virstatin, (E)-4-(8-methylnaphthalen-1-yl) but-3-enoic acid (TL2), and 3-(8-Methyl-1,2,3,4-tetrahydronaphthalen-1-yl) propanoic acid (TL3), were then extracted from the cognate ToxT structures. These ligands were then subjected to blind docking, with 100 replicates, to the published ExsA NTD structure (PDB: 4ZUA) with Autodock 4.2 (Morris et al., 2009) within the PyRx platform (Dallakyan and Olson, 2015). Docking poses within the known ligand binding pocket was taken as a positive result. The same process was undertaken with a ToxT structure (PDB: 5SUX) for validation purposes. The top 100 poses of each ligand ranked by the default predicted binding energy scores (kcal/mol) against the known pockets were visualised and assessed in PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC).

4.2.3 Phylogeny generation

All reviewed member proteins of the same family of helix-turn-helix (HTH) AraC/XylS-type proteins as ExsA were selected from UniProt (PROSITE entry PS01124), and their amino acid sequences were downloaded. A phylogeny was generated from this dataset with the Maximum Likelihood method and JTT matrix-based model, in MEGA X with 100 bootstraps (Kumar *et al.*, 2018). The bootstrap consensus tree that is presented has been inferred from 500 replicates. Where a branch corresponded to less than 50% of replicates the divergence point was removed and branches merged. The percentage of replicates in which the associated taxa clustered together are shown next to the branches, thus a higher number indicates a more certain division.

4.2.4 SNP analysis

The ortholog group of *Pseudomonas aeruginosa* strain PAO1's *exsA* gene (Pseudomonas Ortholog Group POG003676, available: <https://pseudomonas.com/orthologs/list?id=106196>) was taken from pseudomonas.com and aligned to the reference PAO1 *exsA* sequence with a standard protein BLAST [<https://blast.ncbi.nlm.nih.gov/Blast.cgi>] (National Library of Medicine (US) and National Center for Biotechnology Information, 1988). Out of the 320 sequences in the group, 33 sequences which had 100% coverage and <100% but >80% identity were chosen to explore potential non-synonymous single nucleotide polymorphisms (SNPs). The selected alignments were inspected for SNPs which were subsequently visualised using PyMOL, with variants inserted using the build residue function and manually positioned for visibility.

Stability of ExsA with SNP variations was calculated with the I-Mutant 2.0 webservice [http://gpcr2.biocomp.unibo.it/~emidio/I-Mutant2.0/I-Mutant2.0_Details.html] (Capriotti, Fariselli and Casadio, 2005), with the published structure of the N terminus of ExsA (PDB: 4ZUA) and the sequence alone used for residues not covered in this available structure. The parameters were set to pH 7 and a temperature of 37°C.

Finally, a prediction as to whether a SNP affects the function of a protein was made using the SIFT online tool [<https://bio.tools/sift>] (Ng and Henikoff, 2001). The ExsA amino acid sequence for PAO1 (GenBank ascension AAG05102) was used as a reference and each known SNP was used as a query. Unfortunately, all SNPs predicted to affect function suffered from a lack of diversity in the reference sequences, causing the predictions to be of low confidence. Varied search databases and adjustments to parameters were made, however none resolved the issue indicating that it is a result of ExsA's sequence. The database used for the data presented was UniProt-SwissProt 2010_09, and the Median conservation of sequences was

set to 3.00. Sequences more than 90% similar to the query were removed.

4.2.5 Development of methods to copurify ExsA and ExsD

Except for the usage of n-dodecyl- β -D-maltoside (DDM) which was added at a concentration of 0.1% W/V to both the elution and dialysis buffers when indicated, copurification methods followed those outlined in the general methods.

Crude solubility tests were undertaken by diluting 15 mg/ml concentrated copurified ExsA-ExsD to a calculated concentration of 0.15 mg/ml in 1 ml of either pH 7.4 Tris-HCl (50 mM) or pH 6 MES (0.5 M), vortexed and incubated at 4°C overnight. The solutions were then spun for ten seconds in a benchtop microfuge and visually inspected for precipitation, and approximate protein concentration was subsequently calculated from A_{280} on a nanodrop spectrometer.

4.2.6 Crystallisation screening

Screens were conducted with sitting-drop vapour diffusion on screening plates: Classics suite, LMB screen, PEG suite, and Wizard I&II, from the X-ray Crystallography Facility, University of Cambridge. The copurified ExsA-ExsD was added to each plate at a concentration of 10 mg/ml and each plate was tested with two droplets at different buffer concentrations, with droplet two receiving half the concentration of buffer. Each droplet received 0.2 μ L of protein and had a final volume of 150 μ L. Plates were incubated for 34 days with frequent observation.

4.3 ToxT inhibitors do not inhibit ExsA

As mentioned in Chapter 1, ToxT is a homolog of ExsA (with 26.32 % sequence identity) with a comparably critical role in regulating virulence in *V. cholerae*. There are several known ToxT inhibitors, and multiple published full-length structures (both apo and inhibitor bound). Consequently, ToxT was used as a starting point in the present investigation. Two computational approaches were undertaken to investigate the possible utilisation of ToxT ligands for the inhibition of ExsA. The first was a comparison of the ligand binding pockets between ToxT and ExsA as identified by FTsite. This led to the finding that the homologous pocket in ExsA is of a different shape and lined by different amino acids, compared to that of ToxT (Figure 4.1 A). Specifically, the putative ExsA ligand binding pocket contains far more hydrophilic/polar residues compared to ToxT, which would appear to preclude similar ligands binding to both proteins (Figure 4.1 B).

This observation was supported by docking experiments, which found that known ToxT ligands could be readily docked to ToxT structures with a perfect reproducibility (100 poses of 100 replicates), whereas the same methodology yielded no docking of those ligands to the putative ligand binding pocket of ExsA.

Finally, to validate this prediction a commercially available ToxT inhibitor, virstatin (Shakhnovich *et al.*, 2007), was purchased and tested for ExsA inhibition in a *PpcrV*-Lux reporter strain of *P. aeruginosa*. No reduction in luminescence was observed at 200 mM (Figure 4.1 C). The only time point at which the virstatin treated sample differed from the control in a statistically significant manner ($P < 0.05$) was at 4 hours, when it had a higher value than the control.

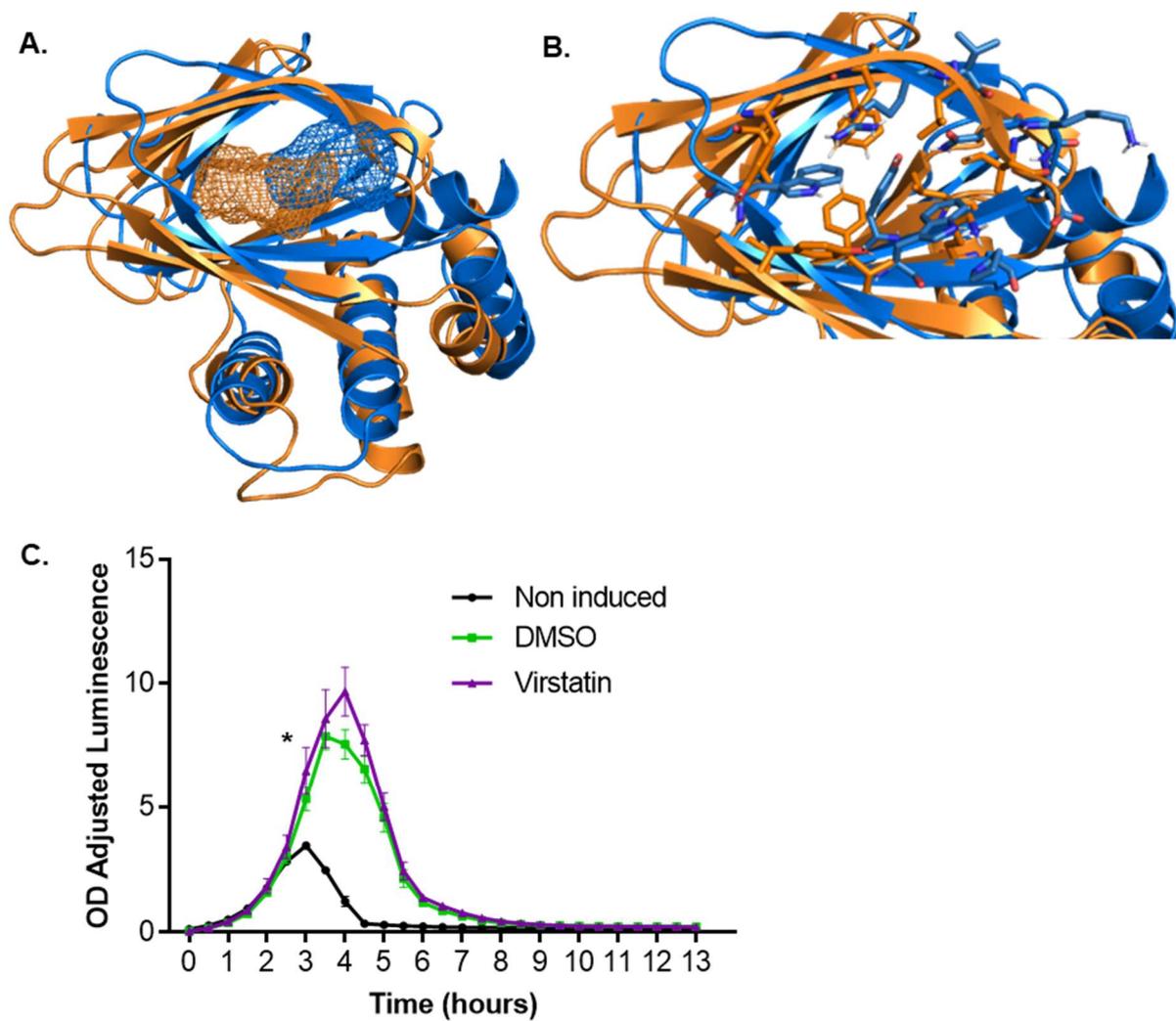


Figure 4.1 Computational and *in vivo* analysis of virstatin and ExsA.

A&B. Overlaid cartoon structures of ExsA (blue, PDB: 4ZUA) and ToxT (orange, PDB:5SUX) with a mesh representation of FtSite-predicted pockets and stick representation of surrounding residues in matching colour. The CTD of the ToxT structure is not shown for clarity.

C. The results of the luciferase reporter of ExsA activity shown by growth adjusted luminescence from *PpcrV-Lux* for virstatin at 200 nM. Each point represents the mean of three biological replicates, with error bars indicating standard error of the mean. * Indicates a statistically significant difference between DMSO and virstatin ($p < 0.05$).

4.4 The close phylogenetic relations of ExsA offer no probable inhibitors

To find more closely related homologs to ExsA, and therefore to find potential inhibitors of ExsA, a phylogeny of the AraC family was constructed. The rationale for this process was the assumption that two proteins sharing the same small molecule binding pockets in terms of shape and electrochemical properties, are likely to accommodate similar ligands in their pockets and thus their functions are likely to be affected by each other's ligands. ExsA was placed on a basally diverging branch of the tree, with four other proteins: VirF, LcrF, NimR, and ChbR (Figure 4.2, a full phylogeny is available in Appendix 2). A literature search was conducted to identify potential lead compounds known to bind to these AraC member proteins. The only member of the cluster with a published ligand which appears to bind to the ligand binding pocket homologous to that of interest in ExsA is ChbR, an *E. coli* protein which is thought to bind chitobiose and promote its metabolism (Plumbridge and Pellegrini, 2004). ChbR's proposed ligands chitobiose and chitobiose phosphate were therefore investigated with docking experiments, with the superficially similar cellobiose used as a negative control. All three docked poorly, with chitobiose and chitobiose phosphate showing no superiority compared to cellobiose (Table 4.1).

Table 4.1 Results of Gold and Vina docking of chitin metabolites and cellobiose to the ExsA NTD.

Compound	ChemPLP Score	Vina Affinity (kcal/mol)
Chitobiose	46.4	-5.6
Chitobiose phosphate	43.77	-6.1
Cellobiose	41.65	-5.6

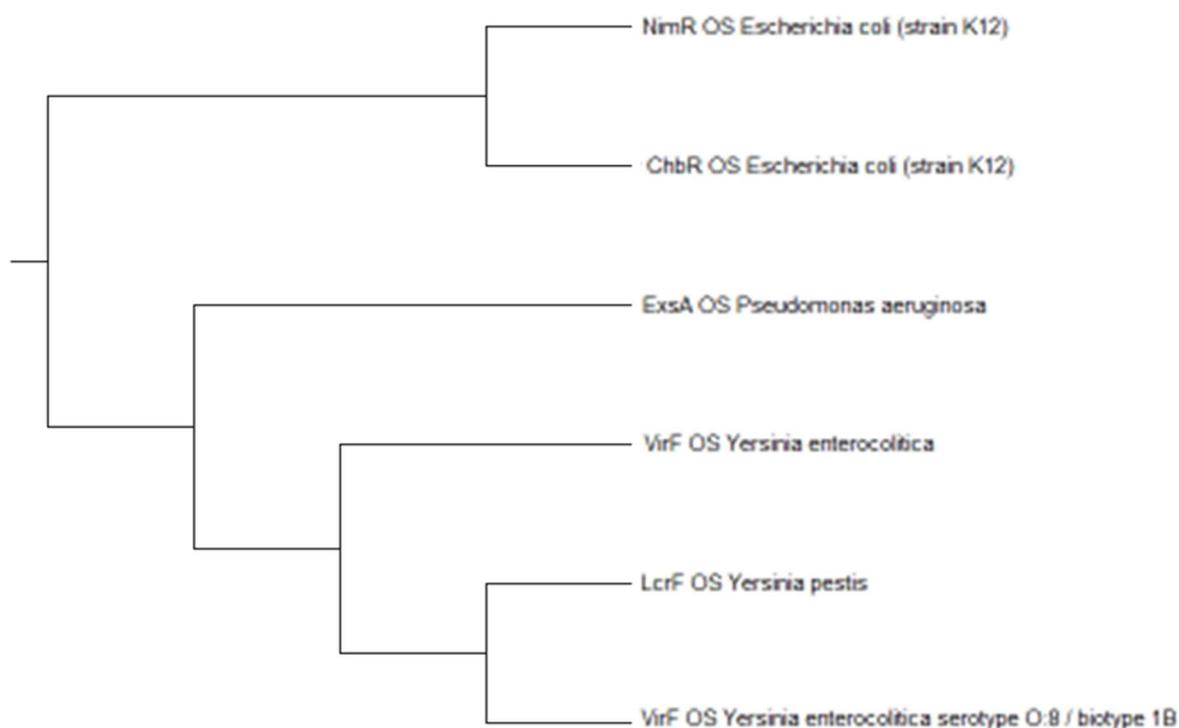


Figure 4.2 The phylogenetic branch of the AraC family containing ExsA.

A basally diverging branch of the AraC family, as per InterPro categorisation, in a phylogeny constructed by MEGA X, on which ExsA was placed. Also present on the subtree are VirF, LcrF, ChbR, and NimR.

4.3 Analysis of single nucleotide polymorphisms (SNPs) identify inconsistent regions across ExsA

Due to the lack of existing inhibitors for the NTD target sites in ExsA it was clear a *de novo* method would be required. To explore the genetic variation around the target sites, analysis of all known non-synonymous single nucleotide polymorphisms (SNPs) in *exsA* was undertaken. Analysis of every reported sequence of *exsA* revealed that the majority (83%) had no SNPs which caused a change in protein sequence. Those SNPs which were identified are listed in Table 4.2. None of the identified SNPs are within or proximal to the putative ligand binding pocket, however one (H131N) is observed on a dimerization helix (Figure 4.3). In the subsequent *in silico* analysis of the SNPs H131N was predicted to have a negative change in change in Gibbs free energy ($\Delta\Delta G$), and to be a tolerated mutation. Other SNPs also appeared to have a stabilising effect on $\Delta\Delta G$, most notably I78T. Several SNPs were predicted to have destabilising effects, however these were generally small, for instances S6F with a $\Delta\Delta G$ of 0.69. Whilst some SNPs were predicted to affect protein function by SIFT, which compares

the SNP to a list of homologous reference sequences and examines variation for the residues in question, the strength of these predictions was insufficient to warrant credibility.

Table 4.2 SNPs detected in ExsA

SNP	Number of Isolates	Structural Location	$\Delta\Delta G$ (I-Mutant 2.0)	SIFT prediction*
S6F	1	NT loop	0.69	Affect
Q11R	1	NT loop	0.21	Tolerated
L57F	1	Beta barrel interior	-1.34	Tolerated
I78T	1	Beta barrel exterior	-3.48	Affect
R101C	1	NTD loop	-0.88	Affect
D103E	7	NTD loop	0.30	Tolerated
E104G	1	NTD loop	-1.22	Tolerated
A114V	3	NTD loop	-0.91	Tolerated
H131N	1	Dimerization helices	-0.19	Tolerated
P156A	12	Small NTD helix	-1.45	Tolerated
T200A	1	CTD	-0.25	Tolerated
R214C	1	CTD	-1.29	Affect
S218L	1	CTD	0.62	Tolerated
T262A	1	CTD	-1.53	Affect

*due to a lack of diversity in reference sequences confidence in SIFT predictions of “Affect [on protein function]” is low

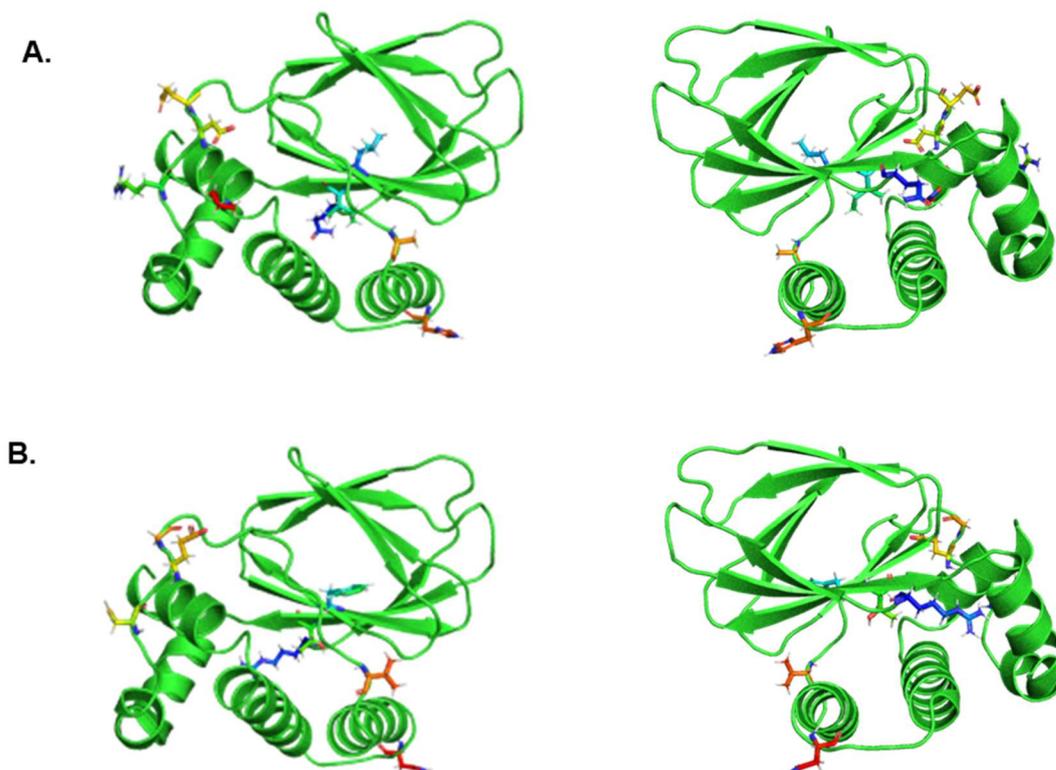


Figure 4.3 Residues at which SNPs have been identified.

All SNPs detected for ExsA which are present in the 4ZUA structure are displayed in stick form, with matching residues in matching colour. **A** and **B** show the PAO1 variant of the SNP and the ortholog variant respectively, both from two perspectives.

4.5 ExsD and ExsA and co-purifiable as a complex but evaded co-crystallisation under any tested condition

ExsA has defied historic experimental attempts at deriving a full-length structure (Shrestha *et al.*, 2015). Whilst the crystal structure of the N terminus of ExsA [PDB: 4ZUA] has been published, a full-length structure would be preferable for virtual screening. Inherent structural flexibility towards the C terminus of ExsA could be a major reason for the crystallisation failure to date. As outline in Chapter 1 (Section 1.5), ExsA is bound by the antiactivator ExsD (Brutinel, Vakulskas and Yahr, 2010). It was therefore conceived of to attempt co-crystallisation of ExsA and ExsD, theorising that the presence of ExsD could stabilise the flexible regions of ExsA and aid crystallisation.

4.5.1 ExsA-ExsD co-purification requires different conditions to ExsA alone

A novel co-expression vector was created (pET-19m-ExsA-ExsD), and purification was undertaken using the established protocol for ExsA within the lab. After purification the protein

was concentrated via ultra-centrifugation; however, during concentration the product precipitated, leaving a negligible yield of soluble protein. This was observed repeatedly, and no protein at a useful concentration could be obtained due to this solubility issue. A detergent, DDM, was subsequently added to the eluted protein prior to concentrating the protein and allowed a high concentration to be obtained. Denaturing gel electrophoresis and Coomassie staining confirmed the presence of a band at the correct size to be both ExsA and ExsD in their monomeric forms, which are too similar in size to be distinguished by this method (Figure 4.4).

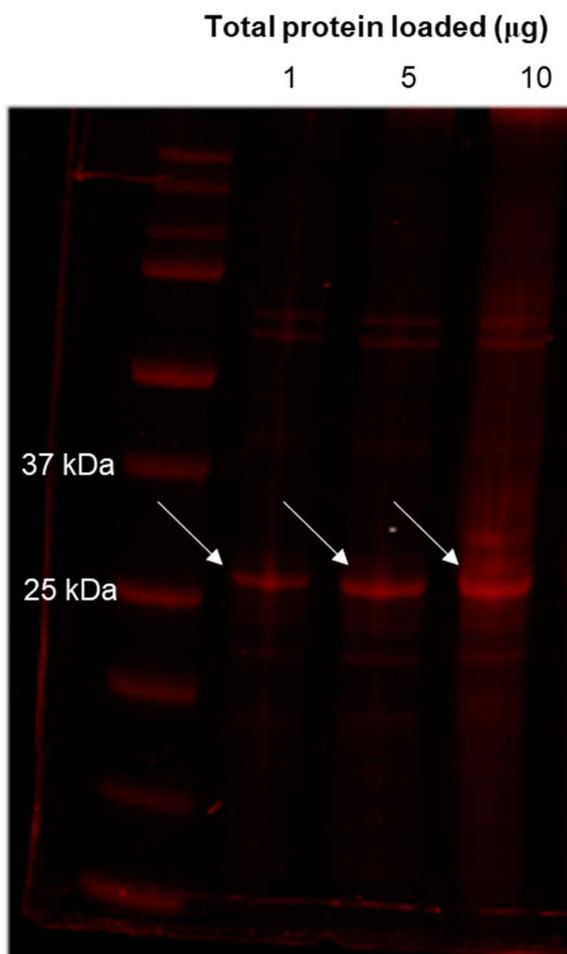


Figure 4.4 Co-purified ExsA and ExsD Coomassie stain.

A Coomassie stained polyacrylamide gel of purified ExsA and ExsD after concentration with DDM. Arrows indicate bands at the correct size to be both ExsA and ExsD, other bands are contaminants.

4.5.2 ExsA-ExsD has extremely low salt tolerance

Due to the expense of DDM, it was deemed impractical to use the quantities necessary to prepare enough co-purified protein for structural investigation. A buffer screen was therefore undertaken. ExsA's exposed hydrophobic regions caused a high background signal which masked the melting point in many samples (Figure 4.5 A). Enough data was obtained to indicate that the purified protein was at its most stable form when the salt was at low concentration or altogether absent (Figure 4.5 B). Tris-HCl at pH 6 without NaCl or glycerol was the only result to differ from the main cluster of melting points around 50°C with a melting temperature of 64°C (Figure 4.5 B). Due to Tris-HCl not buffering at pH 6, a simple experiment to determine if the improved stability was a result of the pH or simply due to a NaCl Tris-HCl buffer was conceived. It was observed that after overnight incubation ExsA-ExsD copurification product was still entirely soluble in Tris-HCl pH 7.4, whilst the protein suffered insolubility and precipitation when in MES buffer at pH 6 (Figure 4.6). A pH 7.4 Tris buffer without NaCl was used for further purifications, with no further solubility issues. It was also found that the addition of 5% of glycerol to the purification as a cryo-preservant did not noticeably affect the stability of the proteins.

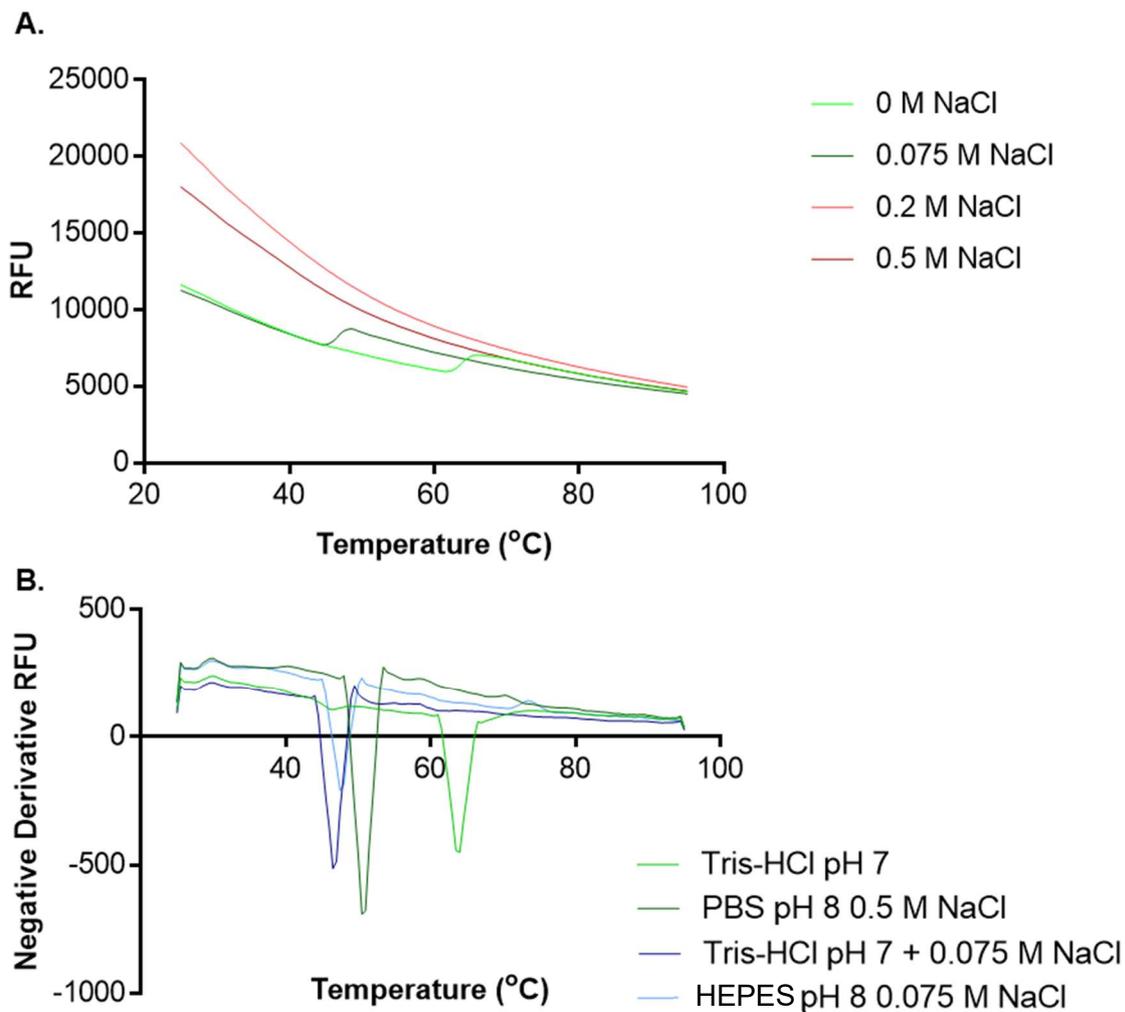


Figure 4.5 ExsA and ExsD copurification buffer screen by thermal shift.

A. The relative fluorescence units detected for All four Tris-HCl pH 6, 0% glycerol, buffers tested for ExsA and ExsD by thermal shift. Two of the samples failed to give a melting point (displayed as red lines), and all four samples are subject to a decaying background signal, which is illustrative of all tested samples. All data is indicative of a single replicate.

B. The negative derivative of selected buffers relative fluorescence units during a thermal shift experiment. The downward peak is indicative of melting temperature. All data is indicative of a single replicate.

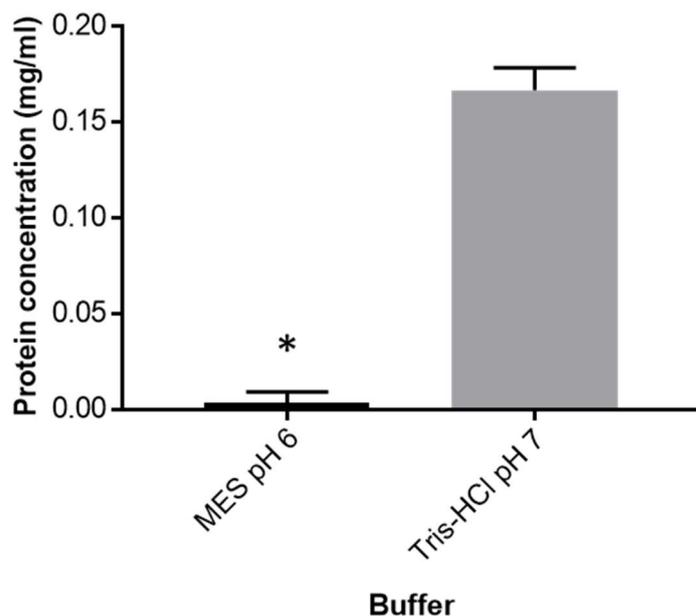


Figure 4.6 Crude estimation of ExsA and ExsD solubility.

The concentration of protein detected via OD_{280} after incubation at 4 °C for 17 hours in the indicated buffer and gentle centrifugation. Samples were diluted to a calculated concentration of 0.15 mg/ml prior to incubation. * indicates the visual observation of a pellet in the samples. Bars are representative of the mean of technical triplicate, error bars indicate standard deviation.

4.5.3 Size exclusion chromatography was necessary for purity ExsA-ExsD, which were confirmed to co-express

The use of size exclusion chromatography dramatically improved purity, with a clear elution peak for the desired product as well as smaller contaminant peaks and the separation of soluble aggregates (Figure 4.7 A). The complex was obtained at high purity when viewed on SDS-PAGE (Figure 4.7 B).

With yield and purity of the co-expressed purification now obtained it seemed highly probable that a heterodimer was being expressed due to the difference in NaCl concentration tolerated, however confirmation of dimerization was sought *via* analytical size exclusion chromatography. The salt free buffer caused poor running of the protein standards in comparison to a buffer with 100 mM NaCl (Figure 4.8 A). The sample elution peak was at a lower elution volume, and thus greater mass, than the 44 kDa standard peak; thereby indicating the sample was a dimer of both proteins with a mass of 64 kDa rather than a monomer with a mass of 32 kDa (Figure 4.8 B). Due to the presence of a shoulder on the 44

kDa standard peak, which was consistent across three independent replicates, instead of a clear 158 kDa standard peak, more precise estimation of the samples mass was not possible. Crystallisation was attempted in a broad range of conditions (four 96 well screening plates) however after 30 days of observation no crystal formation was observed.

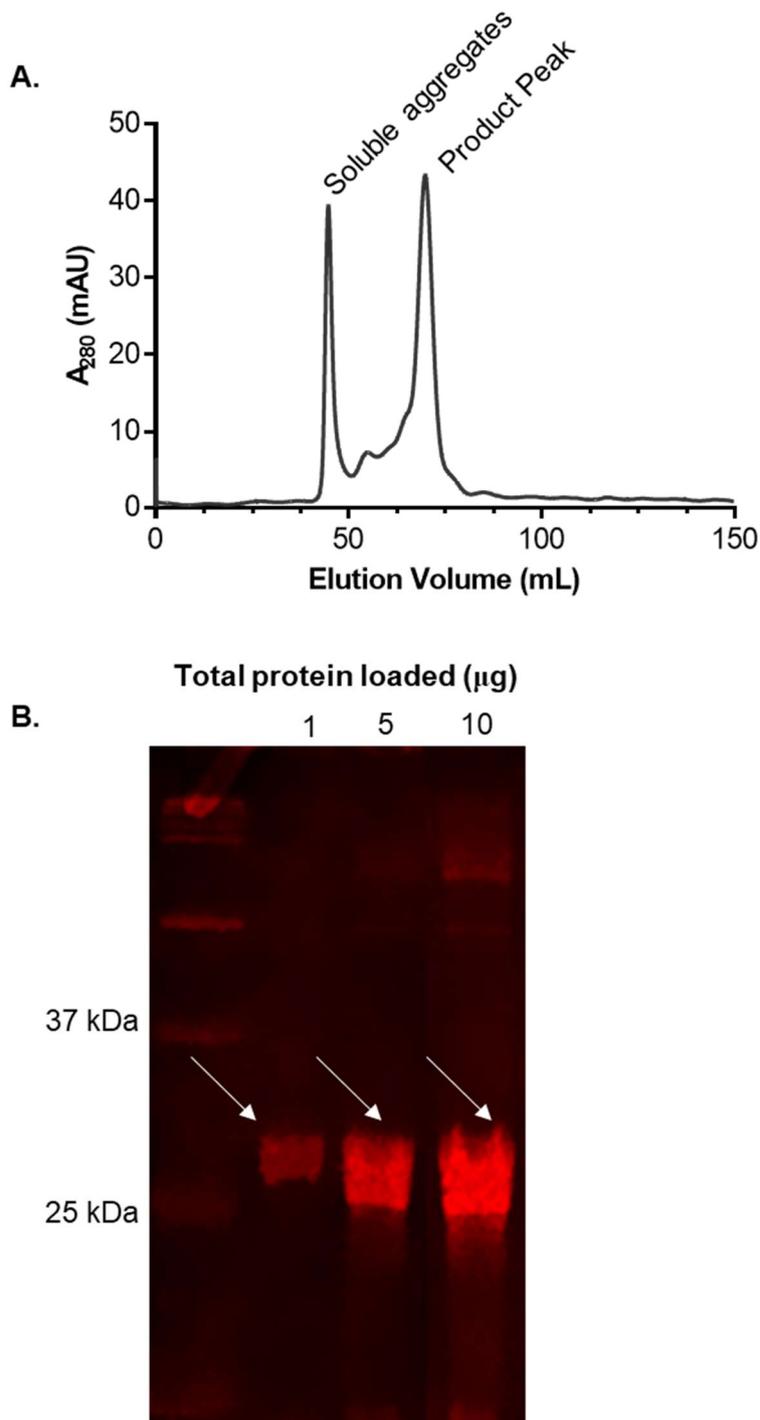


Figure 4.7 Size exclusion chromatography of ExsA-ExsD copurification.

ExsA-ExsD copurification via histrap subsequently concentrated by ultracentrifugation and subject to size exclusion chromatography. **A.** depicts the chromatograph of size exclusion chromatography, with a higher A_{280} reading indicating a higher concentration of eluted protein. Labels indicate the contents of the most significant peaks **B.** Is a coomassie stained polyacrylamide gel of the product eluted at the corresponding volume to the peak marked “product” in A. Arrows indicated bands of the correct size to be ExsA and ExsD (molecular mass 31.6 kDa and 31.4 kDa respectively).

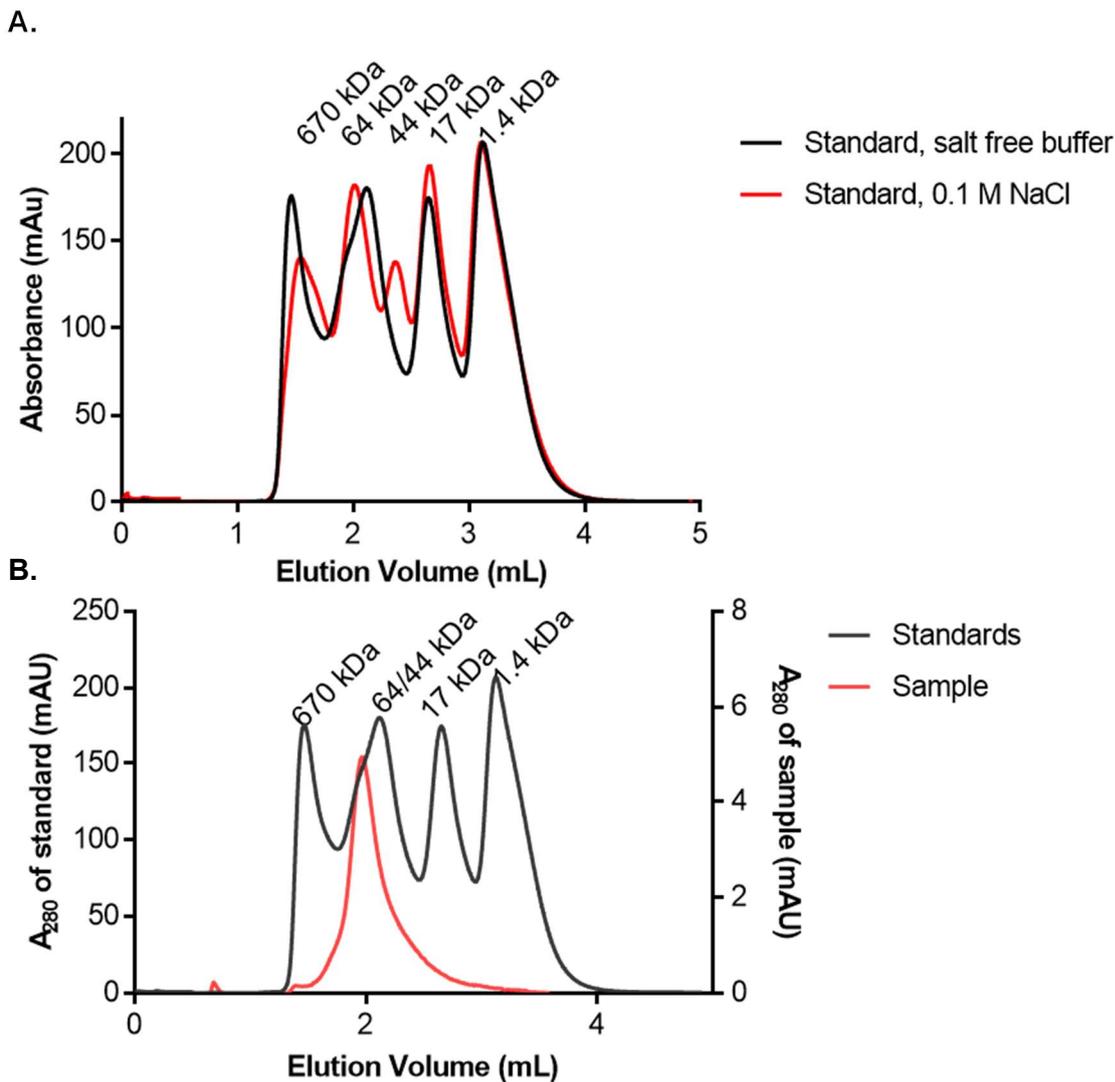


Figure 4.8 Analytical size exclusion chromatography of ExsA-ExsD copurification.

A. A size exclusion chromatograph with the protein standards in both the salt free ExsA-ExsD copurification buffer (black) and another buffer with 0.1 M NaCl (red) overlaid. Comparison illustrates the lack of resolution between the 64 and 44 kDa standards in salt free buffer. Labels indicate the size of the protein standards corresponding to each 0.1 M NaCl buffer peak.

B. Analytical size exclusion chromatograph of ExsA-ExsD copurification after preparatory size exclusion chromatography (red) overlaid with the protein standards in the same buffer (black). Labels indicate the size of the protein standard corresponding to each peak.

4.6 Discussion

FTSite utilises an energy-based method somewhat similar to a blind docking approach. To this end 16 small molecule probes are docked to the protein query, and the free energy of each interaction is calculated (Ngan *et al.*, 2012). Given that this calculation, which is made without any information on homologs, detected a ligand binding pocket in the same location as that found across much of the AraC family it seems highly probable that a functional ligand binding pocket is present in ExsA. Investigation of the comparative pocket on ToxT with FTSite, and *in silico* investigation of the ToxT inhibitor virstatin, revealed marked differences between the cognate ligand binding pockets on the two proteins.

The luciferase assay validation of the computational prediction confirmed that the differences observed *in silico* have a clear impact *in vivo*. This surely did not provide evidence against the theory that targeting the homologous pocket on ExsA could be a potential means for pharmacological inhibition of this protein. However, it was clear that ligands with novel chemical scaffolds that differ from the known ToxT inhibitors are needed to potentially engage with the target pocket on ExsA. Whilst ToxT pocket seems to accommodate its inhibitors such as virstatin and fatty acids largely through hydrophobic interactions (Shakhnovich *et al.*, 2007; Lowden *et al.*, 2010), the cognate pocket on ExsA is more polar, marked with residues that can bind inhibitors through network of hydrogen bonds/or electrostatic interactions and as such, the ExsA pocket appears to be more “druggable” from a conventional point of view (Kozakov *et al.*, 2015).

Phylogenetic analyses led to the grouping of ExsA with several temperature regulated transcription factors, perhaps indicating a similar absence of a native ligand. This is compatible with the current lack of a known ExsA NTD ligand with proven site of binding and the presumed sole dependence on the partner switching cascade known to regulate ExsA’s function – however an absence of evidence is not evidence of an absence. The only member of the group to have a ligand was ChbR which is a transcriptional promoter for the regulation of chitobiose metabolism in *E. coli*. Whilst there is no direct evidence, it seems probable that ChbR binds to the chitobiose (Plumbridge and Pellegrini, 2004). Regulation of ExsA by a metabolite produced in the breakdown of chitin, a major component of the fungal cell wall, could be relevant environmentally and in polymicrobial infections, particularly as there is indirect evidence of *P. aeruginosa* utilising its T3SS to kill fungi (Manavathu, Vager and Vazquez, 2014; Nazir *et al.*, 2017) - a phenomenon better characterised in *Salmonella enterica* serovar Typhimurium (Kim and Mylonakis, 2011). It is also documented that *P. aeruginosa* can metabolise chitin and produce chitobiose (Thompson *et al.*, 2001; Kumar *et al.*, 2017). Therefore, it is conceivable that the presence of chitobiose affects ExsA function to arm *P.*

aeruginosa via T3SS expression in the presence of fungi ChbR's natural ligands. Whilst the lack of structural or binding data prevented an investigation as robust as that of the ToxT ligands, there is no evidence obtained in the present study to support the binding of chitobiose or any similar molecule to ExsA.

VirF and LcrF both have published inhibitors which target the DNA binding interface, and as such have a broad spectrum of activity within the AraC family (Garrity-Ryan *et al.*, 2010; Koppolu *et al.*, 2013). As discussed in the introduction (1.5 The function of the ExsA cascade), an alternative approach is being sought in the present study and DNA binding interface ligands were therefore excluded. One VirF inhibitor was of interest because it has been reported to operate by an alternative mechanism. Unfortunately, the mechanism itself is yet to be elucidated, and no structural binding data is available, rendering repurposing to ExsA inconvenient at the present time (Emanuele and Garcia, 2015).

The final protein within the cluster, NimR, is a poorly characterised transcriptional regulator with no known binding molecules, and thus provided no leads. The phylogenetically related proteins therefore also failed to produce a lead compound for future development into an ExsA inhibitor. Whilst intriguing, the potential function of T3SS against fungi is outside the scope of the present investigation.

With the avenues for utilising existing lead compounds against the NTD of ExsA explored fruitlessly, it is clear a *de novo* approach was needed. There appears to be two potential target sites on the NTD of ExsA - the putative ligand binding pocket that is analogous to the ToxT and few other AraC members, and the dimerization interface. To ensure the validity of these target sites in a broad spectrum of *P. aeruginosa* strains, analysis of all reported code changing SNPs was undertaken. The SNP analysis didn't use a quantitative sample of medical strains due to the lack of an available data set and the impracticality of generating one. An "all available data" approach was instead utilised to have detected any common variants which would invalidate the target site in a common infectious strain. This indicated that the PAO1 variant of *P. aeruginosa* was a suitable choice as a strain to study ExsA and its function. This strain possesses the most common reported variant of ExsA, although this is no doubt biased by PAO1's status as the model strain of *P. aeruginosa*.

None of the SNPs reported in ExsA at the time of writing seemed to affect the putative ligand binding pocket homologous to those present in other AraC family proteins including ToxT. This indicates that it is potentially a viable binding site for developing small molecule drugs with much lesser chance of rapid emergence of resistant strains. The other target site however,

the dimerization interface, does have a SNP reported, H131N. This SNP was identified in the strain LES431, a clinical strain from a chronically infected patients' lung (Salunkhe *et al.*, 2005). LES strains and other strains identified in chronic infections often lasting years are known to inactivate the T3SS via mutation, sometimes within ExsA (Smith *et al.*, 2006; Jeukens *et al.*, 2014). It is therefore possible that there are other unreported dimerization interface SNPs in unsequenced strains of *P. aeruginosa*, inactivating ExsA in long term chronic infections. H131N therefore does not present an issue with targeting the dimerization interface, as any radical mutation of the interface will likely impair dimerization, which is after all the purpose of the inhibitor. It was therefore apparent that no SNPs provide could pose serious concern for either of the targets.

With the target sites chosen and virtual screening as the preliminary method in mind, the next step was to obtain the best possible structure for use. To this end an attempt was made to obtain a crystal structure of ExsA in an inactive conformation but potentially stabilised condition, bound to ExsD. Unfortunately, the crystallisation attempts failed. Whether this was because ExsD did not alter the conformation of ExsA sufficiently to stabilise the flexible region, because the heterodimer did not pack into crystals, or because the right condition was not trialled remains unknown. It is also possible that the uncleaved His tag on the ExsA NTD prevented crystallization. The His tag was not subject to TEV protease cleavage due to buffer conditions in which the complex was stable being incompatible with TEV activity; however there are numerous examples of successful crystallisation with a His tag in place (Kon *et al.*, 2012; Wang *et al.*, 2021), and even instances in which a retained His tag has aided crystallization (Smits *et al.*, 2008).

4.7 Conclusion

It has been demonstrated in this chapter that there appeared to be no viable compounds from the known small molecule binders of the N terminii of ToxT and other AraC members that could be repurposed against PAO1, excepting those with a broad spectrum of activity against AraC family DNA binding domains. It is also evident that the proposed target sites on the NTD of ExsA exhibit no known genetic variation which would invalidate them against any sequenced clinical strain or facilitate the rapid acquisition of resistance. A full-length structure of ExsA remains elusive, and therefore the discovery of a novel inhibitor will proceed using the published NTD structure (PDB: 4ZUA) of ExsA as the receptor for virtual screening to find compounds with a potential to inhibit ExsA.

5. Preliminary identification of ExsA inhibitors targeting the NTD

5.1 Aims and approaches

Given the absence of any known inhibitors of ExsA targeting its NTD, a virtual screening (VS) approach was employed to generate a list of small, lead-like molecules that could potentially bind to the NTD of ExsA. The compounds selected from this screen were subsequently subjected to an *in vivo* test of ExsA inhibition, and attempts were made to further characterise the lead compounds using biophysical and functional analysis.

5.2 Methods

5.2.1 Identification of potential inhibitors of the ExsA ligand pocket

Potential ligands for the conserved AraC ligand binding pocket were identified using IDock (Li, Leung and Wong, 2012). Search criteria were restricted to molecules predicted to be 'drug like' and soluble. Key criteria included: Net charge of 0, Partition coefficient -4 to 4, and molecular weight 100 to 500 gmol⁻¹. A total of 2,400,000 compounds were screened. The top 500 hits from iDock-based VS were then selected for further processing. Autodock Vina was used to dock all 500 ligands to the crystal structure of the NTD of ExsA (PDB: 4ZUA) (Morris *et al.*, 2009; Trott and Olson, 2009), using an unbiased or 'blind docking' approach with three independent repeats (Augustin *et al.*, 2020). The mode 0 results (i.e. the best scored result from each docking iteration) were extracted, and the ligands that docked to the desired ligand binding pocket (Figure 5.1) were identified. Autodock Vina was then performed for these molecules, with 100 replicates. The top scores of these 100 independent docking runs per molecule were then used, and a percentage score of poses which were within the ligand binding pocket was calculated for each ligand.

Parallel to AutoDock Vina-based docking, all 500 hits from the initial iDock-based VS were also docked to the aforementioned NTD structure of ExsA, Gold version 5.3.0 (CCDC, Cambridge, UK) and re-ranking of these hits were done using three different scoring methods available in GOLD namely the ChemPLP, GoldScore and ChemScore (Jones *et al.*, 1997). Ensemble scores incorporating all three algorithms ('Zscores') were calculated for each molecule with the formula: $Zscore = (score - mean) / standard\ deviation$. The mean Zscores were then calculated and used to order the results. This workflow is summarised in Figure 5.2 A.

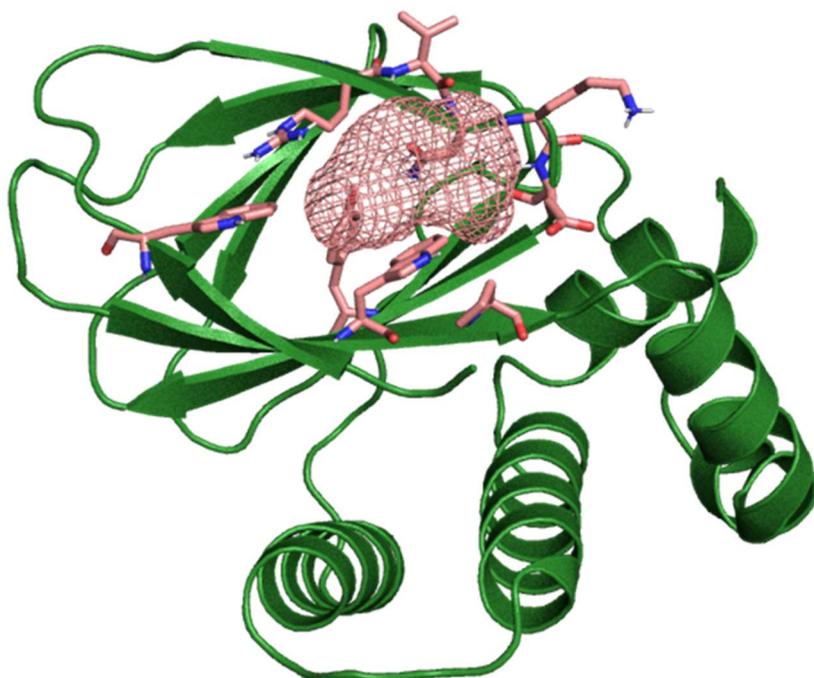


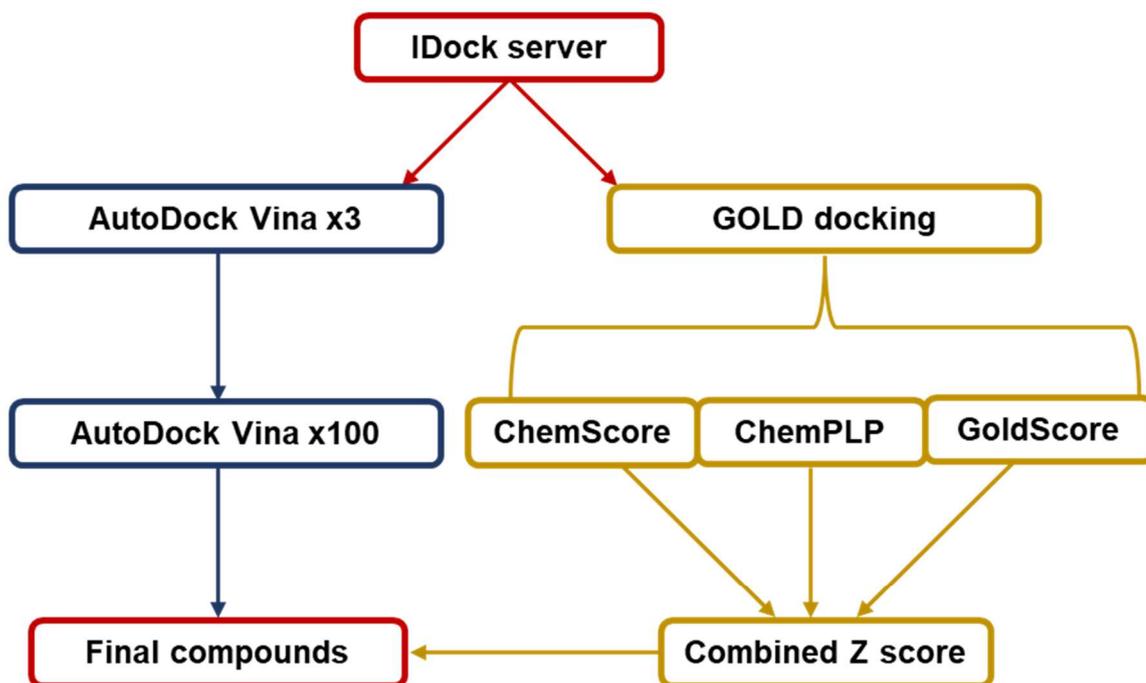
Figure 5.1 Potential ligand binding pocket within the N terminus of ExsA.

The crystal structure (green) of the N terminus of ExsA (PDB: 4ZUA) was subjected to FTSite which identified the best small molecule binding pocket (depicted as a mesh). Residues lining the pocket are shown as pink sticks.

5.2.2 Identification of small molecule ligands that could potentially bind to the ExsA dimer interface

A key region for ExsA dimerization was identified from published mutagenesis experiments (Marsden, Schubot and Yahr, 2014; Shrestha *et al.*, 2015), consisting of amino acids L140, K141, and E143. The peptide spanning residues L140-E143 were therefore extracted from the NTD structure of ExsA (PDB: 4ZUA), and the side chain of I142, which faces away from the dimer interface, was removed. The extracted peptide was converted into a 3D query (Figure 5.3) using vROCS (Open Eye Scientific Software). The latter was then used for a ligand-based VS against a conformer library of the Enamine Discovery Diversity Library (DDS-50, <https://enamine.net/compound-libraries/diversity-libraries>) using Rapid Overlay of Chemical Compounds or ROCS (Open Eye Scientific). The latter ranked molecules using a Tanimoto Combo score that is a sum of 3D shape similarity (the 'Shape Tanimoto' score, maximum value =1) and electrochemical similarity (the 'Colour Tanimoto' score, maximum value =1) (Rahman and Rahman, 2017; Hajbabaie, Harper and Rahman, 2021). The top 500 hits obtained from ROCS-based VS were then re-scored through GOLD 5.3 based focused docking against the desired pocket within the NTD of ExsA using the ChemPLP scoring function.

A.



B.

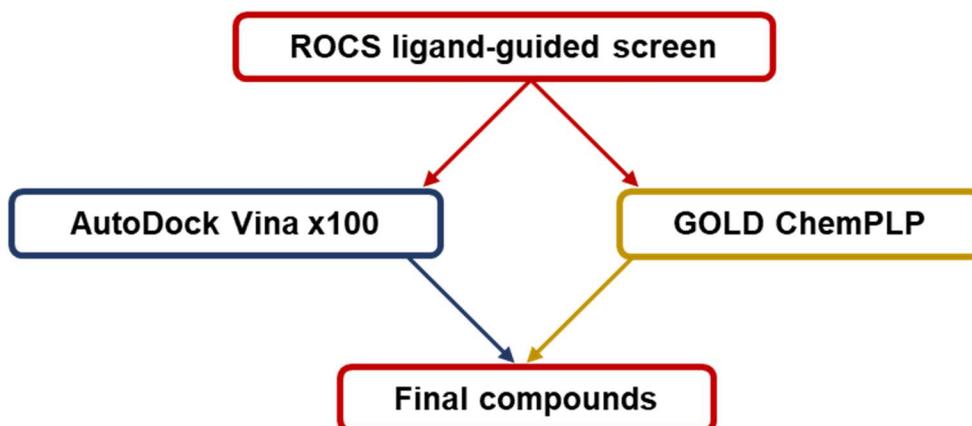


Figure 5.2 Docking workflows.

Schematic depictions of the docking workflows which took place within this chapter. Arrows indicate sequential steps were only the ligands which met a threshold or were in the top bracket of results were carried forward.

5.2.3 Development of a data handling R script

To accelerate and simplify data handling during computational prediction of the next generation of ExsA inhibitors, an R script (available in Appendix 1) was developed. The script took output files from 100 replicate Autodock Vina docking and compared each result to all others at every atomic coordinate, calculating the RMSD for each pose of each molecule. The script then returned the pose which proves most replicable and the pose with the lowest mean RMSD, both with the relevant mean RMSD and percentage values. The script's accuracy was validated by comparison with manual assessment and the RMSD calculator for single replicate data sets within the PyRx platform (<https://pyrx.sourceforge.io/>).

5.2.4 Identification and docking of ROCS hits for L198

The consensus docked poses of L198 (as generated in 5.2.2) were used as a 3D query to search the 2018 eDrugbank library in ROCS, with a stringent Tanimoto combo score (incorporating both shape and electrostatic features) cut off of 1. Three hits met this criterion and were subsequently docked to the NTD of ExsA (PDB: 4ZUA) by AutoDock Vina with 100 replicates and analysed with the aforementioned R script. GOLD docking was performed with the ChemPLP algorithm, utilising the L198 consensus pose as a reference ligand.

5.2.5 Electromobility shift assay

The electromobility shift assay (EMSA) protocol was adapted from several published examples (Brutinel *et al.*, 2008; LaSarre and Federle, 2013; Tian *et al.*, 2019). Mixtures were run on 0.5xTBE, 5% acrylamide gels (5% v/v acrylamide, 10 X Tris-borate-EDTA (Thermo Scientific) adjusted to pH 7 and diluted to 0.5 X, 0.05 % v/v TEMED, 0.08% w/v APS). Gels were run for 2 hours at 100 V in a 4°C room, with 1 X TBE as running buffer. Running buffers and gels were cooled to 4°C prior to use. Gels were imaged on a LICOR at a wavelength of 700 nm. Gels were pre-run for 40 minutes at 100 V.

The probe was designed to mimic the *exoT* promoter via PCR, with the primers AATATCCCATCGGGTTCTCC and GATGATTGACGTCTCCTGATGTTTC used. Both primers were labelled with Cyanine 5.5. After PCR amplification probes were confirmed via agarose electrophoresis and subject to a PCR cleanup with a GeneJET PCR Purification Kit (Thermo Scientific).

Starting reaction mixtures were composed of: 10 mM Tris pH 7.4, 100 mM KCl, 1 mM EDTA, 5% glycerol, 2 mM MgCl₂, 2 mM DTT, 500 ng salmon sperm DNA, 2 µg bovine serum albumin, 0.03105 ng/ µL probe, 0.008 µM purified ExsA. Alterations to this mixture were made as indicated.

5.3 Computational identification of potential ligands against the NTD of ExsA

The pocket within the NTD of ExsA, which is homologous to the binding location of some known ToxT inhibitors (as detailed in Chapter 4 section 4.3.1) was the first targetable site to be explored. FTsite (Ngan *et al.*, 2012) revealed a deep pocket (Figure 5.1) suitable for a structure-guided docking strategy. The compounds were ranked by score for the software at the end of the VS pathway and the top 10 GOLD and top 5 Vina compounds were purchased (Table 5.1). It is noteworthy that one compound, Z980297326 (bold in Table 5.1) scored highly (3rd and 5th for GOLD and Vina respectively) in both screens.

The other potential means for ExsA inhibition is to target the dimerization interface. Utilising published mutagenesis data a key region was established and used as a query in a ligand-based VS (Figure 5.3). A SNP (H131N) reported within the dimer interface was not factored into the selection process, since it is in a different region with no direct contact to the selected region (Figure 5.3 A). The resulting query was used in ROCS to detect drug like compounds with a similar shape and electrostatic profile, and the top 100 compounds were analysed by docking. The top 10 scoring compounds were selected (Table 5.2).

Table 5.1 Top scoring compounds docked to the ExsA ligand binding pocket

Compound name	Gold Combined Z Score⁺	Blind Docking Accuracy* (%)
Z52719569	8.6	62
Z46248170	6.2	
Z980297326	6.1	97
Z27750037	6.1	
Z227103094	6.0	
Z30620096	6.0	
Z1613962124	5.8	
Z46262259	5.7	0
Z1634522876	5.4	
Z30972286	5.3	
Z1273913763	-2.1	100
Z1620689279	-1.1	100
Z1185408250	-1.1	100
Z200099818	-1.2	98
Z809760208	0.1	94

⁺The combined Z score is the mean of function z scores for three separate GOLD docking algorithms compared, where function Z score = (score – mean) / standard deviation.

*Blind docking accuracy indicates the % of docking iterations for which the top scoring pose occupied the desired binding site

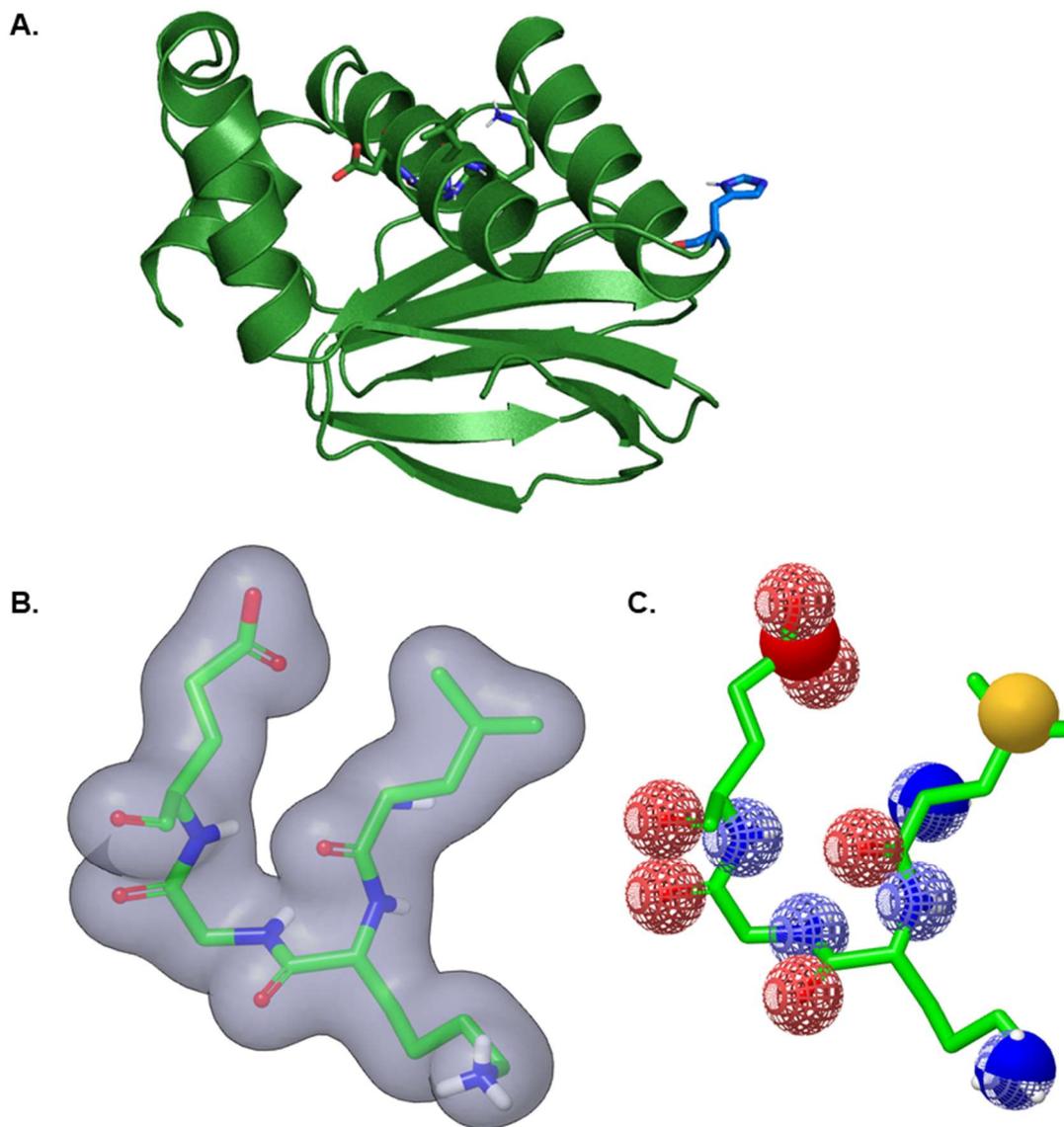


Figure 5.3 The ExsA dimerization interface.

A. The ExsA NTD crystal structure (PDB code: 4ZUA, green) with the two dimerization helices on the top right of the structure. A key region of ExsA for dimerization as chosen for published mutagenesis data is depicted in stick form, whilst in blue H131 is shown in stick form.

B. ROCs query derived from the key residues highlighted in A. The grey surface indicates the search volume.

C. An alternative depiction of the ROCs query, showing the electrostatic features. Red and blue indicate positive and negative areas respectively, whilst the yellow sphere indicates a hydrophobic group.

Table 5.2 Top scoring compounds docked to the ExsA dimerization interface

Compound name	ROCS TanimotoCombo Score	Blind Docking Accuracy (%)*
Z245666992	0.742	100
Z654679278	0.850	98
Z203871970	0.727	95
Z96173580	0.732	67
Z17472212	0.744	66
Z198166104	0.746	61
Z212012366	0.740	52
Z96130645	0.766	51
Z20027124	0.729	51
Z19749967	0.774	49

*Blind docking accuracy indicates the % of docking iterations for which the top scoring pose occupied the desired binding site

5.4 Identification of two small molecules that reproducibly inhibited luminescence of a ExsA dependant Lux reporter

The 28 compounds selected through *in silico* screening were subject to analysis with a *P. aeruginosa* luciferase reporter strain for ExsA activity. The maximum of luminescence was taken and all compounds that caused a reduction of 25% or greater were subject to further analysis, nine candidates met this threshold (Figure 5.4).

Due to the high variability of the results in the preliminary PpcrV-Lux assays (Figure 5.4) further testing on these candidates was performed with four biological replicates, each with 2 technical replicates, all tested at 200 μ M. Of the total 9 compounds tested, compounds Z198166104, Z980297326, and Z1613962124 appeared to have a reduced luminescence (Figure 5.5). Z198166104 demonstrated a statistically significant different OD₆₀₀ adjusted luminescence from 3 to 4 hours ($p < 0.05$). Z980297326 showed a statistically significant difference from the control from 3 hours to 5 hours ($p < 0.05$) whilst Z1613962124 only showed a significant difference ($p < 0.05$) at 5 hours. Maxima analysis (Figure 5.5.C) also showed a reduction in peak Lux values for these three compounds, however only the reduction in Z980297326 was found to be statistically significant ($p < 0.05$, 1-way ANOVA followed by Tukey's post hoc test) (Figure 5.5.C). Z980297326 and Z198166104 were taken forward as lead compounds, referred to from here on as L980 and L198 respectively, due to the greater reduction observed for these compounds compared to Z1613962124.

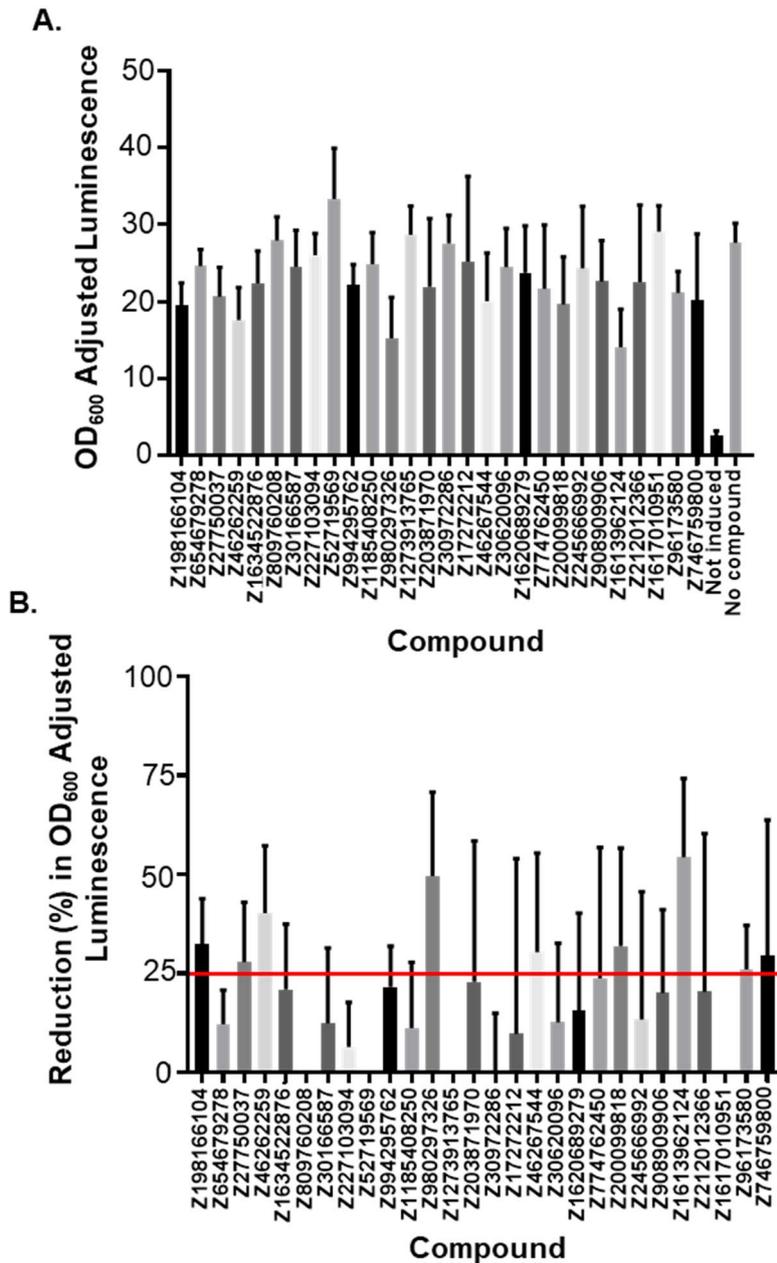


Figure 5.4 Maximum luminescence of *PpcrV* luciferase reporter with putative inhibitors of ExsA.

A luciferase reporter under transcriptional control of an ExsA dependant promoter (the *PpcrV*-Lux reporter strain) was used to find the maximum luminescence in the presence of the putative inhibitory compounds after EGTA induction of ExsA activity. All data is taken from a single time point at the end of exponential growth in *PpcrV*-lux assays with 50 μ M of test compounds and depicts the mean of biological triplicate. Error bars represent standard deviation. **A.** and **B.** show the maximum values and the % reduction compared to the DMSO negative control respectively. A one-way ANOVA indicated significant variation ($p < 0.001$) within the data set; however too many sets are present for post hoc analysis.

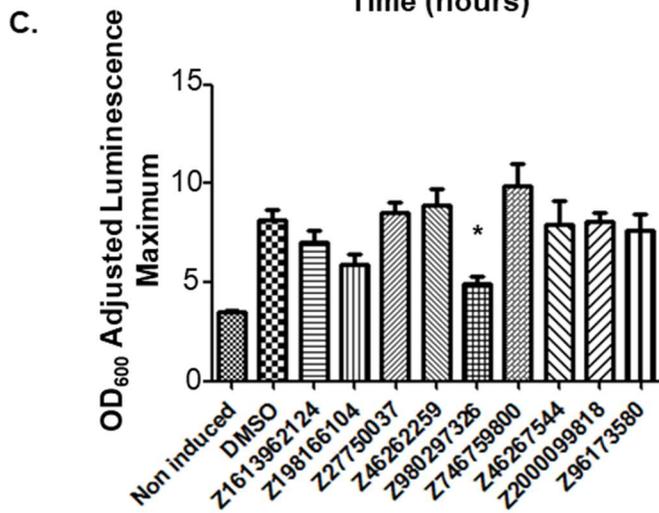
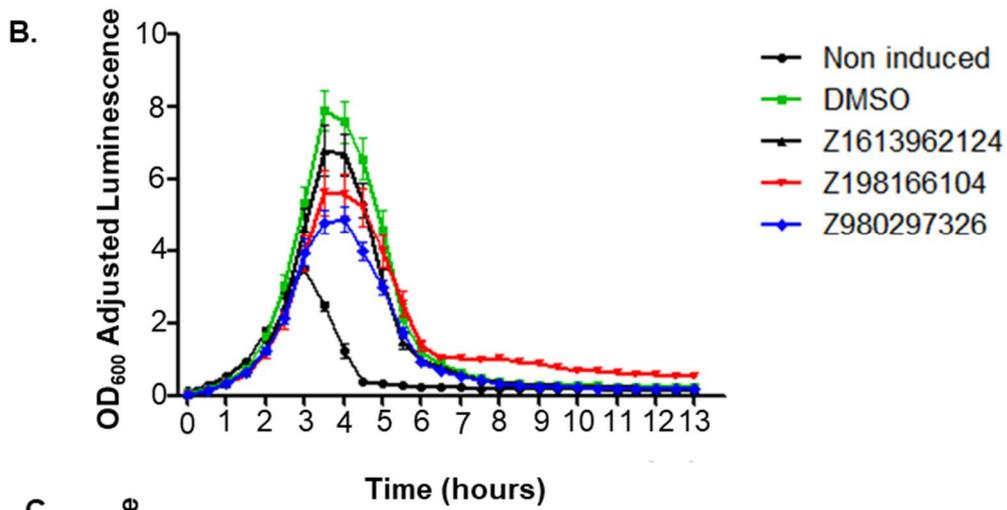
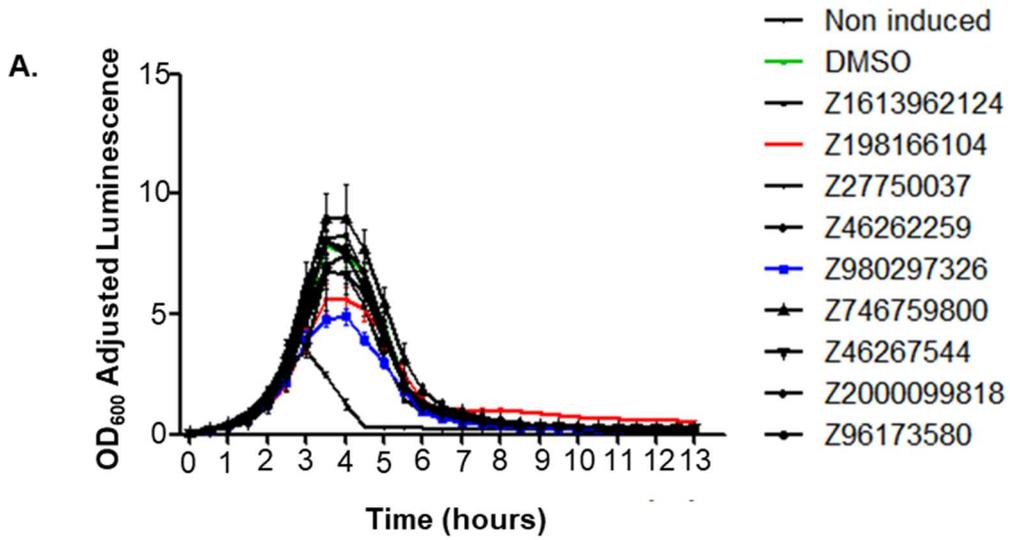


Figure 5.5 Luciferase reporter for selected compounds.

Luminescence readings divided by $OD_{600} \times 10^4$ for the *PpcrV*-Lux reporter strain when grown in AGSY for candidate compounds at 200 μ M, ExsA activity was induced with EGTA as in 5.4. **A.** shows all compounds data, whilst **B.** shows only the compounds which appeared to cause a reduction for greater clarity. All data represent the mean of four biological replicates and are representative of two technical replicates, error bars show standard deviation and are omitted when too small to be visible. Statistical significant deviation from the DMSO negative control at multiple time points was demonstrated for Z198166104 and Z980297326 ($p < 0.05$, 2-way ANOVA with Bonferroni *post hoc* test).

C. All data represent the mean of three biological replicates maxima from the dataset used in A, error bars depict standard deviation, * indicates a significant difference ($p < 0.05$, 1-way ANOVA with Tukey *post hoc* test).

5.5 Compound L198 was lethal to several bacterial species during stationary phase

The growth curves for all but one of the tested compounds appeared to be unchanged from the negative control (Figure 5.6.A). However, L198 had a negative effect at the end of logarithmic growth and the start of stationary phase (Figure 5.6.B). A two-way ANOVA with a Bonferroni post-test gave statistical significance this observation, with every Z198166104 result after 8 hours being significantly different from the DMSO control ($p < 0.001$).

To determine if the observed bactericidal effect was media specific, specific to *P. aeruginosa*, or a wider antimicrobial effect, growth curves in LB media were performed with L198. Three biological replicates were performed, and all statistics represent a two-way ANOVA with a Bonferroni post-test. The same killing effect was observed in wild type PAO1 as in its luciferase reporter derivative; a decline in the L198 treated population at the end of logarithmic growth, which was statistically significant after 7.3 hours ($p < 0.01$) (Figure 5.7). *S. aureus* was unable to grow in the presence of 200 nM L198, and *E. coli*'s growth fell short of its DMSO control when L198 was present, diverging during logarithmic growth. The divergence was statistically significant after 5 hours ($p < 0.01$).

To investigate whether the luminescence inhibition in the *PpcrV*-Lux reporter strain could be caused by the growth inhibitory effect, carbenicillin was used in the assay at sub inhibitory concentrations. Whilst growth was unaffected a slight increase luminescence was observed (Figure 5.8).

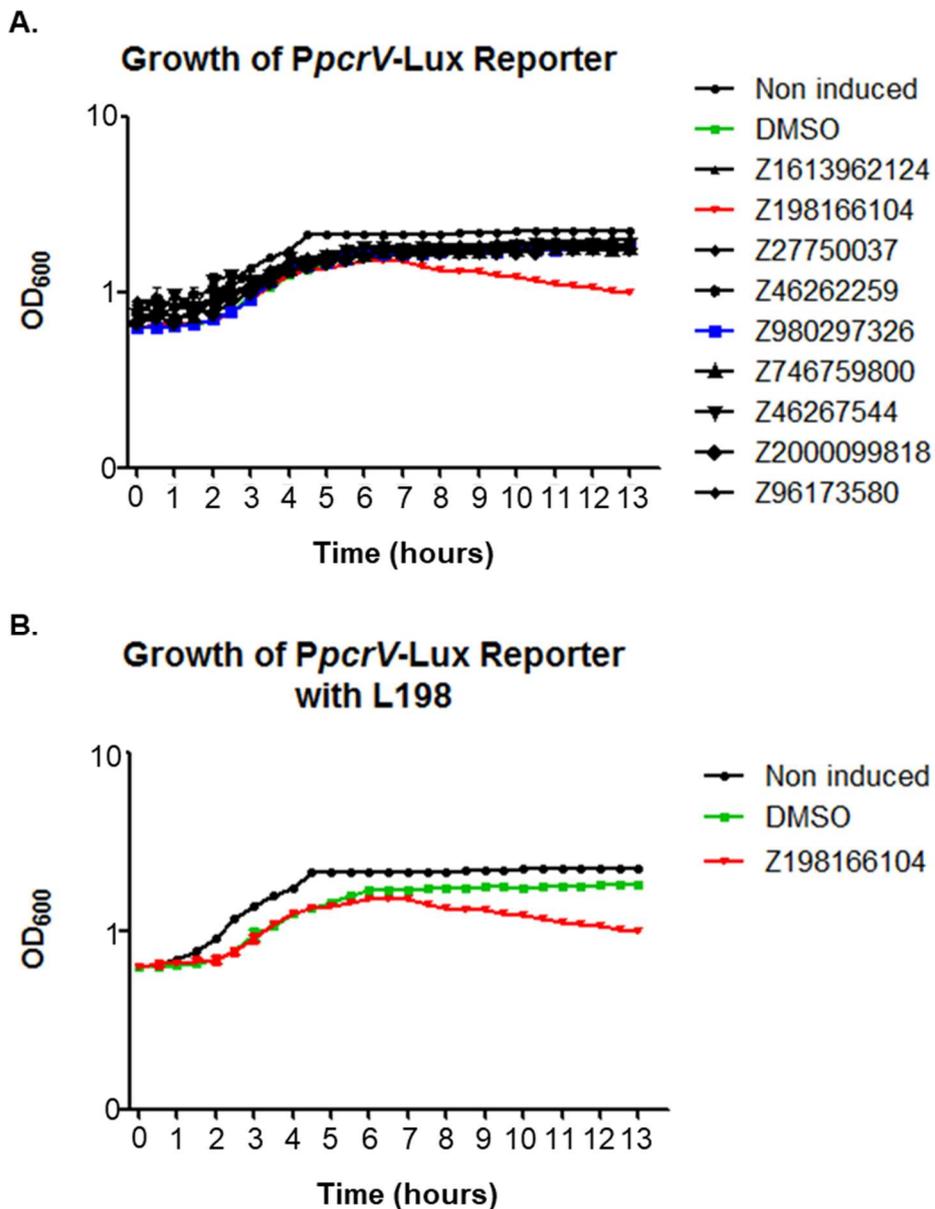


Figure 5.6 Growth curves for candidate compounds with *PpcrV* luciferase reporter strain.

During the luciferase reporter assay OD₆₀₀ is recorded, the growth curves thus obtained and displayed in the present figure are from the same experiment as the data in Figure 5.4. **A.** shows all data collected, whilst **B.** shows the only test compound which appeared to have an effect on growth. All data represent the mean of four biological replicates and are representative of two technical replicates, error bars show standard deviation and are omitted when too small to be visible. Z198166104 diverges from the DMSO control after 8 hours being significantly different from the DMSO control thereafter (2-way ANOVA with Bonferroni *post hoc* test, $p < 0.001$).

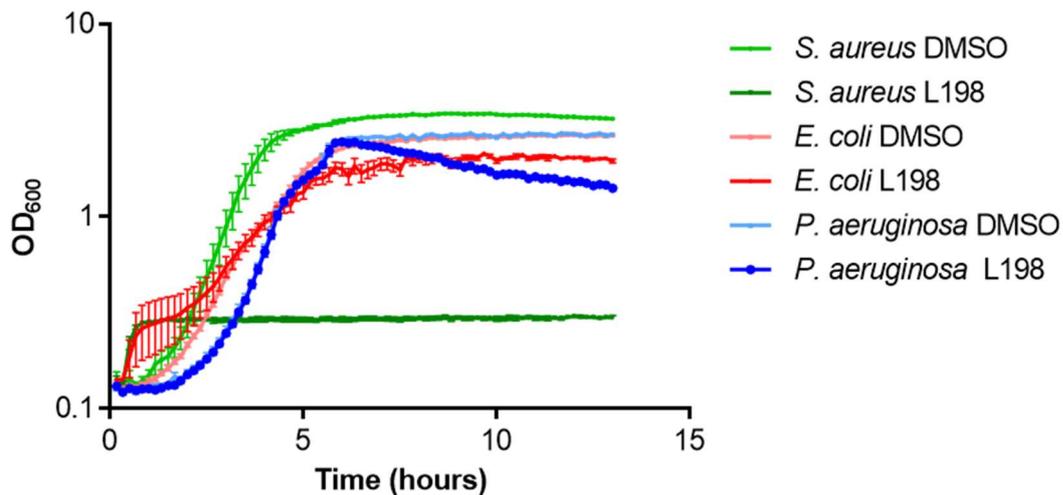


Figure 5.7 Growth of *E. coli*, *S. aureus*, and wild type PAO1 with L198.

Growth curves conducted for *S. aureus*, *E. coli*, and *P. aeruginosa*, with L198 and a DMSO control. Growth curves were performed for 13 hours at 37°C in LB with shaking. Three biological replicates of each were performed and means are plotted, error bars depict the standard error of the mean and are omitted when too small to be visible.

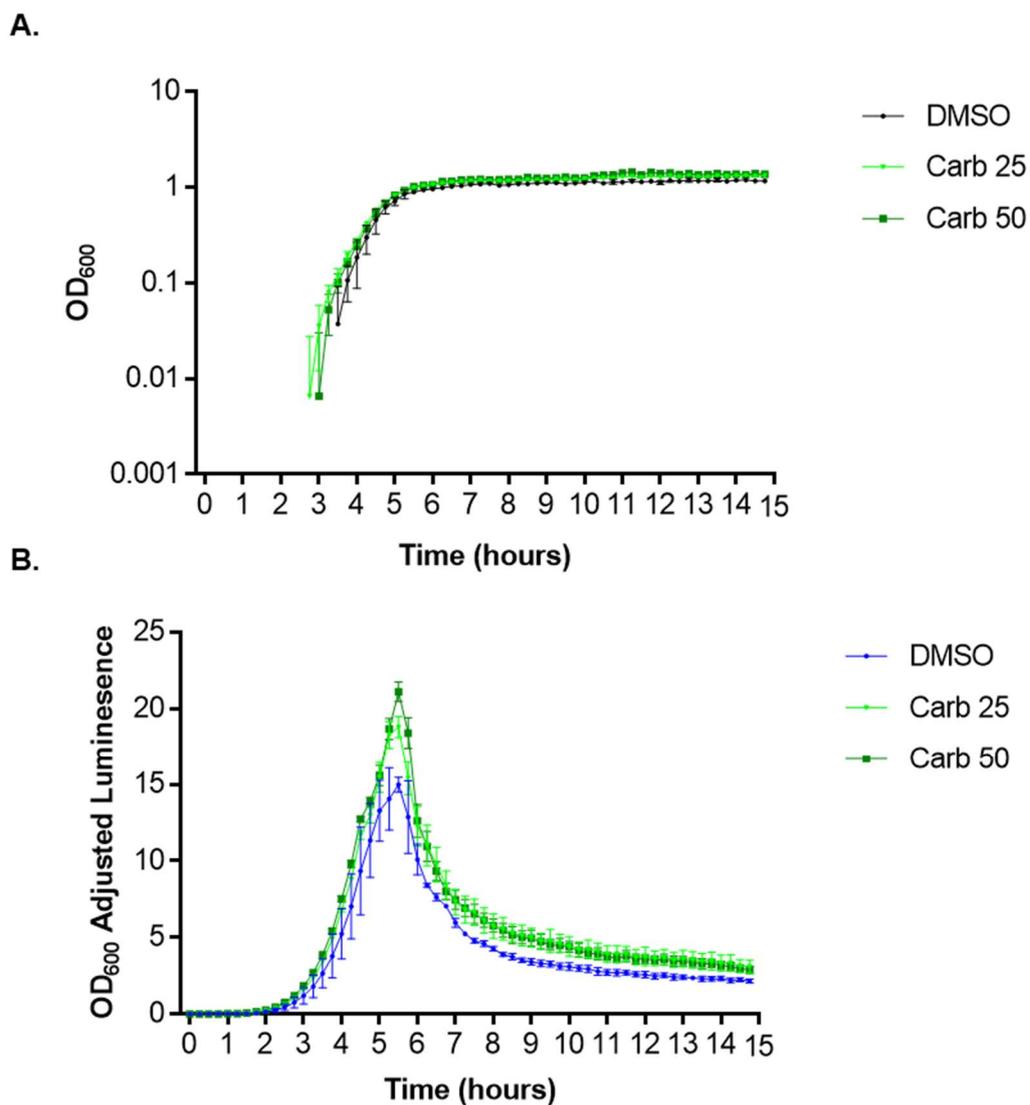


Figure 5.8 Growth and luminescence of *PpcrV* luciferase reporter strain with sub minimal inhibitory concentrations of carbenicillin.

Growth curves (A.) and OD₆₀₀ adjusted luminescence data (B.) for the *PpcrV*-Lux reporter strain of PAO1 with varied carbenicillin concentration in AGSY at 37 °C with shaking. Data points represent the mean of three biological replicates and error bars depict standard error of the mean.

5.6 Compounds L980 and L198 did not inhibit PlacZ-Lux reporter luminescence

To confirm that the inhibition of luminescence observed was a result of the ExsA dependant nature of the promoter controlling Lux expression, a control assay with the same gene cluster under the control of the lacZ promoter was utilised, with the same concentrations, replicates, and statistical analysis as prior. The same killing effect was observed for L198, with a statistically significant decline in the OD₆₀₀ compared to the DMSO control after 10 hours (p<0.001).

The *PlacZ*-Lux adjusted score (luminescence/OD₆₀₀ x 10⁴) also showed a significant difference in L198 at numerous time points, however as the values are higher than those of the DMSO this was likely due to an effect on growth rather than luminescence, and certainly did not indicate inhibition. L980 showed a significant difference at the time point that also showed a growth difference, again indicating the causation is growth rather than luminescence (Figure 5.9).

5.7 Strain YM64 is not suitable for modification as a luciferase reporter

Due to PAO1's efficient efflux systems an efflux deficient reporter strain would be advantageous for detecting functional compounds vulnerable to efflux. A reported strain in the efflux deficient mutant YM64, a strain created by the systematic deletion of efflux pumps from PAO1, was therefore created. However, the luciferase activity of the reporter was highly unusual, with the non-induced samples luminescing to a far greater extent than the induced samples (Figure 5.10).

5.8 No FDA approved drugs are an appropriate replacement for L198

The toxicity of L198 was undesirable in a lead compound for an anti-virulence target, therefore alternatives which preserve the apparent ExsA inhibition *in vivo* without the killing were sought. A ligand-guided screen of FDA approved drugs led to the identification three ligands with a ROCS tanimoto combo score >1. Further docking identified an antipsychotic drug - risperidone which docked with the highest scores and reproducibility to ExsA *in silico* (Table 5.3, Figure 5.11 A). The two other compounds donepezil and nebivolol docked poorly. Despite scoring higher than donepezil in GOLD, nebivolol did not dock in a reproducible manner in Vina (Table 5.3). Donepezil was therefore selected alongside risperidone for further testing. However, neither risperidone nor donepezil had an effect on P_{pcrV} mediated luciferase activity (Figure 5.11 B).

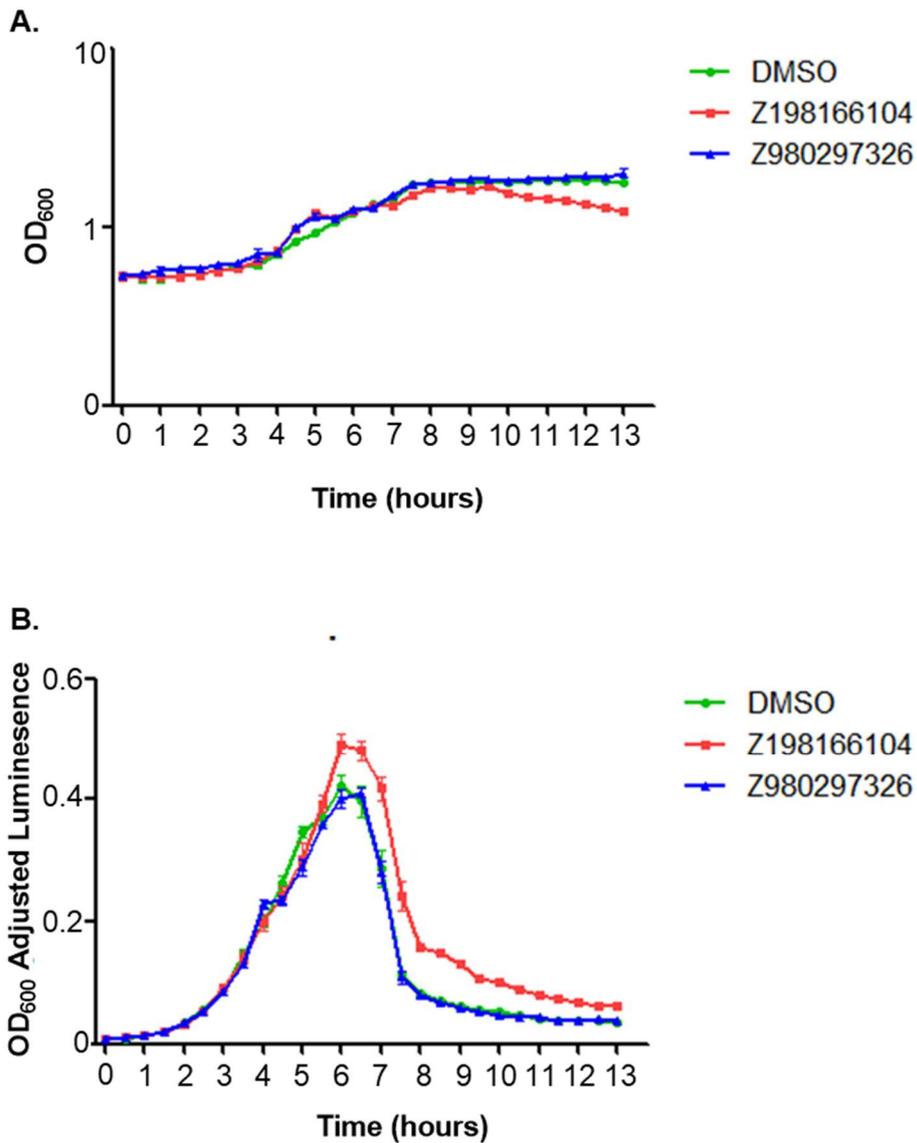


Figure 5.9 The lead compounds do not effect *PlacZ* luciferase reporter luminescence.

Growth curves (A.) and OD₆₀₀ adjusted luminescence data (B.) for the *PlacZ*-Lux reporter strain of PAO1 with L198 and L980 conducted in AGSY at 37°C with shaking. Data points represent the mean of three biological replicates and error bars depict standard deviation.

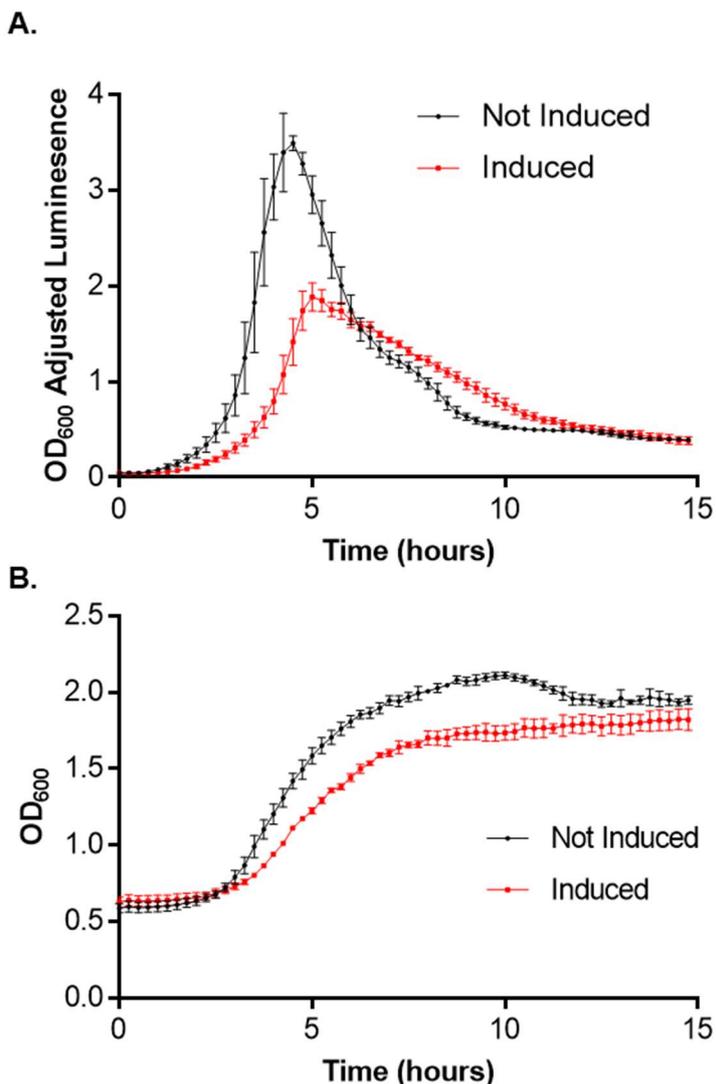


Figure 5.10 YM64 strain is no suitable for reporter strain.

Growth curves (A.) and OD₆₀₀ adjusted luminescence data (B.) for a *PpcrV* luciferase reporter constructed in a YM64 background. The experiment was conducted in AGSY at 37°C with shaking. Data points represent the mean of three biological replicates and error bars depict standard deviation.

Table 5.3 Docking scores for L198 ROCS hits

Drug name	Vina score	Vina docking %	Mean RMSD	ChemPLP score
Risperidone	-8.3	33	0.6	71.16
Donepezil	-7.3	6	0.1	64.44
Nebivolol	-7.7	3	NA	73.60

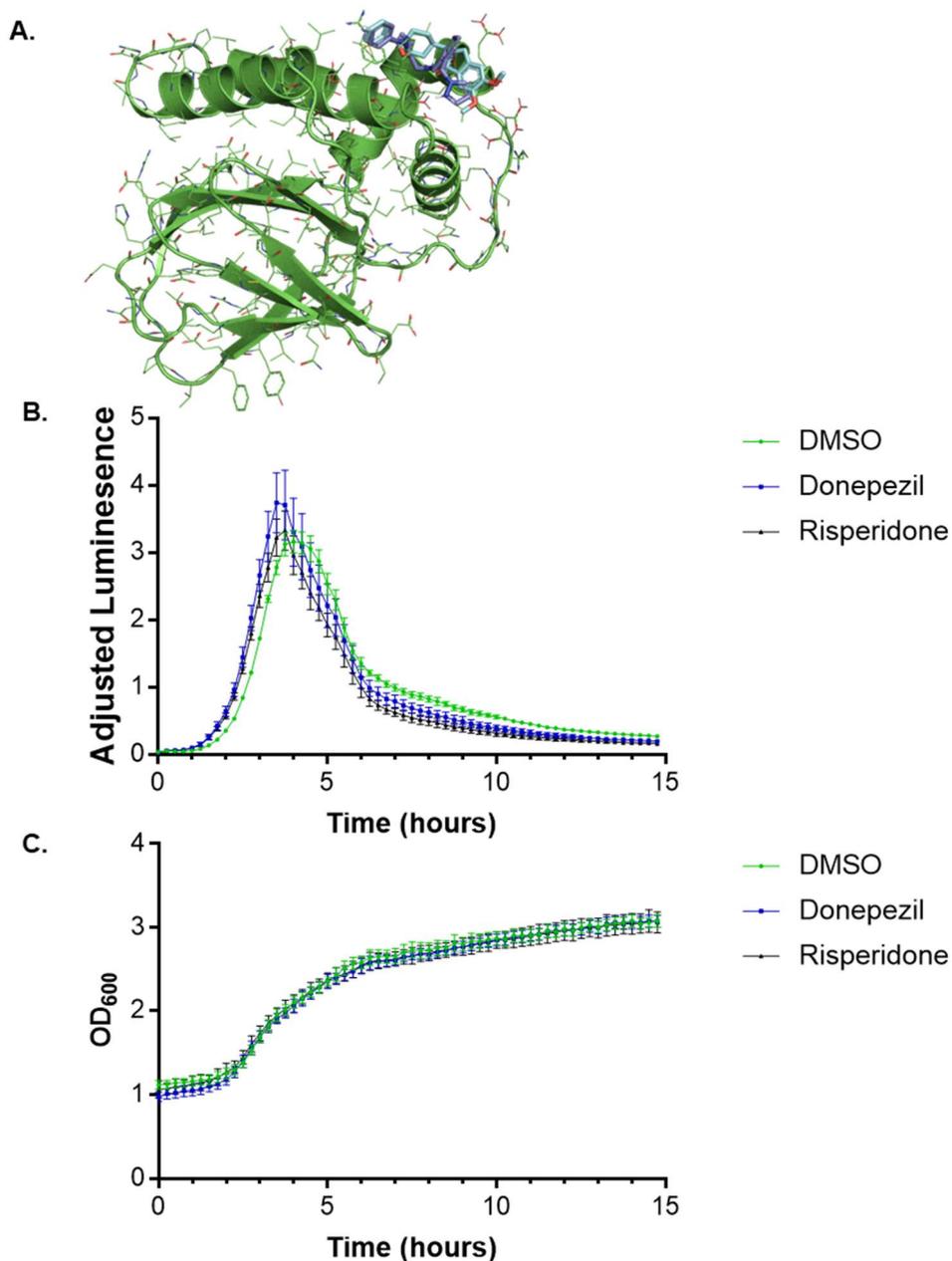


Figure 5.11 Analysis of two FDA approved drugs with potential to replace L198.

A. The crystal structure (green) of NTD of ExsA (PDB: 4ZUA) with the proposed binding modes of donepezil (cyan) and L198 (purple) overlaid.

B. and **C.** show, respectively, OD₆₀₀ adjusted luminescence and OD₆₀₀ growth curves for the *PpcrV* luciferase reporter with the two candidate inhibitors and a DMSO negative control. The experiment was conducted in AGSY at 37 °C with shaking. Data points represent the mean of three biological replicates and error bars depict standard deviation.

5.9 Optimization of the ExsA NTD purification

Due to the full length ExsA proving unamenable to thermal shift (see 4.5.2), NTD alone was sought for thermal shift assays. Initial attempts at purifying the fragment were hampered by low purity and yield. Due to the contaminants appearing to be chaperones an ATP wash was attempted, as some chaperones require ATP to release bound proteins, resulting in an increased purity (Figure 5.12 A), however numerous contaminants remained. Protein sequencing identified these contaminants, a larger band had the ExsA NTD detected within it – indicating another chaperone bound to the product, and the *E. coli* metal binding protein SlyD (Figure 5.12 B, Table 5.4). A novel construct utilising both a His tag and a MBP tag facilitate greatly increased yields and purity.

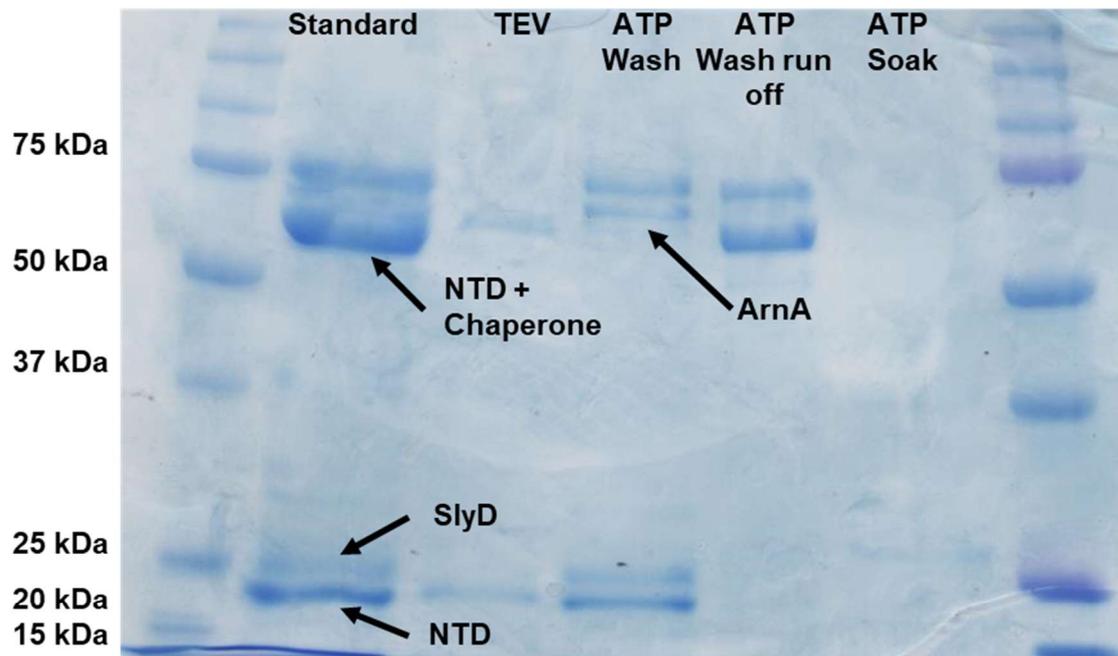
5.10 Thermal Shift Assay for putative ExsA NTD binding compounds

The lead compounds were tested for binding to the NTD of ExsA via the thermal shift assay (Huynh & Partch, 2015), alongside maltose which would bind to the MBP as a positive control. Neither of the compounds shifted the melting temperature, with all melting temperatures (identified by the temperature at which the derivative values were lowest) falling within + or – 0.5°C of the negative control's (DMSO) melting temperature of 52 °C, indicating that the compounds did not bind to the NTD within the assay. The maltose positive control, which binds to the MBP tag increased melting temperature to 58°C (Figure 5.13).

Table 5.4 Mascot search results for ExsA NTD purification products

Band (as labelled in Figure 5.12)	Protein	Mascot Score	% Coverage
1	NTD	2610	85
2	SlyD	1660	47
3	NTD	204	18

A.



B.

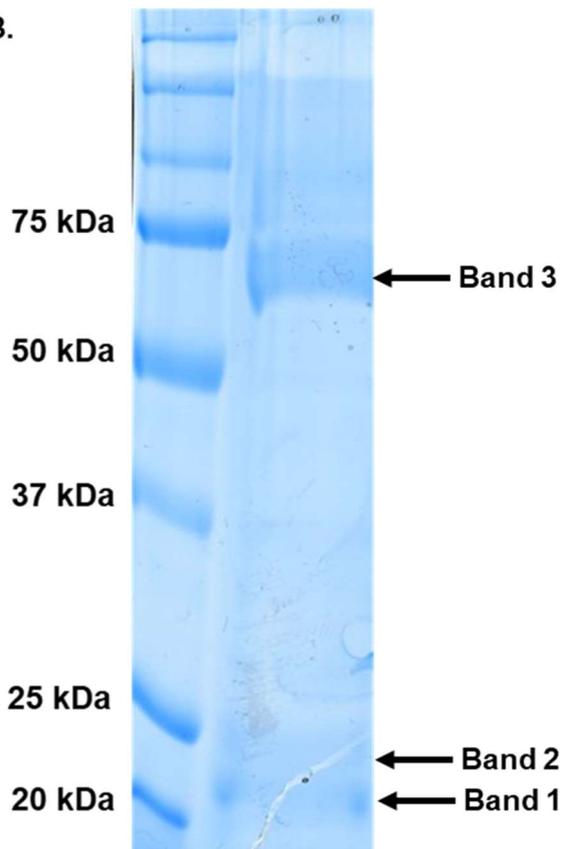


Figure 5.12 Electrophoresis gels for ExsA NTD expression.

Coomassie blue stained acrylamide electrophoresis gels for the purification products of NTD purifications. **A.** shows the elute from the conditions specified, whilst **B.** shows the gel sent for protein sequencing and is labelled to correspond to Table 5.4.

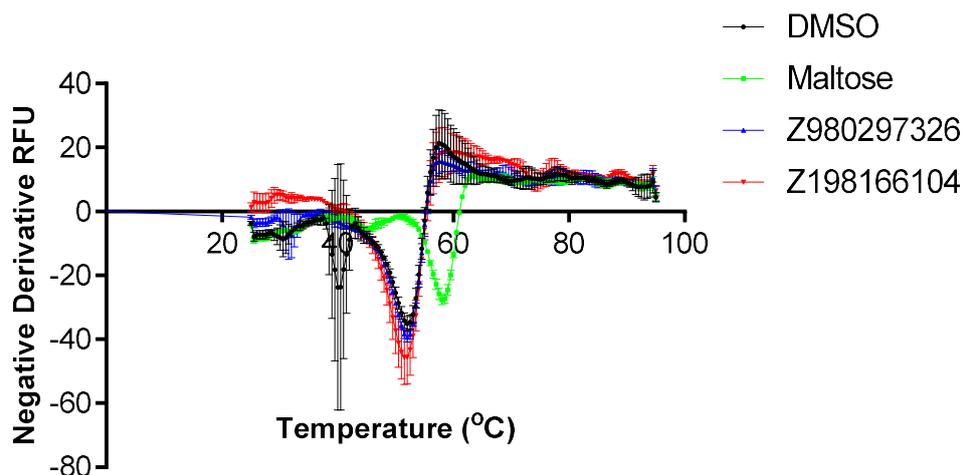


Figure 5.13 Thermal shift assay for lead compounds.

The derivative of relative fluorescence units for thermal shift assays of lead compounds with MBP-NTD fusion protein. Maltose acts as a positive control, and neither compound causes a shift in melting point compared to the DMSO negative control. Error bars depict standard deviation.

5.11 Attempt to establish an electromobility shift assay

As a means to directly evaluate the effect on transcriptional activity of ExsA following binding to any putative inhibitory compound, I decided to establish an EMSA for ExsA-DNA binding. Initial gels were tested to establish a run time, which was non-trivial due to the lack of a suitable loading dye. The gels running on the original protocol suffered excessive smearing, however the addition of ExsA caused addition bands shifted in comparison to DNA alone, and the addition of DMSA had no effect (Figure 5.14 A). A DNA only gel was subsequently run, however excessive smearing was still observed. Salmon sperm was removed from the buffer, and a DNA gel was run again, in which no smearing was observed (Figure 5.14 B). With the addition of ExsA smearing resumed. Trials varying BSA concentrations, or the reintroduction of salmon sperm had no effect, and the smearing was found to be a function of overloading with protein present (Figure 5.15). Due to the lack of an unshifted band the DNA/protein was to be increased next, however due to Covid-19 the experiment could not be pursued further. Attempts to test compounds, or perform unlabelled probe knockdown controls, in the conditions established thus far resulted in unreadable gels.

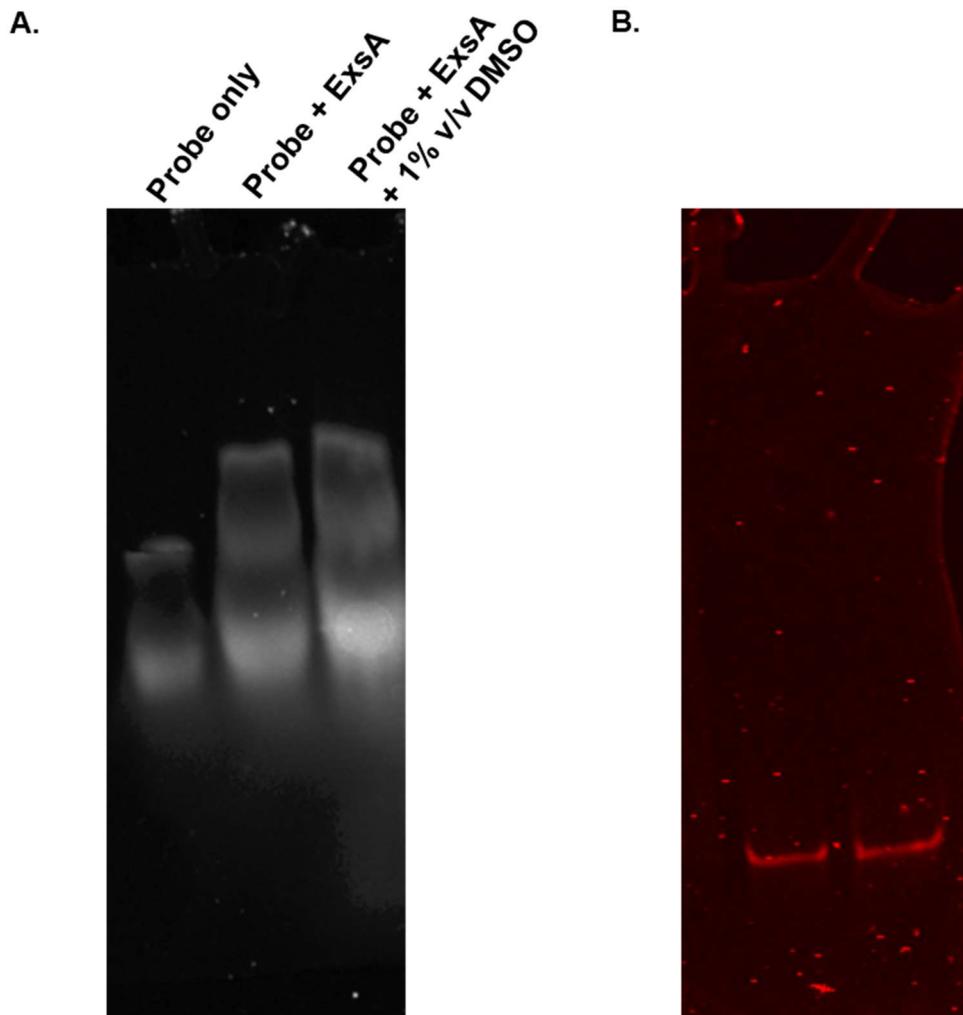


Figure 5.14 Preliminary EMSA gels.

Electromobility shift assay gels for ExsA and a probe mimicking the *exoT* promoter region. Probes were labelled with cyanine 5.5 and gels were imaged at a wavelength of 700 nm, allowing DNA bands to be visualised. **A.** Contains probe only, probe and ExsA, and probe, ExsA, and 1 % v/v DMSO as indicated. The reaction mixtures include salmon sperm. The multiple bands demonstrate shift induced by ExsA-DNA interaction. **B.** Contains only probe in reaction mixture without the addition of salmon sperm DNA, one clean band is formed.

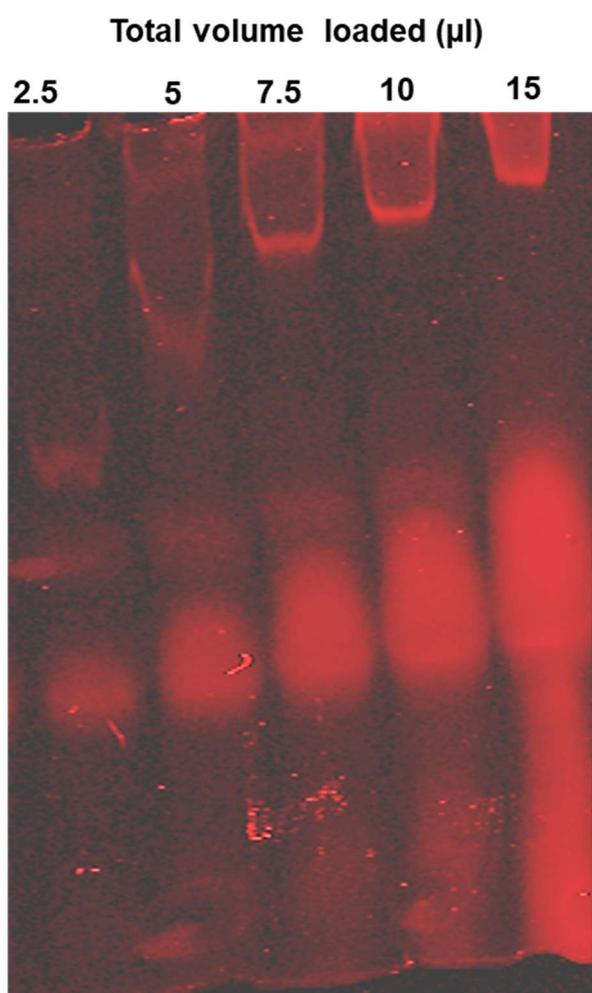


Figure 5.15 EMSA gel for loading volume.

Electromobility shift assay gels for ExsA and a probe mimicking the *exoT* promoter region. Probes were labelled with cyanine 5.5 and gels were imaged at a wavelength of 700 nm. The same reaction mixture, which contained both ExsA and probe but was without salmon sperm, was loaded to each well at the indicated volume.

5.12 Discussion

Initial VS against the FTSite-predicted pocket within the NTD of ExsA yielded interesting putative inhibitors targeting both the ligand binding pocket and the dimerization domain. A high throughput *in vivo* assay was chosen for the subsequent experimental validation of 28 *in silico* hits. This led to the identification of two lead compounds. To further optimise the initial screen, an attempt to make a reporter strain of *P. aeruginosa* in an efflux deficient background was undertaken. Whilst compounds which inhibit ExsA but are susceptible to efflux would not be clinically useful, they could still inform further developments as the initial hit compounds are

improved upon, and contribute to our understanding of the structure-activity relationship. Despite successful genetic modification, the YM64 reporter strain's luminescence did not follow the same pattern of ExsA induced activity as the PAO1 reporter strain. The effect of removing all efflux pumps from the cell envelope could have had unreported consequences in movement across either membrane, or effects on energy metabolism, causing the collapse of luminescence rather than induction upon the addition of EGTA. In any case the YM64 efflux deficient derivative of PAO1 is not suitable and a preliminary reporter able to detect effective compounds vulnerable to efflux was unobtainable.

Of the two compounds, namely L198 and L980, that manifested the greatest efficacy in luminescence assays, neither caused a similar effect in a *PlacZ* reporter strain. This indicated that their effect was specific to the *PpcrV* reporter, and therefore specific to the promoter used, which is known to be regulated by ExsA. This clear indication was however complicated by the observed bacteriocidal effect of L198 against the PAO1 strain of *P. aeruginosa*. During further investigations, similar bacteriocidal effects of the molecule were observed for several bacterial species, including both Gram negatives and a Gram positive. It is conceivable that the reduced luminescence seen in the *PpcrV*-Lux strain was a result of a regulation of ExsA activity in response to the anti-bacterial challenge rather than specific effects. A brief investigation of sub inhibitory concentrations of carbenicillin showed an increase in luminescence rather than a decrease, which supported that the compound L198 had a specific ExsA inhibitory activity.

Whilst the aim of the investigation was to identify anti-virulence compounds, a joint anti-bacterial and anti-virulence compound would also be of interest, albeit losing the presumed reduced pressure towards resistance. At this stage however the compounds in question are only preliminary, and it is conceivable that future structure activity relationship work with L198 could potentially reveal derivatives with only anti-ExsA activity.

Biophysical confirmation of the compounds' interaction with ExsA was desired. Thermal shift assay was selected as the most accessible method, however given ExsA's incompatibility with the method, purification of its NTD only was pursued. To this end an existing plasmid was utilised, however this resulted in insufficient purity. The persistent presence of SlyD, an *E. coli* metal binding protein, proved insurmountable. Due to the absence of the CTD, heparin affinity chromatography could not be utilised to increase purity, and the close similarity in size between SlyD and NTD made size exclusion impractical. A novel construct was therefore made, allowing amylose affinity chromatography to remove contaminants, and as a happy side effect drastically improving yields.

The failure of either lead compound to shift the melting point of ExsA could indicate that the compounds do not bind to the NTD in the assay. There could be two possible explanations of the data, either the results of the luciferase assays were misleading, or the thermal shift assay was not functioning. The luciferase assays indicated that the compounds reduced the transcription of *pcrV*, and whilst this strongly implied an effect on ExsA it does not confirm it. The compounds acting via either the DNA itself or the CTD again seems improbable given their predicted function and that there is two of them. It is possible the NTD was erroneously folded when bound to MBP, or that the MBP occluded the ligand binding sites. To solve this deadlock, an alternative validation was sought in the form of an EMSA.

Progress was made towards developing an EMSA assay, however experimental failure and inexplicable results hampered attempts to improve it. If established, it would allow a simple and clear test of compounds ability to interfere with ExsA's DNA binding by eliminating or reducing the shift induced in the probe by ExsA. If optimised, ExsA should have caused two shifted bands, corresponding to the monomeric and dimeric ExsA binding to DNA (Brutinel *et al.*, 2008). This could thus allow compounds which prevent dimerization to be confirmed as well. Unfortunately, the latter stages of EMSA development were interrupted by the Covid-19 lockdown (see Covid-19 impact statement).

The key limitation of *in silico* strategies the quality of the structure used. In this regard the investigation so far had been deficient, only examining an NTD alone structure rather than a full-length structure. This is unlikely to affect the dimerization interface, which is far removed from the CTD, but the ligand binding pocket is directly adjacent to the CTD. It is possible that a more comprehensive docking strategy with an improved structure could yield more useful compounds.

5.13 Conclusions

In conclusion, both compounds L198 and L980 require further investigation, and the optimisation of an EMSA would settle to conflicting *in vivo* and biophysical data. During the course of this chapter both *in silico* and laboratory methods had been developed, which could allow the more rapid and effective prediction and characterisation of future ExsA inhibitors.

6. Identification of small molecule inhibitors of ExsA - an extended investigation

6.1 Aims and approaches

As outlined in Chapter 5, the NTD of ExsA was subjected to in-depth analysis to identify the potential druggable pocket(s). A preliminary virtual screening campaign led to the identification of two novel scaffolds that proved to be active in a luciferase reporter for an ExsA controlled promoter. Unfortunately, one compound was found to have bactericidal effects against multiple bacterial species, and neither compound manifest any discernible binding against purified ExsA NTD in a thermal shift assay. Nevertheless, the overall learning and the outcomes were encouraging. To identify cleaner hits with definitive evidence of efficacy and target engagement an extended second round of virtual screening against ExsA was undertaken. This time the structure of ExsA was refined, with a full-length model structure replacing the ExsA NTD structure used previously. The virtual screening involved preliminary enrichment of a compound set as potential binders against ExsA through a rapid but somewhat crude virtual screening. This was then followed by more computationally intensive methods to produce a collection of best hits that were then tested with *in vivo* and *in vitro* assays.

6.2 Methods

6.2.1 Generation of protein models

Template identification and initial model generation were carried out in Swiss Model (<https://swissmodel.expasy.org>) (Waterhouse *et al.*, 2018), with PDB structures of 4MLO, 5SUW, 5NLA, and 3OIO used as templates. Models generated on the basis of 4MLO and 3OIO as templates were subsequently refined through Galaxyrefine (<https://www.bio.tools/galaxyrefine>) (Heo, Park and Seok, 2013), then energy minimised and prepared for docking with the global optimization procedure in ICM Pro (Abagyan and Totrov, 1994; Abagyan, Totrov and Kuznetsov, 1994). All visual inspections and alignments were conducted in PyMOL. To construct the full-length model of ExsA, the ExsA CTD was first modelled based on the DNA binding domain of AraC protein from *Chromobacterium violaceum* (PDB: 3OIO). Then the ExsA CTD as well as the published NTD of ExsA (PDB: 4ZUA) were aligned over the full length ToxT protein of *Vibrio cholerae* (PDB: 4MLO) which is a homologous protein belonging to the AraC superfamily of bacterial transcription factor. Following alignment over ToxT, the NTD of ExsA (PDB: 4ZUA) and its modelled C terminus were saved as a continuous, single protein file. This structure of full length ExsA was therefore a hybrid one, comprising of an experimentally solved NTD and modelled C terminus. This

structure was energy minimised and prepared (i.e. removal of steric clashes, addition of missing H atoms etc.) in ICM Pro version 5.3 (Molsoft LLC) and used in all docking experiments described below. FTsite (Ngan *et al.*, 2012) was utilised to identify the residues lining the conserved AraC ligand binding pocket.

The DNA binding helix turn helix motifs within the ExsA CTD were identified by submission of the ExsA protein sequence to the ScanProsite tool (Gattiker, Gasteiger and Bairoch, 2002; Hulo *et al.*, 2006), and manually annotated to the full length ExsA structure model in PyMOL.

6.2.4 Overall docking approach

OpenEye FRED (McGann, 2011) was used targeting the ligand binding pocket with the Enamine 3D diversity library 2019 (Emamine, Ukraine). The top 500 compounds were subjected to further focused docking with GOLD 5.3 suite using the ChemPLP scoring and with GLIDE (Schrödinger Inc.) in its extra-precision ('XP') model. Compounds which fell in the top 10% of any two of these algorithms were subjected to blind docking with Autodock Vina using an exhaustiveness of 16. Ligand interaction diagrams were generated with PoseView (Stierand and Rarey, 2010).

6.2.5 Docking with FRED

A multi-conformer structural database of the Enamine Discovery Diversity Library™(Enamine DDS) was generated with OMEGA (OpenEye Scientific Software, USA) (Hawkins *et al.*, 2010). The 'receptor' for the VS was prepared with MakeReceptor module of FRED (OpenEye Scientific Software) using the ligand binding pocket identified on ExsA structure and no other constraints were imposed. Docking with FRED from the OEDocking suite of programs (McGann, 2011) was conducted as recommended by the proprietor and top 500 hits based on default scoring function (Chemgauss) were obtained.

6.2.6 Docking with GLIDE

The default virtual screening workflow within GLIDE (Friesner *et al.*, 2004) was used to dock Enamine DDS and VITAS Broadway library (<https://vitasmlab.biz/special-databases>) against the chosen pocket on ExsA. The libraries and the receptor were first prepared using the Ligprep and GRID module of GLIDE. Best 10% hits for each library were considered and subjected to visual inspection for binding pose.

6.3 A composite structure of 4ZUA and a CTD model provides the most suitable docking receptor

In order to obtain a full-length model of ExsA for docking, potential template structures were

identified. Two ToxT structures (PDB IDs: 4MLO and 5SUW) and one CuxR structure (PDB: 5NLA) were selected because of their identity and coverage scores (Figure 6.1, Table 6.1). A model of the ExsA CTD was also generated using the structure of an AraC type TF from *Chromobacterium violaceum* (PDB: 3OIO), which had a very high identity, but only partial coverage of ExsA (Table 6.1).

Table 6.1 Candidate template structures for ExsA

Template PDB Code	Identity to ExsA	Coverage of ExsA	Structure GMQE	Model GMQE	Model Qmean	Resolution
5NLA	18.92	0.932	0.51	0.42	-5.68	2.70 Å
4MLO	15.45	0.885	0.54	0.57	-2.75	1.65 Å
5SUW	15.45	0.885	0.54	0.56	-3.42	2.30 Å
3OIO	31.52	0.388*	0.19	0.18	-0.33.	1.65 Å

*3OIO coverage of ExsA is confined to the CTD

The candidate structures were visually compared alongside the truncated ExsA NTD structure (PDB: 4ZUA) and assessed by their scores. The structure derived from the CuxR template (PDB: 5NLA) was disqualified due to both poor score (GMQE: 0.42, Qmean -5.68) and a beta sheet which ran through the target pocket, in stark contrast to the NTD of ExsA (PDB: 4ZUA) (Figure 6.2 A). The other models derived from ToxT structures (PDB IDs: 4MLO and 5SUW) provided very similar structures (Figure 6.2 B), with GMQE scores of 0.57 and 0.56 respectively and Qmean scores of -2.75 and -3.42 respectively, and a reasonable comparison to the ExsA NTD structure (Figure 6.2 C). The model generated the ToxT structure (PDB: 4MLO) as template was selected due to its better score and superior resolution (4MLO = 1.65 Å resolution, 5SUW = 2.30 Å). Whilst the CTD only model generated from the *C. violaceum* AraC protein (PDB: 3OIO), which is homologous to the ExsA CTD could not be compared to the ExsA NTD structure (PDB: 4ZUA) as they have no overlap in protein sequence, scored well with a GMQE of 0.18 and Qmean of -0.33.

The *C. violaceum* AraC protein (PDB: 3OIO) and ToxT structure (PDB: 4MLO) were therefore subjected to refinement and refined models with the best scores were selected (Table 6.2). The ToxT derived model was subjected to a closer comparison to the published ExsA NTD structure (PDB: 4ZUA), focusing on the pocket which would be the target for docking. The positioning of residues lining the pocket varied considerably between the two structures. The residue Tyr33 forms the back of the pocket in the ExsA NTD structure and offers a hydrogen bonding OH group, whereas in the model it is secluded from the exposed surface pocket.

Further, the positions of Trp77 and Arg25 are profoundly different in the model (Figure 6.3 A&B).

Two further models were therefore constructed by aligning 4ZUA and the CTD model constructed from the 3OIO template against the full-length models derived from both the 4MLO and 5NLA models. These models were subjected to visual comparison with respect to information about the original structures. When aligned to the 5NLA model the CTD is much further from the ligand binding pocket compared to the 4MLO aligned model (Figure 6.3 C&D). The 4MLO aligned model was selected for docking.

Upon examination of the model with reference to the SNPs recorded in 4.2.3 it was observed that no SNPs that are reported in the CTD are adjacent to the ligand binding pocket (Figure 6.4 A). One SNP, T200A reported in strain LESlike1, arises in a HTH DNA binding motif (Figure 6.4 B).

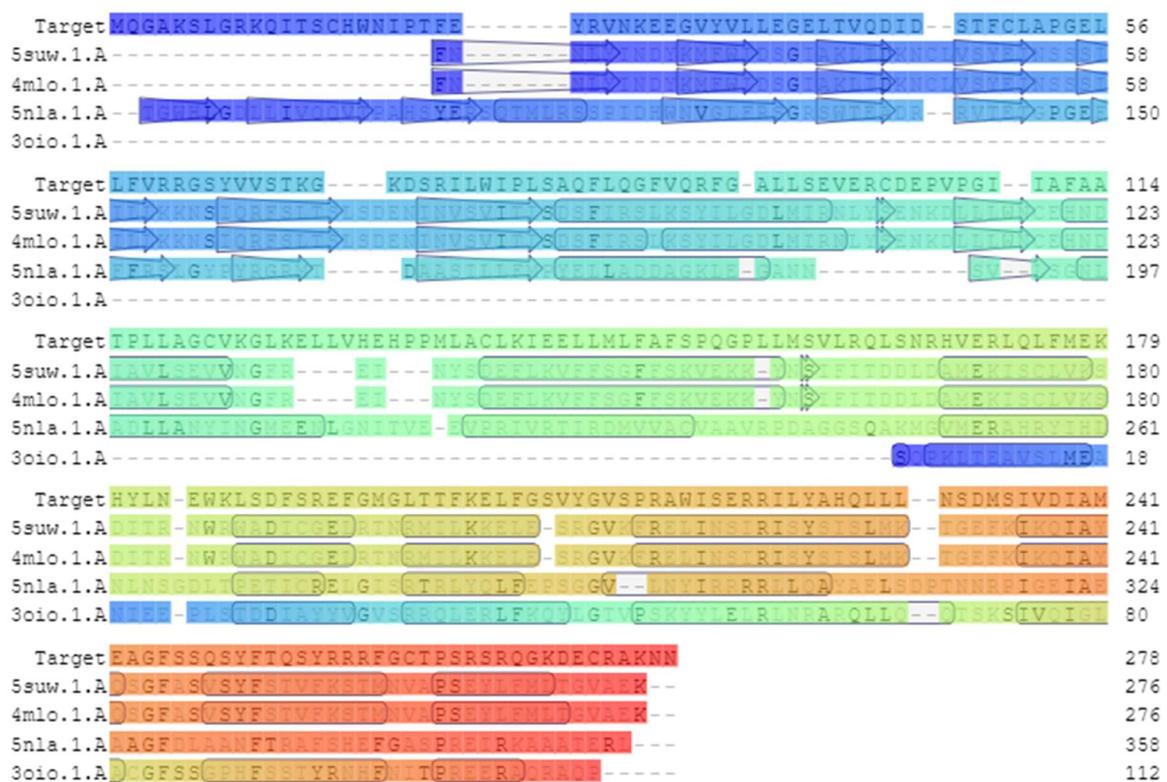


Figure 6.1 Alignment of selected sequences to ExsA.

The alignments of selected template sequences to the full length PAO1 ExsA sequence, as conducted in SwissModel (<https://swissmodel.expasy.org>). "Target" denotes the ExsA sequence.

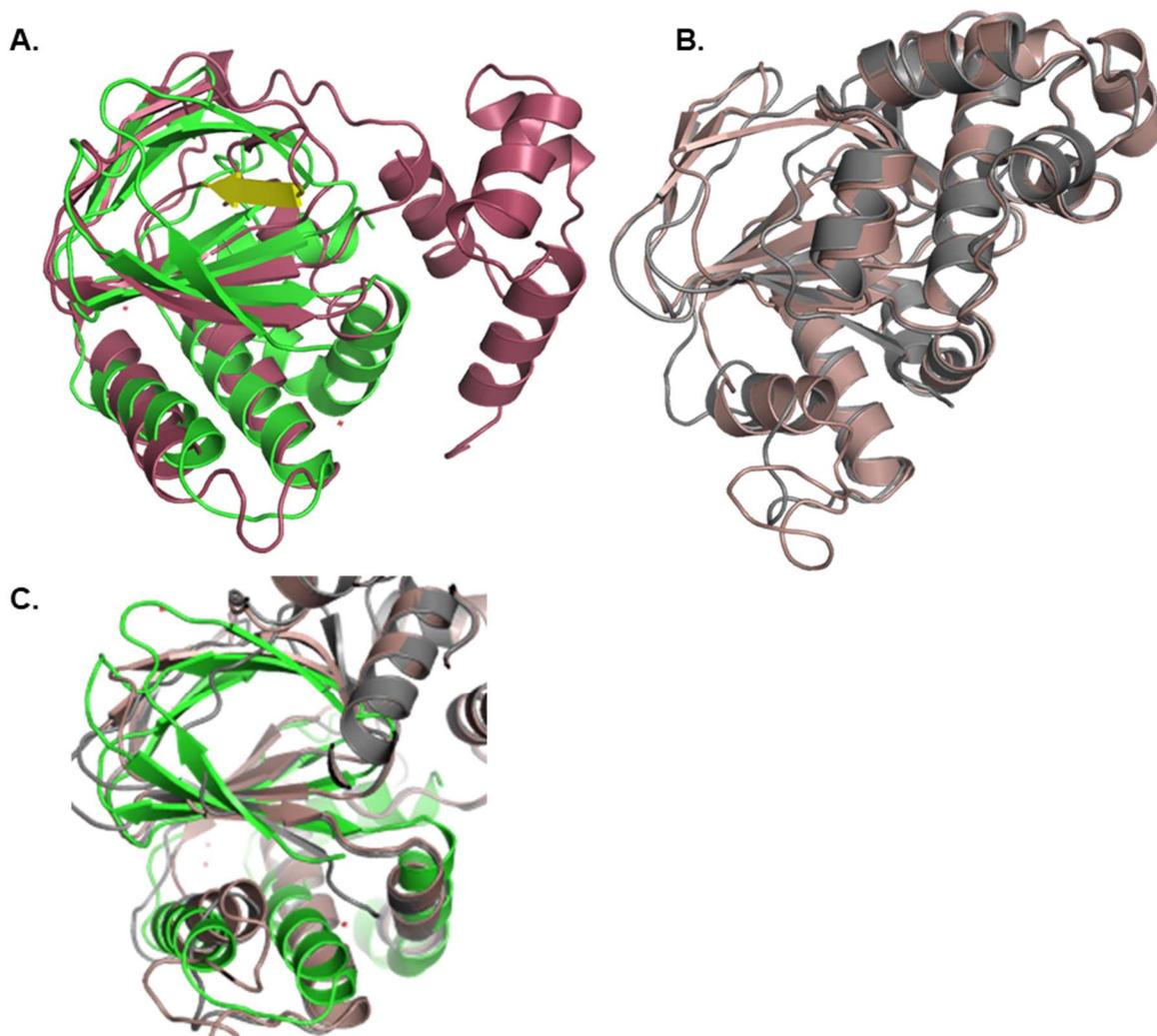


Figure 6.2 Visual Comparison of ExsA Models and 4ZUA.

A. Cartoon depictions of the aligned structures of the NTD of ExsA (PDB: 4ZUA, green) and a model of ExsA using the structure of CuxR (PDB: 5NLA, red) as a template. Highlighted in yellow is a beta sheet within the model structure, which runs through the ligand binding pocket as identified on 4ZUA.

B. Cartoon depictions of the aligned model structures of ExsA which used ToxT structures as a template (template PDB IDs: 4MLO in grey and 5SUW in copper)

C. Cartoon depictions of the aligned structures models which used structures 4MLO and 5SUW as templates (grey and copper respectively) aligned to 4ZUA (green). Part of the model's CTD is not displayed to allow a better view of the NTD including the ligand binding pocket.

Table 6.2 Refinement scores for ExsA models

Template PDB Code	GDT-HA	RMSD	MolProbity	Clash score	Poor rotamers	Rama favored
4MLO	0.9743	0.362	1.825	21.4	0.9	98.5
3OIO	0.9837	0.283	1.686	7.1	2.4	100

Table 6.2 legend

GDT-HA and MolProbability scores are statistical calculations of how probable the structure is. RMSD indicates deviation in Å from the original structure, Clash score indicates the side chains closer to each other than is sterically preferable, whilst poor rotamers and Rama favoured indicate the favourability of bond angles. (Zemla A. LGA: a method for finding 3D similarities in protein structures. *Nucleic Acids Res.* 2003;31:3370–3374. and Chen VB, Arendall WB, III, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 2010;66:12–21).

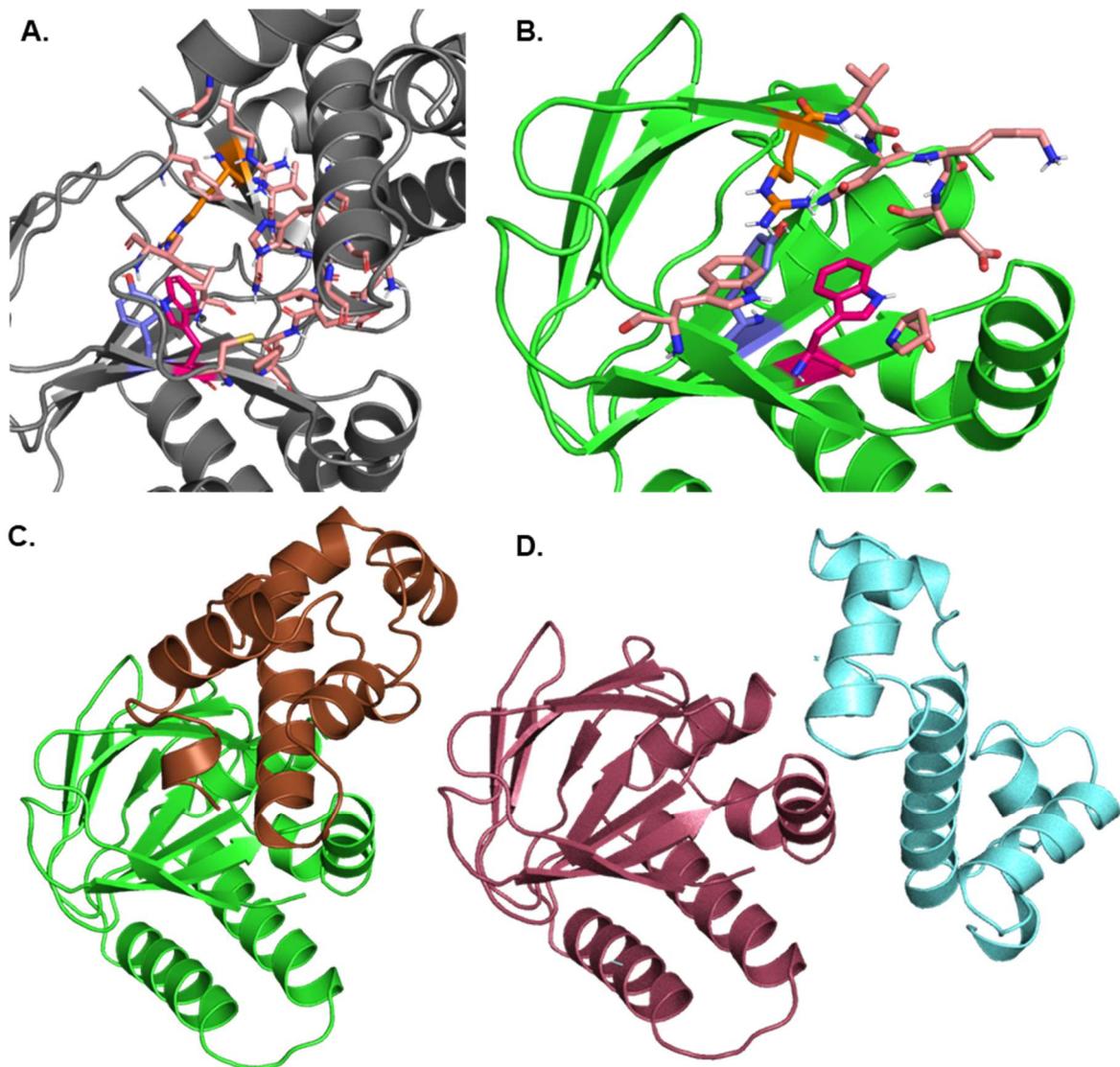


Figure 6.3 Visual comparison of final ExsA models with a focus on the target pocket.

A. The full length ExsA model constructed from a ToxT template structure (PDB: 4MLO), focused on the ligand binding pocket. Residues shown in stick form were identified by Ftsite as lining the ligand binding pocket. Key residues are colour coded as follows: pink = Trp77, purple = Tyr33, orange = Arg25.

B. The ExsA NTD structure (PDB: 4ZUA) focused on the Tyr33 pocket. Residues shown in stick form were identified by Ftsite as lining the ligand binding pocket. Key residues are colour coded as follows: pink = Trp77, purple = Tyr33, orange = Arg25.

C. The combined structures of the ExsA NTD (PDB: 4ZUA, green) and the CTD model derived from a AraC family protein from *C. violaceum* (PDB: 3OIO, bronze) positioned relative to one another by alignment to the final model templated of ToxT (not shown).

D. The combined structures of the ExsA NTD (PDB: 4ZUA, green) and the CTD model (as above) positioned relative to one another by alignment to the model templated on CuxR (PDB: 5NLA, not shown).

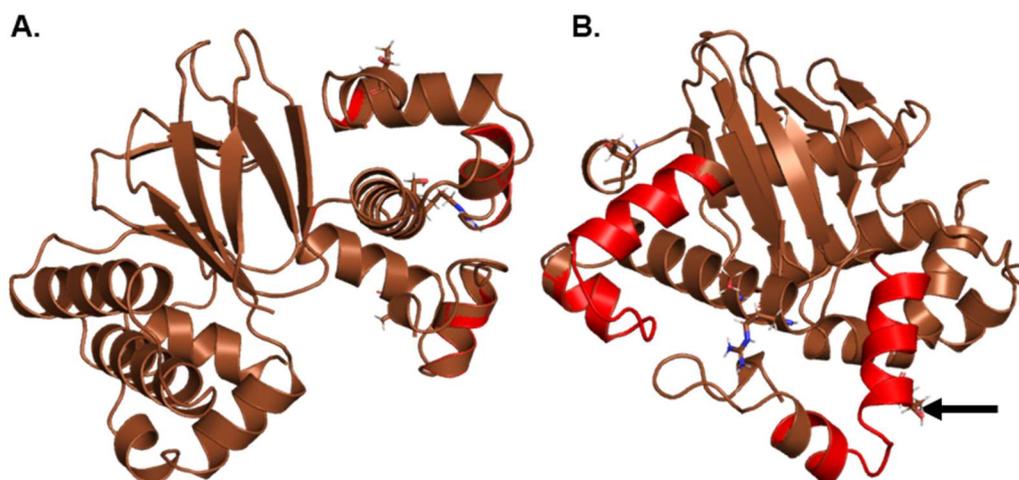


Figure 6.4 SNPs identified on the ExsA CTD.

A cartoon depiction of the ExsA full length model (brown), with the helix turn helix DNA binding motifs as identified by Prosite in red. All residues which have been identified as sites of polymorphism on the CTD are shown in stick form. In **A.** the structure is oriented to give a clear view of the target pocket and the adjacent section of the CTD, whilst **B.** is oriented to show the helix turn helix motifs more clearly. The arrow highlights the residue Thr200.

6.4 Varied docking algorithms show little correlation in score

Docking was conducted using multiple software based on several different algorithms, and their scores for each compound were plotted against each other. It was evident that no general correlation between the scores of different algorithms existed (Figure 6.5).

6.5 Structure-guided screening reveals potential hits

Compounds that achieved scores in the top 10% of two or more of the following scoring systems were collected for further analysis: GOLD ChemPLP, Glidescore, and FRED. The compounds thus identified were subjected to further docking with 100 replicates via Autodock Vina and pose consistency and scoring across all docking algorithms used was assessed. The following pose features were also visually inspected for: occupation of the pocket and formation of at least one polar interaction within the pocket, notably hydrogen bonds to Tyr33.

It was observed that many of the compounds with a good pose consistency and high scoring across multiple methodologies contained the same core 4-hydroxyquinazoline chemical group. For instance, all three compounds to score in the top 10% for all three docking methods contain the group, and almost all compounds that scored highly across multiple software included the group (Table 6.3, Figure 6.6). The same trend was observed with a high proportion of every 10% category possessing the group. 4-hydroxyquinazolinewas therefore assessed by docking in Autodock Vina, and it was found to dock consistently in the same

position whether it was part of a larger compound or not (Figure 6.7). Out of the short-listed compounds it was decided that Z195013052, Z17142402, and Z1756820653 would be tested experimentally alongside 4-hydroxyquinazoline due to scoring and pose reproducibility across multiple software (Figure 6.7 & 6.8). Upon delivery, it was apparent that Z17142402 was insoluble, and therefore it could not be tested.

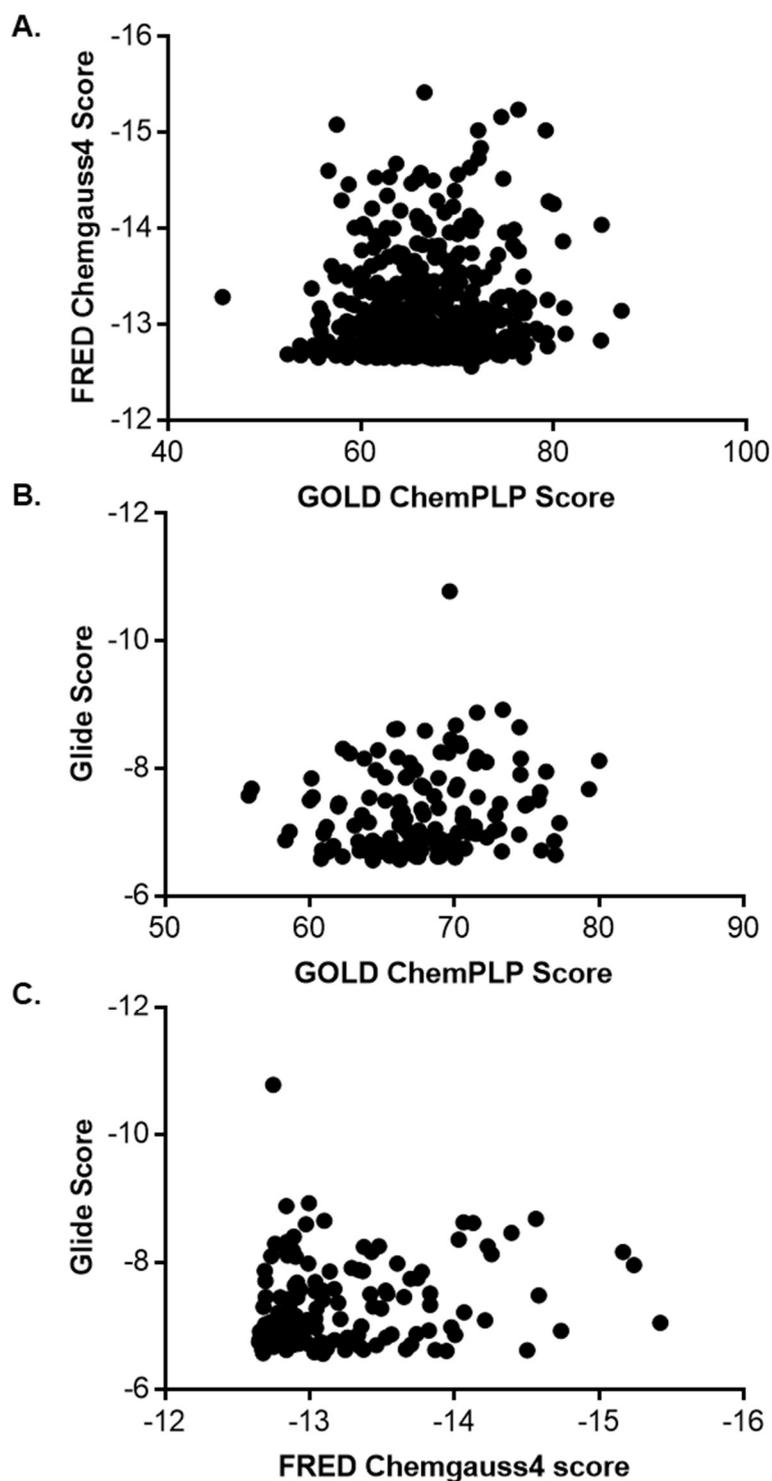


Figure 6.5 Comparison of the top 500 FRED compounds with scoring in GOLD and Glide.

The top 500 scoring compounds docked in FRED to ExsA's ligand binding pocket were docked to the same site in GLIDE and Gold. Each data point represents the docking score of one compound. **A.** Hermes GOLD ChemPLP algorithm to the original FRED Chemgauss4 score. **B.** Hermes GOLD with the ChemPLP algorithm to the final score obtained from a tiered Glide docking approach. **C.** Glide score compared to the FRED Chemgauss4 score.

Table 6.3 A Selected compounds docking scores

Name	Fred Score	ChemPLP score	Glide score*	Methods for which compound is in the top 10%
Z20230049	-15.241922	76.35	-7.957	All
Z28262199	-14.257053	80	-8.129	All
Z20247356	-15.165359	74.58	-8.16	All
Z20242030	-14.2305	69.59	-8.252	ChemPLP, Glide
Z17142402	-15.024	79.17		FRED, ChemPLP
Z195013052	-13.1012	74.5	-8.653	ChemPLP, Glide
Z20242138	-13.8328	75.82	-7.512	FRED, Glide
Z1756820653	-14.5221	74.78		FRED, ChemPLP

*Where a glide score is no displayed the compound did not score highly enough in the early phases of the tiered docking system to receive a final score.

Table 6.4 B Select Vina analysis

Name	Vina Score	Pose replicability (%)	Average RMSD (Å)	4-hydroxyquinazoline group
Z20230049	-9.1	34	0.64	Yes
Z28262199	-9.3	17	1.9	Yes
Z20247356	-9.4	82	3.74467	Yes
Z20242030	-9.3	60	1.5	Yes
Z17142402	-9.2	76	0.383884	Yes
Z195013052	-9.2	50	0.51977	No, but a similar group
Z20242138	-9.3	50	2.28876	Yes
Z1756820653	-8.7	91	4.91301	No

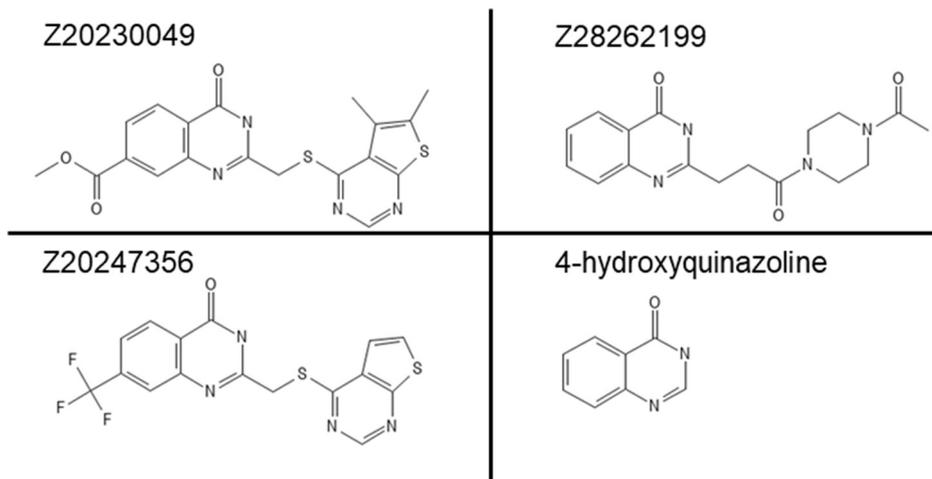


Figure 6.6 An illustration of 4-hydroxyquinazoline ligands.

Two dimensional skeletal formulae for the three compounds which scored in the top 10% of all three docking programs and 4-hydroxyquinazoline.

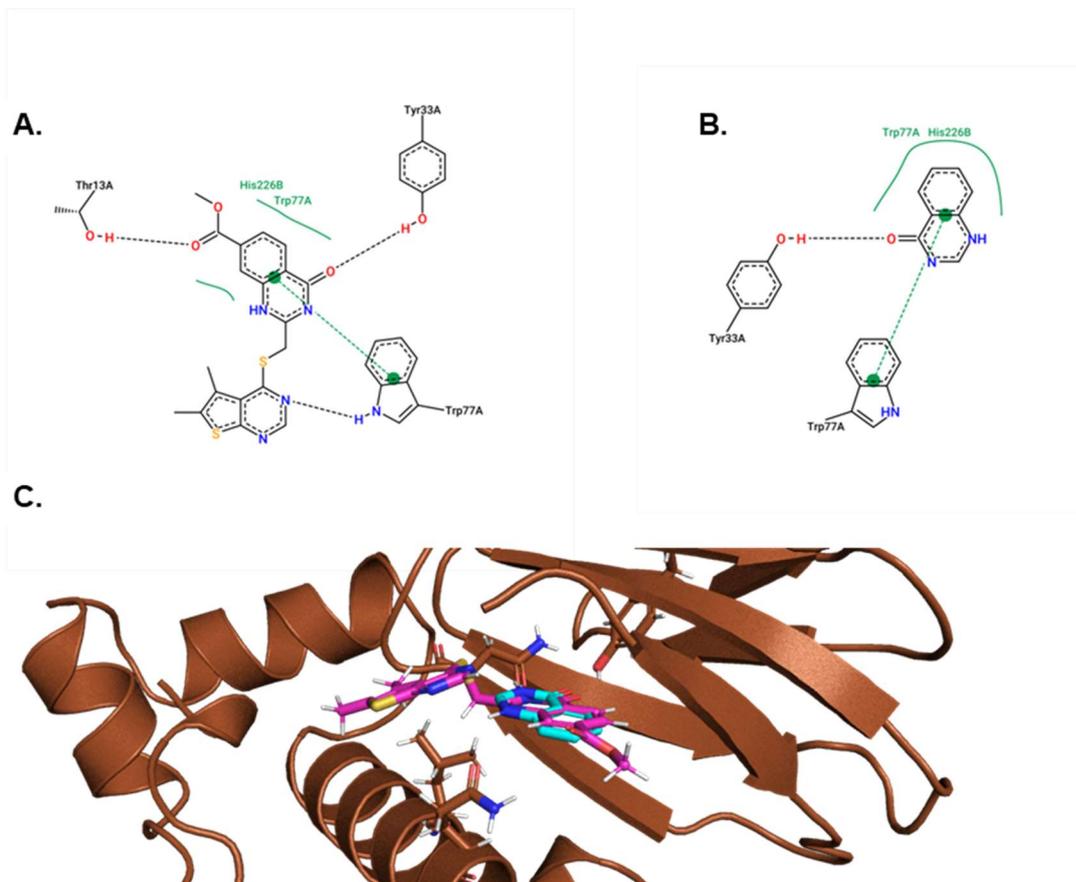


Figure 6.7 Docked poses of Z20230049 and 4-hydroxyquinazoline.

Illustrations of the docked poses of Z20230049 (FRED docking pose) and 4-hydroxyquinazoline (Autodock Vina docking pose) in the Tyr33 pocket of the ExsA full length model. **A.** and **B.** are 2D images rendered in PoseView depicting Z20230049 and 4-hydroxyquinazoline respectively. Green lines identify areas of hydrophobic interaction, and dotted lines depict hydrogen bonds or ring stacking effects when between two green circles. **C.** contains both poses with Z20230049 shown in pink and 4-hydroxyquinazoline in blue. The receptor is shown in brown with residues around the docked site shown in stick form.

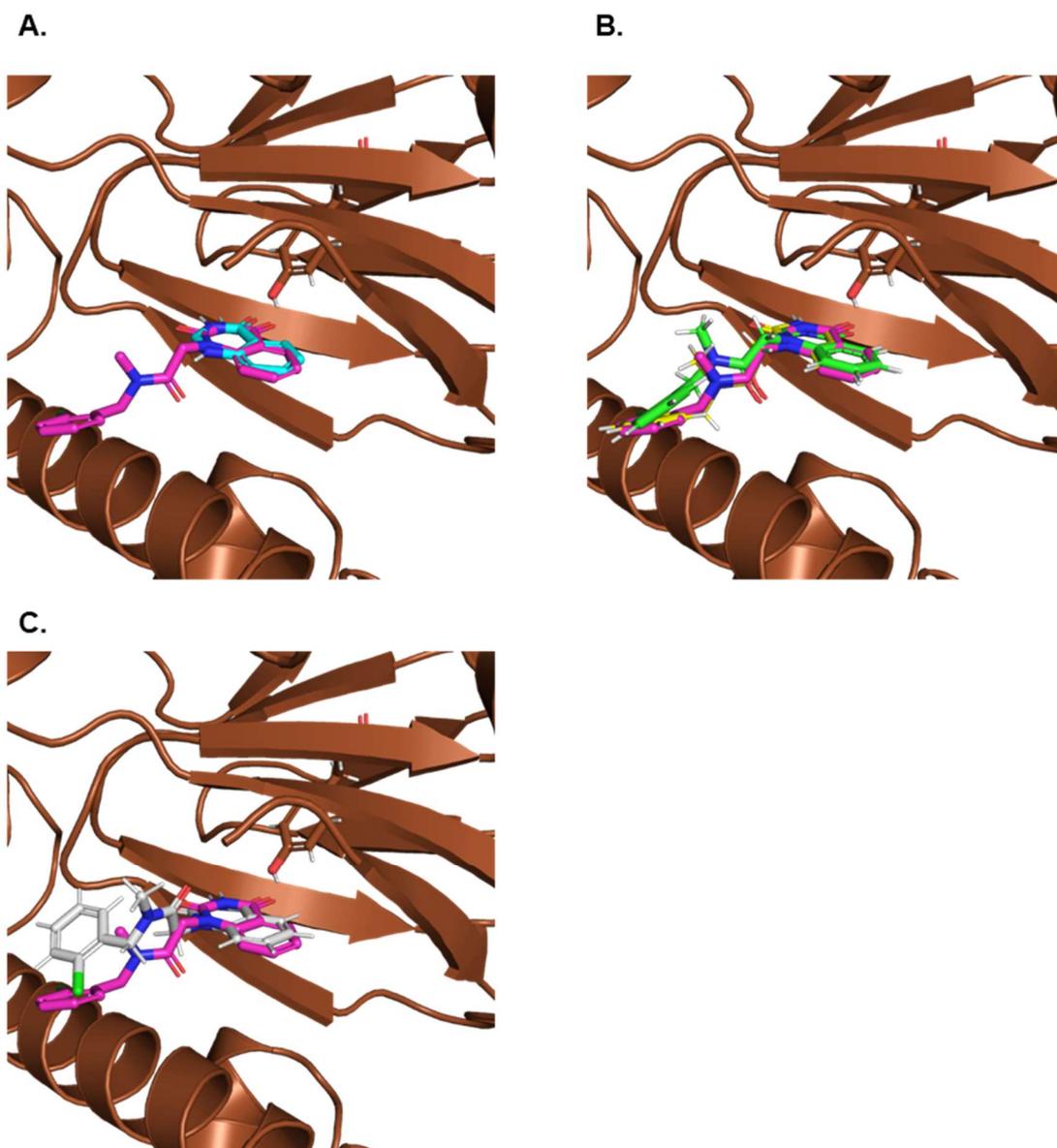


Figure 6.8 A visual comparison of docked poses for Z195013052.

The docked pose of Z195013052 in the Tyr33 pocket of the full length ExsA model (brown throughout). **A.** Z195013052 (pink) and 4-hydroxyquinazoline (blue) Autodock Vina poses selected as the most replicable from a 100 replicates. **B.** Z195013052 Autodock Vina pose as prior (pink), and the top scoring ChemPLP (yellow) and Glide (green) poses. The ChemPLP pose is obscured by the highly precise overlap with the Glide pose. **C.** Z195013052 poses from Autodock Vina as prior (pink) and the top scoring FRED pose (grey).

6.6 Virtual screening and subsequent experimental testing of the in-house putative PARP inhibitors

It was noted that 4-hydroxyquinazoline is a core group of many PARP inhibitors. Whilst of no biological relevance to this project, the lab maintains an in-house library of putative PARP inhibitors, some of which contain 4-hydroxyquinazoline, or a group bearing chemical similarity. Given that the compounds were already available in the lab, and there was therefore no cost to acquiring them, a less stringent threshold was set, and the library was docked with FRED and GOLD ChemPLP. Compounds were then selected for testing according to either docking results or presence of 4-hydroxyquinazoline moiety (Table 6.5).

Table 6.5 Parp Library compounds docking scores

Name	Fred Score	ChemPLP score	4-hydroxyquinazoline
MolPort-000-835-044	-11.809283	69.41	Yes
MolPort-006-806-463'	-10.076071	66.35	Yes
Z1480754091	-11.671803	68.26	No
Z1419498936	-11.127297	78.87	No
MolPort-001-731-331	-11.08763	71.4	No
Rutaecarpine	-11.183145	63.89	No

The identified compounds were tested for binding to NTD-MBP using the thermal shift assay. None of the compounds were found to cause a change in melting temperature (Figures 6.9 & 6.10).

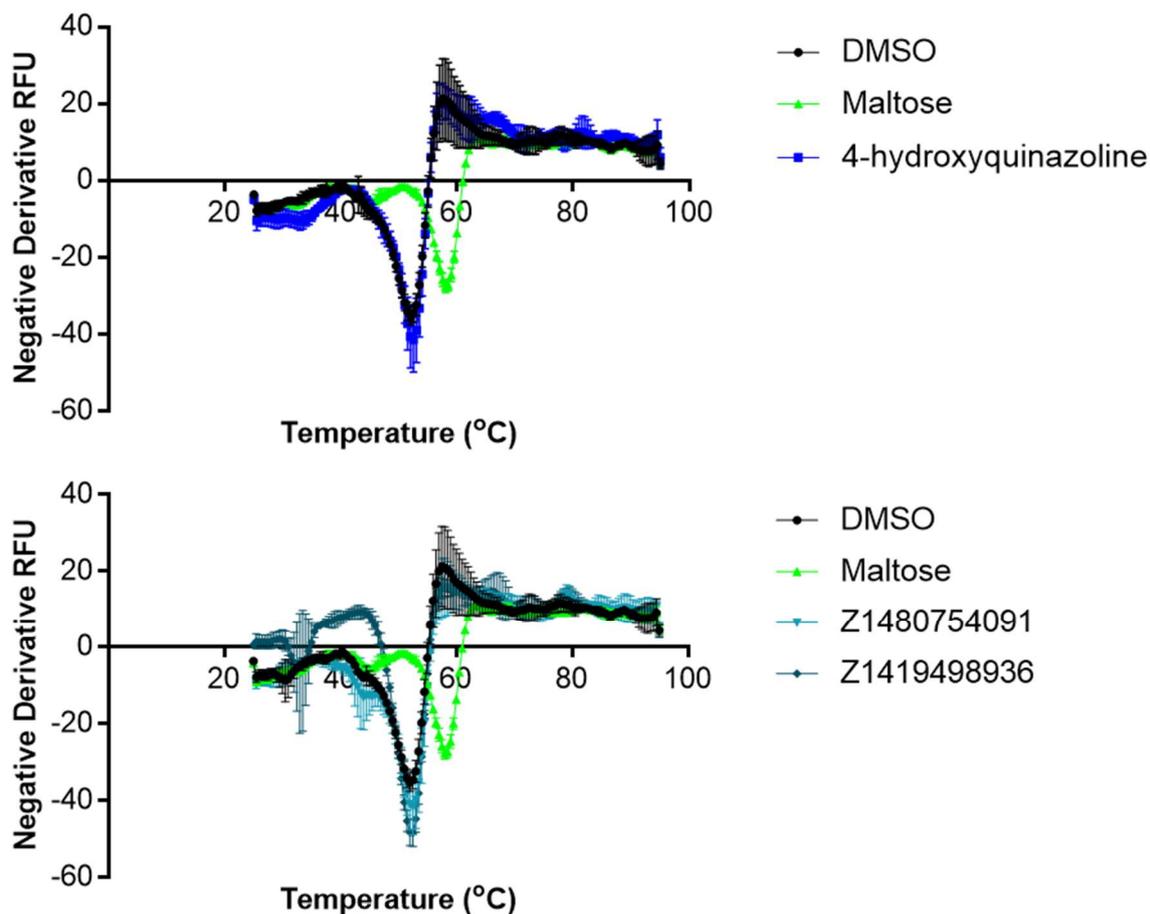


Figure 6.9 Thermal shift assay of compounds from *in silico* docking.

The derivative of relative fluorescence units for thermal shift assays of lead compounds with MBP-NTD fusion protein. Maltose acts as a positive control, whilst DMSO acts as a solvent/negative control. Each data point represents the mean of three technical replicates, error bars depict standard deviation.

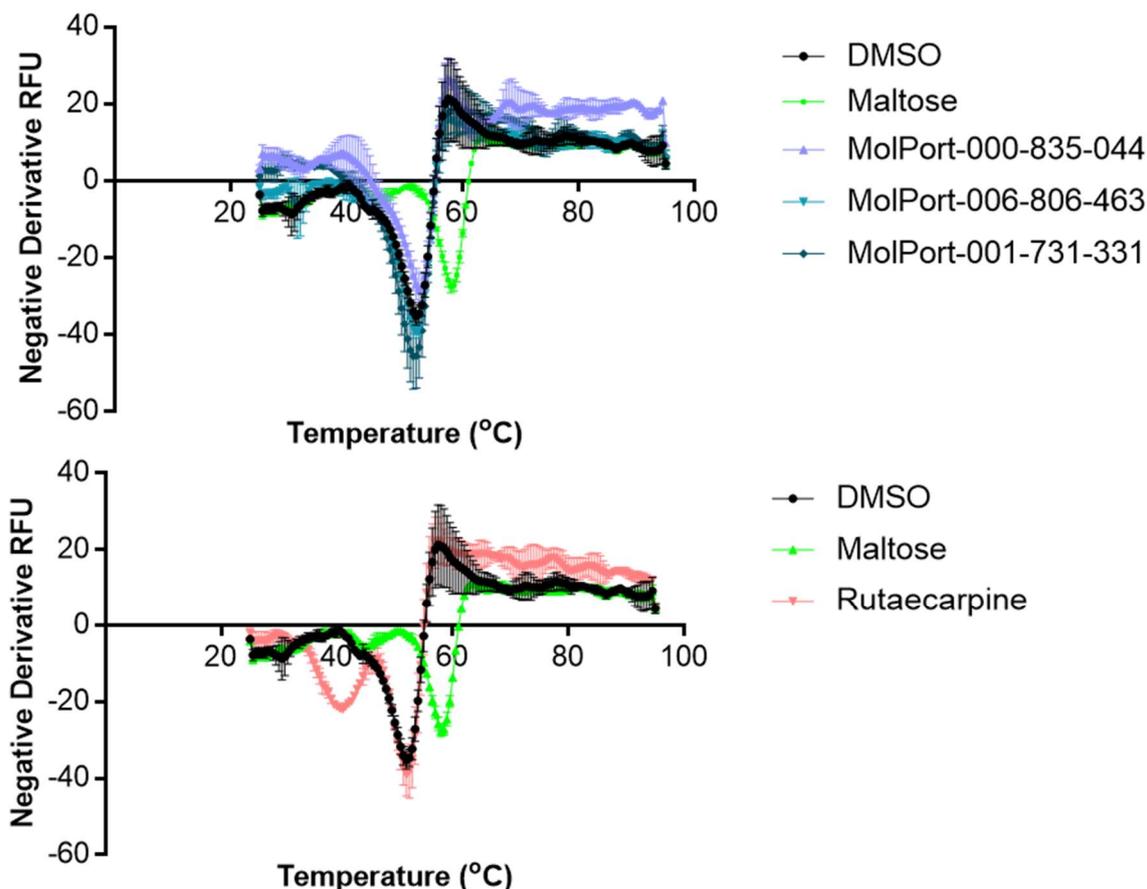


Figure 6.10 Thermal shift assays for compounds from putative PARP inhibitor library.

The derivative of relative fluorescence units for thermal shift assays of lead compounds at 200 μ M with MBP-NTD fusion protein. Maltose acts as a positive control, whilst DMSO acts as a solvent/negative control. Each data point represents the mean of three technical replicates, error bars depict standard deviation.

Rutaecarpine (a member of the putative PARP inhibitor library), Z1756820653, and 4-hydroxyquinazoline all appear to effect *P_{pcrV}*-Lux activity, but not the growth of PAO1. Z1756820653 and 4-hydroxyquinazoline increased luminescence, whilst rutaecarpine decreased it (Figure 6.11). Z195013052 also appeared to increase luciferase activity, whilst none of the other compounds tested had any effect (Figure 6.12). Rutaecarpine, Z1756820653, and 4-hydroxyquinazoline had no effect on *P_{lacZ}*-Lux activity (Figure 6.13). The successful Lux hits were subjected to further *in vivo* test using anti-PcrV Western blotting in EGTA induced PAO1 and the Δ *exsD* “T3SS always on” mutant. All compounds drastically reduced PcrV expression in EGTA induced PAO1 to barely detectable levels, however the effect on the mutant strain was less clear (Figure 6.14). Given the lack of clarity on the

mechanism by which EGTA-dependent Ca^{2+} depletion stimulates ExsA activity the space for inference from this discrepancy is very limited, it could perhaps indicate a more rigorous overexpression of ExsA in the mutant compared to the induced sample. In the ΔexsD mutant a noticeable decrease in PcrV detection is apparent for Z195013052 and rutaecarpine, however their ICD control bands also appear dimmer. 4-hydroxyquinazoline treatment does not appear to affect the quantity of detectable PcrV in the ΔexsD mutant (Figure 6.14).

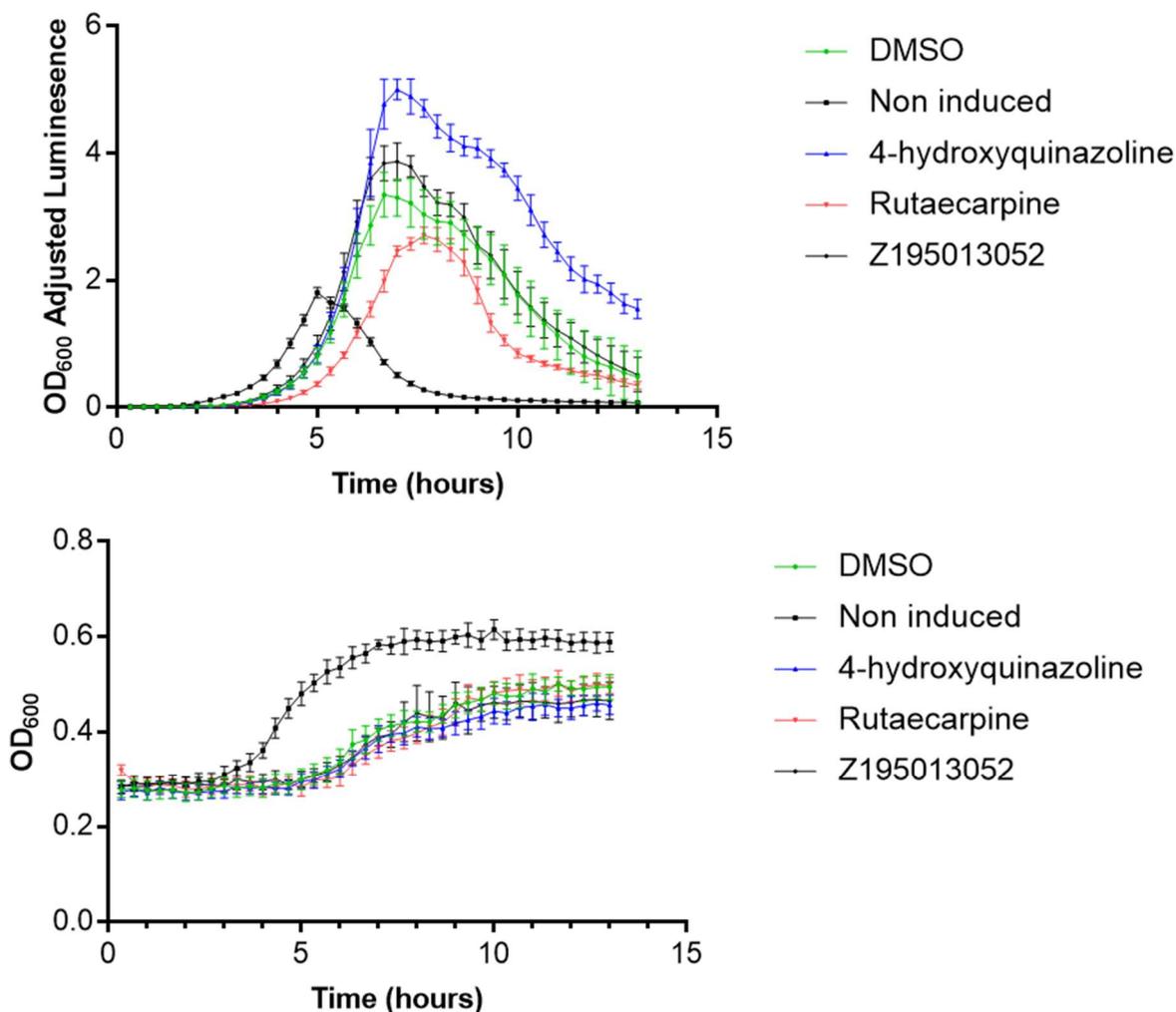


Figure 6.11 Luciferase reporter assays for novel candidate molecules.

OD₆₀₀ adjusted luminescence data (A.) and growth curves (B.) for the *PpcrV*-Lux reporter strain of PAO1 with 200 μ M of the specified compound or an equivalent solvent (DMSO) volume at 37 $^{\circ}$ C with shaking. Data points represent the mean of three biological replicates and error bars depict standard error of the mean. All adjusted luminescence data sets were found to vary from the control in a statistically significant manner ($p < 0.05$) in a 2-way ANOVA and Dunnett's multiple comparison.

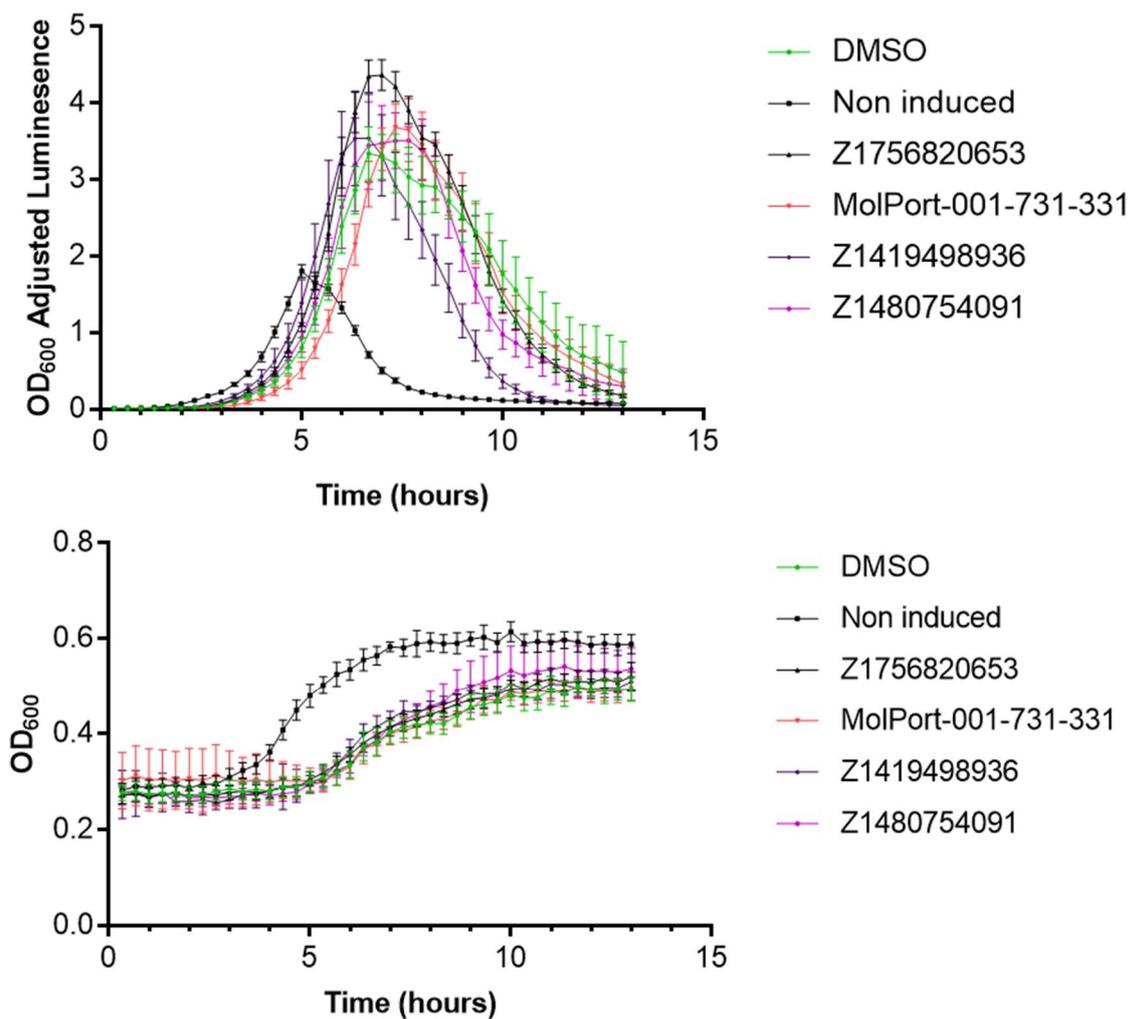


Figure 6.12 Luciferase reporter assays for other screening molecules.

OD₆₀₀ adjusted luminescence data (A.) and growth curves (B.) for the *PpcrV*-Lux reporter strain of PAO1 with 200 μ M of the specified compound or an equivalent solvent (DMSO) volume at 37 °C with shaking. Data points represent the mean of three biological replicates and error bars depict standard error of the mean.

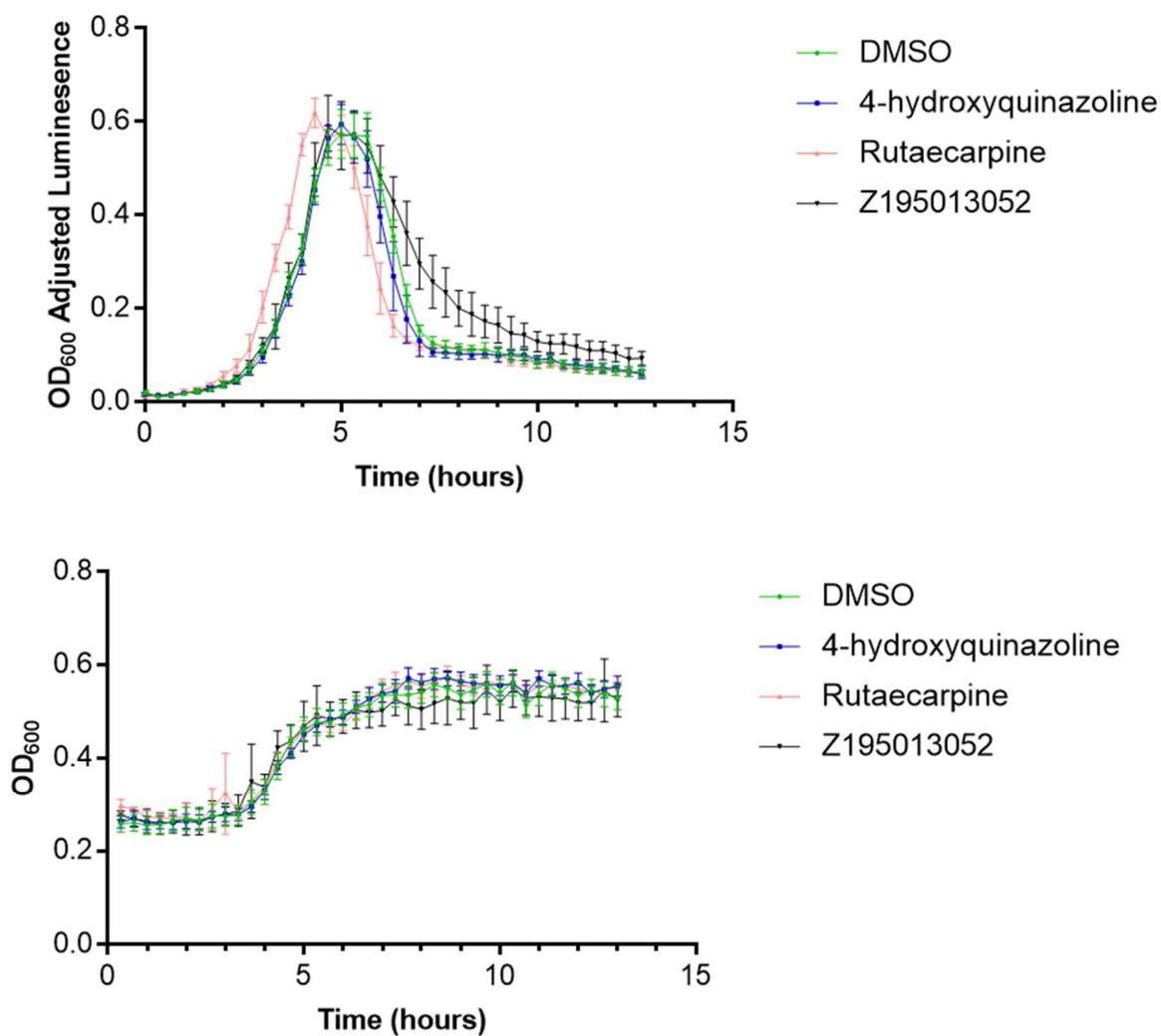


Figure 6.13 Luciferase control assays for hit compounds.

OD₆₀₀ adjusted luminescence data (A.) and growth curves (B.) for the *P/acZ-Lux* reporter strain of PAO1 with 200 μ M of the specified compound or an equivalent solvent (DMSO) volume at 37 °C with shaking. Data points represent the mean of three biological replicates and error bars depict standard error of the mean.

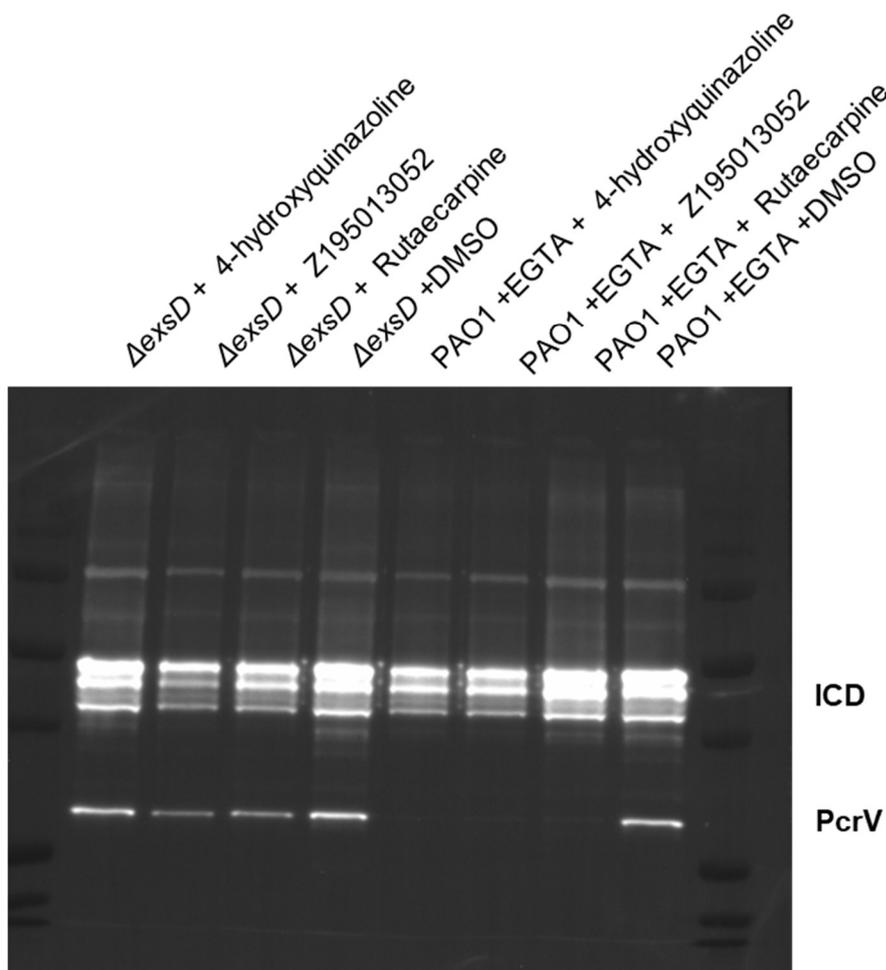


Figure 6.14 PcrV Western blots with lead compounds.

Western blots of AGSY cultures grown to an $OD_{600} = 1$ in the presence of 200 μM of the specified compound or an equivalent volume of solvent, using anti-PcrV and anti-ICD antibodies. A total of 10 μg of protein was loaded into each sample well. The image is representative of two independent biological replicates.

6.7 Discussion

6.7.1 A full length model of ExsA

The utility of a full-length model for docking to the ancestral ligand binding pocket was highlighted by the interactions identified between the CTD and a compound docked to the pocket in Figure 6.8. Given that a full-length structure of ExsA has thus far proven elusive to both my own efforts and that of others the construction of a model was necessary. This work also predates the public availability of the new machine learning approaches to protein structure modelling such as AlphaFold (Jumper *et al.*, 2021). The issue in obtaining full length experimental structures of the AraC family was not restricted to ExsA alone, indeed the titular

protein AraC itself has a number of NTD only structures published (PDB codes: 2ARA, 2ARC, 2AAC). One member of the family, the *E. coli* transcription factor Rob, has a full-length structure which is bound to DNA (PDB code: 1D5Y), however it is a rather atypical AraC protein given that its regulatory domain is the CTD and its DNA binding domain is at the N terminus (Kwon *et al.*, 2000).

Most of the other AraC family structures, including those of ToxT utilised as templates in the present study, are crystallised with a cognate ligand bound in the pocket homologous to the Tyr33 pocket of ExsA. This is not the case for the CuxR structure (PDB: 5NLA), which is not bound to any ligand. There is evidence that CuxR protein is regulated by the action of c-di-GMP, which is thought to bind to two sites on the protein, neither of which is homologous to the ligand binding pocket examined in the present study (Schäper *et al.*, 2017). The more “open” structure observed could therefore be reflective of the entire family when unbound by a ligand, or vary from the structure of ToxT for other reasons. Given the hypothesis that the ligand binding pocket is a site of inhibition on ExsA, the model based on alignment to ToxT (PDB: 4MLO) was selected. If it is the case that CuxR exhibits a more open conformation due to a lack of ligand binding in the homologous region to Tyr33 then the utilisation of the domain positioning of 4MLO allows for a major shift in ExsA’s conformation upon the binding of an inhibitor, which is highly desirable. If the variations between the CuxR and ToxT structures (PDB IDs: 5NLA and 4MLO respectively) are due to a more fundamental biological difference then it was still desirable to use the ToxT aligned model, because CuxR’s regulatory molecule binds to different sites to most of the AraC family’s regulators and have a stimulative rather than inhibitory effect.

The resulting model was suitable for docking, however due to the gaps in the sequence it is not a full model of ExsA, rather it essentially served for a working model for the current investigation. It also facilitated the visualisation of more SNPs identified in section 4.2.3. Of most interest is T200A reported in strain LESlike1, which affects the DNA binding HTH motif. Given the change from a polar to a hydrophobic side change it is conceivable that ExsA’s ability to interact with DNA is likely to be impaired by this substitution. LESlike1 is a clinical isolate from a cystic fibrosis patient’s lung and could well be an example of mutations shutting off the acute virulence machinery during long term chronic infections (Smith *et al.*, 2006). The SNPs do not invalidate the current target site, nor the alternative target site of the DNA binding domain. The fact that mutational inactivation of ExsA occurs does not invalidate ExsA as a therapeutic target for acute infection. Chronic adapted strains would be unable to cause acute infections for which an ExsA inhibitor would be utilised. The compounds generated in this study were therefore unlikely to ever be used against strains such as the LES and LES like

strains which are genetically adapted to chronic virulence.

6.7.2 Small molecule inhibitors of ExsA

The lack of correlation observed in the scoring of the same compound by different methods of docking indicated several things. Firstly, as is probably self-evident, no docking software is perfect and reliance on any one software will result in abundant false positives and false negatives. Secondly, given that all docking software used here have been validated via benchmarking exercises prior to release and used successfully in previous published studies and therefore, providing the structure is suitable, can return true positives – by combining them the success rate should be raised. False negatives from one software could be high scoring in another, and agreement between multiple methods can be seen as reflective of a higher probability of a genuine hit. This observation was in line with the established literature, suggesting consensus scoring method to be more reliable in hit picking from virtual screening outputs (Glaab, 2016).

Given the multitude of compounds containing 4-hydroxyquinazoline were found via this more rigorous docking methodology, and the consistency in its pose against ExsA as either a standalone fragment or part of a larger molecule, such a structural motif offered an interesting potential route towards finding inhibitor against ExsA. If a core fragment or scaffold can be found as initial hit against a structure, the structure activity relationship of the surrounding chemical space can be explored allowing a very good understanding of the pharmacophore to be developed in the absence of structural information. The utilisation of the lab's putative PARP inhibitor library allowed more chemical space to be explored at a negligible cost.

Despite these compounds' promise in *in silico* investigation, it was unfortunate that again a discrepancy between the thermal shift assay and the luciferase assay results was observed, although several of the novel compounds have further *in vivo* validation from Western blotting. This however raised another contradiction. Whilst rutaecarpine lowered *PpcrV* luminescence and *PcrV* expression in the two separate reporting systems, both 4-hydroxyquinazoline and Z195013052 raised luminescence in the reporter system and lowered *PcrV* expression in PAO1. Given that both *PcrV* expression and *Lux* expression within the reporter are mediated by the same promoter region, albeit integrated at different points in the genome, this was unexpected. The mixed result indicated that a change in the screening protocol was warranted, focused more directly on ExsA and its activity using biophysical and functional assays.

The failure to obtain any hits for thermal shift, even for a compound with efficacy shown in two

separate *in vivo* assays indicate that it was providing false negatives and was probably non-functional. It seemed also plausible that the truncated ExsA NTD bound to MBP is not suitable for thermal shift assay based screening. A different assay should therefore be employed, with full length ExsA. Surface plasmon resonance might serve as a potential alternative. If a novel biophysical screening method for ExsA was established, alongside an EMSA as previously discussed, then the validity of the luciferase reporter could be firmly established or dismissed. These three assays, alongside Western blotting, would provide a solid basis for confirming the inhibitory activity of the inhibitors suggested in the present study.

The potential finding that 4-hydroxyquinazoline and several related compounds stimulated ExsA activity was profoundly interesting. It offers an insight into the functional aspects of ExsA's biology, as well as a possible route to identify novel inhibitors. Whilst stimulants of ExsA activity are far from being therapeutically useful, if the finding could be confirmed, especially alongside the activity of putative inhibitors such as rutaecarpine, it offers the chance to build a comprehensive knowledge of the structure activity relationship of these compounds and thus inform further developments.

Rutaecarpine was the best characterised inhibitor in the present study given its consistent effects across both *in vivo* experiments. As well as reducing PcrV expression in EGTA induced PAO1 it seems to reduce PcrV expression in a Δ exsD mutant. The extreme hyperactivity of ExsA in the Δ exsD mutant is presumably responsible for the reduced efficacy of the inhibitors in comparison to EGTA induced PAO1. This is unfortunate as the removal of anti-activator ExsD offers a "cleaner" induction of ExsA than the addition of EGTA. Without a more accurate loading control, such as a total protein method which allows more reliable quantification, effects in this condition cannot be adequately quantified. Of course, further validation of rutaecarpine is required, however it offers an interesting lead compound given its known pharmacological effects.

Rutaecarpine is part of a growing class of compounds which have been isolated as pharmacologically active components of traditional herbal remedies. The compound has been identified as having protective cardiovascular and anti-inflammatory effects in rodent models (Moon *et al.*, 1999; Jia and Hu, 2010; Xu *et al.*, 2017). In acute infection situations such as ventilator associated pneumonia common for *P. aeruginosa*, a combined anti-virulence and anti-inflammatory approach could be beneficial. The potential benefits of anti-inflammatories during pneumonia are well documented, and research on the topic has been intensified around Covid-19 specifically (Meijvis *et al.*, 2012; Stebbing *et al.*, 2020). That rutaecarpine is known to be safe in mammalian models and has already been subjected to animal experimentation

reduces safety concerns about toxicity or potential unknown side effects.

6.8 Conclusion

Several novel ligands for ExsA have been proposed and subjected to a preliminary characterisation in the present chapter. Whilst further work is required both *in vivo* and *in vitro* to confirm the novel inhibitors activity, rutaecarpine offers a promising lead for the future development of ExsA inhibitors.

7. Modelling of full length ExsA using AlphaFold

7.1 Aims and approaches

Given the remarkable advancement of protein structure modelling based on deep learning during the latter stages of the present study, it was decided to comprehensively model ExsA, its dimerization, and its interaction with ExsD using AlphaFold (Jumper *et al.*, 2021). AlphaFold and another comparable approach, RosettaFold (Baek *et al.*, 2021), have revolutionised the structural bioinformatics field since their release. The overall aim was to use this state-of-the-art protein structure modelling approach and to evaluate whether it could lead to better models of full length ExsA as a monomer, a homodimer, and as a heterodimer with its known anti-activator protein ExsD. The MD simulations within this chapter were conducted by Dr Thales Kronenberger (Dept. of Pharmacy, University of Tübingen), who also provided the relevant methods, and Figures 7.14, 7.15, and 7.16 with legends and captions.

7.2 Methods

7.2.1 AlphaFold for protein modelling

Wild type monomeric models were retrieved from the AlphaFold Protein Structure Database (<https://alphafold.ebi.ac.uk/>) (Varadi *et al.*, 2022), all other models were generated using AlphaFold *via* the AlphaFold Colab resource implemented in google cloud, with a AlphaFold variant developed for multimers used where appropriate (AlphaFold.ipynb - Colaboratory; Evans *et al.*, 2021; Jumper *et al.*, 2021). AMBER relaxation was employed on all models generated (Salomon-Ferrer, Case and Walker, 2013). All alignment and visualisation were conducted in PyMOL, for ExsA models only the NTD was utilised for alignment due to the variation in CTD positioning.

7.2.2 Prediction of disorder with IUPred2A

Disordered regions of proteins were detected using IUPred2A (Mészáros, Erdős and Dosztányi, 2018). The setting “IUPred2 short disorder” was used due to all the proteins in question having structured domains and only the linkages between offering potential for disorder, in line with the authors recommendations (Erdős and Dosztányi, 2020).

7.2.3 Prediction of conservation with ConSurf

A prediction of the how conserved each residue is within ExsA and ExsD was undertaken with ConSurf (<https://consurf.tau.ac.il>) (Ashkenazy *et al.*, 2016), using the monomeric AlphaFold model, and the trimeric ExsD crystal structure (PDB: 3FD9) respectively. ExsD elicited the warning: 87 of 256 residues have unreliable conservation scores due to insufficient data in the multiple sequence alignment, whilst ExsA was successfully analysed.

7.2.4 Virtual alanine interface scanning

To scan proposed dimer interfaces, the BAlaS tool was used, generating predicted change in Gibbs free energy change upon protein-protein binding ($\Delta\Delta G$) (Wood *et al.*, 2020).

7.2.5 Model generation with ClusPro

Models of the ExsA homodimer and the ExsA-ExsD heterodimer were generated with the ClusPro protein-protein docking server (ClusPro 2.0: Protein-Protein Docking, n.d.; Kozakov, Beglov, et al., n.d.; Kozakov, Hall, et al., n.d.; Vajda et al., n.d.; Yueh et al., n.d.). The AlphaFold full length monomeric models of ExsA and ExsD, or monomeric units extracted from dimeric structures were used as indicated.

7.2.6 Molecular dynamics simulations

For all structures protonation states of amino acids were optimized with PROPKA (Schrödinger, LLC, New York, NY, 2021), where we selected the most likely ionization state as proposed by the software, and the structures were minimized.

For each system, namely monomeric with (Md) and without DNA (M) and dimeric without DNA (D) simulations were generated. MD simulations were carried out using Desmond (Dror *et al.*, 2010), with the OPLS4 force-field (Lu *et al.*, 2021). The simulated system encompassed the protein-ligand complexes, a predefined water model (TIP3P (Jorgensen *et al.*, 1998)) as a solvent and counterions (Na^+ or Cl^- adjusted to neutralize the overall system charge). The system was treated in a cubic box with periodic boundary conditions specifying the shape and the size of the box as 13 Å distance from the box edges to any atom of the protein. We used a time step of 1 fs, the short-range coulombic interactions were treated using a cut-off value of 9.0 Å using the short-range method, while the smooth Particle Mesh Ewald method handled long-range coulombic interactions (Darden, York and Pedersen, 1998).

Initially, the relaxation of the system was performed using Steepest Descent and the limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithms in a hybrid manner. The simulation was performed under the NPT ensemble for 5 ns implementing the Berendsen thermostat and barostat methods. A constant temperature of 310 K was kept throughout the simulation using the Nose-Hoover thermostat algorithm and Martyna-Tobias-Klein Barostat algorithm to maintain 1 atmosphere of pressure, respectively. After minimization and relaxation of the system, we continued with the production step of at least 1 μs , depending on the system, with sampling every 1,000 ps. For each ligand five independent replicas were run totalling around 12.5 μs for the M_{DNA} , 9.5 μs for M without DNA and 3 μs for dimeric. Trajectories and interaction data are available on Zenodo repository (under the codes: 10.5281/zenodo.6704185, which will not be available prior publication).

Protein secondary structure elements were monitored over the course of the simulation using Maestro tool. Angle and distance calculation were performed employing Maestro event analysis tool (Schrödinger, LLC, New York, NY). Distances between specific secondary structure elements were calculated using their centres of mass, using the script *trj_asl_distance.py*. All the protein-protein measurements were performed using the script *trj_asl_distance.py* having as an argument the atom numbers of the atoms involved in the interaction.

7.3 AlphaFold predicts an open model of monomeric ExsA

The AlphaFold model of monomeric ExsA is, overall, very similar to that constructed by other means in the present study. It retains a NTD structure congruent with that observed in the ExsA NTD crystal structure (PDB: 4ZUA), and a CTD structure like that modelled by homology to the AraC protein from *C. violaceum* (PDB: 3OIO). A key distinction is that the CTD is withdrawn further back from the NTD than in the model selected in Chapter 6, whilst still bearing more similarity to that model, which was created by alignment to the ToxT crystal structure (PDB: 4MLO), rather than the CuxR structure (PDB: 5NLA, Figure 7.1). The same change is not observed from the crystal structure of ToxT and the AlphaFold model of ToxT which are highly congruent except for the carboxy terminus, which is structured into an alpha helix at the terminal residues in the crystal structure but forms an unstructured loop in the model (Figure 7.1 C).

Within the AlphaFold model of ExsA, there are several regions of lower certainty, as indicated by the pLDDT scores for the residues in that part of the sequence. Most notably both the extreme N and C terminal sequences are highly uncertain, and to a lesser extent so are the unstructured linker between the CTD and NTD and the linker between $\beta 8$ and $\alpha 1$ which connects the beta barrel regulatory domain and the dimerization domain within the NTD (Figure 7.1 D, Figure 7.2 A). Each of these areas were also detected as being liable to disorder by IUPred2A, though only the extreme terminal sequences met the score threshold of 0.5 to be classed as a disordered region (Figure 7.2 B). Finally, a peak in the ANCHOR score indicating disordered binding regions was observed around residue Pro134, which corresponds to the unstructured linker between the helices $\alpha 2$ and $\alpha 3$ of the dimerization domain within the NTD, and at both terminal ends of the protein (Figure 7.2 B).

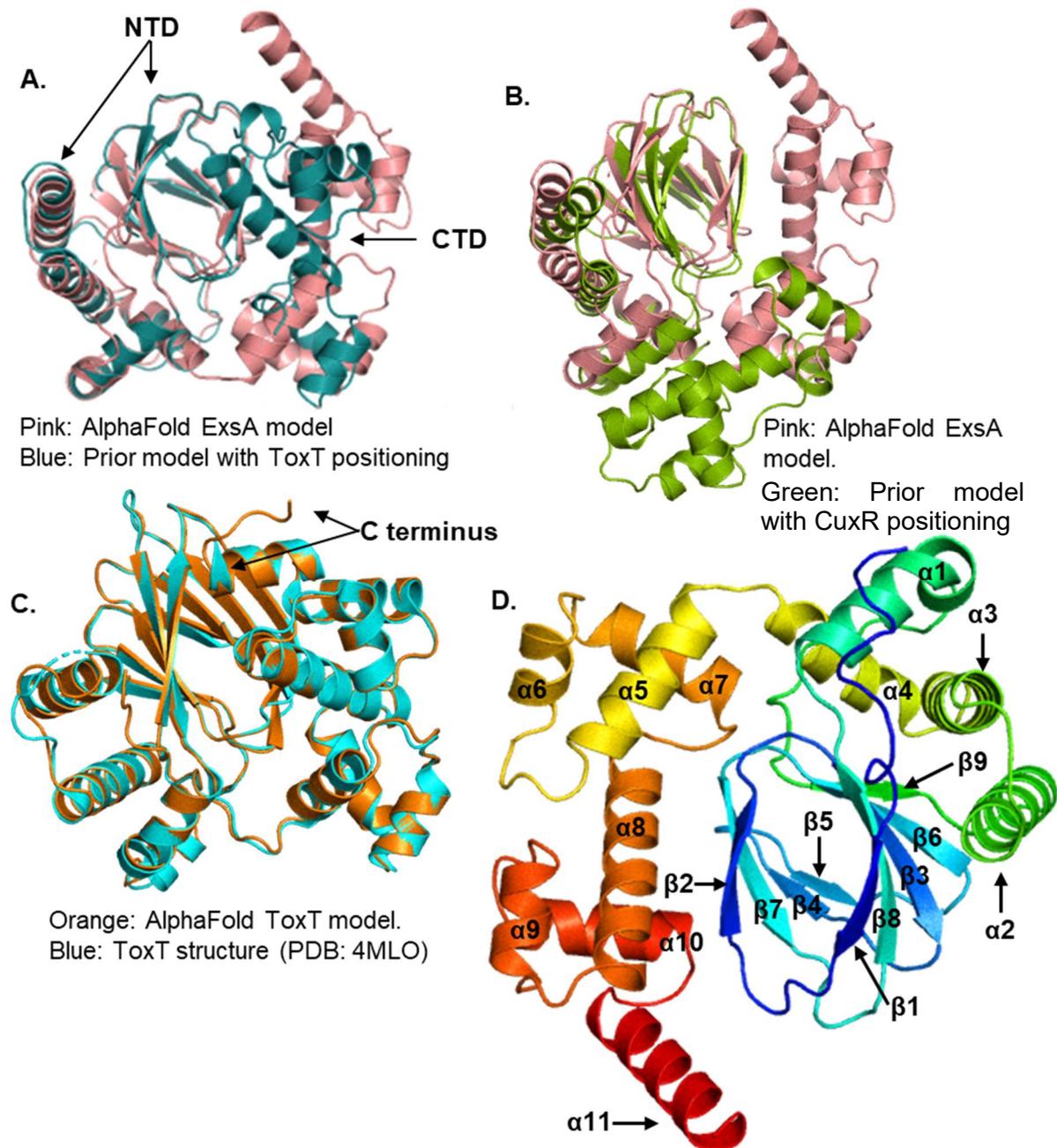


Figure 7.1 AlphaFold models of ExsA and ToxT compared to crystal structures and prior homology modelling.

A&B. Cartoon depictions of the ExsA AlphaFold monomer model (pink) aligned to the models created previously in this study, created by aligning the ExsA NTD structure (PDB: 4ZUA) and a model of the ExsA CTD to crystal structures of ToxT (blue, **A.**) or the 5NLA structure of CuxR (green, **B.**). Alignment was performed with the NTD alone.

C. The ToxT AlphaFold model (orange) and a ToxT (PDB: 4MLO, blue) structure aligned (RMSD=0.37 Å). Dotted sections indicate inadequate electron density for interpretation.

D. The naming system for alpha helices and beta sheets within ExsA illustrated on a spectrum coloured monomeric ExsA AlphaFold model. The numbering begins from the N terminus.

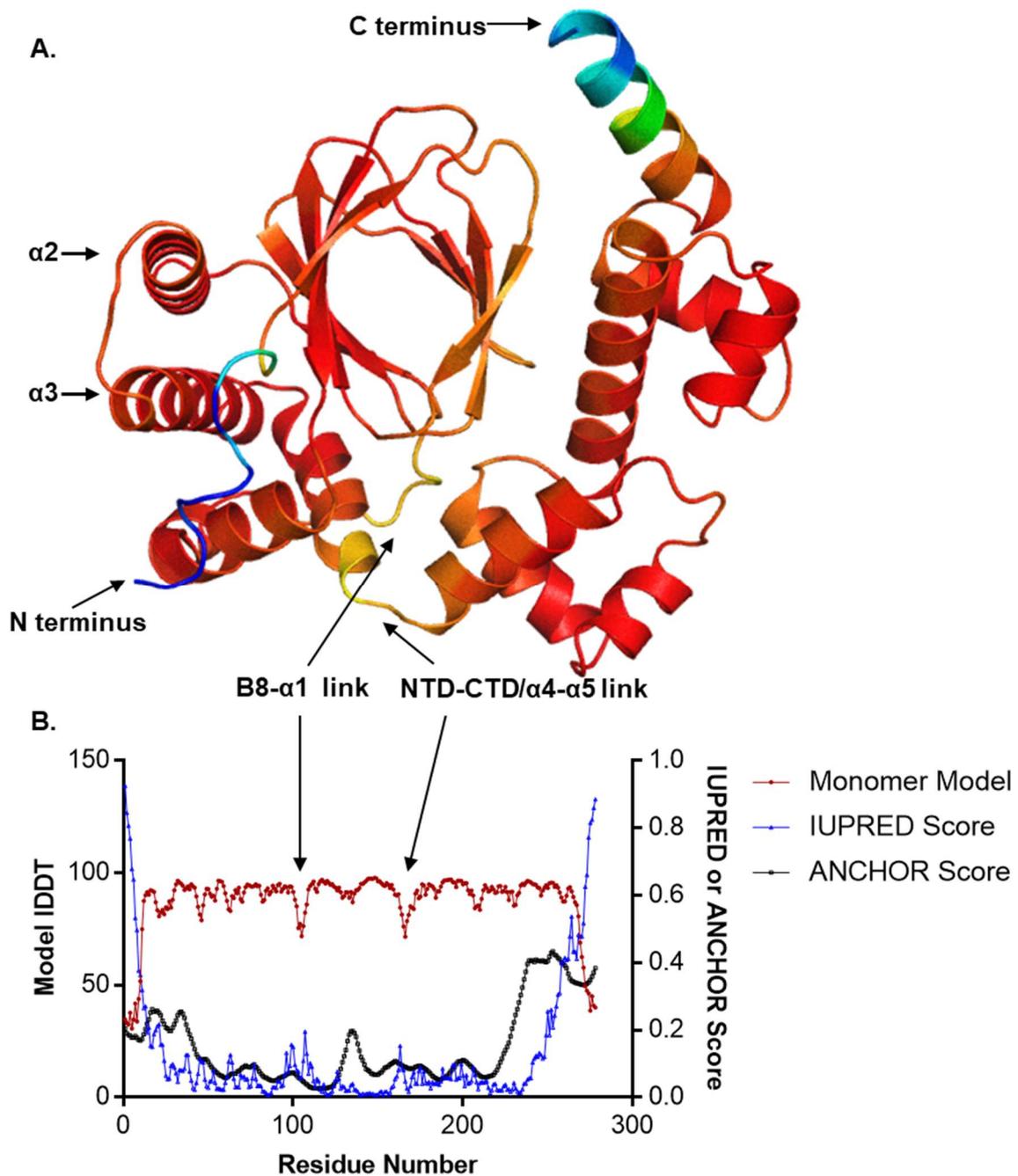


Figure 7.2 IDDT and predicted disorder in the AlphaFold model of ExsA.

A. A cartoon depiction of the ExsA AlphaFold model coloured by IDDT score for each residue. Red indicates a very high score and blue a very low score, with a spectrum between.

B. AlphaFold IDDT score plotted for each residue, above the corresponding IUPRED and ANCHOR scores, which are predictive of disorder and stabilisation of disordered regions during binding respectively, generated from the ExsA protein sequence.

An *in silico* method (ConSurf) for determining the degree of evolutionary conservation of each residue was also undertaken. The CTD is highly conserved, whilst the NTD contains some highly conserved residues, especially on the dimerization helices, it is generally less conserved (Figure 7.3).

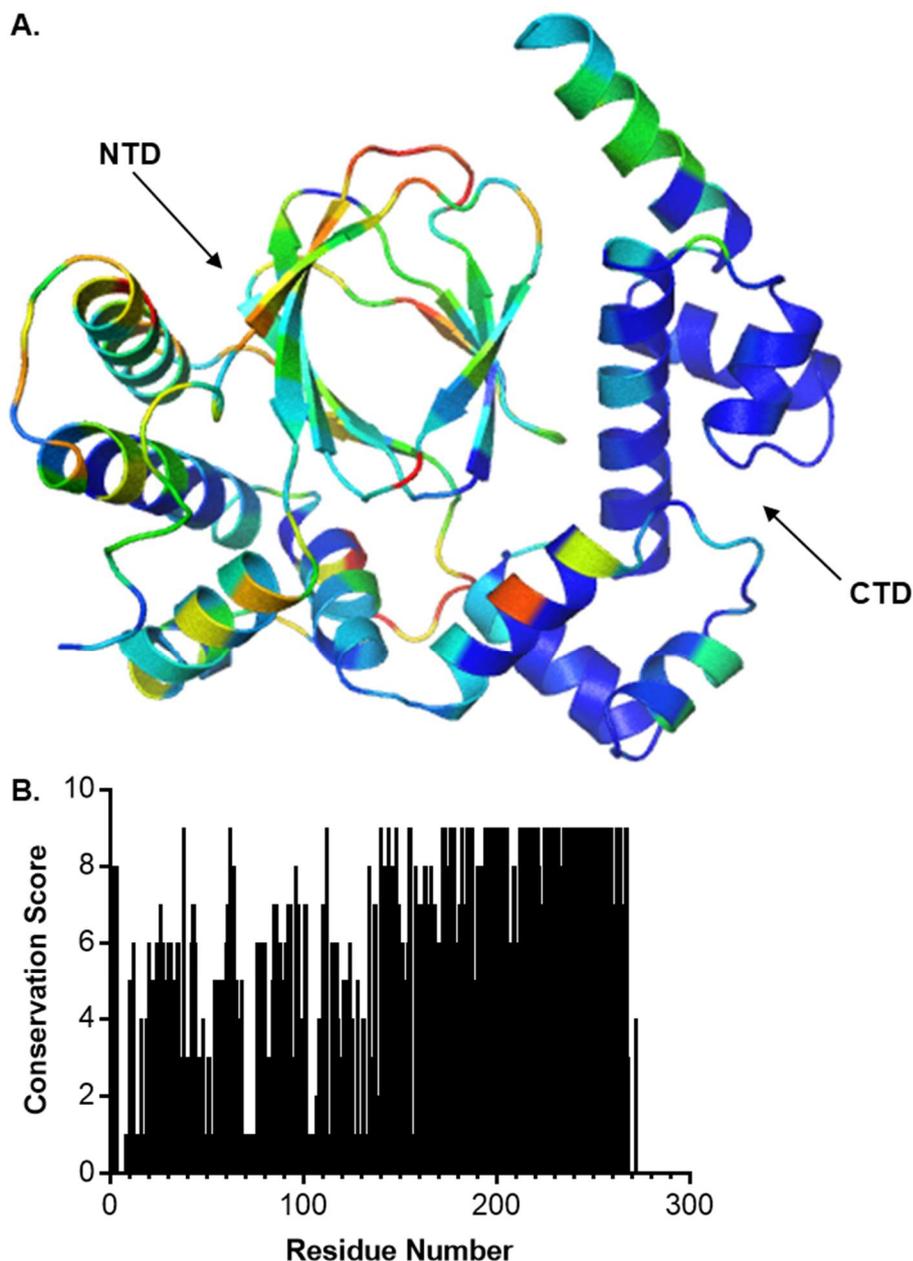


Figure 7.3 Residue conservation in ExsA.

A. A cartoon depiction of the AlphaFold monomeric ExsA structure coloured by conservation score as predicted by ConSurf. Dark blue indicates the maximum score of 9, orange indicates the minimum score of 1, with a spectrum between.

B. A graphic summary of conservation score across ExsA.

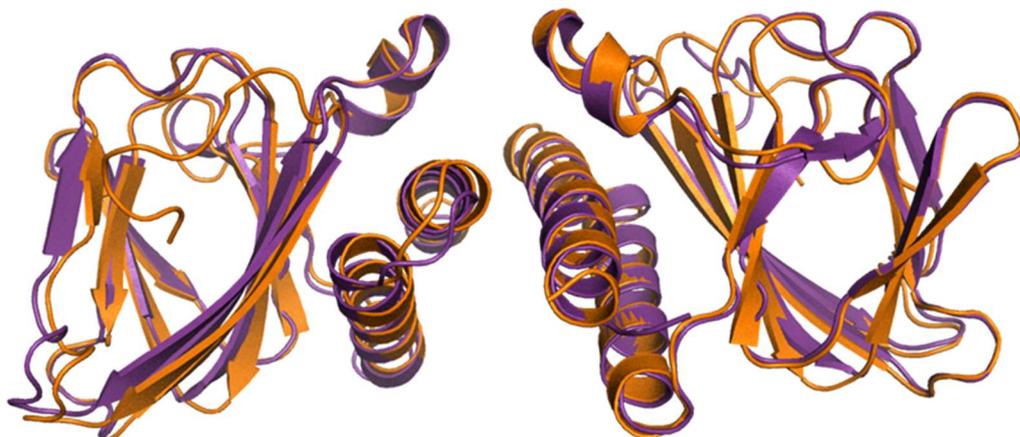
7.4 AlphaFold successfully recreates published dimeric AraC family structures

Three AraC family proteins: AraC, BgaR, and RhaR, which have published dimeric structures homologous to the ExsA NTD (PDB IDs: 1XJA, 6NX3, and 5U9E respectively), were modelled using AlphaFold multimer with the same amino acid sequence as in the published structure. This process was undertaken as a benchmarking exercise to validate AlphaFold's utility at modelling AraC family homodimers. The overall structure and dimerization interface bore a strong resemblance for AraC and BgaR, with all atom RMSDs of 0.69 Å and 0.65 Å respectively for their alignments (Figure 7.4). The recreation of the RhaR structure was not as successful. AlphaFold placed equivalent of ExsA's $\alpha 2$ and $\alpha 3$, the dimerization helices, at the C terminal of each monomer against the N terminal beta barrel, effectively tying the two monomers together, whilst the experimental structure retains cohesive monomers like the rest of the family (Figure 7.5 A&B). Other than this error the alignment showed accurate positioning by AlphaFold, with an all atom RMSD of 2.1 Å which reduced to 0.62 Å when only the helices from Met119 to the carboxy terminus on all dimerization domains were aligned.

To investigate further, a full-length model was sought. However, the corresponding PDB entry 5U9E is labelled as coming from the organism "Bacteria Latreille et al. 1825", in what appears to be an error given that it is listed as a genus of walking sticks or stick insects (Taxonomy browser (Bacteria Latreille et al. 1825)), and the protein is named as RhaR, an AraC family protein found in *E. coli*. Given that the associated paper has never been published, despite the structure being deposited in 2016, and the sequence of the structure differing from the reference sequence for *E. coli* RhaR by one amino acid (a substitution of Tyr30 in *E. coli* for valine in 5U9E), it seems certain that the structure is of a truncated *E. coli* RhaR, with the two residues at the N terminus and the CTD removed. The full-length *E. coli* RhaR monomeric model was therefore retrieved from the AlphaFold database, and a full-length dimeric model was generated, along with a monomeric model using the sequence as seen in the PDB structure 5U9E.

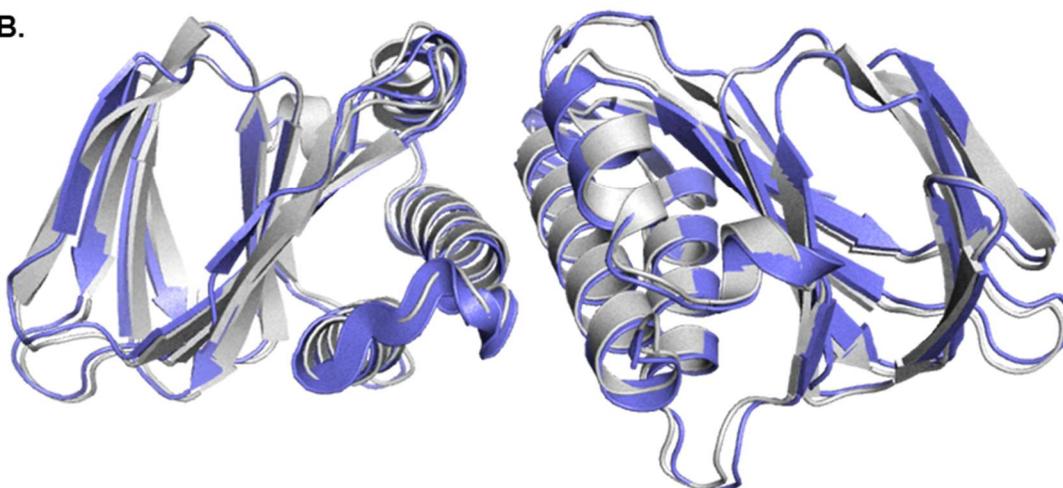
The full-length dimeric model contains the same deviation as when the truncated sequence was used (Figure 7.5 C), however both the full length and truncated monomeric models recreated a single unit of the crystal structure with RMSDs of 0.46 Å and 0.47 Å respectively (Figure 7.6).

A.



Orange: AlphaFold AraC model. Purple: AraC structure (PDB: 1XJA). RMSD = 0.69 Å

B.



Grey: AlphaFold BgaR model. Purple: BgaR structure (PDB: 6NX3). RMSD = 0.65 Å

Figure 7.4 AlphaFold recreations of 1XJA and 6NX3 dimeric structures.

The amino acid sequences of X-ray crystallography structures of published AraC family dimers were used in AlphaFold multimer to recreate the structures. The results were aligned to, and visually compared with, the respective published structure, as shown in the above cartoon representation. **A.** 4XJA (purple) and the respective model (orange), with an all atom RMSD of 0.69 Å. **B.** 6NX3 (blue) and the respective models (grey), with a RMSD of 0.65 Å.

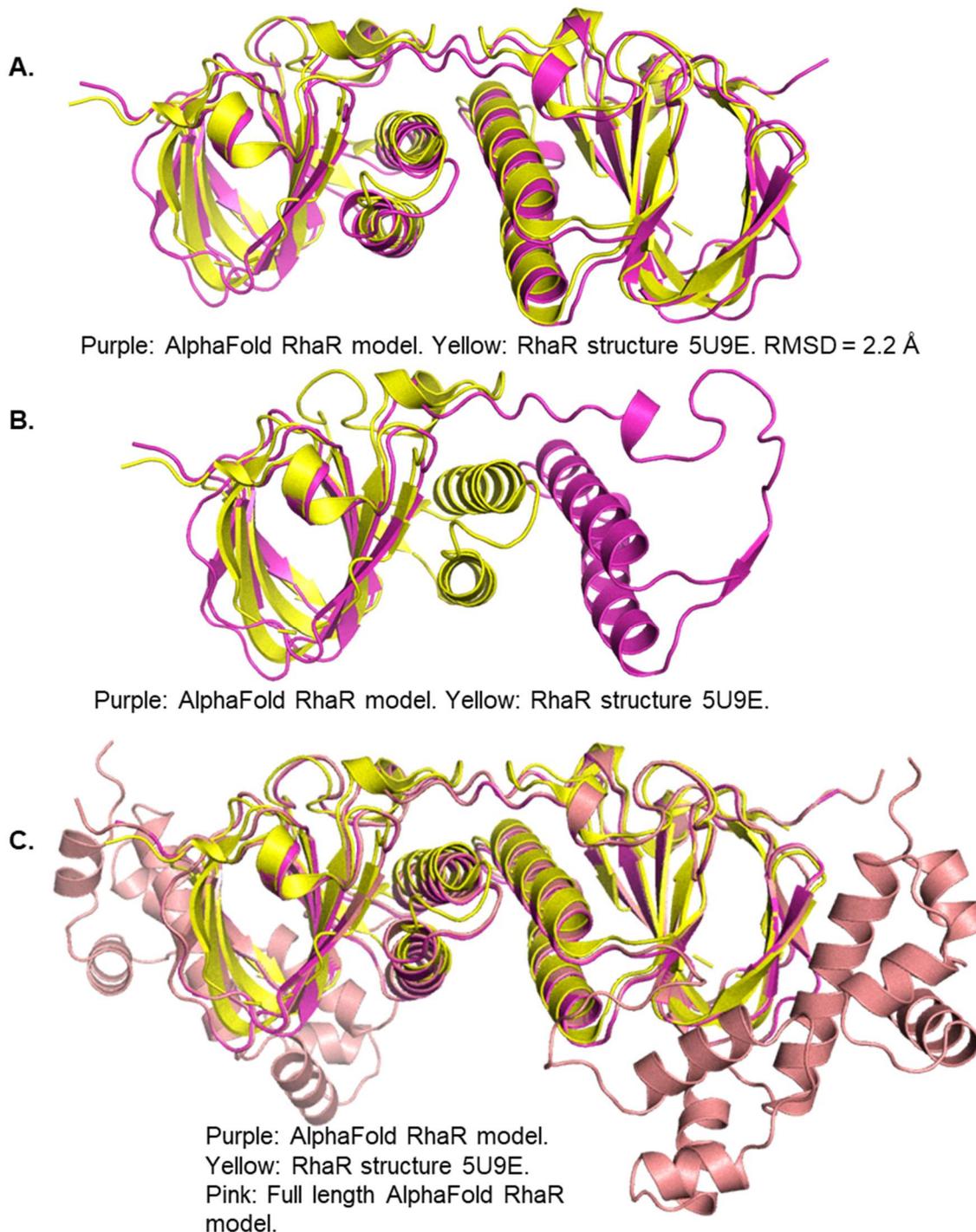


Figure 7.5 Attempts at AlphaFold recreations of 5U9E.

The amino acid sequence of the X-ray crystallography structure 5U9E was used in AlphaFold multimer to recreate the structure. The result (purple) was aligned to and visually compared with 5U9E (yellow) with an RMSD of 2.2 Å. The result can be seen in dimeric form in **A.** and one monomer of each dimer is shown in **B.** Due to the discrepancy between the two, a full length model was generated using the amino acid sequence of *E. coli* RhaR in Alpha fold, and is shown (in pink) alongside the 5U9E and matching truncated sequence dimeric model in **C.** The two models overlap to such an extent that at points only one of them is visible.

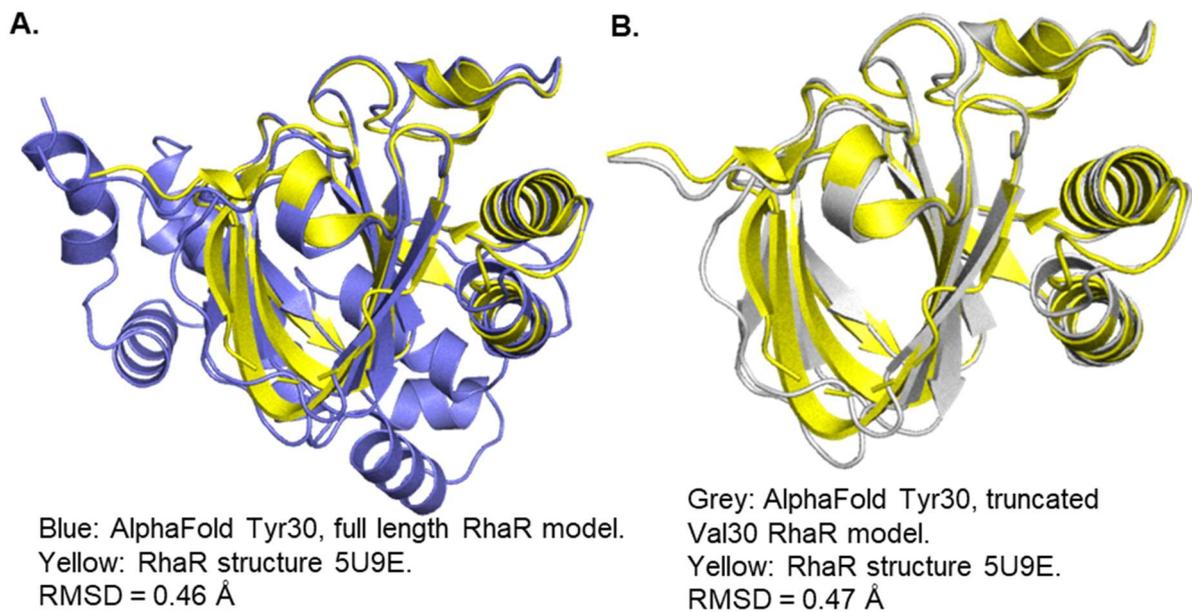


Figure 7.6 Monomeric AlphaFold variants of 5U9E.

Cartoon depictions of monomeric AlphaFold models aligned to a single unit extracted from 5U9E. **A.** 5U9E (yellow) aligned to a monomeric AlphaFold model using the full-length *E. coli* RhaR amino acid sequence (blue) with an RMSD of 0.46 Å. **B.** 5U9E (yellow) aligned to a monomeric AlphaFold model using the same amino acid sequence as 5U9E (grey) with a RMSD of 0.47 Å.

7.5 AlphaFold predicts two different conformations of ExsA dimer

AlphaFold's predictions for the dimeric structure of ExsA fell into two categories: models 2-5 are broadly a dimeric variant of the monomeric model, whilst model 1 had a dramatic conformational change. In model one the end portion of the flexible interdomain linker (from residue Pro156) was structured into $\alpha 5$, the first alpha helix of the CTD, causing it to be projected at a right angle from the NTD (Figure 7.7 A-C). This resolved one of the regions of lesser confidence in the IDDT score, which had a major drop in both the monomeric model and all other dimer models at this point (Figure 7.7 D, Figure 7.2 B). However, having a structured helix to this region conflicted with the elevated IUPRED score assigned to it (Figure 7.2 B). Closer examination of models 1 and 2 demonstrated that the IDDT is overall highly similar, model 2 scored lower in the terminal regions, and more distinctly at the linker between $\alpha 4$ and $\alpha 5$ around Pro156 (Figure 7.8).

A clear distinction between the monomeric model and the dimeric models was observed in the structure of the extreme C terminus. In the monomeric model it was structured into a helix ($\alpha 11$), whilst it is an unstructured loop in all dimeric models (Figure 7.2 A, Figure 7.8 A).

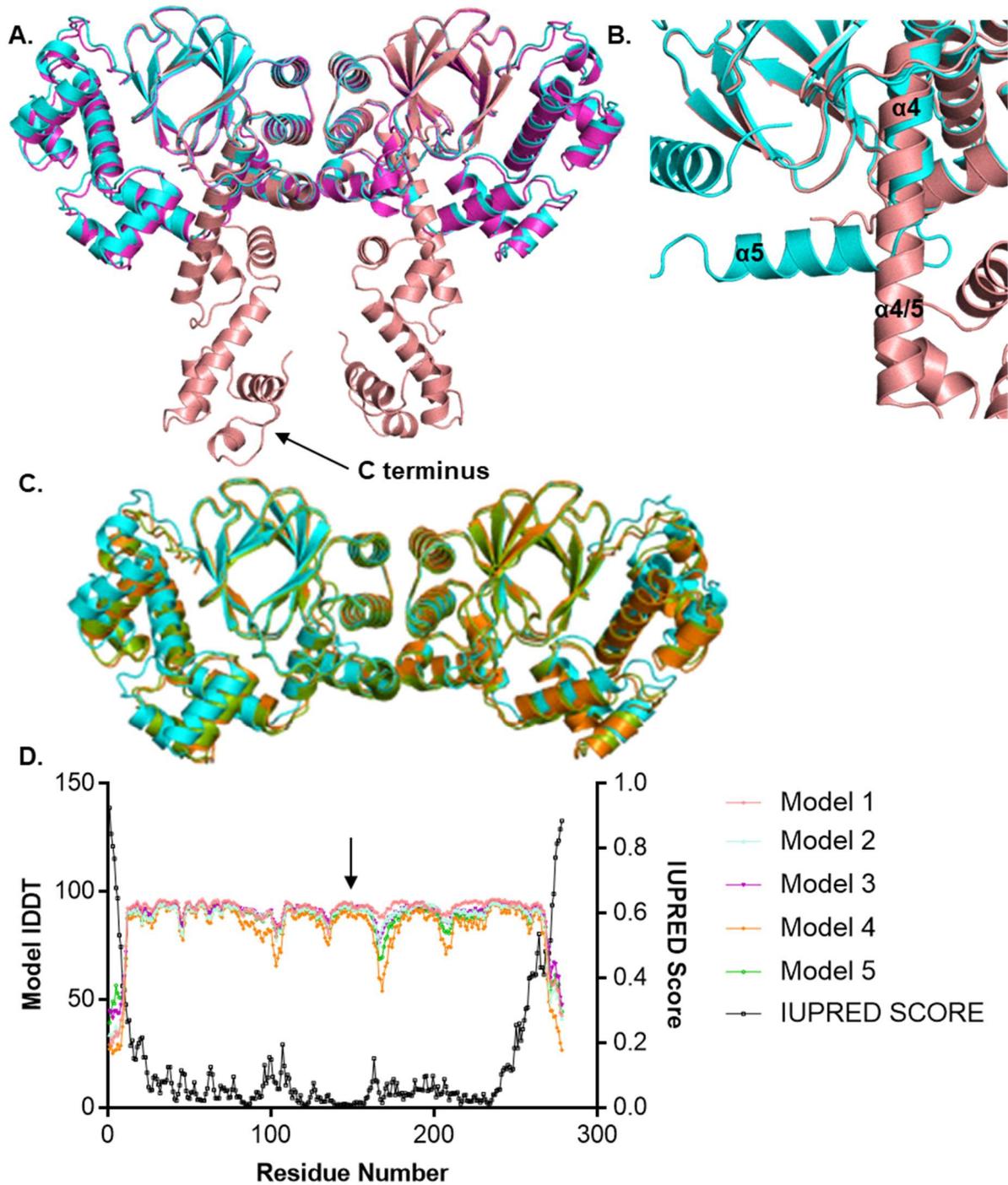


Figure 7.7 The varied AlphaFold models of the ExsA dimer and their IDDT scores.

A-C. Cartoon depictions of the ExsA dimeric AlphaFold models. By scored rank they are model 1 (pink), model 2 (blue), model 3 (purple), model 4 (orange), and model 5 (green).

B. shows they key region around Pro156 which differs between model 1 and the other models due to the fusion of $\alpha 4$ and $\alpha 5$.

D. AlphaFold IDDT score plotted for each residue, above the corresponding IUPRED scores generated from the ExsA amino acid sequence. The arrow indicates the area surrounding Pro156 and the flexible interdomain linker.

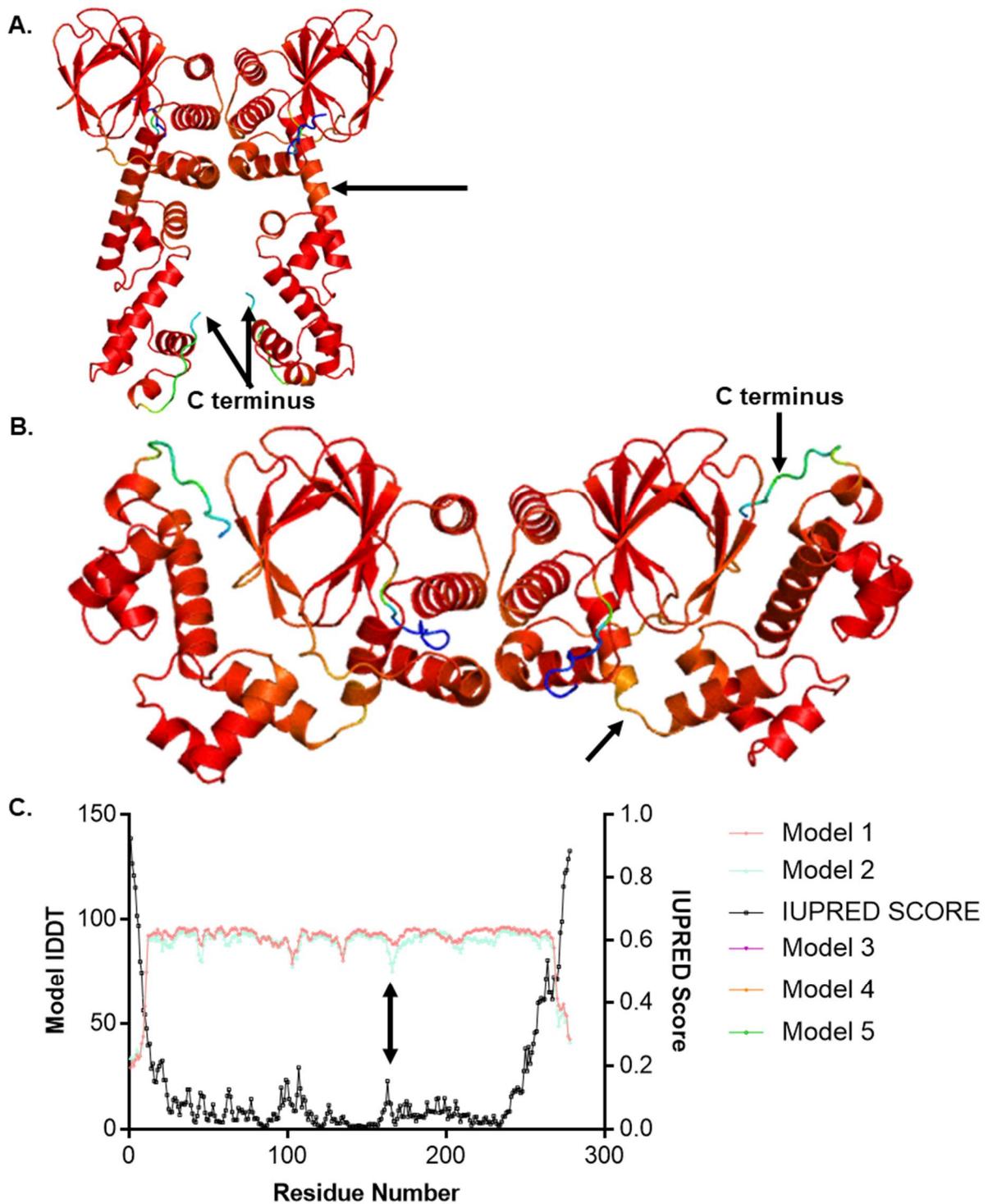


Figure 7.8 AlphaFold ExsA dimer models 1 and 2 compared by IDDT.

Cartoon depictions of AlphaFold ExsA dimer models 1 and 2 are presented in **A.** and **B.** respectively. The models are coloured by IDDT score for each residue. Red indicates a very high score and blue a very low score, with a spectrum between. In **C.** the two models IDDT scores are shown alongside the corresponding IUPRED scores for ExsA. Arrows indicate the area around Ser166 which is part of the interdomain linker in model 2 and part of a CTD alpha helix in model 1 throughout.

7.6 AlphaFold model 2 and ClusPro model 1 of the ExsA homodimer are highly similar

An alternate methodology of *in silico* dimerization was employed using the ClusPro rigid body protein-protein docking system. Two AlphaFold ExsA monomer models were docked to each other; the highest ranked model is in close agreement with the AlphaFold model 2, aligning with an RMSD of 0.49 Å (Figure 7.9 A). ClusPro recreated model 1 when monomers extracted from the model 1 dimeric structure were used as the input (Figure 7.9 B), however this model was ranked eighth, with 23 members of its cluster; the higher ranked models bore no resemblance to any AlphaFold model. The highest scoring cluster, which bore no resemblance to any structure mentioned prior, had 54 members.

7.7 The dimerization interface is broadly conserved between models 1 and 2

The variation in the relative positioning of the NTD and CTD in AlphaFold models one and two has no effect on the position of dimerization helices $\alpha 2$ and $\alpha 3$; the ClusPro model is highly similar also, within the restraints imposed by rigid body docking (Figure 7.9). The overall dimerization interface consists of an antiparallel four-helix bundle, with self-association between the $\alpha 2$ and $\alpha 3$ helices on each monomer. The helix nearer the carboxy terminus end of the sequence bears three leucine residues projecting into the interface. Whilst $\alpha 2$ has few sidechains in the interface itself, it does contain Leu117 which appears to form a leucine triad with Leu148 from $\alpha 3$ of the same monomer, and Leu137 of the $\alpha 3$ of the other ExsA monomer (Figure 7.10).

The importance of these residues is supported by numerical assessment in the form of virtual alanine scanning of AlphaFold models 1 and 2. Each of these three Leu residues received a high score consistently across both chains of both models. The data also validated the greater importance of $\alpha 3$, with higher scores for $\alpha 3$ than $\alpha 2$ in general (Figure 7.11 C). An interesting peak of high scoring residues around residue 92 also occurred (Figure 7.11 C), corresponding to $\alpha 1$ which is adjacent to $\alpha 3$.

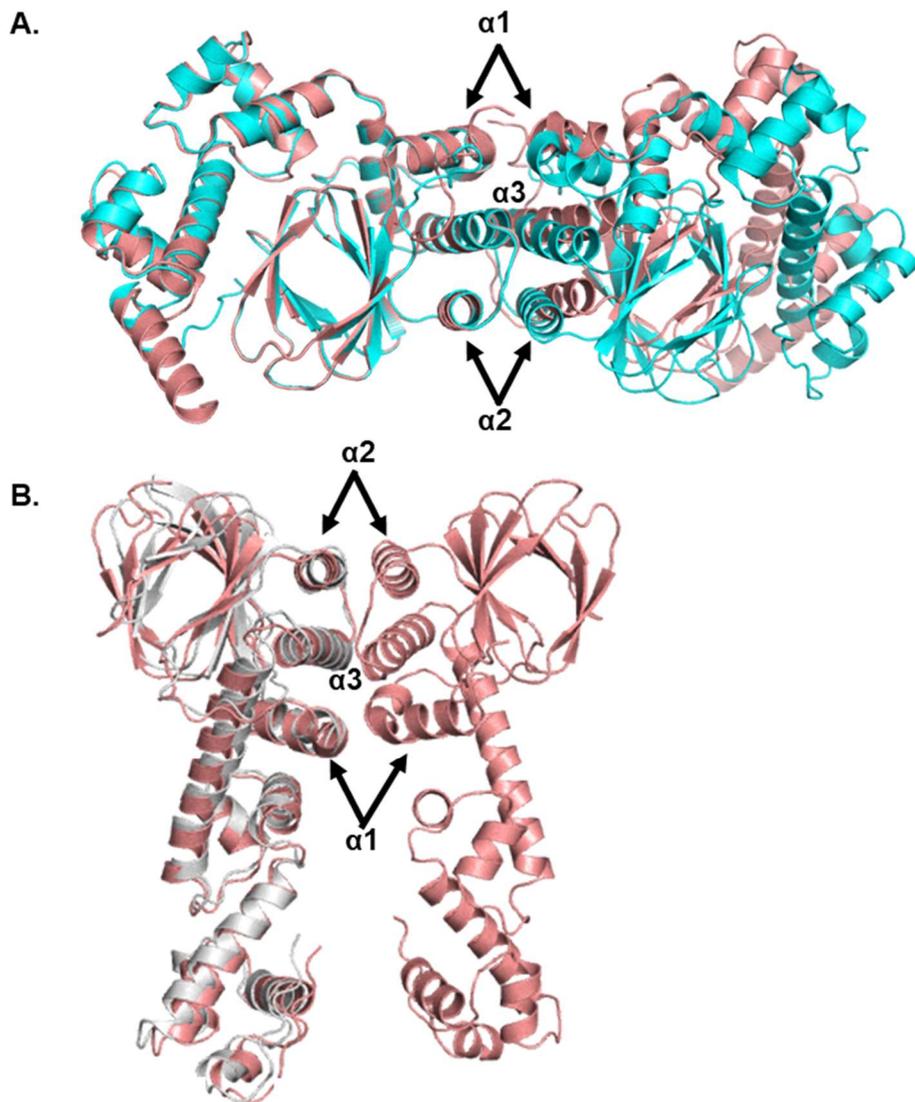


Figure 7.9 ClusPro recreations of AlphaFold models.

A. The ClusPro model of the ExsA dimer, generated from the full length AlphaFold ExsA, model with the highest cluster score (of 122) is depicted in pink, aligned to the AlphaFold model 2 of dimeric ExsA which is depicted in blue with an RMSD of 0.489 Å.

B. The eighth ranked ClusPro model (silver) constructed from the two ExsA units extracted from AlphaFold model 1 aligned to AlphaFold model 1 (pink). Note that the receptor unit from the ClusPro model is not shown as it remains in perfect alignment with the corresponding part of model 1. The ClusPro model shown was ranked eighth, with a cluster score of 23, the higher ranked models bore no resemblance to any AlphaFold model.

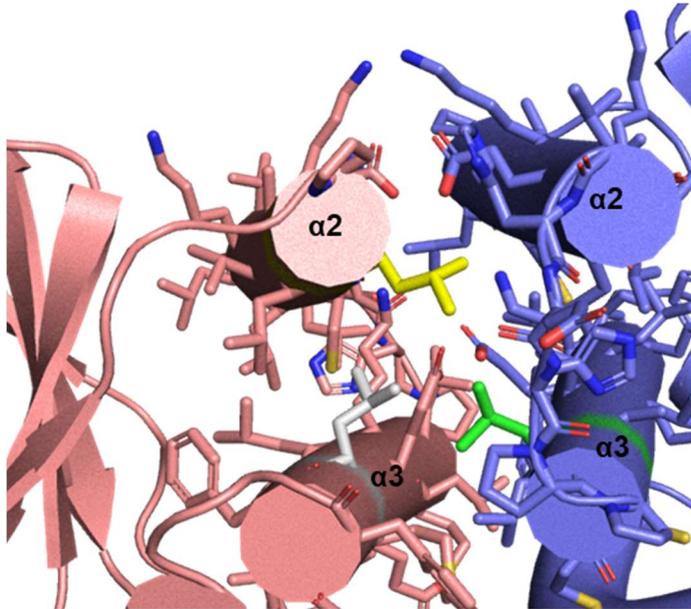


Figure 7.10 ExsA dimeric interface for model 2.

The interface between the two ExsA proteins within AlphaFold ExsA dimer model 2. A stick representation of every residue is overlaid with the cartoon depiction. One monomer is coloured pink, the other is blue. Leu117, Leu148, and Leu137 are coloured yellow, grey, and green respectively.

7.8 Alanine interface scanning highlights discrepancies surrounding Met136 and Glu144 between AlphaFold models 1 and 2

The virtual alanine scanning also predicted several discrepancies between the two models. Model 1 had three residues scored within the CTD, which are not scored in model 2 due to the distance between the CTDs. Conversely model 2 has two scoring residues at the extreme end of the NTD, within the disordered terminal region. Given that these residues are not near the dimerization interface they can be ignored. Within the interface itself Leu140 and Lys141 are predicted to have a very high $\Delta\Delta G$ (>7 across all results) indicating that they play a strong role in dimerization (Figure 7.11 C). Of note Lys141 observed to have a varied side chain position between the monomer and the dimer, whilst the other residue observed to move also scores, albeit not as highly (>2.9 across all models).

Finally, the residues Met136 and Glu144 diverge in score drastically, however, retain very similar positioning between the two models (Figure 7.11). Met136 scores 3.4 in both chains of model 1, however it receives a negative score in both chains of model 2. The positioning of nearby residue K141 is significant, indicating a change in the interactions assigned to Met136 in the two models (Figure 7.11 A). Similarly, the change in score between the two models for Glu144 appears to be mediated by movement of Pro135 and Leu95 (Figure 7.11 B).

No conformational change between the dimer and the monomeric model of ExsA, and a more detailed inspection of residues along the $\alpha 2$ and $\alpha 3$ revealed positional change in the sidechains of Glu127 and Lys141, both of which rotate away from the interface itself (Figure 7.11 A&B).

7.9 Together dimeric models 1 and 2 offer a glimpse of ExsA flexibility

The NTD and CTD of dimer models 1 and 2 are near identical, however their positioning relative to each other varies hugely. In model 1 $\alpha 5$ in the CTD includes sections of what is in model 2 (and prior homology models) a section of the flexible interdomain linker. It is indicated that this is an area of disorder by the IUPRED score, and in model 2 the IDDT score is poor for this region, whereas in model one the IDDT remains high (Figure 7.5).

To explore the feasibility of both models' potential for interaction they were aligned to the DNA bound MarA crystal structure (PDB: 1bl0). MarA is an AraC family protein which consists of only a DNA binding domain homologous to the ExsA CTD. Model 1 sandwiches the DNA when one CTD is aligned with MarA, in a positioning clearly not conducive to binding another site on the DNA (Figure 7.12 A). Model two projects away from the DNA, which when bound to MarA curves underneath the protein at 35° (Figure 7.12 B).

If both of the CTDs present in model 2 have MarA aligned to them the two DNA sites are greater than 100 Å apart and an excessive bending of the DNA would be required to connect the two (Figure 7.13 A). A composite of a monomer from both models 1 and 2 brings the binding sites closer together (61 Å), however, to connect the two DNA fragments a profoundly acute bend would be required (Figure 7.13 B).

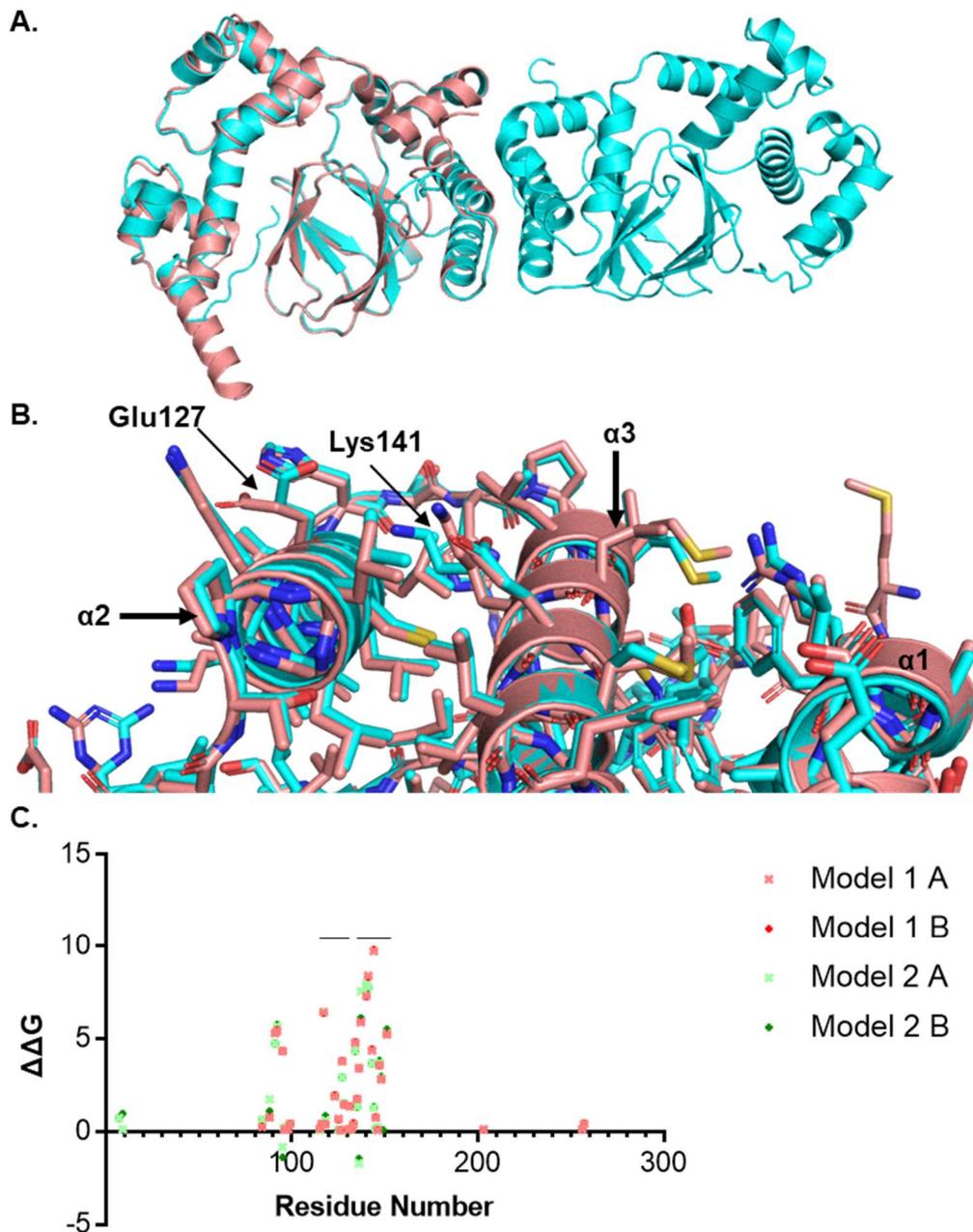
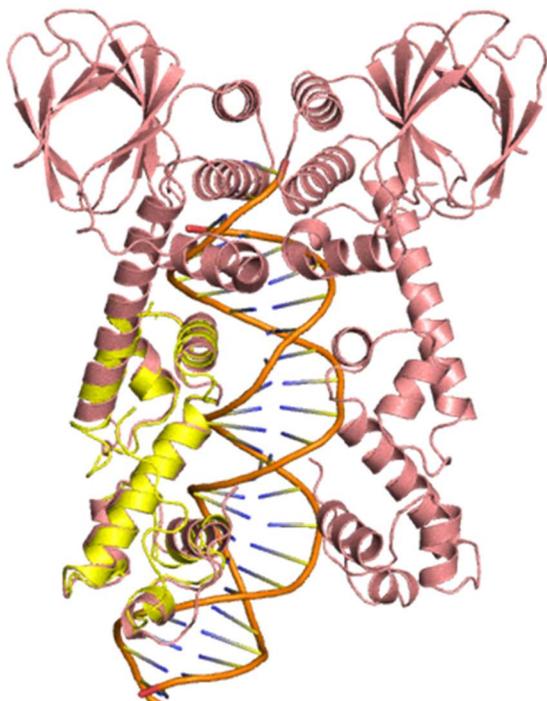


Figure 7.11 AlphaFold models of the ExsA monomer and dimer compared.

A&B. Cartoon depictions of the AlphaFold models of the ExsA monomer (pink) and dimer model 2 (blue). A. shows the overall cartoon structures, whilst B. shows half of the dimer interface and the monomeric model with residues shown in stick form.

C. Virtual alanine interface scanning data for both chains in AlphaFold ExsA dimer models 1 and 2. Only non 0 data is show. $\Delta\Delta G$ is the predicted change in the Gibbs free energy change for complex formation. The black bars indicate the residues within the two alpha helices central to dimerization.

A.



B.

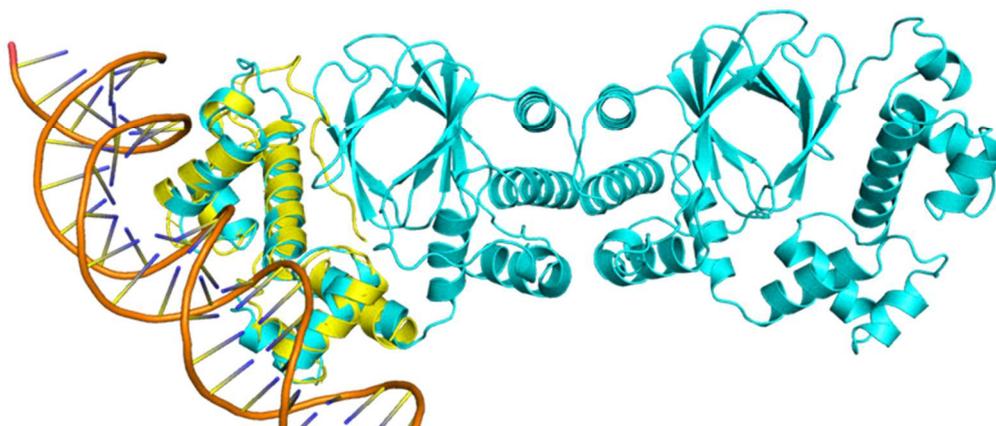


Figure 7.12 AlphaFold ExsA dimer models 1 and 2 with MarA.

Cartoon depictions of AlphaFold ExsA models 1 and 2 (**A**, pink and **B**, blue respectively) with DNA bound MarA (yellow) structure 1BL0 aligned to the models' CTDs.

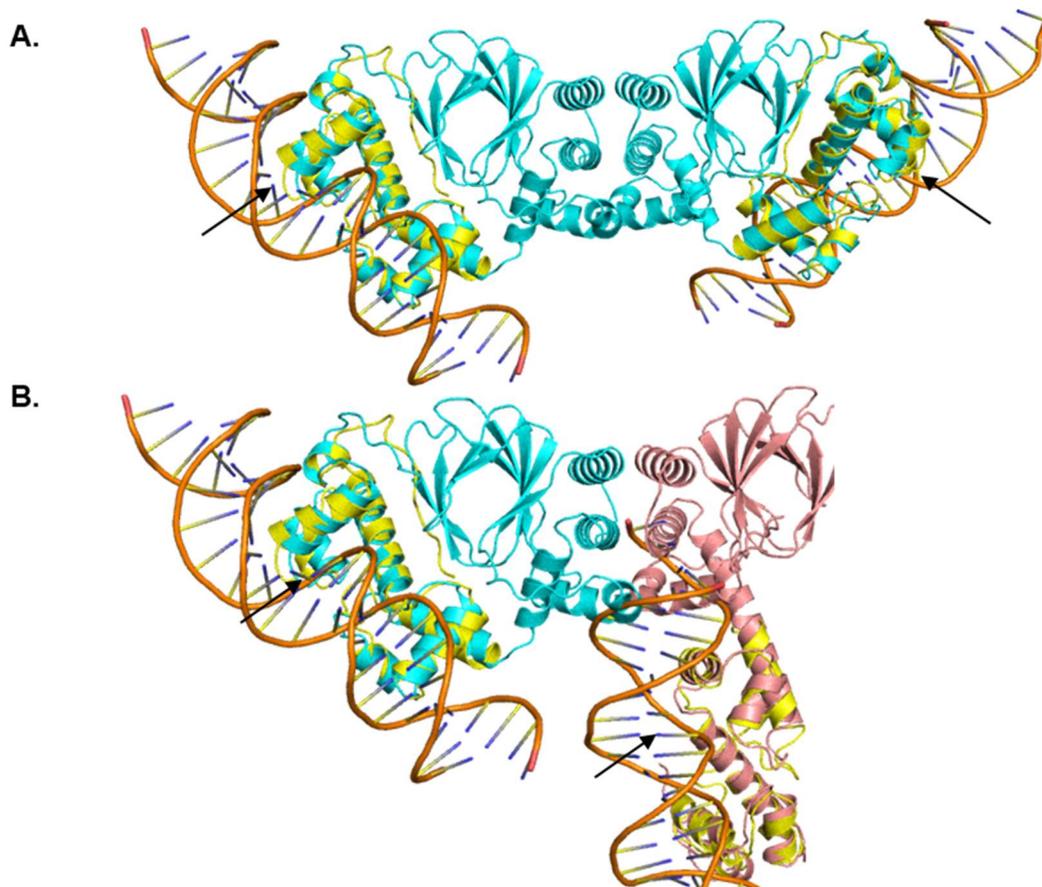


Figure 7.13 AlphaFold ExsA dimer models with two DNA molecules.

Cartoon depictions of AlphaFold ExsA model 2 (**A**, blue) and a composite (**B**) composed of one monomer from with AlphaFold ExsA model 2 (blue) and one monomer AlphaFold ExsA model 1 (pink). DNA bound MarA (yellow) structure 1BL0 is aligned to the models' CTDs. In **A**, the distance between arrows is 101 Å, whilst in **B**, the distance between arrows is 61 Å.

7.10 Molecular dynamics simulations support dimer model 2 and find correlation between dimeric and DNA bound structures

Molecular dynamics simulations were performed for the AlphaFold monomeric model (M), the monomeric model bound to DNA (M_{DNA}) by homology to a MarA DNA bound structure (PDB: 1bl0), and dimeric models 2. Backbone RMSD for each molecule and root mean square fluctuation (RMSF) per residue were recorded, in all models the backbone RMSD fluctuated over time, with greater variance apparent in the M (Figure 7.14 A). On a per residue basis the highest RMSF was seen around the N and C termini in all models, with less significant RMSF peaks corresponding to the flexible linkers around residues 100 and 190, and several other smaller peaks (Figure 7.14 C&D). By adding DNA interaction to the monomer not only the DNA binding domain within the CTD, but also many residues in the NTD had reduced RMSFs (Figure 7.14 C). Little distinction in RMSF was seen between the two chains of the dimeric

model (Figure 7.14 D).

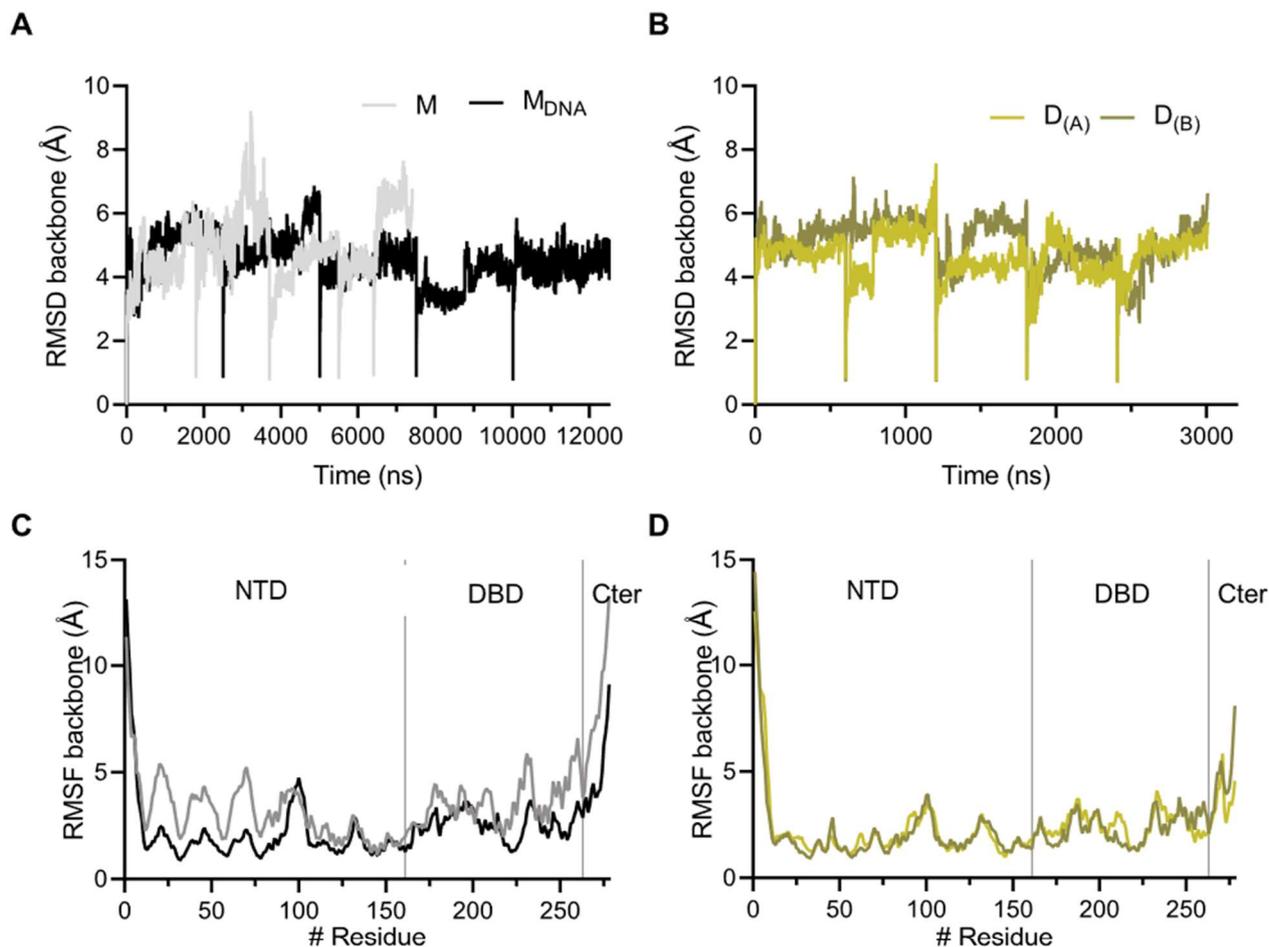


Figure 7.14 MD simulation protein backbone motion

Root-mean-square deviation (RMSD) of the protein backbone for the monomeric systems (A.) and dimeric systems (B.) along the simulated time. Root-mean-square fluctuation (RMSF) of the protein backbone for the monomeric systems (C.) and dimeric systems (D.), calculated by residue number.

Analysis of the proximity of the CTD to the NTD was also undertaken using the angle at the α carbon of Asn167 to Glu23 and Asn231, as well at the distance between Arg25 and Tyr224. Whilst flexibility and change were observed in all four simulated models, M showed far greater flexibility and distance between the NTD and CTD than the other models (Figure 7.15 B). M_{DNA}'s Arg25 to Tyr224 distance profile was very similar to that of the dimeric model, median of 10.9 Å which is between the 11.1 Å and 10.8 Å medians for the two chains of the dimer and considerably lower than the median distance of 13.7 Å of M.

The pre simulation monomeric model's angle of 17.1° and 10.1 \AA ; the median angle of the dimeric model is only slightly larger at 17.3° and 17.5° for the two chains. When simulated M_{DNA} reduced this angle on average, although the distance between Arg25 to Tyr224 remained comparable. In both angle and distance, the DNA bound monomer closely resembled the dimeric model (Figure 7.15 B&C). In both monomeric models the C terminus helix $\alpha 11$ unfolded for the majority of the time and was only present at $<1.5\%$ of timepoints in either chain of the dimeric simulation.

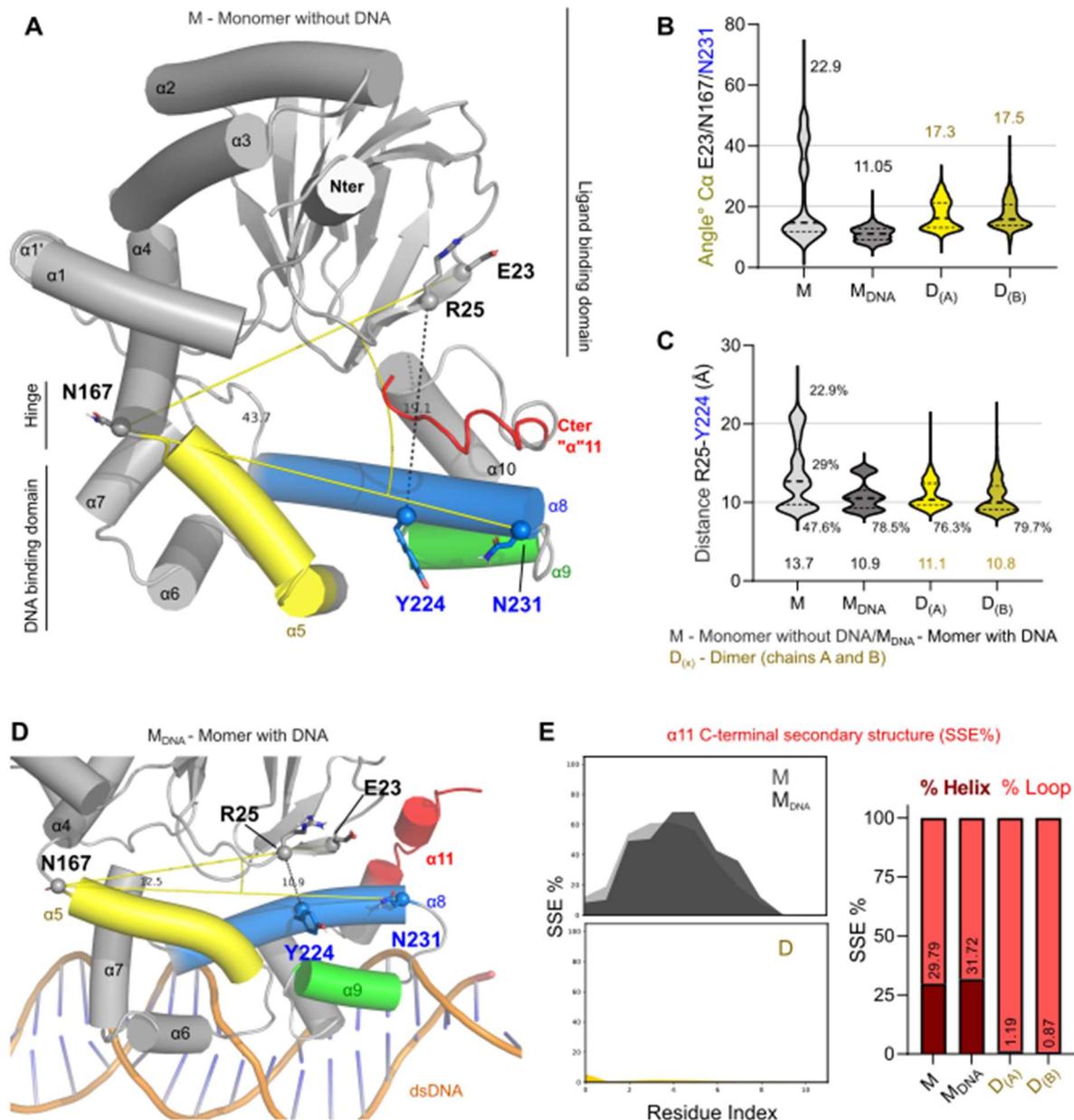


Figure 7.15 Overview of the ExsA model structure MD simulations.

A. Relevant frame from the MD simulations ExsA model. The regions of interest are highlighted with the following colour scheme that is used throughout this chapter: α 5-helix (residues 168–181), yellow; α 8 (residues 214–231), blue; α 9 (residues 236–244), green; α 11/c-termini (residues 268–278), red. Residues employed as reference points for geometrical calculations are depicted as spheres/sticks and labelled.

B. Variation in the opening of the NTD/CTD domains monitored the angle between the carbon alpha's of E23, N167 and N231 (angle depicted in yellow in **A.**)

C. Distances between the centre of mass from R25 and Y224. In both graphics the median values of respective distributions are provided and in multimodal cases, the population frequency is also displayed.

D. Representative snapshots of the CTD domain structure from monomeric simulations with DNA. Secondary structure of α 11 helix appears more stable in monomeric simulations than dimeric counterparts.

E. Area plots (left) represent the observed secondary structure element of the α 11 helix in percentage (right, fur summary) throughout the simulation for each system.

When examining an angle of a HTH within the CTD (His169, Glu178, Asn183) it was found that the dimeric simulation results resembled the results of M_{DNA} , whilst M varied more widely and the median and was less obtuse (Figure 7.16 A&B).

The dynamics of the dimerization interface in dimeric model where explored utilising the % of timepoints at which certain polar contacts were maintained, the distance between interface helices, and the distance between key hydrophobic residues. The distance between the interface helices had the most variance for α 2 between the two chains. The paired α 3 helices demonstrated little movement from their median distance of 12.1 Å apart, remaining in close contact throughout with each other throughout the simulation (Figure 7.16 D). Polar contact between K141 and E144 on the same monomer was maintain at 95% of time points, whilst these same residues made polar contact with each other on alternate monomers only 51% of the time (Figure 7.16 C). Finally hydrophobic interactions between L140 with the matching residue on the alternate ExsA unit, as well as intermolecular interactions between L137 and M147 varied very little in distance, remaining very close to the median distance of 7.3 Å (Figure 7.16 E).

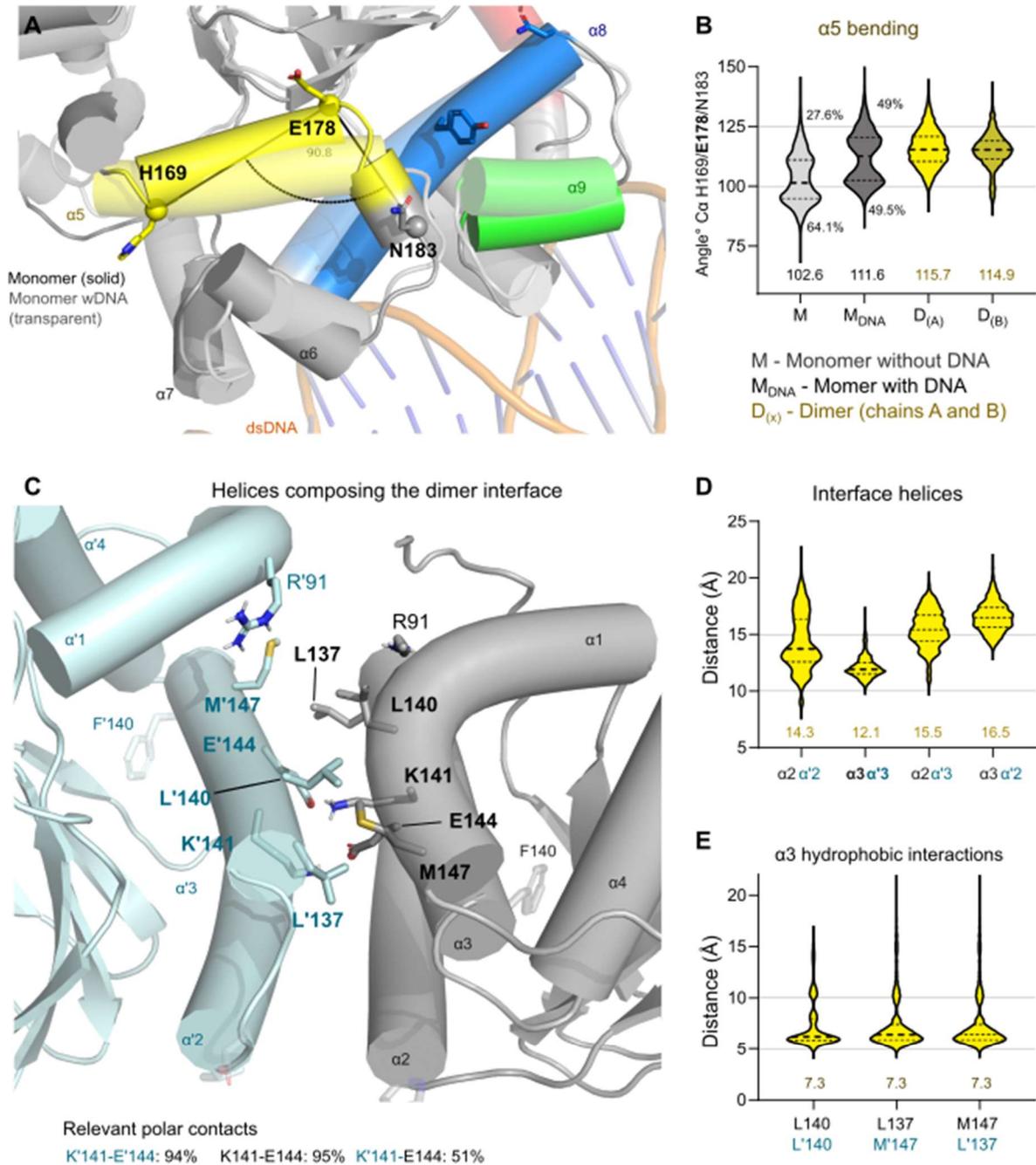


Figure 7.16 CTD and dimerization domain MD simulation overview

Variation in the conformation of CTD's $\alpha 5$ -helix monitored (angle depicted in black, **A.**) by the angle between the alpha carbons of H169, E178 and N183 (**B.**, violin plot with the values). **C.** Representative snapshot of the dimeric simulation displaying the residues stabilizing its interface. **D.** Pairwise distances between the centre of mass for each helix, followed by the distances between specific residues (**E.**). For all graphics the median values of respective distributions are provided and in multimodal cases, the population frequency is also displayed.

7.11 AlphaFold predicts novel ExsA-ExsD binding conformation

In order to determine the structure of the ExsA-ExsD heterodimer both full length sequences were input into AlphaFold multimer. The model produced offers a novel binding pose for ExsA and ExsD (Figure 7.17 A), with both constituents undergoing minimal conformational change compared to the full length ExsA model and the ExsD trimeric crystal structure (PBD: 3FD9) respectively (Figure 7.17 B). In the model ExsD interacts using a region not used in heterotrimer formation (Figure 7.17 C).

Because of the lack of conformational change in the predicted dimer, rigid body docking was used with the monomeric ExsA and ExsD models to provide an external control. None of the models thus generated resembled the AlphaFold heterodimeric model (Figure 7.18 A). The experiment was attempted again utilising monomeric inputs extracted from the AlphaFold dimeric model, however the same conformation was only achieved in the eighth ranked model, with 34 members in its cluster compared to 74 in the highest ranked cluster (Figure 7.18 B).

Given that *in vivo* ExsD interaction with ExsA is governed by the NTD an ExsA NTD-ExsD model was generated in AlphaFold. This model bore no resemblance to the ExsA-ExsD model previously generated, nor any of the ClusPro models (Figure 7.18 C).

7.13 AlphaFold models of ExsA mutations and an ExsD fragment conflict with published *in vitro* data

Whilst there is not published structural data on the ExsA-ExsD heterodimer there have been several experimental investigations of it using mutagenic techniques and other methods. Virtual alanine scanning of the ExsA-ExsD interface in the AlphaFold model highlighted several residues which scored a positive $\Delta\Delta G$ (indicating a contribution to the intramolecular interaction for that residue) *in silico* have previously been reported to have no observable effect on ExsA-ExsD interaction *in vitro* (Table 7.1) (Shrestha *et al.*, 2020). Finally, an AlphaFold model constructed with ExsA and the 40 N terminal residues of ExsD which has been implicated as important for the ExsA-ExsD interaction in the literature, showed no resemblance to the positioning of those residues in any prior models, nor did it the 40 N terminal residues interacted with the N terminal residues of ExsA as implied by prior experiments (Figure 7.19) (Shrestha *et al.*, 2020).

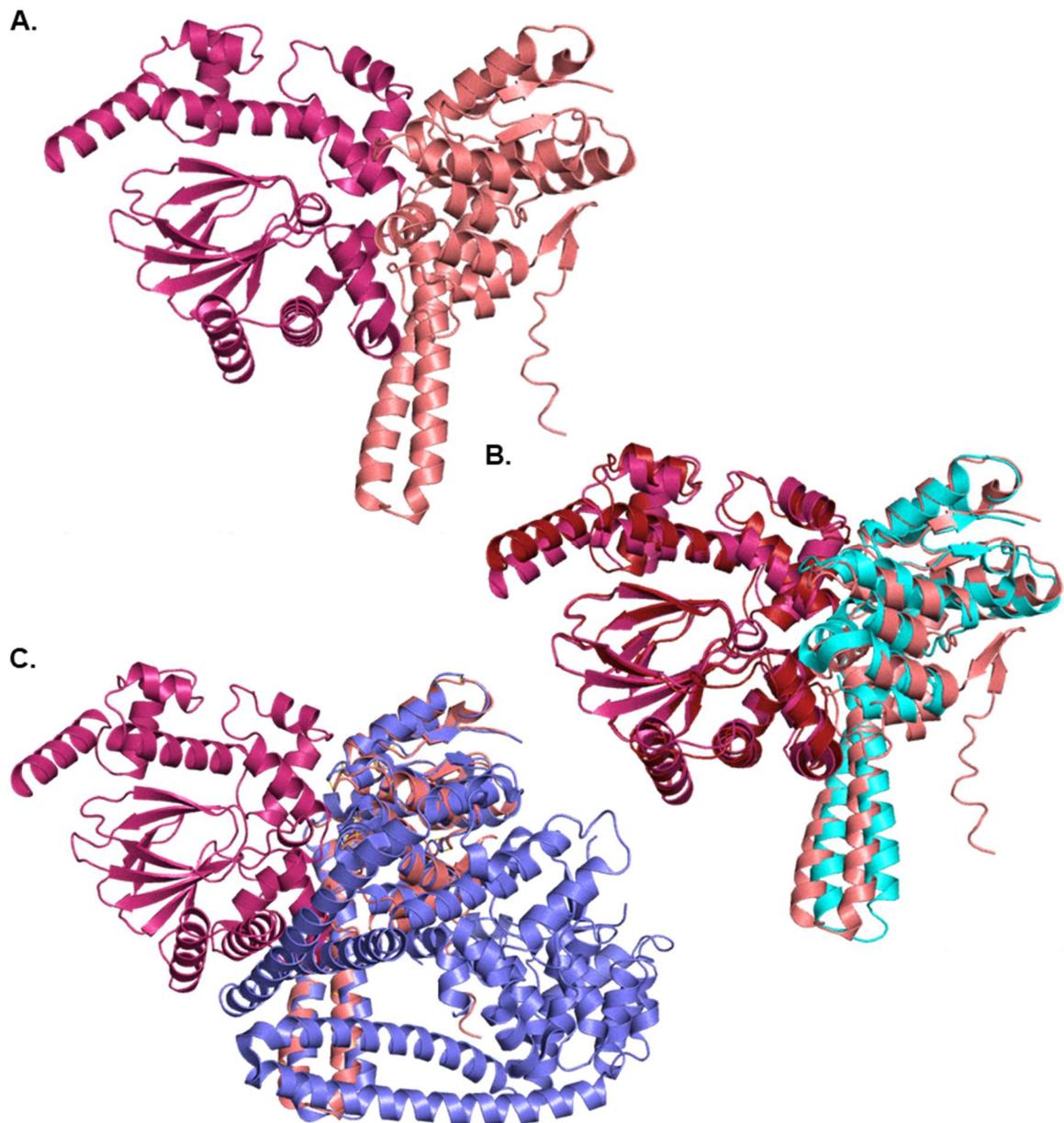


Figure 7.17 An AlphaFold model of ExsA-ExsD interaction.

Cartoon depictions of the AlphaFold ExsA-ExsD dimer model. The ExsA monomer is depicted in purple and the ExsD monomer in pink throughout. **B.** Includes a monomeric ExsD subunit from the PDB structure 3FD9 (turquoise) aligned to ExsD, and the full length AlphaFold model of ExsA (red) aligned to ExsA. **C.** The ExsD trimeric structure (PDB code: 3FD9, lilac) is aligned to ExsD.

Table 7.1 Comparison of virtual and experimental mutagenesis for ExsA-ExsD heterodimer

ExsA Residue	Ala Scan Score	Mutagenesis ExsD interaction in Shrestha et al., 2020
T48A	0	None
Q90A	1.3993	None
L95A	1.465	None
E98A	0.69	None
L129A	0	None
L140A	0	None
E144A	0	None
F151A	0.275	None

*A positive score indicates that the residue makes a contribution to the interaction, with higher scores signifying a greater contribution.



Figure 7.19 Modelling the ExsD 40 N terminal amino acids with ExsA.

An AlphaFold model of the 40 N terminal residues of ExsD (red) with ExsA (silver), overlaid with the prior ExsA (dark pink) and full length ExsD (light pink) model.

7.14 Discussion

7.14.1 Analysis of the AlphaFold-derived model of an ExsA monomer

The monomeric model generated by AlphaFold for ExsA fits well within our established knowledge of the proteins structure. The NTD is highly similar to the published structure (PBD: 4ZUA) and follows the general structure of the model constructed previously in Chapter 6 of the present study, with the caveat that the CTD is in a more open conformation relative to the ligand binding pocket and NTD more generally. This does not seem to be an artifact of AlphaFold as the same is not true of the AlphaFold ToxT model in comparison to the ToxT structure (PBD: 4MLO) from which this study's prior model derived its CTD positioning. This implies that, in fact, the CTD is further from the NTD in ExsA, although ligand induced

conformational changes could alter this conformation which is likely to be flexible. This flexibility was explored in MD simulations, which indicated that a “free floating” ExsA monomer has a larger gap between NTD and CTD, whilst dimerization or DNA binding held a more rigid conformation, with an NTD-CTD distance similar to that proposed in the AlphaFold model. Given that in all simulations the vast majority of variation from the core cluster around the median were towards a larger gap between NTD and CTD it seems unlikely that ExsA adopts a “closed” conformation like that observed in ToxT structures (e.g. PDB: 4MLO), unless induced to by a factor not present in the current simulations such as ligand binding.

The regions of lower confidence, as indicated by IDDT score (Jumper *et al.*, 2021), within the monomeric structure (and indeed all of the dimeric structures) correspond to disordered regions identifiable both computationally and by visual inspection. This is in line with previous findings (Akdal *et al.*, preprint), and implies that the lack of certainty corresponds to a lack of structural rigidity rather than an issue with the model *per se*.

7.14.2 Analysis of the AlphaFold-derived models of ExsA dimers

A notable distinction between the AlphaFold monomeric and dimeric models is the formation of C terminal helix $\alpha 11$ in the monomer, and the same residues consisting of an unstructured loop in all dimeric models. The MD simulations supported the dissolution of $\alpha 11$, with it being present a minority of the time in the monomeric simulations, and only very transiently in the dimeric simulation. This is consistent with the IUPRED prediction of disorder in the region. However, that $\alpha 11$ formed at a significant minority of time points in the monomeric simulation indicates that it is not energetically unfeasible for it to occur. The *in vivo* relevance of the potentially transient $\alpha 11$ is unclear as there is no known function of this portion of ExsA.

The NTD of the dimeric models of ExsA generated in AlphaFold bear resemblance to previously reported members of the AraC family, including AraC itself, utilising the same dimerization helix domain (Soisson *et al.*, 1997). It is also compatible with the assessment made of a symmetry mate dimer model previously investigated (Shrestha *et al.*, 2015). Finally, the $\alpha 3$ has previously been subjected to extensive mutagenesis (Marsden, Schubot and Yahr, 2014). The broad findings of that investigation support the pivotal role played by said helix in the AlphaFold model of ExsA dimerization, and the tight association maintained throughout the MD simulations. The lack of investigation of the $\alpha 2$ in that investigation stems from the looser interaction of dimerised AraC (which was assumed to also be the case for ExsA), in which $\alpha 2$ is only very loosely engaged. The AlphaFold model implied this was not the case, indeed the apparent formation of a leucine triad by Leu117, Leu148, and Leu137 at either end

of the dimerization interface provided a good justification for this tighter interaction for $\alpha 2$ than previously thought. The MD simulations, however, indicated a wider degree of flexibility in movement for $\alpha 2$. It therefore appears that $\alpha 2$ plays an auxiliary role in dimerization, whilst $\alpha 3$ provides main basis for the interaction. It also seemed probable, given the ANCHOR data, that the disordered loop connecting the two helices is stabilised by dimerization. Any attempt at disrupting the dimerization interface of ExsA as an anti-virulence strategy should therefore focus on $\alpha 3$, as was attempted in Chapter 5 of the present investigation.

ClusPro positions two rigid bodies next to each other based on 3D shape complementarity (ClusPro 2.0: protein-protein docking; Kozakov et al., 2017; Vajda et al., 2017), hence the monomers of ExsA make a very similar model to model 2, within the constraints of rigid body docking. The lack of side chain mobility makes detailed analysis of the dimerization interface less useful, however the broad consensus of the two varied docking methods provided an addition validation to model 2.

Comparison of the dimerization interfaces between models 1 and 2 by virtual alanine scanning revealed a sharp distinction at Met136 and Glu144. Met136 scores 3.4 in model 1 but a negative score model 2. This appeared to be the consequence of the movement of surrounding residues. The virtual results assigned to model 2 were more in line with the *in vivo* mutagenesis results previously published which indicated not defect to ExsA activity caused by mutation of these residues (Marsden, Schubot and Yahr, 2014). The MD simulation of the dimerization interface of model 2 indicated several hydrophobic interactions from residues on $\alpha 3$ which are tightly maintained, which again agreed with published mutagenesis data (Marsden, Schubot and Yahr, 2014).

Neither dimer model is in a DNA binding state as such; neither is compatible with the characterised binding of each monomer to a separate binding site 21 base pairs apart from centre to centre (Brutinel *et al.*, 2012). In model 1 this was manifestly apparent by the sandwiching of the DNA between the two CTDs. In model 2 it is apparent both from the extreme bend that would be necessitated in the DNA and the distance between the CTDs which implied a length of DNA far exceeding the approximate 71 Å length of 21 base pairs. This is perhaps unsurprising given that the models are as if in solution, whilst an ExsA dimer forms only when bound to DNA (Brutinel *et al.*, 2008). It is also apparent from the existing literature on ExsA-DNA interaction that upon binding to different promoters different degrees of DNA bending occur, with 20° of bending reported for the P_{exoT} promoter and 78° reported for the P_{exsC} promoter (Brutinel *et al.*, 2008).

Dimeric ExsA bound to DNA therefore retains a large degree of flexibility, however the dimeric MD simulations indicated that the flexibility in solution is less than indicated by the variation in the relative positioning of the CTD and NTD in models 1 and 2. AlphaFold dimeric models 1 and 2 could be conceived of as a likely dimeric structure (model 2), and an extreme which serves to highlight the potential flexibility of ExsA. An ExsA dimer has the flexibility to adopt conformations between the two as the promoter region of DNA to which it is binding dictates. The DNA binding interaction could alleviate the energetic constraints of NTD-CTD distance beyond that seen in dimeric or M_{DNA} bound M simulations.

7.14.3 Analysis of models of the ExsA-ExsD heterodimer

The ExsA-ExsD interaction is well characterised as 1:1 (Thibault *et al.*, 2009), and whilst no structural data is available, there is published mutagenesis experiments available to assist in interpretation of models (Shrestha *et al.*, 2020). It is apparent even within the current study (as was observed with the accurate recreation of published AraC family homodimers) that AlphaFold is capable of solving complex interprotein interactions with no obvious template, however it does not appear to have done so successfully in this instance. It is my opinion that none of the models of ExsA-ExsD interaction produced by AlphaFold or ClusPro present a reliable notion of the ExsA-ExsD interaction. One ClusPro model (Figure 7.14 A) utilised the ExsA homodimerization helices, which, along with the ligand binding pocket, have been demonstrated to not be required for ExsA-ExsD interaction (Brutinel, Vakulskas and Yahr, 2009; Shrestha *et al.*, 2015).

It has also been reported that ExsA with mutations at Y24P and V26P do not respond to ExsD, however they are not relevant to the interaction in any given model, with that section of the protein retaining a free loop structure. Given loose loop secondary structure it also seems highly improbable that either mutation affects the structure of the rest of the protein. The NT20 residues from ExsD are inhibitory to ExsA (Shrestha *et al.*, 2020), however they do not form a noteworthy or consistent interactions with ExsA when modelled alone or as part of ExsD. It seems probable that the previous suggestion of a helix formed by both N terminal regions is correct, and that it is too complex for the modelling methods employed here to determine this structure accurately (Shrestha *et al.*, 2020).

7. 15 Conclusions

In conclusion this chapter presents a model of both monomeric and dimeric ExsA which is supported by MD simulations. The dimeric model is of particular biological interest and, along

with the simulation data, makes a substantial contribution to our understanding of ExsA dimerization. The monomeric simulation without DNA exhibited excessive flexibility in comparison, whilst the DNA bound monomer had more similar molecular dynamics to the dimeric model, and thus seemed to have more reasonable energetic constraints. The data presented here should be used to inform structure choice in any future VS screens against ExsA. Specifically attempts to target the dimerization interface should be focused on $\alpha 3$, whilst attempts to target the ligand binding pocket should carefully consider the NTD-CTD distance of the model employed – ideally searching for molecules which can dock in models with both shorter and wider NTD-CTD distances.

8. Final Conclusions

P. aeruginosa is a pathogen of serious clinical importance, which is at the cutting edge of the AMR crisis because of both innate and acquired antibiotic resistance (World Health Organisation, 2017). Anti-virulence strategies offer the opportunity to target *P. aeruginosa* without the strong selective pressure towards resistance caused by traditional antibiotics (Rasko and Sperandio, 2010). *P. aeruginosa* has two virulence lifestyles, both of which are of clinical importance (Valentini et al., 2018). *P. aeruginosa*'s acute virulence is particularly concerning as a nosocomial infection (Anaissie, Penzak and Dignani, 2002; Micek et al., 2015). Acute infections of *P. aeruginosa* are dominated by the action of its T3SS (Holder, Neely and Frank, 2001; Lee et al., 2005; Vance, Rietsch and Mekalanos, 2005). The master regulator of the T3SS of *P. aeruginosa* is the transcription factor - ExsA (Brutinel et al., 2008). As a drug target ExsA has the added advantage of a wider virulence regulon beyond the T3SS (Tian et al., 2019), though the exact extent of this wider regulon is unclear.

In this thesis, I have furthered our understanding of the ExsA regulon. It has been demonstrated in a proteomics experiment that $\Delta exsD$ mutants, which have hyperactive ExsA, upregulated numerous virulence factors other than the T3SS and associated effectors. The two virulence factors to be confirmed as upregulated in $\Delta exsD$ mutants were pyocyanin and some components of the T6SS machinery, which were both quantified in independent experiments and found to be of increased abundance. The upregulation of pyocyanin is a novel finding, however it is unsurprising given the array of factors already known to cause an increase in pyocyanin expression (Liang et al., 2008; Little et al., 2018; Dong et al., 2021; Thees et al., 2021). Conversely the upregulation of T6SS proteins was unexpected because the T6SS is a chronic virulence trait and therefore reciprocally regulated with acute infection (Francis et al., 2018), and that previous ChIP-seq data indicated that ExsA does not interact with the promoter of one of the upregulated T6SS proteins, HcpA (NCBI Locus Tag PA1512, UniProtKB accession Q9HI36) (Huang et al., 2019). Given that the ExsA regulatory cascade responds to host cell contact (Rietsch et al., 2005), which is a feature of both chronic and acute infections, and that ExsA expression is varied within a population and therefore not beholden to factors which affect the entire population (Lin et al., 2021), this could be taken as evidence that the ExsA cascade operates in parallel with the acute/chronic signalling systems to promote a more general virulence phenotype than was previously thought. This is supported by the slight, but statistically significant, upregulation of QS proteins when ExsA is hyperactive. It is also the case however that ExsA can be mutationally inactivated in chronic infection adapted strains (Smith et al., 2006; Jeukens et al., 2014), and it therefore seems unlikely that ExsA plays an important role in chronic infections except for perhaps during establishment.

Further regulation downstream of ExsA has not been much discussed in previous literature. During a ChIP-seq investigation a regulatory RNA has been identified as transcriptionally regulated by ExsA (Huang *et al.*, 2019), and thereby providing the only prior indication of a downstream regulation. The RNA in question, *phrS*, is active in anaerobic conditions and is unlikely to have been active within the proteomics experiment, which was conducted in aerobic conditions (Sonnleitner *et al.*, 2011). The proteomic data clearly illustrates the fact the ExsA has an indirect impact on the abundance of proteins that are not within its direct regulon due to the large and statistically significant upregulation of HcpA (NCBI Locus TagPA1512, UniProtKB accession Q9HI36) in the “always on” mutants. A possible cause for these secondary regulatory effects at two putative TFs of unknown function which are upregulated in the “always on” mutants.

The regulation of siderophore production and denitrification by the ExsA cascade is also a novel finding. Again, it is possible that they are the result of transcriptional or translational regulation downstream of ExsA, hence their absence from the established literature. In their totality the proteomics results expand the ExsA regulon, especially demonstrating ExsA’s role in promoting the production of virulence factors beyond the T3SS.

Both the T3SS and even individual effectors thereof have been targets for anti-virulence drug development endeavours against *P. aeruginosa* (Sawa *et al.*, 2014; Foulkes *et al.*, 2019). Given both the data within this thesis and the literature discussed above, it is eminently apparent that ExsA offers an attractive drug target, not only placing severe limitations on the expression of the T3SS but also reducing the expression of a variety of other virulence factors. The only known compounds to target and inhibit ExsA however have a broad spectrum of activity against the AraC family of TFs, which is potentially disadvantageous. Whilst the AraC family is only present in bacteria its members have a wide variety of functions, and inhibition of some AraC proteins could increase virulence or have other undesirable effects. A pertinent example is the two *P. aeruginosa* AraC family proteins VqsM and SphR, the inhibition of which is likely to increase virulence (LaBauve and Wargo, 2014; Liang *et al.*, 2014; Okino and Ito, 2016). By targeting the less conserved regulatory NTD, this nonselective antibacterial activity can be avoided, and a compound which acts specifically on ExsA could be developed. Exploration of the literature and phylogenetic data revealed no leads from which to develop an inhibitor.

Three possible target sites or strategies are apparent when considering the NTD of ExsA. These include the dimerization interface, a conserved AraC family ligand binding pocket, and

mimicking the effect of ExsD. The last option is precluded by a dearth of structural information on the ExsA-ExsD interaction. Throughout this thesis both experimental and state of the art modelling approaches were taken to resolve this issue but without a decisive success. It is possible that future developments of modelling methods, or a different experimental approach such as cryogenic electron microscopy, could lead to the elucidation of the ExsA-ExsD structure which is compatible with the data obtained by mutagenesis (Shrestha *et al.*, 2020). Until this information is available however, only the other two sites are targetable by approaches such as VS which require structural data.

Analysis of SNPs did not identify any genetic variation at the targeted section of the dimerization helix or the ligand binding pocket on ExsA. Several of the nearby SNPs, including one in the dimerization domain and the HTH DNA binding domain, were identified in chronic adapted *P. aeruginosa* strains which are unable to cause the acute infections for which an anti-ExsA therapeutic is intended (Salunkheet *et al.*, 2005). Given their suitability for targeting and the absence of any leads from which to work a VS approach was taken to targeting both these sites.

Three rounds of VS were undertaken. Two utilised only the existing structure of the ExsA NTD (PDB: 4ZUA), whilst the final iteration utilised a homology model of full length ExsA. After the VS stage of this project was completed, two new state-of-the-art protein modelling methods namely the AlphaFold (Jumper *et al.*, 2021) and RosettaFold (Baek *et al.*, 2021), both of which make use of deep learning, were released. By utilising AlphaFold, with older rigid body docking methods (e.g. by ClusPro) as comparator, new models of ExsA were proposed. By employing atomistic MD simulations the flexibility of ExsA was exhaustively explored. This model, combined with an understanding of the proteins inherent flexibility revealed through MD simulations, offers a strong basis for future VS attempts against ExsA. The model varied from the one used in the VS conducted in this study most significantly in the positioning of the CTD relative to the NTD. The exact positioning utilised in a VS run could be an important factor as it determines the degree to which the “rear end” of any molecule inserted into the ligand binding pocket will interact with the CTD.

The AlphaFold models of ExsA, both monomeric and dimeric, had a wider gap between the NTD and CTD than the model employed for VS within this study. This indicates that out of the two models of ExsA used for the VS namely a full length ExsA model or an NTD only structure (PDB: 4ZUA), the lack of a CTD in the NTD only structure was not a limitation and may have provided a more realistic structure than the that offered by a model with the CTD. There is however the possibility that ligand binding would induce a narrower gap between NTD and

CTD to a position more similar to that of ToxT, although further MD simulations would be required to support this assertion. Interestingly the DNA bound monomer and dimeric structure both had similar MD profiles, whilst a free monomer proved much more flexible. The MD simulations with ExsA dimer also confirmed the importance of the $\alpha 3$ helix which was targeted by a ligand-based VS screen in the present study. Whilst $\alpha 2$ helix and to a lesser extent $\alpha 1$ helix did appear to make contact at the dimerization interface, they were much more mobile than the $\alpha 3$ helix, which remained in close association.

The VS runs within the present study yielded five compounds of interest, generated in both ligand-guided and structure-guided approaches. All these hits were found to have an effect on the expression of a luciferase reporter when placed after an ExsA controlled promoter region but have no effect of the same system when placed after an unrelated promoter region. Despite this *in vivo* validation of those compounds, no biophysical or functional confirmation could be obtained. This is in part because of their failure in a thermal shift assay, though whether this was a negative result or assay failure was unclear since no known ExsA inhibitor binding to its NTD as positive control was available. Further investigation of these compounds was impaired by the interruptions caused by Covid-19 pandemic.

The most interesting finding from the VS was the recurring presence of a 4-hydroxyquinazoline chemical group in the hits across different software. The core group, which was fragment-sized, alone also docked well and gave a positive outcome when tested *in vivo*. Pending further experimental validation, this fragment could potentially be used as a “warhead” compound and as such be useful to design lead compounds that can bind into the ligand binding pocket. This validation should be undertaken with a biophysical method, for instance surface plasmon resonance, followed by a functional assay such as the EMSA which was attempted within this study. Combined with the *in vivo* data already obtained this would justify infection models using both human tissue culture and insect models such as *Galleria mellonella* larvae. A structure activity relationship for the class of compounds could then be established experimentally. Last but not the least, I argue that using more than one algorithm, use of consensus docking scores and meticulous inspection of possible binding mode(s) of hit molecules against carefully constructed ExsA models together with explicit consideration of structural flexibility, could potentially lead to better outcome in the VS and subsequent rational drug discovery endeavour in future against this anti-virulent drug target.

References

- Abagyan, R. and Totrov, M. (1994) 'Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins', *Journal of molecular biology*, 235(3), pp. 983–1002. Available at: <https://doi.org/10.1006/JMBI.1994.1052>.
- Abagyan, R., Totrov, M. and Kuznetsov, D. (1994) 'ICM—A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation', *Journal of Computational Chemistry*, 15(5), pp. 488–506. Available at: <https://doi.org/10.1002/JCC.540150503>.
- Abraham, E.P. and Chain, E. (1940) 'An Enzyme from Bacteria able to Destroy Penicillin', *Nature*, 146(3713), pp. 837–837. Available at: <https://doi.org/10.1038/146837a0>.
- Aburto-Rodríguez, N.A. *et al.* (2021) 'Anti-Pathogenic Properties of the Combination of a T3SS Inhibitory Halogenated Pyrrolidone with C-30 Furanone', *Molecules*, 26(24). Available at: <https://doi.org/10.3390/MOLECULES26247635>.
- Adla, S.K. *et al.* (2021) 'Neurosteroids: Structure-Uptake Relationships and Computational Modeling of Organic Anion Transporting Polypeptides (OATP)1A2', *Molecules*, 26(18). Available at: <https://doi.org/10.3390/MOLECULES26185662>.
- Aires, J.R. *et al.* (1999) 'Involvement of an active efflux system in the natural resistance of *Pseudomonas aeruginosa* to aminoglycosides', *Antimicrobial Agents and Chemotherapy*, 43(11), pp. 2624–2628. Available at: <https://doi.org/10.1128/aac.43.11.2624>.
- Akdel, M. *et al.* (no date) 'A structural biology community assessment of AlphaFold 2 applications'. Available at: <https://doi.org/10.1101/2021.09.26.461876>.
- Albrecht-Gary, A.M. *et al.* (1994) 'Bacterial Iron Transport: Coordination Properties of Pyoverdine PaA, a Peptidic Siderophore of *Pseudomonas aeruginosa*', *Inorganic Chemistry*, 33(26), pp. 6391–6402. Available at: https://doi.org/10.1021/IC00104A059/SUPPL_FILE/IC00104A059_SI_001.PDF.
- Ali, S.O. *et al.* (2018) 'Phase 1 study of MEDI3902, an investigational anti-*Pseudomonas aeruginosa* PcrV and Psl bispecific human monoclonal antibody, in healthy adults', *Clinical Microbiology and Infection* [Preprint]. Available at: <https://doi.org/10.1016/j.cmi.2018.08.004>.
- Allen, L. *et al.* (2005) 'Pyocyanin production by *Pseudomonas aeruginosa* induces neutrophil apoptosis and impairs neutrophil-mediated host defenses in vivo', *Journal of immunology (Baltimore, Md. : 1950)*, 174(6), pp. 3643–3649. Available at: <https://doi.org/10.4049/JIMMUNOL.174.6.3643>.
- Almohaywi, B. *et al.* (2018) 'Design and synthesis of lactams derived from mucochloric and mucobromic acids as *pseudomonas aeruginosa* quorum sensing inhibitors', *Molecules*, 23(5), p. 1106. Available at: <https://doi.org/10.3390/molecules23051106>.
- AlphaFold.ipynb* - Colaboratory (no date). Available at:

<https://colab.research.google.com/github/deepmind/alphafold/blob/main/notebooks/AlphaFold.ipynb#scrollTo=pc5-mbsX9PZC> (Accessed: 17 February 2022).

van Alst, N.E. *et al.* (2007) 'Nitrate sensing and metabolism modulate motility, biofilm formation, and virulence in *Pseudomonas aeruginosa*', *Infection and Immunity*, 75(8), pp. 3780–3790. Available at: <https://doi.org/10.1128/IAI.00201-07/ASSET/084D4CFA-D852-4C74-854D-861478E7BF92/ASSETS/GRAPHIC/ZII0080768440008.JPEG>.

Altamirano, F.L.G. *et al.* (2022) 'Phage-antibiotic combination is a superior treatment against *Acinetobacter baumannii* in a preclinical study', *eBioMedicine*, 80, p. 104045. Available at: <https://doi.org/10.1016/J.EBIOM.2022.104045>.

Anaissie, E.J., Penzak, S.R. and Dignani, M.C. (2002) 'The hospital water supply as a source of nosocomial infections: a plea for action.', *Archives of internal medicine*, 162(13), pp. 1483–92. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12090885> (Accessed: 17 October 2017).

Angriman, F. *et al.* (2021) 'Interleukin-6 receptor blockade in patients with COVID-19: placing clinical trials into context', *The Lancet Respiratory Medicine*, 9(6), pp. 655–664. Available at: [https://doi.org/10.1016/S2213-2600\(21\)00139-9](https://doi.org/10.1016/S2213-2600(21)00139-9).

Anthouard, R. and DiRita, V.J. (2013) 'Small-molecule inhibitors of toxT expression in *Vibrio cholerae*', *mBio*, 4(4). Available at: <https://doi.org/10.1128/mBio.00403-13>.

Arora, S.K. *et al.* (2005) 'Role of motility and flagellin glycosylation in the pathogenesis of *Pseudomonas aeruginosa* burn wound infections', *Infection and Immunity*, 73(7), pp. 4395–4398. Available at: <https://doi.org/10.1128/IAI.73.7.4395-4398.2005>.

Arya, T. *et al.* (2019) 'Fragment-based screening identifies inhibitors of ATPase activity and of hexamer formation of Cag α from the *Helicobacter pylori* type IV secretion system', *Scientific Reports*, 9(1), p. 6474. Available at: <https://doi.org/10.1038/s41598-019-42876-6>.

Ashkenazy, H. *et al.* (2016) 'ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules', *Nucleic acids research*, 44(W1), pp. W344–W350. Available at: <https://doi.org/10.1093/NAR/GKW408>.

Augustin, T.L. *et al.* (2020) 'Novel Small-Molecule Scaffolds as Candidates against the SARS Coronavirus 2 Main Protease: A Fragment-Guided in Silico Approach', *Molecules*, 25(23). Available at: <https://doi.org/10.3390/MOLECULES25235501>.

B, R. *et al.* (2008) '*Pseudomonas aeruginosa*, cyanide accumulation and lung function in CF and non-CF bronchiectasis patients', *The European respiratory journal*, 32(3). Available at: <https://doi.org/10.1183/09031936.00159607>.

Babin, B.M. *et al.* (2017) 'Selective proteomic analysis of antibiotic-tolerant cellular subpopulations in *pseudomonas aeruginosa* biofilms', *mBio*, 8(5). Available at: <https://doi.org/10.1128/mBio.01593-17>.

Baek, M. *et al.* (2021) 'Accurate prediction of protein structures and interactions using a three-

track neural network', *Science*, 373(6557), pp. 871–876. Available at: https://doi.org/10.1126/SCIENCE.ABJ8754/SUPPL_FILE/ABJ8754_MДАР_REPRODUCIBILITY_CHECKLIST.PDF.

Bassetti, M., Villa, G. and Pecori, D. (2014) 'Antibiotic-resistant *Pseudomonas aeruginosa*: focus on care in patients receiving assisted ventilation', *Future Microbiology*, 9(4), pp. 465–474. Available at: <https://doi.org/10.2217/fmb.14.7>.

Bernhards, R.C. *et al.* (2009) 'Structural evidence suggests that antiactivator ExsD from *Pseudomonas aeruginosa* is a DNA binding protein.', *Protein science: a publication of the Protein Society*, 18(3), pp. 503–13. Available at: <https://doi.org/10.1002/pro.48>.

Bernhards, R.C. *et al.* (2013) 'Self-trimerization of ExsD limits inhibition of the *Pseudomonas aeruginosa* transcriptional activator ExsA in vitro', *The FEBS journal*, 280(4), p. 1084. Available at: <https://doi.org/10.1111/FEBS.12103>.

Bisht, K., Baishya, J. and Wakeman, C.A. (2020) 'Pseudomonas aeruginosa polymicrobial interactions during lung infection', *Current Opinion in Microbiology*. Elsevier Ltd, pp. 1–8. Available at: <https://doi.org/10.1016/j.mib.2020.01.014>.

Bowlin, N.O. *et al.* (2014) 'Mutations in the *Pseudomonas aeruginosa* Needle Protein Gene *pscF* Confer Resistance to Phenoxycetamide Inhibitors of the Type III Secretion System', *Antimicrobial Agents and Chemotherapy*, 58(4), pp. 2211–2220. Available at: <https://doi.org/10.1128/AAC.02795-13>.

Bowser, T.E. *et al.* (2007) 'Novel anti-infection agents: Small-molecule inhibitors of bacterial transcription factors', *Bioorganic and Medicinal Chemistry Letters*, 17(20), pp. 5652–5655. Available at: <https://doi.org/10.1016/j.bmcl.2007.07.072>.

Brandel, J. *et al.* (2012) 'Pyochelin, a siderophore of *Pseudomonas aeruginosa*: physicochemical characterization of the iron(III), copper(II) and zinc(II) complexes', *Dalton transactions (Cambridge, England: 2003)*, 41(9), pp. 2820–2834. Available at: <https://doi.org/10.1039/C1DT11804H>.

Branco-Vanegas, J. *et al.* (2014) 'Glycosylflavonoids from *Cecropia pachystachya* Trécul are quorum sensing inhibitors', *Phytomedicine*, 21(5), pp. 670–675. Available at: <https://doi.org/10.1016/J.PHYMED.2014.01.001>.

Brint, J.M. and Ohman, D.E. (1995) 'Synthesis of multiple exoproducts in *Pseudomonas aeruginosa* is under the control of RhIR-RhII, another set of regulators in strain PAO1 with homology to the autoinducer-responsive LuxR-LuxI family', *Journal of bacteriology*, 177(24), pp. 7155–7163. Available at: <https://doi.org/10.1128/JB.177.24.7155-7163.1995>.

Brutinel, E.D. *et al.* (2008) 'Characterization of ExsA and of ExsA-dependent promoters required for expression of the *Pseudomonas aeruginosa* type III secretion system', *Molecular Microbiology*, 68(3), pp. 657–671. Available at: <https://doi.org/10.1111/j.1365-2958.2008.06179.x>.

- Brutinel, E.D. *et al.* (2012) 'The distal ExsA-binding site in *Pseudomonas aeruginosa* type III secretion system promoters is the primary determinant for promoter-specific properties.', *Journal of bacteriology*, 194(10), pp. 2564–72. Available at: <https://doi.org/10.1128/JB.00106-12>.
- Brutinel, E.D., Vakulskas, C.A. and Yahr, T.L. (2009) 'Functional domains of ExsA, the transcriptional activator of the *Pseudomonas aeruginosa* type III secretion system', *Journal of Bacteriology*, 191(12), pp. 3811–3821. Available at: <https://doi.org/10.1128/JB.00002-09>.
- Brutinel, E.D., Vakulskas, C.A. and Yahr, T.L. (2010) 'ExsD Inhibits Expression of the *Pseudomonas aeruginosa* Type III Secretion System by Disrupting ExsA Self-Association and DNA Binding Activity', *Journal of Bacteriology*, 192(6), pp. 1479–1486. Available at: <https://doi.org/10.1128/JB.01457-09>.
- Bryan-Wilson, J. (2016) 'No time to wait: securing the future from drug-resistant infections', *Artforum International*, 54(10), pp. 113–114.
- Burstein, D. *et al.* (2015) 'Novel type III effectors in *pseudomonas aeruginosa*', *mBio*, 6(2), p. e00161. Available at: <https://doi.org/10.1128/mBio.00161-15>.
- Callejo, G. *et al.* (2020) 'In silico screening of GMQ-like compounds reveals guanabenz and sephin1 as new allosteric modulators of acid-sensing ion channel 3', *Biochemical Pharmacology* [Preprint]. Available at: <https://doi.org/10.1101/704445>.
- Capriotti, E., Fariselli, P. and Casadio, R. (2005) 'I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure', *Nucleic acids research*, 33(Web Server issue). Available at: <https://doi.org/10.1093/NAR/GKI375>.
- Carlson, C.A. and Ingraham, J.L. (1983) 'Comparison of denitrification by *Pseudomonas stutzeri*, *Pseudomonas aeruginosa*, and *Paracoccus denitrificans*', *Applied and Environmental Microbiology*, 45(4), pp. 1247–1253. Available at: <https://doi.org/10.1128/AEM.45.4.1247-1253.1983>.
- Chambonnier, G. *et al.* (2016) 'The Hybrid Histidine Kinase LadS Forms a Multicomponent Signal Transduction System with the GacS/GacA Two-Component System in *Pseudomonas aeruginosa*', *PLoS Genetics*, 12(5), p. e1006032. Available at: <https://doi.org/10.1371/journal.pgen.1006032>.
- Chen, F., Xia, Q. and Ju, L.K. (2006) 'Competition between oxygen and nitrate respirations in continuous culture of *Pseudomonas aeruginosa* performing aerobic denitrification', *Biotechnology and Bioengineering*, 93(6), pp. 1069–1078. Available at: <https://doi.org/10.1002/BIT.20812>.
- Chen, L. *et al.* (2015) 'Composition, function, and regulation of T6SS in *Pseudomonas aeruginosa*', *Microbiological Research*, 172, pp. 19–25. Available at: <https://doi.org/10.1016/j.micres.2015.01.004>.
- Cisek, A.A. *et al.* (2017) 'Phage Therapy in Bacterial Infections Treatment: One Hundred

Years After the Discovery of Bacteriophages.’, *Current microbiology*, 74(2), pp. 277–283. Available at: <https://doi.org/10.1007/s00284-016-1166-x>.

Clay, M.E. *et al.* (2020) ‘Pseudomonas aeruginosa lasR mutant fitness in microoxia is supported by an Anr-regulated oxygen-binding hemerythrin’, *Proceedings of the National Academy of Sciences of the United States of America*, 117(6), pp. 3167–3173. Available at: <https://doi.org/10.1073/PNAS.1917576117/-/DCSUPPLEMENTAL>.

ClusPro 2.0: protein-protein docking (no date). Available at: https://cluspro.bu.edu/dimer_predict/submit.php (Accessed: 23 February 2022).

Cornelis, P. *et al.* (2022) ‘High affinity iron uptake by pyoverdine in Pseudomonas aeruginosa involves multiple regulators besides Fur, PvdS, and Fpvl’, *BioMetals*, pp. 1–7. Available at: <https://doi.org/10.1007/S10534-022-00369-6/FIGURES/3>.

Cornelis, P., Matthijs, S. and van Oeffelen, L. (2009) ‘Iron uptake regulation in Pseudomonas aeruginosa’, *BioMetals*, 22(1), pp. 15–22. Available at: <https://doi.org/10.1007/S10534-008-9193-0/FIGURES/3>.

Cox, C.D. (1986) ‘Role of pyocyanin in the acquisition of iron from transferrin’, *Infection and immunity*, 52(1), pp. 263–270. Available at: <https://doi.org/10.1128/IAI.52.1.263-270.1986>.

Cunrath, O. *et al.* (2020) ‘The pathogen Pseudomonas aeruginosa optimizes the production of the siderophore pyochelin upon environmental challenges’, *Metallomics*, 12(12), pp. 2108–2120. Available at: <https://doi.org/10.1039/D0MT00029A>.

Curtis, M.M. *et al.* (2014) ‘QseC inhibitors as an antivirulence approach for gram-negative pathogens’, *mBio*, 5(6), p. e02165. Available at: <https://doi.org/10.1128/mBio.02165-14>.

Czechowska, K. *et al.* (2014) ‘Cheating by type 3 secretion system-negative Pseudomonas aeruginosa during pulmonary infection’, *The Proceedings of the National Academy of Sciences*, 111(21), pp. 7801–7806. Available at: <https://doi.org/10.1073/pnas.1400782111>.

D’Angelo, F. *et al.* (2018) ‘Identification of FDA-Approved Drugs as Antivirulence Agents Targeting the pqs Quorum-Sensing System of Pseudomonas aeruginosa’, *Antimicrobial Agents and Chemotherapy*, 62(11). Available at: <https://doi.org/10.1128/AAC.01296-18>.

Darden, T., York, D. and Pedersen, L. (1998) ‘Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems’, *The Journal of Chemical Physics*, 98(12), p. 10089. Available at: <https://doi.org/10.1063/1.464397>.

Davis, K.M. and Isberg, R.R. (2019) ‘One for All, but Not All for One: Social Behavior During Bacterial Diseases’, *Trends in microbiology*, 27(1), p. 64. Available at: <https://doi.org/10.1016/J.TIM.2018.09.001>.

Deredjian, A. *et al.* (2014) ‘Low occurrence of Pseudomonas aeruginosa in agricultural soils with and without organic amendment’, *Frontiers in Cellular and Infection Microbiology*, 4, p. 53. Available at: <https://doi.org/10.3389/fcimb.2014.00053>.

Diaz, M.R., King, J.M. and Yahr, T.L. (2011) ‘Intrinsic and Extrinsic Regulation of Type III

Secretion Gene Expression in *Pseudomonas Aeruginosa*', *Frontiers in Microbiology*, 2(APR). Available at: <https://doi.org/10.3389/FMICB.2011.00089>.

Dolan, S.K. *et al.* (2020) 'Contextual flexibility in *pseudomonas aeruginosa* central carbon metabolism during growth in single carbon sources', *mBio*, 11(2). Available at: https://doi.org/10.1128/MBIO.02684-19/SUPPL_FILE/MBIO.02684-19-ST001.DOCX.

Dolan, S.K. and Welch, M. (2018) 'The Glyoxylate Shunt, 60 Years On', *Annual review of microbiology*, 72, pp. 309–330. Available at: <https://doi.org/10.1146/ANNUREV-MICRO-090817-062257>.

Dong, L. *et al.* (2021) 'Ostarine attenuates pyocyanin in *Pseudomonas aeruginosa* by interfering with quorum sensing systems', *The Journal of antibiotics*, 74(12), pp. 863–873. Available at: <https://doi.org/10.1038/S41429-021-00469-4>.

Dortet, L. *et al.* (2018) 'Pore-forming activity of the *Pseudomonas aeruginosa* type III secretion system translocon alters the host epigenome', *Nature Microbiology*, p. 1. Available at: <https://doi.org/10.1038/s41564-018-0109-7>.

Doyle, A.A. and Stephens, J.C. (2019) 'A review of cinnamaldehyde and its derivatives as antibacterial agents', *Fitoterapia*. Elsevier B.V., p. 104405. Available at: <https://doi.org/10.1016/j.fitote.2019.104405>.

Dror, R.O. *et al.* (2010) 'Exploring atomic resolution physiology on a femtosecond to millisecond timescale using molecular dynamics simulations', *The Journal of general physiology*, 135(6), pp. 555–562. Available at: <https://doi.org/10.1085/JGP.200910373>.

Ellis, M.J. *et al.* (2019) 'A macrophage-based screen identifies antibacterial compounds selective for intracellular *Salmonella Typhimurium*', *Nature Communications*, 10(1), p. 197. Available at: <https://doi.org/10.1038/s41467-018-08190-x>.

Emanuele, A.A. *et al.* (2014) 'Potential novel antibiotics from HTS targeting the virulence-regulating transcription factor, VirF, from *Shigella flexneri*', *Journal of Antibiotics*, 67(5), pp. 379–386. Available at: <https://doi.org/10.1038/ja.2014.10>.

Emanuele, A.A. and Garcia, G.A. (2015) 'Mechanism of action and initial, in vitro SAR of an inhibitor of the *Shigella flexneri* virulence regulator virF', *PLoS ONE*, 10(9), p. e0137410. Available at: <https://doi.org/10.1371/journal.pone.0137410>.

Engel, L.S. *et al.* (1998) 'Protease IV, a unique extracellular protease and virulence factor from *Pseudomonas aeruginosa*', *Journal of Biological Chemistry*, 273(27), pp. 16792–16797. Available at: <https://doi.org/10.1074/jbc.273.27.16792>.

Enjapoori, A.K. *et al.* (2014) 'Monotreme Lactation Protein Is Highly Expressed in Monotreme Milk and Provides Antimicrobial Protection', *Genome Biology and Evolution*, 6(10), pp. 2754–2773. Available at: <https://doi.org/10.1093/gbe/evu209>.

Erdős, G. and Dosztányi, Z. (2020) 'Analyzing Protein Disorder with IUPred2A', *Current Protocols in Bioinformatics*, 70(1), p. e99. Available at: <https://doi.org/10.1002/CPBI.99>.

Escoll, P. *et al.* (2013) 'From amoeba to macrophages: Exploring the molecular mechanisms of *Legionella pneumophila* infection in both hosts', *Current Topics in Microbiology and Immunology*, 376, pp. 1–34. Available at: <https://doi.org/10.1007/82-2013-351>.

Evans, R. *et al.* (2021) 'Protein complex prediction with AlphaFold-Multimer'. Available at: <https://doi.org/10.1101/2021.10.04.463034>.

ExPASy - ProtParam tool (no date). Available at: <https://web.expasy.org/protparam/> (Accessed: 22 March 2022).

Ferreira, M.L. *et al.* (2014) 'Pseudomonas aeruginosa bacteraemia: independent risk factors for mortality and impact of resistance on outcome', *Journal of Medical Microbiology*, 63(12), pp. 1679–1687. Available at: <https://doi.org/10.1099/jmm.0.073262-0>.

Fleming, A. (1945) 'Penicillin', in. Stockholm: Nobel Lecture, pp. 83–93.

Fothergill, J.L. *et al.* (2007) 'Widespread pyocyanin over-production among isolates of a cystic fibrosis epidemic strain', *BMC microbiology*, 7. Available at: <https://doi.org/10.1186/1471-2180-7-45>.

Foulkes, D.M. *et al.* (2019) 'Pseudomonas aeruginosa toxin ExoU as a therapeutic target in the treatment of bacterial infections', *Microorganisms*. MDPI AG. Available at: <https://doi.org/10.3390/microorganisms7120707>.

Francis, V.I. *et al.* (2018) 'Multiple communication mechanisms between sensor kinases are crucial for virulence in *Pseudomonas aeruginosa*', *Nature Communications*, 9(1), p. 2219. Available at: <https://doi.org/10.1038/s41467-018-04640-8>.

Francis, V.I., Stevenson, E.C. and Porter, S.L. (2017) 'Two-component systems required for virulence in *Pseudomonas aeruginosa*', *FEMS Microbiology Letters*, 364(11), p. 104. Available at: <https://doi.org/10.1093/FEMSLE/FNX104>.

Frank, D.W. (1997) 'The exoenzyme S regulon of *Pseudomonas aeruginosa*', *Molecular microbiology*, 26(4), pp. 621–629. Available at: <https://doi.org/10.1046/J.1365-2958.1997.6251991.X>.

Friesner, R.A. *et al.* (2004) 'Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy', *Journal of medicinal chemistry*, 47(7), pp. 1739–1749. Available at: <https://doi.org/10.1021/JM0306430>.

Fuqua, W.C., Winans, S.C. and Greenberg, E.P. (1994) 'Quorum sensing in bacteria: The LuxR-LuxI family of cell density-responsive transcriptional regulators', *Journal of Bacteriology*. American Society for Microbiology, pp. 269–275. Available at: <https://doi.org/10.1128/jb.176.2.269-275.1994>.

Gaillard, T. (2018) 'Evaluation of AutoDock and AutoDock Vina on the CASF-2013 Benchmark', *Journal of Chemical Information and Modeling*, 58(8), pp. 1697–1706. Available at: <https://doi.org/10.1021/acs.jcim.8b00312>.

Gallagher, J. (2015) *Analysis: Antibiotic apocalypse*, *BBC News*. Available at:

<https://www.bbc.co.uk/news/health-21702647> (Accessed: 24 May 2022).

Gallagher, L.A. and Manoil, C. (2001) 'Pseudomonas aeruginosa PAO1 kills *Caenorhabditis elegans* by cyanide poisoning', *Journal of Bacteriology*, 183(21), pp. 6207–6214. Available at: <https://doi.org/10.1128/JB.183.21.6207-6214.2001/ASSET/F2492791-E0AE-4986-B79B-18551599B436/ASSETS/GRAPHIC/JB2110672004.JPEG>.

Gao, X. *et al.* (2017) 'Structure-Based Prototype Peptides Targeting the *Pseudomonas aeruginosa* Type VI Secretion System Effector as a Novel Antibacterial Strategy', *Frontiers in Cellular and Infection Microbiology*, 7, p. 411. Available at: <https://doi.org/10.3389/fcimb.2017.00411>.

García-Contreras, R. *et al.* (2013) 'Resistance to the quorum-quenching compounds brominated furanone C-30 and 5-fluorouracil in *Pseudomonas aeruginosa* clinical isolates', *Pathogens and Disease*, 68(1), pp. 8–11. Available at: <https://doi.org/10.1111/2049-632X.12039>.

Garrity-Ryan, L.K. *et al.* (2010) 'Small molecule inhibitors of LcrF, a *Yersinia pseudotuberculosis* transcription factor, attenuate virulence and limit infection in a murine pneumonia model', *Infection and Immunity*, 78(11), pp. 4683–4690. Available at: <https://doi.org/10.1128/IAI.01305-09>.

Gasteiger, E. *et al.* (no date) 'Protein Analysis Tools on the ExPASy Server 571 571 From: The Proteomics Protocols Handbook Protein Identification and Analysis Tools on the ExPASy Server'. Available at: <http://www.expasy.org/tools/>. (Accessed: 22 March 2022).

Gatta, V. *et al.* (2019) 'Targeting quorum sensing: High-throughput screening to identify novel lsrk inhibitors', *International Journal of Molecular Sciences*, 20(12). Available at: <https://doi.org/10.3390/ijms20123112>.

Gattiker, A., Gasteiger, E. and Bairoch, A. (2002) 'ScanProsite: a reference implementation of a PROSITE scanning tool', *Applied Bioinformatics*, 1(2), pp. 107–8. Available at: <http://www.expasy.org/prosite> (Accessed: 3 February 2022).

Gessard, C. (1882) 'Sur les colorations bleue et verte des linges à pansements', *C R Acad Sci Serie D*, 94, pp. 536–538.

Glaab, E. (2016) 'Building a virtual ligand screening pipeline using free software: a survey.', *Briefings in bioinformatics*, 17(2), pp. 352–66. Available at: <https://doi.org/10.1093/bib/bbv037>.

Gordillo Altamirano, F.L. and Barr, J.J. (2019) 'Phage therapy in the postantibiotic era', *Clinical Microbiology Reviews*. American Society for Microbiology. Available at: <https://doi.org/10.1128/CMR.00066-18>.

Greenhalgh, J.C. *et al.* (2020) 'Proposed model of the *Dictyostelium* cAMP receptors bound to cAMP', *Journal of molecular graphics & modelling*, 100. Available at: <https://doi.org/10.1016/J.JMGM.2020.107662>.

Greig, S.L. (2016) 'Obiltoximab: First Global Approval', *Drugs*, 76(7), pp. 823–830. Available

at: <https://doi.org/10.1007/s40265-016-0577-0>.

Gulick, A.M. (2017) 'Nonribosomal peptide synthetase biosynthetic clusters of ESKAPE pathogens', *Natural Product Reports*, 34(8), pp. 981–1009. Available at: <https://doi.org/10.1039/C7NP00029D>.

Hajbabaie, R., Harper, M.T. and Rahman, T. (2021) 'Establishing an Analogue Based In Silico Pipeline in the Pursuit of Novel Inhibitory Scaffolds against the SARS Coronavirus 2 Papain-Like Protease', *Molecules*, 26(4). Available at: <https://doi.org/10.3390/MOLECULES26041134>.

Hall, S. *et al.* (2016) 'Cellular Effects of Pyocyanin, a Secreted Virulence Factor of *Pseudomonas aeruginosa*', *Toxins*, 8(8). Available at: <https://doi.org/10.3390/TOXINS8080236>.

Hammond, J.H. *et al.* (2015a) 'Links between Anr and Quorum Sensing in *Pseudomonas aeruginosa* Biofilms', *Journal of Bacteriology*, 197(17), p. 2810. Available at: <https://doi.org/10.1128/JB.00182-15>.

Hammond, J.H. *et al.* (2015b) 'Links between Anr and Quorum Sensing in *Pseudomonas aeruginosa* Biofilms', *Journal of bacteriology*, 197(17), pp. 2810–2820. Available at: <https://doi.org/10.1128/JB.00182-15>.

Hawkins, P.C.D. *et al.* (2010) 'Conformer generation with OMEGA: Algorithm and validation using high quality structures from the protein databank and cambridge structural database', *Journal of Chemical Information and Modeling*, 50(4), pp. 572–584. Available at: <https://doi.org/10.1021/CI100031X>.

Hawkins, P.C.D., Skillman, A.G. and Nicholls, A. (2007) 'Comparison of shape-matching and docking as virtual screening tools', *Journal of Medicinal Chemistry*, 50(1), pp. 74–82. Available at: <https://doi.org/10.1021/jm0603365>.

Hendrix, H. *et al.* (2022) 'Metabolic reprogramming of *Pseudomonas aeruginosa* by phage-based quorum sensing modulation', *Cell Reports*, 38(7). Available at: <https://doi.org/10.1016/J.CELREP.2022.110372/ATTACHMENT/F01A4838-E4B2-4DEA-ACCE-CE3ADC151507/MMC5.XLSX>.

Heo, L., Park, H. and Seok, C. (2013) 'GalaxyRefine: protein structure refinement driven by side-chain repacking', *Nucleic Acids Research*, 41(Web Server issue), p. W384. Available at: <https://doi.org/10.1093/NAR/GKT458>.

Hmelo, L.R. *et al.* (2015) 'Precision-engineering the *Pseudomonas aeruginosa* genome with two-step allelic exchange', *Nature Protocols*, 10(11), pp. 1820–1841. Available at: <https://doi.org/10.1038/nprot.2015.115>.

Hmelo, L.R. *et al.* (no date) *Cloning vector pEX19Gm, complete sequence - Nucleotide - NCBI*. Available at: <https://www.ncbi.nlm.nih.gov/nuccore/KM887142> (Accessed: 26 March 2021).

Hodgkinson, J.T. *et al.* (2012) 'Design, synthesis and biological evaluation of non-natural

modulators of quorum sensing in *Pseudomonas aeruginosa*', *Organic and Biomolecular Chemistry*, 10(30), pp. 6032–6044. Available at: <https://doi.org/10.1039/c2ob25198a>.

Holder, I.A., Neely, A.N. and Frank, D.W. (2001) 'Type III secretion/intoxication system important in virulence of *Pseudomonas aeruginosa* infections in burns', *Burns*, 27(2), pp. 129–130. Available at: [https://doi.org/10.1016/S0305-4179\(00\)00142-X](https://doi.org/10.1016/S0305-4179(00)00142-X).

Hollingsworth, S.A. and Dror, R.O. (2018) 'Molecular Dynamics Simulation for All', *Neuron*, 99(6), pp. 1129–1143. Available at: <https://doi.org/10.1016/J.NEURON.2018.08.011>.

Hong, Z. *et al.* (2019) 'Azetidine-Containing Alkaloids Produced by a Quorum-Sensing Regulated Nonribosomal Peptide Synthetase Pathway in *Pseudomonas aeruginosa*', *Angewandte Chemie (International ed. in English)*, 58(10), pp. 3178–3182. Available at: <https://doi.org/10.1002/ANIE.201809981>.

Hopkins, K.L. *et al.* (2016) 'SPM-1 metallo- β -lactamase-producing *Pseudomonas aeruginosa* ST277 in the UK', *Journal of Medical Microbiology*. Microbiology Society, pp. 696–697. Available at: <https://doi.org/10.1099/jmm.0.000269>.

Hotchkiss, R.S. and Opal, S. (2010) 'Immunotherapy for sepsis--a new approach against an ancient foe.', *The New England journal of medicine*, 363(1), pp. 87–9. Available at: <https://doi.org/10.1056/NEJMcibr1004371>.

Huang, H. *et al.* (2019) 'An integrated genomic regulatory network of virulence-related transcriptional factors in *Pseudomonas aeruginosa*', *Nature Communications*, 10(1). Available at: <https://doi.org/10.1038/S41467-019-10778-W>.

Hulo, N. *et al.* (2006) 'The PROSITE database.', *Nucleic acids research*, 34(Database issue). Available at: <https://doi.org/10.1093/nar/gkj063>.

Hung, D.T. *et al.* (2005) 'Small-molecule inhibitor of *Vibrio cholerae* virulence and intestinal colonization.', *Science (New York, N.Y.)*, 310(5748), pp. 670–4. Available at: <https://doi.org/10.1126/science.1116739>.

Hurt, J.K. *et al.* (2010) 'High-Throughput screening of the virulence regulator VirF: A novel antibacterial target for shigellosis', *Journal of Biomolecular Screening*, 15(4), pp. 379–387. Available at: <https://doi.org/10.1177/1087057110362101>.

Huynh, K. and Partch, C.L. (2015) 'Analysis of protein stability and ligand interactions by thermal shift assay.', *Current protocols in protein science*, 79, pp. 28.9.1–14. Available at: <https://doi.org/10.1002/0471140864.ps2809s79>.

Jain, R. *et al.* (2017) 'KB001-A, a novel anti-inflammatory, found to be safe and well-tolerated in cystic fibrosis patients infected with *Pseudomonas aeruginosa*', *Journal of Cystic Fibrosis* [Preprint]. Available at: <https://doi.org/10.1016/j.jcf.2017.12.006>.

Jeukens, J. *et al.* (2014) 'Comparative genomics of isolates of a *Pseudomonas aeruginosa* epidemic strain associated with chronic lung infections of cystic fibrosis patients.', *Plos one*, 9(2), pp. e87611–e87611. Available at: <https://doi.org/10.1371/JOURNAL.PONE.0087611>.

Jia, S. and Hu, C. (2010) 'Pharmacological Effects of Rutaecarpine as a Cardiovascular Protective Agent', *Molecules* 2010, Vol. 15, Pages 1873-1881, 15(3), pp. 1873–1881. Available at: <https://doi.org/10.3390/MOLECULES15031873>.

Johnson, S. and Gerding, D.N. (2019) 'Bezlotoxumab', *Clinical Infectious Diseases*, pp. 699–704. Available at: <https://doi.org/10.1093/cid/ciy577>.

Jones, G. *et al.* (1997) 'Development and validation of a genetic algorithm for flexible docking', *Journal of Molecular Biology*, 267(3), pp. 727–748. Available at: <https://doi.org/10.1006/JMBI.1996.0897>.

Jorgensen, W.L. *et al.* (1998) 'Comparison of simple potential functions for simulating liquid water', *The Journal of Chemical Physics*, 79(2), p. 926. Available at: <https://doi.org/10.1063/1.445869>.

Jumper, J. *et al.* (2021) 'Highly accurate protein structure prediction with AlphaFold', *Nature* 2021 596:7873, 596(7873), pp. 583–589. Available at: <https://doi.org/10.1038/s41586-021-03819-2>.

Kamath, K.S. *et al.* (2017) 'Pseudomonas aeruginosa proteome under hypoxic stress conditions mimicking the cystic fibrosis lung', *Journal of Proteome Research*, 16(10), pp. 3917–3928. Available at: https://doi.org/10.1021/ACS.JPROTEOME.7B00561/ASSET/IMAGES/LARGE/PR-2017-005618_0003.JPEG.

Kim, Y. and Mylonakis, E. (2011) 'Killing of *Candida albicans* filaments by *Salmonella enterica* serovar Typhimurium is mediated by *sopB* effectors, parts of a type III secretion system.', *Eukaryotic cell*, 10(6), pp. 782–90. Available at: <https://doi.org/10.1128/EC.00014-11>.

Kirchmair, J. *et al.* (2009) 'How to optimize shape-based virtual screening: Choosing the right query and including chemical information', *Journal of Chemical Information and Modeling*, 49(3), pp. 678–692. Available at: <https://doi.org/10.1021/ci8004226>.

Köhler, T. *et al.* (1997) 'Characterization of MexE-MexF-OprN, a positively regulated multidrug efflux system of *Pseudomonas aeruginosa*', *Molecular Microbiology*, 23(2), pp. 345–354. Available at: <https://doi.org/10.1046/j.1365-2958.1997.2281594.x>.

Kon, T. *et al.* (2012) 'The 2.8 Å crystal structure of the dynein motor domain', *Nature* 2012 484:7394, 484(7394), pp. 345–350. Available at: <https://doi.org/10.1038/nature10955>.

Koppolu, V. *et al.* (2013) 'Small-molecule inhibitor of the shigella flexneri master virulence regulator VirF', *Infection and Immunity*. Edited by S.M. Payne, 81(11), pp. 4220–4231. Available at: <https://doi.org/10.1128/IAI.00919-13>.

Kozakov, D. *et al.* (2013) 'How Good is Automated Protein Docking?', *Proteins*, 81(12), pp. 2159–2166. Available at: <https://doi.org/10.1002/prot.24403>.

Kozakov, D. *et al.* (2015) 'NEW FRONTIERS IN DRUGGABILITY', *Journal of medicinal chemistry*, 58(23), p. 9063. Available at: <https://doi.org/10.1021/ACS.JMEDCHEM.5B00586>.

- Kozakov, D. *et al.* (2017) 'The ClusPro web server for protein-protein docking', *Nature Protocols*, 12(2), pp. 255–278. Available at: <https://doi.org/10.1038/nprot.2016.169>.
- Kumar, A. *et al.* (2017) 'Inactivation and changes in metabolic profile of selected foodborne bacteria by 460 nm LED illumination', *Food Microbiology*, 63, pp. 12–21. Available at: <https://doi.org/10.1016/J.FM.2016.10.032>.
- Kumar, S. *et al.* (2018) 'MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms.', *Molecular biology and evolution*. Edited by F.U. Battistuzzi, 35(6), pp. 1547–1549. Available at: <https://doi.org/10.1093/molbev/msy096>.
- Kwon, H.J. *et al.* (2000) 'Crystal structure of the Escherichia coli Rob transcription factor in complex with DNA', *Nature structural biology*, 7(5), pp. 424–430. Available at: <https://doi.org/10.1038/75213>.
- LaBauve, A.E. and Wargo, M.J. (2012) *Growth and Laboratory Maintenance of Pseudomonas aeruginosa*, *Current Protocols in Microbiology*. Available at: <https://doi.org/10.1002/9780471729259.mc06e01s25>.
- LaBauve, A.E. and Wargo, M.J. (2014) 'Detection of host-derived sphingosine by Pseudomonas aeruginosa is important for survival in the murine lung.', *PLoS pathogens*, 10(1), p. e1003889. Available at: <https://doi.org/10.1371/journal.ppat.1003889>.
- Lam, Y.C. and Crawford, J.M. (2018) 'Discovering antibiotics from the global microbiome', *Nature Microbiology*. Nature Publishing Group, pp. 392–393. Available at: <https://doi.org/10.1038/s41564-018-0135-5>.
- Lamont, I.L., Konings, A.F. and Reid, D.W. (2009) 'Iron acquisition by Pseudomonas aeruginosa in the lungs of patients with cystic fibrosis', *BioMetals*, 22(1), pp. 53–60. Available at: <https://doi.org/10.1007/S10534-008-9197-9/FIGURES/1>.
- Lantz, M. (2005) 'A targeted evaluation of OpenEye's methods for virtual ligand screens and docking'.
- LaSarre, B. and Federle, M.J. (2013) 'EMSA Analysis of DNA Binding By Rgg Proteins.', *Bio-protocol*, 3(15). Available at: <https://doi.org/10.21769/BioProtoc.838>.
- Lau, G.W. *et al.* (2004) 'The role of pyocyanin in Pseudomonas aeruginosa infection', *Trends in molecular medicine*, 10(12), pp. 599–606. Available at: <https://doi.org/10.1016/J.MOLMED.2004.10.002>.
- Lee, J. and Zhang, L. (2015) 'The hierarchy quorum sensing network in Pseudomonas aeruginosa', *Protein & Cell*, 6(1), p. 26. Available at: <https://doi.org/10.1007/S13238-014-0100-X>.
- Lee, V.T. *et al.* (2005) 'Activities of Pseudomonas aeruginosa effectors secreted by the type III secretion system in vitro and during infection', *Infection and Immunity*, 73(3), pp. 1695–1705. Available at: <https://doi.org/10.1128/IAI.73.3.1695-1705.2005>.
- Lewis, K. (2010) 'Persister Cells', *Annual Review of Microbiology*, 64(1), pp. 357–372.

Available at: <https://doi.org/10.1146/annurev.micro.112408.134306>.

Li, H., Leung, K.-S. and Wong, M.-H. (2012) 'idock: A multithreaded virtual screening tool for flexible ligand docking', in *2012 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*. IEEE, pp. 77–84. Available at: <https://doi.org/10.1109/CIBCB.2012.6217214>.

Li, Y. *et al.* (2014) 'Comparative Assessment of Scoring Functions on an Updated Benchmark: 2. Evaluation Methods and General Results', *Journal of Chemical Information and Modeling*, 54(6), pp. 1717–1736. Available at: <https://doi.org/10.1021/ci500081m>.

Li, Y. *et al.* (2018) 'Assessing protein-ligand interaction scoring functions with the CASF-2013 benchmark', *Nature Protocols*, 13(4), pp. 666–680. Available at: <https://doi.org/10.1038/nprot.2017.114>.

Liang, H. *et al.* (2008) 'The YebC Family Protein PA0964 Negatively Regulates the *Pseudomonas aeruginosa* Quinolone Signal System and Pyocyanin Production', *JOURNAL OF BACTERIOLOGY*, 190(18), pp. 6217–6227. Available at: <https://doi.org/10.1128/JB.00428-08>.

Liang, H. *et al.* (2014) 'Molecular mechanisms of master regulator VqsM mediating quorum-sensing and antibiotic resistance in *Pseudomonas aeruginosa*.', *Nucleic acids research*, 42(16), pp. 10307–20. Available at: <https://doi.org/10.1093/nar/gku586>.

Lin, C.K. *et al.* (2021) 'A Primed Subpopulation of Bacteria Enables Rapid Expression of the Type 3 Secretion System in *Pseudomonas aeruginosa*', *mBio*, 12(3). Available at: <https://doi.org/10.1128/MBIO.00831-21>.

LIPSKA, I. (1951) 'Dotychczasowe spostrzeżenia nad leczniczym i zapobiegawczym stosowaniem', *Pediatrics polska*, 26(4), pp. 449–453.

Little, A.S. *et al.* (2018) 'Pseudomonas aeruginosa AlgR Phosphorylation Status Differentially Regulates Pyocyanin and Pyoverdine Production', *mBio*, 9(1). Available at: <https://doi.org/10.1128/MBIO.02318-17>.

Liu, Y. *et al.* (2019) 'Cinnamaldehyde inhibits type three secretion system in *Salmonella enterica* serovar Typhimurium by affecting the expression of key effector proteins', *Veterinary Microbiology*, 239. Available at: <https://doi.org/10.1016/j.vetmic.2019.108463>.

Lowden, M.J. *et al.* (2010) 'Structure of *Vibrio cholerae* ToxT reveals a mechanism for fatty acid regulation of virulence genes', *Proceedings of the National Academy of Sciences*, 107(7), pp. 2860–2865. Available at: <https://doi.org/10.1073/pnas.0915021107>.

Lu, C. *et al.* (2021) 'OPLS4: Improving Force Field Accuracy on Challenging Regimes of Chemical Space', *Journal of chemical theory and computation*, 17(7), pp. 4291–4300. Available at: <https://doi.org/10.1021/ACS.JCTC.1C00302>.

Lu, C. *et al.* (2022) 'A novel antimicrobial peptide found in *Pelophylax nigromaculatus*', *Journal, genetic engineering & biotechnology*, 20(1), p. 76. Available at:

<https://doi.org/10.1186/S43141-022-00366-9>.

Lu, D. *et al.* (2014) 'The structural basis of the Tle4–Tli4 complex reveals the self-protection mechanism of H2-T6SS in *Pseudomonas aeruginosa*', *Acta Crystallographica Section D Biological Crystallography*, 70(12), pp. 3233–3243. Available at: <https://doi.org/10.1107/S1399004714023967>.

Lu, J. *et al.* (2016) 'Molecular Detection of Legionella spp. and their associations with Mycobacterium spp., Pseudomonas aeruginosa and amoeba hosts in a drinking water distribution system', *Journal of Applied Microbiology*, 120(2), pp. 509–521. Available at: <https://doi.org/10.1111/jam.12996>.

Ma, Y.N. *et al.* (2019) 'Identification of Benzyloxy Carbonimidoyl Dicyanide Derivatives as Novel Type III Secretion System Inhibitors via High-Throughput Screening', *Frontiers in Plant Science*, 10. Available at: <https://doi.org/10.3389/fpls.2019.01059>.

Majerczyk, C.D. *et al.* (2014) 'Cross-species comparison of the Burkholderia pseudomallei, Burkholderia thailandensis, and Burkholderia mallei quorum-sensing regulons', *Journal of Bacteriology*, 196(22), pp. 3862–3871. Available at: <https://doi.org/10.1128/JB.01974-14>.

Manavathu, E.K., Vager, D.L. and Vazquez, J.A. (2014) 'Development and antimicrobial susceptibility studies of in vitro monomicrobial and polymicrobial biofilm models with Aspergillus fumigatus and Pseudomonas aeruginosa.', *BMC microbiology*, 14, p. 53. Available at: <https://doi.org/10.1186/1471-2180-14-53>.

Marsden, A.E., King, J.M., *et al.* (2016) 'Inhibition of Pseudomonas aeruginosa ExsA DNA-binding activity by N-hydroxybenzimidazoles', *Antimicrobial Agents and Chemotherapy*, 60(2), pp. 766–776. Available at: <https://doi.org/10.1128/AAC.02242-15>.

Marsden, A.E., Intile, P.J., *et al.* (2016) 'Vfr Directly Activates exsA Transcription To Regulate Expression of the Pseudomonas aeruginosa Type III Secretion System', *Journal of bacteriology*, 198(9), pp. 1442–1450. Available at: <https://doi.org/10.1128/JB.00049-16>.

Marsden, A.E., Schubot, F.D. and Yahr, T.L. (2014) 'Self-association is required for occupation of adjacent binding sites in Pseudomonas aeruginosa type III secretion system promoters.', *Journal of bacteriology*, 196(20), pp. 3546–55. Available at: <https://doi.org/10.1128/JB.01969-14>.

Matsoukas, M.T. *et al.* (2015) 'Identification of small-molecule inhibitors of calcineurin-NFATc signaling that mimic the PxlIT motif of Calcineurin binding partners', *Science Signaling*, 8(382), pp. ra63–ra63. Available at: <https://doi.org/10.1126/scisignal.2005918>.

Maura, D. *et al.* (2016) 'Evidence for Direct Control of Virulence and Defense Gene Circuits by the Pseudomonas aeruginosa Quorum Sensing Regulator, MvfR', *Scientific Reports*, 6. Available at: <https://doi.org/10.1038/SREP34083>.

McGann, M. (2011) 'FRED pose prediction and virtual screening accuracy', *Journal of Chemical Information and Modeling*, 51(3), pp. 578–596. Available at:

https://doi.org/10.1021/CI100436P/SUPPL_FILE/CI100436P_SI_001.PDF.

Medarametla, P. *et al.* (2021) 'Structural Characterization of LsrK as a Quorum Sensing Target and a Comparison between X-ray and Homology Models', *Journal of chemical information and modeling*, 61(3), pp. 1346–1353. Available at: <https://doi.org/10.1021/ACS.JCIM.0C01233>.

Medina-Rojas, M. *et al.* (2020) 'Comparison of *Pseudomonas aeruginosa* strains reveals that Exolysin A toxin plays an additive role in virulence', *Pathogens and disease*, 78(1). Available at: <https://doi.org/10.1093/femspd/ftaa010>.

Meijvis, S.C.A. *et al.* (2012) 'Treatment with anti-inflammatory drugs in community-acquired pneumonia', *Journal of Internal Medicine*, 272(1), pp. 25–35. Available at: <https://doi.org/10.1111/J.1365-2796.2012.02554.X>.

von Mering, C. *et al.* (2005) 'STRING: known and predicted protein-protein associations, integrated and transferred across organisms', *Nucleic acids research*, 33(Database issue). Available at: <https://doi.org/10.1093/NAR/GKI005>.

Mészáros, B., Erdős, G. and Dosztányi, Z. (2018) 'IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding', *Nucleic Acids Research*, 46(W1), pp. W329–W337. Available at: <https://doi.org/10.1093/NAR/GKY384>.

Meyer, J.M. *et al.* (1996) 'Pyoverdinin is essential for virulence of *Pseudomonas aeruginosa*', *Infection and immunity*, 64(2), pp. 518–523. Available at: <https://doi.org/10.1128/IAI.64.2.518-523.1996>.

Micek, S.T. *et al.* (2015) 'An international multicenter retrospective study of *Pseudomonas aeruginosa* nosocomial pneumonia: impact of multidrug resistance.', *Critical care (London, England)*, 19(1), p. 219. Available at: <https://doi.org/10.1186/s13054-015-0926-5>.

Mikkelsen, H., McMullan, R. and Filloux, A. (2011) 'The *Pseudomonas aeruginosa* reference strain PA14 displays increased virulence due to a mutation in *ladS*', *PloS one*, 6(12). Available at: <https://doi.org/10.1371/JOURNAL.PONE.0029113>.

Mishra, N. and Basu, A. (2013) 'Exploring different virtual screening strategies for acetylcholinesterase inhibitors', *BioMed Research International*, 2013. Available at: <https://doi.org/10.1155/2013/236850>.

Mitchell, C.A. *et al.* (2012) 'Structure of PA1221, a nonribosomal peptide synthetase containing adenylation and peptidyl carrier protein domains', *Biochemistry*, 51(15), pp. 3252–3263. Available at: https://doi.org/10.1021/BI300112E/SUPPL_FILE/BI300112E_SI_001.PDF.

Mohanam, L. and Menon, T. (2017) 'Coexistence of metallo-beta-lactamase-encoding genes in *Pseudomonas aeruginosa*', *Indian Journal of Medical Research, Supplement*, 146(Suppl 1), pp. 46–52. Available at: https://doi.org/10.4103/ijmr.IJMR_29_16.

Mokbel, S.A. *et al.* (2020) 'Synthesis of novel 1,2-diarylpyrazolidin-3-one-based compounds and their evaluation as broad spectrum antibacterial agents', *Bioorganic Chemistry*, p.

103759. Available at: <https://doi.org/10.1016/j.bioorg.2020.103759>.

Moon, T.C. *et al.* (1999) 'A new class of COX-2 inhibitor, rutaecarpine from *Evodia rutaecarpa*', *Inflammation research : official journal of the European Histamine Research Society ... [et al.]*, 48(12), pp. 621–625. Available at: <https://doi.org/10.1007/S000110050512>.

Morkunas, B. *et al.* (2012) 'Inhibition of the production of the *Pseudomonas aeruginosa* virulence factor pyocyanin in wild-type cells by quorum sensing autoinducer-mimics', *Organic and Biomolecular Chemistry*, 10(42), pp. 8452–8464. Available at: <https://doi.org/10.1039/c2ob26501j>.

Morris, G.M. *et al.* (2009) 'AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility.', *Journal of computational chemistry*, 30(16), pp. 2785–91. Available at: <https://doi.org/10.1002/jcc.21256>.

Mulcahy, L.R., Isabella, V.M. and Lewis, K. (2014) '*Pseudomonas aeruginosa* biofilms in disease.', *Microbial ecology*, 68(1), pp. 1–12. Available at: <https://doi.org/10.1007/s00248-013-0297-x>.

Murray, C.J. *et al.* (2022) 'Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis', *The Lancet*, 399(10325), pp. 629–655. Available at: [https://doi.org/10.1016/S0140-6736\(21\)02724-0/ATTACHMENT/B227DEB3-FF04-497F-82AC-637D8AB7F679/MMC1.PDF](https://doi.org/10.1016/S0140-6736(21)02724-0/ATTACHMENT/B227DEB3-FF04-497F-82AC-637D8AB7F679/MMC1.PDF).

Murray, T.S., Egan, M. and Kazmierczak, B.I. (2007) '*Pseudomonas aeruginosa* chronic colonization in cystic fibrosis patients', *Current Opinion in Pediatrics*, 19(1), pp. 83–88. Available at: <https://doi.org/10.1097/MOP.0b013e3280123a5d>.

Nair, A. V *et al.* (2015) 'A comparative study of coastal and clinical isolates of *Pseudomonas aeruginosa*.', *Brazilian journal of microbiology: [publication of the Brazilian Society for Microbiology]*, 46(3), pp. 725–34. Available at: <https://doi.org/10.1590/S1517-838246320140502>.

Nakano, M. *et al.* (1997) 'Mutations in the *gyrA* and *parC* genes in fluoroquinolone-resistant clinical isolates of *Pseudomonas aeruginosa*', *Antimicrobial Agents and Chemotherapy*, 41(10), pp. 2289–2291. Available at: <https://doi.org/10.1128/aac.41.10.2289>.

National Library of Medicine (US) and National Center for Biotechnology Information (1988) *National Center for Biotechnology Information (NCBI)[Internet]*. Available at: <https://www.ncbi.nlm.nih.gov/> (Accessed: 26 January 2022).

Naylor, E. *et al.* (2009) 'Identification of a chemical probe for NAADP by virtual screening', *Nature chemical biology*, 5(4), pp. 220–226. Available at: <https://doi.org/10.1038/NCHEMBIO.150>.

Nazir, R. *et al.* (2017) 'The Ecological Role of Type Three Secretion Systems in the Interaction of Bacteria with Fungi in Soil and Related Habitats Is Diverse and Context-Dependent.', *Frontiers in microbiology*, 8, p. 38. Available at: <https://doi.org/10.3389/fmicb.2017.00038>.

- Neill, J. (2016) 'Tackling Drug-Resistance Infections Globally: Final Report and Recommendations. The review on Antimicrobial Resistance Chaired by Jim O'Neil'. [Online ...], pp. 1–84. Available at: [https://amr-review.org/sites/default/files/160518_Final paper_with cover.pdf](https://amr-review.org/sites/default/files/160518_Final%20paper_with%20cover.pdf) (Accessed: 2 April 2020).
- Newman, J. *et al.* (2018) 'Structural characterization of a novel monotreme-specific protein with antimicrobial activity from the milk of the platypus', *Acta Crystallographica Section F Structural Biology Communications*, 74(1), pp. 39–45. Available at: <https://doi.org/10.1107/S2053230X17017708>.
- Ng, P.C. and Henikoff, S. (2001) 'Predicting deleterious amino acid substitutions', *Genome research*, 11(5), pp. 863–874. Available at: <https://doi.org/10.1101/GR.176601>.
- Ngan, C.-H. *et al.* (2012) 'FTSite: high accuracy detection of ligand binding sites on unbound protein structures.', *Bioinformatics (Oxford, England)*, 28(2), pp. 286–7. Available at: <https://doi.org/10.1093/bioinformatics/btr651>.
- Nguyen, V.S. *et al.* (2015) 'Inhibition of type VI secretion by an anti-TssM llama nanobody', *PLoS ONE*, 10(3), p. e0122187. Available at: <https://doi.org/10.1371/journal.pone.0122187>.
- NICOLLE, P. (1952) 'Les applications du bactériophage', *Gazette médicale de France*, 58(22), pp. 1319–1321.
- Nougayrède, J.P. *et al.* (2006) 'Escherichia coli induces DNA double-strand breaks in eukaryotic cells', *Science (New York, N.Y.)*, 313(5788), pp. 848–851. Available at: <https://doi.org/10.1126/SCIENCE.1127059>.
- O'Brien, S. *et al.* (2017) 'Adaptation to public goods cheats in *Pseudomonas aeruginosa*', *Proceedings of the Royal Society B: Biological Sciences*, 284(1859). Available at: <https://doi.org/10.1098/rspb.2017.1089>.
- O'Brien, T.J. and Welch, M. (2019) 'A Continuous-Flow Model for in vitro Cultivation of Mixed Microbial Populations Associated With Cystic Fibrosis Airway Infections', *Frontiers in Microbiology*, 10. Available at: <https://doi.org/10.3389/FMICB.2019.02713/FULL>.
- Okino, N. and Ito, M. (2016) 'Molecular mechanism for sphingosine-induced *Pseudomonas* ceramidase expression through the transcriptional regulator SphR.', *Scientific reports*, 6, p. 38797. Available at: <https://doi.org/10.1038/srep38797>.
- O'Loughlin, C.T. *et al.* (2013) 'A quorum-sensing inhibitor blocks *Pseudomonas aeruginosa* virulence and biofilm formation', *Proceedings of the National Academy of Sciences of the United States of America*, 110(44), pp. 17981–17986. Available at: <https://doi.org/10.1073/PNAS.1316981110/-/DCSUPPLEMENTAL>.
- Olson, M. V. *et al.* (2000) 'Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen.', *Nature*, 406(6799), pp. 959–964. Available at: <https://doi.org/10.1038/35023079>.
- Onaka, H. (2017) 'Novel antibiotic screening methods to awaken silent or cryptic secondary

metabolic pathways in actinomycetes', *Journal of Antibiotics*, 70(8), pp. 865–870. Available at: <https://doi.org/10.1038/ja.2017.51>.

Parfitt, T. (2005) 'Georgia: an unlikely stronghold for bacteriophage therapy', *The Lancet*, 365(9478), pp. 2166–2167. Available at: [https://doi.org/10.1016/S0140-6736\(05\)66759-1](https://doi.org/10.1016/S0140-6736(05)66759-1).

Park, D. *et al.* (2018) 'Visualization of the type III secretion mediated salmonella–host cell interface using cryo-electron tomography', *eLife*, 7. Available at: <https://doi.org/10.7554/eLife.39514>.

Payne, S.M. (1994) 'Detection, isolation, and characterization of siderophores', *Methods in Enzymology*, 235(C), pp. 329–344. Available at: [https://doi.org/10.1016/0076-6879\(94\)35151-1](https://doi.org/10.1016/0076-6879(94)35151-1).

Pena, R.T. *et al.* (2019) 'Relationship Between Quorum Sensing and Secretion Systems', *Frontiers in Microbiology*, 10(JUN). Available at: <https://doi.org/10.3389/FMICB.2019.01100>.

Pier, G.B. (2007) 'Pseudomonas aeruginosa lipopolysaccharide: a major virulence factor, initiator of inflammation and target for effective immunity', *International journal of medical microbiology : IJMM*, 297(5), p. 277. Available at: <https://doi.org/10.1016/J.IJMM.2007.03.012>.

Pirnay, J.-P. *et al.* (2005) 'Global Pseudomonas aeruginosa biodiversity as reflected in a Belgian river', *Environmental Microbiology*, 7(7), pp. 969–980. Available at: <https://doi.org/10.1111/j.1462-2920.2005.00776.x>.

Plumbridge, J. and Pellegrini, O. (2004) 'Expression of the chitobiose operon of Escherichia coli is regulated by three transcription factors: NagC, ChbR and CAP', *Molecular Microbiology*, 52(2), pp. 437–449. Available at: <https://doi.org/10.1111/j.1365-2958.2004.03986.x>.

Pollock, J. *et al.* (2020) 'Alternatives to antibiotics in a One Health context and the role genomics can play in reducing antimicrobial use', *Clinical Microbiology and Infection* [Preprint]. Available at: <https://doi.org/10.1016/j.cmi.2020.02.028>.

Poole, K. *et al.* (1993) 'Multiple antibiotic resistance in Pseudomonas aeruginosa: Evidence for involvement of an efflux operon', *Journal of Bacteriology*, 175(22), pp. 7363–7372. Available at: <https://doi.org/10.1128/jb.175.22.7363-7372.1993>.

Poole, K. *et al.* (1996) 'Overexpression of the mexC-mexD-oprJ efflux operon in nfxB -type multidrug-resistant strains of Pseudomonas aeruginosa', *Molecular Microbiology*, 21(4), pp. 713–725. Available at: <https://doi.org/10.1046/j.1365-2958.1996.281397.x>.

Price, C.T.D. *et al.* (2014) 'Amoeba host-Legionella synchronization of amino acid auxotrophy and its role in bacterial adaptation and pathogenic evolution', *Environmental Microbiology*. NIH Public Access, pp. 350–358. Available at: <https://doi.org/10.1111/1462-2920.12290>.

Pritt, B., O'Brien, L. and Winn, W. (2007) 'Mucoicid Pseudomonas in cystic fibrosis', *American Journal of Clinical Pathology*, 128(1), pp. 32–34. Available at: <https://doi.org/10.1309/KJRPC7DD5TR9NTDM>.

- Projan, S.J. (2003) 'Why is big Pharma getting out of antibacterial drug discovery?', *Current Opinion in Microbiology*. Elsevier Ltd, pp. 427–430. Available at: <https://doi.org/10.1016/j.mib.2003.08.003>.
- Propst, K.L. *et al.* (2010) 'Immunotherapy markedly increases the effectiveness of antimicrobial therapy for treatment of Burkholderia pseudomallei infection.', *Antimicrobial agents and chemotherapy*, 54(5), pp. 1785–92. Available at: <https://doi.org/10.1128/AAC.01513-09>.
- Qu, J. *et al.* (2019) 'Modulating Pathogenesis with Mobile-CRISPRi', *Journal of Bacteriology*, 201(22). Available at: <https://doi.org/10.1128/JB.00304-19>.
- Qu, J. *et al.* (2021) 'Persistent Bacterial Coinfection of a COVID-19 Patient Caused by a Genetically Adapted Pseudomonas aeruginosa Chronic Colonizer', *Frontiers in cellular and infection microbiology*, 11. Available at: <https://doi.org/10.3389/FCIMB.2021.641920>.
- Rahman, S. and Rahman, T. (2017) 'Unveiling some FDA-approved drugs as inhibitors of the store-operated Ca²⁺ entry pathway', *Scientific Reports 2017 7:1*, 7(1), pp. 1–13. Available at: <https://doi.org/10.1038/s41598-017-13343-x>.
- Rasko, D.A. *et al.* (2008) 'Targeting QseC signaling and virulence for antibiotic development', *Science*, 321(5892), pp. 1078–1080. Available at: <https://doi.org/10.1126/science.1160354>.
- Rasko, D.A. and Sperandio, V. (2010) 'Anti-virulence strategies to combat bacteria-mediated disease', *Nature Reviews Drug Discovery*, 9(2), pp. 117–128. Available at: <https://doi.org/10.1038/nrd3013>.
- Rietsch, A. *et al.* (2005) 'ExsE, a secreted regulator of type III secretion genes in Pseudomonas aeruginosa', *Proceedings of the National Academy of Sciences of the United States of America*, 102(22), pp. 8006–8011. Available at: <https://doi.org/10.1073/pnas.0503005102>.
- Rohilla, A., Khare, G. and Tyagi, A.K. (2017) 'Virtual Screening, pharmacophore development and structure based similarity search to identify inhibitors against IdeR, a transcription factor of Mycobacterium tuberculosis', *Scientific Reports*, 7(1). Available at: <https://doi.org/10.1038/s41598-017-04748-9>.
- Rooks, M.G. *et al.* (2017) 'QseC inhibition as an antivirulence approach for colitis-associated bacteria', *Proceedings of the National Academy of Sciences*, 114(1), pp. 142–147. Available at: <https://doi.org/10.1073/pnas.1612836114>.
- Rutherford, S.T. and Bassler, B.L. (2012) 'Bacterial quorum sensing: its role in virulence and possibilities for its control.', *Cold Spring Harbor perspectives in medicine*, 2(11), p. a012427. Available at: <https://doi.org/10.1101/cshperspect.a012427>.
- Saha, R. *et al.* (2013) 'Microbial siderophores: A mini review', *Journal of Basic Microbiology*, 53(4), pp. 303–317. Available at: <https://doi.org/10.1002/jobm.201100552>.
- Salah, I.B., Ghigo, E. and Drancourt, M. (2009) 'Free-living amoebae, a training field for

macrophage resistance of mycobacteria', *Clinical Microbiology and Infection*, pp. 894–905. Available at: <https://doi.org/10.1111/j.1469-0691.2009.03011.x>.

Salomon-Ferrer, R., Case, D.A. and Walker, R.C. (2013) 'An overview of the Amber biomolecular simulation package', *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 3(2), pp. 198–210. Available at: <https://doi.org/10.1002/WCMS.1121>.

Salunkhe, P. *et al.* (2005) 'A cystic fibrosis epidemic strain of *Pseudomonas aeruginosa* displays enhanced virulence and antimicrobial resistance', *Journal of Bacteriology*, 187(14). Available at: <https://doi.org/10.1128/JB.187.14.4908-4920.2005>.

Sample, I. (2018) *Rise in antibiotic resistance must be tackled, says top medic*, *The Guardian*. Available at: <https://www.theguardian.com/society/2018/sep/04/rise-in-antibiotic-resistance-must-be-tackled-says-top-medic> (Accessed: 24 May 2022).

Sanderson, K. *et al.* (2008) 'Bacterial cyanogenesis occurs in the cystic fibrosis lung', *European Respiratory Journal*, 32(2), pp. 329–333. Available at: <https://doi.org/10.1183/09031936.00152407>.

Sass, G. *et al.* (2017) 'Studies of *Pseudomonas aeruginosa* Mutants Indicate Pyoverdine as the Central Factor in Inhibition of *Aspergillus fumigatus* Biofilm', *Journal of bacteriology*, 200(1). Available at: <https://doi.org/10.1128/JB.00345-17>.

Sass, G. *et al.* (2021) 'Under nonlimiting iron conditions pyocyanin is a major antifungal molecule, and differences between prototypic *Pseudomonas aeruginosa* strains', *Medical Mycology*, 59(5), pp. 453–464. Available at: <https://doi.org/10.1093/MMY/MYAA066>.

Sawa, T. *et al.* (2014) 'Anti-PcrV antibody strategies against virulent *Pseudomonas aeruginosa*.', *Human vaccines & immunotherapeutics*, 10(10), pp. 2843–52. Available at: <https://doi.org/10.4161/21645515.2014.971641>.

Schäper, S. *et al.* (2017) 'AraC-like transcriptional activator CuxR binds c-di-GMP by a PilZ-like mechanism to regulate extracellular polysaccharide production', *Proceedings of the National Academy of Sciences of the United States of America*, 114(24), pp. E4822–E4831. Available at: <https://doi.org/10.1073/PNAS.1702435114/-/DCSUPPLEMENTAL>.

Schmidt, S. *et al.* (2018) 'Natural killer cells as a therapeutic tool for infectious diseases - current status and future perspectives', *Oncotarget*, 9(29), pp. 20891–20907. Available at: <https://doi.org/10.18632/oncotarget.25058>.

Schuster, M. *et al.* (2003) 'Identification, timing, and signal specificity of *Pseudomonas aeruginosa* quorum-controlled genes: a transcriptome analysis', *Journal of bacteriology*, 185(7), pp. 2066–2079. Available at: <https://doi.org/10.1128/JB.185.7.2066-2079.2003>.

Shakhnovich, E.A. *et al.* (2007) 'Virstatin inhibits dimerization of the transcriptional activator ToxT.', *Proceedings of the National Academy of Sciences of the United States of America*, 104(7), pp. 2372–7. Available at: <https://doi.org/10.1073/pnas.0611643104>.

Sharifahmadian, M. and Baron, C. (2017) 'Type IV secretion in *agrobacterium tumefaciens*

and development of specific inhibitors', in *Current Topics in Microbiology and Immunology*. Springer Verlag, pp. 169–186. Available at: https://doi.org/10.1007/978-3-319-75241-9_7.

Shaver, C.M. and Hauser, A.R. (2004) 'Relative Contributions of *Pseudomonas aeruginosa* ExoU, ExoS, and ExoT to Virulence in the Lung', *Infection and Immunity*, 72(12), pp. 6969–6977. Available at: <https://doi.org/10.1128/IAI.72.12.6969-6977.2004>.

Sheremet, A.B. *et al.* (2018) 'Small Molecule Inhibitor of Type Three Secretion System Belonging to a Class 2,4-disubstituted-4H-[1,3,4]-thiadiazine-5-ones Improves Survival and Decreases Bacterial Loads in an Airway *Pseudomonas aeruginosa* Infection in Mice.', *BioMed research international*, 2018, p. 5810767. Available at: <https://doi.org/10.1155/2018/5810767>.

Shi, N. *et al.* (2017) 'Interaction between the gut microbiome and mucosal immune system', *Military Medical Research*. BioMed Central Ltd., p. 14. Available at: <https://doi.org/10.1186/s40779-017-0122-9>.

Shlaes, D.M. (2003) 'The abandonment of antibacterials: Why and wherefore?', *Current Opinion in Pharmacology*. Elsevier, pp. 470–473. Available at: <https://doi.org/10.1016/j.coph.2003.04.003>.

Shrestha, M. *et al.* (2015) 'Structural analysis of the regulatory domain of ExsA, a key transcriptional regulator of the type three secretion system in *Pseudomonas aeruginosa*', *PLoS ONE*, 10(8), p. e0136533. Available at: <https://doi.org/10.1371/journal.pone.0136533>.

Shrestha, M. *et al.* (2020) 'Backbone Interactions Between Transcriptional Activator ExsA and Anti-Activator ExsD Facilitate Regulation of the Type III Secretion System in *Pseudomonas aeruginosa*', *Scientific reports*, 10(1). Available at: <https://doi.org/10.1038/S41598-020-66555-Z>.

SIBIRTSEV, G.E., BEL'SKAIA, T.G. and LAVROVA, K. v. (1953) 'Primenenie spetsificheskogo bakteriofaga dlia lechenie differiinykh bol'nykh.', *Pediatriia*, 2, pp. 22–23.

Smith, E.E. *et al.* (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), pp. 8487–8492. Available at: <https://doi.org/10.1073/PNAS.0602138103>.

Smith, M.A. *et al.* (2012) 'Identification of the Binding Site of *Brucella* VirB8 Interaction Inhibitors', *Chemistry & Biology*, 19(8), pp. 1041–1048. Available at: <https://doi.org/10.1016/j.chembiol.2012.07.007>.

Smits, S.H.J. *et al.* (2008) 'Coenzyme- and His-tag-induced crystallization of octopine dehydrogenase', *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, 64(Pt 9), p. 836. Available at: <https://doi.org/10.1107/S1744309108025487>.

Soisson, S.M. *et al.* (1997) 'Structural basis for ligand-regulated oligomerization of AraC', *Science*, 276(5311), pp. 421–425. Available at: <https://doi.org/10.1126/SCIENCE.276.5311.421>.

- Sonnleitner, E. *et al.* (2011) 'The small RNA PhrS stimulates synthesis of the *Pseudomonas aeruginosa* quinolone signal', *Molecular Microbiology*, 80(4), pp. 868–885. Available at: <https://doi.org/10.1111/J.1365-2958.2011.07620.X>.
- Stebbing, J. *et al.* (2020) 'COVID-19: combining antiviral and anti-inflammatory treatments', *The Lancet. Infectious diseases*, 20(4), pp. 400–402. Available at: [https://doi.org/10.1016/S1473-3099\(20\)30132-8](https://doi.org/10.1016/S1473-3099(20)30132-8).
- Stierand, K. and Rarey, M. (2010) 'PoseView -- molecular interaction patterns at a glance', *Journal of Cheminformatics*, 2(S1). Available at: <https://doi.org/10.1186/1758-2946-2-s1-p50>.
- Stilwell, P., Lowe, C. and Buckling, A. (2018) 'The effect of cheats on siderophore diversity in *Pseudomonas aeruginosa*', *Journal of evolutionary biology*, 31(9), pp. 1330–1339. Available at: <https://doi.org/10.1111/JEB.13307>.
- Stintzi, A. *et al.* (1998) 'Quorum-sensing and siderophore biosynthesis in *Pseudomonas aeruginosa*: lasR/lasI mutants exhibit reduced pyoverdine biosynthesis', *FEMS microbiology letters*, 166(2), pp. 341–345. Available at: <https://doi.org/10.1111/J.1574-6968.1998.TB13910.X>.
- Sun, K. *et al.* (2014) 'Screening for inhibition of *Vibrio cholerae* VipA-VipB interaction identifies small-molecule compounds active against type VI secretion', *Antimicrobial Agents and Chemotherapy*, 58(7), pp. 4123–4130. Available at: <https://doi.org/10.1128/AAC.02819-13>.
- Szklarczyk, D. *et al.* (2021) 'The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets', *Nucleic acids research*, 49(D1), pp. D605–D612. Available at: <https://doi.org/10.1093/NAR/GKAA1074>.
- Tabor, D.E. *et al.* (2018) '*Pseudomonas aeruginosa* PcrV and Psl, the Molecular Targets of Bispecific Antibody MEDI3902, Are Conserved Among Diverse Global Clinical Isolates', *The Journal of Infectious Diseases*, 218(12), pp. 1983–1994. Available at: <https://doi.org/10.1093/infdis/jiy438>.
- Takase, H. *et al.* (2000) 'Impact of siderophore production on *Pseudomonas aeruginosa* infections in immunosuppressed mice', *Infection and immunity*, 68(4), pp. 1834–1839. Available at: <https://doi.org/10.1128/IAI.68.4.1834-1839.2000>.
- Taxonomy browser (Bacteria Latreille et al. 1825)* (no date). Available at: <https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=629395&lvl=0> (Accessed: 10 March 2022).
- Taylor, P.L. *et al.* (2008) 'Structure and function of sedoheptulose-7-phosphate isomerase, a critical enzyme for lipopolysaccharide biosynthesis and a target for antibiotic adjuvants', *The Journal of biological chemistry*, 283(5), pp. 2835–2845. Available at: <https://doi.org/10.1074/JBC.M706163200>.
- Tetard, A. *et al.* (2019) 'Cinnamaldehyde induces expression of efflux pumps and multidrug

resistance in pseudomonas aeruginosa', *Antimicrobial Agents and Chemotherapy*, 63(10). Available at: <https://doi.org/10.1128/AAC.01081-19>.

Thees, A. v. *et al.* (2021) 'PmtA Regulates Pyocyanin Expression and Biofilm Formation in Pseudomonas aeruginosa', *Frontiers in microbiology*, 12. Available at: <https://doi.org/10.3389/FMICB.2021.789765>.

Thibault, J. *et al.* (2009) 'Anti-activator ExsD forms a 1:1 complex with ExsA to inhibit transcription of type III secretion operons.', *The Journal of biological chemistry*, 284(23), pp. 15762–70. Available at: <https://doi.org/10.1074/jbc.M109.003533>.

Thompson, S.E. *et al.* (2001) 'Identification and characterization of a chitinase antigen from Pseudomonas aeruginosa strain 385.', *Applied and environmental microbiology*, 67(9), pp. 4001–8. Available at: <https://doi.org/10.1128/AEM.67.9.4001-4008.2001>.

Tian, Z. *et al.* (2019) 'Pseudomonas aeruginosa ExsA Regulates a Metalloprotease, ImpA, That Inhibits Phagocytosis of Macrophages', *Infection and Immunity*. Edited by A.J. Bäuml, 87(12). Available at: <https://doi.org/10.1128/IAI.00695-19>.

Tkaczyk, C. *et al.* (2018) 'Alanine scanning mutagenesis of the MEDI4893 (Suvratoxumab) epitope reduces alpha toxin lytic activity in vitro and staphylococcus aureus fitness in infection models', *Antimicrobial Agents and Chemotherapy*, 62(11). Available at: <https://doi.org/10.1128/AAC.01033-18>.

Trott, O. and Olson, A.J. (2009) 'AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading', *Journal of Computational Chemistry*, 31(2), p. NA-NA. Available at: <https://doi.org/10.1002/jcc.21334>.

Turner, K.H. *et al.* (2014) 'Requirements for Pseudomonas aeruginosa acute burn and chronic surgical wound infection.', *PLoS genetics*, 10(7), p. e1004518. Available at: <https://doi.org/10.1371/journal.pgen.1004518>.

Upadhyay, J., Gajjar, A. and Suhagia, B.N. (2018) 'Combined Ligand-Based and Structure-Based Virtual Screening Approach for Identification of New Dipeptidyl Peptidase 4 Inhibitors', *Current Drug Discovery Technologies*, 16(4), pp. 426–436. Available at: <https://doi.org/10.2174/1570163815666180926111558>.

Urbanowski, M.L., Brutinel, E.D. and Yahr, T.L. (2007) 'Translocation of ExsE into Chinese Hamster Ovary Cells Is Required for Transcriptional Induction of the Pseudomonas aeruginosa Type III Secretion System', *Infection and Immunity*, 75(9), pp. 4432–4439. Available at: <https://doi.org/10.1128/IAI.00664-07>.

Uyttebroek, S. *et al.* (2022) 'Safety and efficacy of phage therapy in difficult-to-treat infections: a systematic review', *The Lancet Infectious Diseases* [Preprint]. Available at: [https://doi.org/10.1016/S1473-3099\(21\)00612-5](https://doi.org/10.1016/S1473-3099(21)00612-5).

Vajda, S. *et al.* (2017) 'New Additions to the ClusPro Server Motivated by CAPRI', *Proteins*, 85(3), pp. 435–444. Available at: <https://doi.org/10.1002/prot.25219>.

- Valentini, M. *et al.* (2018) 'Lifestyle transitions and adaptive pathogenesis of *Pseudomonas aeruginosa*', *Current opinion in microbiology*, 41, pp. 15–20. Available at: <https://doi.org/10.1016/J.MIB.2017.11.006>.
- Vance, R.E., Rietsch, A. and Mekalanos, J.J. (2005) 'Role of the type III secreted exoenzymes S, T, and Y in systemic spread of *Pseudomonas aeruginosa* PAO1 in vivo', *Infection and Immunity*, 73(3), pp. 1706–1713. Available at: <https://doi.org/10.1128/IAI.73.3.1706-1713.2005>.
- Varadi, M. *et al.* (2022) 'AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models', *Nucleic Acids Research*, 50(D1), pp. D439–D444. Available at: <https://doi.org/10.1093/NAR/GKAB1061>.
- van de Veerdonk, F.L. *et al.* (2022) 'A guide to immunotherapy for COVID-19', *Nature Medicine* 2022 28:1, 28(1), pp. 39–50. Available at: <https://doi.org/10.1038/s41591-021-01643-9>.
- Voulhoux, R. *et al.* (2001) 'Involvement of the twin-arginine translocation system in protein secretion via the type II pathway', *EMBO Journal*, 20(23), pp. 6735–6741. Available at: <https://doi.org/10.1093/emboj/20.23.6735>.
- Wagner, S. *et al.* (2016) 'Novel Strategies for the Treatment of *Pseudomonas aeruginosa* Infections', *Journal of Medicinal Chemistry*, 59(13), pp. 5929–5969. Available at: https://doi.org/10.1021/ACS.JMEDCHEM.5B01698/ASSET/IMAGES/MEDIUM/JM-2015-01698D_0029.GIF.
- Wang, D. *et al.* (2022) 'RplI interacts with 5' UTR of *exsA* to repress its translation and type III secretion system in *Pseudomonas aeruginosa*', *PLoS pathogens*, 18(1). Available at: <https://doi.org/10.1371/JOURNAL.PPAT.1010170>.
- Wang, J. *et al.* (2021) 'Structural basis for the biosynthesis of lovastatin', *Nature Communications* 2021 12:1, 12(1), pp. 1–10. Available at: <https://doi.org/10.1038/s41467-021-21174-8>.
- Wang, Y. *et al.* (2011) 'Phenazine-1-carboxylic acid promotes bacterial biofilm development via ferrous iron acquisition', *Journal of bacteriology*, 193(14), pp. 3606–3617. Available at: <https://doi.org/10.1128/JB.00396-11>.
- Wang, Z. *et al.* (2016) 'Comprehensive evaluation of ten docking programs on a diverse set of protein-ligand complexes: The prediction accuracy of sampling power and scoring power', *Physical Chemistry Chemical Physics*, 18(18), pp. 12964–12975. Available at: <https://doi.org/10.1039/c6cp01555g>.
- Waterhouse, A. *et al.* (2018) 'SWISS-MODEL: Homology modelling of protein structures and complexes', *Nucleic Acids Research*, 46(W1), pp. W296–W303. Available at: <https://doi.org/10.1093/nar/gky427>.
- Williams, J.D. *et al.* (2015) 'Synthesis and structure–activity relationships of novel

phenoxyacetamide inhibitors of the *Pseudomonas aeruginosa* type III secretion system (T3SS)', *Bioorganic & Medicinal Chemistry*, 23(5), pp. 1027–1043. Available at: <https://doi.org/10.1016/j.bmc.2015.01.011>.

Winstanley, C., O'Brien, S. and Brockhurst, M.A. (2016) 'Pseudomonas aeruginosa Evolutionary Adaptation and Diversification in Cystic Fibrosis Chronic Lung Infections', *Trends in Microbiology*. Elsevier Ltd, pp. 327–337. Available at: <https://doi.org/10.1016/j.tim.2016.01.008>.

Wood, C.W. *et al.* (2020) 'BALaS: fast, interactive and accessible computational alanine-scanning using BudeAlaScan', *Bioinformatics*, 36(9), pp. 2917–2919. Available at: <https://doi.org/10.1093/BIOINFORMATICS/BTAA026>.

Woodbrey, A.K. *et al.* (2017) 'A new class of inhibitors of the AraC family virulence regulator *Vibrio cholerae* ToxT', *Scientific Reports*, 7, p. 45011. Available at: <https://doi.org/10.1038/srep45011>.

World Health Organisation (2017) *Global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics*. Available at: http://www.who.int/medicines/publications/WHO-PPL-Short_Summary_25Feb-ET_NM_WHO.pdf?ua=1 (Accessed: 28 November 2017).

World Health Organization (2014) 'Antimicrobial resistance: global report on surveillance 2014', *World Health Organization*, pp. 1–257. Available at: <https://doi.org/9789241564748>.

Xu, J. *et al.* (2015) 'Occurrence of antibiotics and antibiotic resistance genes in a sewage treatment plant and its effluent-receiving river', *Chemosphere*, 119, pp. 1379–1385. Available at: <https://doi.org/10.1016/j.chemosphere.2014.02.040>.

Xu, Y. *et al.* (2017) 'Rutaecarpine Inhibits Intimal Hyperplasia in A Balloon-Injured Rat Artery Model', *Chinese Journal of Integrative Medicine* 2018 24:6, 24(6), pp. 429–435. Available at: <https://doi.org/10.1007/S11655-017-2900-3>.

Yan, H. *et al.* (2019) 'Conditional quorum-sensing induction of a cyanide-insensitive terminal oxidase stabilizes cooperating populations of *Pseudomonas aeruginosa*', *Nature Communications* 2019 10:1, 10(1), pp. 1–7. Available at: <https://doi.org/10.1038/s41467-019-13013-8>.

Yang, J.J. *et al.* (2018) 'Commensal staphylococcus aureus provokes immunity to protect against skin infection of methicillin-resistant staphylococcus aureus', *International Journal of Molecular Sciences*, 19(5). Available at: <https://doi.org/10.3390/ijms19051290>.

Yang, X. *et al.* (2022) 'Molecular mechanism of allosteric modulation for the cannabinoid receptor CB1', *Nature Chemical Biology* 2022, pp. 1–10. Available at: <https://doi.org/10.1038/s41589-022-01038-y>.

Yeo, W.S. *et al.* (2018) 'The FDA-approved anti-cancer drugs, streptozotocin and floxuridine, reduce the virulence of *Staphylococcus aureus*', *Scientific Reports*, 8(1). Available at:

<https://doi.org/10.1038/s41598-018-20617-5>.

Yueh, C. *et al.* (2017) 'ClusPro-DC: Dimer Classification by the Cluspro Server for Protein-Protein Docking', *Journal of Molecular Biology*, 429(3), pp. 372–381. Available at: <https://doi.org/10.1016/j.jmb.2016.10.019>.

Żaczek, M. *et al.* (2020) 'Phage Therapy in Poland - a Centennial Journey to the First Ethically Approved Treatment Facility in Europe', *Frontiers in microbiology*, 11. Available at: <https://doi.org/10.3389/FMICB.2020.01056>.

Zarkan, A. *et al.* (2020) 'Inhibition of indole production increases the activity of quinolone antibiotics against *E. coli* persisters', *Scientific reports*, 10(1). Available at: <https://doi.org/10.1038/S41598-020-68693-W>.

Zhang, N. and He, Q.S. (2015) 'Commensal microbiome promotes resistance to local and systemic infections', *Chinese Medical Journal*. Chinese Medical Association, pp. 2250–2255. Available at: <https://doi.org/10.4103/0366-6999.162502>.

Zhang, X. *et al.* (2014) 'Prevalence of Veterinary Antibiotics and Antibiotic-Resistant *Escherichia coli* in the Surface Water of a Livestock Production Region in Northern China', *PLoS ONE*. Edited by T.C. Smith, 9(11), p. e111026. Available at: <https://doi.org/10.1371/journal.pone.0111026>.

Zhang, Y. *et al.* (2018) 'Coumarin Reduces Virulence and Biofilm Formation in *Pseudomonas aeruginosa* by Affecting Quorum Sensing, Type III Secretion and C-di-GMP Levels', *Frontiers in microbiology*, 9(AUG). Available at: <https://doi.org/10.3389/FMICB.2018.01952>.

Zhang, Y. *et al.* (2019) 'Commensal microbes affect host humoral immunity to bordetella pertussis infection', *Infection and Immunity*, 87(10). Available at: <https://doi.org/10.1128/IAI.00421-19>.

Zhao, H. and Caflisch, A. (2013) 'Discovery of ZAP70 inhibitors by high-throughput docking into a conformation of its kinase domain generated by molecular dynamics', *Bioorganic and Medicinal Chemistry Letters*, 23(20), pp. 5721–5726. Available at: <https://doi.org/10.1016/j.bmcl.2013.08.009>.

Zhao, J. *et al.* (2019) 'Widespread existence of quorum sensing inhibitors in marine bacteria: Potential drugs to combat pathogens with novel strategies', *Marine Drugs*, 17(5). Available at: <https://doi.org/10.3390/md17050275>.

Zheng, Z. *et al.* (2007) 'Biochemical Characterization of a Regulatory Cascade Controlling Transcription of the *Pseudomonas aeruginosa* Type III Secretion System', *Journal of Biological Chemistry*, 282(9), pp. 6136–6142. Available at: <https://doi.org/10.1074/jbc.M611664200>.

Zhong, L. *et al.* (2020) 'Attenuation of *Pseudomonas aeruginosa* Quorum Sensing by Natural Products: Virtual Screening, Evaluation and Biomolecular Interactions', *International Journal of Molecular Sciences*, 21(6), p. 2190. Available at: <https://doi.org/10.3390/ijms21062190>.

Zimmermann, A. *et al.* (1991) 'Anaerobic growth and cyanide synthesis of *Pseudomonas aeruginosa* depend on *anr*, a regulatory gene homologous with *fnr* of *Escherichia coli*', *Molecular microbiology*, 5(6), pp. 1483–1490. Available at: <https://doi.org/10.1111/J.1365-2958.1991.TB00794.X>.

Appendixes

Appendix 1: R script for the handling and analysis of Autodock Vina outputs

```
#Notes - Reads in as many files as required as long as the names are given correctly and end
in an unbroken chain of numbers
#Notes - Works for mol2 files as written by MMV - Use prepare function to remove H's.
setwd("E:/Working directory")
#Set variables
#0ing
fscore<-0
score<-0
AllRMSDs<-c()
MaxRMSD<-0
MeanRMSD<-0
deltasum<-0
BestRMSDMean<-500
BestRMSDMax<-500
part1workingscore<-0
XBestRMSDMean<-0
XBestRMSDMax<-0
XBestfscore<-0
#Fill in
TotalAtoms<-21
filebase<-"Name_"
filenum<-100
marg<-1 #Acceptable margin
AllRMSDs<-c()
###Test pose variation
for(i in 1:filenum-1){
  TestPose<-data.frame(read.table(paste0(filebase,i,".mol2"),skip=10,nrows=TotalAtoms))
  PoseName<-paste0(filebase,i,".mol2")
# 0 out again
  MaxRMSD<-0
  MeanRMSD<-0
  fscore<-0
  score<-0
```

```

AllRMSDs<-c()
deltasum<-0
part1workingscore<-0
#read in files
for(i in 1:filenum-1){
  filename<-paste0(filebase,i,".mol2")
  datafile<-data.frame(read.table(filename,skip=10,nrows=TotalAtoms))
#Calculations for score per file
  part2workingscore<-0
  for(i in 1:TotalAtoms){
    coords<-c(TestPose[i,3],TestPose[i,4],TestPose[i,5])
    testcoords<-c(datafile[i,3],datafile[i,4],datafile[i,5])
    part1workingscore<-0
    sub1<-as.numeric(testcoords[1]-coords[1])
    sub2<-coords[1]-testcoords[1]
    sub3<-testcoords[2]-coords[2]
    sub4<-coords[2]-testcoords[2]
    sub5<-testcoords[3]-coords[3]
    sub6<-coords[3]-testcoords[3]
    if((((sub1<=marg)&&(sub1>=0))||((sub2<=marg)&&(sub2>=0)))){
      part1workingscore<-part1workingscore+1
    }
    if((((sub3<=marg)&&(sub3>=0))||((sub4<=marg)&&(sub4>=0)))){
      part1workingscore<-part1workingscore+1
    }
    if((((sub5<=marg)&&(sub5>=0))||((sub6<=marg)&&(sub6>=0)))){
      part1workingscore<-part1workingscore+1
    }
    if(part1workingscore==3){
      part2workingscore<-part2workingscore+1
      #and add to RMSD
      cords<-c(TestPose[i,3],TestPose[i,4],TestPose[i,5])
      testcoords<-c(datafile[i,3],datafile[i,4],datafile[i,5])
      deltaX<-cords[1]-testcoords[1]
      deltaY<-cords[2]-testcoords[2]
      deltaZ<-cords[3]-testcoords[3]
      delta<-((deltaX*deltaX)+(deltaY*deltaY)+(deltaZ*deltaZ))
    }
  }
}

```

```

    deltasum<-deltasum+delta
  }
  RMSDworking<-sqrt(deltasum/TotalAtoms)
}
if(part2workingscore==TotalAtoms){
  fscore<-fscore+1}
AllRMSDs<-c(AllRMSDs,RMSDworking)
}
##Final calculations 1
MeanRMSD<-mean(AllRMSDs)
MaxRMSD<-max(AllRMSDs)
if(MeanRMSD<BestRMSDMean){
  BestRMSDMean<-MeanRMSD
  BestRMSDMax<-MaxRMSD
  BestPose<-PoseName
  Bestfscore<-fscore
}
if(fscore>XBestfscore){
  XBestRMSDMean<-MeanRMSD
  XBestRMSDMax<-MaxRMSD
  XBestPose<-PoseName
  XBestfscore<-fscore
}
}
##Feedback
print("By RMSD")
print(BestRMSDMean)
print(BestRMSDMax)
fb<-paste("Raw score =",Bestfscore)
print(fb)
fb2<-paste("% score =",(Bestfscore/filenum)*100)
print(fb2)
print(BestPose)
print("By score")
print(XBestRMSDMean)
print(XBestRMSDMax)
Xfb<-paste("Raw score =",XBestfscore)

```

```
print(Xfb)
Xfb2<-paste("% score =",(XBestfscore/filenum)*100)
print(Xfb2)
print(XBestPose)
```


Appendix 2: A Phylogeny for ExsA legend

The full phylogeny generated for the ExsA using the PROSITE entry PS01124 grouping of AraC family proteins. Numbers at junctions indicate the percentage of replicates which contained the shown division. In order to reduce the height of the figure it has been split in two, with an overlap between the bottom of the left hand panel and the top of the right hand panel.