



The University of  
**Nottingham**

**RsmN – a new atypical RsmA homologue  
in *Pseudomonas aeruginosa***



Laura C Lovelock, M.Sci.

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

July 2012

## **DECLARATION**

Unless otherwise acknowledged, the work presented in this thesis is my own. No part has been submitted for another degree at the University of Nottingham or any other institute of learning.

Laura Lovelock

<b>TABLE OF CONTENTS</b>	
<b>DECLARATION</b>	<b>2</b>
<b>TABLE OF FIGURES</b>	<b>12</b>
<b>TABLE OF TABLES</b>	<b>17</b>
<b>ACKNOWLEDGMENTS</b>	<b>19</b>
<b>ABSTRACT</b>	<b>20</b>
<b>ABBREVIATIONS</b>	<b>22</b>
<b>1 INTRODUCTION</b>	<b>27</b>
<b>1.1 Bacterial Virulence</b>	<b>27</b>
1.1.1 Bacterial Pathogenicity	27
1.1.2 Virulence of <i>Pseudomonas aeruginosa</i>	27
<b>1.2 REGULATION OF VIRULENCE</b>	<b>29</b>
1.2.1 Virulence regulation at the transcriptional level	29
1.2.1.1 Bacterial cell-to-cell communication	29
1.2.1.2 QS-dependent control of gene expression	30
1.2.1.3 Quorum sensing signalling molecules	31
1.2.1.4 Transcriptional virulence regulation in <i>P. aeruginosa</i>	32
1.2.1.5 The GacS/GacA two-component system	35
1.2.1.6 Regulation by the Csr/Rsm System	37
1.2.1.6.1 Role of RsmA in gene expression	37
1.2.1.6.2 RsmA Structure	39
1.2.1.6.3 The <i>P. aeruginosa</i> Regulatory RNAs, RsmZ and RsmY	41
1.2.1.6.4 Additional control of the Rsm system by RetS and LadS	42
1.2.1.7 Regulatory RNA structures	44
1.2.1.8 Target mRNAs	46
1.2.2 Gene regulation by sRNAs	50
1.2.2.1 sRNA Regulation	50
1.2.2.2 Cis-encoded natural Antisense RNA (asRNA)	55
1.2.2.2.1 Previous limitations of the study of asRNA transcription	56
1.2.2.2.2 Types of antisense transcripts in bacteria	57
1.2.2.2.3 Mechanisms of asRNA action	58
<b>1.3 Research outline and aims of the presented work</b>	<b>65</b>

<b>2</b>	<b>MATERIALS AND METHODS</b>	<b>66</b>
<b>2.1</b>	<b>Bacterial strains</b>	<b>66</b>
<b>2.2</b>	<b>Plasmids</b>	<b>68</b>
<b>2.3</b>	<b>Oligonucleotides</b>	<b>71</b>
<b>2.4</b>	<b>Plasmid and strain construction</b>	<b>73</b>
2.4.1	Construction of plasmids	73
2.4.1.1	Plasmids made by PCR-based point mutagenesis	73
2.4.1.2	Construction of arginine-alanine substitution mutants	74
2.4.1.3	Construction of the <i>E. coli</i> overexpression plasmid pHT:: <i>rsmN</i>	74
2.4.1.4	Construction of suicide plasmid pDM4:: <i>lacI</i> <sup>Q</sup> P <sub>tac</sub> - <i>rsmN</i> (pLT10)	74
2.4.1.5	<i>rsmN</i> deletion mutant.	75
2.4.1.6	<i>rsmN</i> conditional mutant (PALT11)	77
2.4.1.7	Construction of a <i>gacA</i> mutant ( $\Omega$ Sm/Sp)	77
2.4.1.8	Construction of a sense <i>rsmN-lux</i> transcriptional reporter fusion (pLT1)	77
2.4.1.9	Construction of an antisense <i>nmsR-lux</i> transcriptional fusion (pLT2)	78
<b>2.5</b>	<b>General chemicals</b>	<b>78</b>
2.5.1	Antibiotics	78
2.5.2	Synthetic quorum sensing signal molecules	79
<b>2.6</b>	<b>Growth Media</b>	<b>79</b>
2.6.1	Luria Bertani media (LB)	79
2.6.2	Peptone Tryptone Soy Broth (PTSB)	80
2.6.3	King's B Medium	80
2.6.4	Swarming motility agar	80
2.6.5	Kornberg medium	81
2.6.6	Pyocyanin medium	81
2.6.7	M9 Minimal Medium Protein Expression	81
<b>2.7</b>	<b>Growth &amp; storage of bacteria</b>	<b>81</b>
2.7.1	Bacterial growth conditions	81

2.7.2	Long term storage of bacterial strains	81
<b>2.8</b>	<b>Protocols</b>	<b>82</b>
2.8.1	Transformation of bacterial strains	82
2.8.1.1	Preparation of electrocompetent <i>E. coli</i> cells	82
2.8.1.2	Electroporation of electrocompetent <i>E. coli</i> cells	82
2.8.1.3	Preparation of electrocompetent <i>P. aeruginosa</i> cells	83
2.8.1.4	Electroporation of electrocompetent <i>P. aeruginosa</i> cells	83
2.8.1.5	<i>P. aeruginosa</i> transformation using CaCl <sub>2</sub>	83
2.8.2	Quantifying DNA, RNA and protein concentrations	84
2.8.3	DNA manipulation	84
2.8.3.1	Isolation of chromosomal DNA	84
2.8.3.2	Isolation of plasmid DNA	85
2.8.3.3	CTAB mini-prep for plasmid purification	85
2.8.3.4	Isolation of large quantities of plasmid DNA	86
2.8.3.5	Precipitation of DNA/RNA	86
2.8.3.6	Polymerase chain reaction (PCR) amplification	87
2.8.3.7	DNA Clean and Concentrate (Zymoclean)	88
2.8.3.8	DNA agarose gel electrophoresis	88
2.8.3.9	DNA molecular weight markers	89
2.8.3.10	Agarose gel extraction using the Qiaquick method.	89
2.8.3.11	Agarose gel extraction using Zymoclean™	89
2.8.3.12	Phenol/chloroform purification of DNA	90
2.8.3.13	DNA restriction enzymes	90
2.8.3.14	Dephosphorylation of DNA	91
2.8.3.15	DNA ligation	91
2.8.3.16	Klenow fill-in	91
2.8.3.17	DNase digestion	92
2.8.4	DNA sequencing	92
2.8.4.1	DNA sequencing	92
2.8.4.2	DNA sequence analysis	92
2.8.5	Gene replacement in <i>P. aeruginosa</i>	93
2.8.5.1	Conjugation of plasmid DNA into <i>P. aeruginosa</i>	93
2.8.5.2	Sucrose counter-selection	93

2.8.6	RNA work	94
2.8.6.1	<i>In vitro</i> transcription	94
2.8.6.2	RNA extraction (phenol-chloroform)	94
2.8.6.3	Total RNA extraction (Qiagen)	95
2.8.6.4	RNA Cleanup	96
2.8.6.5	RNA molecular weight markers	96
2.8.7	Protein Methods	96
2.8.7.1	Protein expression	96
2.8.7.2	Purification using hexahistidine tags and Ni-NTA chromatography	97
2.8.7.3	Protein purification - HisPur™ cobalt resin	98
2.8.7.4	Scale up	99
2.8.7.5	Protein expression in M9 minimal medium	99
2.8.7.6	Thrombin cleavage	99
2.8.7.7	Desalting	100
2.8.7.8	Ionic exchange	100
2.8.7.9	HiTrap™ heparin affinity column	101
2.8.7.10	Gel filtration	102
2.8.7.11	Superloop	102
2.8.7.12	Freeze-drying	102
2.8.7.13	Anionic exchange	102
2.8.7.14	Circular dichroism spectroscopy (CD)	103
2.8.7.15	Estimation of protein concentration using the Bradford assay	103
2.8.7.16	Tricine SDS-PAGE	104
2.8.7.17	Tricine-SDS-PAGE for Western blot	105
2.8.7.18	Coomassie staining	106
2.8.7.19	Western blotting	106
2.8.7.19.1	Detection of Proteins after Western blotting	107
2.8.7.19.2	PVDF membrane dye	108
2.8.7.19.3	Stripping immunoblots	108
2.8.7.19.4	Peptide mass fingerprinting	108
2.8.8	Protein-RNA interactions	109

2.8.8.1	Electrophoretic mobility shift assay (EMSA)	109
2.8.8.2	Detection of RNA on nylon membranes	110
2.8.8.3	Deep-Seq analysis	110
2.8.8.4	Protein-RNA experiments	113
2.8.8.4.1	Ni-NTA column	113
2.8.8.4.2	Ni-NTA magnetic beads	115
2.8.8.4.3	RNA extraction after overexpression of <i>rsmA</i> and <i>rsmN</i> .	117
2.8.8.4.4	RNA purification and recovery	118
2.8.9	Determination of bioluminescence and growth using a microtitre well plate assay	119
2.8.10	<i>rsmA/N</i> complementation assays	119
2.8.10.1	Swarming motility assays	120
2.8.10.2	Pyocyanin assay	120
2.8.10.3	Kornberg assay	121
2.8.10.4	Elastase assay	121
2.8.10.5	Exoprotease assay.	122
2.8.10.6	Skimmed milk protease assay	123
2.8.10.7	Transformation efficiency–restriction assay	123
2.8.11	Molecular modelling	124
<b>2.9</b>	<b>Protein Analysis</b>	<b>124</b>
2.9.1	Electrospray ionisation mass spectrometry (ESI-MS)	124
2.9.2	Circular dichroism spectroscopy (CD)	125
2.9.3	UV-Vis spectroscopy	126
2.9.4	Equilibrium fluorescence spectroscopy	127
2.9.5	Nuclear magnetic resonance spectroscopy (NMR)	128
<b>3</b>	<b>PURIFICATION AND BIOPHYSICAL ANALYSIS OF RSM A</b>	<b>129</b>
<b>3.1</b>	<b>Introduction</b>	<b>129</b>
<b>3.2</b>	<b>Results and discussion</b>	<b>135</b>
3.2.1	RsmA–Protein expression and purification	135
3.2.1.1	Thrombin cleavage - Gel filtration chromatography	141
3.2.1.2	Major contaminant in Ni-NTA purification of RsmA.	142

3.2.2	Expression and purification of RsmA from M9 minimal growth medium for NMR experiments.	146
3.2.2.1	Protein expression	146
3.2.2.2	Protein purification	146
3.2.3	Electrospray ionization mass spectrometry	146
3.2.3.1	His <sub>6</sub> -Thr-RsmA Purification Comparison.	146
3.2.4	Circular dichroism analysis of RsmA	149
3.2.4.1	Spectra and temperature melting of cleaved RsmA	149
3.2.5	Equilibrium Fluorescence	156
3.2.5.1	RsmA tryptophan substitution mutants	156
3.2.5.2	Prospective tryptophan substitution mutants	159
3.2.6	Impact of temperature, denaturant and pH on the structure of RsmA using NMR analysis	163
3.2.6.1	Comparison of RsmA purified by Ni-NTA agarose and HisPur™ cobalt resin	163
3.2.6.2	Temperature Study	164
3.2.6.3	Denaturant	167
3.2.6.4	The effect of pH on RsmA.	169
<b>3.3</b>	<b>Conclusions</b>	<b>171</b>
3.3.1	Expression and purification of RsmA	171
3.3.2	Biophysical Methods	172
<b>4</b>	<b>IDENTIFICATION OF A NOVEL RSM A HOMOLOGUE IN P. AERUGINOSA AND ITS IMPACT ON THE REGULATION OF VIRULENCE DETERMINANTS</b>	<b>176</b>
<b>4.1</b>	<b>Introduction</b>	<b>176</b>
<b>4.2</b>	<b>Results and Discussion</b>	<b>178</b>
4.2.1	Identification of <i>RsmN</i>	178
4.2.2	Sequence comparison of RsmN and RsmA	179
4.2.3	Structural comparisons of RsmN and RsmA	182
4.2.4	Transcriptional analysis	184
4.2.5	Construction of strains used in this chapter	185
4.2.5.1	mini-CTX:: <i>lux</i> promoter fusions	186

4.2.5.2	<i>rhlI</i> , <i>lasI</i> and <i>pqsA</i> promoter fusions	189
4.2.5.3	Sense and antisense <i>rsmN</i> and <i>nmsR</i> fusions in $\Delta rhlR$ , $\Delta lasR$ and $\Delta pqsA$	189
4.2.6	<i>rsmN</i> and <i>nmsR</i> gene expression	190
4.2.6.1	Construction of RsmN Arginine-62-Alanine (R62A) mutants	191
4.2.6.2	Construction of an <i>rsmN</i> mutant (PALT16)	191
4.2.6.3	Construction of a conditional, inducible <i>rsmN</i> mutant (strain PALT11)	192
4.2.6.4	Construction attempts for a $\Delta rsmA\Delta rsmN$ double mutant strain.	192
4.2.6.5	Western Blot	193
4.2.7	The influence of RsmN and RsmA on Quorum Sensing (QS)	194
4.2.7.1	Influence of RsmN and RsmA on <i>lasI</i> transcription	194
4.2.7.2	Influence of RsmN and RsmA on <i>rhlI</i> transcription	194
4.2.7.3	Influence of RsmN and RsmA on <i>pqsA</i> transcription	194
4.2.7.4	Influence of <i>lasR</i> , <i>rhlR</i> and QS signalling molecules on <i>rsmN</i> expression	194
4.2.7.4.1	Influence of LasR on <i>rsmN</i> and <i>nmsR</i> transcription	194
4.2.7.4.2	Influence of RhlR on <i>rsmN</i> transcription	194
4.2.7.4.3	Influence of PQS signalling on <i>rsmN</i> expression.	194
4.2.8	Phenotypic characterisation of the <i>rsmN</i> mutant	194
4.2.8.1	Swarming	195
4.2.8.1.1	<i>rsmN</i> mutant	195
4.2.8.2	Glycogen accumulation in <i>E. coli</i>	198
4.2.8.3	Restriction assay	199
4.2.8.4	Control of secondary metabolite production	201
4.2.8.4.1	Elastase Assay	201
4.2.8.4.2	Protease Assay	203
4.2.8.4.3	Pyocyanin Assay	205
<b>4.3</b>	<b>Conclusions</b>	<b>219</b>
<b>5</b>	<b>RELATIONSHIP BETWEEN RSMN, RSMA, AND THE GAC SYSTEM</b>	<b>225</b>

<b>5.1 Introduction</b>	<b>225</b>
<b>5.2 Results and Discussion</b>	<b>227</b>
5.2.1 Strains constructed in this Chapter	227
5.2.1.1 Mini-CTX:: <i>lux</i> promoter fusions	227
5.2.1.2 Construction of <i>gacA</i> mutant PALT40	227
5.2.1.3 Chromosomal transcriptional fusions	228
5.2.1.4 P <sub><i>rsmN</i></sub> and P <sub><i>nmsR</i></sub> fusions in <i>rsmA</i> and <i>rsmN</i> mutants	228
5.2.1.5 P <sub><i>rsmN</i></sub> and P <sub><i>nmsR</i></sub> fusions in $\Delta$ <i>retS</i> mutant	229
5.2.1.6 P <sub><i>rsmN</i></sub> and P <sub><i>nmsR</i></sub> fusions in $\Delta$ <i>ladS</i> mutant	229
5.2.1.7 P <sub><i>rsmN</i></sub> and P <sub><i>nmsR</i></sub> fusions in $\Delta$ <i>gacA</i> mutant	229
5.2.2 Impact of RsmA and RsmN on <i>rsmN</i> and <i>nmsR</i> expression	230
5.2.2.1 The control of expression of <i>rsmN</i> and <i>nmsR</i> by RsmA and RsmN	230
5.2.3 Impact of <i>retS</i> , <i>ladS</i> and <i>gacA</i> on <i>rsmN</i>	233
5.2.3.1 Impact of RetS on <i>rsmN</i> and <i>nmsR</i> transcription	233
5.2.3.2 Impact of LadS on <i>rsmN</i> and <i>nmsR</i> expression	234
5.2.3.3 Impact of GacA on <i>rsmN</i> and <i>nmsR</i> expression	236
<b>5.3 Conclusions</b>	<b>238</b>
<b>6 IDENTIFICATION OF RSMN AND RSMA RNA TARGETS</b>	<b>241</b>
<b>6.1 Introduction</b>	<b>241</b>
<b>6.2 Results and Discussion</b>	<b>245</b>
6.2.1 Strains	245
6.2.1.1 Construction of RsmA and RsmN arginine substitution mutants	245
6.2.2 RNA binding experiments	246
6.2.2.1 Protein-RNA binding using total RNA from <i>P. aeruginosa</i>	246
6.2.2.1.1 Ni-NTA agarose Purifications	246
6.2.2.2 RNA extraction from RsmA and RsmN overexpressed in PAO1	249
6.2.3 RNA Deep-sequencing results	250
6.2.3.1 RNA transcript identification	250
6.2.3.2 RsmN transcript analysis	252
6.2.3.2.1 RNAs enriched by binding to RsmN	252

6.2.3.2.2 Other potential RsmN targets	259
6.2.3.2.3 RNAs depleted when RsmN is overexpressed	260
6.2.3.2.4 RNAs enriched by binding to RsmA	261
6.2.3.2.5 Depleted Transcripts of RsmA	265
<b>6.3 Conclusions</b>	<b>271</b>
<b>7 GENERAL CONCLUSIONS</b>	<b>273</b>
<b>8 BIBLIOGRAPHY</b>	<b>279</b>
<b>9 ANNEX</b>	<b>297</b>
<b>9.1 Appendix I</b>	<b>298</b>
<b>9.2 Appendix II</b>	<b>301</b>
<b>9.3 Appendix III</b>	<b>305</b>
<b>9.4 Appendix IV</b>	<b>309</b>

## TABLE OF FIGURES

Figure 1.1: Quorum sensing signal molecules in <i>P. aeruginosa</i> . .....	33
Figure 1.2: Proposed model for the influence of RhlR on the las regulon. ....	35
Figure 1.3: RNA-binding domain structure comparison. ....	40
Figure 1.4: Model of the GacA/RsmA signal transduction pathway in <i>P. aeruginosa</i> PAO1.....	42
Figure 1.5: Summary of gene regulation in <i>P. aeruginosa</i> .....	43
Figure 1.6: Predicted secondary structure of regulatory RNAs RsmY and RsmZ.....	44
Figure 1.7: Predicted secondary structures of representative selected RNA ligands. ....	46
Figure 1.8: Genetic organization of the <i>hcnA</i> 5' untranslated mRNA.....	47
Figure 1.9: Representations of CsrA-RNA binding combinations. ....	50
Figure 1.10: Widely accepted modes of Hfq activity. ....	52
Figure 1.11: The <i>isiA/IsiR</i> pair of <i>Synechocystis</i> .....	59
Figure 1.12: Inhibition of translation through SymR. ....	60
Figure 1.13: Transcription termination by bacterial asRNAs in <i>Vibrio anguillarum</i> .....	61
Figure 1.14: Transcription interference by collision in the <i>ubiG-mccBA</i> operon in <i>Clostridium acetobutylicum</i> .....	63
Figure 1.15: Promoter occlusion mechanism in $\lambda$ phage $P_R$ and $P_{RE}$ promoters. ....	63
Figure 1.16: Sitting duck transcriptional interference in bacteriophage 186...	64
Figure 2.1: Schematic representation of pLT10, the suicide plasmid for the construction of inducible <i>rsmN</i> strains. ....	75

Figure 2.2: Representation of the steps required to make the <i>rsmN</i> mutant strain.....	76
Figure 2.3: Chemiluminescence production by the ECL detection system. ..	107
Figure 2.4: Integration of SOLiD system barcodes into the library construction workflow .....	112
Figure 2.5: Schematic diagram for the RNA extraction from PAO1 pRsmA and PAO1 pRsmN.....	117
Figure 3.1: MALDI-TOF mass spectrum of CsrA.....	130
Figure 3.2: NMR solution structure of the RsmE- <i>hcnA</i> RNA complex. ....	132
Figure 3.3: Surface potential of the CsrA structure. ....	133
Figure 3.4: Schematic representation of intermolecular RsmE- <i>hcnA</i> interactions.....	134
Figure 3.5: Sequences of plasmids for RsmA over-expression in <i>E. coli</i> .....	137
Figure 3.6: SDS-PAGE Tricine gel of successful Ni-NTA purification of His <sub>6</sub> -Thr-RsmA. ....	140
Figure 3.7: Gel filtration trace of cleaved His <sub>6</sub> -Thr-RsmAY48W.....	141
Figure 3.8: SDS-PAGE tricine gel of contaminants in Ni-NTA purification. ....	142
Figure 3.9: Contaminant removal gel of Ni-NTA purification.....	143
Figure 3.10: ESI mass spectra of His <sub>6</sub> -Thr-RsmA.....	148
Figure 3.11: CD spectra of pure protein secondary structures.....	149
Figure 3.12: Comparison CD spectra of RsmA wild type and tryptophan substitution mutants cleaved and uncleaved. ....	151
Figure 3.13: CD temperature melts of RsmA wild type and tryptophan substitutions mutants cleaved and uncleaved. ....	153

Figure 3.14: CD comparison spectra of purification resins. ....	155
Figure 3.15: Excitation spectra of RsmAY48W. ....	157
Figure 3.16: Emission spectra of His <sub>6</sub> -Thr-RsmAY48W. ....	159
Figure 3.17: Prospective tryptophan mutants chosen for site-directed mutagenesis. ....	160
Figure 3.18: Prospective RsmAT19W mutant. ....	160
Figure 3.19: Close up of RsmAT19W predicted structure. ....	161
Figure 3.20: Swarming assays of RsmA and the RsmA tryptophan mutants, RsmAL23W and N23W. ....	162
Figure 3.21: 1D NMR proton spectra purification method comparison. ....	164
Figure 3.22: 1D NMR spectra of the effect of temperature on cleaved RsmAY48W. ....	166
Figure 3.23: 1D NMR WG proton spectra of the effect of chemical denaturant on cleaved RsmAY48W. ....	168
Figure 3.24: 1D NMR proton spectra of effect of pH on His <sub>6</sub> -Thr-RsmA stability. ....	170
Figure 4.1: Restoration of swarming in <i>P. aeruginosa</i> <i>rsmA</i> mutants by clones identified as carrying <i>rsmN</i> . ....	179
Figure 4.2: Structure-based amino acid sequence alignments of RsmN/RsmA/CsrA homologues. ....	180
Figure 4.3: Possible salt bridge in RsmN. ....	181
Figure 4.4: RsmA and RsmN molecular models and schematics. ....	182
Figure 4.5: Molecular model of RsmN. ....	183
Figure 4.6: CD comparison spectra of wild type His <sub>6</sub> -Thr-RsmA and His <sub>6</sub> -Thr-RsmN. ....	184

Figure 4.7: Genetic context of <i>rsmN</i> .....	185
Figure 4.8: Diagrammatic representation of the <i>rsmN</i> and <i>nmsR</i> miniCTX:: <i>lux</i> promoter gene fusions.....	186
Figure 4.9: Chromosomal constructs made in <i>P. aeruginosa</i> PAO1.....	187
Figure 4.10: PCR products for P <sub><i>rsmN</i></sub> and P <sub><i>nmsR</i></sub> construction. ....	188
Figure 4.11: Expression of <i>rsmN</i> and <i>nmsR</i> promoters in <i>P. aeruginosa</i> PAO1 (Nottingham) as a function of growth.....	190
Figure 4.12: Western blot analysis of RsmN production.....	194
Figure 4.13: Swarming motility of <i>P. aeruginosa rsmA</i> and <i>rsmN</i> mutants complemented by RsmN variants. ....	196
Figure 4.14: Swarming motility of the inducible <i>P. aeruginosa rsmN</i> mutant. ....	197
Figure 4.15: Repression of glycogen synthesis in <i>E. coli</i> by RsmA but not RsmN. ....	199
Figure 4.16: Restriction Assay for <i>rsmN</i> and <i>rsmA</i> complemented PAO1 strains. ....	201
Figure 4.17: Elastin-congo red assay to investigate the impact of RsmN on elastase production.....	202
Figure 4.18: Impact of RsmN on exoprotease. ....	204
Figure 4.19: Pyocyanin production in <i>rsmA</i> and <i>rsmN</i> mutants. ....	205
Figure 4.20: Pyocyanin production by of PAO1 wild type, $\Delta rsmA$ and $\Delta rsmN$ mutants complemented with RsmN variants. ....	207
Figure 4.21: Expression of <i>lasI</i> in <i>rsmA</i> and <i>rsmN</i> mutants using chromosomal reporter <i>lux</i> fusions. ....	209

Figure 4.22: Expression of <i>rhlI</i> in <i>rsmA</i> and <i>rsmN</i> mutants using chromosomal reporter <i>lux</i> fusions. ....	211
Figure 4.23: Expression of <i>pqsA</i> in <i>rsmA</i> and <i>rsmN</i> strains using chromosomal reporter <i>lux</i> fusions. ....	213
Figure 4.24: Impact of LasR on the expression of <i>rsmN</i> and <i>nmsR</i> . ....	215
Figure 4.25: Impact of RhlR on the expression of <i>rsmN</i> and <i>nmsR</i> . ....	216
Figure 4.26: Impact of 2-alkyl-4-quinolone signalling on the expression of <i>rsmN</i> and <i>nmsR</i> . ....	218
Figure 4.27: <i>rsmN</i> expression in a <i>pqsA</i> mutant in the presence or absence of PQS. ....	219
Figure 5.1: A model for the convergence of the signalling pathways during reciprocal regulation of virulence factors by LadS, RetS, and GacS through transcription of the small regulatory RNA RsmZ (Ventre et al., 2006). ....	226
Figure 5.2: Effect of RsmA and RsmN on the <i>rsmN</i> (A and C) and <i>nmsR</i> (B and D) promoters. ....	232
Figure 5.3: Effects of RetS on the <i>rsmN</i> (A) and <i>nmsR</i> (B) promoters. ....	233
Figure 5.4: Effect of LadS on the <i>rsmN</i> (A) and <i>nmsR</i> (B) promoters. ....	235
Figure 5.5: Effects of GacA on the <i>rsmN</i> (A) and <i>nmsR</i> (B) promoters. ....	237
Figure 6.1: Agilent bioanalyzer traces for RNA samples extracted from RsmA bound to a Ni-NTA column and magnetic beads. ....	248
Figure 6.2: Interpretation of RNA genetic arrangements .....	251

## TABLE OF TABLES

Table 1.1: Advantages of RNA-seq compared with other transcriptomic methods(Wang et al., 2009).....	53
Table 2.1: Bacterial strains used in this study.....	66
Table 2.2: Plasmids used in this study.....	68
Table 2.3: Oligonucleotides used in this study.....	71
Table 2.4: Tricine-SDS-PAGE separating and resolving gel solution components.....	104
Table 2.5: Tricine-SDS-PAGE separating and resolving gel solution components for Western Blotting.....	105
Table 2.6: Interaction Buffer B to optimise protein-RNA binding (Volume dependent on volume of RNA used).....	114
Table 3.2: Conditions used for the optimization of contaminant removal from RsmA bound to either Ni-NTA agarose or HisPur™ Cobalt columns.....	145
Table 4.2: <i>rhlI</i> , <i>lasI</i> and <i>pqsA</i> promoter fusions.....	189
Table 4.3: <i>rsmN</i> and <i>nmsR</i> promoter fusions in $\Delta rhlR$ , $\Delta lasR$ and $\Delta pqsA$ .....	190
Table 6.1: <i>P. aeruginosa</i> strains for RNA-binding experiments.....	245
Table 6.2: Plasmids for RNA-binding experiments.....	245
Table 6.3: Quantity of identified transcripts for RsmN.....	253
Table 6.4: RsmN-enriched Target Transcripts.....	254
Table 6.5: Undetermined RsmN targets.....	260
Table 6.6: Depleted RsmN Transcripts.....	261
Table 6.7: Quantity of identified transcripts for RsmA.....	261
Table 6.8: RsmA-enriched Target Transcripts.....	263

Table 6.9: Depleted RsmA target transcripts.....	267
Table 6.10: Comparison of selected RsmA and RsmN data.....	269

## ACKNOWLEDGMENTS

I would like to express my appreciation to Prof. Mark Searle for providing me with the opportunity to carry out a Ph.D. project and to Prof Paul Williams, Prof. Miguel Cámara and Dr Stephan Heeb for their unwavering support. I cannot begin to express my gratitude for their supervision and belief which has been sincerely appreciated.

I could not have finished my Ph.D. without the assistance and guidance of Dr. Stephan Heeb, to whom I would like to thank for all his help, advice and enthusiasm, his support has made this work possible.

To everybody in the Searle Research group, many thanks for imparting their knowledge and support, especially Dr Jed Long and Dr Huw Williams. For their wonderful friendship, a special thank you to Anita Rea, Vicki Thurston, Graham Balkwill and my Ph.D. brother, Tom Garner. Thank you to Liz Morris, who is continuing the biophysical work on RsmN, for her collaboration.

I would also like to thank everybody in the *Pseudomonas* Research group for making the past three years very special. For all their advice, help and support, special mention goes to Sarah Kuehne, Jeni Lockett and Hannah Patrick. There have been so many people who made a special effort to help answer my questions. Thank you to all my friends on B-Floor for their camaraderie, I know I will be missed almost as much as my cakes.

I would like to acknowledge the BBSRC for funding and a special acknowledgement to Victoria Wright and Dr Jo Rowsell for their knowledge and assistance with the RNA DeepSeq work.

Finally I would like to thank my family for always supporting me, to my brother Gareth, sister Helen and my parents Colin and Margaret for their unwavering love and belief. Last but not least, thank you to my husband Kevin for his love and encouragement.

## ABSTRACT

### **RsmN – new atypical RsmA homologue in**

#### *Pseudomonas aeruginosa*

The RsmA/CsrA family of global post-transcriptional regulators are small RNA-binding proteins involved in the regulation of a large number of genes such as those involved in quorum sensing, virulence factor production, secondary metabolism, motility and biofilm formation. They bind to target mRNAs and hence modulate their stability and translation rates. Their effects are antagonised by small non-coding regulatory RNAs. The control of expression of target genes via this post-transcriptional regulatory network is mostly operated in *Pseudomonas* spp. via the GacS/GacA two component system. This study aimed to perform a biophysical analysis of RsmA and to obtain a preliminary understanding of the structure, function and regulation of RsmN, a new atypical RsmA homologue from *Pseudomonas aeruginosa*.

RsmA was purified and biophysical analysis confirmed that RsmA exists as a dimer and is highly stable at high temperatures (75 °C) and low pH (5.2). Although RsmN was found to be structurally similar to RsmA, no functional phenotypes have been identified. Consequently, *rsmN* was mutated and transcriptional fusions to *rsmN* and its anti-sense transcript were constructed for expression studies. Phenotypic analysis indicated that RsmN was not involved in the control of swarming, pyocyanin, elastase and protease production or glycogen accumulation. Unlike RsmA, RsmN does not have a control on the restriction modification system of *P. aeruginosa*. Transcriptional fusions revealed RetS, LadS and GacA all appear to have a

significant effect as activators of both the *rsmN* and *nmsR* promoters. 2-Alkyl-4(1*H*)-quinolone (AQ) signalling also modulate *rsmN* expression possibly via the iron chelating properties of 2-alkyl-3-hydroxy-4(1*H*)-heptyl-quinolone (PQS). RsmN targets identified from Deep Sequencing include those required for structural outer membrane proteins, transcriptional regulators as well as genes involved in motility, secretion, flagellar structure and biofilms. RsmA, RsmZ and RsmY were all identified as targets together with the small RNAs RgsA (indirectly *gac*-controlled) and the antagonistic RNA CrcZ (represses catabolite repression control protein Crc). Targets common to both RsmN and RsmA include the transcriptional regulators Vfr, PqsR, MvaT and Anr, regulatory RNAs RsmZ and RsmY together with transcripts corresponding to the *pqsABCDE* operon, LasB, LecA/B, RhII, LasR/I, Crc and CrcZ.

The identification of many mRNA targets for RsmN which are shared with targets of RsmA provides further evidence that RsmN is involved in global-post-transcriptional regulation of gene expression.

## ABBREVIATIONS

C4-HSL	<i>N</i> -butanoyl-L-homoserine lactone
3-oxo-C12-HSL	<i>N</i> -(3-oxododecanoyl)-L-homoserine lactone
μl	Micro litre
AHLs	<i>N</i> -Acyl-Homoserine Lactones
ANR	Arginine fermentation transcription factor
Ap <sup>R</sup>	Ampicillin resistant
APS	Ammonium persulfate
AQs	2-alkyl-4(1 <i>H</i> )-quinolones
asRNA	Antisense ribonucleic acid
bp	Base Pair
CAP	Catabolite Gene Activator Protein
CD	Circular Dichroism
Cfu	Colony forming units
cDNA	Complementary deoxyribonucleic acid
CDS	Coding Sequence
Cm <sup>R</sup>	Chloramphenicol resistant
CSR	<u>C</u> arbon <u>S</u> torage <u>R</u> egulator
CTAB	Cetyl trimethylammonium bromide
Deep-seq	Deep Sequencing
DEPC	Diethyl pyrocarbonate
DIG	Digoxigenin
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic Acid

DNase	Deoxyribonuclease
dNTP	Deoxyribonucleotide triphosphate
DoF	Degrees of Freedom
dsRNA	Double stranded ribonucleic acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic Acid
EMSA	Electrophoretic mobility shift assays
ESI-MS	Electrospray Ionization Mass Spectroscopy
FPLC	Fast Performance Liquid Chromatography
g	Gram
<i>g</i>	Relative centrifugal force
Gac	<u>G</u> lobal <u>a</u> ctivator of secondary metabolism
GacA	<u>G</u> lobal activator of <u>a</u> ntibiotic and <u>c</u> yanide production
GdCl	Guanidinium Chloride
GF	Gel Filtration or Size-Exclusion Chromatography (SEC)
Gm <sup>R</sup>	Gentamicin resistant
h	Hour (s)
HCl	Hydrochloric acid
HCN	Hydrogen Cyanide
HD	Heterodimer
HHQ	2-heptyl-4-quinolone
HSL	Homoserine lactone
HSQC	Heteronuclear Single Quantum Coherence
HPLC	High Pressure Liquid Chromatography
IPTG	Isopropyl- $\beta$ -D-Thiogalactopyranoside

ITC	Isothermal titration microcalorimetry
kDa	kilo Daltons
KP	Potassium Phosphate
L	Litre
LadS	<u>L</u> ost <u>a</u> dherence
LB	Luria Broth
lecA	Lectin PA-IL
M	Molar
mA	milli ampere
MAD	Multiple wavelength Anomalous Diffraction
MCS	Multiple cloning site
min	Minute (s)
ml	milli litre
MOPS	4-Morpholinepropanesulfonic acid
mRNA	Messenger Ribonucleic Acid
N	Native or folded state
Ni-NTA	Nickel – nitrilotriacetic acid
NMR	Nuclear Magnetic Resonance
NOEs	Nuclear Overhauser Effect
Nt	Nucleotide (s)
OD	Optical Density
o/n	overnight
ORF	Open reading frame
PAGE	Polyacrylamide Gel Electrophoresis
PAP	Poly(A) polymerase

PBS	Phosphate-buffered saline
P <sub>cons</sub>	Constitutive promoter
PCR	Polymerase Chain Reaction
PDB	Protein data bank
PGA	Polysaccharide adhesion
P <sub>ind</sub>	Inducible promoter
pL	Lysogenic-phase promoter
pMol	Pico mol
PQS	Pseudomonas quinolone signal (2-heptyl-3-hydroxy-4(1H)-quinolone)
pR	Lytic-phase promoter
p.s.i	Pounds per square inch pressure
PTSB	Peptone tryptone soy broth
QS	Quorum Sensing
RLU	Relative light units
rpm	Revolutions Per Minute
RBS	Ribosome binding site
RetS	<u>R</u> egulator of <u>e</u> xopolysaccharide and <u>t</u> ype III secretion
RNA	Ribonucleic Acid
RNase	Ribonuclease
rNTP	ribonucleotide triphosphate
RSM	<u>R</u> egulator <u>S</u> econdary <u>M</u> etabolites
s	Seconds
SD	Shine Dalgarno
SDev	Standard Deviation

SDM	Site Directed Mutagenesis
SDS	Sodium Dodecyl Sulphate
SELEX	Systematic Evolution of Ligands by Exponential
Enrichment	
SEC	Size Exclusion Chromatography or Gel Filtration (GF)
Sm <sup>R</sup> /Sp <sup>R</sup>	Streptomycin/spectinomycin resistant
sRNA	Small Ribonucleic Acid
SSC	Sodium Chloride / Sodium Citrate
STET	Tris-HCl/EDTA
TAE	Tris-Acetate-EDTA
TBE	Tris base, boric acid and EDTA
TBS	Tris-buffered saline
TEMED	N,N,N',N'-Tetramethylethylenediamine
Tet <sup>R</sup>	Tetracycline resistant
Thr	Thrombin
tRNA	Transfer RNA
U	Unfolded or denatured state
UTR	Untranslated region
UV	Ultraviolet
V	Volts
Vol	Volume
v/v	Volume per volume
v/w	Volume per weight
X-gal	5-bromo-chloro-3 indoyl β-D-galactoside
wt	wild type

# **1 INTRODUCTION**

## **1.1 BACTERIAL VIRULENCE**

### **1.1.1 Bacterial Pathogenicity**

Pathogenicity is the ability of a pathogen to cause an infectious disease in a host organism. The virulence of a microorganism is a measure of the severity of the disease it causes and can be investigated using genetic, biochemical and/or structural elements that promote disease production. The means by which pathogenic bacteria cause acute disease is characterised by two mechanisms. The first is invasiveness, encompassing the mechanisms of colonization, the production of extracellular substances that facilitate invasion (invasins) and the ability to circumvent host defence mechanisms (Niemann et al., 2004). The second is toxigenesis, the ability of the pathogen to produce toxins, which can act at the site of invasion or on other tissues sites away from the bacterial growth.

### **1.1.2 Virulence of *Pseudomonas aeruginosa***

*P. aeruginosa* is a Gram-negative, aerobic rod-shaped bacterium which inhabits a diverse range of environments such as soil, water, plants and animals (including humans). It is an opportunistic human and plant pathogen which has been extensively studied. In humans *P. aeruginosa* is a leading cause of nosocomial infections, especially in immuno-compromised hosts such as burn victims and cancer patients (Van Delden and Iglewski, 1998). It is also the predominant cause of morbidity and mortality in cystic fibrosis patients,

whose abnormal airway epithelia allow long-term colonization of the lungs causing serious and often fatal complications (Stover et al., 2000, Fagerlind et al., 2005). *P. aeruginosa* also colonises medical equipment and forms biofilms on catheters, contact lenses and many other devices; this organism is very problematic because of a resistance to many drug classes and its ability to acquire resistance after exposure to antimicrobial agents. It has been noted that multi-antibiotic resistance is rapidly increasing (Van Eldere, 2003). Most antibiotics were developed to either kill bacteria (bactericidal) or stop them from dividing (bacteriostatic), however more recently strategies to control bacterial infections have involved the attenuation of virulence (Camara et al., 2002, Finch et al., 1998).

Bacteria have a phenomenal ability to adapt to their environment which is why infections are often persistent and treatments frequently unsuccessful. They can survive in many different ecological niches, a factor which is enhanced by their ability to utilise different energy sources (Lyczak et al., 2000). The genome of a number of *P. aeruginosa* strains have been sequenced e.g. (Stover et al., 2000), revealing a genome size of ~6 million base pairs (bp) coding for over 5,500 genes, of which up to 10 % are dedicated to regulation. This suggests a high order of complexity which may explain the versatility that this organism shows.

#### *1.1.2.1 Motility in P. aeruginosa*

The different modes of motility of *P. aeruginosa* enhance the ability to mobilize, colonize a wide range of environments, attachment of bacteria to surfaces and biofilm formation, influencing the virulence of the bacterium (O'Toole and Kolter, 1998). *P. aeruginosa* is capable of three different types

of motility: flagellum-mediated swimming in aqueous environments and at low agar concentrations (<0.3% [wt/vol]); type IV pilus-mediated twitching on solid surfaces or interfaces; and swarming on semisolid (viscous) media (0.5 to 0.7% [wt/vol] agar)(Déziel et al., 2003, Köhler et al., 2000, Rashid and Kornberg, 2000). Swarming is described as a social phenomenon involving the coordinated and rapid movement of bacteria across a semisolid surface, often typified by a dendritic-like colonial appearance. Recently, it was shown that swarming of *P. aeruginosa* is dependent on both flagella and type IV pili as well as the presence of rhamnolipids and it is induced under nitrogen limitation and in response to certain amino acids (e.g., glutamate, aspartate, histidine, or proline) when provided as the sole source of nitrogen (Köhler et al., 2000, Overhage et al., 2007). *P. aeruginosa* swarmer cells are elongated and can possess two polar flagella (Rashid and Kornberg, 2000). In addition to these physical changes, swarmer differentiation can also be coupled to increased expression of important virulence determinants in some species (Fraser and Hughes, 1999, Kim et al., 2003, Rather, 2005).

## **1.2 REGULATION OF VIRULENCE**

### **1.2.1 Virulence regulation at the transcriptional level**

#### **1.2.1.1 Bacterial cell-to-cell communication**

The production of extracellular products, most of which act as virulence factors, is positively controlled in *P. aeruginosa* via a quorum sensing (QS) system. Quorum sensing is a bacterial communication system using small, diffusible signal molecules. This class of cell-to-cell communication is

population-density dependent, whereby the detection of accumulated signal molecules at a threshold concentration enables a single bacterial cell to sense population density. The QS mechanism is used by bacteria to co-ordinate their behaviour towards environmental changes to enhance survival. These responses include adaptation to availability of nutrients, defence against other microorganisms and the avoidance of potentially dangerous toxic compounds. This response mechanism is very important for pathogenic bacteria during infection as it enables them to co-ordinate the expression of virulence genes in order to overcome host immune responses and subsequently to establish a successful infection.

Bacteria produce and release QS signals (sometimes termed “autoinducers”) into the surrounding medium until a “quorum”, or minimum concentration threshold is reached. When this occurs the QS signal molecules interact with their respective cognate receptors, which in turn activate or repress the transcription of genes coding for example for secondary metabolites and virulence factors (Winzer et al., 2000). Processes controlled by QS are often those that are unproductive when undertaken by an individual bacterial cell, which become effective only when undertaken by the population. These processes include competence and luminescence (see below), but also virulence factor expression and secretion, biofilm formation and sporulation.

#### 1.2.1.2 QS-dependent control of gene expression

Intercellular communication within a bacterial population was first postulated in the 1960s from studies of genetic competence in *Streptococcus pneumoniae* by Tomasz (previously known as *Pneumococcus*) and on bioluminescence in

*Vibrio fischeri* by Hastings (Tomasz, 1965, Nealson et al., 1970). QS has been extensively studied in the symbiotic Gram-negative marine bacterium *V. fischeri*, in which it controls bioluminescence. Hastings demonstrated that light was produced at high cell population densities but not in dilute suspensions, and that light production could be stimulated by the exogenous addition of cell-free culture fluids. The chemical responsible, was called an autoinducer, and was later identified as an *N*-acyl-homoserine lactone (Eberhard, 1972).

#### 1.2.1.3 Quorum sensing signalling molecules

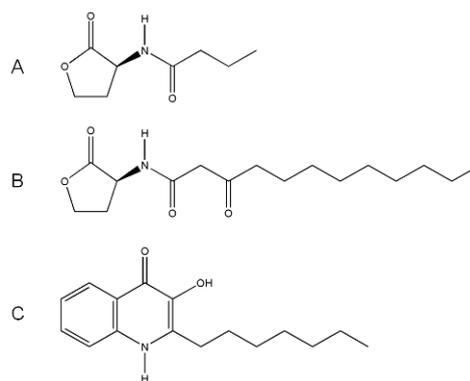
Gram-negative bacteria such as *V. fischeri* produce *N*-acyl-L-homoserine lactones (AHLs), which are the products of autoinducer synthases, which are usually homologues of the LuxI protein originally found in *V. fischeri*. When the bacterial population increases and the signal molecule concentration reach a minimum threshold, the signals are detected by LuxR, a response regulator protein. The interaction of LuxR with a cognate signal molecule leads to the formation of a complex that binds to a specific DNA sequence present in the promoters of target genes, the so-called *lux* box, thereby increasing transcription. In contrast, Gram-positive bacteria, such as *Staphylococcus aureus* and *Bacillus subtilis*, employ small peptides that often contain chemical modifications as QS signalling molecules (Okada et al., 2005, Kleerebezem et al., 1997). AHLs and peptides represent the two major classes of known bacteria cell to cell signalling molecules.

#### 1.2.1.4 Transcriptional virulence regulation in *P. aeruginosa*

The production of extracellular products, many of which act as virulence factors, is regulated in *P. aeruginosa* via two main linked QS systems termed *las* and *rhl* (Bassler, 2002, Heurlier et al., 2004). Two major AHLs are produced as QS signal molecules by *P. aeruginosa* that are involved in these two systems. These AHLs activate the transcriptional regulators LasR and RhlR respectively, which in turn induce the AHL synthase LasI (Gambello et al., 1993) or RhlI (Latifi et al., 1995). The two pairs of transcriptional regulators and AHL synthases are homologues of, respectively, LuxR and LuxI from *V. fischeri*.

LasI directs the synthesis of *N*-(3-oxododecanoyl)-L-homoserine lactone (3-oxo-C12-HSL) whereas RhlI is responsible for the synthesis of *N*-butanoyl-L-homoserine lactone (C4-HSL) (Pesci et al., 1997, Winson et al., 1995). These AHLs can bind and subsequently activate their cognate receptor proteins LasR and RhlR, respectively, which in turn bind to the promoters of the AHL synthase genes and increase their transcription. 3-oxo-C12-HSL and C4-HSL are both autoinducers because they are responsible for stimulating their own synthesis via a positive feedback system (Fig. 1.1)(Seed et al., 1995).

In addition to the AHL-based QS systems, a third, distinct autoinducer regulatory system has also been identified in *P. aeruginosa*, based on the 2-alkyl-4(1*H*)-quinolones (AQs).



**Figure 1.1: Quorum sensing signal molecules in *P. aeruginosa*.**

A) C4-HSL, *N*-butanoyl-L-homoserine lactone, (B) 3-oxo-C12-HSL, *N*-(3-oxododecanoyl)-L-homoserine lactone and (C) PQS, *Pseudomonas* quinolone signal, 2-heptyl-3-hydroxy-4(1*H*)-quinolone.

This directly activates two operons (*phnAB* and *pqsABCDE*) which are required for the biosynthesis of 2-alkyl-4-quinolones (AQs), including molecules involved in AQ signalling and the activation of QS-controlled genes via *pqsE* (Deziel et al., 2005, Lépine et al., 2003). The *pqsABCDE* operon (PA0996-PA1000) is adjacent to the anthranilate synthase genes *phnAB* (PA1001-1002) and *pqsR* (*mvfR*, PA1003). The genes *pqsH* (PA2587) and *pqsL* (PA4190) are also involved in AQ biosynthesis but are located separately elsewhere on the chromosome. Among the AQs is the Pseudomonas Quinolone Signal (PQS) which acts as an activator of PqsR, inducing a positive feedback loop typical of many QS systems (Heeb et al., 2011, Xiao et al., 2006). The PQS precursor, 2-heptyl-4-quinolone (HHQ) has been shown to act as an autoinducer (Diggle et al., 2007) in addition to PQS, other longer alkyl chain AQs can induce PqsR-dependent gene expression but more weakly (Xiao et al., 2006). HHQ been suggested to induce a conformational change in PqsR as its presence enhances the binding of PqsR to the *pqsA* promoter *in vitro*.. PQS has been reported to be more than 100 times more potent at inducing the *pqsA* promoter than HHQ (Xiao et al., 2006, Diggle et al., 2007).

The regulation of virulence factors by AQS was first demonstrated by the positive impact of PQS on the *lasB* (elastase) gene). The presence of PQS is required for the expression of *lecA* and pyocyanin production. The synthesis of PQS requires the *pqsABCDE* operon to remain intact, however *pqsE* mutants produce parental levels of AQS but do not exhibit any PQS-associated phenotypes (Gallagher et al., 2002, Diggle et al., 2003). PqsE is concluded to facilitate the response to PQS and is therefore essential for the expression of genes such as *lecA* and the *phz* pyocyanin biosynthesis (Fletcher et al., 2007).

The involvement of AQS in regulation is highly complex as both RhlR and RpoS are essential for *lecA* expression, as the addition of PQS to the corresponding mutants failed to restore *lecA* transcription (Winzer et al., 2000). Diggle et al., (2003) demonstrated that PQS can overcome the repression of *lecA* by the H-NS-type protein, MvaT and the post-transcriptional regulator, RsmA. It has also been shown that PQS, but not HHQ, can induce transcription of the small regulatory RNA, RsmZ. Therefore PQS can act on the expression of virulence genes at both the transcriptional and post-transcriptional levels (Heeb et al., unpublished data).

There is a hierarchy between the *las* and *rhl* QS systems (Antunes et al., 2010) where LasR has been defined as the master regulator (Fig. 1.2). The *las* system directly regulates the *rhl* system, exerting transcriptional control over *rhlR* and *rhlI* (Latifi et al., 1995, Winzer et al., 2000). The QS cascade in *P. aeruginosa* involves some additional regulatory factors, such as the PQS (Heeb et al., 2011, Diggle et al., 2003) which provides a supplementary link between the *las* and the *rhl* systems (Juhás et al., 2005). Additional factors can modulate QS activity in *P. aeruginosa*. For example, QscR is an orphan LuxR



potential two-component systems, being one of the largest number present in any bacterial genome sequenced so far and reflecting the significant adaptability that *P. aeruginosa* has to a variety of environmental niches (Rodrigue et al., 2000). A two-component system typically consists of a sensor kinase and a cognate response regulator.

The GacS/GacA two-component system is conserved in *Pseudomonas* spp. and other Gram-negative bacteria, where “gac” designates ‘global activator of secondary metabolism’. GacS/GacA homologues have been identified in *E. coli* (BarA/UvrY), *Salmonella* (BarA/SirA), *Erwinia* (ExpS/ExpA) and *Vibrio* (VarS/VarA) as well as in the following Pseudomonads: *P. fluorescens*, *P. aeruginosa*, *P. syringae* and *P. aureofaciens* (Laville et al., 1992, Reimmann et al., 1997, Hrabak and Willis, 1992, Chancey et al., 1999).

GacS was first described in the plant pathogen *Pseudomonas syringae* B728a as LemA and identified as an essential factor for lesion manifestation on bean leaves, where inactivation of the *gacS* gene resulted in the loss of virulence (Hrabak and Willis, 1992, Hirano et al., 1997). GacA, the cognate response regulator, was first identified as a global activator of antibiotic and cyanide production in *P. fluorescens* CHA0 (Laville et al., 1992).

The GacS/GacA system is characterised by autophosphorylation, receiver and histidine phosphotransfer (Hpt) output domains (Rodrigue et al., 2000). GacS is activated by an as yet unknown signal, leading to auto-phosphorylation and then phosphoryl group transfer onto the response regulator GacA. GacS/GacA positively control the expression of genes involved in the production of a variety of secondary metabolites, extracellular products and virulence factors in *P. aeruginosa* (Reimmann et al., 1997, Pessi and Haas, 2001). QS

molecules are also regulated by this system in some pseudomonads, as demonstrated by the production of C4-HSL in *P. aeruginosa* (Reimann et al., 1997).

GacA, like other response regulators, has a C-terminal helix-turn-helix DNA-binding domain, however the DNA binding sequence that is recognised by phosphorylated GacA and its directly controlled target genes is still unknown. The GacS/GacA two-component system acts at a post-transcriptional level controlling target genes indirectly, with a region near to or at the RBS (ribosome binding site) of some target genes having been identified as necessary for GacA and RsmA control (Blumer et al., 1999).

The mechanism by which this two component system controls the expression of target genes is via a post-transcriptional network involving RNA-binding proteins and the transcription of small, untranslated regulatory RNAs.

#### 1.2.1.6 Regulation by the Csr/Rsm System

##### 1.2.1.6.1 Role of RsmA in gene expression

The CsrA/RsmA family of RNA-binding proteins are global post-transcriptional regulators that bind to target mRNAs, affecting their translation and/or their stability and mediating the resulting changes in gene expression. This function is modulated by small, untranslated RNAs that are able to titrate out the RNA binding proteins away from the target mRNAs, and via this mechanism control translation and mRNA stability.

The Csr (carbon storage regulator) system was first discovered in *E. coli* and characterised as a negative regulator of glycogen metabolism and glycolysis

and a positive regulator of motility, modulating expression of the *flhDC* operon, responsible for the control of flagellar biosynthesis (Romeo et al., 1993, Yang et al., 1996, Wei et al., 2001). CsrA has also recently been shown to inhibit translation initiation of *hfq*, a gene encoding an RNA chaperone that mediates sRNA-mRNA interactions (Baker et al., 2007).

In *Erwinia* spp., the CsrA homologue RsmA (repressor of secondary metabolites) was identified as a global repressor of the production of extracellular enzymes, AHL molecules and pathogenicity (Cui et al., 1995). Flagellar formation and bacterial movement are regulated in many enterobacteria by the master regulator of flagellar genes *flhDC* and *fliA*, a flagellum-specific  $\sigma$  factor. Recent work has demonstrated that motility in *E. carotovora* subsp. *carotovora* is positively regulated by the quorum-sensing signal, *N*-3-(oxohexanoyl)-L-homoserine lactone (3-oxo-C6-HSL), and negatively regulated by RsmA (Chatterjee et al., 2010, Chatterjee et al., 1995). Members of the Csr/Rsm family play important functional roles in post-transcriptional regulation in many other bacterial genera. These include regulating gene expression required for host-cell interactions and environmental adaptation in *Salmonella typhimurium* (Altier et al., 2000), for swarming motility in *Serratia marcescens* (Ang et al., 2001), for transmissibility, cytotoxicity and efficient macrophage infection in *Legionella pneumophila* (Fettes et al., 2001), for swarming motility and virulence in *Proteus mirabilis* (Liaw et al., 2003) and for lipooligosaccharide production in *Haemophilus influenzae* (Wong and Akerley, 2005). The importance of this family of post-transcriptional regulators is further highlighted by the fact that it is present in the highly adapted human gastric pathogen *Helicobacter pylori*,

which has relatively few transcriptional regulators and where it controls virulence and the stress response (Barnard et al., 2004).

RsmA together with a second RNA-binding protein RsmE (72 % identity), is involved in the post-transcriptional control of secondary metabolism regulated by the GacS/GacA system in *P. fluorescens* CHA0, controlling negatively the production of exoenzymes and antifungal secondary metabolites such as hydrogen cyanide. In *P. aeruginosa*, RsmA can act as both a positive and a negative regulator. RsmA negatively regulates the production of hydrogen cyanide, pyocyanin, LecA (PA-IL) lectin and AHLs, whereas it positively regulates swarming motility, lipase and rhamnolipid production (Heurlier et al., 2004).

#### 1.2.1.6.2 RsmA Structure

The structure of the *Yersinia enterocolitica* RsmA has been solved using X-ray crystallography, revealing a novel RNA-binding site (Heeb et al., 2006). Many RNA-binding proteins contain a KH domain and many, but not all members of the RsmA family contain a sequence (VLGVKGXXVR) similar to the KH motif. On comparison of the structural data, it was demonstrated that the RsmA family members contain a novel structural motif (Fig. 1.3, Heeb *et al.* 2006).



**Figure 1.3: RNA-binding domain structure comparison.**

Comparison of the *Y. enterocolitica* binding protein RsmA (A) and the KH-domain eukaryote neuronal protein Nova (B). Amino acids conserved between the two proteins with respect to the KH domain are shown in red (Heeb *et al.*, 2006).

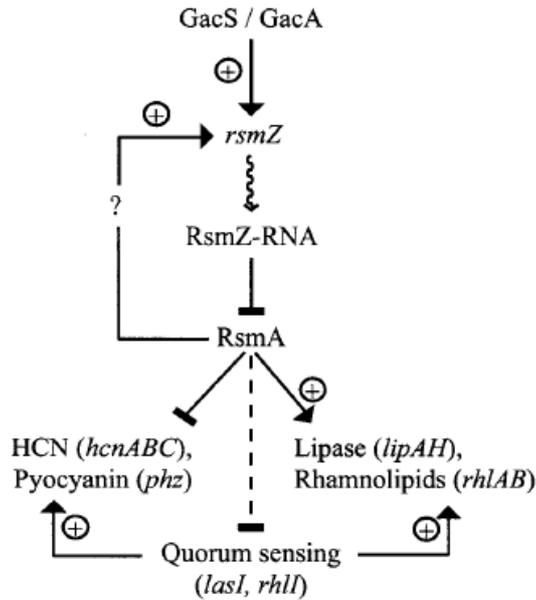
The functional unit of RsmA is a dimer with each subunit consisting of five-stranded antiparallel  $\beta$ -sheets and an  $\alpha$ -helix. The three central strands form the hydrophobic core by hydrogen-bonding to each other in the order 2-3-4 with extensive hydrophobic residues throughout the core. The other two  $\beta$ -strands are peripheral, where  $\beta$ 1 is hydrogen bonded to  $\beta$ 4 of the other strand, and  $\beta$ 5 is hydrogen bonded to  $\beta$ 2 in the other monomer. The  $\alpha$ -helices project out from the  $\beta$  sheets, the N-terminal of which interacts with the rest of the protein and are important for retention of structure. The R44 residue was unequivocally demonstrated to be the key residue involved in target RNA binding and is strictly conserved in all RsmA/CsrA sequences. It is close to other solvent exposed residues such as R7, L26 and R36. As RNA-binding sites often contain positively charged amino acids, therefore this domain in the protein is a good candidate for an RNA-binding site.

In *P. fluorescens* the NMR solution structure of RsmE, an RsmA homologue, was obtained in complex with a target RNA (Schubert *et al.*, 2007). The importance of R44 residue is confirmed by demonstration that the phosphate

backbone of the target RNA hexanucleotide loop is stabilized by four positively charged lysine and arginine side chains (Arg31, Lys38, Arg44 and Arg50).

#### 1.2.1.6.3 The *P. aeruginosa* Regulatory RNAs, RsmZ and RsmY

RsmA and its homologue CsrA have previously been shown to act as post-transcriptional regulators by binding to target mRNAs: this mechanism controls the transcription and stability of the mRNAs. RsmA can be sequestered by either of the small, untranslated regulatory RNAs RsmZ and RsmY, whose functions are analogous to those of CsrB and CsrC in *E. coli*, therefore antagonising RsmA activity (Fig. 1.4)(Kay et al., 2006, Liu et al., 1997, Weilbacher et al., 2003). The effects of RsmA depend on the GacS/GacA two-component system, as this system controls the expression of *rsmZ* and *rsmY* (Heurlier et al., 2004). These non-coding RNAs are also activated by RsmA which results in a negative feedback loop, affecting RsmA activity (Kay et al., 2006, Bejerano-Sagie and Xavier, 2007). Activation of the GacS/GacA system results in RsmA inactivation.



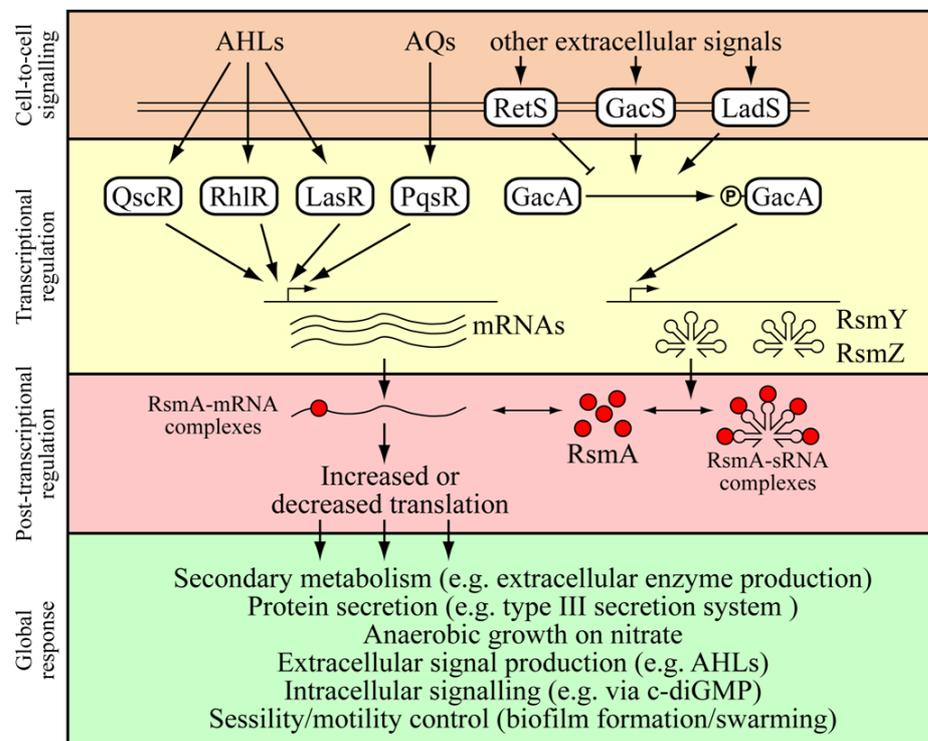
**Figure 1.4: Model of the GacA/RsmA signal transduction pathway in *P. aeruginosa* PAO1.**

Expression of the untranslated regulatory RNA, RsmZ depends on the presence of GacA. The function of RsmZ is to antagonize the action of the small RNA-binding protein RsmA. RsmA positively controls *rsmZ* expression, thus forming a negative autoregulatory circuit whose mechanism is not understood at present. RsmA also negatively controls AHL-dependent QS as well as a number of QS-dependent genes, some of which code for secondary metabolites and virulence determinants; these are regulated indirectly at the transcriptional level via QS but probably also directly at the translational level, as is the case for *hcnA* (Pessi and Haas, 2001). Lipase and rhamnolipid production are controlled positively by RsmA, independently of the quorum-sensing control. Dotted line, modulating negative effect; solid bar, negative effect; arrow, positive effect (Heurlier et al., 2004).

#### 1.2.1.6.4 Additional control of the Rsm system by RetS and LadS

In addition to the GacS sensor kinase, two additional elements have been identified that control, together with GacA, the transcription of *rsmY* and *rsmZ*. These consist of unconventional sensor kinase-response regulator hybrid proteins, which have their sensor domains in the periplasm linked by a transmembrane region to the cytoplasmic histidine kinase and receiver domains. LadS (Lost adherence) was described as acting similarly to GacS and promoting biofilm formation which is generally more associated with chronic, persistent infections and simultaneously repressing the type III secretion system which is most needed in the acute stage of infection (Ventre et al.,

2006). The second regulator discovered was RetS (for regulator of exopolysaccharide and type III secretion) and interestingly this hybrid sensor kinase function opposes the effects of LadS and GacS (Goodman et al., 2004). Like the two other sensor kinases, RetS seems to constitute an environmentally sensitive switch, but activating acute virulence characteristics such as the type III secretion system and repressing the production of exopolysaccharides necessary for biofilm formation. Both, LadS and RetS regulate transcription of *rsmZ* and *rsmY*; LadS like GacS positively controls their expression whereas RetS exerts negative control (Fig. 1.5).

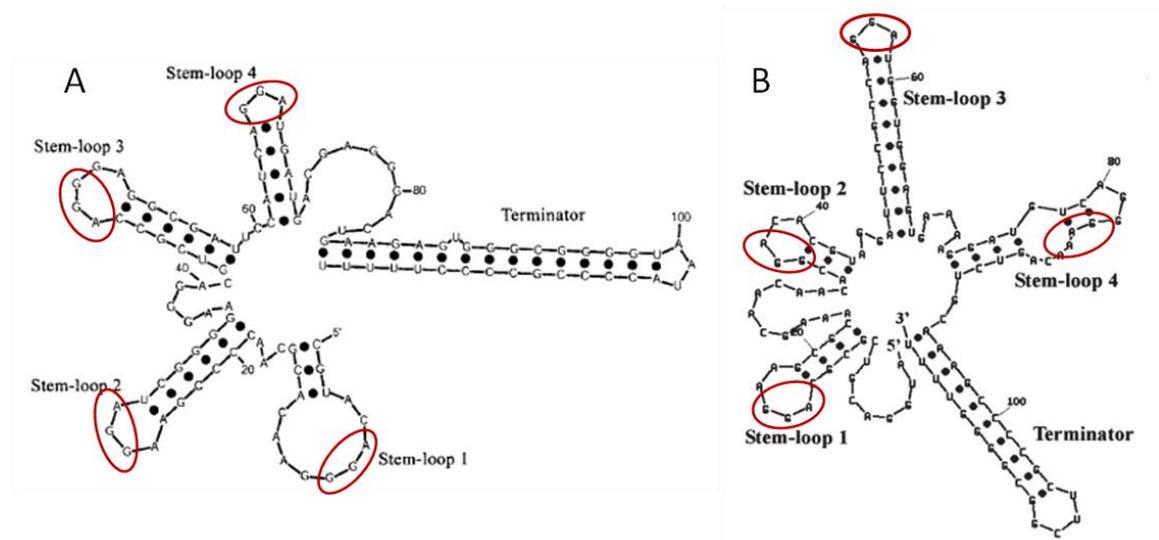


**Figure 1.5: Summary of gene regulation in *P. aeruginosa*.**

Gene regulation in *P. aeruginosa* is complex and functions at several levels. This diagram aims to display the links between the different levels without being exhaustive. Cell-cell signalling molecules involve the QS molecules (AHLs and AQS) but also some unidentified signal(s) stimulating the regulators RetS, LadS and GacS which in turn activate or repress the response regulator GacA which activates transcription of the regulatory RNAs RsmY and RsmZ. The QS signal molecules bind to the regulators RhlR, LasR and QscR, activating the two QS systems in *P. aeruginosa*, which regulate expression of many genes. The regulatory RNAs (RsmY and RsmZ) control the post-transcriptional regulator RsmA that in turn represses or activates target mRNAs, which leads to increased or decreased translation. Affected are amongst others many secondary metabolites, anaerobic growth, signal molecule production, motility, biofilm formation and also restriction (S Heeb, personal communication).

### 1.2.1.7 Regulatory RNA structures

Between a variety of Pseudomonads, the nucleotide sequence conservation of RsmZ is only about 45 %, however some highly conserved predicted secondary structures suggest they have analogous modes of action (Heurlier et al., 2004). The small regulatory RNAs RsmZ and RsmY of *P. aeruginosa* and *P. fluorescens*, CsrB and CsrC of *E. coli* (Liu et al., 1997, Weilbacher et al., 2003) and RsmX of *P. fluorescens* (Valverde et al., 2003, Kay et al., 2005) all have a conserved secondary structure in spite of low sequence homology. The RNA structures are elaborate and their length varies from 112 to approximately 345 nucleotides, while the retention of the characteristic GGA motifs located in the loops of stem-loops structures is constant (Fig. 1.6). These repeated motifs enable multiple RsmA units to be sequestered by a single RNA transcript.



**Figure 1.6: Predicted secondary structure of regulatory RNAs RsmY and RsmZ.**

Predicted secondary structures of (A) RsmZ from *P. aeruginosa* at 37 °C (Heurlier et al., 2004) and (B) RsmY from *P. fluorescens* at 30 °C (Valverde et al., 2003) using the M-Fold (Zuker, 1989).

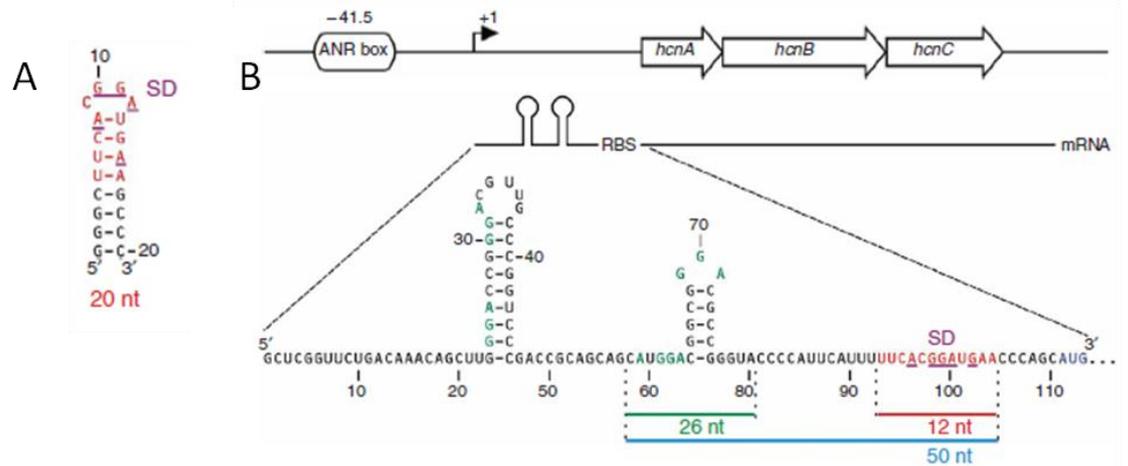
The optimal binding of CsrA to some sRNAs has been investigated (Dubey et al., 2005), using high-affinity RNA ligands containing a single CsrA binding site by systematic evolution of ligands by exponential enrichment (SELEX). This study revealed a consensus sequence (RUACARGGAUGU) where the ACA and GGA motifs were 100 % conserved and the GU sequence present in all but one of the experimental ligands. The majority of ligands contained GGA in the loop of short hairpins within the most stable predicted structure, the same as natural predicted CsrA binding sites (Fig. 1.7). The CsrA binding site consensus sequence for CsrC and CsrB is CAGGGAUG compared to the SELEX-derived sequence. Not all natural CsrA binding sites contain the GGA motif, in CsrB four are replaced with a GGG, while GGA is replaced with AGA in one of the *pgaA* binding sites. The *pgaA* gene is required for the synthesis of the polysaccharide adhesin (PGA), which plays an important role in biofilm formation in *E. coli* (Wang et al., 2005).

Part of the binding consensus sequence is found in the stem, therefore it was suggested that the hairpin structure partially melts after initial recognition, leading to additional base-specific contacts allowing interaction with the full consensus sequence (Dubey et al., 2005). This study did not however determine whether the CsrA dimer interacted with one or two binding sites.



solution structure of RsmE was determined as a complex with a target RNA containing the ribosome-binding site of the *hcnA* gene (encoding hydrogen cyanide synthase subunit A)(Schubert et al., 2007).

A 12-nucleotide sequence containing the RBS of the *hcnA* gene was used for the primary NMR experiments. The free RNA didn't form a stable stem loop structure, with base pairs only formed upon binding with RsmE (Fig. 1.8A). Transcription of the *P. fluorescens hcnABC* operon is under control of the anaerobic regulator of nitrate respiration and arginine fermentation (ANR) transcription factor (Fig. 1.8B). In order to obtain an NMR structure, the RNA sequence was extended to 20 nt, enabling the formation of a stem loop in the free RNA, resembling that of the other high-affinity ligands that bind to CsrA.



**Figure 1.8: Genetic organization of the *hcnA* 5' untranslated mRNA.**

A) Predicted secondary structure of the 20-nucleotide *hcnA* sequence used for structure determination of the RsmE–RNA complex. (B) Transcription of the *P. fluorescens hcnABC* operon is under control of the anaerobic regulator of nitrate respiration and arginine fermentation (ANR) transcription factor, which binds the ANR box. Highlighted in red is the 12-nucleotide *hcnA* sequence involved in RsmE binding, in green the other potential RsmE-binding sites, and in blue the AUG *hcnA* start codon; underlined, Shine-Dalgarno sequence (SD) of the RBS (Schubert et al., 2007).

The Heteronuclear Single Quantum Coherence (HSQC) spectra altered substantially upon binding with RNA, allowing excellent recognition of the

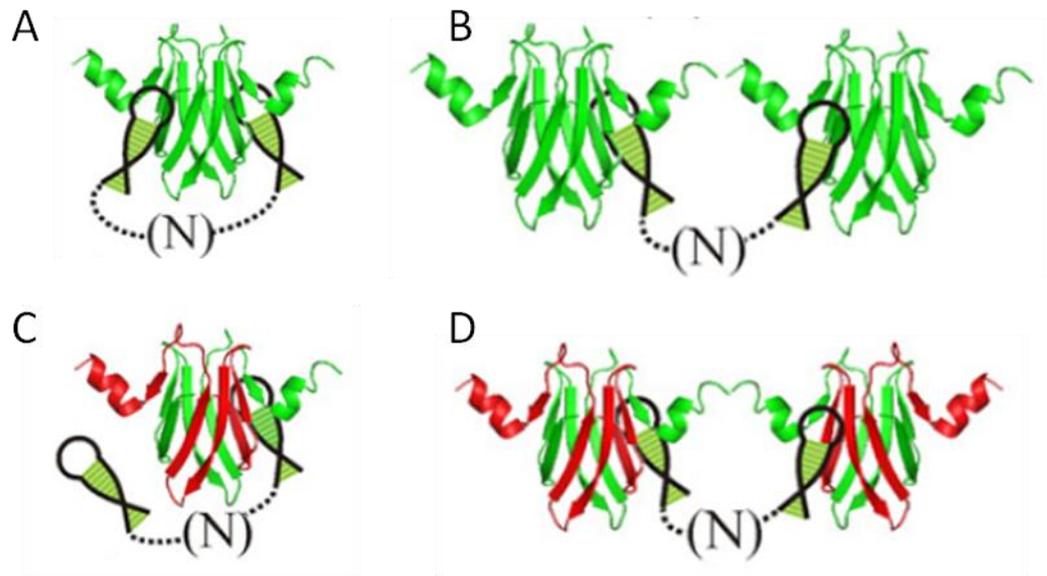
residues involved in binding. The RsmE homodimer has two binding sites and makes optimal contact with a 5'-<sup>A</sup>/<sub>U</sub>CANGGANG<sup>U</sup>/<sub>A</sub>-3' sequence within the RNA. When bound to RsmE the ANGGAN core folds into a loop structure, favouring the formation of a 3-base-pair stem. By binding specifically to the 5' <sup>A</sup>/<sub>U</sub>CANGGANG<sup>U</sup>/<sub>A</sub>-3' consensus sequence which closely matches the ideal 5'-AAGGAGGU-3' Shine Dalgarno (SD) sequence, the proteins of the RsmA/CsrA family can globally regulate the expression of numerous genes at the level of translation.

Five nucleotides of the *hcnA* SD sequence ACGGAUG are buried in the complex, either by contacts with the RsmE protein (ACGGAUG) or by base-pairing in the stem induced by protein binding (ACGGAUG). In the 5' untranslated region (5' untranslated region (UTR)) of *hcnA* in *P. fluorescens*, there are 4 GGA motifs upstream of the SD site. When all 4 motifs are mutated, translational regulation of *hcnA* by the Gac/Rsm system is abolished (K. Lapouge and D. Haas, unpublished data) It can be surmised that the upstream motifs as well as the motif overlapping the Shine-Dalgarno sequence are required for effective regulation by the Gac/Rsm system.

The SELEX method has also been used to probe the higher order binding properties of CsrA (Mercante et al., 2009). Using electrophoretic mobility shift assays (EMSA), the binding of CsrA to model RNAs demonstrated the formation of two complexes. The faster-minor consisted of CsrA with two bound RNAs and a slower-major complex of CsrA bound to a single RNA. CsrA can simultaneously bind at two target sites within a transcript when the sites are located as close together as 10 nt or as distant as 63 nt. The optimum

intersite distance was predicted to be 18 nt, with enough space to compensate for defects in either a secondary RNA target site or a CsrA binding surface, but not both. Below 18 nt, the spacing was detrimental for tight bridging sterically and binding to one of the target sites was easily displaced by the addition of excess CsrA, forming two CsrA dimers joined by a single RNA molecule. When the intersite distance was  $\geq 18$  nt, RNAs formed a stable bridge complex in wild type CsrA and neither of the bound target sites could be displaced by excess free CsrA. This result was found using model RNAs with the targets sited in a stable hairpin loop and might vary for unstructured or alternatively structured RNAs. CsrA binding at one site almost certainly leads to a cooperative interaction at an adjacent site under physiological conditions.

The study by Mercante *et al.*, 2009 also represented the first experimental demonstration of the function of dual RNA-binding sites of CsrA in regulation (Fig. 1.9). As well as the wild type (WT), a heterodimer was used (HD), where one of the binding surfaces had an alanine mutation at the R44 site, previously shown to be required for biological function (Heeb *et al.*, 2006). CsrA binds to the 5'-untranslated leader sequence of target transcripts and alters their translation and/or stability. The example used was the *glgCAP* 5'-leader, which has four RNA binding sites, only two of which had been previously characterized.



**Figure 1.9: Representations of CsrA-RNA binding combinations.**

The wild type (WT) is represented in green and the heterodimer (HD-R44A) in red, using high-affinity RNA ligands. Models depict the following; A: WT-CsrA bound to two target sites on same RNA, B: Two WT-CsrA molecules are joined by a bridging RNA, C: HD-CsrA where one RNA target site binds to the WT-functional surface and D: Two HD-CsrA molecules, where one RNA binds each target site to a functional binding site (Mercante et al., 2009).

Compared to the WT-CsrA, the HD-CsrA had only a third of the affinity for a single target. The heterodimeric CsrA, was ~14 fold less effective at repression using a *glgC*'-'*lacZ* reporter fusion. When a GGA site upstream of the RNA target was deleted, the difference in the HD-CsrA was unchanged, but relative to the WT-CsrA regulation decreased by 7 fold.

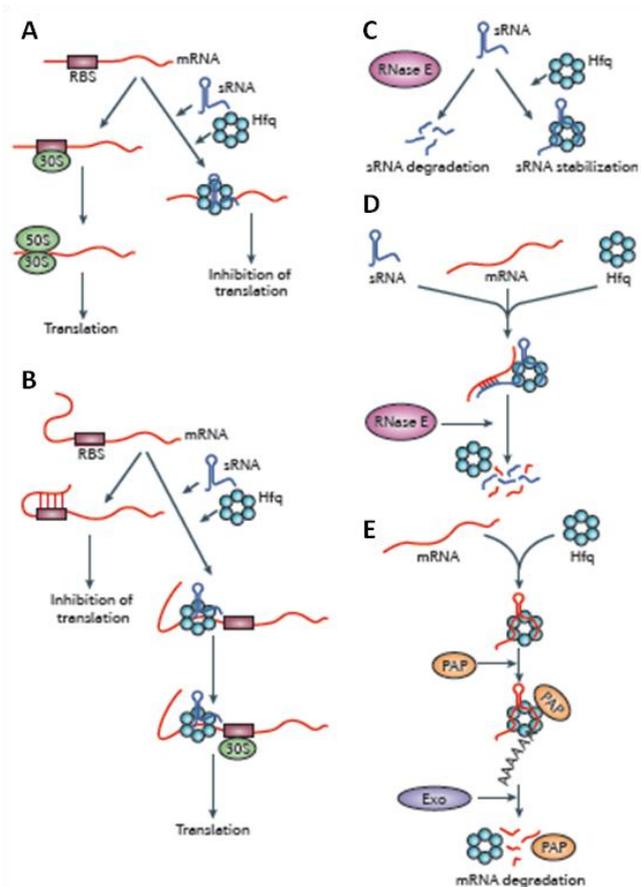
## 1.2.2 Gene regulation by sRNAs

### 1.2.2.1 sRNA Regulation

sRNAs can exert their action by base pairing with target transcripts and regulate gene expression post-transcriptionally, influencing translation or mRNA stability. The two major classes of sRNAs are *cis*-encoded and *trans*-encoded. *Cis*-encoded are encoded at the same genetic location as their target

but on the opposite strand to the RNAs they act upon. Trans-encoded sRNAs are normally found in a different chromosomal location and do not exhibit perfect base-pairing potential with their targets, with additional proteins often required in order to form a complex with their target.

The mechanisms for regulation, as mentioned above, are commonly of two types either influencing translation or effecting mRNA stability, although the precise mechanism of action depends on the structural information encoded in the RNA molecules. The RNA-binding protein Hfq mediates regulation using numerous mechanisms (Vogel and Luisi, 2011), demonstrating the complexity of sRNA regulation (Fig. 1.10). In the first mechanism Hfq can suppress protein synthesis by aiding a cognate sRNA to bind the 5' region of its target mRNA. This subsequently renders this 5' region inaccessible for translation initiation (Fig. 1.10A). Alternatively Hfq can enhance translation by guiding a sRNA to the 5' region of its target mRNA in order to disrupt a secondary structure that would otherwise inhibit ribosome binding (Fig. 1.10B). A third method of regulatory control occurs prior to the target recognition where Hfq can protect sRNAs from ribonuclease cleavage (Fig. 1.10C) or present some RNAs in such a way as to promote mRNA cleavage (Fig. 1.10D). In the last known mechanism Hfq can promote RNA turnover by rendering the 3' ends accessible for polyadenylation and subsequent 3'-to-5' exonucleolytic degradation (Fig. 1.10E).



**Figure 1.10: Widely accepted modes of Hfq activity.**

**A)** In association with a small RNA (sRNA) Hfq may sequester the ribosome-binding site (RBS) of a target mRNA, thus blocking binding of the 30S and 50S ribosomal subunits and repressing translation. **B)** Secondary structure in the 5' UTR can mask the RBS (Kozak, 2005) and inhibit translation. A complex formed by Hfq and a specific sRNA may activate the translation of one of these mRNAs by exposing the translation initiation region for 30S binding (Fröhlich and Vogel, 2009, Soper et al., 2010). **C)** Hfq may protect some sRNAs from ribonuclease cleavage, which is carried out by ribonuclease E (RNase E) in many cases. **D)** Hfq may induce the cleavage (often by RNase E (Massé et al., 2003, Morita et al., 2005, Pfeiffer et al., 2009) of some sRNAs and their target mRNAs. **E)** Hfq may stimulate the polyadenylation of an mRNA by poly(A) polymerase (PAP), which in turn triggers 3'-to-5' degradation by an exoribonuclease (Exo) (Mohanty et al., 2004, Hankins et al., 2010). In *E. coli*, the exoribonuclease can be polynucleotide phosphorylase, RNase R or RNase II (Vogel and Luisi, 2011).

As a consequence of advances in understanding sRNA regulation, it has become apparent the some fundamental mechanistic features are as yet undiscovered or approaches are just being made. Recently the number of known cellular targets of Hfq has increased, demonstrating the ability of Hfq to interact with numerous RNA species, with an evolutionarily conserved preference *in vivo* for sRNA and mRNA partners (Wassarman et al., 2001,

Zhang et al., 2003, Sittka et al., 2008). In addition to the modes of action, the behaviour of the sRNAs themselves are potentially more complex than previously believed. Whereas these RNAs were previously thought to be specific to a single target, increasing numbers have been shown to act on multiple mRNAs and consequently more mRNAs are emerging as shared targets of multiple cognate sRNAs (Beisel and Storz, Papenfort and Vogel, 2009).

### 1.2.2.2 RNomic Methods

High-throughput RNomic methods are providing new insights of the interplay between proteins and regulatory RNAs and the effect on the genome. RNA-Seq has several advantages over existing technologies, including that it is not limited to detecting transcripts that correspond to existing genomic sequences and can reveal the precise location of transcription boundaries, to a single base resolution (Comparison in Table 1.1).

Technology	Tiling microarray	cDNA or EST sequencing	RNA-Seq
<i>Technology specifications</i>			
Principle	Hybridization	Sanger sequencing	High-throughput sequencing
Resolution	From several to 100 bp	Single base	Single base
Throughput	High	Low	High
Reliance on genomic sequence	Yes	No	In some cases
Background noise	High	Low	Low
<i>Application</i>			
Simultaneously map transcribed regions and gene expression	Yes	Limited for gene expression	Yes
Dynamic range to quantify gene expression level	Up to a few-hundredfold	Not practical	>8,000-fold
Ability to distinguish different isoforms	Limited	Yes	Yes
Ability to distinguish allelic expression	Limited	Yes	Yes
<i>Practical issues</i>			
Required amount of RNA	High	High	Low
Cost for mapping transcriptomes of large genomes	High	High	Relatively low

**Table 1.1: Advantages of RNA-seq compared with other transcriptomic methods (Wang et al., 2009)**

Short RNA reads from 30 bp can provide information on how two or multiple exons are connected. A second advantage of RNA-Seq relative to DNA microarrays is that RNA-Seq has minimal background signal and no upper

limit for quantification. It has a large dynamic range of expression levels over which transcripts can be detected: in a study that analysed 16 million mapped reads in *Saccharomyces cerevisiae* a greater than 9,000-fold range was estimated (Nagalakshmi et al., 2008). RNA-Seq has also been shown to be highly accurate for quantifying expression levels, as determined using quantitative PCR (qPCR)(Nagalakshmi et al., 2008) and spike-in RNA controls of known concentration(Mortazavi et al., 2008).The results of RNA-Seq also show high levels of reproducibility, for both technical and biological replicates(Nagalakshmi et al., 2008, Cloonan et al., 2008). RNA-Seq also requires less RNA sample due to no cloning steps.

A major limitation of traditional sequencing for the discovery of small RNAs by cloning is that it is extremely challenging to identify small RNAs that are expressed at a low level, in restricted cell-types, or at very specific stages (Lu et al., 2007).

The generation of specialized cDNA libraries method for cloning ncRNAs, often by employing an antibody against the RNA-binding protein of interest to isolate entire populations of ncRNAs by immunoprecipitation, has disadvantages by the fact that it might not always be possible to reverse transcribe an ncRNA into cDNA because of its structure or modification (e.g. base or backbone modifications) and therefore will not reflect all ncRNAs present or their relative abundances (Vitali et al., 2003, Huttenhofer and Vogel, 2006). Also, some size-selected cDNA libraries might not identify all ncRNAs as the cut-off by size (e.g. 20–500 nt) will prohibit identification of longer ncRNAs. A cDNA expression library is only a true representation at a

particular developmental stage not taking into account all possible growth and nutrient conditions.

Alternatively, identification by enzymatic or chemically sequencing requires electrophoretic fractionation of the labelled fragments on denaturing polyacrylamide gels, followed by autoradiography which allows determination of the RNA sequence of interest (Sambrook and Russell, 2001, Bruce and Uhlenbeck, 1978). Disadvantages of this method are that, for identification, ncRNAs have to be highly abundant to be visible as single bands in ethidium-bromide stained gels and no other ncRNAs in the same size range should be present in the total RNA population, since it would hamper isolation of a single RNA species resulting in ambiguous sequencing data. Also results in sequencing data that are difficult to interpret, as well as limited to RNAs sized to the most, a couple of hundred nucleotides.

RNA-Seq is therefore the first sequencing based method that allows the entire transcriptome to be surveyed in a very high-throughput and quantitative manner. This method offers both single-base resolution for annotation and ‘digital’ gene expression levels at the genome scale, often at a much lower cost than either tiling arrays or large-scale Sanger EST sequencing.

These newer technologies constitute various strategies that rely on a combination of template preparation, sequencing and imaging, and genome alignment and assembly methods.

### 1.2.2.3 Cis-encoded natural Antisense RNA (asRNA)

High-throughput RNomic methods are providing new insights of the interplay between proteins and regulatory RNAs and the effect on the genome. The

regulation of gene expression via cis-encoded RNAs adds a further layer of complexity of control in bacteria. Naturally occurring anti-sense RNAs (asRNAs) were first observed in bacteria over thirty years ago (Itoh and Tomizawa, 1980, Lacatena and Cesareni, 1981). Antisense transcription has been observed in mice, *Saccharomyces cerevisiae* and *Drosophila melanogaster* (Group et al., 2005, David et al., 2006, Xu et al., 2009, Zhang et al., 2006).

#### 1.2.2.3.1 Previous limitations of the study of asRNA transcription

The deficiency of information regarding antisense transcription in bacteria from systematic genome wide analysis has been due to three technical problems, experimental and interpretational. The lack of robust bioinformatic algorithms to specifically predict asRNAs has been a hindrance together with the fact that the measurement of antisense transcription in microarray analyses was incorrectly identified as an experimental artefact generated during complementary DNA (cDNA) synthesis. The difficulty interpreting experimental data occurred as only low levels of transcription was reported to occur throughout the genome, leading to the conclusion that it was difficult to differentiate transcriptional noise from the asRNAs with regulatory functions (Selinger et al., 2000). Direct labelling of the RNA instead of cDNA prior to hybridization on tiled microarrays avoided unintentional second strand synthesis, and the stringent comparison of experimental results to computer predictions further strengthened the observation of asRNAs. These criteria, together with concentrating on highly expressed asRNAs, allowed for the confirmation that in a model cyanobacterium, *Synechocystis* PCC6803 the

experimentally confirmed highly expressed asRNAs increased from 1 to 73 (Dühring et al., 2006). The advance in high-throughput RNomics methods such as tiling microarrays, direct RNA-labelling and especially RNA deep sequencing, has changed the view of how antisense transcription can be investigated.

Recent studies have found that antisense transcription rates, for the respective transcriptomes have been determined to be approximately 4.7 % for *Vibrio cholerae*, 2.2 % for *Pseudomonas syringae* and 1.3 % for *Staphylococcus aureus* (Georg and Hess, 2011). Data from the examination of the compact genome of *Helicobacter pylori* found asRNAs for 46 % of all annotated ORFs, revealing antisense transcription to be an active, non-random process (Sharma et al., 2010).

#### 1.2.2.3.2 Types of antisense transcripts in bacteria

Bacterial asRNAs can only be roughly classified based on their location, as there is no conserved feature due to the diversity of bacterial asRNAs, apart from transcription occurring from the antisense strand of a known transcriptional unit. The categories are divided into 5'-overlapping (divergent, head to head), 3'-overlapping (convergent, tail to tail) or internally located asRNAs. Regulatory connections between neighbouring genes can occur with transcripts from protein-coding genes with long 5' or 3' untranslated regions (UTRs), which overlap substantially with the mRNAs originating from other genes. The size of asRNAs are diverse ranging from 100 nt (e.g., GadY (Opdyke et al., 2004)) to substantially larger at 700 – 3,500 nt or longer, even overlapping multiple genes (Stazic et al., 2011).

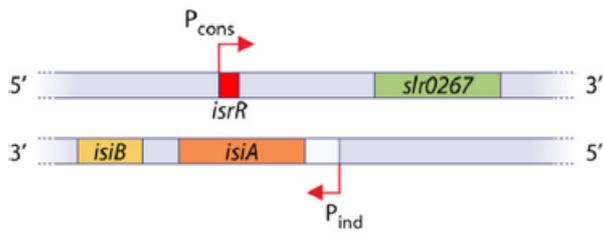
### 1.2.2.3.3 Mechanisms of asRNA action

Rapid progress is being made in the identification of chromosomally located cis-antisense RNAs, however knowledge of the molecular mechanisms by which these asRNAs act is only increasing slowly. Experimental analysis has revealed functional characteristics for phage- and plasmid-encoded asRNAs and multiple *trans*-acting non-coding RNAs (Brantl, 2007, Wagner and Simons, 1994).

#### 1.2.2.3.3.1 Alteration of target RNA stability

There are four broad categories which describe these mechanisms, the first of which acts by the alteration of target RNA stability. The interaction of an asRNA with its target RNA results in a duplex formation of double-stranded RNA (dsRNA) by alteration of the secondary structure of both molecules. These changes affect the stability of RNAs with a variety of possible outcomes. There can be rapid and complete degradation of both RNAs, a yield of a translationally inactive mRNA or a mature or stabilized form of mRNA. An example of codegradation is the *isiA*/IsrR sense/antisense pair in *Synechocystis* PCC6803 (Dühning et al., 2006). Regulation of *isiA* is tightly controlled by IsrR as the IsiA protein is involved in the iron stress response regulon and the expression of IsiA subsequently results in a massive reorganisation of the photosynthesis apparatus (Fig. 1.11).

### Co-Degradation: IsrR/IsiA



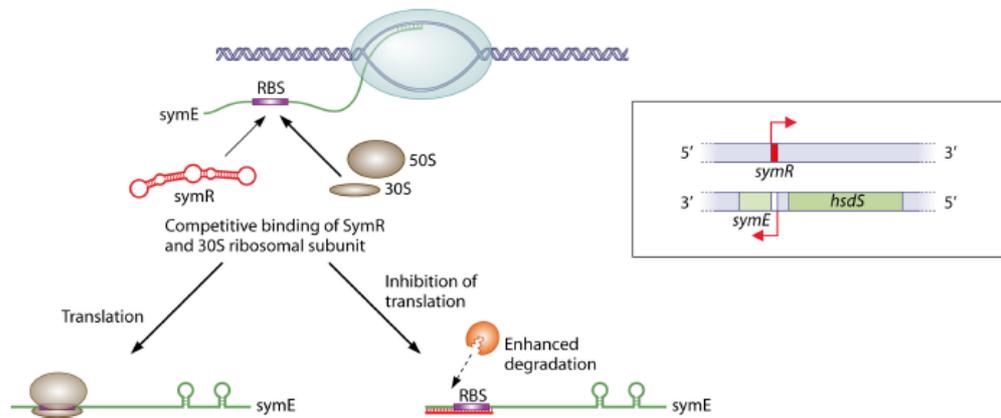
**Figure 1.11: The *isiA/IsiR* pair of *Synechocystis***

The asRNA *IsrR* originates from the central part of the *isiA* gene from a constitutive promoter ( $P_{cons}$ ). The *isiA* gene is under the control of the inducible promoter ( $P_{ind}$ ). Under early-stress conditions, *isiA* transcription becomes activated. Both transcripts are codegraded. The mRNA cannot accumulate as long as  $IsrR > isiA$ , and no protein is made.

The accumulation of transcripts is inversely related with both RNAs existing as almost exclusive species. When both species are expressed concurrently they form an RNA duplex which is immediately degraded, although the mechanism by which this occurs is unknown. The mRNA can only accumulate when the number of *isiA* mRNA molecules titrates out the number of asRNA molecules.

#### 1.2.2.3.3.2 Modulation of translation

Whereas the degradation/stabilization of RNA is of primary importance for the previous example, this becomes of secondary consequence to the suppression of gene expression. The regulation of the SOS response-inducible *SymE* protein in enterobacteria is an example of this type of mechanism (Kawano et al., 2005, Georg and Hess, 2011). This protein is believed to be a toxin-like RNA endonuclease which is under a strictly controlled and complex regulation. The asRNA *SymR* has been shown to be necessary for at least three repression mechanisms. This asRNA overlaps the 5' end of the *symE* mRNA, inclusive of the ribosome binding site and the AUG start codon.



**Figure 1.12: Inhibition of translation through SymR.**

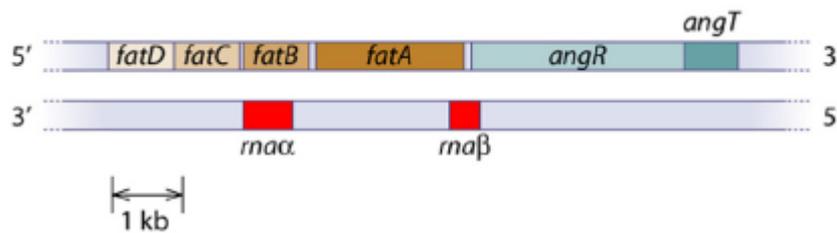
SymR is complementary over its full length to the *symE* 5' UTR, including the ribosome binding site (RBS), and probably causes a block in ribosome binding and, to a lesser extent, enhanced degradation of the untranslated mRNA. GadY and SymR are drawn according to their RNA fold maximum free energy (mfe) secondary structures (Georg and Hess, 2011).

Both SymR and the 30S ribosomal subunit competitively bind at the RBS on *symE* (Fig 1.12). The *symE* mRNA/SymR duplex formed is incompatible with the binding of the 30S RNA, subsequently preventing the initiation of SymE translation. In a *symR* mutant protein levels were shown to increase by more than 7-fold, however the mRNA level increased by only 3-fold in comparison. The cause of the enhanced degradation of the *symE* mRNA is unclear, either a direct result of the binding of the asRNA or a secondary effect due to the absence of the translating ribosomes on the mRNA.

The regulation of translation inhibition for trans-acting non-coding RNAs has recently been shown that involvement of the RBS may not be obligatory. The binding of a regulatory RNA after the start codon (Beiter et al., 2009) as well as upstream of the RBS (repression of *istR* (Darfeuille et al., 2007)), induction of *dsrA* (Majdalani et al., 1998)) have also been found to effect ribosome binding.

### 1.2.2.3.3 Transcription Termination

In addition to posttranscriptional mechanisms, other mechanisms exist which directly influence the transcription of target genes. The iron transport-biosynthesis operon in *Vibrio anguillarum* contains four ferric siderophore transport genes (*fatDCBA*) and two siderophore biosynthesis genes (*angR* and *angT*), as well two asRNAs (RNA $\alpha$  and RNA $\beta$ ) (Fig. 1.13) (Chen and Crosa, 1996, Salinas et al., 1993, Waldbeser et al., 1993, Waldbeser et al., 1995).



**Figure 1.13: Transcription termination by bacterial asRNAs in *Vibrio anguillarum*.** Organization of the *Vibrio anguillarum* iron transport-biosynthesis operon. The asRNA RNA induces transcription termination at a predicted stem-loop after the *fatABCD* part of the mRNA (Stork et al., 2007).

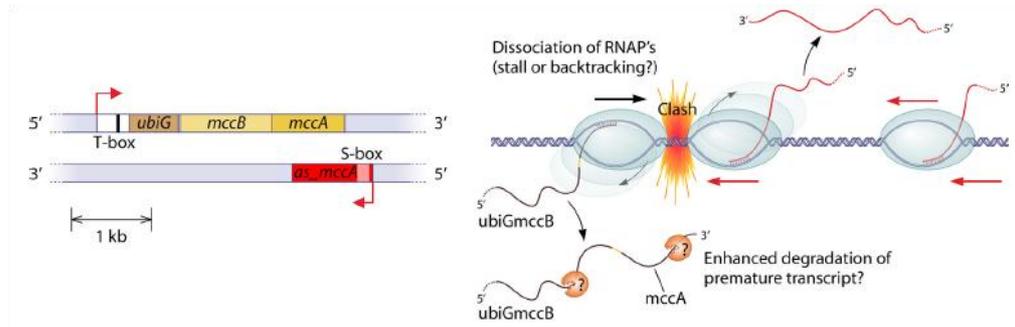
The asRNAs act co-operatively, with RNA $\alpha$  repressing *fatA* and *fatB* expression under iron-rich conditions and RNA $\beta$  causing the differential transcription of the full length *fatDCBA* operon and a shortened *fatDCBA* message (Stork et al., 2007). As the short form is 17 times more abundant than the full length version, when RNA $\beta$  binds to the growing polycistronic *fatDCBA* message, this leads to transcription termination at a potential hairpin which is located close to the *fatA* stop codon.

#### 1.2.2.3.3.4 Transcriptional Interference

Transcriptional interference mechanisms involve the effects of divergently or tandemly transcribed promoters on each other. The process of transcription is the point at which regulation takes place and therefore the resulting RNA could be a side effect. There are three mechanisms which contribute to the various interference effects observed, collision, promoter occlusion and sitting duck.

The collision of two divergent elongating RNA polymerase complexes results in the premature termination of one or both transcription events. This is more likely to be a long distance electrostatic interaction or as a result of the bow wave of positively super-coiled DNA in front of an elongating RNA polymerase rather than a direct steric interaction (Crampton et al., 2006). After this interaction, the outcome for the RNA polymerase includes the dissociation of one or both complexes, the backtracking of one complex or a stalling of the polymerases (Crampton et al., 2006, Sneppen et al., 2005). An example of this interference mechanism is illustrated in the transcription of the *ubiG-mccBA* operon in *Clostridium acetobutylicum* (Fig. 1.14).

This operon contains genes responsible for converting methionine to cysteine, the expression of which is upregulated in the presence of methionine and down regulated in the presence of cysteine. The asRNA mediating this regulation, *as\_mccA*, is up to 1,000 nt long with an additional three major fragments of 700, 400 and 200 nt lengths and is regulated in response to sulphur availability.

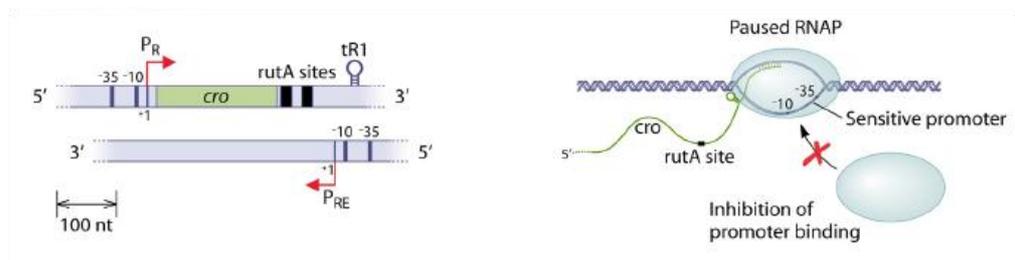


**Figure 1.14: Transcription interference by collision in the ubiG-mccBA operon in *Clostridium acetobutylicum*.**

Proposed collision mechanism for the *ubiG-mccABas\_mccA* system (*as\_mccA* stands for *mccA* antisense RNA). The two divergently elongating RNA polymerases, transcribing the asRNA and the *ubiG-mccAB* operon, collide and give rise to the 1,000-nt fragment for *as\_mccA*, which represents the sole known mechanism of termination. Short fragments for the mRNA were not detected, indicating rapid degradation of the prematurely terminated transcript (Georg and Hess, 2011).

Due to the lack of correlation between the longer transcript ends with obvious terminator structures and no change in the RNase fragmentation patterns, an alternative termination mechanism and not codegradation, was concluded to be taking place.

The next transcription interference mechanism is promoter occlusion, which occurs when an elongating RNA polymerase from an “aggressive” promoter passes over a “sensitive” promoter. This prevents the formation of an initiation complex at the “sensitive” promoter (Fig. 1.15).

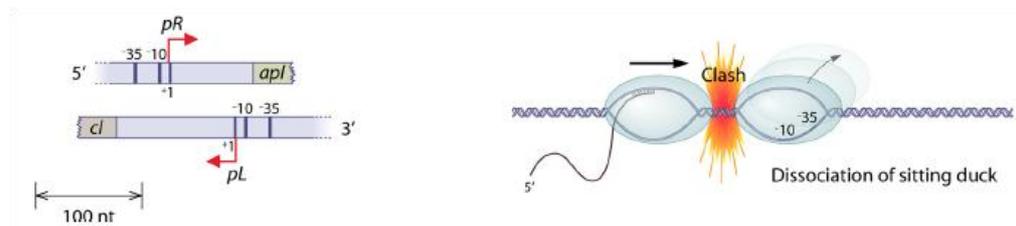


**Figure 1.15: Promoter occlusion mechanism in  $\lambda$  phage  $P_R$  and  $P_{RE}$  promoters.**

Promoter binding is inhibited by the pausing of RNA polymerase opposite the “sensitive” promoter, enhancing interference at the  $\lambda$  phage promoters  $P_R$  and  $P_{RE}$  (Palmer et al., 2009).

The interference by occlusion of the divergent phage promoters  $P_R$  and  $P_{RE}$  in  $\lambda$  phage demonstrated that the pausing of RNA polymerase at a tR1 site opposite the “sensitive” promoter causes interference to be strongly enhanced (Palmer et al., 2009).

The third transcriptional interference mechanism is ‘sitting duck’ interference, where a bound RNA polymerase at an open complex of the “sensitive” promoter is removed by the collision of another elongating RNA polymerase complex, occurring prior to the first polymerase proceeds to elongation (Fig. 1.16).



**Figure 1.16: Sitting duck transcriptional interference in bacteriophage 186.**

Sitting duck transcriptional interference is the major mechanism in bacteriophage 186 between the lytic-phase promoter (pR) and the lysogenic-phase promoter (pL), where “sensitive” bound RNA polymerase is removed by collision with another polymerase complex.

An example of this type of interference is recognised as the major mechanism between the lytic-phase promoter (pR) and the lysogenic-phase promoter (pL) in bacteriophage 186 (Callen et al., 2004, Sneppen et al., 2005). Computational modelling concluded this to be strongest interference mechanism when promoters are located close together and of moderate strength (Sneppen et al., 2005).

### 1.3 RESEARCH OUTLINE AND AIMS OF THE PRESENTED WORK

This study aimed to obtain a preliminary understanding of the structure, function and regulation of RsmN, a new atypical RsmA homologue in *Pseudomonas aeruginosa*. The role of RsmA as a global post-transcriptional regulator has been extensively studied with respect to its structure, regulation, and its binding mechanisms towards regulatory as well as target RNAs and the interplay between its structure and function. To elucidate the structure and function of RsmN and gain further insights into that of RsmA, various complementary strategies were devised and implemented experimentally:

- Biophysical techniques were used to characterise the solution structure of RsmA and RsmN, mechanism of self-assembly and the nature of the RNA binding interaction (in collaboration with Prof. Mark Searle and Elizabeth Morris).
- A DNA fragment containing the *rsmN* gene from *P. aeruginosa* was cloned and inserted into an *E. coli* based overexpression plasmid in order to perform protein expression and purification experiments.
- A series of plasmid and chromosomal *rsmN* and *rsmN* promoter DNA constructs were made to facilitate the construction of *rsmN* mutants, strains for *rsmN* inducible overexpression and strains for investigating *rsmN* transcription.
- Impact of *rsmN* mutation or overexpression on PAO1 virulence factors.
- RNA targets for RsmN and RsmA in *P. aeruginosa* using were identified using RNA-protein binding experiments.

## 2 MATERIALS AND METHODS

### 2.1 BACTERIAL STRAINS

All bacterial strains used in this study are listed in Table 2.1.

**Table 2.1: Bacterial strains used in this study.**

All the *P. aeruginosa* strains in this list are derived from PAO1-N unless stated otherwise.

Strain	Genotype/Characteristics	Reference/Source
<i>E. coli</i> :		
DH5 $\alpha$	F <sup>-</sup> <i>endA1 hsdR17</i> (r <sub>K</sub> <sup>-</sup> m <sub>K</sub> <sup>+</sup> ) <i>supE44 thi-1</i> $\Delta$ <i>recA1 gyrA96 relA1 deoR</i> $\Delta$ ( <i>lacZYA-argF</i> )-U169 $\phi$ 80 <i>dlacZ</i> $\Delta$ M15	(Grant et al., 1990)
S17-1 $\lambda$ pir	<i>recA, thi, pro, hsdR17</i> (r <sub>K</sub> <sup>-</sup> , m <sub>K</sub> <sup>+</sup> ), RP4-2-Tc::Mu-Km::Tn7, $\lambda$ pir	(Simon et al., 1983)
C41 (DE3)	F- <i>ompT gal hsdSB</i> (r <sub>B</sub> -m <sub>B</sub> <sup>-</sup> ) <i>dcm lon</i> $\lambda$ DE3 and an uncharacterised mutation described in Miroux and Walker, 1996	(Miroux and Walker, 1996)
TR1-5	<i>csrA</i> ::Kan <sup>r</sup> , <i>rpoS</i> (Am)	(Romeo et al., 1993) <i>rpoS</i> mutation described in (Wei et al., 2000)
<i>P. aeruginosa</i> :		
PAO1-N	Wild type, Nottingham strain	Holloway collection
PAO1-L	Wild type, Lausanne strain	ATCC 15692
PAZH13	$\Delta$ <i>rsmA</i> mutant	(Pessi et al., 2001)
PASK10	<i>lacI</i> <sup>Q</sup> , P <sub><i>tac</i></sub> - <i>rsmA</i> ; inducible <i>rsmA</i> , (Sm <sup>R</sup> /Sp <sup>R</sup> )	Sarah Kuehne thesis
PACP10	$\Delta$ <i>rhlR</i> mutant, in frame deletion	(Rampioni et al., 2010)
PASDP233	$\Delta$ <i>lasR</i> mutant::Gm insertional mutant-N	(Pessi and Haas, 2000)
PASDP123	$\Delta$ <i>pqsA</i> mutant, in frame deletion	(Aendekerck et al., 2005)
PAKR52	$\Delta$ <i>retS</i> mutant, in frame deletion	K. Righetti, Thesis
PAKR45	$\Delta$ <i>ladS</i> mutant, in frame deletion	K. Righetti, Thesis
PALT40	$\Delta$ <i>gacA</i> :: $\Omega$ Sm/Sp mutant	This work
PALT1	PAO1::(miniCTX::P <sub><i>rsmN</i></sub> - <i>lux</i> ) transcriptional fusion	This work
PALT2	PAO1::(miniCTX::P <sub><i>nmsR</i></sub> - <i>lux</i> ) transcriptional fusion	This work
PALT3	PASK10::(miniCTX::P <sub><i>rsmN</i></sub> - <i>lux</i> )	This work
PALT4	PAO1::(miniCTX::P <sub><i>nmsR</i></sub> - <i>lux</i> )	This work
PALT5	PALT16::(miniCTX::P <sub><i>rsmN</i></sub> - <i>lux</i> )	This work
PALT6	PALT16::(miniCTX::P <sub><i>nmsR</i></sub> - <i>lux</i> )	This work

<b>Strain</b>	<b>Genotype/Characteristics</b>	<b>Reference/Source</b>
PALT7	PAZH13::(miniCTX::P <sub>rsmN</sub> - <i>lux</i> )	This work
PALT8	PAZH13::(miniCTX::P <sub>nmsR</sub> - <i>lux</i> )	This work
PALT11	<i>laqI</i> <sup>0</sup> , P <sub>tac</sub> - <i>rsmN</i> , inducible <i>rsmN</i> , (Sm <sup>R</sup> /Sp <sup>R</sup> )	This work
PALT13	<i>laqI</i> <sup>0</sup> , P <sub>tac</sub> - <i>rsmA</i> , inducible <i>rsmA</i> , (Sm <sup>R</sup> /Sp <sup>R</sup> )	This work
PALT16	$\Delta$ <i>rsmN</i> mutant	This work
PALT22	PAO1::(miniCTX::P <sub>pqsA</sub> - <i>lux</i> )	This work
PALT23	PAO1::(miniCTX::P <sub>rhlI</sub> - <i>lux</i> )	This work
PALT24	PAO1::(miniCTX::P <sub>lasI</sub> - <i>lux</i> )	This work
PALT25	PALT11::(miniCTX::P <sub>pqsA</sub> - <i>lux</i> )	This work
PALT26	PALT11::(miniCTX::P <sub>rhlI</sub> - <i>lux</i> )	This work
PALT27	PALT11::(miniCTX::P <sub>lasI</sub> - <i>lux</i> )	This work
PALT28	PALT16::(miniCTX::P <sub>pqsA</sub> - <i>lux</i> )	This work
PALT29	PALT16::(miniCTX::P <sub>rhlI</sub> - <i>lux</i> )	This work
PALT30	PALT16::(miniCTX::P <sub>lasI</sub> - <i>lux</i> )	This work
PALT31	PAZH13::(miniCTX::P <sub>pqsA</sub> - <i>lux</i> )	This work
PALT32	PAZH13::(miniCTX::P <sub>rhlI</sub> - <i>lux</i> )	This work
PALT33	PAZH13::(miniCTX::P <sub>lasI</sub> - <i>lux</i> )	This work
PALT34	PALT11::(miniCTX::P <sub>rsmN</sub> - <i>lux</i> )	This work
PALT35	PALT11::(miniCTX::P <sub>nmsR</sub> - <i>lux</i> )	This work
PALT44	PASK10::(miniCTX::P <sub>pqsA</sub> - <i>lux</i> )	This work
PALT45	PASK10::(miniCTX::P <sub>rhlI</sub> - <i>lux</i> )	This work
PALT46	PASK10::(miniCTX::P <sub>lasI</sub> - <i>lux</i> )	This work
PALT49	PACP10::(miniCTX::P <sub>rsmN</sub> - <i>lux</i> )	This work
PALT50	PACP10::(miniCTX::P <sub>nmsR</sub> - <i>lux</i> )	This work
PALT51	PASDP123::(miniCTX::P <sub>rsmN</sub> - <i>lux</i> )	This work
PALT52	PASDP123::(miniCTX::P <sub>nmsR</sub> - <i>lux</i> )	This work
PALT53	PASDP233::(miniCTX::P <sub>rsmN</sub> - <i>lux</i> )	This work
PALT54	PASDP233::(miniCTX::P <sub>nmsR</sub> - <i>lux</i> )	This work
PALT55	PACP10::(miniCTX:: <i>lux</i> ), negative control	This work
PALT56	PASDP123::(miniCTX:: <i>lux</i> ), negative control	This work
PALT57	PASDP233::(miniCTX:: <i>lux</i> ), negative control	This work
PALT63	PAO1 pRsmA (L), C-terminal hexahistidine tag	This work
PALT64	PAO1 pRsmN (L), pRsmN = pLT28, N-terminal hexahistidine tag	This work

## 2.2 PLASMIDS

All plasmids used in this study are listed in Table 2.2

**Table 2.2: Plasmids used in this study**

Plasmid	Characteristics	Reference/Source
pBLS	pBluescript KS cloning vector; ColE1 replicon (Ap <sup>R</sup> )	Stratagene
pUC6S	Small cloning vector (Ap <sup>R</sup> )	(Vieira and Messing, 1991)
pDM4	Suicide vector with <i>sacBR</i> genes for sucrose counter-selection (Cm <sup>R</sup> )	(Milton et al., 1996)
pZH13	pDM4 carrying $\Delta rsmA$ (Cm <sup>R</sup> )	Zoë Hindle, used in (Pessi et al., 2001)
pHP45 $\Omega$	Transcription and translation termination signal(Sm <sup>R</sup> , Sp <sup>R</sup> , Ap <sup>R</sup> )	(Prentki and Krisch, 1984)
miniCTX:: <i>lux</i>		(Becher and Schweizer, 2000)
pGEM <sup>®</sup> -T Easy	Cloning vector (Ap <sup>R</sup> ), <i>lacZ</i> gene with internal MCS.	Promega
pME6001	Cloning vector derived from pBBR1MCS (Gm <sup>R</sup> )	(Blumer et al., 1999)
pME6032	<i>lacI</i> <sup>Q</sup> -P <sub>tac</sub> expression vector; pVS1-p15A shuttle vector (Tet <sup>R</sup> )	(Heeb et al., 2002)
	miniCTX:: <i>P<sub>lasI</sub>-lux</i>	G. Rampioni
	miniCTX:: <i>P<sub>rhlI</sub>-lux</i>	G. Rampioni
	miniCTX:: <i>P<sub>pqsA</sub>-lux</i>	(Diggle et al., 2007)
pRsmA	pME6032:: <i>rsmA</i> (Tet <sup>R</sup> )	(Heeb et al., 2006)
pRsmN	pME6032:: <i>rsmN</i> (Tet <sup>R</sup> )	This Work
pSK11	Suicide plasmid based on pDM4 to replace <i>rsmA</i> by an inducible <i>lacI</i> <sup>Q</sup> P <sub>tac</sub> - <i>rsmA</i> allele	S. Kuehne, Thesis
pMM31	pBLS upstream RsmN 544 bp fragment <i>XbaI-EcoRI</i> for construction of pMM33	M. Messina, Thesis
pMM32	pBLS downstream RsmN 544 bp fragment <i>EcoRI-XhoI</i> for construction of pMM33	M. Messina, Thesis
pMM33	Suicide plasmid pDM4-based carrying $\Delta rsmN$ (Cm <sup>R</sup> )	M. Messina, Thesis
pME6111	Suicide plasmid $\Omega$ Sm/Sp inserted into <i>gacA</i> , ColE1 pME3088-based	(Reimann et al., 1997)

Plasmid	Characteristics	Reference/Source
pHLT	Modification of the expression vector pRSETA (Invitrogen) including a hexahistidine tag, followed by a lipoyl domain and a thrombin cleavage site (Ap <sup>R</sup> )	(Heeb et al., 2006)
pHLT:: <i>rsmA</i>	pHLT with <i>rsmA</i> , cloned in with <i>EcoRI</i> and <i>BamHI</i> (Ap <sup>R</sup> )	(Heeb et al., 2006b) (Heeb et al., 2006)
pHT	Modification of the expression vector pRSETA (Invitrogen) including a hexahistidine tag and a thrombin cleavage site (Ap <sup>R</sup> )	This work
pHT:: <i>rsmAV40W</i>	pHT with <i>rsmA</i> , tryptophan mutant V40W	This work
pHT:: <i>rsmAY48W</i>	pHT with <i>rsmA</i> , tryptophan mutant Y48W	This work
pHT:: <i>rsmAL23W</i>	pHT with <i>rsmA</i> , tryptophan mutant L23W	This work
pHT:: <i>rsmAN35W</i>	pHT with <i>rsmA</i> , tryptophan mutant N35W	This work
pLT1	miniCTX:: $P_{rsmN}$ - <i>lux</i>	This work
pLT2	miniCTX:: $P_{nmsR}$ - <i>lux</i>	This work
pLT3	pHT with <i>rsmA</i> , cloned in with <i>EcoRI</i> and <i>BamHI</i> (Ap <sup>R</sup> )	This work
pLT4	pHT with <i>rsmN</i> , cloned in with <i>EcoRI</i> and <i>BamHI/BglIII</i> (Ap <sup>R</sup> )	This work
pLT5	pBLS:: <i>rsmNa</i> (amplified from RSMNPA3 and RSMNPA4); intermediate step for the construction of pLT10	This work
pLT6	pBLS:: <i>rsmNd</i> (amplified from RSMNPA1 and RSMNPA2); intermediate step for the construction of pLT10	This work
pLT7	pBLS:: <i>rsmNab</i> intermediate step for the construction of pLT10 from pLT5 with cloned $lacI^O P_{tac}$ from pME6032 ( <i>EcoRI</i> , <i>BamHI</i> )	This work
pLT8	pBLS:: <i>rsmNabc</i> intermediate step for the construction of pLT10 from pLT7 with inserted $\Omega$ -Sp cassette ( <i>BamHI</i> )	This work
pLT9	pBLS:: <i>rsmNabcd</i> intermediate step for the construction of pLT10 from pLT8 with <i>rsmN</i> containing fragment cloned in from pLT6 ( <i>EcoRI</i> , <i>XhoI</i> )	This work
pLT10	Suicide plasmid based on pDM4 to replace <i>rsmN</i> by an inducible	This work

Plasmid	Characteristics	Reference/Source
	<i>lacI</i> <sup>Q</sup> P <sub>tac</sub> - <i>rsmN</i> construction ( <i>XhoI</i> , <i>XbaI</i> fragment from pLT9)	
pLT15	pHT with <i>rsmAR44A</i> arginine mutation, cloned in with <i>EcoRI</i> and <i>BamHI</i> (Ap <sup>R</sup> )	This work
pLT16	pHT with <i>rsmNR62A</i> arginine mutation, cloned in with <i>EcoRI</i> and <i>BamHI/BglIII</i> (Ap <sup>R</sup> )	This work
pLT25	<i>rsmN</i> in pGEM-T using <i>EcoRI</i> - <i>ClaI</i>	This work
pLT26	<i>H<sub>6</sub>rsmN</i> in pGEM-T using <i>EcoRI-XhoI</i>	This work
pLT30	<i>rsmNR62A</i> in pGEM-T using <i>EcoRI-ClaI</i>	This work
pLT27	<i>rsmN</i> in pME6032 using <i>EcoRI</i> - <i>ClaI</i>	This work
pLT28	<i>H<sub>6</sub>rsmN</i> in pME6032 using <i>EcoRI-XhoI</i>	This work
pLT31	<i>rsmNR62A</i> in pME6032 using <i>EcoRI-ClaI</i>	This work

## 2.3 OLIGONUCLEOTIDES

Oligonucleotides were synthesised by Sigma Genosys Biotechnologies, Cambridge, UK.

**Table 2.3: Oligonucleotides used in this study**

Oligonucleotide	Sequence (5' to 3')	Function
rsmA1 (S)	CTGGCCAAGGAAAGCATCAAC	Screening of PAO1 $\Delta rsmA$
rsmA2 (S)	CTCCGCAACCCGGGGCGCATG	Screening of PAO1 $\Delta rsmA$
Ptac (S)	CGGCTCGTATAATGTGTGGA	Sequence multiple cloning site in pME6032
P6032 (S)	CCCTCACTGATCCGCTAGTC	Sequence multiple cloning site in pME6032
T3(S)	ATTAACCCTCACTAAAGGG	Sequence multiple cloning site in pBluescript
T7t(S)	TATGCTAGTTATTGCTCAGCGG	Sequence multiple cloning site in pBluescript
Ctx (S)	CATGCTCTTCTCTAATGCGTGA	Sequence miniCTX:: <i>lux</i> plasmid
RSMNPR1	TATCTGCAGGTGTGGAGGGATGGTCACAG	Reverse primer to make miniCTX:: <i>lux</i> promoter fusion with <i>rsmN</i> sense promoter
RSMNPF1	TATCTCGAGCTTGCTCTGGGCTACCTGAT	Forward primer to make miniCTX:: <i>lux</i> promoter fusion with <i>rsmN</i> sense promoter
RSMNPR2	TATGAATTCGTTTCGCGGGGCTTTTACACATCAG	Reverse primer to make miniCTX:: <i>lux</i> promoter fusion with <i>rsmN</i> antisense promoter
RSMNPF2	TATAAGCTTCTCTCCTGGTAATCGCGTTC	Forward primer to make miniCTX:: <i>lux</i> promoter fusion with <i>rsmN</i> antisense promoter
rsmNA	CGCGAAGGCGGCATCCGGATCCTGGTCACC	DIG-labelled oligonucleotide probe for antisense analysis of <i>rsmN</i> transcripts
rsmNS	GGTGACCAGGATCCGGATGCCGCTTCGCG	DIG-labelled oligonucleotide probe for sense analysis of <i>rsmN</i> transcripts
HT_RSMNPR1	TATGAATTCTCAGCCTTTTCGGTGCCGTTT	Reverse primer to amplify <i>rsmN</i> to produce His-tagged RsmN proteins, <i>EcoRI</i>
HT_RSMNPF1	TATAGATCTATGGGTTTCCTGATACTCTCC	Primer to amplify <i>rsmN</i> to produce His-tagged RsmN proteins, <i>BglII</i> .
RSMNPA1	TATGAATTCATGGGTTTCCTGATACTCTC	Primer to make suicide plasmid to integrate inducible and constitutively expressed <i>rsmN</i> in the chromosome
RSMNPA2	TATCTCGAGGGCGACTCCACCAAGACC	Primer to make suicide plasmid to integrate inducible and constitutively expressed <i>rsmN</i> in the chromosome
RSMNPA3	TATTCTAGACCAGGTTGAGCTGATTGAGG	Primer to make suicide plasmid to integrate inducible and constitutively expressed <i>rsmN</i> in the chromosome

Oligonucleotide	Sequence (5' to 3')	Function
RSMNPA4	TATGGATCCCCTTTGGTGAATGAAATGGTGT	Primer to make suicide plasmid to integrate inducible and constitutively expressed <i>rsmN</i> in the chromosome
HisThrFor	TATGCACCATCACCATCACCATCTGGTGCCGCGCG	Primer to make pHT vector by removal of lipoyl domain
HisThrRev	GATCCGCGCGGCACCAGATGGTGATGGTGATGGTGCA	Primer to make pHT vector by removal of lipoyl domain
L23W_F	GTCACCGTGACGGTACTGGGTGTCAAAGGG	Forward primer to introduce L23W mutation in RsmA
L23W_R	CCCTTTGACACCCCATACCGTCACGGTGAC	Reverse primer to introduce L23W mutation in RsmA
N35W_F	CGCATGGGCGTCAACGCGCCGAAGGAAGTC	Forward primer to introduce N35W mutation in RsmA
N35W_R	GACTTCCTTCGGCGCCAGACGCCGATGCG	Reverse primer to introduce N35W mutation in RsmA
R44A_F	GCCGTACACGCGGAGGAAATT	Forward primer to introduce R44A mutation in RsmA
R44A_R	AATTTCTCCGCGTGTACGGC	Reverse primer to introduce R44A mutation in RsmA
R62A_F	CTGATCGTTGCGGACGAGTTG	Forward primer to introduce R62A mutation in RsmN
R62A_R	CAACTCGTCCGCAACGATCAG	Reverse primer to introduce R62A mutation in RsmN
gacA1	TAAGGTTGCCGAAATCTCCTG	Primer to identify PAO1 $\Delta$ <i>gacA</i>
gacA2	CTTCTCGAAGATGCGGTAGC	Primer to identify PAO1 $\Delta$ <i>gacA</i>
pMNF2	TATGAATTCATGGGTTTCCTGATACTC	Primer to introduce <i>EcoRI</i> site at the start of <i>rsmN</i> in pME6032 based constructs.
pMNR	TATATCGATTGAGCCTTTTCGGTGCCGTTT	Primer to introduce <i>ClaI</i> site at the end of <i>rsmN</i> in pME6032 based constructs.
pME_NR	TATCTCGAGTCAGCCTTTTCGGTGCCGTTT	Primer to introduce <i>XhoI</i> site at the end of <i>rsmN</i> in pME6032 based constructs.
HT_pME_NF	TATGAATTCACCATCACCATCACCATAAGCTTATGGGTTTCC	Primer to introduce 6xHistidine tag at start of <i>rsmN</i> flanked by <i>EcoRI</i> and <i>HindIII</i>
Fw_RsmN_up	TATTCTAGATGTGCGAACGACCGTATTTTC	Forward primer to insert downstream RsmN fragment into pMM32 (Primer to identify PAO1 $\Delta$ <i>rsmN</i> )
Rv_RsmN_dw	TATCTCGAGTACTGGACCAGCTTGTTCCG	Reverse primer to insert upstream RsmN fragment into pMM31 (Primer to identify PAO1 $\Delta$ <i>rsmN</i> )
Fw_RsmN_dw	TATGAATTCACCCATGTTCCGCGTCCTT	Forward primer to insert upstream RsmN fragment into pMM31
Rv_RsmN_up	TATGAATTCGGCTGACGAACGGTAGAAA	Reverse primer to insert downstream RsmN fragment into pMM32

## 2.4 PLASMID AND STRAIN CONSTRUCTION

### 2.4.1 Construction of plasmids

For all plasmids and strains constructed in this thesis, cloned PCR products were sequenced to verify the absence of unwanted nucleotide substitutions.

#### 2.4.1.1 Plasmids made by PCR-based point mutagenesis

Primers were designed to introduce a tryptophan mutation into the wild type RsmA gene (L23W, N35W, V40W and Y48W). Using the Stratagene Quick Change Site-Directed Mutagenesis kit<sup>®</sup>, the PCR reaction were carried out as follows. Components required are: 10× reaction buffer (100 mM KCl, 100 mM (NH<sub>4</sub>)SO<sub>4</sub>, 200 mM Tris-HCL pH 8.8, 20 mM MgSO<sub>4</sub>, 1 % Triton X-100, 1 mg/ml nuclease free bovine serum albumin), a DNA plasmid template (50 ng/μl), forward and reverse primers (125 ng/μl), dNTP mix (0.1 mM) and ddH<sub>2</sub>O (40 μl). Last of all *Pfu Turbo*<sup>®</sup> DNA polymerase (0.05 U/μl) was added to the reaction mixture. The reactions were carried out in a Techne Thermal Cycler (Progene).

The reaction mixes were then stored on ice before digestion. Prior to further use the PCR product was subjected to *DpnI* endonuclease (0.2 U/μl), which digests parental DNA due to the specificity for methylated and hemi methylated DNA. The mixture was centrifuged for 1 min and incubated at 37 °C for 1 h.

After the PCRs, the product was digested and cloned into the pHT vector using the *EcoRI* and *ClaI* sites. The plasmids pHT::*rsmAL23W/N35W/V40W* and Y48W (Table 2.2) were constructed using this strategy.

#### 2.4.1.2 Construction of arginine-alanine substitution mutants

Primers were designed to introduce an arginine-alanine substitution into the wild type *rsmN* (R62A) and *rsmA* (R44A) genes using the Stratagene Quick Change Site-Directed Mutagenesis kit<sup>®</sup> as above. The extension time for the PCR for using template DNA for *rsmA* was 8 min and 5 min for *rsmN*.

PCR mutagenesis for mutation of R62A in pME6032::*rsmN* was repeatedly unsuccessful, possibly due to large size of pME6032 plasmid (8 - 9 kb). The experiment was repeated successfully using pGEM-T::*rsmN* DNA (3015 bp empty vector) as the template for the PCR reaction prior to insertion in pME6032.

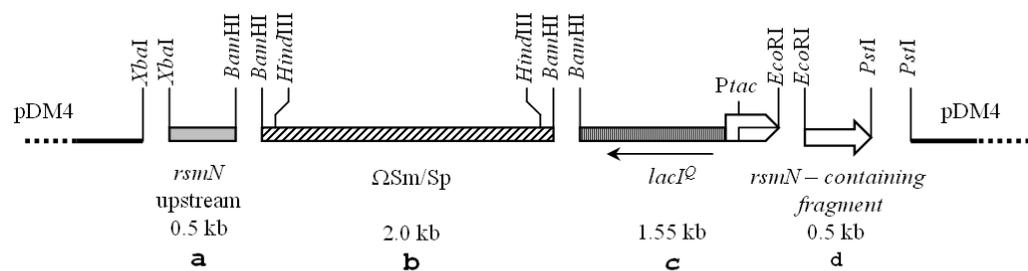
#### 2.4.1.3 Construction of the *E. coli* overexpression plasmid pHT::*rsmN*

A histidine-tagged *rsmN* gene was constructed by the amplification of a fragment from PAO1 genomic DNA using primers HT\_RSMNPF1 and HT\_RSMNPR1. The plasmid pHT::*rsmA* was opened (*Bam*HI, *Eco*RI) and the 264-bp product was inserted. The *rsmN* gene contains a *Bam*HI site within its DNA sequence, therefore *Eco*RI and the *Bam*HI compatible enzyme *Bgl*III were used to digest the *rsmN* PCR product to form pHT::*rsmN*.

#### 2.4.1.4 Construction of suicide plasmid pDM4::*lacI*<sup>Q</sup> P<sub>tac</sub>-*rsmN* (pLT10)

A 632-bp fragment containing *rsmN* was amplified from the PAO1 genomic DNA using primers RSMNPA3 and RSMNPA4 and cloned into pBLS to give pLT6. Another 572-bp fragment containing the downstream region of *rsmN* was amplified similarly using primers RSMNPA1 and RSMNPA2 to give

pLT5. The plasmid pLT5 was linearised (*EcoRI*, *BamHI*) and *lacI*<sup>Q</sup> P<sub>tac</sub> (from pME6032, *EcoRI*, *BamHI*) was introduced to give pLT7. The Ω-cassette (2.0 kb) was excised from pHP45Ω (*BamHI*) and cloned into pLT7 (cut with *BamHI* and dephosphorylated) to give pLT8. The plasmid pLT8 was then digested (*EcoRI*, *XhoI*) and the 632-bp fragment containing *rsmN* from pLT6 were cloned to give pLT9. The final construct was subcloned into pDM4 (*XhoI*, *XbaI*) to give the suicide plasmid pLT10 (Figure 2.1). pDM4 is a suicide vector derived from pNQ705, containing a chloramphenicol resistance marker, the conditionally lethal *sacBR* gene from *Bacillus subtilis* and a modified multicloning site (Milton et al., 1996).



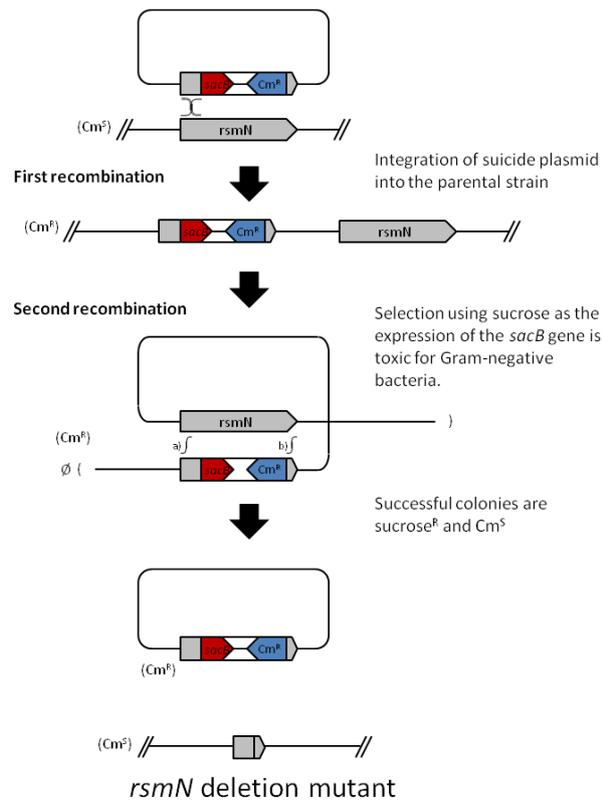
**Figure 2.1: Schematic representation of pLT10, the suicide plasmid for the construction of inducible *rsmN* strains.**

The suicide plasmid consists of four fragments where fragment **a**) contains the upstream fragment of *rsmN*; **b**) contains the omega cassette (ΩSm/Sp) from pHP45Ω, **c**) consists of the *lacI*<sup>Q</sup>P<sub>tac</sub> from pME6032 and **d**) is the *rsmN* containing fragment.

#### 2.4.1.5 *rsmN* deletion mutant.

An *rsmN* in-frame deletion mutant was made using a two-step procedure where the suicide plasmid pMM33 (Table 2.2) underwent conjugation with recipient PAO1. The pDM4-based suicide plasmid pMM33 was constructed using the pBluescript cloning vectors pMM31 (upstream RsmN fragment 544 bp, *XbaI*-*EcoRI*) and pMM32 (downstream RsmN fragment 544 bp, *EcoRI*-*XhoI*), resulting in a 206 bp deletion of the 216 bp RsmN. pMM33 was grown in *E.*

*coli* S17-1  $\lambda$ pir which supplies R6K replication functions and the *tra* genes for efficient conjugation. Firstly, the entire plasmid is integrated into the chromosome by a single cross-over between one of the two homologous regions, producing duplication within the chromosome (Figure 2.2). The chloramphenicol resistance marker of the pDM4-based suicide plasmid facilitates the selection. The mating was performed as described, with selection for nalidixic acid (15  $\mu$ g/ml) and chloramphenicol (300  $\mu$ g/ml).



**Figure 2.2: Representation of the steps required to make the *rsmN* mutant strain.**

The suicide plasmid is integrated into the chromosome by a single cross-over between one of the two homologous regions, producing a duplication in the chromosome. The suicide vector and one of the alleles are removed after the second homologous recombination. The example above is one of the two possibilities leading to the same final product.

Secondly, the suicide vector and one of the alleles are removed during a second homologous recombination event. After single colonies were grown on nalidixic acid/chloramphenicol plates followed by culturing in LB medium overnight, batches were sub-cultured into LB containing 10 % sucrose in the

absence of chloramphenicol. Sucrose induces the *sacBR* gene which encodes levansucrase that converts sucrose to levan. This compound is toxic and prevents the clones that still carry the suicide plasmid from multiplying, enriching the population in exconjugants that have lost the plasmid. The successful clones together with revertants to wild type should be chloramphenicol-sensitive, and these are then identified by PCR.

#### 2.4.1.6 *rsmN* conditional mutant in wt (PALT11) and $\Delta rsmA$ (PALT13)

In order to acquire the conditional mutant, the pDM4-based suicide plasmid pLT10 underwent transconjugation with recipients PAO1 and PAZH13 (*rsmA* mutant) to give the conditional *rsmN* strains PALT11 (PAO1::*lacI*<sup>Q</sup> P<sub>tac-rsmN</sub>) and PALT13 (PAZH13::*lacI*<sup>Q</sup> P<sub>tac-rsmN</sub>). The mating and selections were done as described in 2.4.1.5.

#### 2.4.1.7 Construction of a *gacA* mutant ( $\Omega$ Sm/Sp)

The *gacA* mutant was constructed by conjugation of the ColE1 pME3088-based suicide plasmid pME6111 (omega cassette disruption ( $\Omega$ Sm/Sp)) into the PAO1 wild type (Reimann et al., 1997).

#### 2.4.1.8 Construction of a sense *rsmN-lux* transcriptional reporter fusion (pLT1).

To construct the *rsmN* sense promoter fusion carried by pLT1, PAO1 genomic DNA was amplified using primers RSMNPF1 and RSMNPR1 to produce a

331-bp product with part of the sense promoter flanked by *XhoI* and *PstI* restriction sites. The miniCTX::*lux* plasmid was opened (*XhoI,PstI*) and the 331-bp product was inserted. Following ligation, the DNA was transformed into *E. coli* S17-1  $\lambda$ *pir* cells. Successful fusions were identified on the transformation plates using a Berthold Luminograph LB980. Bacterial colonies that successfully incorporated the fusions emitted light.

#### 2.4.1.9 Construction of an antisense *nmsR-lux* transcriptional fusion (pLT2)

To construct the antisense promoter fusion pLT2, PAO1 genomic DNA was amplified using primers RSMNPF2 and RSMNPR2 to produce a 452-bp product with part of the sense promoter and flanking *HindIII* and *EcoRI* restriction sites. The miniCTX::*lux* plasmid was linearised (*HindIII,EcoRI*) and the 452-bp product was subcloned into it.

Following ligation, the DNA was transformed into *E. coli* S17-1  $\lambda$ *pir* cells. Successful fusions were identified as previously explained in section 2.4.1.8.

## 2.5 GENERAL CHEMICALS

Unless otherwise stated, all chemicals were obtained from Sigma (Poole, UK).

### 2.5.1 Antibiotics

Stock solutions of antibiotics were prepared according to standard protocols (Sambrook et al., 1989) and stored at -20 °C. Ampicillin was used from a 50 mg/ml in 50 % v/v EtOH stock, tetracycline from 100 mg/ml in MeOH, kanamycin from 30 mg/ml in dH<sub>2</sub>O, chloramphenicol from 50 mg/ml in EtOH,

carbenicillin from 50 mg/ml in dH<sub>2</sub>O and streptomycin from 50 mg/ml in dH<sub>2</sub>O.

### **2.5.2 Synthetic quorum sensing signal molecules**

Synthetic 3O-C12-HSL, C4-HSL and PQS were made by A. Truman at the School of Molecular Medical Sciences, University of Nottingham and kept as 10 mM stocks in methanol (PQS) or acetonitrile (3O-C12-HSL, C4-HSL) as described by (Chhabra et al., 2003) and (Pesci et al., 1999). Compounds were stored at -20 °C.

## **2.6 GROWTH MEDIA**

Media were prepared using deionised water and autoclaved at 121 °C for 20 min at 15 pound-force per square inch (p.s.i.).

### **2.6.1 Luria Bertani media (LB)**

LB broth was prepared as previously described (Sambrook et al., 1989) and consisted of 10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl and NaOH to pH 7.2.

LB agar was prepared by addition of 0.8 % (w/v) Technical Agar No. 3 (Oxoid) to LB broth.

### **2.6.2 Peptone Tryptone Soy Broth (PTSB)**

An alternative to LB for the overnight cultures used subsequently in phenotypic assays was prepared as described (Ohman et al., 1980). PTSB consists of 5 % w/v peptone (Difco) and 0.25 % w/v tryptone soy broth (Merck).

### **2.6.3 King's B Medium**

King's B medium is used as the base medium for a skimmed milk protease assay. The medium was prepared as previously described (King et al., 1954) using 20 g/l proteose peptone No. 3 (Difco), 10 g/l glycerol, 1.5 g/l  $K_2HPO_4 \cdot 3H_2O$  and 17 g/l bacto agar (Difco) with a final pH 7.2 - 7.4. Prior to use,  $MgSO_4$  was added from an autoclaved 1 M stock solution for a final concentration of 6 - 7 nM. At the same time, a solution of 50 % wt/vol skimmed milk was added to give a final concentration of 5 %.

### **2.6.4 Swarming motility agar**

Swarming motility agar was prepared according to a previously published method (Rashid and Kornberg, 2000). This consisted of 0.5 % (w/v) Bacto agar (Difco) and 0.8 % (w/v) Nutrient broth No. 2 (Oxoid) in distilled water. After autoclaving, filter sterilised D-glucose (Sigma) in distilled water was added to a final concentration of 0.5 % (w/v).

### **2.6.5 Kornberg medium**

Kornberg medium was prepared as previously described (Romeo et al., 1993) and consisted of 1.1 % (w/v)  $K_2HPO_4$ , 0.85 % (w/v)  $KH_2PO_4$ , 0.6 % (w/v) yeast extract, 0.5 % (w/v) glucose, and 1.5 % (w/v) agar.

### **2.6.6 Pyocyanin medium**

Pyocyanin medium consisted of 4 g D/L-alanine, 9.2 ml glycerol 87 % (v/v), 0.056 g  $K_2HPO_4$ , 5.68 g  $Na_2SO_4$ , 0.04 g citric acid, pH 7.0 in a total of 388 ml  $H_2O$  + 8 ml  $MgCl_2 \cdot 6H_2O$  (2.3 g / 10 ml) + 4 ml  $FeCl_3$  (0.06 g/10 ml) (Frank and Demoss, 1959).

## **2.7 GROWTH & STORAGE OF BACTERIA**

### **2.7.1 Bacterial growth conditions**

Routine liquid cultures were grown in LB or PTSB in a shaking incubator (Gallenkamp Ltd., UK or New Brunswick Scientific, USA) with agitation at 200 rpm at 37 °C, unless otherwise stated. Growth of bacterial cultures was monitored by absorbance at a wavelength of 600 nm using a Novospec II visible spectrophotometer (Pharmacia LKB Ltd., Cambridge, UK).

### **2.7.2 Long term storage of bacterial strains**

To allow long-term storage of bacterial strains, 0.75 ml of a bacterial culture grown overnight (o/n) was mixed thoroughly with 0.75 ml 50 % (v/v) glycerol prepared by filtration through 0.2  $\mu m$  filter membrane. The cell suspension was

then transferred into 2 ml Micro tubes (Sarstedt, Germany) and stored at -80 °C.

## **2.8 PROTOCOLS**

### **2.8.1 Transformation of bacterial strains**

#### 2.8.1.1 Preparation of electrocompetent *E. coli* cells

To prepare competent *E. coli* cells, a 1 % (v/v) inoculum from an overnight *E. coli* culture was added to 200 ml of sterile LB in a 1 l conical flask and grown at 37 °C with shaking at 200 rpm to an OD at 600 nm of 0.4 - 0.6 (reached approximately 6 h after inoculation). Cells were harvested by centrifugation at 6,000 rpm (JA-14, Beckman) for 10 min at 4 °C and washed four times in sterile ice-cold 1 mM MOPS with 10 % (v/v) glycerol before being resuspended in 1 ml of the same buffer. Cells were aliquoted into 50 µl samples in microcentrifuge tubes, flash frozen in liquid nitrogen and stored at -80 °C until required.

#### 2.8.1.2 Electroporation of electrocompetent *E. coli* cells

For electroporation of DNA into *E. coli* cells, salts were removed from the DNA solution by filter dialysis through 0.025 µm millipore filters (Millipore Corporation, USA) for 20 min. Electroporation was performed in 0.2 cm electrode gap Gene Pulser cuvettes (BioRad, UK) containing 50 µl of competent cells and 2 µl dialysed DNA. An electroporation pulse of 2.5 kV (25 µF, 200 Ω) was delivered using the BioRad Gene Pulser connected to a

BioRad pulse controller (BioRad, UK). A 0.75 ml volume of NYB broth was added to the cells which were then incubated for 1 h at 37 °C in the absence of antibiotics before plating aliquots onto LB agar plates containing appropriate antibiotics to select for transformants which grew o/n at 37 °C. Negative controls of electroporated cells with no plasmid were also similarly prepared.

#### 2.8.1.3 Preparation of electrocompetent *P. aeruginosa* cells

*P. aeruginosa* cells competent for electroporation were prepared from 1.5 ml of culture, grown o/n in LB at 42 °C, followed by centrifugation at 13,000 g for 3 min. The cells were washed four times with progressively smaller volumes of ice-cold 1 mM MOPS containing 10 % (v/v) glycerol. The pellet obtained was resuspended in 50 µl ice cold 1 mM MOPS with 10 % (v/v) glycerol, and the resulting electrocompetent cells immediately used for electroporation.

#### 2.8.1.4 Electroporation of electrocompetent *P. aeruginosa* cells

Transformation of *P. aeruginosa* cells was performed as for *E. coli* using electrocompetent cells prepared as described in 2.8.1.2.

#### 2.8.1.5 *P. aeruginosa* transformation using CaCl<sub>2</sub>

Calcium-competent *P. aeruginosa* cells for transformation were prepared by diluting an culture, grown o/n in LB at 42 °C, 1:100 and growing it at 37 °C until OD<sub>600</sub> 0.8. Forty ml of the culture was centrifuged at 8,000 g for 10 min at 4 °C. The pellet was resuspended in ice-cold 100 mM CaCl<sub>2</sub>, 20 % (v/v) glycerol and left on ice for 30 min before centrifuging. The pellet was

resuspended in 1.6 ml of the same ice-cold solution. To transform, 200  $\mu$ l of the cells were mixed with 200 ng of plasmid DNA, incubated on ice for 30 min and then heat-shocked at 42 °C for 2 min, prior to addition of 0.75 ml LB broth and further treatment as described after electroporation in 2.8.1.2.

## **2.8.2 Quantifying DNA, RNA and protein concentrations**

The NanoDrop® ND-1000 (Nanodrop Technologies) was used to measure DNA, RNA and protein concentrations. 1 to 2  $\mu$ l of sample was used to determine characteristic absorbance and concentrations. Whole spectra of the samples could also be measured to assess purity.

## **2.8.3 DNA manipulation**

### **2.8.3.1 Isolation of chromosomal DNA**

Genomic DNA extraction was performed following a modification of a previously described procedure (Gamper et al., 1992). Bacteria were grown overnight in LB, 1.5 ml of the culture was centrifuged for 2 min at 10,000 *g* and the pellet was washed once in TE before resuspension in 400  $\mu$ l Tris-EDTA (TE) buffer (1 mM EDTA and 10 mM Tris-HCl, pH 8.0), 50  $\mu$ l proteinase K (2.5 mg/ml), 50  $\mu$ l SDS 10 % (w/v) and 20  $\mu$ l RNaseA (5mg/ml). Cell lysis was achieved after incubation at 37 °C for 3 h. Afterwards the suspension was drawn 5 times into a syringe with a needle. The total volume was increased to 600  $\mu$ l with TE and the DNA was repeatedly extracted with phenol:chloroform (1:1) until the aqueous phase appeared clear. To precipitate the DNA, 2.5 volumes of cold EtOH 100 % (v/v) were added and the sample

was spun for 10 min at 14,000 *g*. After washing with EtOH 70 % (v/v), the DNA was dried and finally resuspended in 100 µl H<sub>2</sub>O.

#### 2.8.3.2 Isolation of plasmid DNA

Plasmid DNA isolation was performed using the Qiagen Miniprep kit (Qiagen Ltd., Surrey, UK) according to the manufacturer's protocol. Briefly, cells were pelleted from 1-10 ml of an o/n bacterial culture were subjected to alkaline lysis, neutralised and centrifuged at 13,000 *g* for 10 min to remove denatured and precipitated cellular debris. Lysate was loaded onto a silica-gel column, washed and plasmid DNA was eluted into 30-50 µl HPLC grade H<sub>2</sub>O (Fisher Scientific, UK).

#### 2.8.3.3 CTAB mini-prep for plasmid purification

For rapid extractions during routine screening, purification of plasmids was carried out using the CTAB mini-prep method (Del Sal et al., 1989). Briefly, cultures were grown o/n and 1.5 ml was centrifuged at 14,000 *g* for 3 min after which the pellet was resuspended in 200 µl of STET (8 % w/v sucrose, 50 mM Tris-HCl, pH 8.0, 50 mM EDTA) supplemented with lysozyme to a final concentration of 1 µg/ml. After incubation at room temperature for 5 min the cultures were boiled for 45 s and subsequently centrifuged for 10 min at 14,000 *g*. The pellet was removed with a toothpick and 8 µl of 5 % (w/v) hexadecyl-trimethyl-ammonium bromide (CTAB) were added to precipitate the nucleic acids. After brief centrifugation the pellet was resuspended in 300 µl NaCl (1.2 M), 750 µl of cold EtOH 100 % (v/v) was added and

centrifugation carried out at 14,000 *g* for 10 min. After washing with cold EtOH 70 % (v/v) the pellet was dried and finally resuspended in 19.5  $\mu$ l H<sub>2</sub>O and 0.5  $\mu$ l RNaseA (10 mg/ml).

#### 2.8.3.4 Isolation of large quantities of plasmid DNA

Preparation of microgram quantities of low copy number plasmids was performed using the Qiagen Midiprep kit (Qiagen Ltd., Surrey, UK) according to the manufacturer's protocol. Briefly, cells were pelleted from 100 ml of an o/n bacterial culture were subjected to alkaline lysis, neutralised and centrifuged at 10,000 rpm (10,285 *g*) in a Beckman Avanti 30 centrifuge, rotor C0650 for 30 min to remove denatured and precipitated cellular debris and then centrifuged for another 15 min. Lysate was then loaded onto a pre-equilibrated anion-exchange resin column, washed and plasmid DNA eluted with 4 ml of high-salt buffer. Finally the DNA was precipitated with isopropanol, desalted by washing with EtOH 70 % (v/v) and resuspended in 50 - 100  $\mu$ l HPLC grade H<sub>2</sub>O (Fisher Scientific, UK).

#### 2.8.3.5 Precipitation of DNA/RNA

DNA precipitation was routinely performed by adding 2.5 volumes of 100 % ethanol to the sample and 0.1 volumes of 3 M NaOAc, pH 5.2. This was then left at -20 °C for at least 20 min or o/n before centrifugation at 14,000 *g*, 20 min, 4 °C. The pellet was washed with cold 70 % (v/v) ethanol and centrifuged at 14,000 *g*, 10 min, 4 °C. The ethanol was carefully removed and

the pellet dried. The DNA was then resuspended in an appropriate volume of HPLC grade H<sub>2</sub>O.

For RNA, essentially the same protocol was used, allowing the samples to precipitate for at least 20 min or overnight at -80 °C and finally resuspending the dried pellet in DEPC-treated H<sub>2</sub>O.

#### 2.8.3.6 Polymerase chain reaction (PCR) amplification

PCR amplifications were performed according to previously described methods (Saiki et al., 1985) in a final volume of 20 µl unless otherwise stated. The reaction mix contained 0.75 µl *taq* or *pfu* polymerase (5 U/µl) and 2 µl of 10× buffer (Promega, UK), plus 20 pmol of each primer, 1 µl MgCl<sub>2</sub> 25 mM (for *taq* reactions), 2 µl of 2.5 mM dNTPs and DNA template, with optional addition of 8 % (v/v) DMSO for colony PCR. The DNA template used was either from whole cells transferred from a fresh colony or 1 µl of a (diluted if appropriate) chromosomal or plasmid preparation. Reactions were carried out in a Techne Thermal Cycler (Progene) for a total of 30 cycles. Briefly, the DNA template was initially denatured at 95 °C for 5 min, followed by 30 cycles of denaturation at 95 °C for 30 s, annealing at 50 - 55 °C for 30 s and extension at 72 °C for 30-70 s. Reaction tubes were cooled to 4 °C until needed. Annealing temperatures and extension times were adjusted to each specific pair of primers and product size respectively.

#### 2.8.3.7 DNA Clean and Concentrate (Zymoclean)

PCR products and restriction enzyme reactions were purified using Zymoclean™ DNA Clean and Concentrator (Cambridge Biosciences) as described in the manufacturer's instructions. Briefly, 2 volumes of DNA buffer were added to and mixed with each volume of DNA sample. The sample was applied to a Zymo-spin™ column and centrifuged at  $\geq 10,000$  g for 30 s and the flowthrough discarded. The column was then washed twice with 0.2 ml ethanol-containing wash buffer and centrifuged at  $\geq 10,000$  g for 30 s between washes. The flowthrough was discarded and the column was then placed in a clean Eppendorf tube and the DNA eluted with 30 - 50  $\mu$ l of distilled water or elution buffer and centrifuging for 30 s.

#### 2.8.3.8 DNA agarose gel electrophoresis

DNA loading buffer (5 $\times$  stock: 40 % (w/v) sucrose, 0.4 % (w/v) Orange G in 1 $\times$  TAE buffer (40 mM Tris-acetate, pH 8.0; 1 mM EDTA)) was added to the DNA samples and analysed on 0.6 - 2 % (w/v) agarose gels using a horizontal gel apparatus (Biorad, UK). The gels were prepared using the method described by Sambrook *et al.*, (1989) using analytical grade agarose (Promega, UK) in 1 $\times$  TAE buffer with the addition of ethidium bromide to a final concentration of 10  $\mu$ g/ml. The gels were run in 1 $\times$  TAE buffer and electrophoresis was performed at 70 - 120 V. DNA fragments were visualised on a UV transilluminator with Vision Works software (UVP, USA).

#### 2.8.3.9 DNA molecular weight markers

To establish the size of DNA fragments, 1 µg of 1 kb Plus Ladder (Invitrogen, UK) in DNA loading buffer were loaded on agarose gels.

#### 2.8.3.10 Agarose gel extraction using the Qiaquick method

PCR products were excised from agarose gels and purified using Qiaquick kits (Qiagen Ltd., Surrey, UK) as described in the manufacturer's instructions. Briefly, 3 volumes of QG buffer were added to 1 volume of gel slice which was then melted at 50 °C. 1 sample volume of isopropanol was added, mixed well, and the contents of the tube were applied to a Qiaquick column. The column was centrifuged at 13,000 g for 1 min and the flowthrough discarded. The column was then washed with 0.5 ml QG buffer and then 0.75 ml of PE buffer and centrifuged for a further 1 min. The flowthrough was discarded and the column centrifuged for an additional 1 min. The column was then placed in a clean Eppendorf tube and the DNA eluted with 50 µl of distilled water or elution buffer and centrifuging for 1 min.

#### 2.8.3.11 Agarose gel extraction using Zymoclean™

PCR products and restriction enzyme reactions were purified from agarose gels using Zymoclean™ DNA Recovery Kit (Cambridge Biosciences) as described in the manufacturer's instructions. Briefly, 3 volumes of ADB buffer were added to each volume of agarose excised from the gel in a clean eppendorf (e.g. for 100 µl (mg) of agarose gel slice 300 µl of ADB was added). The eppendorf containing the buffer and gel slice was then incubated at 37 - 55 °C

for 5 - 10 min until the gel was completely dissolved. The sample was applied to a Zymo-spin™ column and centrifuged at  $\geq 10,000 g$  for 30 - 60 s and the flowthrough discarded. The column was washed twice with 0.2 ml ethanol-containing wash buffer and centrifuged at  $\geq 10,000 g$  for 30 s between washes. The flowthrough was discarded and the column was then placed in a clean Eppendorf tube and the DNA eluted with  $\geq 6 \mu\text{l}$  of distilled water or elution buffer and centrifuging for 30 - 60 s.

#### 2.8.3.12 Phenol/chloroform purification of DNA

An equal volume of phenol equilibrated with TE buffer was added to chloroform to obtain a 1:1 mixture. This was added to the nucleic acids to be purified, vortexed and centrifuged at  $13,000 g$  for 3 min. The aqueous phase was transferred to a fresh tube, the procedure repeated as required and finally an equal volume of pure chloroform added, mixed and centrifuged as above to remove traces of phenol. The aqueous phase was again collected and 0.1 volume of 3 M NaOAc (pH 5.2) and 2.5 volumes of 100 % (v/v) EtOH were added. Nucleic acids were pelleted by centrifugation at  $13,000 g$  for 10 min. After washing with 70 % (v/v) ethanol, the nucleic acid was dried at room temperature and resuspended in TE buffer or water.

#### 2.8.3.13 DNA restriction enzymes

Restriction enzymes were purchased from Promega (UK) or New England Biolabs (UK) and were used according to the manufacturer's instructions. Reactions generally contained 0.05 - 1  $\mu\text{g}$  DNA, 0.5 - 1  $\mu\text{l}$  restriction

endonuclease and 1× restriction buffer made to a final volume of 20 µl with ddH<sub>2</sub>O and incubated at the appropriate temperature for a minimum of 1 h or until the digestion was complete. Reactions were analysed on agarose gels (0.6 - 2 %, depending on product size) and the appropriate bands cut out prior to DNA extraction.

#### 2.8.3.14 Dephosphorylation of DNA

Dephosphorylation of cleaved ends of vector DNA for ligations was carried out when required using calf intestinal alkaline phosphatase (Promega). 0.5 µl of enzyme was added to the digested DNA (~ 100 ng) which was then incubated for further 30 min at 37 °C.

#### 2.8.3.15 DNA ligation

DNA ligations were performed using 1:10 ratios of vector to insert where possible. Reactions were carried out using 0.75 µl T4 ligase (3 U/µl, Promega or NEB, USA) and 2 µl 10× T4 ligation buffer in a final volume of 20 µl. Ligations were incubated on melting ice in a Styrofoam container at room temperature o/n.

#### 2.8.3.16 Klenow fill-in

When required, overhanging DNA ends were filled in with the Klenow fragment of DNA polymerase to create blunt ends. DNA (1 µg) was incubated with 6 U Klenow fragment (Promega) and 2.5 mM dNTPs for 30 min at 37 °C.

#### 2.8.3.17 DNase digestion

A DNase digestion was performed by addition of TURBO™ DNase (Ambion) using 1 U/μg template and 10 x TURBO™ DNase buffer. The reaction was left at 37 °C for 15 - 30 min.

### 2.8.4 DNA sequencing

#### 2.8.4.1 DNA sequencing

Routine DNA sequencing was conducted by the DNA Sequencing Laboratory, Queens Medical Centre, University of Nottingham, using the Applied Biosystems BigDye® Terminator v3.1 Cycle Sequencing Kit and 3130xl Genetic Analyzer.

#### 2.8.4.2 DNA sequence analysis

Analysis of DNA sequences was performed using the Lasergene computer package (DNASTar, Ltd) or Vector NTI (Invitrogen) in combination with the BLAST programs available from the NCBI web site (<http://www.ncbi.nlm.nih.gov/>). *P. aeruginosa* sequences were analysed using the *P. aeruginosa* Genome Sequence database (<http://www.pseudomonas.com>).

## 2.8.5 Gene replacement in *P. aeruginosa*

### 2.8.5.1 Conjugation of plasmid DNA into *P. aeruginosa*

Plasmid transfer from *E. coli* donor strains to *P. aeruginosa* recipient cells was carried out by bacterial mating. Both donor and recipient cells were grown o/n in 5 ml of LB with shaking. *P. aeruginosa* recipient strains were grown at 42 °C to reduce the activity of the restriction-modification system which degrades incoming foreign DNA whilst *E. coli* donor strains were grown at 37 °C. 1.5 ml of each culture were centrifuged at 13,000 g for 5 min and washed twice with 1 ml fresh LB broth. Pellets were resuspended in 0.5 ml LB broth, mixing donor cells with recipient bacteria in a sterile eppendorf. The resulting 1 ml of bacterial mix was centrifuged at 13,000 g for 5 min and the resulting pellet resuspended in its equivalent volume of fresh LB. Conjugations were achieved by spotting the mixed bacteria onto an LB agar plate, allowing drying before incubating at 37 °C for 4 - 8 h. Cells from the plate were then harvested, resuspended in 1 ml of LB broth and plated onto PIA agar plates containing antibiotics to select for *P. aeruginosa* transconjugants. Plates were incubated between 24 and 48 h at 37 °C.

### 2.8.5.2 Sucrose counter-selection

Suicide plasmids used to perform gene replacements during this study carried the *sacBR* locus that allows its counter-selection. Single colonies from the first cross-over were re-streaked and grown o/n in LB broth. Then they were diluted  $10^6\times$  in LB broth containing 20 % (w/v) sucrose and allowed to grow o/n to counter-select for cells having achieved the second cross-over. Dilutions were

then plated onto sucrose plates to obtain single colonies. Colonies that grew were checked for loss of the suicide plasmid by screening for antibiotic sensitivity.

## **2.8.6 RNA work**

To minimise RNase contamination, all RNA work was carried out in designated clean areas. Separate pipette tips and microcentrifuge tubes were used and when possible solutions were treated with 1 % (v/v) DEPC and autoclaved.

### *2.8.6.1 In vitro* transcription

*In vitro* transcription of DNA fragments was performed using the RiboMAX™ Large Scale RNA production system (Promega) according to the manufacturer's manual. Briefly, 4 µl of 5 × transcription buffer, 6 µl of rNTPs (ATP, CTP, GTP and UTP mix, 25 mM each), 8 µl of template DNA and 2 µl of enzyme mix from the kit were mixed together at room temperature and incubated at 37°C for 3.5 h. The reaction was subsequently subjected to DNase digestion (2.8.3.17: 1 U DNase per µg template). The reaction was left at 37 °C for 15 - 30 min).

### *2.8.6.2* RNA extraction (phenol-chloroform)

RNA transcribed *in vitro* was purified by phenol:chloroform extraction using acidified phenol:chloroform premixed with isoamyl alcohol (125:25:1)

saturated with citrate buffer (citric acid) at pH 4.5. The extraction was repeated twice and then the sample was extracted once with chloroform. The RNA was desalted using a Sephadex Mini Quick Spin Column (Roche Diagnostics,) and precipitated with 0.1 volume of 3 M NaOAc, pH 5.2 and 2.5 volumes of 100 % (v/v) EtOH. Finally the RNA was resuspended in DEPC-H<sub>2</sub>O and stored at -80 °C.

#### 2.8.6.3 Total RNA extraction (Qiagen)

RNA was purified from a 1 L growth of LB (2 L flask) grown at 37 °C, 180 rpm inoculating with 1:1000 ratio of inoculant. RNA samples were taken in triplicate at the exponential and late-exponential growth phases and immediately treated with RNA Bacteria Protect solution (Qiagen). The total RNA samples extracted from the growth were added to 5 ml (2 vol) of RNA Bacteria Protect Reagent. The samples were vortexed for 5 s and left at room temperature for 5 min. The samples were then centrifuged at 3000 - 5000 g for 10 min before removing the supernatant. The pellets were stored at -20 °C.

RNA was extracted using the RNeasy Midi kit, eluting in 2 x 150 µl elutions for a final volume of approx 230 µl. A DNase digestion was performed by addition of 25 µl 10 x TURBO DNase buffer and 5 µl of TURBO DNase. The reaction was left at 37 °C for 30 min. The RNA was recovered using the RNeasy MinElute kit (Qiagen), eluting with 16 µl nuclease-free water for a final volume of 14 µl.

#### 2.8.6.4 RNA Cleanup

RNA was purified after DNase digestions using the RNeasy MinElute Cleanup kit (Qiagen) according to the manufacturer's instructions. Briefly, the sample volume was adjusted to 100  $\mu$ l using nuclease-free water before the addition of 350  $\mu$ l of RLT Buffer (contains guanidine thiocyanate, to which 10  $\mu$ l  $\beta$ -mercaptoethanol is added per ml of RLT). To this 250  $\mu$ l of 96-100 % ethanol was added before transferring to an RNeasy MinElute spin column (Qiagen). The column was washed with 500  $\mu$ l of Buffer RPE (contains ethanol) followed by 500  $\mu$ l of 80 % ethanol. After transferring to a fresh collection tube, the spin column was opened and dried by centrifugation at high speed for 5 min. The RNeasy MinElute spin column was transferred to a 1.5 ml eppendorf tube and 14  $\mu$ l nuclease-free water was pipetted onto the centre of the column membrane. The column was left for 1 min before elution by centrifugation at high speed for 1 min.

#### 2.8.6.5 RNA molecular weight markers

To establish the size of RNA fragments, 1.5  $\mu$ g of RNA ladder, low range (Fermentas, UK) in 1 $\times$  urea loading buffer were treated like the samples and simultaneously loaded onto the gels.

### **2.8.7 Protein Methods**

#### 2.8.7.1 Protein expression

RsmA proteins (wild type and modified variants) from *P. aeruginosa* and likewise RsmA homologues from various organisms were expressed from

plasmids either based on pME6032 or pHLT in the *E. coli csrA* mutant strain TR1-5, in the laboratory strain C41 (DE3), in the *P. aeruginosa* wild type strain PAO1 or the *rsmA* mutant strain PAZH13, and purified by nickel-loaded nitrilo-triacetic (Ni-NTA) affinity chromatography as previously described (Heeb et al., 2002). Briefly, 2 ml of an o/n culture of the overproducing strain were used to inoculate 200 ml of LB broth and grown for 3 h at 37 °C to early exponential phase ( $OD_{600} \sim 0.3$ ). Then, IPTG was added to a final concentration of 1 mM. The culture was grown for further 6 h, centrifuged and the pellet was stored at -80 °C.

#### 2.8.7.2 Purification using hexahistidine tags and Ni-NTA chromatography

This method of purification is based on the selectivity and affinity of the nickel nitrilotriacetic acid (Ni-NTA) metal-affinity chromatography for biological molecules which have been tagged with six consecutive histidine residues. When needed, the pellet was thawed and resuspended in 4 ml of lysis buffer (50 mM  $NaH_2PO_4$ , 300 mM NaCl, 10 mM imidazole, pH 8.0). Lysozyme was added to a final concentration of 1 mg/ml and the suspension was incubated on ice for 1 h. Cells were sonicated on ice ( $9 \times 10$  s, with 10 s cooling intervals). The lysate then was drawn 5 times through a syringe with needle and centrifuged for 30 min at 10,000 rpm in a Beckman Avanti 30 centrifuge, rotor C0650. To 4 ml of the clear supernatant, 1 ml of 50 % (w/v) Ni-NTA slurry (Qiagen) was added and binding of the hexahistidine-tagged proteins allowed to occur for 1 h at 4 °C with gentle shaking. The sample was loaded onto an empty 1 ml column and washed once with 5 ml of lysis buffer. Washing was

performed by running  $4 \times 5$  ml of washing buffer (50 mM  $\text{NaH}_2\text{PO}_4$  buffer pH 8.0, 300 mM NaCl, and 10 - 100 mM imidazole) through the column. Elution was done for each column by running and collecting separately  $4 \times 500$   $\mu\text{l}$  of elution buffer (50 mM  $\text{NaH}_2\text{PO}_4$  buffer pH 8.0, 300 mM NaCl, 300 mM imidazole).

#### 2.8.7.3 Protein purification - HisPur<sup>TM</sup> cobalt resin

Cobalt resin is used to purify proteins from total soluble protein extract using a cobalt-charged tetradentate chelator immobilized on 6 % cross linked agarose. The resin has a binding capacity of  $\sim 10$  mg at  $> 90$  % purity of a 28 kDa His-tagged protein per millilitre of resin.

To purify, a cell pellet was removed from the  $-80$  °C freezer and allowed to defrost at room temperature for 60 - 75 min before being resuspended in 3 ml of 1 x IMAC buffer (20 mM NaP 0.5 NaCl pH 7.4) with DNase added (100  $\mu\text{l}$  of 10 mg/ml in 1M  $\text{MgCl}_2$  and 0.1M  $\text{MnCl}_2$ ). The sample was transferred to a sonication glass container along with 1 ml 1 x IMAC of washings and sonicated (10 times 30 s with cooling periods over ice every 30 s). The lysate was transferred to a plastic centrifuge tube and centrifuges for 30 min at 30000 - 40000 g at 4 °C depending on viscosity of pellet. The clear supernatant was then added to 2 ml of HisPur<sup>TM</sup> cobalt resin (Pierce) and equilibrated for 30 min tumbling slowly at 5 °C. After collecting the flowthrough, various washing stages are used, each of 20 ml. First washed with 1 x IMAC and 1 mM imidazole (in 1 x IMAC), proceeded with washes of ddH<sub>2</sub>O, 2M NaCl, ddH<sub>2</sub>O, 1 % Triton X-100 (non-ionic surfactant). For each of these, 2 ml is eluted down

the column and collected before washing in 4 - 5 ml stages. The water, salt and Triton X-100 washes were repeated twice more. After a further wash using ddH<sub>2</sub>O, elution was carried out using 1 M imidazole. 2 ml was collected in a 2 ml eppendorf before putting the stopper on the bottom of the column and adding 5ml 1M imidazole followed by the lid. The column was left tumbling slowly at 5 °C for 10 min before eluting. This was repeated three more times. After elution, the column was washed with progressively lower imidazole washes (500, 250, 100 and 50 mM) before washing with ddH<sub>2</sub>O, 1 x IMAC and ddH<sub>2</sub>O and storage in 20 % v/v EtOH.

#### 2.8.7.4 Scale up

When large amounts of protein were required, expression was scaled up using essentially the protocol described in section 2.8.7.1, with the difference that only high expression constructs that incorporate a thrombin cleavage site between the wanted protein and the hexahistidine tag were used. When lysing the cells DNase was added to reduce the viscosity. All volumes were increased proportionally except for the Ni-NTA slurry, which was kept the same as its binding capacity is up to 10 mg of protein per ml.

#### 2.8.7.5 Thrombin cleavage

Purified proteins containing a thrombin cleavage site were cleaved by adding 10 - 40 units of thrombin (bovine  $\alpha$ -thrombin, Cambridge Biosciences Ltd. Haematologic Technologies Inc.) per mg of fusion protein and leaving the sample shaking o/n at room temperature.

#### 2.8.7.6 Desalting

Protein solutions were either desalted using Zeba Spin Desalting Columns (Pierce) according to the manufacturer's instructions, or using a HiTrap™ column (Amersham Bioscience). The latter were equilibrated, protein was applied to the column and eluted with thrombin cleavage buffer (20 mM Tris, 150 mM NaCl, 2.5 mM CaCl<sub>2</sub>, pH 8.4). Fractions of 10 ml were collected and the absorbance at 280 nm recorded. When desalting into water, the column was equilibrated with 3 × column volumes of autoclaved ddH<sub>2</sub>O prior to injection. The protein was eluted into water using the same method as above.

#### 2.8.7.7 Ionic exchange

An FPLC system (ÄKTA prime, Amersham Pharmacia) was used to purify RsmA and the RsmA mutant proteins. Anion exchange was carried out using a HiTrap™ Q Sepharose 5ml High Performance column (Amersham Pharmacia). The sample was then loaded onto an anionic exchange column, which had previously been equilibrated with buffer A (50 mM K<sub>2</sub>HPO<sub>4</sub> pH 8.8, filtered through cellulose nitrate membrane filters). The bound protein was eluted by increasing the salt concentration by introducing a step gradient to 10 % B, followed by a linear gradient from 10 % to 100 % buffer B (50 mM K<sub>2</sub>HPO<sub>4</sub>, 2 M NaCl, pH 7.8 filtered) over 100 ml. Upon completion of this gradient the column was flushed with 100 % B to ensure complete removal of any residual protein. Samples taken from the anion exchange step were analysed by SDS-

PAGE, and those identified as containing the protein of interest were freeze dried.

For cationic exchange, the same procedure was carried out using a HiTrap<sup>TM</sup> SP 5 ml High Performance column. Buffer A was 1 x TAE (Tris -Acetate-EDTA pH 4.0, filtered) and Buffer B was 1 x TAE 2M NaCl (filtered).

#### 2.8.7.8 HiTrap<sup>TM</sup> heparin affinity column

HiTrap<sup>TM</sup> heparin affinity columns are used for the separation of many proteins including DNA binding proteins. Heparin consists of alternating units of uronic acid and D-glucosamine substituted with one or two sulphate groups, which is covalently coupled to cross-linked agarose beads. The ligand used was sulphated glucosaminoglycan. The column has a binding capacity of ~3 mg antithrombin III (bovine) per millilitre of medium.

The sample was then loaded onto the heparin column, which had previously been equilibrated with buffer A (50 mM K<sub>2</sub>HPO<sub>4</sub> pH 7, filtered through cellulose nitrate membrane filters). The sample, typically 4 - 8 ml, was diluted using buffer A to ~ 50 ml and loaded onto the column using a Superloop (GE Healthcare). The bound protein was eluted by increasing the salt concentration by introducing a linear gradient from 0 % to 60 % buffer B (50 mM K<sub>2</sub>HPO<sub>4</sub>, 2 M NaCl, pH 7 ) over 240 ml. The fractions containing the identified protein were collected and freeze dried.

#### 2.8.7.9 Gel filtration

Gel filtration (size exclusion chromatography) was carried out using a high load Superdex™ 200 10/300 GL (Amersham Biosciences) onto which the filtered sample was injected. Prior to loading, the column was equilibrated with buffer A of a high pH (50 mM K<sub>2</sub>HPO<sub>4</sub> pH 8) or low pH buffer (50mM NaAc pH 4.5) through cellulose nitrate membrane filters. The protein was eluted in 10 ml fractions, identified using SDS-PAGE from samples taken prior to being freeze-dried.

#### 2.8.7.10 Superloop

In order to load the larger samples for the anionic exchange, a 50 ml Superloop (GE Healthcare) was used instead of an ordinary sample loop (e.g. 5 ml).

#### 2.8.7.11 Freeze-drying

Protein solutions were frozen in liquid nitrogen and then freeze dried overnight (~12 - 16 h) under vacuum using the MicroModulyo freeze drier from Thermo Scientific according to the manufacturer's instructions.

#### 2.8.7.12 Anionic exchange

To separate protein and tag after thrombin cleavage, anionic exchange chromatography was performed using a HiTrap™ Q Sepharose 5ml High Performance column (Amersham Pharmacia) on an FPLC system (ÄKTA

prime, Amersham Pharmacia). A salt gradient (50 mM  $K_2HPO_4$ , pH 8.0, 0 - 2 M NaCl) was used to separate the different polypeptides.

#### 2.8.7.13 Circular dichroism spectroscopy (CD)

CD spectra were recorded on an Applied Photophysics Pi-Star-180 Spectrophotometer. The temperature was regulated using a Neslab RTE-300 circulating programmable water bath and a thermoelectric temperature controller (Melcor). A correction for the CD spectra was made for the buffer. The sample was read in a cuvette of path length 1 mm. Spectra were recorded from 300 nm to 180 nm to characterise the secondary structure content in 2 - 4 nm steps and 4.0 nm entrance and exit slit widths. The absorbance readings are given in molar ellipticity (millidegrees).

#### 2.8.7.14 Estimation of protein concentration using the Bradford assay

To estimate the protein concentration of a sample, the Bradford assay (Bradford, 1976) was used. In a 1 ml cuvette, the solution to be assayed was added in a volume of 1 - 50  $\mu$ l and made up to 800  $\mu$ l with the appropriate buffer solution or  $H_2O$ . 200  $\mu$ l of Bradford reagent (Sigma, UK) was added and the cuvette incubated at room temperature for 5 min. Absorbance was then read at 595 nm ( $A_{595}$ ). Protein concentration was estimated using a standard curve of bovine serum albumin (BSA) concentrations vs  $A_{595}$  in buffer solution.

### 2.8.7.15 Tricine SDS-PAGE

To achieve greater resolution of low molecular weight proteins, Tricine SDS-polyacrylamide gels were used. A separating gel of the appropriate percentage acrylamide was cast and overlaid with a 4 % (w/v) acrylamide stacking gel (Table 2.4).

	Separating gel		Stacking gel
	10 %	18 %	4 %
30 % (w/v) Acrylamide: Bisacrylamide	3.3 ml	6 ml	670 $\mu$ l
Gel buffer	3.3 ml	3.3 ml	1.25 ml
dH <sub>2</sub> O	3.4 ml	0.7 ml	3.03 ml
10 % APS	50 $\mu$ l	50 $\mu$ l	50 $\mu$ l
TEMED	10 $\mu$ l	10 $\mu$ l	10 $\mu$ l

**Table 2.4: Tricine-SDS-PAGE separating and resolving gel solution components**

Gel buffer: 3M Tris-HCl; 0.5% (w/v) SDS; pH 8.45

An appropriate volume of Tricine sample buffer (20  $\mu$ l  $\beta$ -mercaptoethanol added to 980  $\mu$ l of buffer (50 mM Tris-HCl, pH6.8; 100 mM DTT; 2 % (w/v) SDS; 0.1 % (w/v) bromophenol blue; 10 % (v/v) glycerol)) was added to samples and heated at 90 °C for 2 min. Aliquots of 5-15  $\mu$ l of the samples were loaded onto the gel. Electrophoresis was performed using an anode buffer of 0.1 M Tris-HCl, pH 8.9 and a cathode buffer of 0.1 M Tris-HCl, 0.1 M Tricine, 0.1 % SDS. The samples were run through the gel with a voltage of 150 V - 200 V. Precision Plus Protein All Blue Standard (BioRad, UK) was used as a molecular weight marker.

#### 2.8.7.16 Tricine-SDS-PAGE for Western blot

Tricine-SDS-PAGE is the preferred gel system for the resolution of proteins smaller than 30 kDa, however by making up the acrylamide and bis-acrylamide solutions separately allows the adaptation of the conditions to the experiment (Table 2.5) (Schagger, 2006).

		Separating Gel 4 %	Resolving Gel 16 %/6 M Urea
AB-3	(ml)	1	5
Gel buffer (3×)	(ml)	3	5
Urea	(g)		5.4
Add dH <sub>2</sub> O to final volume	(ml)	12	15
10 % APS	(μl)	90	100
TEMED	(μl)	9	10

**Table 2.5: Tricine-SDS-PAGE separating and resolving gel solution components for Western Blotting.**

AB-3 (49.5 % T, 3 % C): Dissolve 48 g acrylamide and 1.5 g bisacrylamide in 100 ml final volume of water.

Gel Buffer: 3 M Tris, 1 M HCl, 0.3% SDS pH 8.45.

An appropriate volume of Tricine loading buffer (50 mM Tris-HCl pH 6.8, 12.5 mM EDTA; 1 % β-mercaptoethanol; 2 % (w/v) SDS; 0.02 % (w/v) bromophenol blue; 10 % (v/v) glycerol) was added to samples and heated at 90 °C for 2 min. Aliquots of 5 - 15 μl of the samples were loaded onto the gel. Electrophoresis was performed using an anode buffer of 0.1 M Tris-HCl, pH 8.9 and a cathode buffer of 0.1 M Tris-HCl, 0.1 M Tricine, 0.1 % SDS, pH ~8.25. The samples were run through the gel using an initial 30 V until the samples had passed into the resolving gel. The voltage was then increased to 150 V for approximately 4 h. Colour Marker Ultra Low Range (1,060–26,600 MW) (Sigma Aldrich, UK) was used as molecular weight markers.

#### 2.8.7.17 Coomassie staining

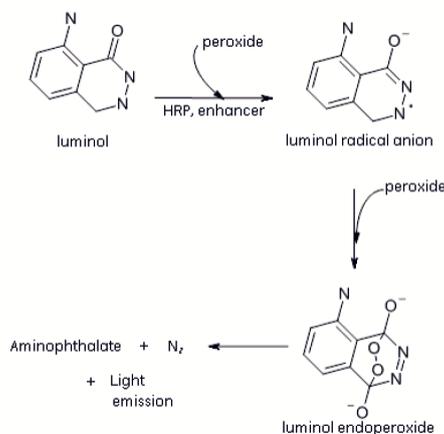
Protein gels were stained with Coomassie Brilliant Blue solution (45 % (v/v) methanol, 10 % (v/v) acetic acid, 0.025 % (w/v) Coomassie Brilliant Blue R250) and destained with 15 % (v/v) isopropanol, 10 % (v/v) acetic acid.

#### 2.8.7.18 Western blotting

To detect proteins of interest, proteins were transferred onto Immobilon-P<sup>SQ</sup> PVDF membranes (Millipore) using a Trans-Blot SD semi-dry transfer cell (Biorad, UK). Blotting was carried out in transfer buffer (12 mM Tris base; 10 mM glycine; 0.04 % SDS and 10 % (v/v) methanol) at 15 V for 15 - 20 min at room temperature. To block the membrane, TBS (10 mM Tris Base, 50 mM NaCl pH 7.6) with 0.1 % (v/v) Tween 20 (abbreviated to TBST) and 1 % (w/v) casein hydrolysate (Sigma) was added overnight with shaking at 4 °C. The primary antibody was diluted as appropriate in TBST-1 % w/v casein blocking solution and incubated with the membrane for 1 h shaking at room temperature. The membrane was washed 3 × 15 min in TBST. The secondary antibody was diluted as appropriate in TBST-1 % wt/v casein and incubated with the membrane for 1 h at room temperature with shaking. The membrane was washed 3 × 5 min, 2 x 15 min and 3 x 5 min in TBST before developing the blot using the ECL Advance Western Blotting Detection System (Amersham Biosciences, UK) as described in section 2.8.7.18.1. Blots were exposed to Hyperfilm<sup>TM</sup> chemiluminescence film (Amersham Biosciences).

### 2.8.7.18.1 Detection of Proteins after Western blotting

Western blots were developed using the ECL Advance Western Blotting Detection System (Amersham Biosciences, UK) according to the manufacturer's instructions (Kricka, 2003). This method is for the detection of immobilized specific antigens conjugated to Horseradish Peroxidase (HRP) labelled antibodies. Briefly, for each blot, 500  $\mu$ l of solution A (luminol solution) was mixed with 500  $\mu$ l of solution B (peroxide solution).



**Figure 2.3: Chemiluminescence production by the ECL detection system.**

The peroxide-catalyzed oxidation of luminol generates weak chemiluminescence at 425 nm. With Amersham ECL Prime detection reagent, incorporation of a redox mediator, or enhancer, into the buffer improves the enzyme turnover and increases the equilibrium concentration of the luminol radical anion.

The peroxidase acts as a catalyst for the oxidation of luminol, generating chemiluminescence at 425 nm (Fig. 2.3). The detection reagents include an enhancer which improves enzyme turnover and increase the equilibrium concentration of the luminol radical ion. This shift improves both the signal intensity and duration.

Excess reagent was drained off and the blot placed protein side up inside a plastic shield in an X-ray cassette. In a dark room using safety lights a sheet of

autoradiography film (Amersham Hyperfilm ECL) was placed on top of the membrane. The cassette was closed and exposed for 2 - 15 min. The film was removed and developed in a tray using X-ray film processing developer. The film was washed in water and fixed using a fixing solution (Kodak, GBX solutions).

#### 2.8.7.18.2 PVDF membrane dye

After electroblotting it is possible to visualise the proteins on a wet PVDF membrane prior to blocking by staining for 5 min (25 % methanol, 10 % acetic acid and 0.02 % Coomassie blue G-250 dye). The gel was then destained twice for 10 min (25 % methanol, 10 % acetic acid). If the membrane was needed to complete the western blot, the dye was removed with 100 % methanol and the membrane washed thoroughly in water before continuing.

#### 2.8.7.18.3 Stripping immunoblots

Western blots can be stripped after development for re-probing with a different antibody or for visualization using the PVDF stain.

The blot was rinsed in water before immersing in 3 % w/v trichloroacetic acid (TCA) with shaking for 10 min. The blot was washed in water for 2 × 20 min with shaking before washing the blot using running water for 5 min.

#### 2.8.7.18.4 Peptide mass fingerprinting

Protein samples are derived from SDS-PAGE and after subsection to chemical modification. The proteins are cut into several fragments using proteolytic

enzymes. The resulting peptides are extracted with acetonitrile and dried under vacuum. The peptides are then dissolved in a small amount of distilled water prior to mass spectrometric analysis.

MALDI-TOF MS (Matrix assisted laser desorption ionisation-time of flight, coupled to mass spectrometry) uses energy from a laser directed at the sample mixed with a chemical matrix (usually an organic acid derivative) in order to generate ions which mass are then determined in a time-of-flight type analyser. The ions generated via this process represent the intact peptides resulting from the tryptic digest of the target protein (peptide mass fingerprint).

These values may then be used as a data set to challenge databases containing lists of expected peptide masses that would result from the theoretical tryptic digestion of proteins currently held in Swiss-Prot and TrEMBL databases.

The procedure was performed in the Post-Genomics Technologies Facility, Queens Medical Centre, University of Nottingham using a Micromass M@aldi MS (BSAU).

## **2.8.8 Protein-RNA interactions**

### **2.8.8.1 Electrophoretic mobility shift assay (EMSA)**

The wild type RsmN protein was assayed for its capacity to bind to *rsmZ* transcribed *in vitro* as previously described where the detection was performed after electrotransfer of the RsmN-RNA complexes to a Hybond-N (nylon) membrane followed by Northern hybridisation with an DIG-labelled DNA probe (Heeb et al., 2002, Heeb et al., 2006). Briefly, binding reactions with a total volume of 10 µl were set up (1 µl gel-shift buffer 10 × (10 mM

Tris-acetate, pH 7.5, 10 mM MgCl<sub>2</sub>, 50 mM NaCl, 50 mM KCl, 5 % (v/v) glycerol), 1 µl DTT 100 mM, 1 µl yeast tRNA (30 ng), 1 µl RNase inhibitor (4 U), 1 µl RNA 200 nM, 1 µl H<sub>2</sub>O and 4 µl protein of appropriate dilution) and incubated at 30 °C for 30 min. The samples were run on a native polyacrylamide gel (1 ml 1× TBE, 5.5 ml DEPC-H<sub>2</sub>O, 3.5 ml acrylamide-bisacrylamide 40 % (19:1), APS and TEMED) in 1× TBE at 100 - 150 volts for 2 - 4 h. Gels were blotted onto nylon membrane in 1 × TBE at 30 volts for 30 min. After rinsing the membrane in 2 × SSC the nucleic acids were cross linked to the membrane by UV-light.

#### 2.8.8.2 Detection of RNA on nylon membranes

Blotted membranes were pre-hybridised for 1 h in high SDS pre-hybridisation buffer (formamide 50 %, SSC 5×, sodium phosphate buffer 50 mM, pH 7.0, blocking reagent 2 % (w/v), *N*-laurylsarcosine 0.1 % (w/v), SDS 7 % (w/v)) and then hybridised (same buffer including the DIG-labelled probe) overnight at 50 °C. Stringency washes were carried out at room temperature after the hybridisation (2× 15 min in 2× SSC, 0.1 % (w/v) SDS and 2× 15 min in 0.5× SSC, 0.1 % (w/v) SDS) and the detection procedure followed as described in section 2.8.7.19.1.

#### 2.8.8.3 Deep-Seq analysis

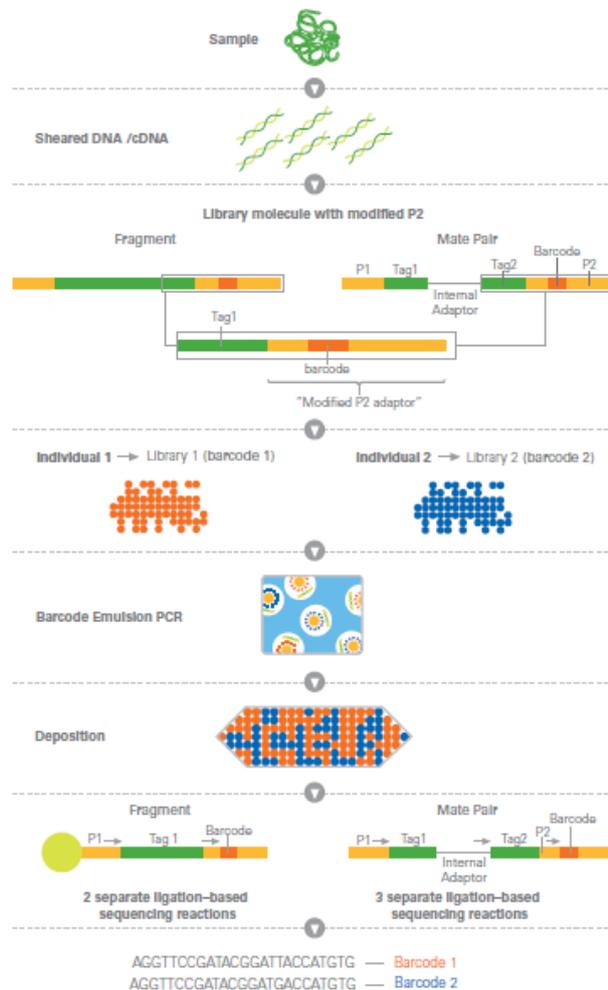
The amplified RNA libraries were obtained by first hybridization and ligation of the RNA. cDNA synthesis and amplification were performed according to the supplier's protocol using the SOLiD Total RNA sequencing kit. The yield

and size distribution of the amplified DNA was confirmed using the Agilent 2100 Bioanalyzer with the DNA 1000 kit (Agilent). The resulting libraries were assigned a specific barcode and pooled. Each library template was clonally amplified on SOLiD P1 DNA beads by emulsion PCR. After PCR the templates are denatured and bead enrichment was performed. The modified beads were deposited on a glass slide, prior to sequencing by ligation using fluorescently labelled probes. Data analysis was performed using SOLiD Bioscope software 1.3.1 (Applied Biosystems.) and a whole transcriptome pipeline was run for each of the eight samples individually. The output files were alignment BAM files which had been checked for possible PCR duplicates.

#### 2.8.8.3.1 Barcoding

SOLiD system barcodes contain unique sequences designed for optimal multiplexing (Parameswaran et al., 2007). Sixteen different barcodes were selected based on uniform melting temperature, low error rate and orthogonal sequences that are unique in colour space. Barcodes are added to the 3' end of the target sequence using a modified version of the P2 adaptor (Figure 2.4). SOLiD system barcoding enables the assignment of a unique identifier to the template beads that are made from one individual library. Once these identifiers are assigned, multiple batches of template beads may be pooled together and sequenced in a single flowcell run. The combination of two sequencing slides with eight segments each and the capacity of sixteen different barcodes enables the interrogation of up to 256 samples in a single run. Data analyses can then trace the sequence data back to a specific sample using its respective identifier. Following sequencing of the target DNA,

additional rounds of ligation based sequencing are performed using the primer sets complimentary to the barcode. The resulting reads can then be sorted by the barcode and aligned into groups to the reference sequence.



**Figure 2.4: Integration of SOLiD system barcodes into the library construction workflow** Barcodes are added to the 3' end of the target sequence using a modified version of the P2 adaptor. Once assigned, multiple batches of template beads may be pooled together and sequenced in a single flowcell run. Data analyses can then trace the sequence data back to a specific sample using its respective identifier.

### 2.8.8.3.2 Analysis

The Whole Transcriptome Analysis (WTA) in BioScope™ Software aligns to a reference genome. Using the mapping results, WTA counts the number of tags aligned with exons, and can convert the \*.bam file to a Wiggle File (\*.wig).

Analysis was performed by S. Heeb on the \*.wig files containing reads of RNA with the rRNA retained. As there was no internal standard that can be used to compare the total RNAs (samples 2, 4, 6 and 8) with the samples enriched in RNAs that bind RsmN or RsmA (samples 1, 3, 5, and 7), the data in the wiggle files was first be normalised to the average of their values.

To calculate the averages, all the values in a file were added up and divided by the length of the chromosome. Average reads per nucleotide and standard deviations were calculated for each of these files. Once the normalised wiggle files had been created, enrichment factors between RNAs extracted with RsmN or RsmA versus the corresponding total RNAs were calculated for each nucleotide in the genome. For practical purposes this factor was multiplied by 100, so that it will be greater than this number if there had been enrichment in a particular nucleotide, or smaller if there had been depletion. To avoid division by zero errors, the arbitrary value of 9999 was used instead for undetermined enrichment factors (*i.e.*, every time that a nucleotide produced reads in the enriched but not in the corresponding total RNA sample). The program to do this also uses the genomic position of the nucleotide and the strand from which its reading originated to obtain additional information about its genomic context.

#### 2.8.8.4 Protein-RNA experiments

##### 2.8.8.4.1 Ni-NTA column

Protein from a 250 ml culture was obtained by sonication and bound to 2 ml of Ni-NTA agarose suspension as previously described in sections 2.8.7.1 and

2.8.7.2. The normal sequential washes were performed with 2 x 5 ml H<sub>2</sub>O, 2 M NaCl, H<sub>2</sub>O, 1 % Triton X-100. The second wash of H<sub>2</sub>O is to remove concentrated NaCl prior to the detergent wash. These were repeated 4 times. The column was then washed with normal lysis buffer pH 8 and subsequently stored in this buffer overnight prior to RNA binding. The column was then washed with 2 × 5 ml of 1 × Interaction buffer (10 mM Tris-acetate pH 7.5, 10 mM MgCl<sub>2</sub>, 50 mM NaCl, 50 mM KCl, 10 mM imidazole and 5 % (w/v) glycerol) to enable total buffer exchange prior to RNA binding studies.

Interaction Buffer	1 ×	75 µl of 10 × stock
β mercaptoethanol	10 mM	75 µl of 100 mM solution
Yeast tRNA	30 ng/µl	75 µl of 1:333 dilution
RNase inhibitor	4 U/µl	75 µl of 1:10 dilution
<b>RNA</b>	53 µg	75 µl
DEPC H <sub>2</sub> O		375 µl

**Table 2.6: Interaction Buffer B to optimise protein-RNA binding (Volume dependent on volume of RNA used).**

The column was plugged at the bottom prior to the RNA containing solution (Interaction buffer B, Table 2.6) being added. A further 2 ml of 1× Interaction Buffer was added to ensure the tumbling of the protein-RNA mixture was of sufficient volume to occur. The column lid and bottom were sealed with parafilm and tumbled for 1 h at 4 °C.

The flow through was collected as were the subsequent washes. The washes consisted of 10 ml each of Interaction buffer A, Interaction buffer C (A + 1 % Triton X-100), Interaction buffer A.

The elutions followed consisting of  $1 \times 500 \mu\text{l}$  Interaction buffer D (A + 1 M NaCl),  $8 \times 500 \mu\text{l}$  Elution buffer (50 mM NaP, 300 mM NaCl, 300 mM imidazole pH 8) and  $2 \times 500 \mu\text{l}$  1 M imidazole.

The column was cleaned by agitation with 0.5 M NaOH for 30 min at RT and stored by washing with water prior to storage in 25 % ethanol.

#### 2.8.8.4.2 Ni-NTA magnetic beads

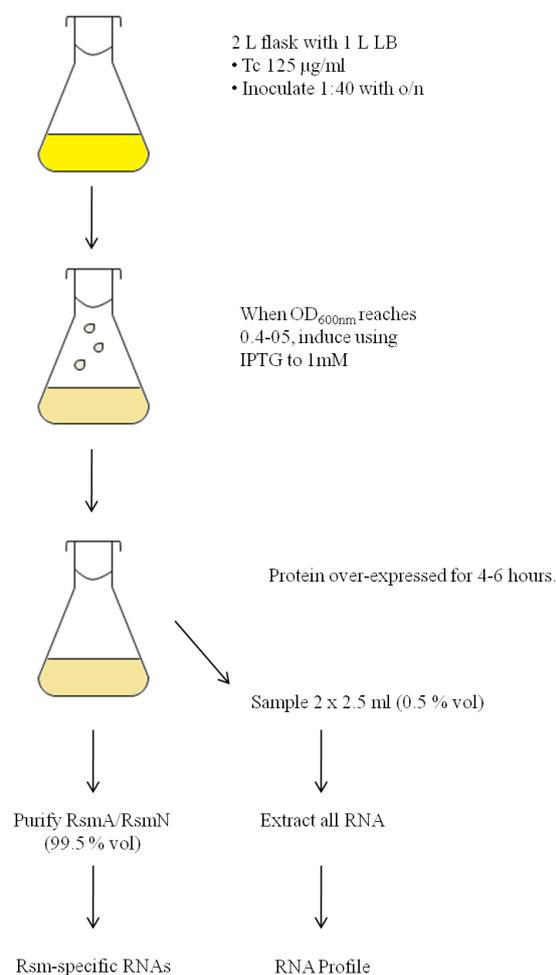
The magnetic beads work using the same principle as the Ni-NTA resin, involving the capture of the 6xHis-tagged proteins followed by washing, binding of interaction partners, further washing, and finally elution of the interacting partner from the still immobilized 6xHis-tagged protein or elution of the interacting partner-6xHis-tagged protein complex. Between each step, the beads are collected by attracting them to the side of the vessel, after placing near a magnet, enabling removal of the solutions. This separation holds the protein on the sides of the wells while the buffers are exchanged to wash or elute the 6xHis-tagged proteins. The beads are easily resuspended by agitation. The advantages of using the magnetic beads include adjusting the amount of the magnetic beads and therefore binding capacity allows flexibility when tailoring the amount of protein purified for a particular experiment. Elution of smaller volumes,  $500 \mu\text{l}$  magnetic bead elution compared to 5 ml resin elution, is preferable for limiting RNA loss when collecting RNA from the eluted samples. The experiment is fast, allowing a high throughput and can be used without prior protein purification if required.

Using purified samples of RsmA and RsmN, the proteins were bound to the Ni-NTA magnetic beads (Qiagen) using 1 × interaction buffer. This was done by measuring 0.9 mg of each protein and resuspension in 1.5 ml of 10 × Interaction buffer. As 100 µl of the 5 % (v/v) magnetic bead suspension has a maximum binding capacity of 30 µg protein (based on 6xHis-tagged dihydrofolate reductase (DHFR, approximately 12.5 nmol per ml, molecular weight: 24 kDa), this is the total protein needed in the 500 µl volume used. 100 µl of each 10× protein solution was aliquoted to a new eppendorf and 900 µl of HPLC H<sub>2</sub>O was added to give 1 ml of 1 × protein solution.

The protein was bound to the beads by incubation on an end-over-end shaker for 1 h. The supernatant was retained after the tube was placed on a magnetic separator for 1 min. Following a wash with Interaction buffer A to allow for buffer exchange, Interaction buffer B containing the RNA sample was added to each protein sample. These were incubated like the column with the lid sealed with parafilm and tumbled for 1 h at 4 °C. A series of washes were used to remove non-specific RNAs from the total RNA sample, consisting of 500 µl of Interaction buffer A, Interaction buffer C (A + 1 % Triton X-100), A repeat. In order to elute, 50 µl Interaction buffer D (as above) was mixed with the beads, quickly vortexed to ensure thorough mixing, pulsed in centrifuge and allowed to incubate at room temperature for 1 min. The solution was then removed following magnetic separation. This was repeated 8 times with normal elution buffer followed twice more using 1 M imidazole.

### 2.8.8.4.3 RNA extraction after overexpression of *rsmA* and *rsmN*

Strains PAO1/pRsmA and PAO1/pRsmN were grown in 1 L LB in 2 L flasks at 37 °C with shaking at 180 rpm (inoculated 1:40 ratio). Expression of RsmA/N was induced by adding IPTG to a final concentration of 1 mM when OD<sub>600nm</sub> reached 0.4 - 0.6 and the cloned RsmA or RsmN left to express for 4 to 6 h. The whole culture was then centrifuged and the pellet stored at -80 °C until needed.



**Figure 2.5: Schematic diagram for the RNA extraction from PAO1 pRsmA and PAO1 pRsmN.**

The RsmA- or RsmN-specific RNAs were purified using a His-tagged RsmA or RsmN respectively immobilised on a Ni-NTA column. The pellet was

resuspended in 6 ml of lysis buffer (1 mg/ml lysozyme) and 5 µl Turbo DNase (RNase-free, 10 U) on ice for 1 h. The sample underwent sonication on ice (15 × 15 s with 15s of cooling in between). The lysate was then drawn through a syringe with needle 10 times and centrifuged at 10,000 rpm (10,285 g) in a Beckman Avanti 30 centrifuge, rotor C0650 at 4 °C for 90 min. The supernatant was added to 2 ml Ni-NTA suspension in a column which was sealed and equilibrated for 1 h at 4 °C. The flowthrough was collected and the column washed consecutively with lysis buffer, water, 1 M NaCl, water, 0.5 % Triton X-100 and water (10 ml of each, repeated 4 times). The samples were eluted with 10 x 500 µl elution buffer (1 M imidazole).

The control RNAs for the experiments were sampled just before the bulk of the culture was centrifuged after induction for 4 - 6 h (Figure 2.5). At least two samples were taken from identical cultures which were subsequently purified as in section 2.8.6.3.

#### 2.8.8.4.4 RNA purification and recovery

RNA was purified by phenol:chloroform extraction (overlaid with citrate buffer at pH 4.5). phenol:chloroform was added in 1:1 volume to the sample and the mixture was vortexed thoroughly, centrifuged for 30 min, 8,000 rpm 4 °C and the upper aqueous layer extracted using a Pasteur pipette. Extraction was repeated twice and then the sample was once extracted with chloroform only.

The RNA was precipitated by the addition of 2.5 vol of 100 % cold EtOH with 0.1 vol of NaOAc (3 M, pH 5.2), overnight at -20 °C. Then the sample was

centrifuged at 4 °C, 13 k rpm for 30 min. The supernatant was quickly removed and the pellet washed with 70 % v/v EtOH, to remove any residual salt. The centrifugation was repeated, the supernatant removed and the tube left to dry at 37 °C for 15 min. The pellet was resuspended in 500 µl nuclease free H<sub>2</sub>O.

The sample was subjected to an additional DNase digestion (20 µl 10 × Buffer, 20 µl of Turbo DNase) at 37 °C for 15 min with shaking to ensure the absence of DNA in the samples. The RNA was recovered using the RNeasy Mini Elute cleanup kit, eluting in 16 µl of nuclease-free H<sub>2</sub>O. The RNA samples were stored at -80 °C.

### **2.8.9 Determination of bioluminescence and growth using a microtitre well plate assay**

To measure bioluminescence throughout growth, light levels and OD<sub>600nm</sub> were monitored in 96-well microtitre plates using the Anthos LUCY1 combined photometer/ luminometer controlled by the Stingray software (Dazdaq). O/N cultures were diluted to a starting OD<sub>600</sub> 0.01 in LB broth, with antibiotics where appropriate, in a total volume of 200 µl. The assay was performed at 37 °C. The program measures OD<sub>600</sub> and luminescence from the wells every 30 min for 24 h. Readings were analysed using Microsoft Excel.

### **2.8.10 *rsmA/N* complementation assays**

Analysis of swarming, lipase and pyocyanin production in *P. aeruginosa* PAO1 or the *rsmA* mutant, PAZH13 carrying various plasmids containing *rsmN* and modified variants was performed as previously described (Pessi et

al., 2001, Heurlier et al., 2004). The ability of these plasmids to complement the *csrA* mutation in the *E. coli* strain TR1-5 by repressing glycogen overproduction (Romeo et al., 1993) was also assayed. When required, 1 mM IPTG was added to cultures to induce expression and the empty expression vector pME6032 was used as a negative control.

#### 2.8.10.1 Swarming motility assays

Swarming motility of bacterial strains was assessed by adapting previously published methods (Rashid and Kornberg, 2000). Briefly, a 5 µl aliquot of an overnight culture of *P. aeruginosa* was inoculated onto the surface of a swarm plate (section 2.6.4: 0.5 % (w/v) Bacto agar (Difco), 0.8 % (w/v) Nutrient broth No. 2 (Oxoid) and 0.5 % (w/v) D-glucose (Sigma) and incubated overnight at 37 °C. The ability to swarm was assessed by the distance of swarming from the central inoculation site.

#### 2.8.10.2 Pyocyanin assay

Pyocyanin levels were measured according to a previously published method (Essar et al., 1990). Briefly, overnight cultures were standardised to OD<sub>600nm</sub> 1.0 and subcultured into pyocyanin medium (section 2.6.6: 4 g D/L-alanine, 9.2 ml glycerol 87 % (v/v), 0.056 g K<sub>2</sub>HPO<sub>4</sub>, 5.68 g Na<sub>2</sub>SO<sub>4</sub>, 0.04 g citric acid, pH 7.0 in a total of 388 ml H<sub>2</sub>O + 8 ml MgCl<sub>2</sub>·6H<sub>2</sub>O (2.3 g/10 ml) + 4 ml FeCl<sub>3</sub> (0.06 g/10 ml) in a total volume of 20 ml and incubated for 16-24 h at 37 °C, with shaking. To 5 ml of culture, 3 ml of chloroform were added, mixed well, and the tubes centrifuged for 10 min at 3,000 rpm, after which 2 ml of the

chloroform phase were transferred to a tube containing 1.5 ml HCl 0.2 M and mixed well. After separation, the OD<sub>520</sub> of the HCl aqueous phase was measured. The amount of pyocyanin produced, expressed as µg of pyocyanin produced per ml of culture per OD<sub>600</sub> unit, was calculated using Equation 2.1:

$$\text{pyocyanin}(\mu\text{g per ml per OD}_{600}) = \frac{\text{OD}_{520} \times 1.5 \times 0.66 \times 17.072}{\text{OD}_{600}}$$

**Equation 2.1: Calculation of Pyocyanin concentration.**

Where the factor of 1.5 corresponds to the volume of HCl used (ml), the 0.66 deriving from the use of only 2 of the 3 ml of chloroform extract, and the 17.072 being a constant derived from the extinction coefficient of pyocyanin.

#### 2.8.10.3 Kornberg assay

Glycogen overproduction (Romeo et al., 1993) has been assayed in the *E. coli* strain TR1-5 with various plasmids. The relevant TR1-5 strains were streaked onto Kornberg media and grown o/n. Colonies were then stained with iodine stain (0.1 M I<sub>2</sub>, 0.03 M ICl). Glycogen shows as a dark brown colouration.

#### 2.8.10.4 Elastase assay

The elastin congo-red assay was used to quantify elastase production in *P. aeruginosa* strains complemented by *rsmN* and its variants (Caballero et al., 2001, Klinger, 1983).

For each strain (performed in triplicate), 1 ml of overnight culture was centrifuged for 10 min at 13,000 rpm. 100 µl of the supernatant was transferred to a new 2 ml eppendorf containing 1 ml of the buffer (100 mM Tris, 1 mM CaCl<sub>2</sub>, pH 7.5) and 5 mg of elastin-Congo red (insoluble). The samples were incubated at 37 °C with shaking. The reaction was stopped by the addition of 100 µl 120 mM EDTA after 2 h.

The samples were centrifuged and 1 ml of supernatant was transferred to a plastic cuvette. The absorbance at 495 nm was recorded.

For both the elastase and protease quantitative assays, when reading the absorbance a blank un-inoculated growth medium control is required and subtracted from the wild type absorbance. This will be either LB or PTSB depending on which broth was used for the overnight cultures. The control was treated as the rest of the samples by adding 100 µl to the relevant reagent.

#### 2.8.10.5 Exoprotease assay

This assay was used to quantify the levels of exoprotease in a *P. aeruginosa* strains complemented by *rsmN* and its variants (Swift et al., 1999, Iversen and Jørgensen, 1995).

For each strain (performed in triplicate), 1 ml of overnight culture was centrifuged for 10 min at 13,000 rpm. 100 µl of the supernatant was transferred to a new 2 ml eppendorf containing 1 ml of the buffer (100 mM Tris, 1 mM CaCl<sub>2</sub>, pH 7.5) and 5 mg of azocasein (soluble). The samples were incubated at 37 °C with shaking. The reaction was stopped by the addition of 500 µl 10 % TCA after 15 min.

The samples were centrifuged and 1 ml of supernatant was transferred to a plastic cuvette. The absorbance at 400 nm was recorded. In some cases the supernatant had to be diluted before measuring the  $A_{400\text{nm}}$ .

#### 2.8.10.6 Skimmed milk protease assay

The skimmed milk assay is a qualitative assay that enables the comparison of samples by giving a visual result. The amount of protease produced by a particular sample corresponds to the translucent zone of proteolysis created around the inoculation site (King et al., 1954).

As described in section 2.6.3, 1.4 ml of 1M  $\text{MgSO}_4$  and 20 ml of 50 % skimmed milk solution were added to 180 ml melted King's B medium (20 g/l proteose peptone No. 3 (Difco), 10 g/l glycerol, 1.5 g/l  $\text{K}_2\text{HPO}_4 \cdot 3\text{H}_2\text{O}$  and 17 g/l bacto agar (Difco) with a final pH 7.2 - 7.4) and 25 ml plates were poured and left to set before moving. They were dried for 30 min in a room temperature ventilated cabinet before use. The plates were inoculated using 2 - 5  $\mu\text{l}$  of overnight culture of the relevant strains and left overnight at 37°C without inverting the plates. All experiments were performed in triplicate.

#### 2.8.10.7 Transformation efficiency–restriction assay

The plasmid pME6001 was extracted from either *E. coli* DH5 $\alpha$  or *P. aeruginosa* PAO1 using standard protocols as described in section 2.8.3.2. 50 ng of the appropriate plasmid preparation was used to transform 100  $\mu\text{l}$  of competent cells produced by  $\text{CaCl}_2$  treatment (section 2.8.1.5). Although this method does not produce the maximum efficiency of transformation, the

results are substantially more constant and reproducible than electroporation. After incubation the number of colonies was counted on LB plates containing gentamicin (300 µg/ml) and the transformation efficiencies, expressed as colony forming units (CFU) per µg of DNA were calculated.

### **2.8.11 Molecular modelling**

Molecular modelling was carried out using several programs briefly described below. To investigate the molecular dynamics of RsmA, the Protein Data Bank (PDB) file corresponding to the determined crystal structure (PDB accession code 1VPZ) was modified so that it could be processed by AMBER, a package of molecular simulation programs which was used to run molecular dynamics simulations (Case et al., 2005). The editor program Emacs was used to modify PDB files before reading into the xLEaP program, to prepare the molecules for simulation in AMBER by introduction of missing protons. Neutralization by chloride ions and explicit solvent was added (TIP3P Water) using a truncated octahedron salvation geometry. Parameters for the system were taken from the *parm99* force field. Molecular mechanics calculations were performed using the *sander* module of AMBER. After a preliminary energy minimisation step, molecular dynamics simulation was carried out for 2 and 5 nanoseconds.

## **2.9 PROTEIN ANALYSIS**

### **2.9.1 Electrospray ionisation mass spectrometry (ESI-MS)**

Electrospray ionisation mass spectrometry (ESI-MS) was used for determination of the mass and purity of the protein. ESI-MS was performed on

an SYNAPT™ electrospray ionisation, high definition mass spectrometry (HDMS™) system with a Triwave™ ion mobility separation cell and a quadrupole time-of-flight (qTOF) mass analyser (Waters). The machine was calibrated using horse heart myoglobin by Neil Oldham (University of Nottingham) and then altered for optimization of RsmA. Samples were injected using a 100 µl syringe (Hamilton) at 10 µl/min with a mechanically driven injector. Instrument control and initial data analysis was performed using the Masslynx™ software (Waters). Samples were dissolved in 1 ml 25 mM ammonium acetate pH 7.0.

Mass and purity of the protein samples were analysed with a capillary voltage of 2.5 kV, desolvation gas flow of 200 L/h, trap and transfer collision energy of 7 V, trap gas flow of 4.5 ml/min 1.88 mbar backing pressure.

Using the Masslynx™ software the apparent molecular mass was calculated from the mass to charge ratios recorded in the positive ion mode. Each mass to charge ratio can be used to calculate the molecular mass using Equation 2.2.

$$W = (MZ) - Z$$

**Equation 2.2: Molecular mass calculation.**

This is where W is the molecular weight, M is the measured mass to charge ratio of the ion and Z is the charge state of that ion.

### **2.9.2 Circular dichroism spectroscopy (CD)**

The CD spectra were recorded on a Pi-Star-180 Spectrophotometer (Applied Photophysics), using inbuilt software (Applied Photophysics) on an Acorn Archimedes computer. The optical system was configured with a 75 W Xenon

lamp, circular light polarizer and end mounted photomultiplier. The temperature was regulated using a RTE-300 circulating programmable water bath (Neslab) and a thermoelectric temperature controller (Melcor). A correction for the CD spectra was made for the buffer (not including temperature melts). The sample was read in a cuvette of path length 1mm. Spectra were recorded from 300 nm to 200 nm to characterize the secondary structure content in 2 - 4 nm steps and 4.0 nm entrance and exit slit widths. The absorbance readings are given in molar ellipticity (millidegrees).

### 2.9.3 UV-Vis spectroscopy

To measure the concentrations of protein samples their absorbance at 280nm was recorded and the Beer-Lambert law used to determine concentration.

$$A = \varepsilon \times c \times l$$

#### Equation 2.3: Beer-Lambert Law.

Where, A = absorbance, c = concentration, l = path length and  $\varepsilon$  = molar extinction co-efficient in Equation 2.3. The molar extinction coefficient is calculated from the content of the following residues in the protein: tryptophan ( $5690 \text{ M}^{-1} \text{ cm}^{-1}$ ) and tyrosine ( $1280 \text{ M}^{-1} \text{ cm}^{-1}$ ). For wt:  $\varepsilon = 1490 \text{ M}^{-1} \text{ cm}^{-1}$ , V40W:  $\varepsilon = 6990 \text{ M}^{-1} \text{ cm}^{-1}$  and Y48W:  $\varepsilon = 5500 \text{ M}^{-1} \text{ cm}^{-1}$ . The presence of non-protein chromophores can increase the absorbance at  $A_{280}$ . Nucleic acids strongly absorb at 260 nm and Equation 2.4 can be applied in order to give an accurate estimation of the protein content by removing the contribution to absorbance by nucleotides (Aitken and Learmonth, 1996).

$$\text{Protein (mg/ml)} = (1.55 A_{280}) - (0.76 A_{260})$$

**Equation 2.4: Protein concentration calculation.**

#### **2.9.4 Equilibrium fluorescence spectroscopy**

Equilibrium fluorescence spectroscopy was conducted in order to investigate the unfolding behaviour and thermodynamics of the tryptophan mutants. This was carried out using a Luminescence Spectrometer LS50B (Perkin Elmer) with a circulating water bath which maintained the temperature at 298 K. Two protein stock solutions were prepared each containing 1  $\mu\text{M}$  protein, 25 mM potassium phosphate at pH 7.0. Solution A did not contain GdCl, whilst solution B contained 8 M GdCl. The two solutions were mixed in the cuvette to achieve the required concentration. Exact concentrations of GdCl were calculated using an Abbe 60 hand refractometer (Bellingham & Stanley) through use of equation 2.5.

$$[\text{GdCl}] = 57.147(\Delta N) + 38.68(\Delta N)^2 - 91.60(\Delta N)^3$$

**Equation 2.5: [GdmCl] calculation**

Where  $\Delta N$  is the difference in refractive index between GdCl and water. An experiment was conducted in order to determine the correct wavelength at which to excite the protein. The emissions were recorded between 300 - 400 nm at a scan speed of 200 nm per min.

The denaturant used in the experiments in this report was guanidinium chloride, the properties of which were first observed by Greenstein (Greenstein, 1938, Greenstein, 1939). It is a good chaotropic agent which

denatures a protein by disrupting the three dimensional structure. Chaotropic agents disturb the stabilizing intra-molecular interactions of the non-covalent forces such as hydrogen bonds, van der Waal forces and hydrophobic effects.

### **2.9.5 Nuclear magnetic resonance spectroscopy (NMR)**

NMR experiments were run to confirm the structure of protein produced using an Advance™-600 MHz (1.41 field strength) NMR spectrometer (Bruker). This instrument recorded 1D proton spectra at a variety of experimental conditions of varying temperature, pH and concentration of denaturant. The solution contained H<sub>2</sub>O and 10 % deuterated solvents. Water solvent suppression was achieved using the WATERGATE pulse sequence. Guanidinium chloride suppression was achieved using a WET solvent suppression with off resonance pre-saturation using a seduce pulse sequence. All spectra were referenced internally in the proton dimension to the methyl peak of 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS). The data was processed using TOPSIN (Bruker).

Lyophilised protein was dissolved in 600 µl of NMR buffer used (50 mM NaCl, 0.6 mM K<sub>2</sub>HPO<sub>4</sub>, 0.3 mM KH<sub>2</sub>PO<sub>4</sub>, 0.02 % NaN<sub>3</sub> and 10 % D<sub>2</sub>O at pH 7.0), subjected to centrifugation at 13,000 rpm for 1 min at room temperature prior to loading in a standard 5 mm 528-PP-7 NMR tube (Wilmad).

### 3 PURIFICATION AND BIOPHYSICAL ANALYSIS OF RSMA

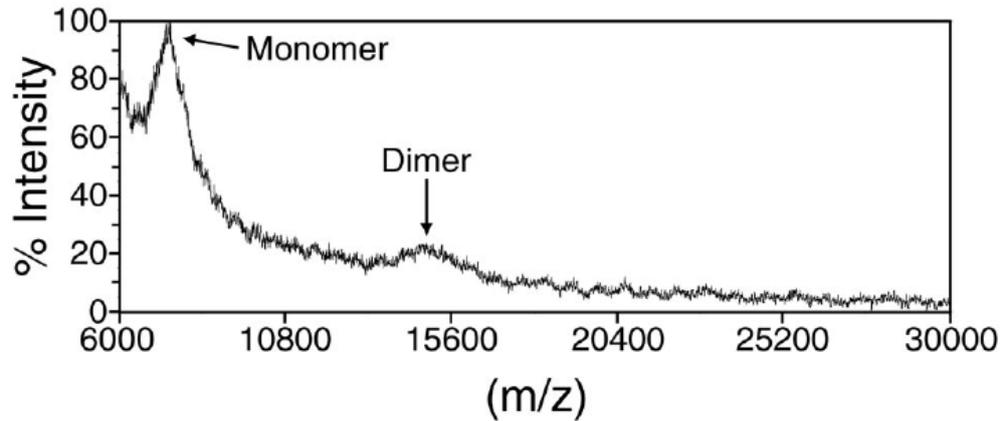
#### 3.1 INTRODUCTION

The role of the RNA-binding proteins belonging to the CsrA/RsmA family in global post-transcriptional regulation in pseudomonads, *E. coli*, *Erwinia carotovora* and other bacterial genera has been well documented (section 1.2.1.6.1). Biochemical and structural data indicates that CsrA/RsmA functions as a homodimer (Dubey et al., 2003) and it has been shown that certain residues are required for maintaining structure and functionality (Heeb et al., 2006).

Although there have been studies into the biophysical nature of CsrA and RsmE, a second RsmA homologue in *P. fluorescens*, the research on RsmA itself is minimal, possibly due to the difficulties represented by the low yields when purifying this protein on a large (1-2 L) scale. RsmA has been purified before for use in Western blots and EMSA assays, with protein yields in the ng to µg scales, far from the mg quantities required for NMR experiments. RsmA has been successfully purified for crystallization from *P. aeruginosa* using an *E. coli* based vector (pMH4) (Rife et al., 2005).

Initial studies using MALDI-TOF mass spectrometry have been performed on CsrA from *E. coli*. Using a CsrA-CsrB complex this method revealed a molecular mass of 7677.7 Da, differing by less than 3 Da from the predicted value of CsrA-H6. Fifteen cycles of Edman degradation yielded the 15 N-terminal residues identical to that of the deduced amino acid sequence of CsrA. This indicated that the polypeptide was not covalently modified, except for the deformylation of the N-terminal methionine residue (Liu et al., 1997). A

later study used glutaraldehyde cross-linked CsrA to confirm that CsrA exists as a dimer of identical subunits (Fig. 3.1).



**Figure 3.1: MALDI-TOF mass spectrum of CsrA.**

Mass spectrum of CsrA of intensity mass to charge ratio (m/z) after cross-linking with glutaraldehyde for 60 min. CsrA was confirmed to exist as a dimer after monomer and dimer peaks were identified (Dubey et al., 2003).

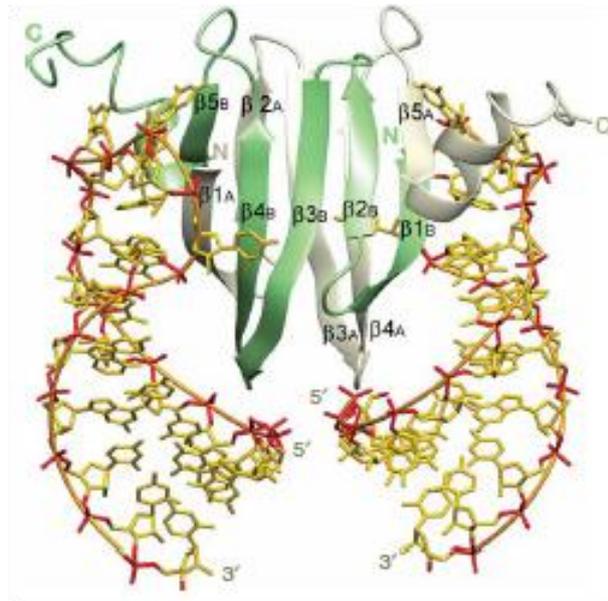
Apparent equilibrium binding constants have been obtained for CsrA-RNA complexes from radioactive bands of free and complexed species in EMSA assays using ImageQuant Software (Molecular Dynamics). No equilibrium fluorescence studies have been performed on CsrA, RsmA or any tryptophan substitution mutants. Suitable residues for tryptophan mutation can be identified by the likelihood they will undergo a change in environment upon unfolding or denaturation of the RsmA protein. The binding of a 5'-end-labeled 16-nucleotide RNA probe (containing a high affinity binding site) to CsrA mutant proteins exhibiting regulatory defects was studied (Mercante et al., 2006b), revealing the apparent binding equilibrium constants ( $K_d$ ) were increased from 10 - 150 fold in comparison with the wild type. The binding affinities of the proteins *in vitro* were roughly correlated with their ability to regulate gene expression *in vivo*. This method was also used to establish that

CsrA binds specifically to both *ycdT* and *ydeH* mRNA transcripts, genes which are controlled post-transcriptionally by CsrA and which code for GGDEF proteins(cyclic di-GMP cyclases) involved in regulating bacterial motility and attachment (Jonas et al., 2008).

Crystallographic structures have been obtained using X-ray diffraction for RsmA from *P. aeruginosa* at 2.05 Å resolution (pdb:1VPZ (Rife et al., 2005)) and RsmA from *Y. enterocolitica* 8081 at 2.5 Å resolution (pdb:2BTI (Heeb et al., 2006)) both of which were over expressed in *E. coli*. Both confirmed a homodimer as the biologically relevant form by size-exclusion chromatography, each monomer consisting of 5 consecutive antiparallel sheets followed by an alpha helix.

The solution NMR structures have been solved for CsrA from *E. coli* (pdb:1Y00 (Gutiérrez et al., 2005)), CsrA from *B. subtilis* (pdb:1T30 (Koharudin et al., Not published)) and RsmE from *P. fluorescens* (pdb:2JPP (Schubert et al., 2007)). Although sequence similarity predicted a KH domain fold ( $\beta\alpha\alpha\beta\beta\alpha$ ), which binds RNA and can function in RNA recognition (García-Mayoral et al., 2007), neither proteins are a member of that family. Interestingly, in the unbound form the solution structure for CsrA was obtained at pH 4.5, as at physiological pH, concentrations of the protein above 0.1 mM led to aggregation. RsmE was chosen for NMR studies particularly for its solubility properties. For both RNA-binding titration experiments, the NMR structure was obtained at pH 7.2-7.5. Both studies confirmed that the target RNA binds in a 1:1 ratio at 2 RNA strands per homodimer (Fig. 3.2). For CsrA, the target RNAs only bound if they were in a stable stem-loop structure (CAP leader mRNA sequence-based), unlike for RsmE when the stem loop

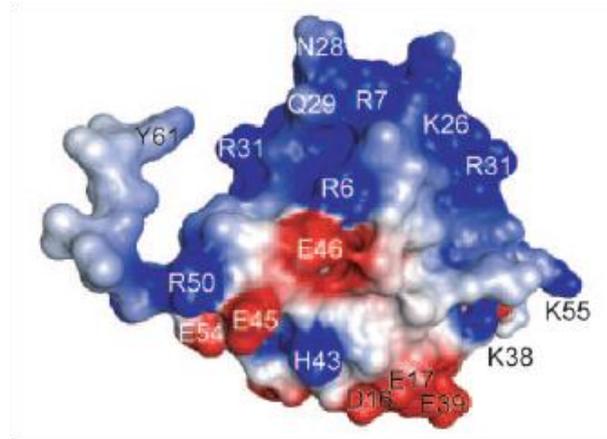
was formed upon binding with the protein (*hcnA*-based). All RNA targets contained the conserved GGA consensus element.



**Figure 3.2: NMR solution structure of the RsmE-*hcnA* RNA complex.**

Solution structure of the 2:2 complex of RsmE with the 20-nucleotide *hcnA* sequence. Protein ribbons for each monomer are shown in green and grey. Heavy atoms of the two RNAs are shown in yellow and red. An orange ribbon linking the phosphates is also shown. The complex has C2 symmetry and consists of the protein dimer with two RNA molecules bound at spatially separated sites. The RNAs are bound on a highly positively charged surface formed by the edges of the  $\beta$ -sandwich, the  $\beta 1_A/\beta 5_B$  and the  $\beta 1_B/\beta 5_A$  edge, and the region around the  $\beta 3$ - $\beta 4$  and  $\beta 4$ - $\beta 5$  loops (Schubert et al., 2007).

Gutierrez *et. al.*, concluded that CsrA is likely to need two domains in order to recognise correct transcripts to be bound, and thus regulated, and that the surface exposed residues R6, R7, E10, N28, Q29, V30 and R31 were most likely to be important in the recognition of the RNA GGA signature (Fig. 3.3).



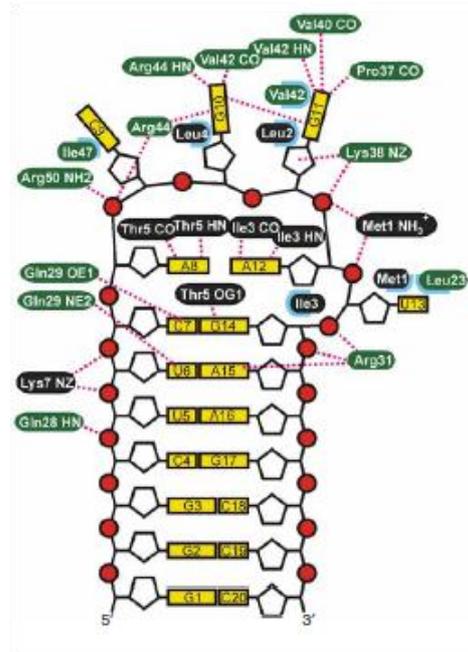
**Figure 3.3: Surface potential of the CsrA structure.**

Blue and red colours indicate positive and negative electrostatic potential respectively. The charged residues in CsrA are grouped into well-defined clusters on the protein surface where the main basic patch comprises residues R6, R7, K26, R31, and the side chain amides of N28 and Q29, defining a putative RNA-binding site. Residues E10, E45, and E46 and D16, E17, and E39 give rise to well-defined acidic patches located on the side and bottom of the CsrA molecule. Electrostatic interactions between these basic and acidic patches may explain CsrA aggregation at high concentrations. (Gutiérrez et al., 2005).

Subsequent experimental data has shown that E10 is not required for the biological function or interaction of RsmA with RsmZ and that R44 (residues 40-50 were unable to be assigned by Gutierrez *et al.*) is a key residue involved in RNA binding (Heeb et al., 2006).

Work by Schubert *et al.*, furthers the hypothesis, that by binding specifically to the 5'-<sup>A</sup>/<sub>U</sub>CANGGANG<sup>U</sup>/<sub>A</sub>-3' consensus sequence which closely matches the ideal Shine Dalgarno sequence 5'-AAGGAGGU-3' complementary to the 16S ribosomal RNA, the RsmA/CsrA family of proteins can globally regulate the expression of numerous genes (Fig. 3.4). Further work indicated that RNA targets with more than one GGA binding site for the protein have a greater affinity than a target with just one binding site. The RsmA/CsrA-RNA recognition of targets depends on at least two RNA-recognition sequences as well as their spatial arrangement and binding cooperativity.

The aim of this chapter was to establish a robust purification protocol for RsmA in order to obtain enough protein for the biophysical experiments. This was desired due to wealth of information that could be obtained regarding the protein structure and stability, with a further aim of looking at the binding of RsmA to small RNAs using ESI-MS and NMR.



**Figure 3.4: Schematic representation of intermolecular RsmE–*hcnA* interactions.**

RsmE recognises the 5'-<sup>A</sup>U<sub>C</sub>CANGGANG<sup>U</sup>/<sub>A</sub>-3' consensus sequence. Black and green: side chains and backbone of RsmE monomers A and B, respectively; magenta dashed lines: possible hydrogen bonds. cyan: hydrophobic interactions (Schubert et al., 2007).

## 3.2 RESULTS AND DISCUSSION

The aim of this chapter is to express and purify RsmA to a high degree of purity and at a quantity to allow the conduction of biophysical experiments from which information regarding the protein structure and stability can be obtained.

### 3.2.1 RsmA–Protein expression and purification

The methods used to purify RsmA or other homologues of the CsrA family vary greatly as they depend on the subsequent experimental use. The wild type protein RsmA and the tryptophan substitution mutants V40W and Y48W were successfully over-produced using the plasmids pHT::*rsmAV40W* and pHT::*rsmAY48W* by induction with 1 mM IPTG (Isopropyl- $\beta$ -D-Thiogalactopyranoside) when the OD<sub>600nm</sub> reached 0.4-0.6 (early exponential phase). Unless otherwise stated, (o/n) and over-expression cultures were grown in LB (see section 2.6.1) with ampicillin to a final concentration 100  $\mu$ g/ml). An o/n culture of the overproducing strain was used to inoculate sterile LB broth. Over-expression of RsmAY48W was problematic at this 37 °C due to the formation of inclusion bodies. Successful over-expression was obtained by conducting the growth at 37 °C until 30 min prior to induction, whereby the incubator temperature was lowered to 20 °C for the remainder of the growth period (from early exponential phase onwards for 4-6 hrs).

The original over-expression plasmid used was the pHLT::*rsmA* (S. Kuehne University of Nottingham Ph.D. Thesis), derived from the pRSETA expression vector (Invitrogen) to produce a translational fusion consisting of an

hexahistidine tag, followed by a lipoyl domain, a thrombin cleavage site (Ap<sup>R</sup>) and then the *rsmA* reading frame. Successful expression of *rsmA* was obtained with small scale experiments ( $\leq 200$  ml LB), using the method as described above prior to purification using Ni-NTA agarose resin (Qiagen). When scaling-up the protein purification procedure to larger volumes, problems were encountered with the reduced efficiency of thrombin cleavage and an increase in number of contaminants of the eluted protein (Table 3.1). Upon submission of the protein to anionic exchange (equilibration buffer: 50 mM K<sub>2</sub>HPO<sub>4</sub> pH 8; elution buffer 50 mM K<sub>2</sub>HPO<sub>4</sub>, 2 M NaCl pH 8) the cleaved protein was found to not be separating fully from the uncleaved form with the lipoyl domain. Cationic exchange was also attempted without success (equilibration buffer: 1 x AE, pH 5.2; elution buffer: 1 x AE 2 M NaCl pH 5.2). The problem was most likely due to the lipoyl domain interfering with the thrombin cleavage. The lipoyl domain coding sequence was therefore removed in the pHLT::*rsmA* construct to form the new pHT::*rsmA* plasmid (Fig. 3.5). This was performed by PCR amplification using primers to incorporate the histidine tag and thrombin cleavage site only, introducing an internal *Bam*HI site prior to the RsmA start codon (Section 2.3 Table 2.3, primers HisThrFor and HisThrRev).

A) pHLT::*rsmA*

```

NdeI           6xHistidine Tag           Lipoyl Domain
|             |                           |
|             |                           |
CATATGCACCATCACCATCACCATAGCGGCGCTTTGAATTTAAGCTGCCGGACATTGGCGAAGGCATCCACGAAGGTGAAATTTGTCAAATGG
TTTGTGAAACCGGGCGATGAAGTGAACGAAGACGATGTATTGTGCGAAGTGCAAATGACAAGGCGGTTGTCGAAATTCCTCCCGGTCAAAA
GGGAAAGTGCTTGAAATCCTCGTCCCGGAGGGAACAGTGGCAACGGTCGGGCAAACGCTCATCACGCTCGATGCGCCGGGTTATGAAAACATG
ACGAGCAGCGGCCTGGTGCCGCGGGATCCATGCTGATTCTGACTCGTGGGTCGGAGAGACCCCTGATGGTAGGTGACGACGTCACCGTGACG
GTACTGGGTGCAAAGGGAACAGGTGCGCATCGGCGTCAACGCGCCGAAGGAAGTCGCCGTACACGGGAGGAAATTTACCAGCGCATCCAG
AAGAGAAAGATCAAGAGCCAAACCATTAAGAATTC
|
EcoRI

```

B) pHT::*rsmA*

```

NdeI           6xHistidine Tag Thrombin Cleavage BamHI
|             |                 |                 |
|             |                 |                 |
CATATGCACCATCACCATCACCATCTGGTGCCGCGGGATCCGGATTCATGGTTTGTGAAACCGGGCGATGAAGTGAACGAAGACGATGTATT
GTGCGAAGTGCAAATGACAAGGCGGTTGTCGAAATTCCTCCCGGTCAAAGGAAAGTGCTTGAAATCCTCGTCCCGGAGGGAACAGTGGC
AAGGTCGGGCAAACGCTCATCACGCTCGATGCCCGGGTTATGAAAACATGACGAGCAGCGGCCTGGTGCCGCGCAATTC
|
EcoRI

```

**Figure 3.5: Sequences of plasmids for RsmA over-expression in *E. coli*.**

Where A) pHLT::*rsmA* and B) pHT::*rsmA*. Restriction sites (black); hexahistidine tag (blue), *rsmA* gene (red), lipoyl domain (green) and thrombin cleavage site (purple). The pHT construct has had the lipoyl domain (L) removed in order to facilitate purification using size-exclusion chromatography.

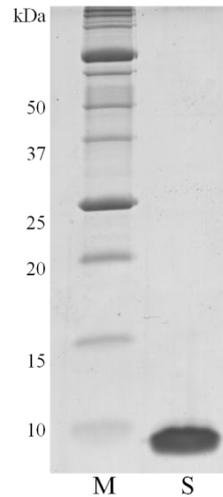
**Table 3.1: Methods and conditions used for the purification of RsmA after elution from the Ni-NTA agarose column (50 mM K<sub>2</sub>HPO<sub>4</sub>, 300 mM Imidazole, 300 mM NaCl).**

Purification methods with a superscript number correspond to protein purified by this method which was used for experimental results included in this thesis. In the caption of the experiment results the methodology type will be referred to.

<b>Small Scale Purifications (200 ml)</b>						
<b>Clone</b>	<b>Ni-NTA Elutions</b>	<b>BE</b>	<b>T/C</b>	<b>IEX (Buffer B)</b>	<b>GF (Buffer C)</b>	<b>SDS-PAGE Gel</b>
pHLT::rsmA- $\alpha$	Pure, low contaminants	Buffer A	o/n RMT	Multiple peaks of varying intensity.	n/a	Cleaved protein was eluted followed by lipoyl domain. A mixture of cleaved protein and uncleaved was then eluted followed by uncleaved protein. ⇒ <b>Inefficient cleavage – protein also truncated. Try GF to separate by size.</b>
pHLT::rsmA	High contamination	Buffer A	o/n RMT	n/a	Unable to separate lipoyl domain from protein as equal in size	Cleaved protein and lipoyl domain present in same eluted fractions. ⇒ <b>Remove lipoyl domain</b>
pHT::rsmA	High contamination	Buffer A	o/n RMT	Peak with shoulder <sup>1</sup>	n/a	Mixture of cleaved and uncleaved products, but majority cleaved. ⇒ <b>Test Thrombin cleavage conditions and try GF.</b>
<b>Thrombin Cleavage Test Conditions</b>						
<b>Clone</b>	<b>Ni-NTA Elutions</b>	<b>BE</b>	<b>T/C</b>	<b>GF (Buffer C)</b>	<b>Hep (Buffer D)</b>	<b>SDS-PAGE Gel</b>
pHT::rsmA	High contamination	Buffer A	o/n RMT	Peak with shoulder <sup>1</sup>		Inefficient cleavage ⇒ <b>Try running elution gradient over longer time to try to separate peaks.</b>
		Buffer A	o/n RMT	Two peaks visible		Inefficient cleavage ⇒ <b>Longer cleavage time</b>
			wk/end RMT	Clean peak <sup>2</sup>	Clean peak <sup>3</sup>	Ran gels of T/C after both GF and heparin. ⇒ <b>Majority of protein degraded</b>
			wk/end RMT Repeat	Clean peak at higher vol <sup>4</sup>	Protein did not bind	Inefficient cleavage ⇒ <b>Test temperatures and times for cleavage</b>
			5°C	n/a	n/a	Inefficient cleavage
			RMT	n/a	n/a	Inefficient cleavage

			37°C	n/a	n/a	Inefficient cleavage
			RMT 4 hr	n/a	n/a	Inefficient cleavage
			RMT 8 hr	n/a	n/a	Inefficient cleavage
			RT 12 hr	n/a	n/a	Inefficient cleavage ⇒ <b>Remove thrombin cleavage step as can be retained for downstream experiments (Schubert et al., 2007)</b>
<b>Purification of His<sub>6</sub>-Thr-RsmA (1L samples)</b>						
Clone	Ni-NTA Elutions	BE	T/C	GF (Buffer C)	Hep (Buffer D)	SDS-PAGE Gel
pHT::rsmA	Low contamination	n/a	n/a	Strong peak eluted high MW, series of peaks at less intensity lower MW <sup>5</sup>		RsmA present in first elution peak, but experiment was not reproducible with no protein being detected on subsequent runs ⇒ <b>Use lower pH buffer further from RsmA pI of 7.4</b>
		n/a		Buffer E – three well separated peaks <sup>6</sup>		Minimal protein eluted in first peak eluted. ⇒ <b>Use heparin to check protein is binding and to separate from contaminants.</b>
		n/a	n/a		Direct elution off column, separate peak eluted after salt gradient started <sup>7</sup>	Protein present in flowthrough off column, did not bind probably due to high salt concentration. ⇒ <b>Repeat with buffer exchange prior to column.</b>
		Buffer B	n/a		Smaller initial peak. Larger salt gradient peak, shoulder <sup>8</sup>	RsmA only in salt gradient peak and not in flowthrough. ⇒ <b>Experiment not reproducible.</b>
BE	Buffer Exchange Thrombin Cleavage	DS Buffer	Desalt	20 mM Tris, 150 mM NaCl, 2.5 mM CaCl <sub>2</sub> , pH 8.4	DS	Desalt
T/C		A Buffer	Thrombin cleavage buffer	10 mM K <sub>2</sub> HPO <sub>4</sub> pH 7 (+ 2 M NaCl for elution)	wk/end Buffer	Weekend
GF	Gel Filtration	B Buffer	Ionic Exchange equilibration buffer	50 mM K <sub>2</sub> HPO <sub>4</sub> pH 8 (+ 300 mM NaCl for elution)	D Buffer	Heparin equilibration buffer 10 mM K <sub>2</sub> HPO <sub>4</sub> pH 7 (+ 2 M NaCl for elution)
HEP	Heparin Column	C Buffer	Gel Filtration equilibration Buffer		E Buffer	Gel Filtration Equilibration Buffer low pH 50 mM NaAc 300 mM NaCl pH 4.5

The pHT::*rsmA* plasmid was constructed as well as the corresponding mutants and the proteins were shown to be well expressed (Figure 3.6).

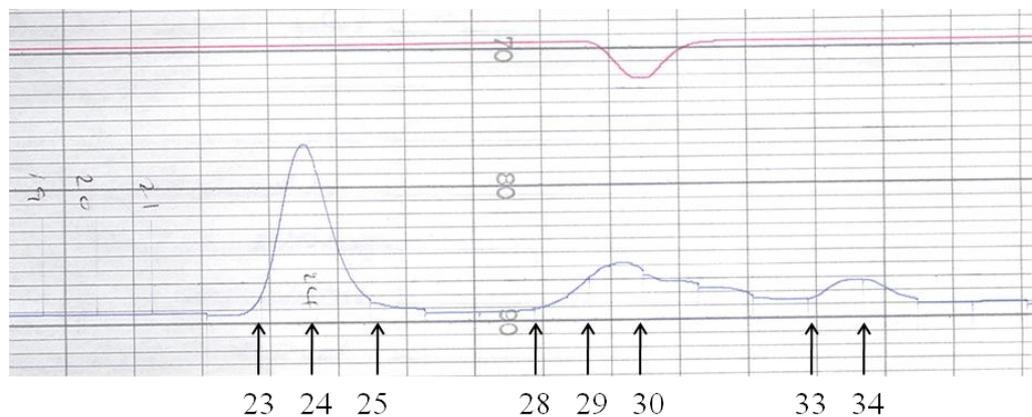


**Figure 3.6: SDS-PAGE Tricine gel of successful Ni-NTA purification of His<sub>6</sub>-Thr-RsmA.** M: Markers and S: RsmA. His-tagged RsmA was overexpressed by IPTG induction and purified using Ni-NTA agarose resin. The protein was eluted (50 mM NaH<sub>2</sub>PO buffer pH 8.0, 300 mM NaCl, 300 mM imidazole) and sampled to run on an 18 % SDS-PAGE Tricine gel.

The protein was successfully desalted into a thrombin cleavage buffer (20 mM Tris, 150 mM NaCl, 2.5 mM CaCl<sub>2</sub>, pH 8.4) after difficulties desalting into water. Little protein was eluted from the anionic exchange column most likely due to high salt concentration. Loading using a 50 ml Superloop (GE Healthcare) improved the yield especially when using on the heparin column (HiTrap<sup>TM</sup>, Amersham Pharmacia, section 2.8.7.9) with buffer A (50 mM K<sub>2</sub>HPO<sub>4</sub> pH 7 (buffer B: A+2 M NaCl). This was due to the lowering of the initial sample salt concentration to allow binding of the protein to the column. Cationic exchange was attempted using a HiTrap<sup>TM</sup> SP 5 ml High Performance column, but the protein did not bind at the lower pH (Buffer A is 1 x TAE (Tris -Acetate-EDTA pH 4.0) and Buffer B: A+2 M NaCl (section 2.8.7.8)).

### 3.2.1.1 Thrombin cleavage - Gel filtration chromatography

Due to the removal of the lipoyl domain and after the lack of success using ion exchange chromatography, it was attempted to separate His<sub>6</sub>-Thr-RsmA from its thrombin-cleaved form using gel filtration (GF or size-exclusion chromatography, section 2.8.7.10). The column used was a high load Superdex™ 200 10/300 GL (Amersham Biosciences) using Buffer A: 50 mM K<sub>2</sub>HPO<sub>4</sub> pH 8 and Buffer B: A + 300 mM NaCl for elution.



**Figure 3.7: Gel filtration trace of cleaved His<sub>6</sub>-Thr-RsmAY48W.**

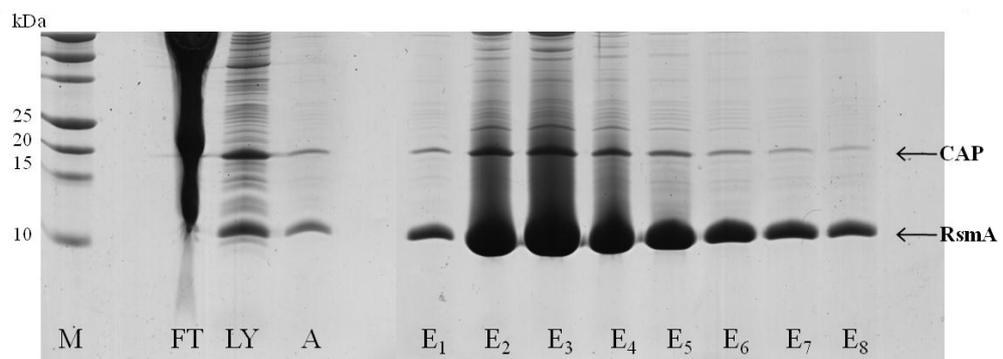
The protein was eluted into 10 mM K<sub>2</sub>HPO<sub>4</sub> 300 mM NaCl with fractions collected labelled along the horizontal axis, measured in time. The vertical axis is an arbitrary measure of intensity in set at 0.5 absorbance units' full scale (AUFS).

A sample of His<sub>6</sub>-Thr-RsmAY48W which had previously been buffer-exchanged into thrombin cleavage buffer (20 mM Tris, 150 mM NaCl, 2.5 mM CaCl<sub>2</sub>, pH 8.4) and allowed to undergo cleavage at 37 °C overnight, was freeze-dried (~ 12 h) and re-suspended in 4 ml. Prior to injection upon a gel filtration column, all samples were sterile-filtered through 0.22 µm filters. The fractions were collected (Figure 3.7) and run on a Tricine SDS-PAGE gel. The protein was eluted in fractions 22 to 26, the time of elution representative of a dimer, however bands also appeared in fractions 28 to 32. As the protein

appears to be eluting at a higher elution volume, this possibly means that some degradation had occurred. An SDS-PAGE gel of thrombin cleavage products confirmed that this was very inefficient (data not shown). Once the protein was desalted into water, the samples were run on an SDS-PAGE gel. His<sub>6</sub>-Thr-RsmA was present along with contaminants.

### 3.2.1.2 Major contaminant in Ni-NTA purification of RsmA

Major contaminants appeared to be retained after the washing stages on the Ni-NTA agarose column, prior subjecting sample to SEC, during purification of the full length His<sub>6</sub>-Thr-RsmA protein at 20 and 25 kDa (Figure 3.8). Elutions (E<sub>1</sub>-E<sub>8</sub>) are the eluted fractions from the Ni-NTA column after the wash stages, where the elution buffer is 50 mM K<sub>2</sub>PO<sub>4</sub>, 300 mM imidazole, 300 mM NaCl at pH 8.0.



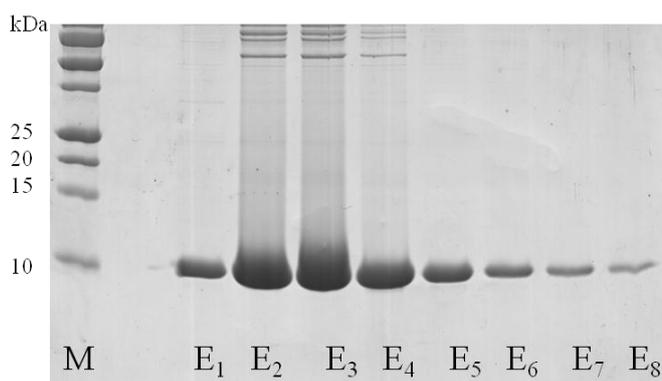
**Figure 3.8: SDS-PAGE tricine gel of contaminants in Ni-NTA purification.**

18 % SDS-PAGE Tricine gel of His<sub>6</sub>-Thr-RsmAY48W where M:Marker, FT: Flowthrough, LY: Lysis buffer (10 mM imidazole, 50 mM K<sub>2</sub>HPO<sub>4</sub>, 300 mM NaCl pH 8), A:20 mM Imidazole, and E<sub>1</sub>-E<sub>8</sub>:Elutions (50 mM K<sub>2</sub>HPO<sub>4</sub>, 300 mM imidazole, 300 mM NaCl pH 8). CAP = catabolite gene activator protein.

To identify the nature of this contamination, the relevant bands from the gel were sent for peptide mass fingerprinting analysis. The major contaminant was

identified as the catabolite gene activator protein (CAP) by peptide mass fingerprinting (Section 2.8.7.19.4). When RsmA was overexpressed in *E. coli* C41 cells, the CAP protein (23.5 kDa) co-purified with RsmA on the Ni-NTA agarose column. RsmA in *P. aeruginosa* shares 92 % identity with the analogous CsrA protein in *E. coli* (Liu and Romeo, 1997, Romeo, 1998, Ahmer, 2004). It has previously been found that CsrA regulates CAP through an interaction with CAP's mRNA (Gutiérrez et al., 2005). Due to high sequence similarity these proteins could be binding to each other or to a mutual partner, possibly RNA, DNA, lipid or another protein (Mulcahy et al., 2006).

The stage to remove this protein must be during the Ni-NTA purification or otherwise it will remain as a contaminant during subsequent steps. Various methods were used to try to remove this contaminant (Table 3.2), after which pure protein was obtained from the Ni-NTA purification as revealed by SDS-PAGE analysis (Fig. 3.9).



**Figure 3.9: Contaminant removal gel of Ni-NTA purification.**

18 % SDS-PAGE Tricine gel of His<sub>6</sub>-Thr-RsmAV40W, clearly demonstrating no contaminants at 20 – 25 kDa, where M:Marker and E<sub>1</sub>-E<sub>6</sub>:Elutions in 50 mM K<sub>2</sub>HP0<sub>4</sub>, 300 mM NaCl, 300 mM imidazole pH 8.

In order to remove CAP, a new purification protocol was implemented replacing the Ni-NTA agarose with HisPur™ cobalt resin. This resin was

chosen because although Ni<sup>2+</sup> chelate resins achieve high protein yields, the purity can be lower, requiring further optimization of wash and elution steps. According to the manufacturers, upon comparison with Ni-NTA, cobalt gives a good protein yield and purity with less need for further optimization (Thermo Fisher Scientific, Rockford, IL)(Postis et al., 2008). Column washes used (final concentrations) included 2 M NaCl to disrupt any contaminants involved in electrostatic interactions with His<sub>6</sub>-Thr-RsmA and 1 % (w/w in ddH<sub>2</sub>O) Triton X-100 which is a non-ionic surfactant used to disrupt non-ionic interactions. This protocol combined with greater wash volumes succeeded in disrupting the binding of the CAP contaminant and removing it in a wash step prior to the elution of His<sub>6</sub>-Thr-RsmA. These washes also improved purification when using Ni-NTA as well as the HisPur™ resin.

When His<sub>6</sub>-Thr-RsmA was applied onto the gel filtration column (Superdex™ 200 10/300 GL, Amersham Biosciences) eluting in a buffer at pH 7.0, no protein peak appeared to be eluted (Buffer A: 50 mM K<sub>2</sub>HPO<sub>4</sub> pH 8 and Buffer B: A + 300 mM NaCl for elution). As previous work suggested, the next step was elution in a pH 4.5 buffer which successfully eluted His<sub>6</sub>-Thr-RsmA (A: 50mM NaAc pH 4.5, B: A + 300 mM NaCl). Lowering the pH increases the number of protonated residues in His<sub>6</sub>-Thr-RsmA and as the state of ionization changes, the ionic bonds which determine the 3D shape and structure of the protein can be altered. This disrupted the electrostatic interactions with the 1 M imidazole eluent which resulted in successful purification(Hart et al., 2002). Although all the the purification methods used had problems with reproducibility, this method would be the one selected for further purification work.

<b>Ni-NTA Agarose</b>		RsmA Eluted	Contaminants		GF (Buffer C unless stated)	Superloop	Hep	Gel
Increase [Imidazole]	10 mM	yes	no effect	Increasing concentration of imidazole did remove some of contamination, but still present in elutions.		All elutions diluted 50 ml using Buffer B	Peak eluted, still shoulder	
	20 mM	yes	no effect					
	40 mM	yes	no effect					
	100 mM	yes	no effect					
Increase wash volume	20 ml each wash	minimal	most removed	Vast majority of contaminants removed. One ~ 20 kDa no effect on.	Clean peaks	yes	Peak eluted with shoulder	On a gel, both GF and HEP samples contaminated.
	0-60% 2 M NaCl Hep gradient					yes	Two distinct peaks, not separated	
	As above sample				Clean peaks	yes		contaminated
<b>HisPur Cobalt</b>								
New washes	2 M NaCl, 1% Triton X, 300 mM imidazole	yes	no	Most contaminants removed in first wash step. Lot of RsmA eluted in 300 mM imidazole wash				
[Imidazole]	50 mM	yes	no					
	100 mM	yes	no					
	150 mM	yes	no					
	200 mM	yes	no					
	1 M	yes	no		No RsmA peak, broad imidazole peak.			No RsmA on gel.
	1 M	yes	no		Peak eluted correct volume (Buffer D)			Pure RsmA. Still no T/C

**Table 3.2: Conditions used for the optimization of contaminant removal from RsmA bound to either Ni-NTA agarose or HisPur™ Cobalt columns.**

### 3.2.2 Electrospray ionization mass spectrometry

In order to verify the purity of the His<sub>6</sub>-Thr-RsmA preparation from the new purification protocol using HisPur™ cobalt resin, the samples of both purification methods were analysed using ESI-Mass Spectrometry (Fenn et al., 1990, Smith and Light-Wahl, 1993, Loo, 1997).

#### 3.2.2.1 His<sub>6</sub>-Thr-RsmA Purification Comparison

Ionised molecules are separated by the ESI-mass spectrometer on the basis of their mass ( $m$ ) to charge ratio ( $z$ ), or  $m/z$ . Therefore each peak on the spectrum corresponds to a different charge state of the protein. As the protein denatures more multiple states are visible on the spectrum and at a lower  $m/z$  due to smaller fragment size as the protons can attach to more sites.

Figure 3.10 shows the spectra of His<sub>6</sub>-Thr-RsmA purified by the Ni-NTA agarose (A) and HisPur™ cobalt (B) methods. In spectrum A, multiple charge states are present, the  $m/z$  ratio of the +5 (monomer) and +10 (dimer) charged species is 1713.26. From this the molecular weight was calculated to be 8,561 Da ( $\pm 0.19$ ) for the monomer and 17,122 Da ( $\pm 1.13$ ) for the dimer. These are higher than the predicted molecular weight of 8,530 Da and 17,060 Da for monomer and dimer respectively. The broad peaks indicate the sample still has a high salt content in relation to the protein concentration; however, the mass peaks are still visible. If a sample was fully denatured, an ESI-MS would not be expected to produce any peaks corresponding to the dimer; however, these peaks are visible. This could be due to the sample not

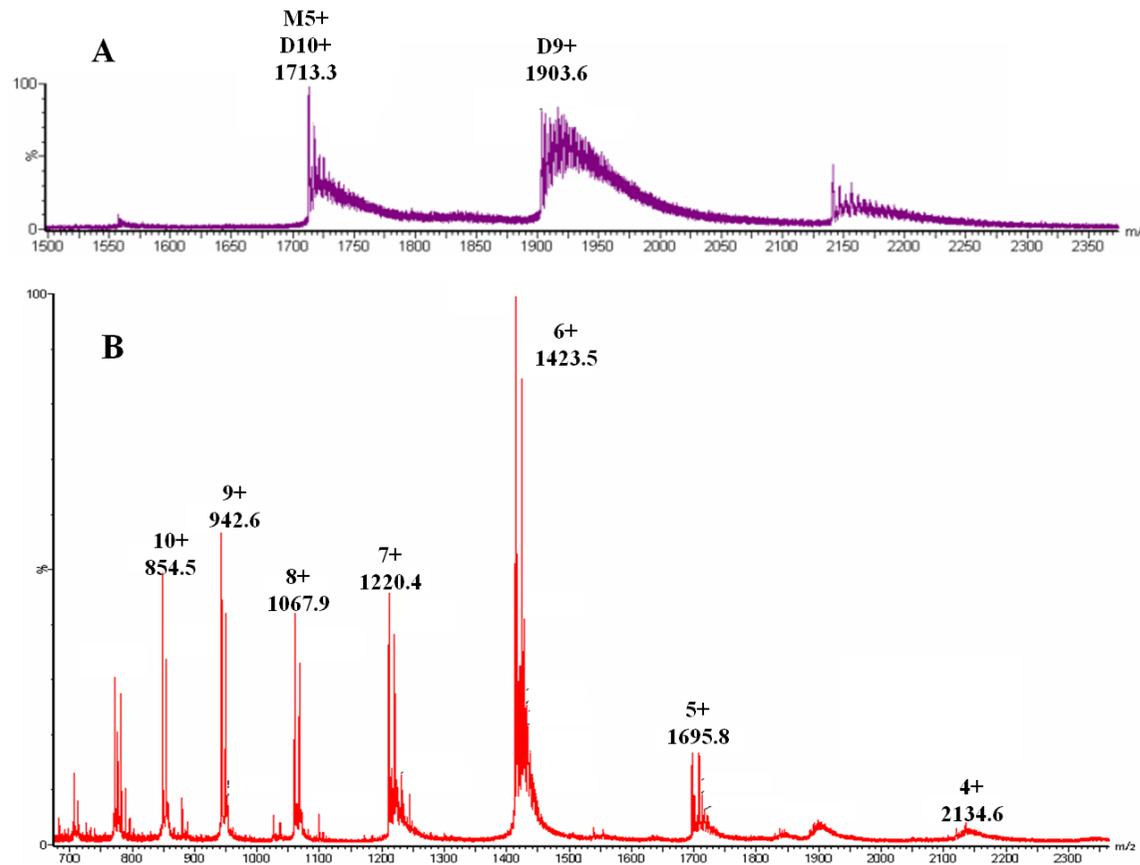
being left long enough to denature prior to the conducting the experiment, or the high salt concentration preventing complete denaturation.

In previous work, the tetramer-dimer equilibrium has been investigated (Huang et al., 2005). It was found that under denaturing conditions the instrument parameters significantly affected the ratio of detected tetramer/dimer in ESI mass spectra. The harshest conditions, including high desolvation voltages and pressures in the collision cell, led to enhanced detection of the tetramer. This was attributed to the pressure in the first pumping stage of the ESI influencing the ion abundance of large non-covalent complexes, greatly enhancing their detection. The increased pressure contributes to a shorter distance between two successive collisions so that more frequent but less energetic, less destructive collisions are generated to enhance collisional cooling of the protein assembly.

Spectrum B in Figure 3.10 is of His<sub>6</sub>-Thr-RsmA purified by the newer method utilising HisPur™ cobalt resin. Notably, the spectrum has much cleaner and sharper peaks than seen in the protein sample purified by the Ni-NTA resin. Multiple charge states were present and the m/z ratio of the +6 charged species was 1423.53. From this the molecular weight was calculated to be 8,511 Da ( $\pm 30.78$ ), which correlates very well with the predicted weight of 8,530 Da of the RsmA monomeric unit.

ESI-MS therefore confirmed that spectrum B of HisPur™ purified protein has much cleaner and sharper peaks than Ni-NTA purified protein (Spectrum A) with greater accuracy of monomer and dimer size predictions.

After sample purity determination, the effect of inserting tryptophan mutations upon protein stability needs to be ascertained by Circular Dichroism.



**Figure 3.10: ESI mass spectra of His<sub>6</sub>-Thr-RsmA.**

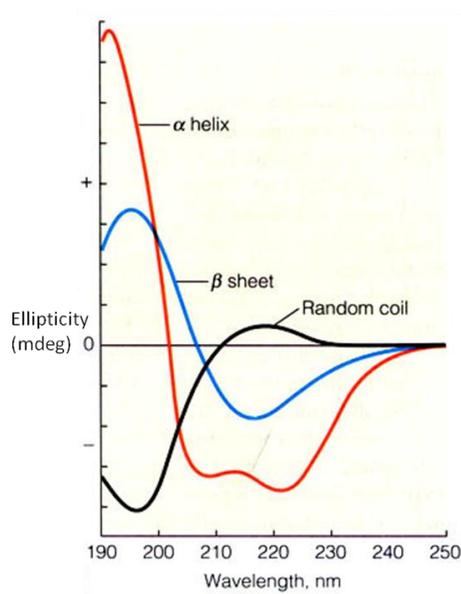
The protein samples were purified using different methodologies in spectrum A: Ni-NTA and B: HisPur™ Cobalt. Marked are the m/z ratios based on either monomer or dimer. Spectra were recorded of the protein in 25 mM ammonium acetate pH 7.0 with a capillary voltage of 2.5 kV, desolvation gas flow of 200 L/hr, trap and transfer collision energy of 7 V, trap gas flow of 4.5 ml/min 1.88 mbar backing pressure and displayed as Intensity (100 % corresponding to highest intensity peak with remaining peaks as a % relative to the 100 % peak) vs m/z. Spectrum B of HisPur™ purified protein has much cleaner and sharper peaks than Ni-NTA purified protein (Spectrum A) with greater accuracy of monomer and dimer size predictions. Protein purified method 5 (Table 3.1).

### 3.2.3 Circular dichroism analysis of RsmA

The aim of the CD experiments was to determine the effect of the tryptophan mutations on protein stability. Also if the change in CD as a function of temperature is reversible, analysis of the data may be used to determine the van't Hoff enthalpy ( $\Delta H$ ) and entropy ( $\Delta S$ ) of unfolding, the midpoint of the unfolding transition ( $T_M$ ) and the free energy ( $\Delta G$ ) of unfolding.

#### 3.2.3.1 Spectra and temperature melting of cleaved RsmA

A simple CD scan can demonstrate quickly and with a very low sample concentration ( $\mu M$ ) whether the protein present has secondary structure (Gore, 2000).



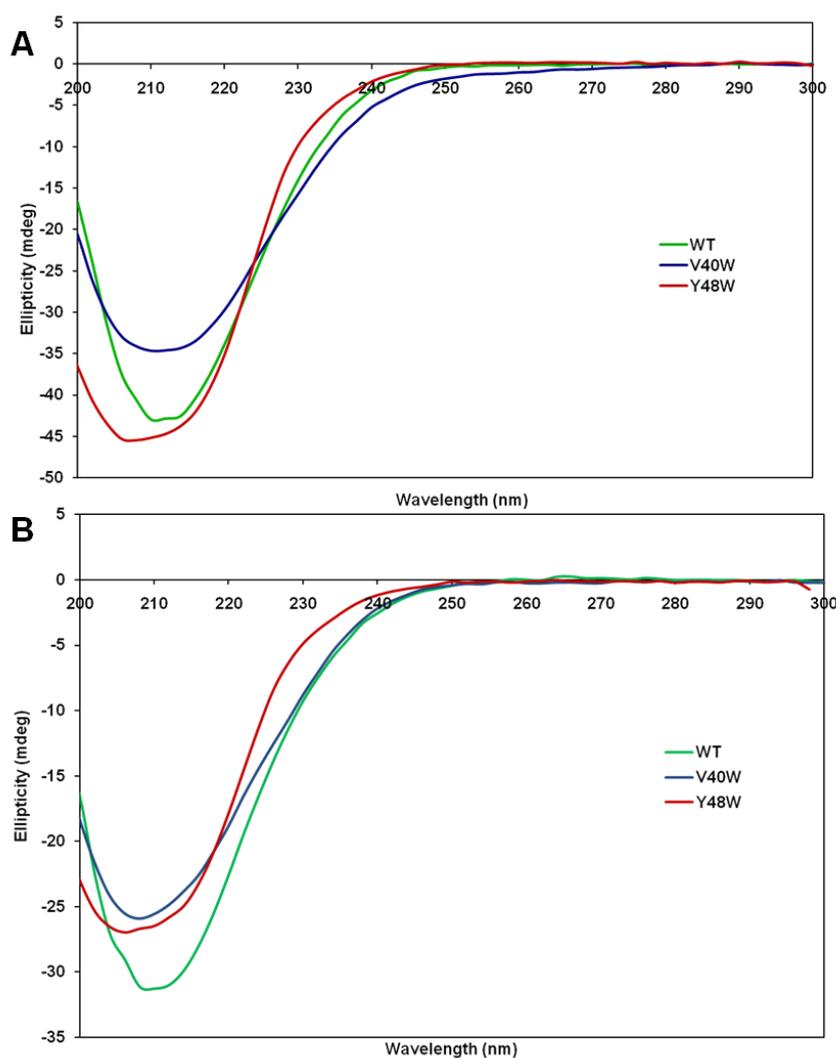
**Figure 3.11: CD spectra of pure protein secondary structures.**

Example CD spectra of ellipticity (mdeg) wavelength (nm) of pure protein structures of  $\alpha$ -helix character (red),  $\beta$ -sheet (blue) and random coil (black).

The CD spectra of  $\alpha$ -helices are characterized by a negative band with separate minima of similar magnitude at 222 nm and 208 nm (Fig 3.11). The magnitude of the CD signal can be dependent on variations in the helix, helix

length and the interactions with neighbouring structural units. The spectra of  $\beta$ -sheets generally have a negative band at approximately 216 nm and a positive band near 195 nm. Random coils have their CD maxima at similar wavelengths and of the opposite sign from those of sheets'.

CD spectra of the RsmA protein and the tryptophan substitution mutants (V40W and Y48W) are shown (Fig. 3.12) where the hexahistidine tag has been removed by thrombin cleavage together with the uncleaved proteins. This was to ensure that the additional hexahistidine tag and thrombin cleavage site which were previously removed did not affect the RsmA secondary structure.



**Figure 3.12: Comparison CD spectra of RsmA wild type and tryptophan substitution mutants cleaved and uncleaved.**

CD spectra measures ellipticity (mdeg) vs wavelength (nm). Cleaved spectra (A) and uncleaved (B) where RsmA wt (green: A) 100  $\mu$ M, B) 120  $\mu$ M), substitution mutants V40W (blue: A) 80  $\mu$ M, B) 95  $\mu$ M) and Y48W (red: A&B) 106  $\mu$ M) in 10 mM  $K_2HPO_4$  pH 7.0 at 25  $^{\circ}$ C. Similar secondary structures were observed for the wild type and tryptophan mutants, with comparable traces observed between the cleaved and uncleaved spectra. Method of purification WT:1, V40W: 2 and Y48W: 3 (Table 3.1).

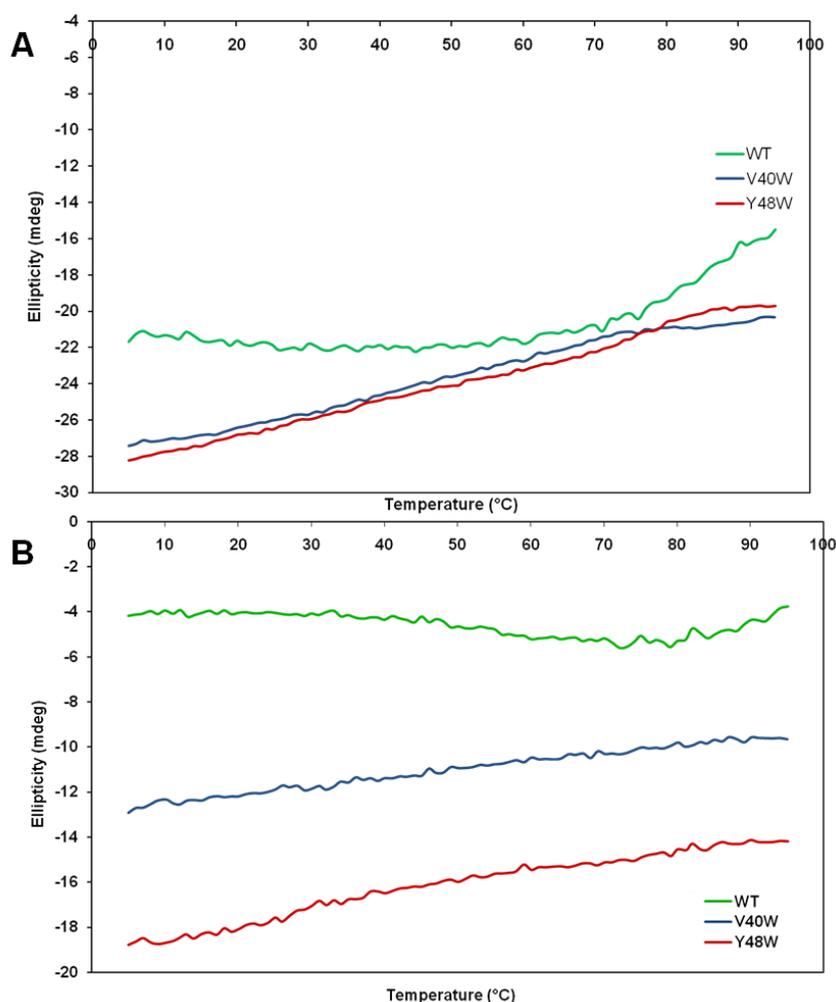
The CD traces of the cleaved RsmA wild type protein and RsmAV40W and RsmAY48W mutants were recorded at 25  $^{\circ}$ C (Fig 3.12 (A)). There is a minimum at around 210 nm that indicates the wild type protein is composed predominantly of beta sheets with very similar spectra for the mutants. There does not appear to be a clear secondary minimum at 222 nm as shown with

alpha helices, but it is important to remember that the spectrum is additive of the types of secondary structure, therefore if the alpha helical content is low it can be masked by the stronger beta sheet signal. Although of slightly differing concentrations compared to the cleaved proteins, the uncleaved His<sub>6</sub>-Thr-RsmA wild type and tryptophan mutants proteins (Fig. 3.12 (B)) overall have a very similar shape, leading to a confirmation that the structure of RsmA has been unaffected by the thrombin cleavage.

The CD temperature melts of cleaved and uncleaved RsmA wild type and the two mutants from 5 °C to 95 °C were measured at 208 nm in 10 mM potassium phosphate buffer at pH 7.0 (Fig. 3.13). The melt traces for cleaved RsmAV40W and Y48W (Fig. 3.13 (A)) show a steady gradual increase in ellipticity indicating a pre-melting transition is occurring, but no melting transition has been reached before 95 °C. The wild type spectra is similar, although it appears that when reaching the higher temperatures of 80 - 90 °C the protein could be about to start the melting transition. A complete melting transition would display a sigmoidal curve shape and if the reaction was reversible then the melting temperature would be directly related to conformational stability. However this complete melting transition did not occur, therefore these properties cannot be calculated for RsmA.

Despite the uncleaved proteins being of different concentrations which resulting in the traces being of differing ellipticity ranges (Fig. 3.13 (B)), the results are very similar to those of the cleaved proteins (Fig. 3.13 (A)). Neither the mutants nor wild type have undergone a melting transition from folded to unfolded protein. The wild type experiment again shows a slight increase in

ellipticity at higher temperature, however the instrumentation is limited to 95 °C, the maximum temperature available for experimental work.



**Figure 3.13: CD temperature melts of RsmA wild type and tryptophan substitution mutants cleaved and uncleaved.**

CD spectra measuring ellipticity (mdeg) vs temperature (°C) of cleaved (A) and uncleaved (B) proteins. RsmA wt (green): A) 100  $\mu$ M, B) 15  $\mu$ M, substitution mutants V40W (blue: A) 95  $\mu$ M, B) 34  $\mu$ M and Y48W (red: A) 106  $\mu$ M, B) 64  $\mu$ M in 10 mM  $K_2HPO_4$  pH 7.0 at 208 nm. None of the cleaved or uncleaved proteins underwent melting transitions over the temperature range examined. Method of purification WT: 7, V40W: 7 and Y48W: 5 (Table 3.1).

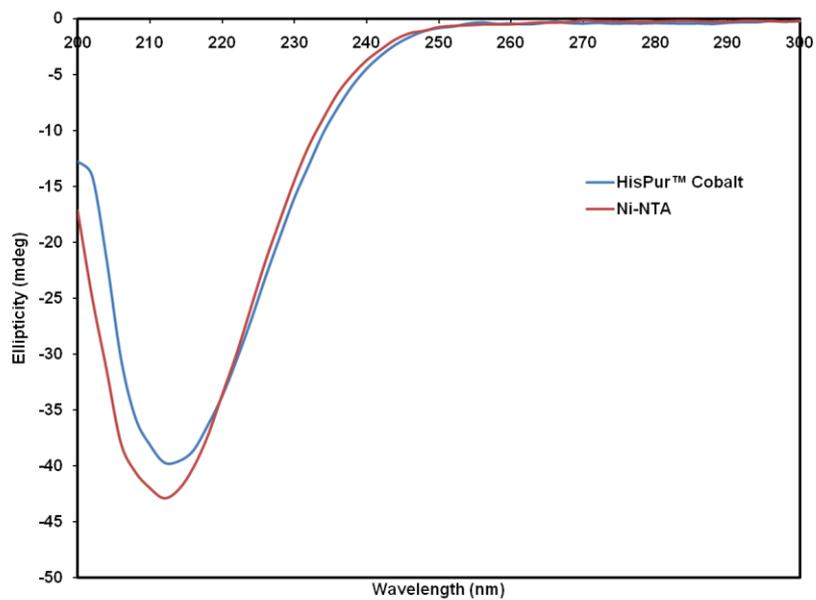
CD spectra were run on the wild type His<sub>6</sub>-Thr-RsmA before the temperature melt at 5 °C, just after the melt had completed at 95 °C and after the reverse melt at 5 °C (data not shown). A reverse melt was performed in order to slow down the refolding as much as possible to limit the likelihood of mis-folds

occurring. Both spectra pre- and post-melt are very similar indicating  $\beta$ -sheet prevalence, however the CD spectra obtained after the temperature melt appears to have decreased in intensity and the minimum has shifted from 208 to 206 nm. This could be because the protein is not undergoing a complete melting transition, the increase in temperature has disrupted some of the non-covalent bonding between the two monomers leading to less  $\beta$ -sheet character being present or due to protein aggregation effects.

Comparison CD spectra of the His<sub>6</sub>-Thr-RsmA wild type protein purified using two different purification methods, each using a different metal affinity resin were also run (Fig. 3.14). The characteristics of both spectra are identical demonstrating that His<sub>6</sub>-Thr-RsmA purified by the new cobalt resin produces the same CD spectra as that purified from the Ni-NTA agarose.

The CD experiments have confirmed that there is no observable change in structure when the hexahistidine tag is removed or a tryptophan mutant is introduced. The temperature ramps observed no melting transition, indicating a highly stable protein, with no difference in spectra when purified using different affinity columns.

An alternative method for monitoring protein stability is equilibrium unfolding. This monitors the effect on the fluorescence signal due to the change in environment of the tryptophan chromophore within the protein structure as unfolding occurs as a result of the addition of chemical denaturants.



**Figure 3.14: CD comparison spectra of purification resins.**

CD spectra measuring ellipticity (mdeg) vs wavelength (nm) of uncleaved His<sub>6</sub>-Thr-Rsma purified by HisPur™ cobalt (92 μM) and Ni-NTA agarose (100 μM) in 10 mM K<sub>2</sub>HPO<sub>4</sub> pH 7.0 at 25°C. Proteins purified by each method have comparable secondary structures. Protein purified by method 5 (Table 3.1).

### 3.2.4 Equilibrium Fluorescence

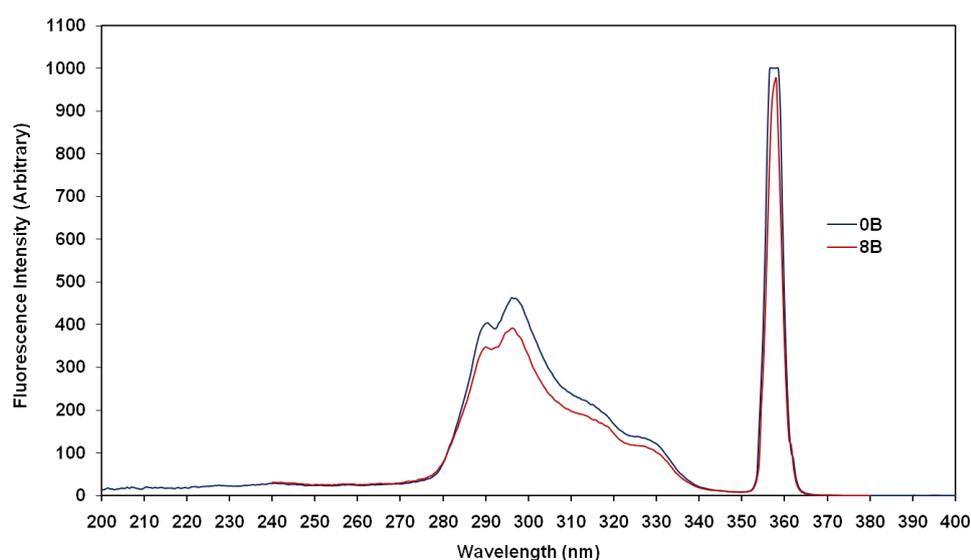
Equilibrium unfolding is the process of unfolding a protein by gradually changing its environment, for example, by changing the temperature or the addition of chemical denaturants. As the equilibrium of the sample is maintained at each step, the process is reversible. Monitoring the effect on fluorescence signal due to the change in environment of the tryptophan chromophore allows determination of the conformational stability of the molecule.

#### 3.2.4.1 RsmA tryptophan substitution mutants

Due to the absence of native tryptophan residues in RsmA, two tryptophan mutants were made by S Keuhne, RsmAV40W and RsmAY48W. Preliminary experiments were undertaken to elucidate the optimal excitation wavelength to be used for the emission spectra as the wavelength varies for different proteins. Excitation spectra were run at a variety of fixed emission wavelengths for both 0 and 8 M guanidinium chloride (GdmCl). It was found that the lower the emission wavelength ( $\lambda_{EM}$ ) used, the lower the signal intensity was observed. The maximum intensity was given by the emission wavelength when fixed at 358 nm. Across the spectrum four different signals could be observed. Two broad peaks are due to the protein and two sharper ones at higher wavelengths corresponding to the emission wavelength used (Fig. 3.15).

The difference in fluorescence between the extreme concentrations of GdmCl does not appear to be very large, (~f 80 fluorescence units), but this was still the greatest change upon comparison with the other fixed emission

wavelengths. The effect of pH was also studied from pH 4 to pH 7, with the greatest signal intensity being observed at pH 7.0. The intensity and wavelength of the maximum fluorescence emission of tryptophan is very solvent-dependent, typically maximal absorption occurring at 280 nm. However, the maximum fluorescence peak observed had a double maximum at 289 nm and 295 nm (Fig. 3.15). The fluorescence spectrum shifts to a longer wavelength as the polarity of the solvent surrounding the tryptophan residue increases. The tryptophan fluorescence may be partly quenched by two neighbouring glutamic acid residues. Tryptophan residues which are buried in the hydrophobic core of proteins can have spectra which are shifted by 10 to 20 nm compared to residues on the surface of the protein. This could explain these observations, however it is an unlikely scenario as the Y48W residue is expected to be solvent-exposed. The tryptophan could be in its own micro-environment, shielding it from the solvent.

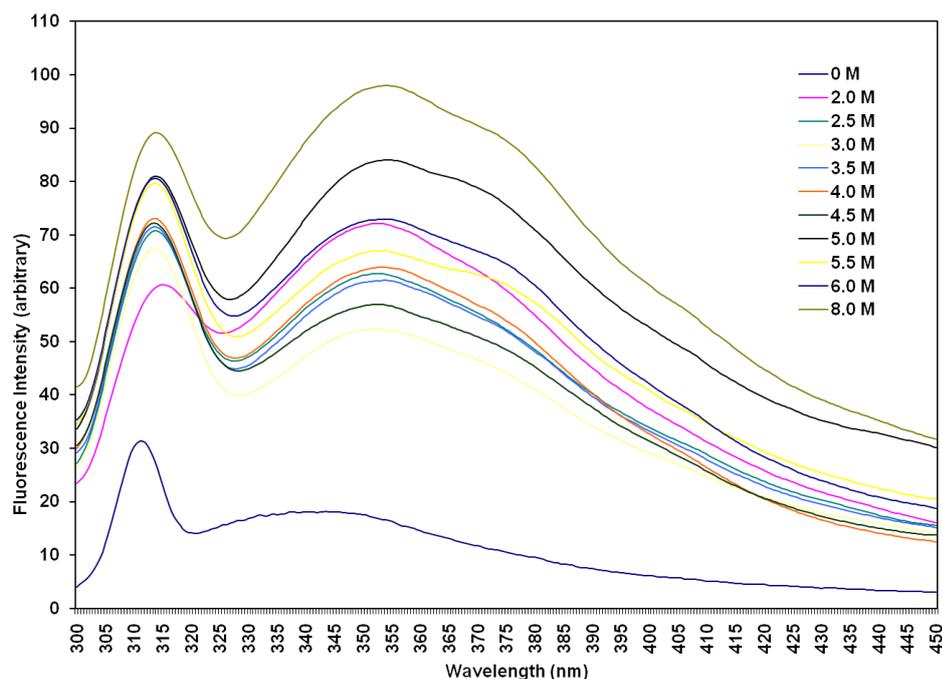


**Figure 3.15: Excitation spectra of RsmAY48W.**

Excitation spectra where fluorescence intensity (arbitrary) vs wavelength (nm). RsmAY48W (30  $\mu$ M) was measured at an emission wavelength of  $\lambda_{EM} = 358$  nm at 0 and 8 M GdmCl in 25 mM  $K_2HPO_4$  pH 7.0. Spectra were collected using a scan speed of 11, entrance and exit slits of 5.0 mm and 5 accumulated scans. Purification by method 2 (Table 3.1).

Emission spectra of His<sub>6</sub>-Thr-RsmAY48W were obtained with an excitation wavelength of 297 nm, with samples present in a variety of denaturant concentrations (Fig. 3.16). The samples were prepared separately in order to increase the equilibration time, a necessary precaution if the rate of unfolding was very slow. A continual increase in fluorescence intensity from 0 to 8 M was not found as might have been expected, but instead variations in fluorescence were observed. The same behaviour occurred with the His<sub>6</sub>-Thr-RsmAV40W mutant. Although the fluorescence trend is to increase with the concentration of denaturant, the increase is minimal, most likely indicating that upon moving from the folded to the unfolded state, only a very small change in the environment of the tryptophan is taking place. This could be due to the residues being too solvent-exposed in the native conformation to make a difference, no quenching residues residing near in space in the native form, or a combination of both.

Neither tryptophan mutant underwent a significant change in environment to enable the calculation of the dissociation equilibration constant ( $K_d$ ) for reasons described above. Therefore new candidates for tryptophan mutation need to be identified.

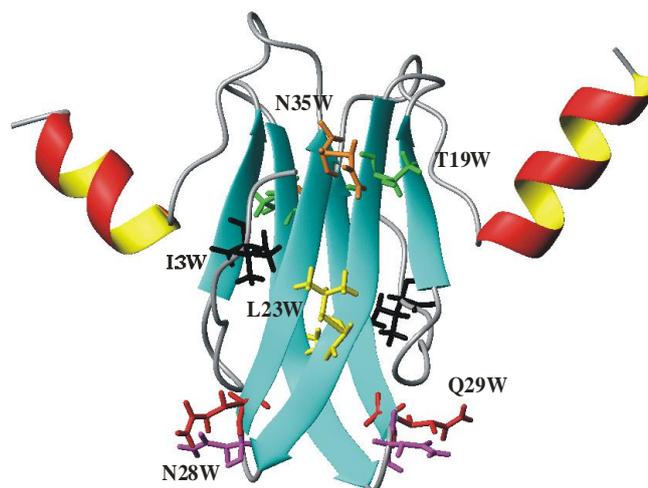


**Figure 3.16: Emission spectra of His<sub>6</sub>-Thr-RsmAY48W.**

Emission spectra with fluorescence intensity (arbitrary) vs wavelength (nm). His<sub>6</sub>-Thr-RsmAY48W (36 μM) was measured at an emission wavelength of  $\lambda_{EX} = 297$  nm at 0 and 8 M GdmCl in 25 mM K<sub>2</sub>HPO<sub>4</sub> pH 7.0. Spectra were collected using a scan speed of 11, entrance and exit slits of 5.0 mm and 5 accumulated scans. Protein purified by method 5 (Table 3.1).

#### 3.2.4.2 Prospective tryptophan substitution mutants

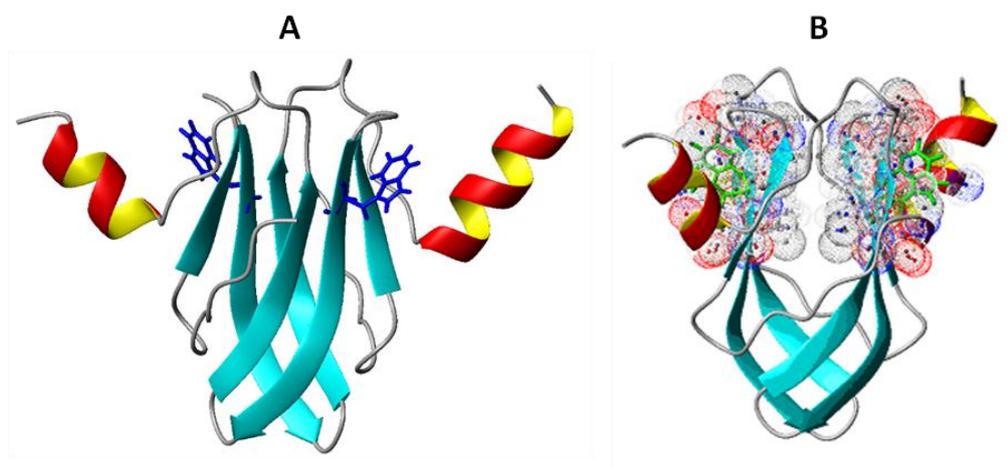
Since neither RsmA mutants V40W or Y48W undergo a change in fluorescence upon unfolding, additional tryptophan mutants were required. Using the same principle as described before, residues were chosen so that when unfolded, a change in the intensity is observed. Therefore residues that were either buried near the core of the structure, close to quenching residues, or a combination of both were chosen for substitution so that the fluorescence intensity observed would be reduced, but when unfolded by GdCl the tryptophan would become exposed, causing the fluorescence intensity to increase. Prospective tryptophan mutants were identified as I3W, T19W, L23W, N28W, Q29 and N35W (Fig. 3.17).



**Figure 3.17: Prospective tryptophan mutants chosen for site-directed mutagenesis.**

Predicted representation of the prospective tryptophan mutants using MolMol (Koradi et al., 1996), where the RsmA back bone is displayed as a ribbon, with the  $\alpha$ -helices (red and yellow) and  $\beta$ -sheets (blue). The tryptophan mutants I3W (black), T19W (green), L23W (yellow), N28W (purple), Q29W (red) and N35W (orange) are labelled.

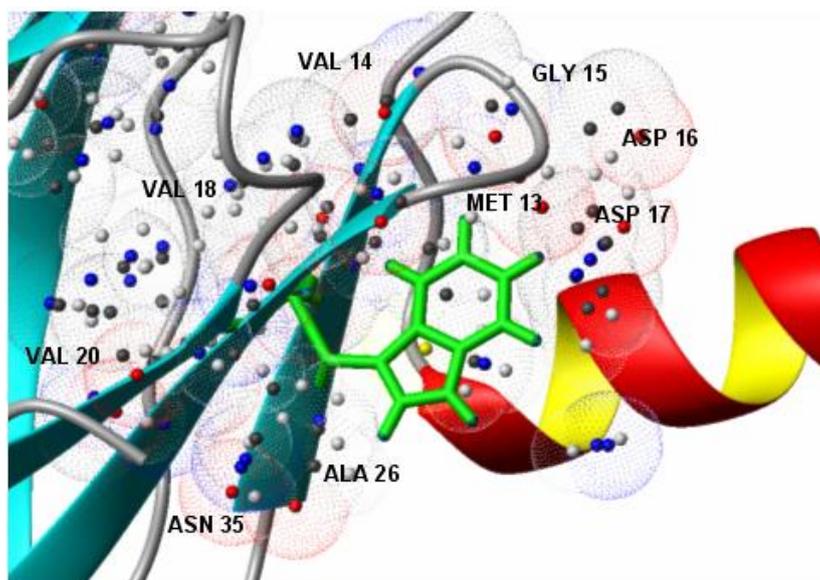
The RsmAT19W mutant is displayed below with the tryptophan residue in blue neon form (Fig. 3.18 A) and with the residues (coloured mesh) surrounding this tryptophan (green neon) within 5 Å (Fig. 3.18 B). Within these surrounding residues the oxygen (red), carbon (grey) and nitrogen (blue) atoms are identified.



**Figure 3.18: Prospective RsmAT19W mutant.**

Predicted representation of the T19W mutant using MolMol (Koradi et al., 1996), where the RsmA back bone is displayed as a ribbon, with the  $\alpha$ -helices (red and yellow) and  $\beta$ -sheets (blue). The tryptophan residue in blue neon form (A) and with defined residues 5Å surrounding the green neon tryptophan residue (B).

As demonstrated in the close up of the predicted structure of RsmAT19W (Figure 3.19), residues within 5 Å of the tryptophan include Asp-16 and Asn-35. Aspartic acid and asparagine are both side chains that quench fluorescence by excited state electron transfer. Although the aspartic acid residue would still be relatively close to the tryptophan when unfolded, in the native conformation it is much closer. The asparagine would be in a completely different environment to the tryptophan when unfolded, eliminating the effect of its quenching. Therefore this is a good candidate for mutation. This study was done for each prospective new tryptophan mutant.



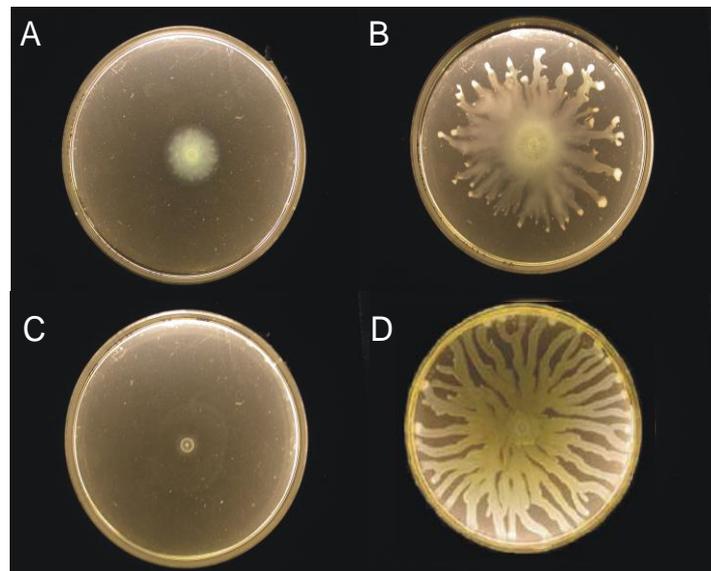
**Figure 3.19: Close up of RsmAT19W predicted structure.**

Representation of the T19W mutant using MolMol (Koradi et al., 1996), where the RsmA back bone is displayed as a ribbon, with the  $\alpha$ -helices (red and yellow) and  $\beta$ -sheets (blue). The tryptophan residue in green neon form and with defined residues 5Å surrounding the tryptophan residue, including fluorescence quenching residues Asp-16 and Asn-35.

For two of these, L23W and N35W, the mutants were constructed using site-directed mutagenesis. A phenotypic assay using swarming was carried out using the *P. aeruginosa rsmA* mutant PAZH13. Swarming was chosen to

characterise the mutants as it is positively regulated by RsmA and this assay was used to assess the biological activity of the mutants. The *rsmA* mutant PAZH13/pME6032 is deficient in swarming. Swarming activity was restored in PAZH13/pRsmA and PAZH13/pRsmAN35W (Fig. 3.20) but not PAZH13/pRsmAL23W.

Therefore the conclusion is that from the swarming assays PAZH13/pRsmAL23W does not retain biological activity, whereas PAZH13/pRsmAN35W is biologically active and could be used for future biophysical experiments.



**Figure 3.20: Swarming assays of RsmA and the RsmA tryptophan mutants, RsmAL23W and N23W.**

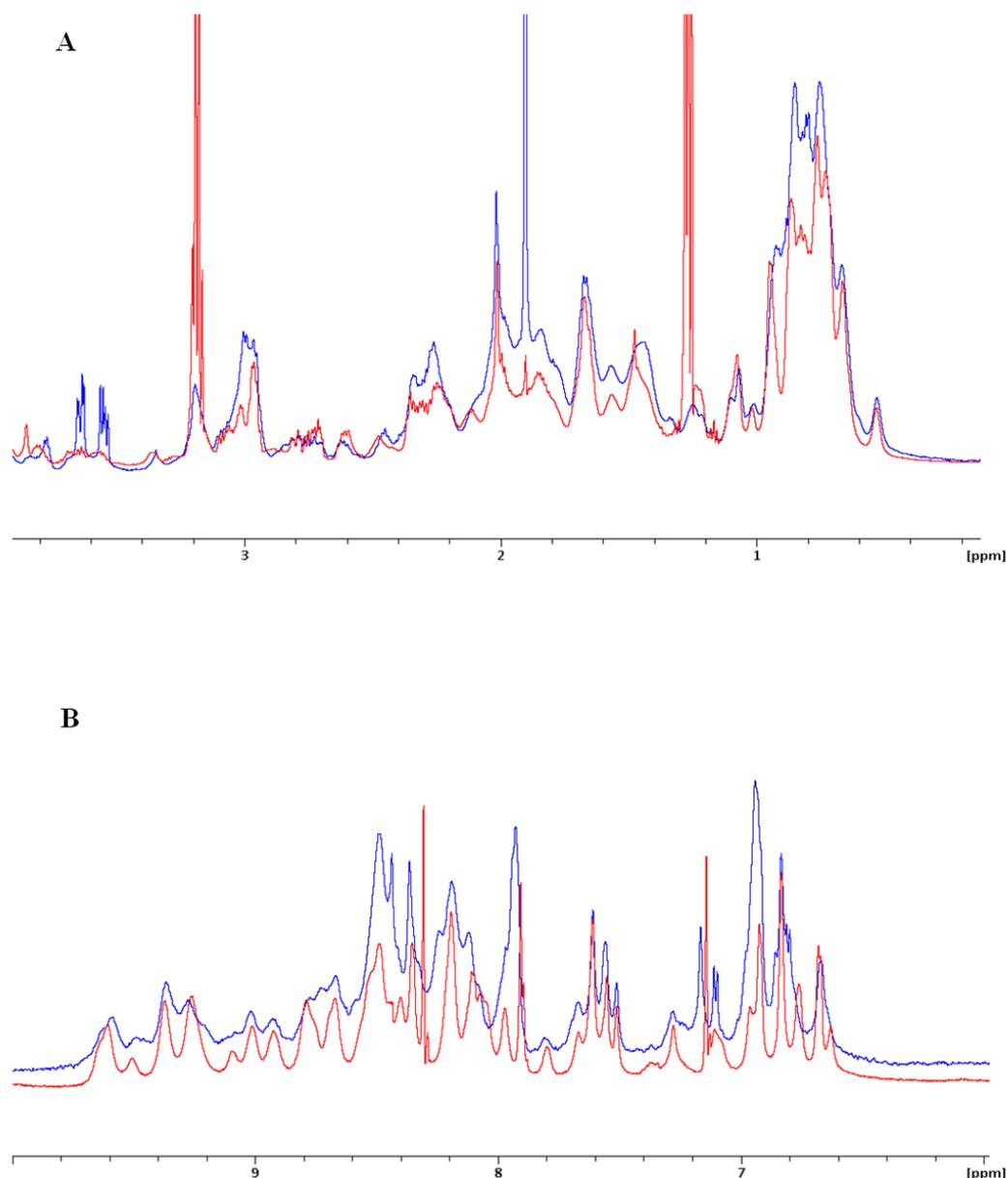
5  $\mu$ l cultures of the strains (A) PAZH13/pME6032, (B) PAZH13/pRsmA, (C) PAZH13/pRsmAL23W, (D) PAZH13/pRsmAN35W, were deposited in the middle of the swarming plates and incubated overnight at 37 °C.

### 3.2.5 Impact of temperature, denaturant and pH on the structure of RsmA using NMR analysis

NMR experiments were conducted with the aim of providing an assigned structure of RsmA, which residues are involved when binding to small RNA molecules and the effect of changing conditions (*e.g.*, temperature, pH and purification method) upon the protein structure and stability.

#### 3.2.5.1 Comparison of RsmA purified by Ni-NTA agarose and HisPur™ cobalt resin

A comparison of 1D NMR spectra of His<sub>6</sub>-Thr-RsmA wild type purified either by using Ni-NTA agarose or HisPur™ cobalt resin was performed (Fig. 3.21). Spectra A is of the amide proton region and spectra B of the methyl protons. The peaks on Ni-NTA spectrum (red) show a decrease in line width compared to the HisPur™ cobalt resin (blue) purification sample, probably due to the difference in concentration. Broadening of the peaks could also be due to a greater amount of buffer salts in Ni-NTA sample, or if the concentration caused aggregation in the HisPur™ cobalt sample. Although visible on spectra A, the comparison of the two samples is more noticeable on spectra B, where the Ni-NTA sample has well resolved, cleaner peaks. In the sample prepared with the cobalt resin, there appears to be a double peak at 3.5 and 3.7 ppm. in the spectrum which is not present in the sample prepared with the Ni-NTA agarose. This chemical shift indicates that something is attached to the histidine tag, which could easily be removed through repeating the freeze-drying process. However even with different sample concentrations, the two samples are definitely folded the same, leading to conclusion both methods produce pure proteins.



**Figure 3.21: 1D NMR proton spectra purification method comparison.**

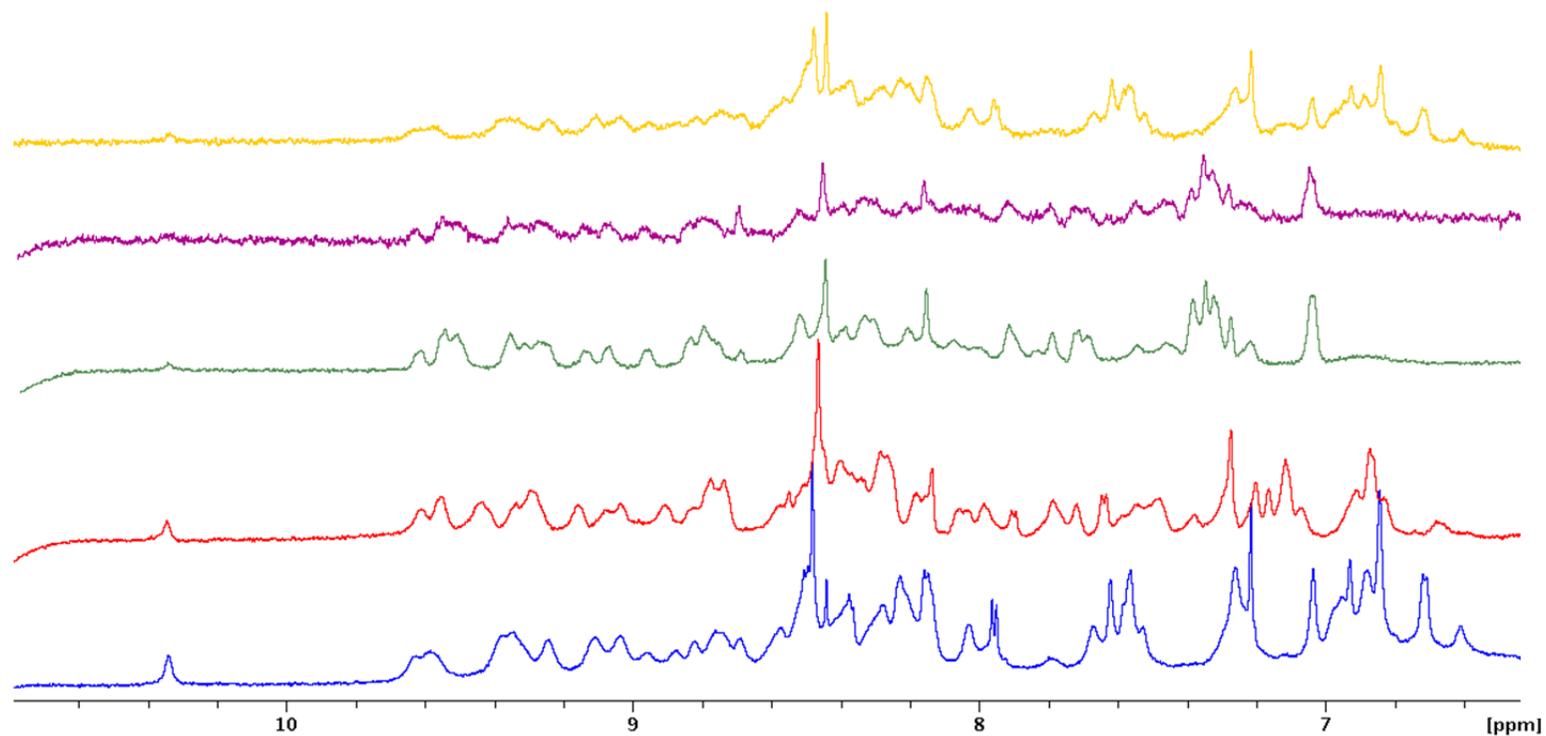
1D NMR sample buffer used contains 50 mM NaCl, 0.6 mM  $K_2HPO_4$ , 0.3 mM  $KH_2PO_4$ , 0.02 %  $NaN_3$  and 10 %  $D_2O$  at pH 7.0, using DSS as an internal standard. Purification of His<sub>6</sub>-Thr-RsmA was by Ni-NTA resin (red) at 0.41 mM protein concentration and HisPur™ cobalt resin (blue) at 0.2 mM with spectra A showing the amide proton region and spectra B for the alkyl protons. The folding is the same between both spectra, indicating both purification methods produce pure protein, however the Ni-NTA sample has well resolved, cleaner peaks. Method of purification NiNTA:1 and HisPur: 5.(Table 3.1).

### 3.2.5.2 Temperature Study

CD experiments demonstrated that RsmA is a highly stable protein up to 95 °C (section 3.2.3). Monitoring change in the fluorescence of RsmA using

the RsmAV40W and RsmAY48W mutants showed only a very small change in the environment of the tryptophan which could be due to the residues being too solvent exposed in the native conformation, no quenching residues residing near in space in the native form, or a combination of both. By use of NMR techniques, a detailed study was conducted in order to verify and understand why these results had been obtained (section 2.9.5).

The 1D NMR spectra of RsmAY48W (cleaved) recorded at a range of temperatures was analysed (Fig. 3.22). The spectra recorded at 298 K (blue) shows well distributed, sharp peaks indicating a folded protein, with the tryptophan peak present at 10.24 ppm. As the temperature increases to 323 K (red), the peaks have started to broaden out and disappear due to a loss of structure and fast exchange with the solvent caused by the increase in temperature. The temperature was further increased to 348 K (green) and 353 K (purple) where nearly all protein signal is lost. The temperature was then reduced back down to 298 K (yellow) where a minimal structure has been recovered. These data suggest that an increase in temperature causes the loss of RsmA quaternary structure. There is still the  $\beta$ -sheet character but not properly folded, with the secondary structure mostly lost by 353 K. The reaction appeared to be non-reversible. The sample would need to be exposed to higher temperatures for a more definitive answer.



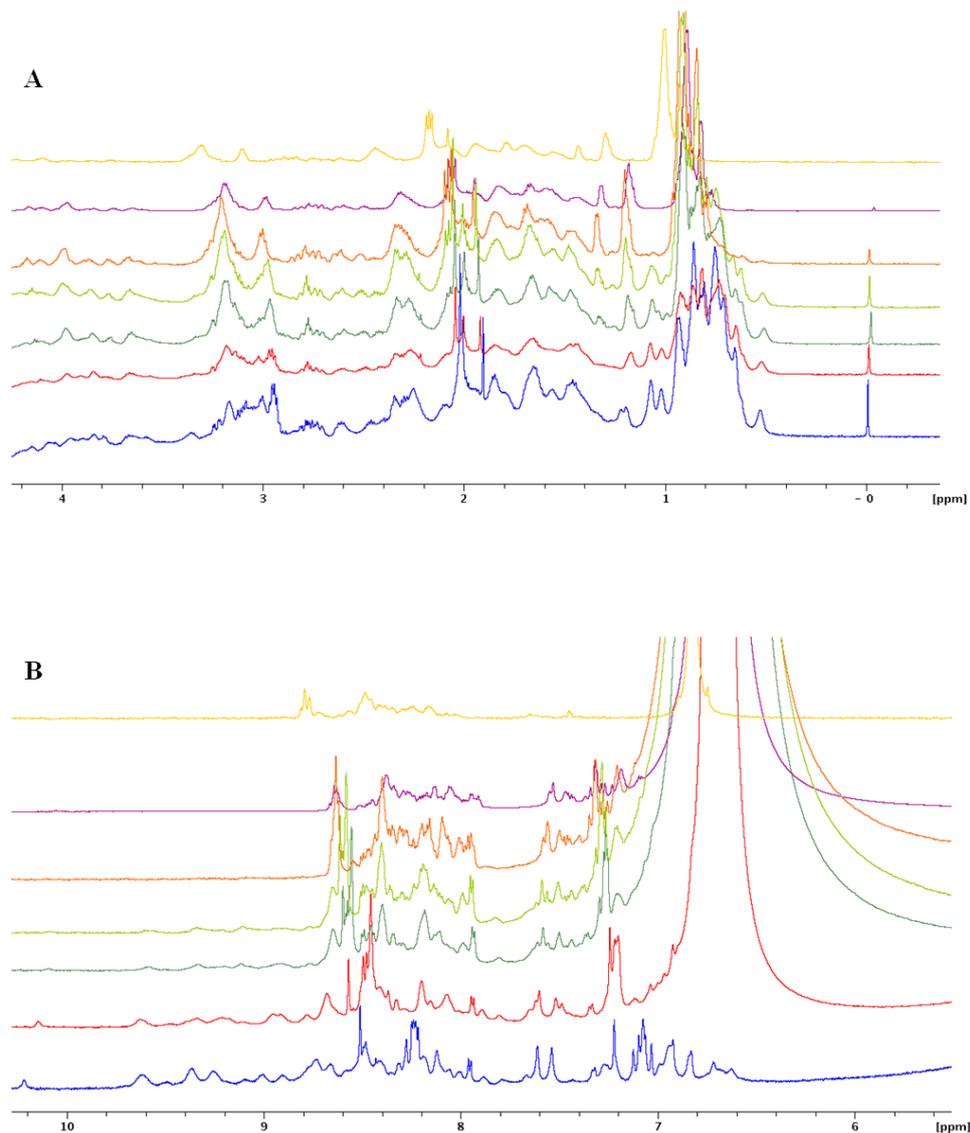
**Figure 3.22: 1D NMR spectra of the effect of temperature on cleaved RsmAY48W.**

1D NMR proton spectra of RsmAY48W with the sample buffer containing 50 mM NaCl, 0.6 mM  $K_2HPO_4$ , 0.3 mM  $KH_2PO_4$ , 0.02 %  $NaN_3$  and 10 %  $D_2O$  at pH 7.0, using DSS as an internal standard. The 0.24 mM sample was run at the following temperatures, 298 K (blue), 323 K (red), 348 K (green), 353 K (purple) and 298 K (yellow) cooled sample post-heating, with RsmAY48W experiencing loss of stability as temperature is increased. Protein purification method 2 (Table 3.1).

### 3.2.5.3 Denaturant

The 1D NMR spectra derived from the amide region (A) and the methyl region (B) of RsmA Y48W (cleaved) was obtained using the normal watergate (WG) sequence with additional WET solvent suppression due to the effect of using GdmCl as denaturant (Fig. 3.23). Each spectrum represents the protein in different concentrations of chemical denaturant GdmCl. The spectrum of the 0 M denaturant solution (blue) shows well resolved and dispersed proton peaks, indicating a good quality protein sample that is folded. Denaturant was then titrated into the sample. Upon the addition of denaturant from 0.84 M (red) the NH peaks broaden, until peaks are mostly gone at 4.4 M (purple) and completely lost from the spectrum at 5.2 M (yellow) showing the protein was fully denatured (unfolded) and NH protons were participating in fast exchange with the solvent environment. In the spectrum of 5.2 M denaturant (yellow), the peak remaining was due to the protons in the denaturant, GdmCl.

In Spectra B, as the denaturant concentration increases, the methyl residual structure starts to reduce. At 1.95 M (green) the structure is still there but starting to disappear, whereas from 2.25 M (lime) up to 4.4 M (purple) the folding collapses between 2.25 M (lime) and 3 M (orange). By watching the progress of the up field methyl group at 1.05 ppm the gradual decrease in intensity is representative of the unfolding event.



**Figure 3.23: 1D NMR WG proton spectra of the effect of chemical denaturant on cleaved RsmAY48W.**

NMR sample buffer used 50 mM NaCl, 0.6 mM  $K_2HPO_4$ , 0.3 mM  $KH_2PO_4$ , 0.02 %  $NaN_3$  and 10 %  $D_2O$  at pH 7.0, using DSS as an internal standard. Spectra A illustrates the effect of denaturant concentration increasing on the 0.4 mM sample in the amide proton region and Spectra B for the alkyl protons. Denaturant concentrations corresponding on the spectra are 0 M (blue), 0.8 M (red), 1.95 M (green), 2.25 M (lime), 3.0 M (orange), 4.4 M (purple) and 5.2 M (yellow). As the [GdmCl] increases the NH peaks broaden from 0.8 M, are mostly gone by 4.4 M and complete denaturation is observed at 5.2 M. Upon addition of GdmCl the methyl residual structure starts to reduce and from 2.25 M (lime) up to 4.4 M (purple) the main structure has gone completely. Protein purification method 2 (Table 3.1).

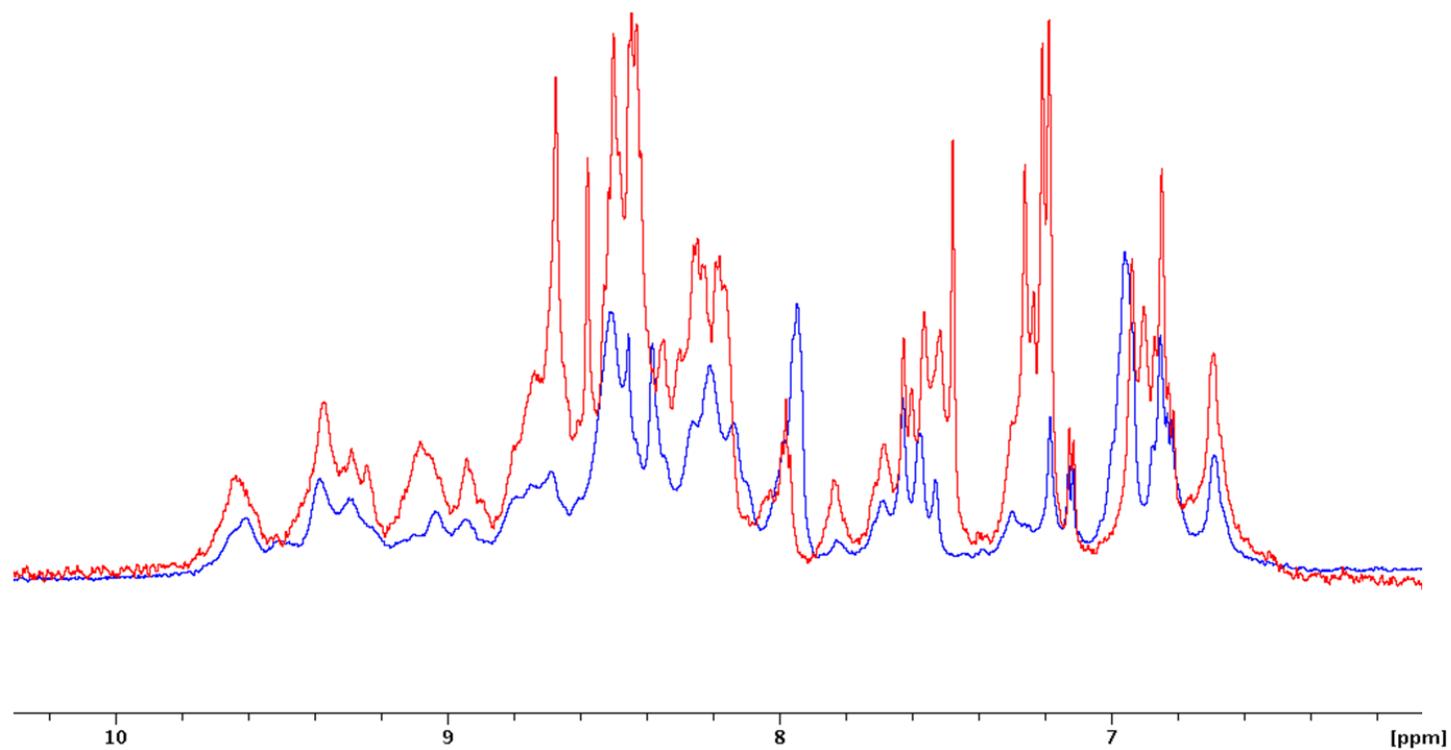
#### 3.2.5.4 The effect of pH on RsmA

By changing the pH of a solution, the proton environment and therefore the solubility of a protein can be altered. Solubility of His<sub>6</sub>-Thr-RsmA was assessed by decreasing the pH from 7.2 to 5.2 (Fig. 3.24). This reduction appeared to slow the exchange rate, leading to sharper and stronger peaks. However it is uncertain whether the increase in peak resolution and strength is due to the change in exchange rate or due to an actual change of structure.

The non-exchangeable alkyl region of both spectra will be less susceptible to solvent and pH effects. Comparison shows no significant difference in protein signals with only minor differences in buffer impurities reflecting the different preparation methodologies. The methyl regions of both spectra show a number of up field shifted signals suggesting the presence of packing interactions and therefore folding.

This comparison of spectra has an interesting significance. Better quality spectra can be obtained at lower pHs in order to obtain a structure assignment for the protein. However, in order to undergo binding studies with various RNAs, whether to run the spectra at the lower pH for optimal signal intensity or at a more biologically relevant pH is an important question for consideration (Gutiérrez et al., 2005, Schubert et al., 2007).

These NMR experiments have confirmed the purity of protein purified and stability of RsmA *in vitro* using temperature, denaturant and pH as probes.



**Figure 3.24: 1D NMR proton spectra of effect of pH on His<sub>6</sub>-Thr-RsmA stability.**

NMR sample buffer used 50 mM NaCl, 0.6 mM K<sub>2</sub>HPO<sub>4</sub>, 0.3 mM KH<sub>2</sub>PO<sub>4</sub>, 0.02% NaN<sub>3</sub> and 10% D<sub>2</sub>O at pH 7.0, using DSS as an internal standard. The buffers used were potassium phosphate pH 7.2 (blue) at 0.41 mM and sodium acetate pH 5.2 (red) at 0.23 mM protein concentration. Protein purification method 5.

### **3.3 CONCLUSIONS**

#### **3.3.1 Expression and purification of RsmA**

One of the main focuses of this study was to optimize the RsmA purification protocol in order for it to be reproducible and to enable the production of high yields of pure protein. The inefficient thrombin cleavage was removed from the protocol to enhance reproducibility. The introduction of the Superloop component enabled the loading of diluted samples, avoiding high salt concentrations which prevented the protein binding to the heparin and gel filtration column. Two peaks present on the heparin trace were separated by eluting over a greater volume, but analysis by CD and ESI-mass spectroscopy (not shown) revealed them to be identical.

A high molecular weight contaminant was identified by peptide mass fingerprinting as the catabolite gene activator protein (CAP) (Passner et al., 2000, Zhou et al., 1993). The CAP protein co-purified with RsmA and both were eluted from the Ni-NTA column. RsmA and CAP could be binding to each other or to a mutual partner, but this was not investigated further. A new purification protocol was implemented replacing the Ni-NTA agarose with HisPur™ cobalt resin, chosen because although cobalt resin gives a lower protein yield than Ni-NTA agarose, the protein is of a higher purity and the requirement for further optimization is reduced. This protocol combined with greater wash volumes succeeded in disrupting the binding of the CAP contaminant and removing it in a wash step prior to the elution of RsmA. The new washes were also used on a Ni-NTA purification resulting in the greater removal of contaminants. Therefore either resin is suitable. The final

purification method selected was gel filtration using low pH buffer (A: 50mM NaAc pH 4.5, B: A + 300 mM NaCl). The purification methods were often nonreproducible, although reducing the number of freeze-thaw stages during purification also resulted in a better yield.

### 3.3.2 Biophysical Methods

Prior to conducting biophysical analyses, a phenotypic assay was conducted based upon swarming activity in order to confirm the biological activity of the mutants in comparison to the wild type. Both V40W and Y48W retained biological activity (Figure 3.20).

ESI-mass spectroscopy comparison spectra of His<sub>6</sub>-Thr-RsmA demonstrated that using the cobalt HisPur™ resin gave samples of higher purity (Fig. 3.10). The molecular weight was calculated to be  $8,511 \pm 30.78$  Da which correlates very well with the predicted weight of 8,530 Da of the His<sub>6</sub>-Thr-RsmA monomeric unit. Further experimental testing would be necessary in order to elucidate the ideal conditions for the samples due to the complex behaviour of RsmA monomer-dimer.

CD spectroscopy comparisons of the purified RsmAV40W and RsmAY48W cleaved mutants with the wild type protein displayed identical traces of predominantly  $\beta$ -sheet character (Fig. 3.12 (A)). The experiments were repeated with the uncleaved His<sub>6</sub>-Thr-RsmA wild type protein and mutants (Fig. 3.12 (B)) with the same characteristic  $\beta$ -sheet character observed. The temperature melt profiles of all three cleaved proteins indicated that the proteins were not melting at 95 °C, although with the wild type protein the transition could be beginning at this temperature (Fig. 3.13 (A)). The

temperature melts of the uncleaved proteins displayed the same behaviour (Fig. 3.13 (B)). RsmA displayed identical degrees of  $\beta$ -sheet character before and after the temperature melt, confirming that unfolding had not occurred. A scan run at 95 °C at the end of the temperature melt does show a decrease in signal due to a loss of some of the beta sheet character. However upon cooling, the native conformation re-formed. Identical spectra were obtained when comparing protein prepared by either purification method (Fig. 3.14).

Equilibrium fluorescence spectroscopy was used to study the behaviour of the RsmA tryptophan mutants. Plasmids expressing RsmA with a V40W or a Y48W substitution were constructed and the proteins purified. However, neither the V40W nor the Y48W substitution mutants exhibited useful shifts in fluorescence upon denaturation. The lack of change in fluorescence suggests that only very small alterations in the environments of the substituted tryptophan are taking place. This could be due to the residues being too solvent exposed in the native conformation, that no quenching residues residing near in space in the native form, or to a combination of both. Prospective tryptophan mutants were identified, analysed and preliminary work started. These constructs could prove valuable for further fluorescence spectroscopy analysis. The purification methods were compared using NMR and revealed that RsmA from the Ni-NTA sample has well resolved, cleaner peaks compared with purification using the HisPur™ cobalt resin (Fig. 3.21). However even with different sample concentrations, RsmA in the two samples are definitely folded the same, leading to conclusion both methods produce pure protein.

1D NMR spectra of RsmAY48W (cleaved) at a range of temperatures (Fig. 3.22) reveals the transition from a folded protein with well distributed and sharp peaks, to an unfolded protein at 353 K where nearly all signal was lost. The reaction was shown to be non-reversible, contradicting the results found using circular dichroism, which indicated the protein to be stable at 80 °C. This is a very useful result as the concentrations needed in CD range are in the  $\mu\text{M}$  rather than in the nM range required for NMR. At lower concentrations, the protein is more solvent exposed, enabling greater stabilizing electrostatic interactions with the solvent. At higher concentrations, the overall change in conformation would result in an increase in stability. Both experiments would need to be run at higher temperatures to obtain a more definitive result.

The addition of GdmCl as denaturant (Fig. 3.23) caused a total loss of secondary structure by 5.2 M, indicating complete denaturation of the protein and participation of NH protons in fast exchange with the solvent environment. The increase in concentration of GdmCl demonstrates the collapse of folding, where the main structure is lost by 4.4 M.

The folded structure of RsmA is also highly similar at both a high and low pH (Fig. 3.24).

RsmA has been successfully over-expressed and purified in large quantities. However, further optimization for culture on a large scale and in a minimal, defined medium would be needed. The biophysical methods have confirmed that RsmA is a highly stable protein, although with conflicting results as to the temperature at which unfolding occurs. The use of CD, ESI-MS and NMR have confirmed the effect of purification on the protein, with cleaner peaks

from the Ni-NTA agarose than the HisPur™ cobalt resin and the successful introduction of the TritonX-100 and NaCl wash steps for contaminant removal. The molecular weight of RsmA was verified and the effects of denaturant and pH on RsmA were revealed. GdmCl successfully denatured RsmA and the folded structure was retained at both high and low pH.

Insights into the stability and structure of RsmA need to be combined with knowledge regarding its function and role within *P. aeruginosa*. A possible RsmA homologue RsmN was identified, therefore together with protocols set in place during this chapter, the characterization of RsmN and its impact on the regulation of virulence determinants with the aim of identification of an RsmN phenotype will be discussed in the next chapter.

## **4 IDENTIFICATION OF A NOVEL RSMA HOMOLOGUE IN *P. AERUGINOSA* AND ITS IMPACT ON THE REGULATION OF VIRULENCE DETERMINANTS**

### **4.1 INTRODUCTION**

Most pseudomonas species have at least a second *rsmA*-like gene often termed *rsmE*, and the genomes of some strains contain additional, still uncharacterised, potential *rsmA/rsmE* homologues (Reimmann et al., 2005). Expression of *rsmA* occurs earlier in growth than that of *rsmE*, and it is probable that they are differentially regulated. In homology dendrogram analysis RsmA/RsmE from some pseudomonas species are clustered in four groups according to degree of conservation in their primary sequence with respect to *E. coli* CsrA. RsmA sequences cluster together (75-85 % conservation), whilst the RsmE sequences (69-77%) and members of a third and fourth cluster distinctively at lower degrees of conservation (45-69 % and <45 % respectively) (Heeb et al., 2006). This variability is useful for the identification of conserved residues involved in structure maintenance and RNA binding, and less conserved residues that may be responsible for some specificity towards different target RNAs. Each group has a characteristic pattern of conserved residues in the putative RNA-binding site. It has previously been discussed that substitutions in the conserved residues of RsmA homologues are likely to have a significant effect on RNA binding, either lowering affinity or altering specificity (Heeb et al., 2006), therefore it is likely these substitutions will incur a similar effect on RsmE.

As well as being differentially regulated these homologues are therefore also likely to be functionally distinct and could bind preferentially to different

targets and/or regulatory RNAs. It is also possible that, if expressed at the same time, these proteins could form heterodimers that are likely to have distinct properties.

Antisense transcription in bacteria has only recently become a focus of genome wide analyses, conquering the traditional idea of bacterial transcriptomes composed mainly of protein-coding genes. A transcriptional profiling of *P. syringae* has recently revealed that antisense transcription occurs in 2.2 % of the known genes, as 124 genes were revealed to be significantly transcribed on both strands. This area of research has suggested that the regulation of gene expression can occur through *cis*-encoded asRNAs, giving rise to a previously unrecognised, distinct layer of regulatory control in bacteria (Filiatrault et al., 2010, Georg and Hess, 2011).

The aim of this chapter is to begin to evaluate the recently discovered gene encoding RsmN, a new potential RsmA homologue in *P. aeruginosa* (by M Messina and S Heeb). Initial analysis conducted by performing a sequence and structure comparison with RsmA revealed conserved residues which are good candidates for point mutations and subsequent phenotypic assays. The effect of RsmN on the expression of various genes of interest was assessed, and the impact of the AHL- and PQS-dependent QS systems on the expression of *rsmN* investigated.

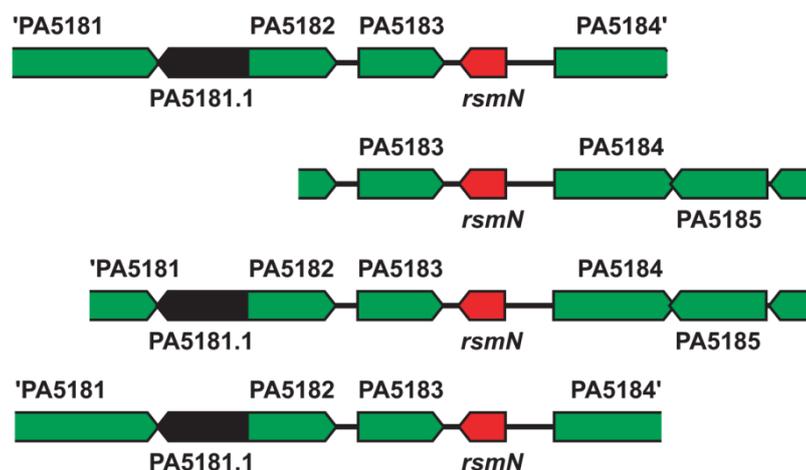
RsmA regulates negatively or positively the expression of various target genes at the post-transcriptional level by binding to the corresponding mRNAs. In *P. aeruginosa*, RsmA negatively regulates the production of a range of exoenzymes, secondary metabolites and virulence factors, including hydrogen

cyanide, pyocyanin, the staphylolytic enzyme, LasA and the galactophilic lectin, LecA, as well as the production of the QS molecules, 3-oxo-C12-HSL and C4-HSL (Pessi et al., 2001). In contrast, swarming motility, lipase and rhamnolipid production are positively regulated by RsmA in this organism (Heeb et al., 2004). The role of RsmN was investigated to determine whether *rsmN* could control virulence determinant and secondary metabolite production. These findings are of interest as together with RsmA, global post-transcriptional regulators such as RsmN could potentially become targets for novel antimicrobial drugs against *P. aeruginosa*.

#### **4.1.1 Identification of *RsmN***

Characteristic effects of an *rsmA* mutation in *Pseudomonas aeruginosa* PAO1 include a reduction of rhamnolipid production and of swarming motility (Heurlier et al., 2004). The mechanism by which RsmA exerts a positive effect on these phenotypes remains unclear. To clarify this, different systematic approaches were followed. Screening of random transposon mutants and genomic banks for the restoration of swarming in an *rsmA* mutant were conducted to identify novel elements mediating these regulations (M Messina, PhD thesis). This phenotype was used as the swarming deficiency in an *rsmA* mutant can be restored by complementation.

After the screening of a genomic bank consisting of a broad-host range vector carrying random 2 - 4 kb chromosomal DNA fragments, 14 plasmids were found to restore swarming. Four clones of interest shared the same unannotated intergenic region, between the annotated genes PA5183 and PA5184 which code for hypothetical proteins (Fig. 4.1).



**Figure 4.1: Restoration of swarming in *P. aeruginosa* *rsmA* mutants by clones identified as carrying *rsmN*.**

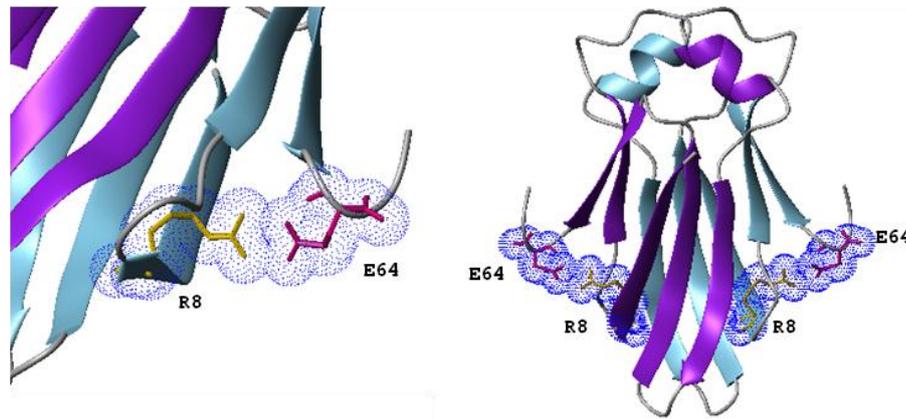
Diagram of the clones identified which restored the swarming phenotype and which carried *rsmN*, the new *rsmA* homologue. The *rsmN* gene is identified in red, sense genes to *rsmN* are in black and antisense genes to *rsmN* are identified in green. Surrounding ORFs encode; PA5181 (probable oxidoreductase), PA5181.1 (P34 ncRNA), PA5182-PA5184 (hypothetical proteins) and PA5185 (conserved hypothetical protein).

*In silico* analysis of this region revealed the presence of an unidentified open reading frame that encodes a protein of the RsmA/CsrA family. This ORF, which has been designated *rsmN*, encodes a 7.8 kDa protein which shares 34 % identity and 52 % similarity with the 6.9 kDa protein RsmA. RsmN (pI = 8.7) is a more basic protein than RsmA (pI = 7.4). The identification of RsmN and the transcriptional analysis was performed by M Messina and S Heeb.

#### 4.1.2 Sequence comparison of RsmN and RsmA

Sequence comparisons between RsmN and RsmA/CsrA homologues enabled the identification of strictly conserved residues (Fig. 4.2). The residues important for maintaining structure include Arg8 and Glu64 (Fig. 4.3), the corresponding residues (Glu 46 in RsmA) of which form an inter-chain salt-bridge in RsmA (Heeb et al., 2006). The representation of these residues by a





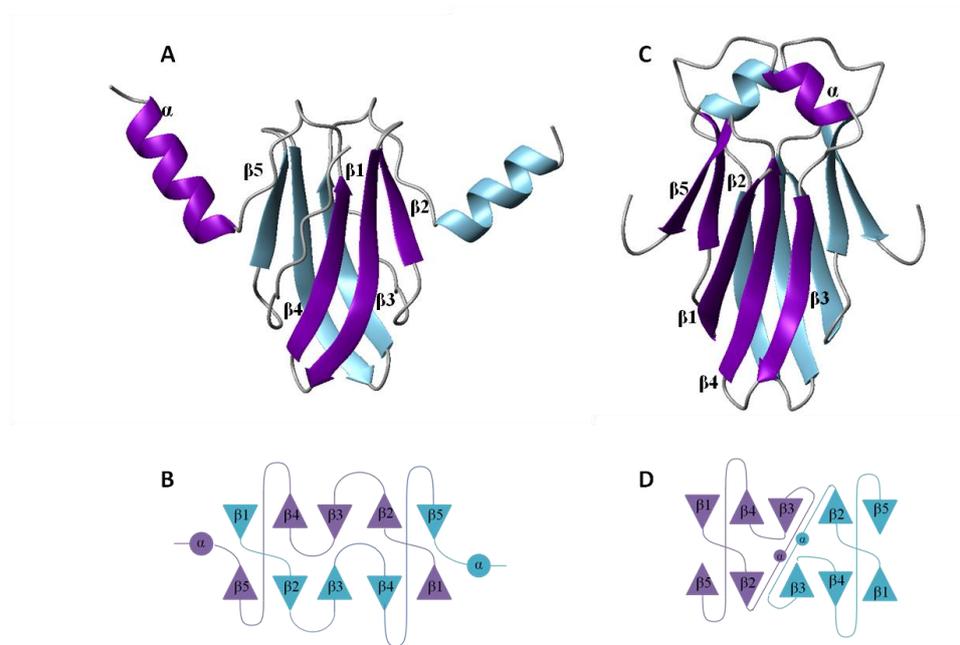
**Figure 4.3: Possible salt bridge in RsmN.**

Representation of a possible salt bridge in RsmN between arginine 8 (yellow neon residue) and glutamine 64 (pink neon residue) using the MolMol program (Koradi et al., 1996). RsmN is displayed in ribbon form, where the monomers are turquoise and purple. The close proximity is highly suggestive that there is a salt bridge between these R8 and E64 residues.

The two solvent-exposed residues Glu10 and Arg62 (Arg44 in RsmA) are also conserved (Fig. 4.3). Previous studies have shown that R44 is required for retention of biological function as the *rsmAR44A* mutant, in contrast to the wild type, is unable to restore the swarming in the *P. aeruginosa rsmA* mutant or to repress glycogen synthesis in an *E. coli csrA* mutant (Heeb et al., 2006). Thus, in RsmA this residue is essential for biological activity *in vivo* and RNA-binding *in vitro* (Heeb et al., 2006). RNA-binding proteins often contain discrete RNA binding modules such as the KH domain found in the mammalian neuro-oncological ventral antigen protein Noval (Lewis et al., 1999). Many, but not all, members of the RsmA family contain a sequence (VLGVKGGXXVR) that has been reported to be similar to a motif found in the KH domain (Romeo, 1998). This sequence is not conserved in RsmN or in RsmE (*P. fluorescens*) and is clearly not involved in RNA-binding in RsmA (Heeb et al., 2006).

### 4.1.3 Structural comparisons of RsmN and RsmA

The structure of RsmN was obtained by E. Morris (University of Nottingham, unpublished data), using X-ray crystallography (Fig. 4.4).

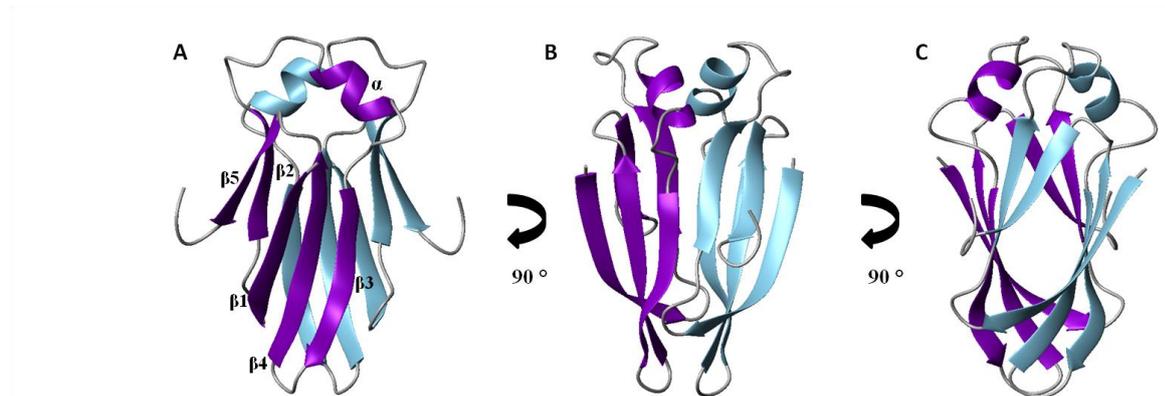


**Figure 4.4: RsmA and RsmN molecular models and schematics.**

Molecular model of RsmA (A) with the corresponding schematic of the dimer secondary structure (B) and the molecular model of RsmN (C) with the corresponding schematic of the dimer secondary structure (D). The molecular models are displayed in ribbon form where each monomer is represented in turquoise and purple. The purple monomer is labelled according to the order of secondary structure within the strand. The  $\alpha$ -helices (circles) and  $\beta$ -sheets (triangles) are represented in the schematics which also illustrate the spatial arrangement of the monomers within the dimer.

Both RsmA and RsmN are dimeric proteins, where the former contains two, five-stranded anti-parallel  $\beta$  sheets with  $\alpha$  helices projecting outwards from the C terminals (Heeb et al., 2006), and the monomers interact by hydrogen bonding between the separate strands, forming an intertwined structure. The RsmN protein also contains five  $\beta$  sheets and an  $\alpha$  helix, but the order is different. Instead, the monomers in RsmN form a clam-like structure, only interacting at one surface plane as demonstrated using a molecular model rotated around the vertical axis (Fig. 4.5). The helix, shorter compared with

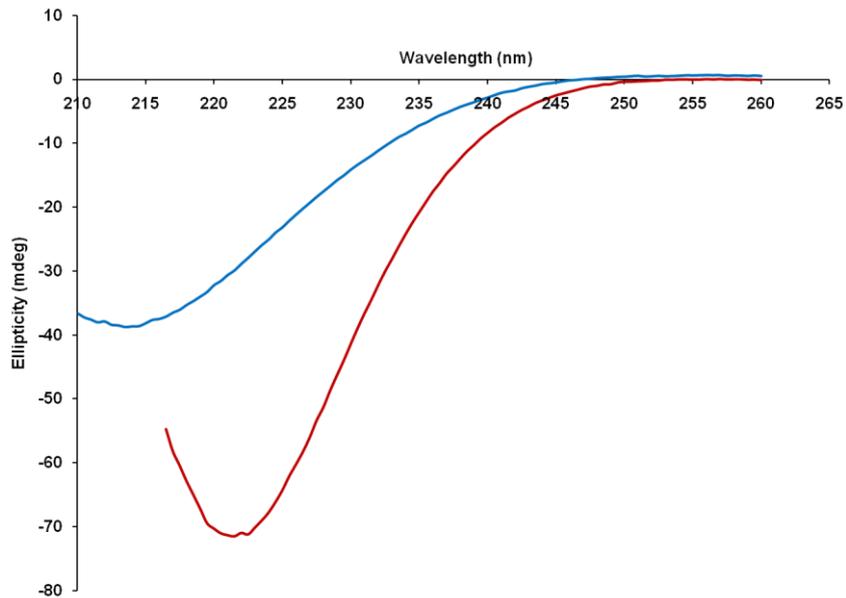
that of RsmA, is located between  $\beta 2$  and  $\beta 3$  at the interacting surface plane with the second monomer instead of projecting out from the protein at the C terminus. The interaction between the  $\beta$  sheets appears to occur between  $\beta 2$  and  $\beta 3$  of opposite strands and the helices.



**Figure 4.5: Molecular model of RsmN.**

The molecular model views of RsmN are displayed in ribbon form where each monomer is represented in turquoise and purple. The purple monomer is labelled with the order of secondary structure within that strand. View A is rotated  $90^\circ$  clockwise around the vertical axis for view B and  $90^\circ$  further for view C.

Native CD spectra were run using wild type His<sub>6</sub>-Thr-RsmA and His<sub>6</sub>-Thr-RsmN, both were purified using the pHT vector in the C41 cell line (Fig. 4.6). The proteins were purified using the cobalt HisPur™ resin followed by gel filtration, desalting and lyophilising. The minimum wavelength has shifted from 205 nm for RsmA to 220 nm for RsmN. The main conclusion from this data is that RsmN has greater  $\alpha$  helical content and that RsmA has more unstructured polypeptide chain than RsmN (E. Morris, personal communication).

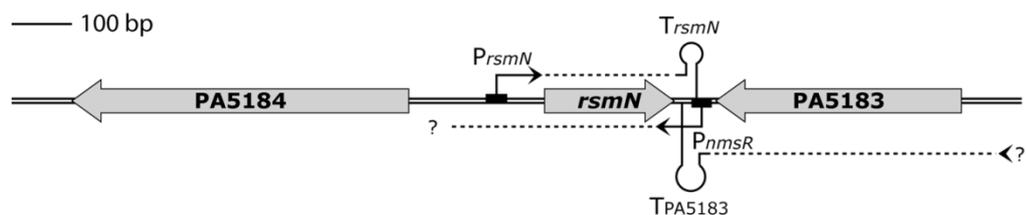


**Figure 4.6: CD comparison spectra of wild type His<sub>6</sub>-Thr-RsmA and His<sub>6</sub>-Thr-RsmN**  
 Samples were dissolved in a buffer of 25 mM K<sub>2</sub>HP0<sub>4</sub>, 50 mM NaCl, pH 7.0 at 25 °C, of His<sub>6</sub>-Thr-RsmA (– blue line) and His<sub>6</sub>-Thr-RsmN (– red line), both at 200 μM. RsmN has greater α helical content and RsmA has more unstructured polypeptide chain than RsmN.

Work is currently being conducted by E. Morris to optimise NMR conditions to generate solution state structural information. Together with folding studies, it will elucidate how the RsmN dimer is formed and which residues are necessary for structural and biological functions.

#### 4.1.4 Transcriptional analysis

To investigate the expression of *rsmN*, transcriptional analysis was performed. Intriguingly, in addition to the sense promoter P<sub>*rsmN*</sub>, a possible second promoter in the antisense direction P<sub>*nmsR*</sub> can be predicted (Fig. 4.7).



**Figure 4.7: Genetic context of *rsmN*.**

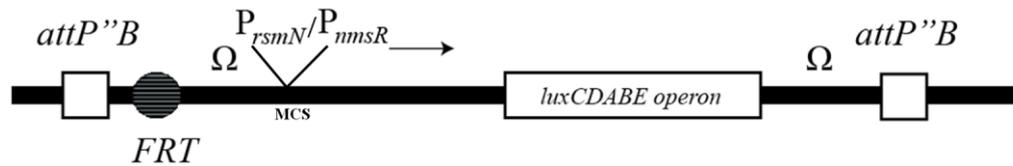
Genetic context and location of the *rsmN* gene (PAO1 genome nucleotides 5836776-5836991), which is antisense to PA5184 (PAO1 genome, nucleotides 5,836,910-5837467) and PA5183 (5835994-5836401). The location of the predicted promoter and terminator for *rsmN* are shown in the intergenic regions between PA5184-*rsmN* and *rsmN*-PA5183 respectively. The promoter for *nmsR* is located in the intergenic *rsmN*-PA5183 region in the antisense strand. The *rsmN* terminator, PA5183 terminator and the *nmsR* promoter are closely located next to each other with the *rsmN* terminator overlapping the *nmsR* promoter. The asRNA could potentially affect not only the expression of *rsmN* but also PA5183 (hypothetical protein).

## 4.2 RESULTS AND DISCUSSION

The aims of this chapter are to identify a *rsmN* phenotype, determine whether RsmN is a RsmA homologue and if so, what its role is within the Gac regulatory system.

### 4.2.1 Construction of strains used in this chapter

A variety of strains were constructed including transcriptional *lux* reporter fusions for both the sense  $P_{rsmN}$  and the anti-sense  $P_{nmsR}$  predicted promoters. Other promoter fusions made included those for *rhII*, *lasI* and *pqsA*. The miniCTX:*lux* vector was chosen as it contains a modified *lux* gene cluster from *Xenorhabdus luminescens* (Fig. 4.8) (Colepicolo et al., 1989, Becher and Schweizer, 2000). This bioluminescence operon allows for the monitoring of gene expression from the promoter inserted with no exogenous substrate required for light emission.



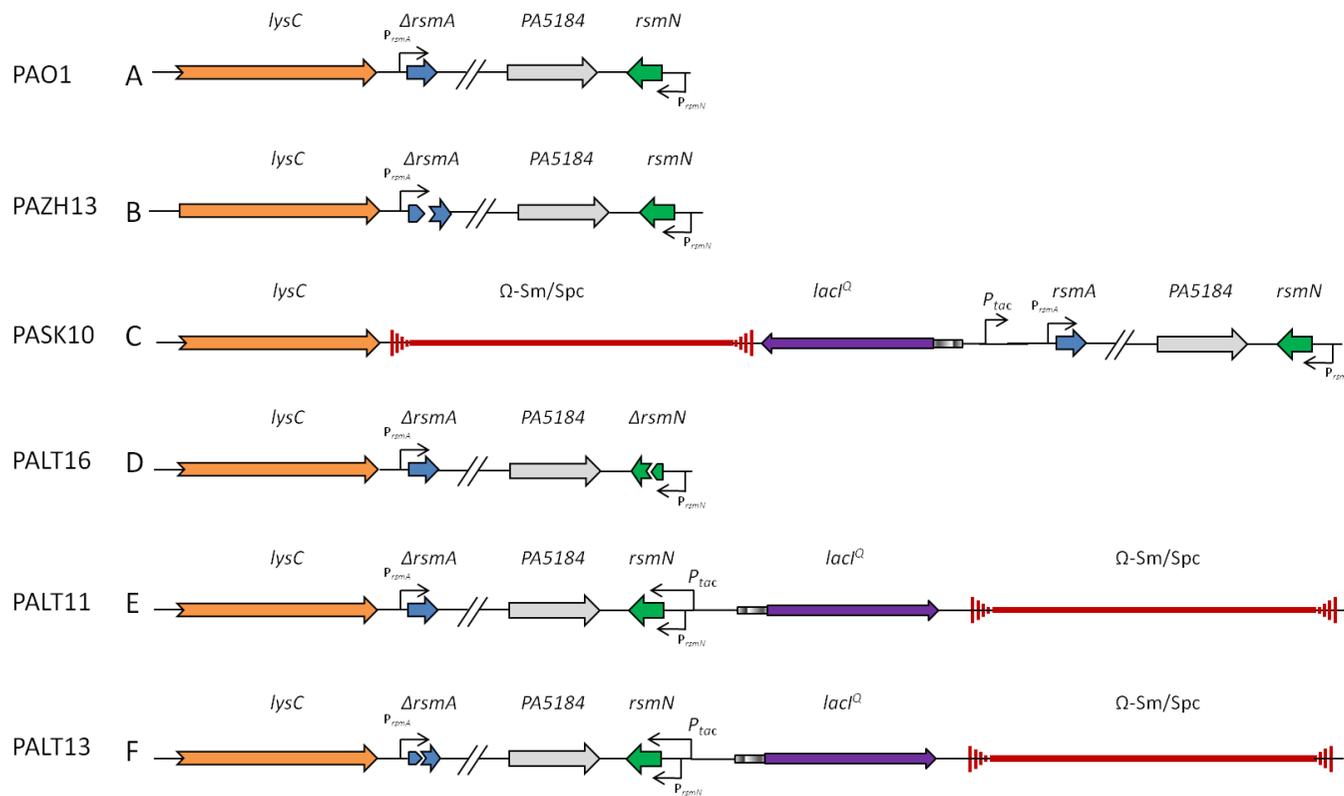
**Figure 4.8: Diagrammatic representation of the *rsmN* and *nmsR* miniCTX::*lux* promoter gene fusions.**

The *rsmN* or *nmsR* predicted promoter was inserted into the multiple cloning site (MCS) upstream of the *luxCDABE* operon and between two  $\Omega$ -cassettes. An engineered FRT site is present to remove any unwanted plasmid sequences from the genome. The integration of the *attP*-containing suicide plasmid occurs at the *attB* site in the *P. aeruginosa* recipient strain.

Chromosomal fusions were made in the *P. aeruginosa* strains PAO1 (wild type), PAZH13 ( $\Delta rsmA$ ), PASK10 (inducible *rsmA*), PALT16 ( $\Delta rsmN$ ) and PALT11 (inducible *rsmN*). These strains are described in section 2.4.1 and schematic representations of their genotypes are provided in Fig. 4.9. All strains constructed in this chapter were made using the Nottingham strain.

#### 4.2.1.1 mini-CTX::*lux* promoter fusions.

The sense and antisense promoter fusions,  $P_{rsmN}$  (pLT1) and  $P_{nmsR}$  (pLT2), were constructed as described in sections 2.4.1.8 and 2.4.1.9 respectively. For pLT1, the primers RSMNPF1 and RSMNPR1 (Table 2.3) were used to amplify a 331 bp product from the PAO1 wild type Lausanne strain genome with part of the sense promoter and flanking *XhoI* and *PstI* restriction sites (Fig. 4.10A). This was repeated with the primers RSMNPF2 and RSMNPR2 to produce a 452 bp product with part of the antisense promoter and flanking *HindIII* and *EcoRI* restriction sites for pLT2 (Fig. 4.10B). The mini-CTX::*lux* plasmid was then linearised with the required enzymes and the relevant product inserted. Following ligation the DNA was transformed into *E. coli* S17-1  $\lambda$ pir cells.



**Figure 4.9: Chromosomal constructs made in *P. aeruginosa* PAO1.**

Constructs were made using suicide plasmids, where (A) represents the wild type PAO1, (B) PAZH13, *rsmA* mutant, (C) PASK10, inducible *rsmA*, (D) PALT16, *rsmN* mutant, (E) PALT11, inducible *rsmN* and (F) PALT13, inducible *rsmN*, *rsmA* mutant. The genes are drawn to scale with the correct orientations: *lysC* (orange), *rsmA* (blue), PA5184 (grey), *rsmN* (green), *lacI*<sup>Q</sup>P<sub>tac</sub> (purple) and Ω-cassette (Ω-Sm/Spc in red).



#### 4.2.1.2 *rhlI*, *lasI* and *pqsA* promoter fusions

Promoter fusions using the mini-CTX::*P<sub>rhlI</sub>-lux*, mini-CTX::*P<sub>lasI</sub>-lux* (G. Rampioni, private communication), and mini-CTX::*P<sub>pqsA</sub>-lux* (Diggle et al., 2007) plasmids were made using the donor strains PA01, PAZH13, PASK10, PALT16 and PALT11. The resulting strains are shown in Table 4.2, taken from Table 2.1.

**Table 4.2: *rhlI*, *lasI* and *pqsA* promoter fusions.**

PA Number	Genotype/Characteristics
PALT22	PA01::(miniCTX:: <i>P<sub>pqsA</sub>-lux</i> )
PALT23	PA01::(miniCTX:: <i>P<sub>rhlI</sub>-lux</i> )
PALT24	PA01::(miniCTX:: <i>P<sub>lasI</sub>-lux</i> )
PALT25	PALT11::(miniCTX:: <i>P<sub>pqsA</sub>-lux</i> )
PALT26	PALT11::(miniCTX:: <i>P<sub>rhlI</sub>-lux</i> )
PALT27	PALT11::(miniCTX:: <i>P<sub>lasI</sub>-lux</i> )
PALT28	PALT16::(miniCTX:: <i>P<sub>pqsA</sub>-lux</i> )
PALT29	PALT16::(miniCTX:: <i>P<sub>rhlI</sub>-lux</i> )
PALT30	PALT16::(miniCTX:: <i>P<sub>lasI</sub>-lux</i> )
PALT31	PAZH13::(miniCTX:: <i>P<sub>pqsA</sub>-lux</i> )
PALT32	PAZH13::(miniCTX:: <i>P<sub>rhlI</sub>-lux</i> )
PALT33	PAZH13::(miniCTX:: <i>P<sub>lasI</sub>-lux</i> )
PALT44	PASK10::(miniCTX:: <i>P<sub>pqsA</sub>-lux</i> )
PALT45	PASK10::(miniCTX:: <i>P<sub>rhlI</sub>-lux</i> )
PALT46	PASK10::(miniCTX:: <i>P<sub>lasI</sub>-lux</i> )

#### 4.2.1.3 Sense and antisense *rsmN* and *nmsR* fusions in $\Delta rhlR$ , $\Delta lasR$ and

##### $\Delta pqsA$

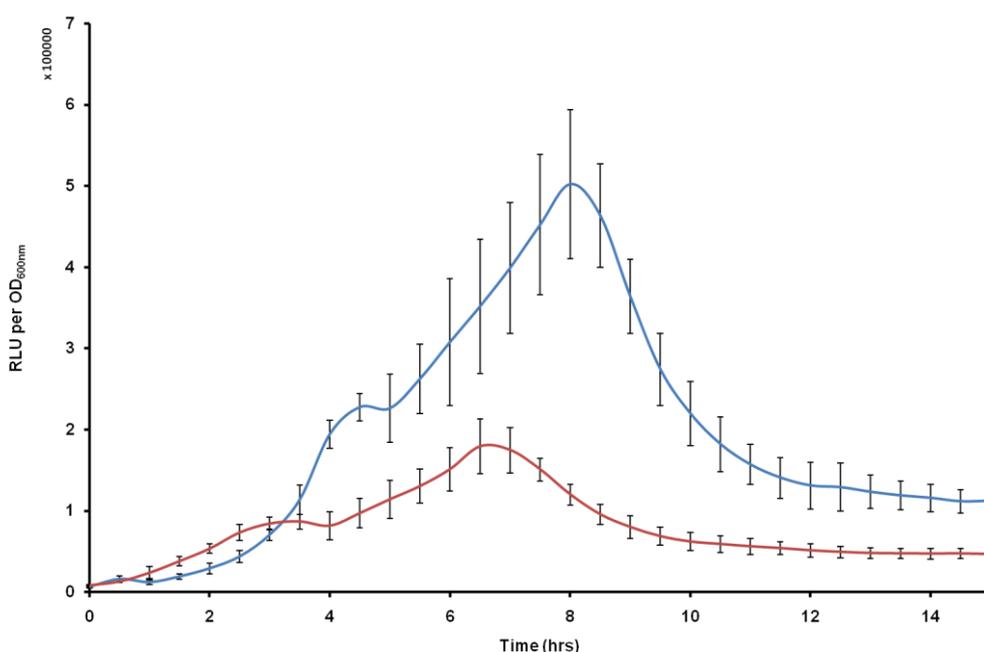
Sense and antisense promoter fusions were made using *P<sub>rsmN</sub>* (pLT1) and *P<sub>nmsR</sub>* (pLT2) in the *P. aeruginosa* strains PACP10 ( $\Delta rhlR$ ), PASDP233 ( $\Delta lasR$ ) and PASDP123 ( $\Delta pqsA$ ) by conjugation, shown in Table 4.3 (taken from Table 2.1). Control strains were also made using the mutant strains containing the empty mini-CTX::*lux*.

**Table 4.3: *rsmN* and *nmsR* promoter fusions in  $\Delta rhIR$ ,  $\Delta lasR$  and  $\Delta pqxA$** 

PA Number	Genotype/Characteristics
PALT49	PACP10::(miniCTX:: <i>P<sub>rsmN</sub>-lux</i> )
PALT50	PACP10::(miniCTX:: <i>P<sub>nmsR</sub>-lux</i> )
PALT51	PASDP123::(miniCTX:: <i>P<sub>rsmN</sub>-lux</i> )
PALT52	PASDP123::(miniCTX:: <i>P<sub>nmsR</sub>-lux</i> )
PALT53	PASDP233::(miniCTX:: <i>P<sub>rsmN</sub>-lux</i> )
PALT54	PASDP233::(miniCTX:: <i>P<sub>nmsR</sub>-lux</i> )
PALT55	PACP10::(miniCTX:: <i>lux</i> ), negative control
PALT56	PASDP123::(miniCTX:: <i>lux</i> ), negative control
PALT57	PASDP233::(miniCTX:: <i>lux</i> ), negative control

#### 4.2.2 *rsmN* and *nmsR* gene expression

The transcription levels are low but can effectively occur from both *rsmN* and *nmsR* promoters, with the activity of the sense promoter nearly three times that of the antisense promoter (Fig. 4.11).



**Figure 4.11: Expression of *rsmN* and *nmsR* promoters in *P. aeruginosa* PAO1 (Nottingham) as a function of growth.**

A dilution of an o/n culture adjusted to OD<sub>600nm</sub> 1.0 of 1:1000 was used to inoculate sterile LB. The experiment was run in 96 well plates using a GENios Tecan for 15 h at 37 °C measuring OD<sub>600nm</sub> and luminescence. *P<sub>rsmN</sub>* (-) = Sense promoter fusion, *P<sub>nmsR</sub>* (-) = Antisense promoter fusion. Technical replicates where N = 9. All error bars used in this thesis are ± 1 standard deviation (SDev).

The comparison is statistically significant to 5% with a t value at 8 hours of 2.83 using a critical t value of 1.746 (16 degrees of freedom (DoF)). Details of how the t test was performed are available in Appendix I.

#### 4.2.2.1 Construction of RsmN Arginine-62-Alanine (R62A) mutants

Primers were designed to introduce an alanine replacement mutation into the wild type *rsmN* gene to generate *rsmNR62A* using the Stratagene Quick Change Site-Directed Mutagenesis kit<sup>®</sup> as described in section 2.4.1.2. This R62A mutant was made in *rsmN* as the corresponding conserved residue R44A in *rsmA* is essential for biological activity *in vivo* and RNA-binding *in vitro* (Heeb et al., 2006). PCR mutagenesis to introduce the R62A mutation in pLT27 (pME6032::*rsmN*) using the primers R62A\_F and R62A\_R was unsuccessful, therefore the experiment was repeated using pLT25 (pGEM-T::*rsmN*) DNA (3015 bp empty vector) as the template for the PCR reaction. The *rsmNR62A* fragment was removed from pLT30 (pGEM-T::*rsmNR62A*) using *EcoRI-ClaI* and inserted into pME6032 to produce pLT31 (pME6032::*rsmNR62A*).

#### 4.2.2.2 Construction of an *rsmN* mutant (PALT16)

An *rsmN* deletion mutant was made using a two step homologous recombination procedure where the pDM4-based suicide plasmid pMM33 (M Messina, personal communication) was mobilised by conjugation from *E. coli* into the recipient PAO1 strain. The suicide plasmid pMM33 was maintained in

*E. coli* S17-1  $\lambda$ pir, which also supplies the *tra* genes for efficient conjugation and mobilisation of the plasmid into *P. aeruginosa* (section 2.4.1.5).

#### 4.2.2.3 Construction of a conditional, inducible *rsmN* mutant (strain PALT11)

To produce an inducible, conditional *rsmN* mutant, the pDM4-based suicide plasmid pLT10 was mobilised by conjugation from *E. coli* into the recipient strains PAO1 and PAZH13 ( $\Delta$ *rsmA*) to produce strains PALT11 (*lacI*<sup>Q</sup>, P<sub>tac</sub>-*rsmN*) and PALT13 ( $\Delta$ *rsmA*, *lacI*<sup>Q</sup>, P<sub>tac</sub>-*rsmN*), respectively. The conditional mutants were made using the same method as the *rsmN* mutant PALT16 (section 2.4.1.5) using the plasmid pLT10 (section 2.4.1.4).

#### 4.2.2.4 Construction attempts for a $\Delta$ *rsmA* $\Delta$ *rsmN* double mutant strain

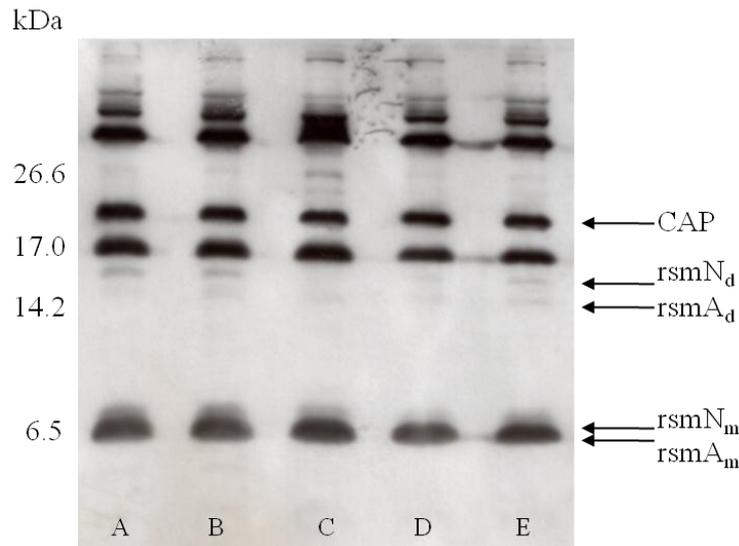
Two different approaches were used in attempting to produce a double  $\Delta$ *rsmA* $\Delta$ *rsmN* mutant. The first was to perform a conjugation using the *rsmN* mutant PALT16 and the suicide plasmid pZH13 (pDM4 carrying  $\Delta$ *rsmA*, (Pessi et al., 2001)). In the second approach a conjugation was performed using the *rsmA* mutant PAZH13 and the suicide plasmid pMM33 (pDM4 carrying  $\Delta$ *rsmN*). After performing the conjugations overnight, the samples were resuspended in LB and plated on LB agar containing nalidixic acid (to counter-select *E. coli*) and chloramphenicol (selecting for suicide plasmid integration in the chromosome), and incubated at 37 °C overnight. Although colonies were always present at this stage, the conjugation using pMM33 normally had to be left two days instead of one for the colonies to be of a suitable size to sample. It has been noted in previous work (S. Kuehne, PhD thesis) that *rsmA* mutant

strains have poor transformation efficiencies in comparison with the wild type. After the sucrose and antibiotic selection there were normally greater than 100 colonies from the conjugation PALT16 × pZH13 but only ~ 20 for the conjugation of PAZH13 × pMM33. When checked by PCR in comparison to the wild type none of the clones had successfully accomplished the desired allelic exchange. The experiment was repeated numerous times with varying conditions in order to try to optimise the recombination events. Conjugations were performed from 1 to 24 h, a greater number of colonies were sampled and the sucrose concentration used for selection was increased from 5 to 10 %. None of the variations successfully selected for the double mutant.

#### 4.2.2.5 Western Blot

Expression of the *rsmN* gene product was investigated by Western blot analysis using the *P. aeruginosa* strains PA01, PALT16 (*rsmN* mutant), PALT13 (inducible *rsmN* in *rsmA* mutant strain and PAZH13 (*rsmA* mutant) (Figure 4.12).

A combination of RsmA and RsmN strains were used in order to try to visualise the separate protein bands. Unfortunately, this was unsuccessful as all strains have double bands that can be visualised around 6-8 kDa, possibly of both RsmA and RsmN monomers. There are multiple cross-reactive bands visible at higher molecular weights including two the could correspond to the RsmA and RsmN dimers.



**Figure 4.12: Western blot analysis of RsmN production.**

Western blot detection of RsmN produced in the strains; A: PAO1 (wild type, wt), B: PALT16 ( $\Delta rsmN$ ), C: PALT13 - IPTG ( $\Delta rsmArsmN$ ), D: PALT13 + IPTG ( $\Delta rsmArsmN^{++}$ ) and E: PAZH13 ( $\Delta rsmA$ ). The proteins were sampled from whole cell lysate taken from 1 ml of o/n cultures for gel electrophoresis after which the proteins were transferred to a PVDF membrane by electroblotting and probed for RsmN using a polyclonal antibody raised against RsmN-.

### 4.2.3 Phenotypic characterization of the *rsmN* mutant

RsmA is involved in the post-transcriptional regulation of a range of secondary metabolites, virulence factors and swarming motility (Pessi et al., 2001, Heurlier et al., 2004, Heeb et al., 2006). Phenotypes were compared using the wild type PAO1, PALT16 ( $\Delta rsmN$  mutant), PAZH13 ( $\Delta rsmA$  mutant) and PALT11 (inducible *rsmN*) strains, complemented with *rsmN*. Analysis of swarming, elastase, protease and pyocyanin production in *P. aeruginosa* (Pessi et al., 2001, Heurlier et al., 2004) and glycogen synthesis in the *E. coli* strain TR1-5 (Romeo et al., 1993) were assayed.

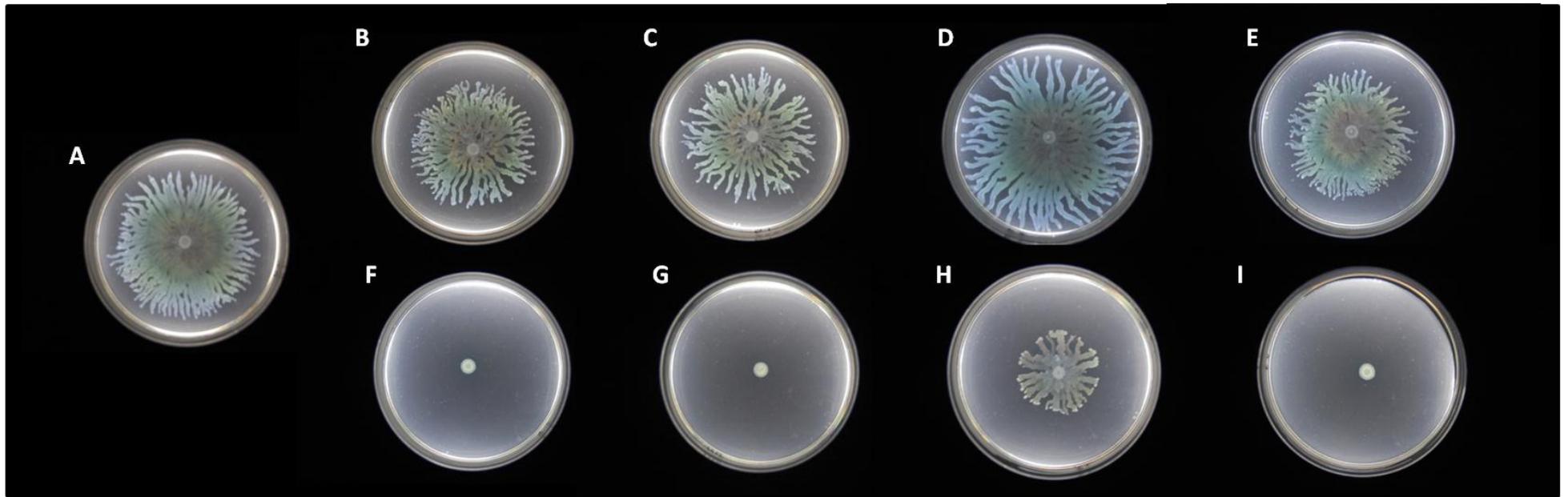
#### 4.2.3.1 Swarming

##### 4.2.3.1.1 *rsmN* mutant

The *rsmN* mutant is not defective in swarming motility (Fig. 4.13). However, swarming appeared to be enhanced by a plasmid containing hexahistidine tagged *rsmN*. The plasmid containing the arginine substitution mutant R62A in *rsmN* did not affect the swarming. The same plasmids were transformed into the *rsmA* mutant strain. As expected, the *rsmA* mutant is deficient in swarming, but this phenotype was not restored when *rsmN* was used for complementation. The phenotype could be partially restored in an *rsmA* mutant when a plasmid expressing a hexahistidine-tagged version of *rsmN* is transformed. There was no change when the *rsmA* mutant was complemented with the arginine substitution mutant compared with the wild type allele of *rsmN*.

Further complementation with a hexahistidine-tagged *rsmN* arginine mutant would determine if the arginine mutation only has an effect when the histidine tag is present.

RsmN is therefore not necessary for swarming. In contrast, the hexahistidine tagged RsmN produces an interesting behaviour with both the *rsmN* and *rsmA* mutants, in that swarming appears to be enhanced and induced, respectively. As this effect is due to a difference of 6 amino acids at the N terminal, this consequence could be due to a stabilisation of the transcript, or the tag could be interfering with the possible effects of the antisense gene *nmsR*.



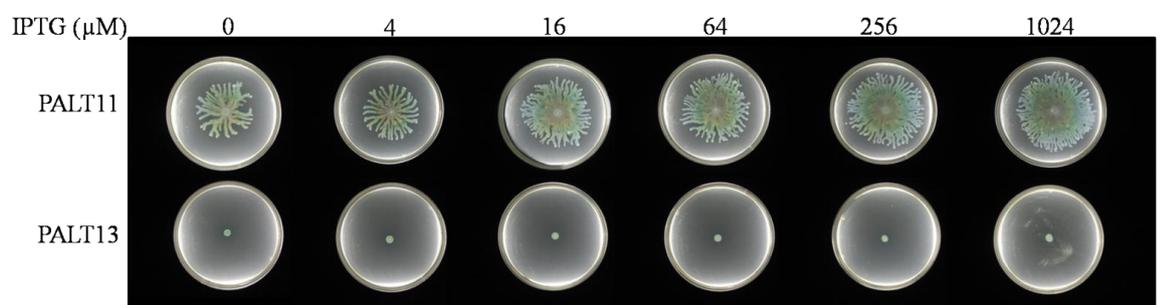
**Figure 4.13: Swarming motility of *P. aeruginosa* *rsmA* and *rsmN* mutants complemented by RsmN variants.**

Strains used were PAO1 (wt), PALT16 ( $\Delta rsmN$  mutant) and PAZH13 ( $\Delta rsmA$  mutant). Culture droplets of strains A: PAO1/pME6032, B) PALT16/pME6032, C) PALT16/pRsmN, D) PALT16/pH<sub>6</sub>RsmN, E) PALT16/pRsmNR62A, F) PAZH13/pME6032, G) PAZH13/pRsmN, H) PAZH13/pH<sub>6</sub>RsmN, I) PAZH13/pRsmNR62A, were spotted onto the middle of swarming plates and incubated o/n at 37 °C (Rashid and Kornberg, 2000). Swarming was not disrupted in the PALT16 strains, an increase is seen when complemented by the hexahistidine tagged-RsmN containing plasmid. As expected, the *rsmA* mutant strain PAZH13 was negative for swarming motility. Complementation with pH<sub>6</sub>RsmN partially restores swarming, however the other RsmN-containing plasmids had no visible effect.

Importantly, as RsmN is unable to fully complement the swarming phenotype, this poses a contradiction as this was the phenotype which was originally used to identify RsmN (Section 4.1.1). It could be that an antisense mechanism that is necessary for complementation has been inactivated due to just the *rsmN* gene being cloned into the complementation plasmids. Another consideration is that RsmN was not the target that complemented the original phenotype. Looking back to Figure 4.2 there are two partial and one full ORF which were consistent between the clones which restored the swarming phenotype, PA5182, PA5183 and PA5184, all three of which are hypothetical proteins. None of these hypothetical proteins show sequence similarity to RsmA.

Although RsmN has shown to not complement this phenotype, the similarity of both the sequence and folding to RsmA is striking and therefore worthy of continuing the study into deciphering its possible role within *Pseudomonas aeruginosa*.

Using the conditional inducible *rsmN* locus and the inducible *rsmN rsmA* mutant, the swarming motility assay was repeated, using a gradual increase in the concentration of IPTG when inducing the expression of *rsmN* (Fig. 4.14).



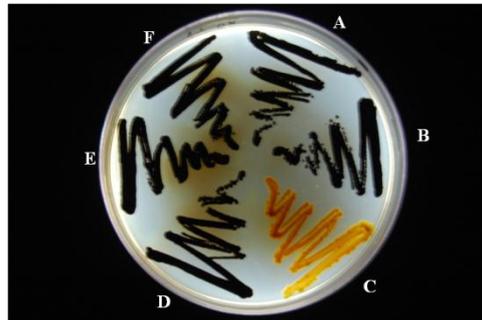
**Figure 4.14: Swarming motility of the inducible *P. aeruginosa rsmN* mutant.** Culture droplets of *rsmN* inducible strains PALT11 (*rsmN*<sup>ind</sup>) and PALT13 (*rsmA* mutant *rsmN*<sup>ind</sup>) were spotted in the middle of swarming plates with IPTG present in 6 concentrations from 0–1024 μM and incubated o/n at 37 °C. Swarming occurs with and without addition of IPTG to PALT11, however as the concentration of IPTG increases, so does the degree of swarming.

IPTG was added to the plates from 0 - 1024  $\mu$ M to compare swarming in the absence of IPTG up to an excess of this inducer. When *rsmN* is not expressed (at 0  $\mu$ M IPTG) there is swarming present in the otherwise wild type background (where *rsmA* is present) but not in the *rsmA* mutant. As the level of IPTG increases, the swarming phenotype of PALT11 appears to increase slightly. In the *rsmA* mutant strain, when *rsmN* is induced there is no restoration of swarming. These results suggest that RsmN can enhance swarming but only in the presence of RsmA. While *rsmN* does not act as an *rsmA* homologue in the swarming assays, the introduction of the hexahistidine tagged *rsmN* partially restores swarming motility in a *rsmA* mutant. The insertion of the tag is after the *rsmN* promoter but prior to the *rsmN* gene. A tag of such a small size would be expected to have limited effect on the RsmN protein sterically, but the addition of six basic, polar residues might be relevant, especially given their location close to the R62 residue. However, these results are surprising as the *rsmN* locus was identified multiple times in a genomic bank for its capacity to restore the swarming in an *rsmA* mutant. Further investigation of the expression levels of *rsmN* obtained with these plasmids, together with a better understanding of the role of the antisense *nmsR* gene are required.

#### 4.2.3.2 Glycogen accumulation in *E. coli*

The *rsmA* gene from *P. aeruginosa* can complement a *csrA* mutation in the *E. coli* strain TR1-5 that causes glycogen overproduction (Romeo et al., 1993, Pessi et al., 2001). The effect of the *rsmN* gene and its variants on the

repression of glycogen production was examined. Kornberg medium and iodine staining were used to reveal glycogen accumulation in *E. coli* strains expressing *rsmN* (Fig. 4.15). As expected, strains TR1-5 and TR1-5/pME6032 showed glycogen accumulation as indicated by a dark iodine staining. As expected, the *rsmA* gene from *P. aeruginosa* was capable of fully complementing the *csrA* mutation, *i.e.*, there was no iodine staining revealing no glycogen accumulation. However, the *rsmN* gene did not complement the *csrA* mutation. In the *E. coli* TR1-5 strain, the plasmids expressing the hexahistidine-tagged RsmN protein and the R62A arginine substitution mutant protein presented the same phenotype as strains TR1-5 or TR1-5/pME6032, indicating that no complementation was obtained with any of these constructs.



**Figure 4.15: Repression of glycogen synthesis in *E. coli* by RsmA but not RsmN.**

Kornberg medium and iodine staining were used to reveal glycogen accumulation in the *E. coli* TR1-5 strain complemented with RsmN variants and RsmA where; A) TR1-5, B) TR1-5/pME6032 (empty vector), C) TR1-5/pRsmA, (D) TR1-5/pRsmN, (E) TR1-5/pH<sub>6</sub>RsmN and (F) TR1-5/pRsmNR62A. Single colonies were streaked onto the prepared plates and incubated o/n at 37 °C. A dark brown colour indicates abnormal glycogen accumulation (Romeo et al., 1993). RsmA complemented the *csrA* mutation, however none of the RsmN variants were active.

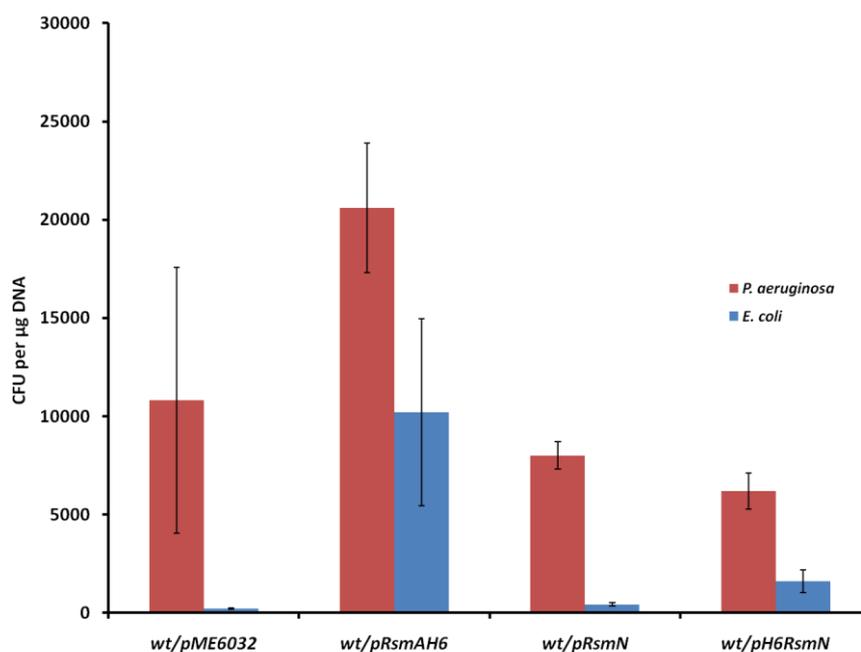
#### 4.2.3.3 Restriction assay

As noted in previous work (S. Kuehne, PhD thesis, University of Nottingham) the transformation efficiency of the inducible *rsmA* mutant strain is comparable for plasmids extracted from PAO1 under both, induced and non-induced

conditions. The same efficiency is obtained when the plasmid originates from *E. coli* and *rsmA* is induced. However, efficiency dropped radically when the cells were grown in the absence of IPTG, *i.e.*, when *rsmA* is not expressed. This suggested that a restriction-modification system in *P. aeruginosa* might be strongly controlled by RsmA. Restriction systems help the bacteria to protect themselves from invasion of foreign DNA as these systems represent a barrier, for example, against detrimental bacteriophage infection.

To investigate the effect of RsmN on restriction, the wild type PAO1 strain was transformed with plasmids overexpressing *rsmAH<sub>6</sub>*, *rsmN*, *H<sub>6</sub>rsmN*, and with the empty expression vector pME6032 as a control. DNA of the broad host range plasmid pME6001 (Gm<sup>R</sup>) was extracted from both, *E. coli* and *P. aeruginosa*, and separately transformed into each strain using chemically competent cells. The plasmids containing the *rsmN* and *rsmA* genes express these under the control of the inducible P<sub>tac</sub> promoter, therefore the strains were grown with the addition of IPTG.

All of the strains were efficiently transformed with DNA extracted from *P. aeruginosa* (Fig. 4.16). However, the strain containing the overexpressed RsmA performed to a higher efficiency than the wild type or the *rsmN*-overexpressing strains. When DNA from *E. coli* was transformed, the efficiency was very poor for the wild type and *rsmN*-overexpressing strains, but of a reasonable efficiency in the *rsmA*-overexpressing strain.



**Figure 4.16: Restriction Assay for *rsmN* and *rsmA* complemented PAO1 strains.** PAO1 wild type chemically competent cells consisting of wt/pME6032, wt/pRsmAH<sub>6</sub>, wt/pRsmN and wt/pH<sub>6</sub>RsmN underwent transformation with 50 ng of pME6001 DNA extracted from either *E. coli* or *P. aeruginosa*. After recovery in LB, a series of dilutions were plated in triplicate and incubated o/n at 37 °C. The colony forming units (CFUs) were counted and the average taken. DNA from *P. aeruginosa* performed with good efficiency, notably when containing the pRsmAH<sub>6</sub> plasmid. Transformation efficiency of plasmids from *E. coli* was generally poor, however a reasonable efficiency was observed with the pRsmAH<sub>6</sub> containing strain. Technical replicates where N = 3, error bars are ± 1 SDev.

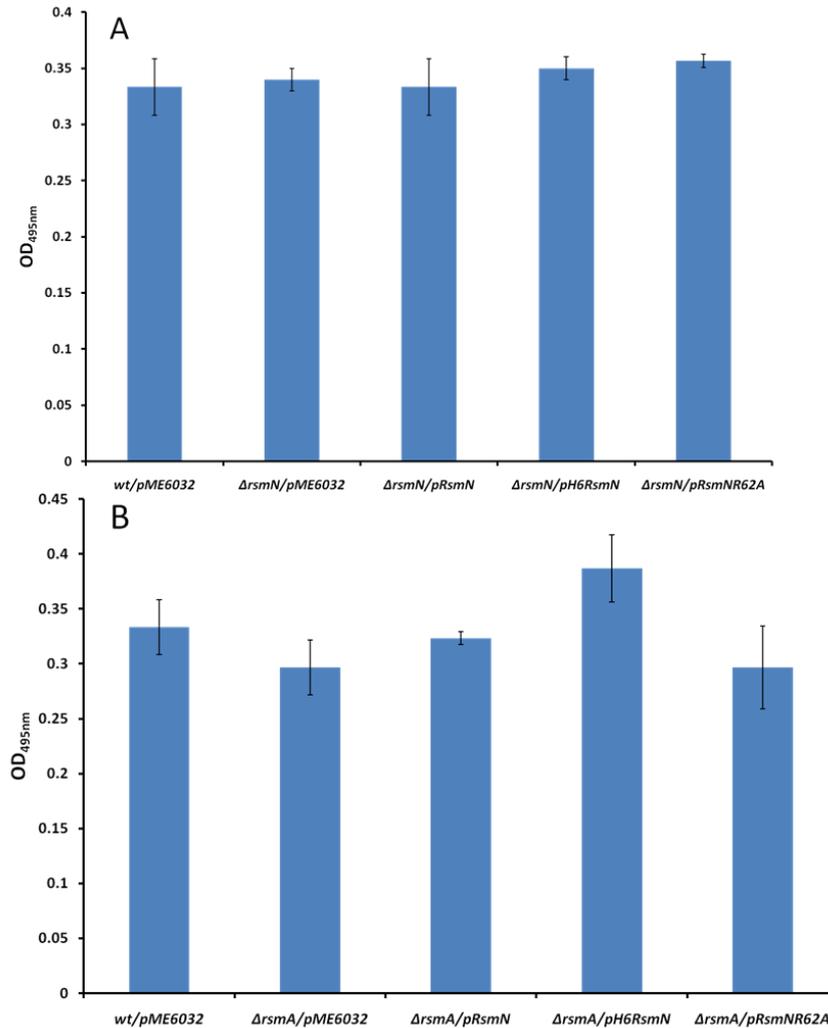
The results suggest that RsmN, unlike RsmA, does not appear to have control on the restriction-modification system of *P. aeruginosa*. However, repeating the experiments would be beneficial by comparing transformation efficiencies for the inducible *rsmN* strain PALT11.

#### 4.2.3.4 Control of secondary metabolite production

##### 4.2.3.4.1 Elastase Assay

The elastin-congo red based elastase experiments were conducted using the wild type PAO1, PAZH13 ( $\Delta rsmA$  mutant) and PALT16 ( $\Delta rsmN$  mutant). The plasmids pRsmN, pH<sub>6</sub>RsmN and pRsmNR62A were transformed separately

into these strains, as well as the empty plasmid pME6032. All complementation plasmids containing the *rsmN* variants were inserted into the *rsmA* mutant as well as the *rsmN* mutant strain, to enable a comparison of the effect of the uncharacterised *rsmN* with *rsmA*.



**Figure 4.17: Elastin-congo red assay to investigate the impact of RsmN on elastase production.**

Panel (A) compares the  $\Delta rsmN$  mutant strains and panel (B) compares the  $\Delta rsmA$  mutant strains. The supernatants of each strain were incubated with 5 mg of elastin congo-red for 2 h at 37 °C. The reaction was stopped by the addition of 120 mM EDTA and the OD<sub>495</sub> of the supernatant recorded after centrifugation. Technical replicates where N = 3, error bars are  $\pm$  1 SDev.

The average OD readings for all the strains are very similar. The large standard deviations mean that there is a minimal change in elastase activity between the

strains (Fig. 4.17, panel A). Therefore under these conditions RsmN is not involved in elastase production.

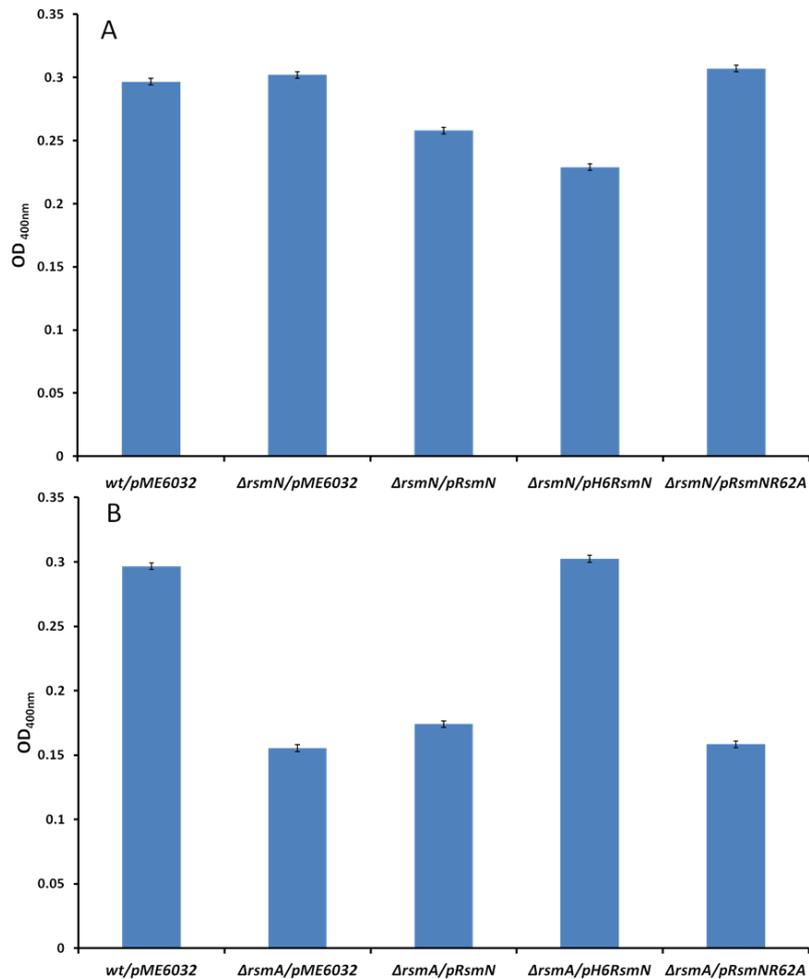
It has previously been shown that overproduction of RsmA causes a reduction in the levels of elastase production and that the wild type levels were similar to that of the mutant (Pessi et al., 2001). As well as using the *rsmA* and *rsmN* mutant strains, the experiment could be repeated using the wild type strain complemented by pRsmA and pRsmN.

#### 4.2.3.4.2 Protease Assay

The azocasein based protease experiments were conducted using the wild type PAO1, PAZH13 ( $\Delta rsmA$  mutant) and PALT16 ( $\Delta rsmN$  mutant). The plasmids pRsmN, pH<sub>6</sub>RsmN and pRsmNR62A were transformed separately into these strains, as well as the empty plasmid pME6032.

The average OD<sub>400</sub> readings for the *rsmN* mutant and wild type strains containing pME6032 are very similar at 0.3 (Fig 4.18A). The complemented strain shows a reduction to 0.25 and the strain complemented with pH<sub>6</sub>RsmN demonstrates a further drop in the protease levels to 0.2.

The protease production in the *rsmA* mutant is half that of the wild type, 0.3 compared with 0.15 OD<sub>400nm</sub> (Fig. 4.18B). Complementation with RsmN sees a minimal increase but the complementation with the histidine tagged protein returns the protease levels to the wild type.



**Figure 4.18: Impact of RsmN on exoprotease.**

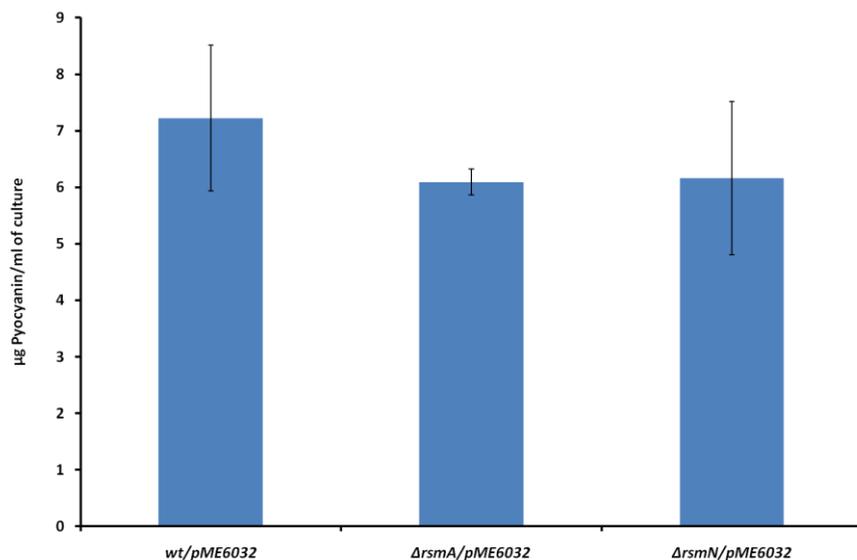
Panel (A) compares the *rsmN* mutant strains and panel (B) compares the *rsmA* mutant strains. The supernatants of each strain were incubated with 5 mg of azocasein for 15 min at 37 °C. The reaction was stopped by the addition of 10 % Trichloroacetic acid (TCA) and the optical density of the supernatant read after centrifugation at 400 nm. Technical replicates where N = 3, error bars are ± 1 SDev.

Complementation with pRsmN62A does not restore exoprotease to wild type levels when introduced into the *rsmA* mutant in contrast to pH<sub>6</sub>RsmN. A comparison by complementation with pH<sub>6</sub>RsmNR62A would also be required for further work. These data are comparable with that obtained for swarming motility in that RsmN has little effect on exoprotease unless histidine-tagged.

#### 4.2.3.4.3 Pyocyanin Assay

RsmA negatively regulates pyocyanin, a virulence factor of *P. aeruginosa* (Pessi et al., 2001), therefore a larger quantity of pyocyanin is produced in strain PAZH13 in comparison with wild type PAO1. To determine whether *rsmN* has a role in the regulation of pyocyanin, experiments were conducted using the wild type PAO1, PAZH13 ( $\Delta rsmA$  mutant), PALT16 ( $\Delta rsmN$  mutant) and the relevant complementing plasmids.

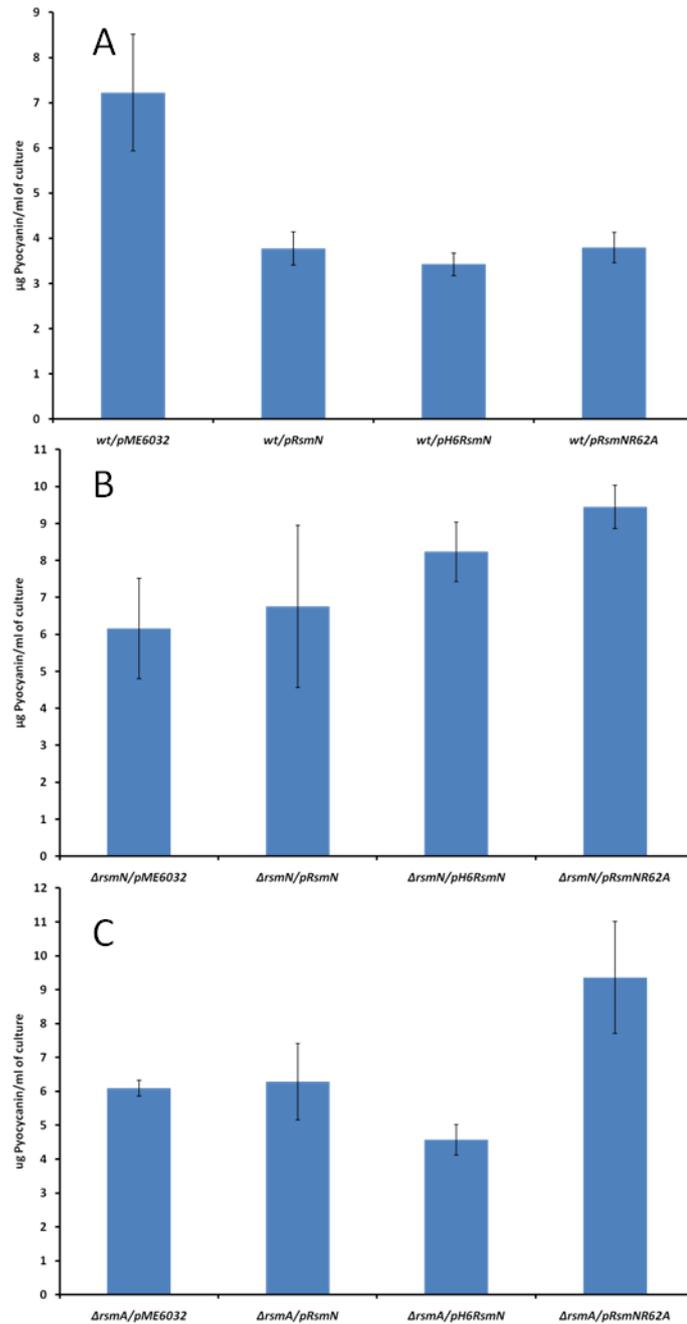
The quantities of pyocyanin/ml found in the wild type PAO1 strain were compared with the *rsmA* and *rsmN* mutants with the empty plasmid pME6032 (Fig. 4.19). Both mutants appear to have reduced levels of pyocyanin in relation to the wild type, however the standard deviation error bars for the *rsmN* value is quite large and overlaps with the wild type. The experiments were performed in triplicate.



**Figure 4.19: Pyocyanin production in *rsmA* and *rsmN* mutants.**

Pyocyanin production was assayed for PALT16 ( $\Delta rsmN$  mutant) and PAZH13 ( $\Delta rsmA$  mutant) carrying the empty vector and plotted as pyocyanin measured in  $\mu\text{g/ml}$  bacterial culture using a previously published method (Essar et al., 1990). There is no significant effect on pyocyanin in the mutant strains in comparison to the wild type. Technical replicates where  $N = 3$ , error bars are  $\pm 1$  SDev.

The levels of pyocyanin in the wild type PAO1, PAZH13 ( $\Delta rsmA$  mutant) and PALT16 ( $\Delta rsmN$  mutant) with the overexpressing *rsmN* variants are compared (Fig. 4.20). In the wild type, when *rsmN* is overexpressed, the level of pyocyanin is reduced by two-fold (Fig. 4.20A). The overexpression of the other *rsmN* variants has no further effect. The overexpression of the RsmN variants in the *rsmN* mutant strain stimulates a rise in pyocyanin levels (Fig. 4.20B). Whereas the pyocyanin level of the empty vector is comparable with the  $\Delta rsmN$ /pRsmN variant, transformation with pH<sub>6</sub>RsmN causes an increase of a third and the addition of the R62A mutant by a half. There is no change in the level of pyocyanin when pRsmN is overexpressed in the *rsmA* mutant (Fig. 4.20C), but there is a reduction by ~30 % when the H<sub>6</sub>RsmN plasmid is present. Overexpression of the R62A mutant causes an increase by 50 % in comparison with both the *rsmA* mutant with or without overexpressing pRsmN. Therefore the conclusions are that while pRsmN has no effect on the *rsmA* and *rsmN* mutants, both the pH<sub>6</sub>RsmN and pRsmNR62A variants causes a change in the pyocyanin production. This provides further evidence that the histidine tagged RsmN in contrast to the native untagged RsmN exerts a minor effect on multiple virulence factors.



**Figure 4.20: Pyocyanin production by of PAO1 wild type,  $\Delta rsmA$  and  $\Delta rsmN$  mutants complemented with RsmN variants.**

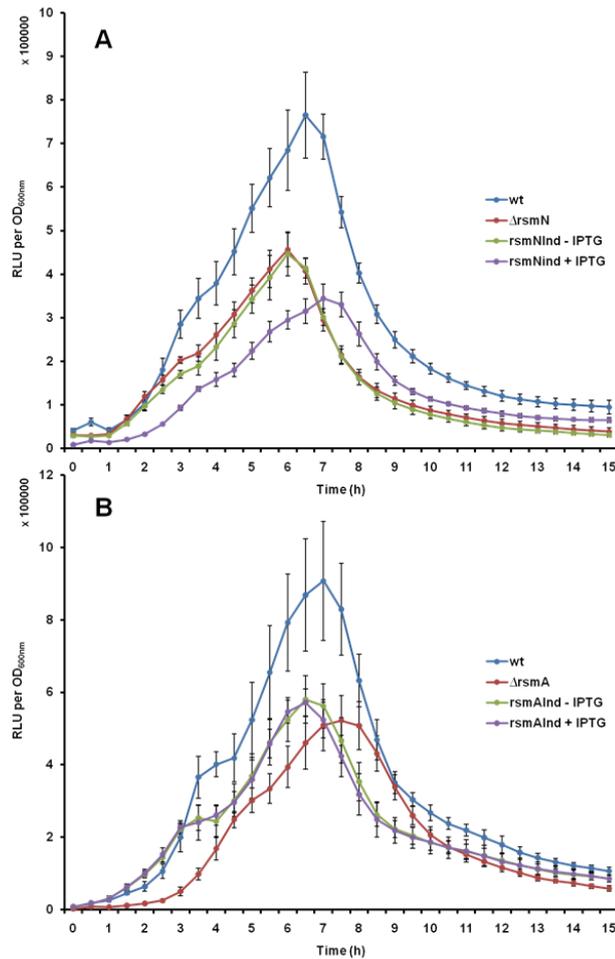
Pyocyanin production was assayed for A: PAO1 (wild type, wt), B: PALT16 ( $\Delta rsmN$  mutant) and C: PAZH13 ( $\Delta rsmA$  mutant) carrying the pRsmN variants and plotted as pyocyanin measured in  $\mu\text{g/ml}$  bacterial culture using a previously published method (Essar et al., 1990). Technical replicates where  $N = 3$ , error bars are  $\pm 1$  SDev.

#### 4.2.4 The influence of RsmN and RsmA on Quorum Sensing (QS)

*P. aeruginosa* possesses two main AHL-dependent quorum sensing systems, the *las* and *rhl* systems which comprise of the LuxRI homologues LasRI (Gambello and Iglewski, 1991) and RhlRI (Ochsner et al., 1994, Latifi et al., 1995) respectively. LasI directs the synthesis of *N*-(3-oxododecanoyl)-L-homoserine lactone (3-oxo-C12-HSL, (Passador et al., 1993, Pearson et al., 1994)) whereas RhlI directs the synthesis of *N*-butanoyl-L-homoserine lactone (C4-HSL, (Winson et al., 1995)) (section 1.2.1.4). In addition to 3-oxo-C12-HSL and C4-HSL, a third QS system exists based in that *P. aeruginosa* releases 2-heptyl-3-hydroxy-4(1*H*)-quinolone, termed the *Pseudomonas* Quinolone Signal (PQS) (Pesci et al., 1999). Transcriptional fusions were made to probe the influence RsmN or RsmA might have on the expression of key genes in these quorum-sensing systems.

##### 4.2.4.1 Influence of RsmN and RsmA on *lasI* transcription

The engineered recombinant transcriptional fusion plasmids underwent chromosomal integration with the PAO1 (wild type, wt), PALT16 ( $\Delta rsmN$  mutant), PAZH13 ( $\Delta rsmA$  mutant), PALT11 (*rsmN* inducible, *rsmN*<sup>Ind</sup>) and PASK10 (*rsmA* inducible strains, *rsmA*<sup>Ind</sup>) to produce the desired transcriptional fusion strains. These were then used to measure growth and bioluminescence over time by monitoring expression of the gene of interest.



**Figure 4.21: Expression of *lasI* in *rsmA* and *rsmN* mutants using chromosomal reporter *lux* fusions.**

A dilution of an o/n culture adjusted to OD<sub>600nm</sub> 1.0 of 1:1000 was used to inoculate sterile LB. The experiment was run in 96 well plates using a GENios Tecan for 15 h at 37 °C measuring OD<sub>600nm</sub> and luminescence. The *lasI* promoter fusions were made in PAO1 (wt), PALT16 ( $\Delta rsmN$ ), PALT11 (*rsmN*Ind), PAZH13 ( $\Delta rsmA$ ) and PASK10 (*rsmA*Ind). Fusions in the *rsmN* strains are shown in panel A and *rsmA* fusions in panel B. Technical replicates where N = 9 and error bars are  $\pm 1$  SDev.

The level of *lasI* expression in the *rsmN* mutant decreased from wild type by over a third, reaching a maximum at 6 h after inoculation, half an hour earlier than wild type, suggesting that RsmN could act positively on the *las* quorum sensing system (Fig. 4.21). Expression is identical in the  $\Delta rsmN$  and in the non-induced *rsmN*<sup>Ind</sup> strains. However upon induction of *rsmN*, the expression of *lasI* drops paradoxically by a factor of >2 (compared with wild type), with the maximum level of expression delayed to 7 h from inoculation. This delay

could be due to the timing and magnitude of *rsmN* expression. The expression of *rsmN* using a  $P_{rsmN}::lux$  promoter follows the same profile in the wild type strain as the inducible strain but with a third greater level of expression (Fig. 5.2). A comparison between the wt and  $\Delta rsmN$  is statistically significant to 5% with a t value at 7 hours of 2.85 using a critical t value of 1.746 (16 DoF).

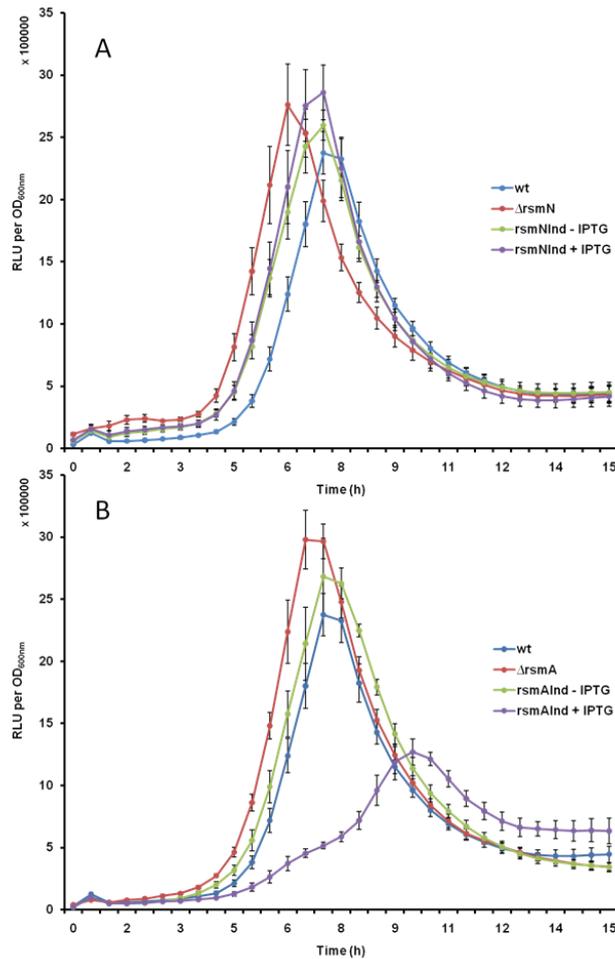
The effect of *rsmA* on *lasI* is shown in Fig. 4.14B. In the *rsmA* mutant, *lasI* expression is reduced and slightly delayed. The uninduced and induced RsmA appears to have little effect on *lasI* expression, with no difference between the inducible strains. This is unexpected as there is a reduction in expression between the wild type and  $\Delta rsmA$  mutant strains. The growth ( $OD_{600}$  nm) of all strains was identical with respect to time. A comparison between the wt and  $\Delta rsmA$  is statistically significant to 5% up to 7.5 hours after inoculation with a t value of 2.16 at 7 hours using a critical t value of 1.746 (16 DoF).

In the literature, it is not yet clear what the role of RsmA has on the *las* QS system as it has previously been reported that *lasI* translation is increased in an RsmA mutant, yet this might not be necessarily reflected by increased transcription (Pessi et al., 2001, Reimann et al., 1997). Both RsmA and RsmN appear to be acting on the transcription of *lasI*, however the method by which this is occurring is unknown.

#### 4.2.4.2 Influence of RsmN and RsmA on *rhlI* transcription

The level of activity of the  $P_{rhlI}$  promoter in the *rsmN* and *rsmA* strains is higher than that of  $P_{lasI}$  by up to a factor of 5 (Fig. 4.22). Both the *rsmN* mutant

and induced  $rsmN^{Ind}$  overexpression strains demonstrate a very slight increase in the level of expression of  $rhII$  compared to the wild type. Under these growth conditions, RsmN has no impact on  $rhII$  expression as shown by the minimal differences in expression between strains.



**Figure 4.22: Expression of  $rhII$  in  $rsmA$  and  $rsmN$  mutants using chromosomal reporter  $lux$  fusions.**

A dilution of an o/n culture adjusted to  $OD_{600nm}$  1.0 of 1:1000 was used to inoculate sterile LB. The experiment was run in 96 well plates using a GENios Tecan for 15 h at 37 °C measuring  $OD_{600nm}$  and luminescence. The  $rhII$  promoter fusions were made in PAO1 (wt), PALT16 ( $\Delta rsmN$ ), PALT11 ( $rsmN^{Ind}$ ), PAZH13 ( $\Delta rsmA$ ) and PASK10 ( $rsmA^{Ind}$ ). Fusions in the  $rsmN$  strains are shown in panel A and  $rsmA$  fusions in panel B. Technical replicates where  $N = 8$  and error bars are  $\pm 1$  SDev.

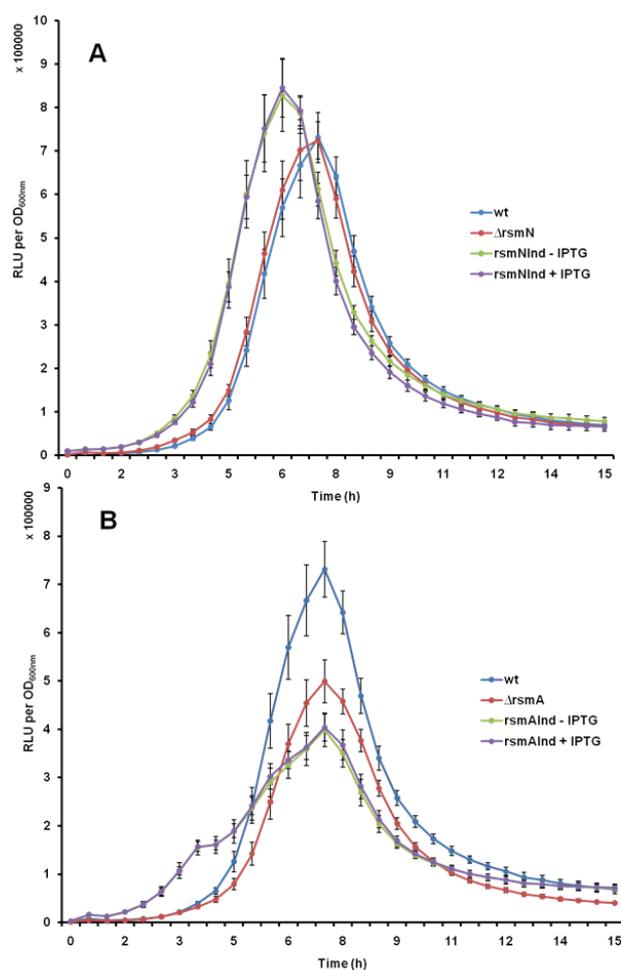
The level of  $rhII$  is slightly elevated in the  $rsmA$  mutant and non-induced  $rsmA^{Ind}$  strain, showing an increase in expression compared to the wild type.

Upon induction of *rsmA* with IPTG, the levels of expression of *rhII* decreased by more than 50 %. The comparison of the wt with the induced *rsmA<sup>Ind</sup>* strain is statistically significant to 5% with a t value at 8 hours of 22.35 using a critical t value of 1.753 (15 DoF).

The growth (OD<sub>600</sub> nm) of all strains was identical with respect to time. This is consistent with previous reports that RsmA is a negative regulator of *rhII*, particularly when *rsmA* is overexpressed. It has been suggested that the mechanism by which RsmA inhibits *rhII* translation is by binding directly to its mRNA transcript (Kay et al., 2006, Pessi and Haas, 2000, Pessi et al., 2001).

#### 4.2.4.3 Influence of RsmN and RsmA on *pqsA* transcription

Deletion of *rsmN* appears to have no effect on the level of *pqsA* expression compared with wild type (Fig. 4.23). The level increases in the inducible *rsmN* mutant by ~10%, with the peak expression occurring an hour earlier. There is no observed difference in expression of *pqsA* between the *rsmN<sup>Ind</sup>* strains prior or after induction. The effect of the *rsmA* mutant on *pqsA* is more striking, with a reduction by ~30%. This expression level is reduced further in the *rsmA* inducible strain to ~50% that of the wild type. The expression of *pqsA* seems to be bi-modal with increases at both 3.5 and 6 h after inoculation. This is not due to differences in growth between the strains and therefore could be indicative of other factors positively regulated by the over-expression of RsmA that have a subsequent effect on *pqsA*. A comparison of the wt and  $\Delta$ *rsmA* is statistically significant to 5% with t value at 7 hours of 4.48 using a critical t value of 1.753 (15 DoF).



**Figure 4.23: Expression of *pqsA* in *rsmA* and *rsmN* strains using chromosomal reporter *lux* fusions.**

A dilution of an o/n culture adjusted to OD<sub>600nm</sub> 1.0 of 1:1000 was used to inoculate sterile LB. The experiment was run in 96 well plates using a GENios Tecan for 15 h at 37 °C measuring OD<sub>600nm</sub> and luminescence. The *pqsA* promoter fusions were made in PAO1 (wt), PALT16 ( $\Delta rsmN$ ), PALT11 (*rsmNInd*), PAZH13 ( $\Delta rsmA$ ) and PASK10 (*rsmAInd*). Fusions in the *rsmN* strains are shown in panel A and *rsmA* fusions in panel B. Technical replicates where N = 8 and error bars are  $\pm 1$  SDev.

It has previously been reported that levels of transcription of the *pqsABCDE* operon, which encodes enzymes required for PQS biosynthesis, did not appear to be altered in microarray analysis of PAO1 wild type compared to the *rsmA* mutant. This result was validated using a *pqsA-lacZ* transcriptional fusion which confirmed there was no significant difference in the transcription of the *pqsABCDE* operon (Burrowes et al., 2006). Although the results in this thesis

show that the *rsmA* mutant has lower levels of *pqsA* expression compared to the wild type, this is not supported by the inducible strain results, with no difference in *pqsA* expression with or without RsmA.

#### 4.2.4.4 Influence of *lasR*, *rhlR* and QS signalling molecules on *rsmN* expression

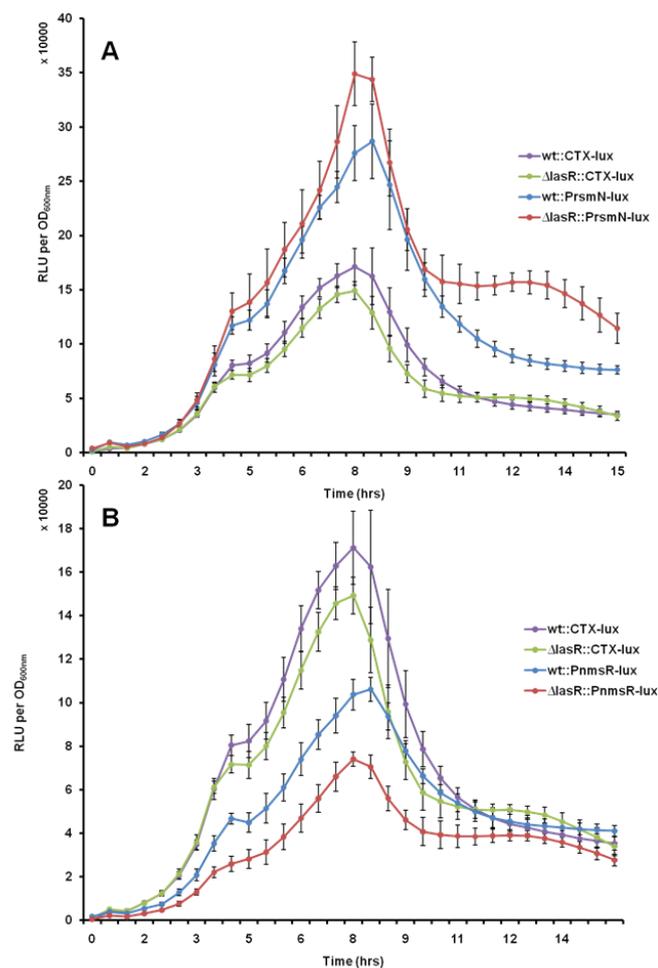
LasR and RhlR exist in a hierarchy where by LasR/3-oxo-C12-HSL regulates the transcription of *rhlR* and consequently both systems are required for the regulation for many virulence determinants. Transcriptional fusions were made to probe the influence of the QS systems upon the expression of *rsmN* and *nmsR*.

##### 4.2.4.4.1 Influence of LasR on *rsmN* and *nmsR* transcription

To determine whether *lasR* has an effect on the expression of *rsmN*, transcriptional reporter fusions were made in a *lasR* mutant strain. All strains labelled in the figures as wt::CTX-*lux*,  $\Delta$ *lasR*::CTX-*lux*,  $\Delta$ *rhlR*::CTX-*lux* or  $\Delta$ *pqsA*::CTX-*lux* are negative controls which contain the miniCTX::*lux* reporter without a promoter inserted in the chromosome.

The PAO1 and *lasR* mutant strains with the empty miniCTX::*lux* promoter fusions were run as controls. The  $P_{rsmN}$ -*lux*' fusions in the *lasR* mutant strain show a slight increase in expression of *rsmN* compared with the wild type by a sixth (Fig. 4.24). *lasR* has a minimal and probably insignificant effect as a repressor of *rsmN* transcription which is confirmed with a t value at 8 hours of 0.78 ( $P_{rsmN}$ ) using a critical t value of 1.734 (18 DoF).

The levels of expression are lower by a factor of two in the *nmsR* promoter fusions compared to that of the  $P_{rsmN}$  fusions, with a reduction in the transcription of *nmsR* by nearly a third in the *lasR* mutant compared with the wild type fusion. Therefore *lasR* has a moderate effect as an activator of *nmsR* and is statistically significant to 5% with a t value at 8 hours of 6.51 using a critical t value of 1.734 (18 DoF).

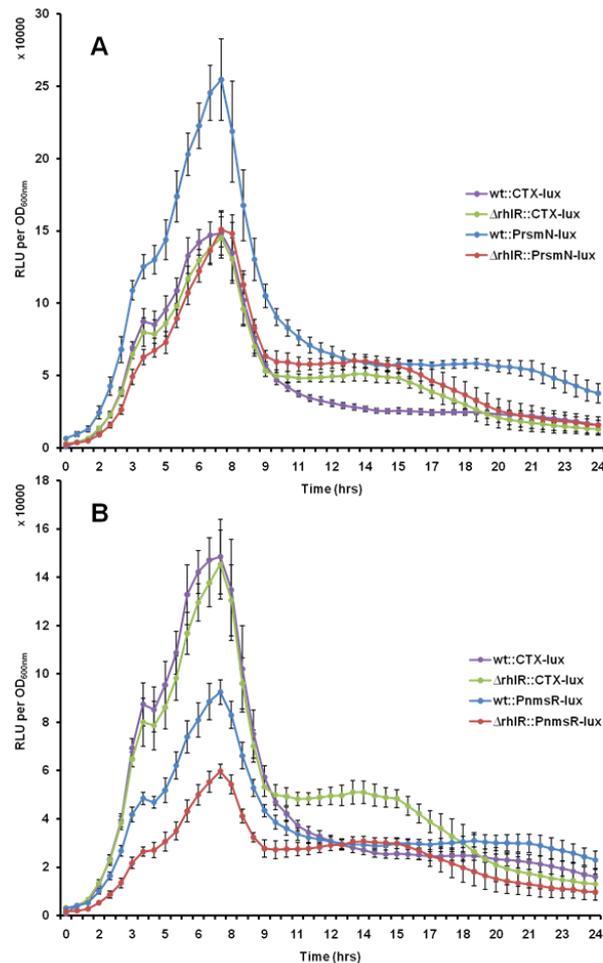


**Figure 4.24: Impact of LasR on the expression of *rsmN* and *nmsR*.**

A dilution of an o/n culture adjusted to  $OD_{600nm}$  1.0 of 1:1000 was used to inoculate sterile LB. The experiment was run in 96 well plates using a GENios Tecan for 15 h at 37 °C measuring  $OD_{600nm}$  and luminescence. The *rsmN* and *nmsR* promoter fusions were made in PAO1 (wt), PASDP233 ( $\Delta lasR$ ). Fusions using the *rsmN* promoter are shown in panel A and *nmsR* promoter fusions in panel B. Technical replicates where  $N = 10$  and error bars are  $\pm 1$  SDev.

#### 4.2.4.4.2 Influence of RhIR on *rsmN* transcription

The  $P_{rsmN}$ -*lux'* fusions in the *rhIR* mutant strain show a 50 % reduction of transcription from the *rsmN* promoter compared with the wild type (Fig. 4.25), suggesting that RhIR has a possible effect acting as an activator on the *rsmN* promoter.



**Figure 4.25: Impact of RhIR on the expression of *rsmN* and *nmsR*.**

A dilution of an o/n culture adjusted to OD<sub>600nm</sub> 1.0 of 1:1000 was used to inoculate sterile LB. The experiment was run in 96 well plates using a GENios Tecan for 24 h at 37 °C measuring OD<sub>600nm</sub> and luminescence. The *rsmN* and *nmsR* promoter fusions were made in PAO1 (wt), PACP10 ( $\Delta rhIR$ ), Fusions using the *rsmN* promoter are shown in panel A and *nmsR* promoter fusions in panel B. Technical replicates where N = 10, error bars are  $\pm 1$  SDev.

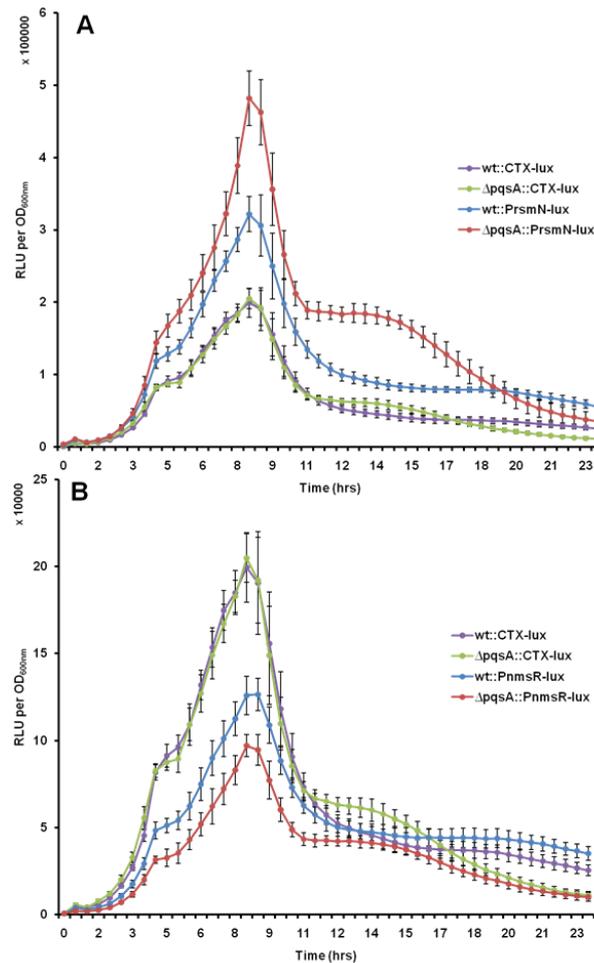
Comparing the  $P_{nmsR}$  fusions, the expression has decreased by 33% from the wild type to the *rhIR* mutant, therefore *rhIR* might also have an effect as a possible activator on the *nmsR* promoter. Both comparisons are statistically

significant to 5% with t values at 7 hours of 4.80 ( $P_{rsmN}$ ) and 5.39 ( $P_{nmsR}$ ) using a critical t value of 1.734 (18 DoF).

#### 4.2.4.4.3 Influence of PQS signalling on *rsmN* expression

The  $P_{rsmN}$ -*lux'* fusions in a *pqsA* mutant strain reveal an increase in the expression of *rsmN* compared to the wild type by ~30 % (Fig. 4.26), the deletion of *pqsA* thus having a moderate positive effect on *rsmN*. The effect of the *pqsA* mutation on the promoter of *nmsR* reveals a decrease in activity by a quarter in the *pqsA* mutant compared to the wild type, therefore *pqsA* has a moderate effect as an activator on the *nmsR* promoter. Both comparisons are statistically significant to 5% with t values at 8 hours of 5.05 ( $P_{rsmN}$ ) and 3.23 ( $P_{nmsR}$ ) using a critical t value of 1.734 (18 DoF).

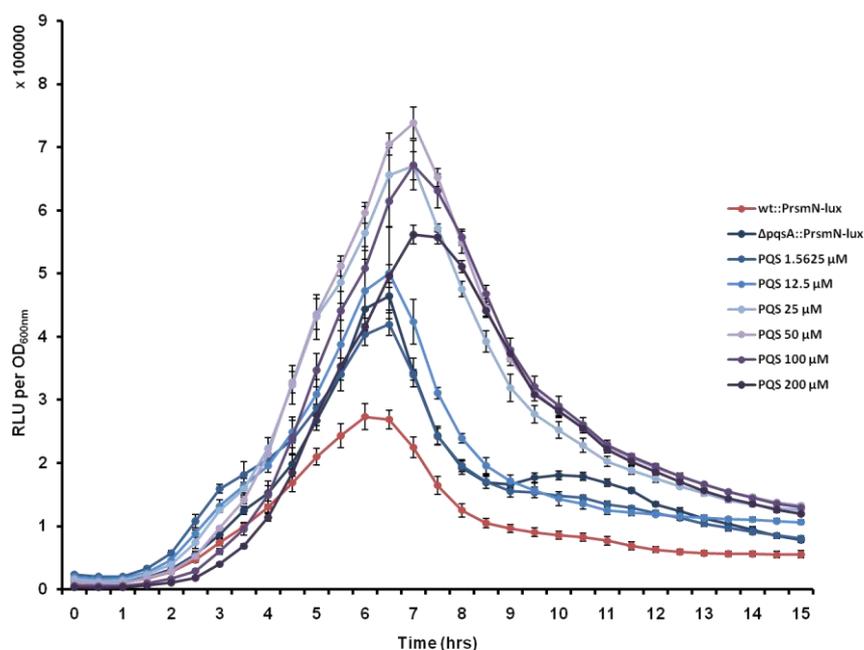
The  $\Delta pqsA$  mutant strain exhibits an increased level of expression of *rsmN* by ~30 %, with a maximum of nearly double after addition of 50  $\mu$ M PQS (Fig. 4.27). The effect of PQS on the expression of *rsmN* in a  $\Delta pqsA$  mutant has a positive affect up to the addition of 50  $\mu$ M PQS and consequent higher PQS concentrations partially reverse this trend, however they remain well above the wild type level. This effect was unexpected as addition of PQS to a *pqsA* mutant would expect levels of expression to fall towards wild type levels, however this is very complex data and would need futher replicates to get more consistent data and increase significance.



**Figure 4.26: Impact of 2-alkyl-4-quinolone signalling on the expression of *rsmN* and *nmsR*.**

A dilution of an o/n culture adjusted to OD<sub>600nm</sub> 1.0 of 1:1000 was used to inoculate sterile LB. The experiment was run in 96 well plates using a GENios Tecan for 24 h at 37 °C measuring OD<sub>600nm</sub> and luminescence. The *rsmN* and *nmsR* promoter fusions were made in PAO1 (wt), PASDP123 ( $\Delta pqsA$ ), Fusions using the *rsmN* promoter are shown in panel A and *nmsR* promoter fusions in panel B. Technical replicates where N = 10 and error bars are  $\pm$  1 SDev.

However mutation of *pqsA*, which is the first enzyme in HHQ biosynthesis (the immediate PQS precursor (Diggle et al., 2006)), results in an increase in *rsmN* expression. Therefore *rsmN* expression may be increased by the action of the other quinolones (HHQ, the AQ N-oxides or dihydroxyquinoline (DHQ)) the synthesis of which depends on *pqsA* or the response regulator PqsE.



**Figure 4.27: *rsmN* expression in a *pqsA* mutant in the presence or absence of PQS.**

A dilution of an o/n culture adjusted to OD<sub>600nm</sub> 1.0 of 1:1000 was used to inoculate sterile LB. The experiment was run in 96 well plates using a GENios Tecan for 15 h at 37 °C measuring OD<sub>600nm</sub> and luminescence. The *rsmN* promoter fusions were made in PAO1 (wt) and PASDP123 ( $\Delta pqsA$ ). A PQS containing solution of a range of concentrations (0-200  $\mu$ M) was added to the inoculated media of the *pqsA* mutant strains. Technical replicates where N = 4 and error bars are  $\pm$  1 SDev.

### 4.3 CONCLUSIONS

The aim of this chapter was to investigate the biological function of RsmN. RsmN was discovered from *in silico* analysis of an intergenic region common to 4 clones found using genomic bank screening (M. Messina, PhD thesis) where the clones were identified as capable of restoring the swarming-deficient phenotype of an *rsmA* mutant. RsmN is a 7.8 kDa protein which shares 34 % identity and 52 % similarity with the 6.9 kDa protein RsmA. The sequence comparison revealed some conserved residues, Arg6, Ala54, Pro55 and Glu64, the corresponding residues of which in RsmA are important for maintenance of structure. The solvent-exposed residue Arg62 was also conserved, where

previous study has shown the corresponding residue in RsmA, R44, is required for retention of biological function (Heeb et al., 2006).

Although both RsmN and RsmA are dimeric proteins, the RsmN dimer forms a clam-like structure. Circular dichroism data confirmed that RsmN has greater alpha helical content and that RsmA has more unstructured polypeptide chain than RsmN.

Transcriptional reporter fusions revealed that transcription around *rsmN* occurred from both, sense and antisense promoters, with the activity of the sense promoter  $P_{rsmN}$  nearly three times that of the antisense promoter  $P_{nmsR}$ .

Identification RsmN by western blot analysis was impossible and it is uncertain if the anti-RsmN antibody was cross-reactive with RsmA. There were multiple reactive bands at higher molecular weights which are probably proteins which are cross-reactive with the RsmN polyclonal antibody or cross-reacting background proteins from the rabbit serum. Any further elucidation from the western blot is not possible due to the concentrations and resolution of the bands of interest. To improve this experiment, the blot could be stripped and re-probed using an anti-RsmA antibody. This could help identify which bands are due to RsmA out of the bands which the anti-RsmN antibody detected.

RsmN had no obvious effect on the transcription of *lasI*, *pqsA* and *rhII* under the growth conditions employed Results suggest RsmA acts as a concentration dependent regulator of *rhII*, however the effect of RsmA on *pqsA* is unclear. The results show that the *rsmA* mutant has lower levels of *pqsA* expression compared with the wild type. However this is not supported by the inducible

strain results, with no difference in *pqsA* expression with or without RsmA. Repeating the experiments using a wide range of IPTG concentrations from 0 to 1000  $\mu$ M could help elucidate the effect of RsmN and RsmA at a range of concentrations on the *lasI*, *rhlI* and *psqA* promoter fusions.

Subsequently experiments were undertaken to examine the effect of RhlR, LasR and PqsA on the *rsmN* and *nmsR* promoter fusions. LasR has no significant regulatory effect on the expression of the *rsmN* or *nmsR* promoters, whereas RhlR possibly has a minor effect on *rsmN* transcription.

Mutation of *pqsA*, the first enzyme in AQ biosynthesis results in the loss not only of PQS but also the other AQS including the immediate precursor of PQS, HHQ (which itself a QS signal molecule) as well the AQ N-oxides and DHQ (Heeb et al, 2011). While the *pqsA* mutant exhibited higher *rsmN* expression levels, paradoxically the addition of PQS to the *pqsA* mutant also resulted in enhanced *rsmN* expression. Thus it is possible that the other AQS or the AQ effector protein PqsE may also modulate *rsmN* expression or that the iron chelating properties of PQS (Diggle et al 2007) are responsible for the observed increase in *rsmN* transcription. This could be investigated by examining the impact of HHQ, HQNO and DHQ added exogenously to the *pqsA* mutant or by restricting the iron content of the growth medium.

No evidence could be found that RsmN acts as an RsmA homologue in the swarming assay as this phenotype is not repressed in a  $\Delta$ *rsmN* mutant (Fig. 4.13). The introduction of the hexahistidine tagged RsmN partially restores the swarming activity in the  $\Delta$ *rsmA* mutant and causes hyper swarming when complementing the  $\Delta$ *rsmN* mutant. The insertion of the tag is after the *rsmN*

promoter but prior to the *rsmN* gene. A tag of such a small size would have limited effect on the RsmN protein sterically, but the addition of six basic, polar residues might be relevant. The effect could be due to either a disruption in the transcription of the gene sequence, the tag could be acting as a blocker to external effects from the possible antisense gene *nmsR*, or it could be interfering with the R62 which is sterically positioned close to the histidine tag. Repeating the results together with complementation of a histidine tagged *rsmN* arginine substitution mutant could help elucidate the role of the tag and arginine mutation.

When RsmN is induced using the conditional mutant (*rsmN*<sup>Ind</sup>), there was no effect upon the  $\Delta$ *rsmA* mutant strain and a gradual increase in the degree of swarming in the wild type strain (Fig. 4.14), suggesting that RsmN can enhance swarming but only in the presence of RsmA.

The *rsmN* gene was not capable of complementing the *csrA* mutation in the *E. coli* TR1-5 glycogen accumulation assay. The restriction assay results indicate that RsmN, unlike RsmA, does not control restriction modification in *P. aeruginosa*. However, repeating these experiments would be beneficial in order to compare transformation efficiencies of the inducible *rsmN* strain PALT11 when induced or not induced.

In the elastase assay (Fig. 4.17), when the *rsmA* mutant strains are transformed with the RsmN-containing plasmids, the only strain which is atypical from the wild type allele is that containing the histidine tagged RsmN. Therefore the  $\Delta$ *rsmN* mutant has no effect on elastase activity, however when used to

complement a  $\Delta rsmA$  mutant, the histidine tagged RsmN appears to increase the elastase production.

The protease assay (Fig. 4.18) demonstrates that although a  $\Delta rsmN$  mutation has no effect on protease activity, when complemented by either RsmN or the H<sub>6</sub>RsmN containing plasmids, activity is reduced. Complementation of the  $\Delta rsmN$  with the arginine R62A RsmN mutant restores the activity to the mutant and wild type levels. The protease assay using the  $\Delta rsmA$  mutant strain provides some interesting results. The mutation of *rsmA* results in a reduction in protease activity which is not restored by complementation with RsmN or RsmNR62A, however, activity is restored with H<sub>6</sub>RsmN.

The  $\Delta rsmA$  mutant strain demonstrated that RsmA possibly has a positive regulatory effect on pyocyanin whereas the *rsmN* mutant has no effect. However this effect is probably insignificant due to the overlap of error bars, so although this study was unable to reproduce the published results of a negative effect (Pessi et al., 2001). An improvement would be to repeat the experiment using a glycerol-alanine medium to promote high levels of pyocyanin production (Pessi and Haas, 2000). Complementation of the wild type strain with *rsmN*-containing plasmids causes a decrease in pyocyanin levels, however complementation of the  $\Delta rsmN$  mutant had minimal effect. There is no change in the level of pyocyanin when the  $\Delta rsmA$  mutant was complemented by *rsmN*, but there is a reduction when complemented by the H<sub>6</sub>RsmN plasmid. This provides further evidence that the histidine tagged RsmN has an effect on the activity of RsmA. All of the phenotypic assays would benefit from repeating using a plasmid complementation of a histidine tagged RsmN arginine substitution mutant.

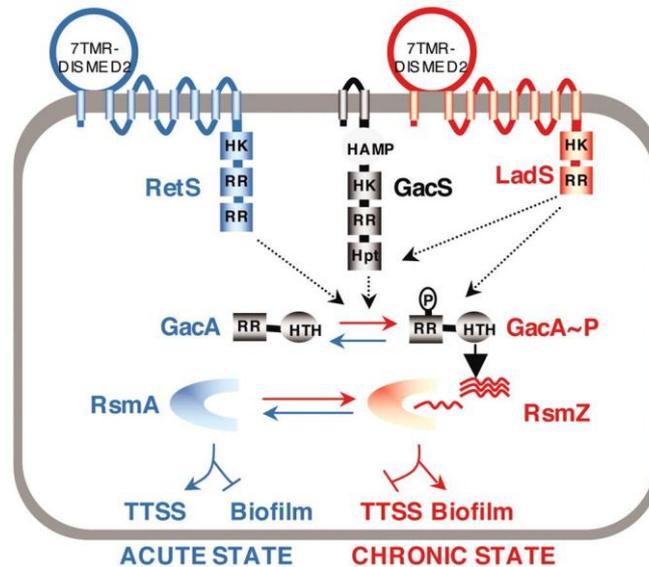
The aim of this chapter to discover a phenotype for *rsmN* which has so far proved elusive, however experiments performed with pH<sub>6</sub>RsmN in conjunction with the  $\Delta rsmA$  mutant strain have yielded some interesting results. It is therefore unlikely that RsmN is involved in the control of any of the phenotypes investigated in this chapter. A different approach to these phenotypic assays is to use chromosomal transcriptional fusions to explore whether RsmN is involved in the Gac signalling cascade.

## 5 RELATIONSHIP BETWEEN RSMN, RSMA, AND THE GAC SYSTEM

### 5.1 INTRODUCTION

RetS (for regulator of exopolysaccharide and type III secretion) and LadS (for lost adherence) are membrane-bound hybrid sensor kinases present in a variety of Pseudomonads (Ventre et al., 2006, Humair et al., 2009, Records and Gross, 2010). Deletion of RetS results in the overexpression of the *pel* and *psl* genes required for the formation of polysaccharides and biofilm development. Strains having mutations in *retS* are unable to respond to host-cell contact or media-derived signals that normally activate the expression of genes encoding the type III secretion system (TTSS). RetS has been implicated as a regulator of bacterial behaviour during infection due to this reciprocal relationship between TTSS expression and biofilm formation. RetS and LadS share domain organisation and downstream targets, but act in a reciprocal manner on a shared set of positively and negatively regulated virulence determinants. Both signalling pathways function by influencing the levels of the small regulatory RNAs RsmY and RsmZ by regulating the cascade at the level of GacA phosphorylation. It has been found that RetS inhibits and LadS activates the activity of the Gac pathway, but the mechanisms by which these sensors communicate with one another and subsequently determine the output of the system are not known (Ventre et al., 2006). There is however evidence that both RetS and LadS physically interact with GacS (Workentine et al., 2009, Goodman et al., 2009). If LadS and subsequently GacA are activated in the

signal cascade, the latter increases *rsmZ* transcription, which leads to more RsmA being sequestered.



**Figure 5.1: A model for the convergence of the signalling pathways during reciprocal regulation of virulence factors by LadS, RetS, and GacS through transcription of the small regulatory RNA RsmZ (Ventre et al., 2006).**

The three sensors are anchored into the cytoplasmic membrane via their transmembrane domains. Unknown signals received by the input domains (7TMR-DISMED2 and HAMP) of the sensor kinases activate or repress the expression of genes specifying factors necessary for acute or chronic infection. The signalling cascade going through RetS and resulting in TTSS activation and biofilm repression is represented in blue. The signalling cascade going through LadS and resulting in TTSS repression and biofilm activation is represented in red. The small RNA RsmZ is represented by a curved line, which can form a complex with RsmA, resulting in biofilm.

The expression of the Rsm system RNAs is therefore potentially regulated by at least three different regulatory systems which can probably respond to and integrate at least three different signals.

Therefore this chapter focuses on using chromosomal transcriptional fusions to explore whether the newly identified RsmA homologue RsmN is involved in this signalling cascade, and if control of RsmN is exerted by RsmA.

## 5.2 RESULTS AND DISCUSSION

### 5.2.1 Strains constructed in this Chapter

These strains were constructed in order to use chromosomal transcriptional fusions to help elucidate if RsmA and RsmA have an effect upon each other and also what affect the Gac signalling pathway has on *rsmN* and *nmsR* by constructing Rets, LadS and GacA transcriptional fusions.

#### 5.2.1.1 Mini-CTX::*lux* promoter fusions

The sense and antisense promoter fusions, P<sub>*rsmN*</sub> (pLT1) and P<sub>*nmsR*</sub> (pLT2), were constructed as described in sections 2.4.1.8 and 2.4.1.9 respectively. For pLT1, the primers RSMNPF1 and RSMNPR1 were used to amplify a 331 bp product from the PAO1 wild type Lausanne genome with part of the sense promoter and flanking *XhoI* and *PstI* restriction sites (Section 4.2.1.1, Fig. 4.10). This was repeated with the primers RSMNPF2 and RSMNPR2 to produce a 452 bp product with part of the antisense promoter and flanking *HindIII* and *EcoRI* restriction sites for pLT2. The mini-CTX::*lux* plasmid was then linearised with the required enzymes and the relevant product inserted. Following ligation the DNA was transformed into *E. coli* S17-1  $\lambda$ pir cells.

#### 5.2.1.2 Construction of *gacA* mutant PALT40

A *gacA* mutant was made using a two step homologous recombination procedure where the suicide plasmid pME6111 (Reimmann et al., 1997) underwent conjugation with recipient PAO1, inserting an omega cassette into the *gacA* gene. The suicide plasmid pME6111 was maintained in *E. coli* S17-1

$\lambda$ pir, which also supplies the *tra* genes for efficient mobilisation into *P. aeruginosa*.

### 5.2.1.3 Chromosomal transcriptional fusions

Chromosomal fusions were made (section 2.8.5.1). by conjugation of pLT1 and pLT2 donors in *E. coli* S17-1  $\lambda$ pir for delivery into the chromosome of the recipient strain.

### 5.2.1.4 $P_{rsmN}$ and $P_{nmsR}$ fusions in *rsmA* and *rsmN* mutants

Promoter fusions using pLT1 and pLT2 were made with the donor strains PAO1 (wild type), PAZH13 (*rsmA* mutant), PASK10 (inducible *rsmA*), PALT16 (*rsmN* mutant) and PALT11 (inducible *rsmN*), resulting in the strains listed in Table 5.1 (taken from Table 2.1.).

**Table 5. 1: Sense and antisense promoter fusions in *P. aeruginosa rsmA* and *rsmN* strains.**

PA Number	Genotype/Characteristics
PALT1	PAO1::(miniCTX:: $P_{rsmN}$ - <i>lux</i> )
PALT2	PAO1::(miniCTX:: $P_{nmsR}$ - <i>lux</i> )
PALT3	PASK10::(miniCTX:: $P_{rsmN}$ - <i>lux</i> )
PALT4	PASK10::(miniCTX:: $P_{nmsR}$ - <i>lux</i> )
PALT5	PALT16::(miniCTX:: $P_{rsmN}$ - <i>lux</i> )
PALT6	PALT16::(miniCTX:: $P_{nmsR}$ - <i>lux</i> )
PALT7	PAZH13::(miniCTX:: $P_{rsmN}$ - <i>lux</i> )
PALT8	PAZH13::(miniCTX:: $P_{nmsR}$ - <i>lux</i> )
PALT34	PALT11::(miniCTX:: $P_{rsmN}$ - <i>lux</i> )
PALT35	PALT11::(miniCTX:: $P_{nmsR}$ - <i>lux</i> )

#### 5.2.1.5 $P_{rsmN}$ and $P_{nmsR}$ fusions in $\Delta retS$ mutant

Promoter fusions using pLT1, pLT2 and the empty mini-CTX::*lux* plasmid were made with the donor strains PAO1 and PAKR52 ( $\Delta retS$  mutant), resulting in the strains listed in Table 5.2 (taken from Table 2.1.).

**Table 5. 2: Sense and antisense promoter fusions in PAO1 and  $\Delta retS$  strains.**

PA Number	Genotype/Characteristics	Comment
PAKR52	$\Delta retS$ in frame deletion mutant	
PALT36	PAKR52::(miniCTX:: $P_{rsmN}$ - <i>lux</i> )	<i>rsmN</i> promoter fusion in $\Delta retS$
PALT37	PAKR52::(miniCTX:: $P_{nmsR}$ - <i>lux</i> )	<i>nmsR</i> promoter fusion in $\Delta retS$
PALT41	PAO1::(miniCTX:: <i>lux</i> )	Empty miniCTX:: <i>lux</i> in wild type
PALT42	PAKR52::(miniCTX:: <i>lux</i> )	Empty miniCTX:: <i>lux</i> in $\Delta retS$

#### 5.2.1.6 $P_{rsmN}$ and $P_{nmsR}$ fusions in $\Delta ladS$ mutant

Promoter fusions using pLT1, pLT2 and the empty mini-CTX::*lux* plasmid were made with the donor strains PAO1 and PAKR45 ( $\Delta ladS$  mutant), resulting in the strains displayed in Table 5.3 (taken from Table 2.1.).

**Table 5. 3: Sense and antisense promoter fusions in PAO1 and  $\Delta ladS$  strains.**

PA Number	Genotype/Characteristics	Comment
PAKR45	$\Delta ladS$ in frame deletion mutant	
PALT38	PAKR45::(miniCTX:: $P_{rsmN}$ - <i>lux</i> )	<i>rsmN</i> promoter fusion in $\Delta ladS$
PALT39	PAKR45::(miniCTX:: $P_{nmsR}$ - <i>lux</i> )	<i>nmsR</i> promoter fusion in $\Delta ladS$
PALT41	PAO1::(miniCTX:: <i>lux</i> )	Empty miniCTX:: <i>lux</i> in wild type
PALT43	PAKR45::(miniCTX:: <i>lux</i> )	Empty miniCTX:: <i>lux</i> in $\Delta ladS$

#### 5.2.1.7 $P_{rsmN}$ and $P_{nmsR}$ fusions in $\Delta gacA$ mutant

Promoter fusions using pLT1, pLT2 and the empty mini-CTX::*lux* plasmid were made with the donor strains PAO1 and PALT40 ( $\Delta gacA$  mutant), resulting in the strains displayed in Table 5.4 (taken from Table 2.1.).

**Table 5. 4: Sense and antisense promoter fusions in PAO1 and  $\Delta gacA$  strains.**

PA Number	Genotype/Characteristics	Comment
PALT40	$\Delta gacA$ :: $\Omega$ Sm/Sp mutant	
PALT41	PAO1::(miniCTX:: <i>lux</i> )	Empty miniCTX:: <i>lux</i> in wild type
PALT58	PALT40::(miniCTX:: <i>lux</i> )	Empty miniCTX:: <i>lux</i> in $\Delta gacA$
PALT59	PALT40::(miniCTX::P <sub><i>rsmN</i></sub> - <i>lux</i> )	<i>rsmN</i> promoter fusion in $\Delta gacA$
PALT62	PALT40::(miniCTX::P <sub><i>nmsR</i></sub> - <i>lux</i> )	<i>nmsR</i> promoter fusion in $\Delta gacA$

## 5.2.2 Impact of RsmA and RsmN on *rsmN* and *nmsR* expression

### 5.2.2.1 The control of expression of *rsmN* and *nmsR* by RsmA and RsmN

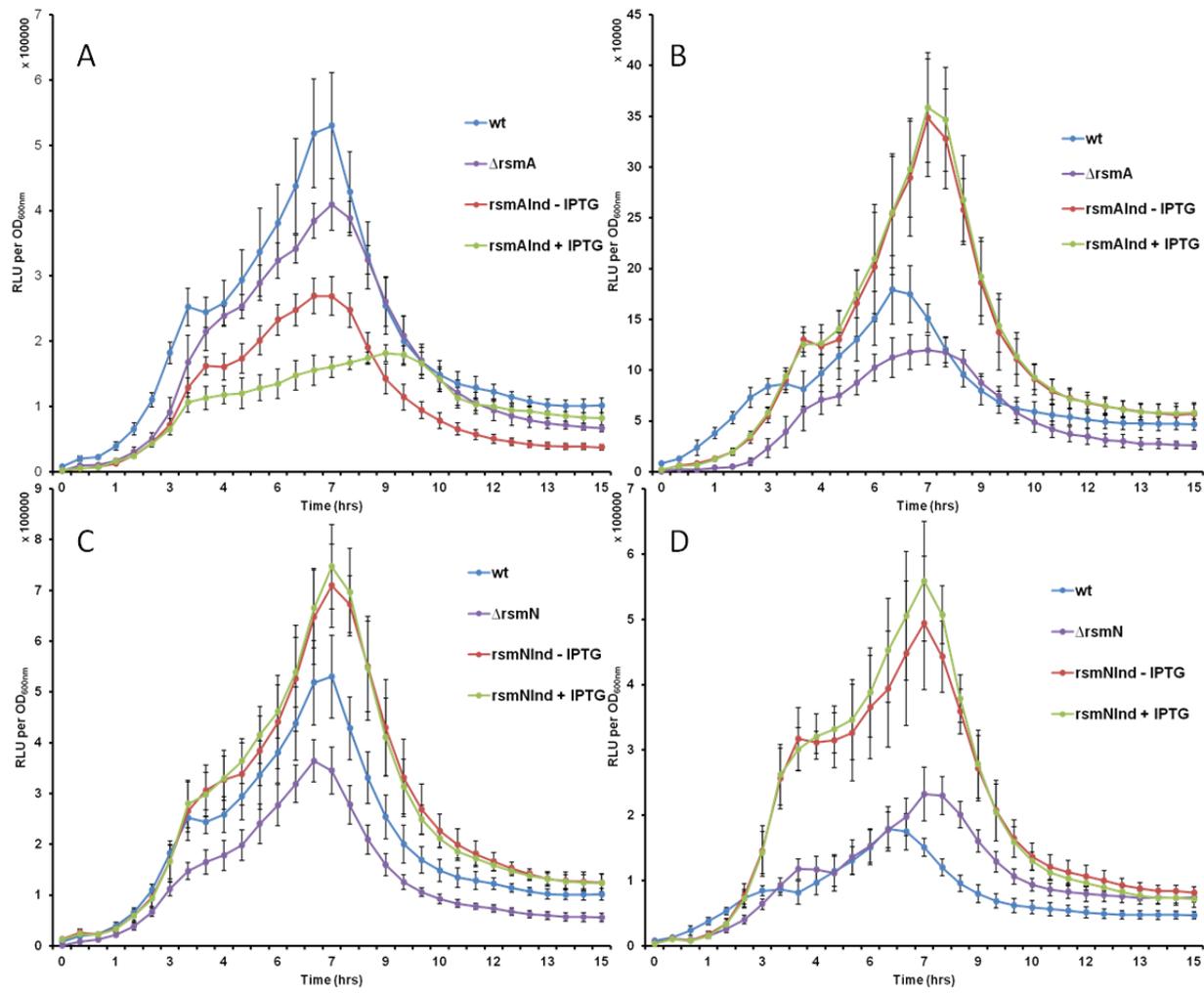
The activity of the *rsmN* promoter is reduced in the *rsmA* mutant by ~20 % compared with that of the wild type (Fig. 5.2A). The activity in the inducible *rsmA* strain is even lower when not induced, to ~50 % that of the wild type. These results suggest RsmA acts as a positive regulator of *rsmN* transcription, likely in an indirect manner. Over-production of RsmA results in further reduction in transcription of the *rsmN* promoter, with a delay to maximum expression of the reporter to 9 h after inoculation compared to the wild type maximum at 7 h. This delay in expression is not due to a growth effect.

Expression levels of the antisense *nmsR* promoter in the *rsmA* strains are all reduced by a factor of 10 compared with the *rsmN* reporter (Figs. 5.2 A and B). In this case the absence of *rsmA* expression again triggers a decrease in the activity of the *nmsR* promoter compared to the wild type. However, the inducible strain, with or without the overexpression of *rsmA*, exhibits a two-fold increase in *nmsR* transcription compared with that of the wild type strain. This supports the role of RsmA as a positive regulator of *nmsR*. Although both the *rsmN* and *nmsR* promoters appear to act under a positive effect of RsmA, the expression levels of the *rsmN* reporter reaches levels twice that of *nmsR*.

Both comparisons between the wt and  $\Delta$ rsmA are statistically significant to 5% with t values at 7 hours of 2.12 ( $P_{rsmN}$ ) and 2.24 ( $P_{nmsR}$ ) using a critical t value of 1.753 (15 DoF).

The experiment would need to be repeated with a range of concentrations of IPTG added to PASK10, the inducible RsmA strain to try to understand the effect of RsmA on *rsmN* and check that the PASK10 strain is not leaky for rsmA expression. A direct comparison of the wild type strain with the *rsmA* inducible strain also presents a problem. The induced strains were not included in the statistical calculations due to the necessity of repeats. The induction of *rsmA* in PASK10 is at time point 0 h, whereas in the wild type the expression of *rsmA* is controlled. Initial expression is low with a three-fold enhancement in the stationary phase, therefore a time-dependent induction of *rsmA* should be examined.

The effect of RsmN on the activity of the *rsmN* and *nmsR* promoters (Fig. 5.2C and D) follows the same pattern as the effect of RsmA on *nmsR* expression (Fig. 5.2B). A small reduction in expression is observed in the *rsmN* mutant from wild type levels which is restored to greater than that of wild type in the *rsmN* inducible strains. The expression of *nmsR* and *rsmN* in the induced RsmN strain is more than 2-fold that of the wild type. As previously mentioned, there is no difference in expression between the uninduced and induced RsmN strains. The comparison between the wt and  $\Delta$ rsmN are statistically significant to 5% with t values at 7 hours of 2.35 for  $P_{rsmN}$  and not significant (0.80) for  $P_{nmsR}$  using a critical t value of 1.746 (16 DoF). As both experiments have this feature, no convincing conclusion can be made.

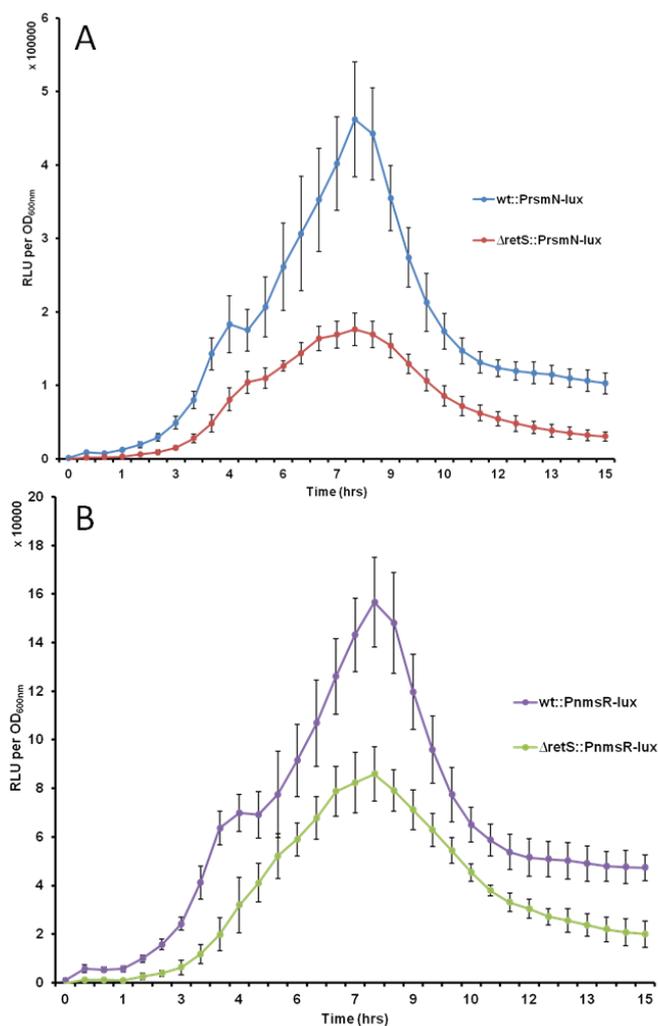


**Figure 5.2: Effect of RsmA and RsmN on the *rsmN* (A and C) and *nmsR* (B and D) promoters.** A dilution of an o/n culture adjusted to OD<sub>600</sub> 1.0 of 1:1000 was used to inoculate sterile LB. The experiment was run in 96 well plates using a GENios Tecan for 15 h at 37 °C measuring OD<sub>600</sub> and luminescence. The *rsmN* and *nmsR* promoter fusions were made in PAO1 (wt), PAZH13 ( $\Delta$ *rsmA*), PALT16 ( $\Delta$ *rsmN*), PASK10 (*rsmA*Ind) and PALT11 (*rsmN*Ind). Panel A: effect of RsmA on *rsmN* expression, B: effect of RsmA on *nmsR* expression, C: effect of RsmN on *rsmN* expression and D: effect of RsmN on *nmsR* expression. The variations of IPTG concentration and timing of induction would need to be repeated with these strains. Technical replicates where N = 8 and error bars are ± 1 SDev.

### 5.2.3 Impact of *retS*, *ladS* and *gacA* on *rsmN*

To elucidate whether there is a link between *rsmN* and the *gac* system, fusions using the *rsmN* and *nmsR* promoters were made in PAO1 (wild type, wt), PAKR54 ( $\Delta retS$  in frame deletion mutant), PAKR45 ( $\Delta ladS$  in frame deletion mutant) and PALT40 ( $\Delta gacA:\Omega Sm/Sp$  mutant) strains.

#### 5.2.3.1 Impact of RetS on *rsmN* and *nmsR* transcription



**Figure 5.3: Effects of RetS on the *rsmN* (A) and *nmsR* (B) promoters.**

A dilution of an o/n culture adjusted to OD<sub>600</sub> 1.0 of 1:1000 was used to inoculate sterile LB. The experiment was run in 96 well plates using a GENios Tecan for 15 h at 37 °C measuring OD<sub>600</sub> and luminescence. The *rsmN* and *nmsR* promoter fusions were made in PAO1 (wt) and PAKR52 ( $\Delta retS$ ). Technical replicates where N = 10 and error bars are  $\pm 1$  SDev.

The mutation of *retS* resulted in an ~2-fold reduction in the expression of both *rsmN* and *nmsR*, indicating RetS has a significant effect acting as an activator on both promoters (Fig. 5.3A and B). Both comparisons are statistically significant to 5% with t values at 8 hours of 4.98 ( $P_{rsmN}$ ) and 4.61 ( $P_{nmsR}$ ) using a critical t value of 1.734 (18 DoF).

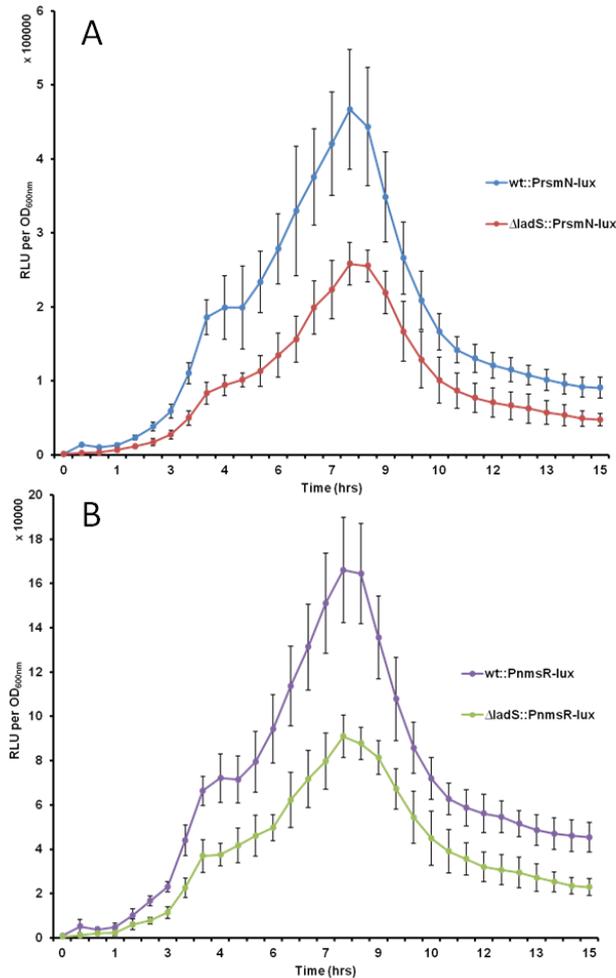
#### 5.2.3.2 Impact of LadS on *rsmN* and *nmsR* expression

Near identical behaviours of the *rsmN* and *nmsR* promoters in the *LadS* mutant (Fig. 5.4) as in the *RetS* mutant strains were observed (Fig. 5.3). The  $\Delta ladS$  mutation resulted in a ~2-fold reduction in the expression of *rsmN* and *nmsR*, indicating LadS has a significant effect acting as an activator on both the promoters (Fig. 5.4A and B). Therefore both RetS and LadS appear to act as activators of the *rsmN* and *nmsR* promoters, which contradicts publications that RetS and LadS act differentially (Ventre et al., 2006). If RsmN was acting as an RsmA homologue in this situation, expression of *rsmN* would be expected to increase when RetS is activated. This would subsequently inhibit both GacA and RsmZ, leading to an increase in RsmA. However by the same hypothesis, *rsmN* expression would decrease when LadS and GacA are activated, increasing *rsmZ* transcription leading to more RsmA being sequestered.

Both comparisons are statistically significant to 5% with t values at 8 hours of 3.44 ( $P_{rsmN}$ ) and 4.13 ( $P_{nmsR}$ ) using a critical t value of 1.734 (18 DoF).

The strains carrying a  $P_{rsmN}$ -*luxCDABE* reporter produce bioluminescence, however removing an activator will not necessarily lead to a decrease in the

activity of a promoter is down to the point that absolutely no bioluminescence is ever made (if it is even made in the first place), and inversely, removing a repressor may not cause a nuclear explosion.



**Figure 5.4: Effect of *LadS* on the *rsmN* (A) and *nmsR* (B) promoters.**

A dilution of an o/n culture adjusted to  $OD_{600}$  1.0 of 1:1000 was used to inoculate sterile LB. The experiment was run in 96 well plates using a GENios Tecan for 15 h at 37 °C measuring  $OD_{600}$  and luminescence. The *rsmN* and *nmsR* promoter fusions were made in PAO1 (wt), PAKR45 ( $\Delta$ *ladS*). Technical replicates where  $N = 10$  and error bars are  $\pm 1$  SDev.

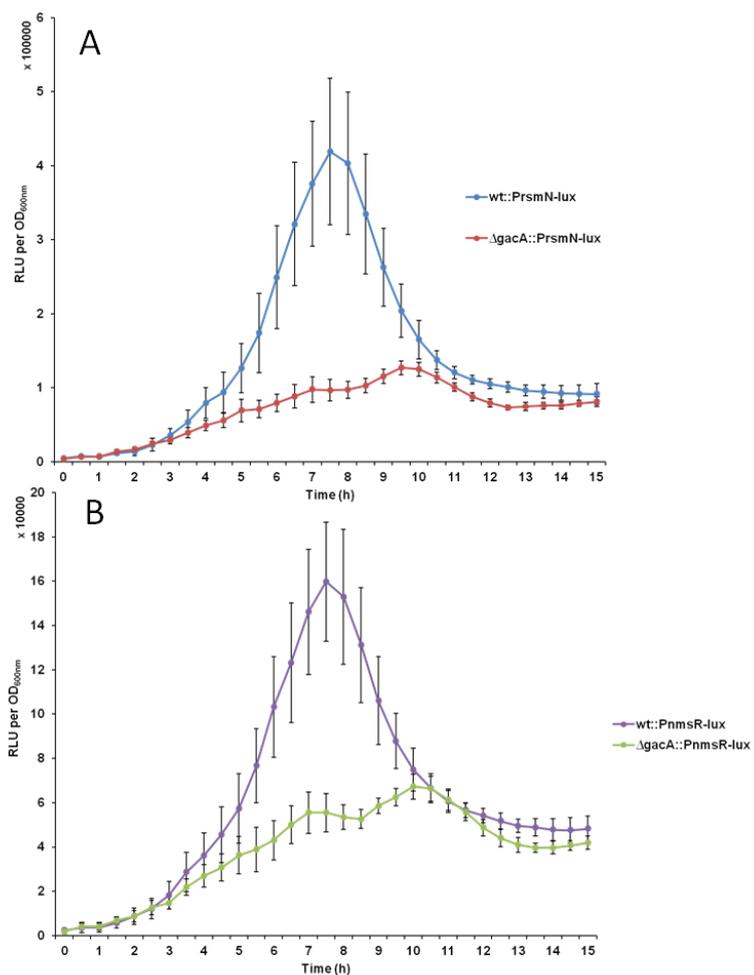
Sometimes removing a repressor will not have any dramatic effect because, under the conditions of the experiment, the system might have been nearly totally derepressed. In that special case what needs to be done is to overexpress

the suspected repressor and if it is indeed a repressor that should dramatically decrease the activity of the promoter (but never to an absolute zero level).

### 5.2.3.3 Impact of GacA on *rsmN* and *nmsR* expression

The behaviour of the *rsmN* and *nmsR* promoters in the *GacA* mutant (Fig. 5.5) reproduces those already seen in the *RetS* (Fig. 5.3) and *LadS* mutant strains (Fig. 5.4). The  $\Delta gacA$  mutation resulted in a 75 % reduction in the expression of *rsmN* and a 66 % reduction of *nmsR*, indicating GacA has a significant effect acting as an activator on both the promoters (Fig. 5.5A and B). Both comparisons are statistically significant to 5% with t values at 8 hours of 4.46 ( $P_{rsmN}$ ) and 4.13 ( $P_{nmsR}$ ) using a critical t value of 1.734 (18 DoF).

As activation of GacA increases *rsmZ* transcription, therefore leading to more RsmA being sequestered, it would suggest that RsmN does not act as an RsmA homologue with respect to GacA.



**Figure 5.5: Effects of GacA on the *rsmN* (A) and *nmsR* (B) promoters.**

A dilution of an o/n culture adjusted to OD<sub>600</sub> 1.0 of 1:1000 was used to inoculate sterile LB. The experiment was run in 96 well plates using a GENios Tecan for 15 h at 37 °C measuring OD<sub>600</sub> and luminescence. The *rsmN* and *nmsR* promoter fusions were made in PAO1 (wt), PALT40 ( $\Delta$ *gacA*). Technical replicates where N = 10 and error bars are  $\pm$  1 SDev.

### 5.3 CONCLUSIONS

Expression of *rsmN* appears to be weakly affected by the levels of RsmA as expression of an *rsmN-lux* reporter in the  $\Delta rsmA$  mutant, in the uninduced  $rsmA^{Ind}$  and in the induced  $rsmA^{Ind}$  are all reduced compared to that in the wild type.

In the case of the *nmsR* reporter, the mutation of *rsmA* again triggers a decrease in expression of *nmsR* compared to the wild type, however, the inducible strain, without and with overexpression of RsmA, causes a two-fold increase in *nmsR* expression compared to that of the wild type strain. Looking at just the wild type and mutant strains, RsmA appears to be acting as a positive regulator, however both the inducible strain produces different results. If the conditional mutant is leaking expression of *rsmA* in the absence of IPTG or whether there is a concentration-dependent effect of IPTG on both promoters could explain some of these results. Further experiments could be performed using a range of concentrations of IPTG in the conditional mutants could illuminate this situation. However, the expression of *rsmA* in the inducible strain does not directly mirror the kinetics of *rsmA* induction in the wild type since it is induced earlier and at a higher level than in the wild type (Pessi et al., 2001).

The results were very similar for the role of RsmN on expression of *rsmN* and *nmsR* reporters. To elucidate the effect that RsmN has on *rsmA* expression, more transcriptional fusions would need to be constructed in the wild type,  $\Delta rsmN$  mutant and conditional *rsmN* mutant strains with an *rsmA* promoter.

RetS, LadS and GacA all appear to have a significant effect as activators on both the *rsmN* and *nmsR* promoters, which contradicts the theory of RsmN acting as an RsmA homologue as when GacA is activated, subsequently increasing *rsmZ* transcription, more RsmA is sequestered. Repression of GacA results in a decrease of *rsmZ*, increasing the amount of free RsmA. Further elucidation could be obtained by the construction of additional transcriptional fusions, for example looking at the effect of RsmA and RsmN on *rsmZ* and *rsmY* expression.

In *P. fluorescens* CHA0, genetic evidence has indicated that RsmA is not the only negative control element in the GacS/GacA cascade (Blumer et al., 1999). When the chromosomal *rsmA* gene is inactivated in a *gacS* mutant background, the effect of the *gacS* mutation on an *aprA*'-*lacZ* fusion is only partially suppressed. This indicates that RsmA is not the only negative regulator in the Gac/Rsm cascade. Reimmann *et al.*, identified RsmE, a homologue of RsmA and provided evidence that both proteins are required together for maximum translational repression of the GacS/GacA target genes *hcnA* (HCN), *aprA* (AprA) and *phlA* (2,4-diacetylphloroglucinol, antibiotic)(Reimmann et al., 2005). Testing the effect of *rsmA*, *rsmN* and *gacS* mutants, as well as double and triple mutants on target gene expression in a *gacS* mutant background, could provide insight to the effect of RsmN in concert with RsmA.

Obtaining an expression profile of RsmN would be of interest in comparison to RsmA. The observation that RsmE levels were highest at the end of growth in *P. fluorescens* CHA0 suggested that RsmE could play a role in the termination of GacA-controlled gene expression (Reimmann et al., 2005), shows how the knowledge of expression profiles can provide important links.

After the ‘top-down’ approach, a ‘bottom – up’ design can also be used to glean information regarding RsmN. By identifying the sRNAs that RsmN binds to, this could provide areas for further study into the role and mechanism of RsmN within *P. aeruginosa*. By comparison with RsmA this will also provide an evaluation of the purification and attainment techniques used.

## 6 IDENTIFICATION OF RSMN AND RSMA RNA TARGETS

### 6.1 INTRODUCTION

The Rsm/Csr family of proteins specifically recognize and bind to a conserved GGA trinucleotide located in the 5' leader sequence of target mRNAs, preferably with the motif exposed in the loops of stem-loop structures (Lapouge et al., 2007, Dubey et al., 2005, Baker et al., 2007, Lapouge et al., 2008, Schubert et al., 2007). Multiple copies of the GGA motif may be found in the leader sequence of a target mRNA but one GGA element must overlap the ribosome binding site (RBS) sequence (Baker et al., 2007, Blumer et al., 1999, Baker et al., 2002).

The mechanisms of how the Rsm/Csr RNA-binding proteins control regulation of various phenotypes is largely unknown, with only a few direct targets identified, the rest probably indirectly affected by Rsm/Csr via Rsm/Csr influence on various regulatory systems.

CsrA in *E. coli* has been shown to directly bind and regulate translation of mRNAs encoding the RNA chaperone Hfq, enzymes involved in carbon starvation and glycogen synthesis, proteins responsible for the production of a biofilm polysaccharide (Baker et al., 2002, Baker et al., 2007, Dubey et al., 2003, Wang et al., 2005). Two proteins with GGDEF domains involved in the regulation of motility have also recently been identified (Jonas et al., 2008).

CsrA is also involved in the positive regulation of flagellar motility, where CsrA binds to the 5' region of the *flhDC* mRNA (Wei et al., 2001).

Direct regulation by RsmA in *P. aeruginosa* and *P. fluorescens* has been demonstrated for the hydrogen cyanide synthesis (*hcn*) transcript (Pessi and Haas, 2001, Lapouge et al., 2008).

Using co-purification of mRNAs with RsmA in *P. aeruginosa* PAK, genes identified to be directly upregulated by RsmA include those involved in hydrogen cyanide synthesis (*hcnABC* operon), a predicted Zn-dependent protease (PA0277), fatty acid and phospholipid metabolism (PA2541 operon), cell division and chromosome partitioning (PA3728 in the PA3732 operon), a hypothetical protein (PA4492) and T6S novel bacterial secretion system genes (PA0081/PA0082) (Brenic and Lory, 2009).

In order to identify targets for the novel RsmA-homologue RsmN, a variety of different approaches may be taken. The use of phenotypic assays as performed in Chapter 4, such as swarming, can give clear and unequivocal proof of regulatory control. Even working under the assumption that RsmN is an RsmA homologue, it might not have an effect, whether direct or indirect, on the same phenotypes. If a phenotype is not identified using RsmA targets there is a multitude of phenotypes that could be tested such as secondary metabolite and virulence factor production, motility and biofilm formation. When taken into consideration that RsmN might require an external factor to function, or be dependent on growth phase, a more global targeted approach is required.

Microarrays have been used in numerous experiments on differing scales and are often utilized to identify differentially expressed genes. Situations studied in *P. aeruginosa* have included transcriptome comparisons of *P. aeruginosa* strains grown under iron starvation conditions (Ochsner et al., 2002), *las/rhl*

regulatory mutants (Wagner et al., 2003, Hentzer et al., 2003, Schuster et al., 2003), cellular responses to hydrogen peroxide (Chang et al., 2005) and genes differentially expressed in mucoid strains (Firoved and Deretic, 2003). RNA profiling methods such as microarray analysis of transcriptomes have previously been non-strand specific and therefore unable to accurately identify antisense transcripts, determine the transcribed strand of non-coding RNAs or identify the boundaries of closely situated or overlapping genes.

RNA-seq uses novel high-throughput sequencing technologies to sequence cDNA produced from whole transcriptomes, but at a lower cost and running a greater number of samples than traditional sequencing, with an enhanced range of nucleotide sequence sizes. The technology used in this work was Next-Generation SOLiD sequencing (Applied Biosystems) as described in section 2.8.8.3, utilizing a novel barcoding approach. This allows cDNA from independent RNA samples to be pooled and sequenced. Data analyses can trace the sequence data back to a specific sample using its specific barcode (Section 2.8.8.3.1). System accuracy up to 99.99 % is achieved, based on sequencing control synthetic beads and reference-free data analysis.

During the past five years the use of these RNA-seq platforms has enabled the acquisition of large datasets in numerous models such as mouse embryonic stem cells (Cloonan et al., 2008), *Vibrio vulnificus* (Gulig et al., 2010), and mRNA sample isolated from *Bacillus anthracis*, applied using a mapping program for SOLiD platform data to a reference genome (Ondov et al., 2008), all using SOLiD platform and *Bacillus subtilis* (Irnov et al., 2010), *Helicobacter pylori* (Sharma et al., 2010) using the Roche FLX platform.

A recent transcriptome analysis based on Illumina sequencing confirmed that widespread antisense transcription also occurs in *E. coli* by identifying about 1,000 different asRNAs (Dornenburg et al., 2010).

During the writing of this thesis, Dötsch *et. al.*, published the first transcriptome study on *P. aeruginosa* that employs RNA sequencing technology and provides insights into the expression of small RNAs in *P. aeruginosa* biofilms using the Illumina platform (Dotsch et al., 2012). In this study qualitative analysis of the RNA-seq data revealed more than 3000 putative transcriptional start sites (TSS) and by the use of rapid amplification of cDNA ends (5'-RACE) they provided confirmation of the presence of three different TSS associated with the *pqsABCDE* operon, two in the promoter of *pqsA* and one upstream of the second gene, *pqsB*. These studies emphasise not only the power and versatility of the RNA-seq platforms, but the novelty of their use in providing qualitative and quantitative insights into bacterial transcriptomes.

The aims of this chapter are to identify RsmN targets with the use of Deep-Sequencing, together with an evaluative comparison of an RsmA dataset to provide context and assess stringency.

## 6.2 RESULTS AND DISCUSSION

### 6.2.1 Strains

The strains described in this chapter are all derived from the PAO1 Lausanne (Table 6.1).

**Table 6.1: *P. aeruginosa* strains for RNA-binding experiments.**

PA Number	Genotype/Characteristics
PAO1	PAO1 wild type Lausanne (L)
PALT63	PAO1 pRsmA (L)
PALT64	PAO1 pRsmN (L)

The plasmids used in the following experiments are described below (Table 6.2). Those for use in *P. aeruginosa* were based on pME6032, plasmids for use in *E. coli* were based on the pHT vector.

**Table 6.2: Plasmids for RNA-binding experiments.**

PA Number	Genotype/Characteristics
pLT3	pHT with <i>rsmA</i> , <i>Bam</i> HI and <i>Eco</i> RI (Ap <sup>R</sup> )
pLT4	pHT with <i>rsmN</i> , <i>Bam</i> HI/ <i>Bgl</i> III and <i>Eco</i> RI (Ap <sup>R</sup> )
pLT15	pHT with <i>rsmAR44A</i> arginine mutation, <i>Bam</i> HI and <i>Eco</i> RI (Ap <sup>R</sup> )
pLT16	pHT with <i>rsmNR62A</i> arginine mutation, <i>Bam</i> HI/ <i>Bgl</i> III and <i>Eco</i> RI (Ap <sup>R</sup> )
pRsmA	pME6032:: <i>rsmA</i> (Tet <sup>R</sup> ) C terminal hexahistidine tag
pRsmN	pME6032:: <i>rsmN</i> (Tet <sup>R</sup> ) N-terminal hexahistidine tag

#### 6.2.1.1 Construction of RsmA and RsmN arginine substitution mutants

Primers were designed to introduce an alanine mutation into the wild type *rsmN* and *rsmA* genes using the Stratagene Quick Change Site-Directed Mutagenesis kit<sup>®</sup> as described in section 2.4.1.1 and cloned into the pHT

vector. pHT has a modification of the expression vector pRSETA (Invitrogen) including a hexahistidine tag and a thrombin cleavage site (Ap<sup>R</sup>).

## **6.2.2 RNA binding experiments**

### 6.2.2.1 Protein-RNA binding using total RNA from *P. aeruginosa*

The first RNA binding experiments were attempted by the addition of pre-extracted RNA from PAO1-L to either RsmN (pLT4/pLT16) or RsmA (pLT3/pLT15) purified from *E. coli*. The RNA was extracted as described in section 2.8.7.3 and submitted to DNase digestion. The samples were cleaned using RNeasy MinElute cleanup kit.

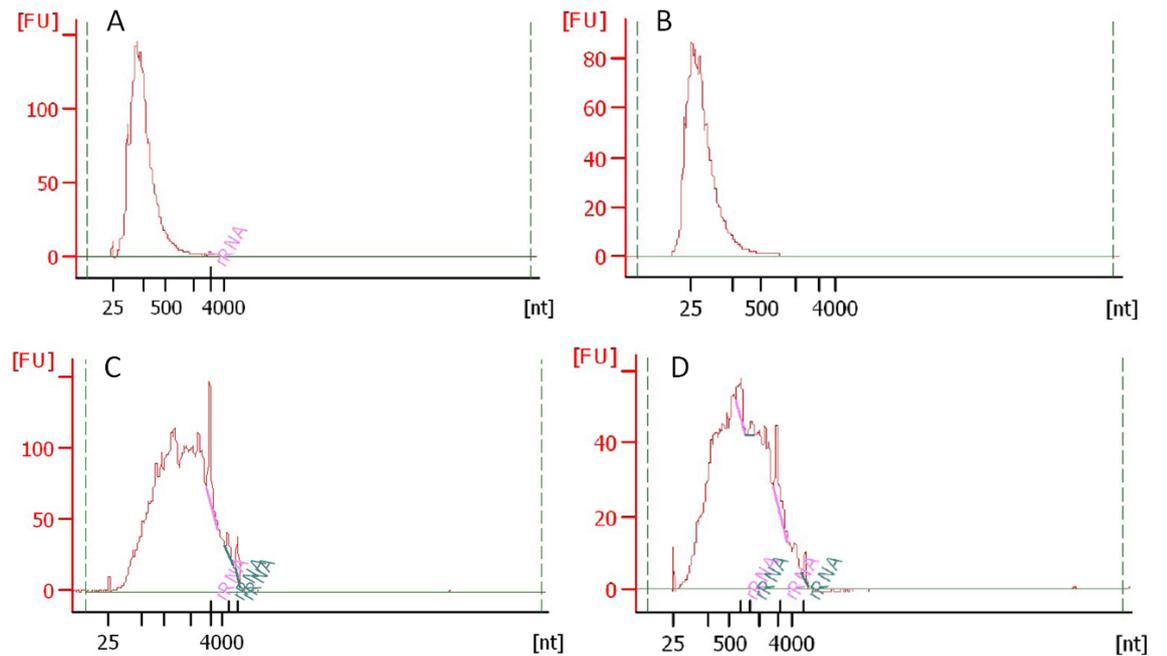
#### 6.2.2.1.1 Ni-NTA agarose Purifications

The RNA-protein binding experiments using a Ni-NTA agarose column is described in full in section 2.8.8.4.1. Both Ni-NTA and HisPur<sup>TM</sup> resin eluted highly pure protein when using the new wash stages as described in section 3.2.1.2. Ni-NTA was chosen as it would allow data comparison with Ni-NTA magnetic beads. For the magnetic bead experiments, the protein (RsmA or RsmA) was purified separately and bound to the beads by measuring 0.9 mg of protein and resuspended in 1.5 ml of 10 × Interaction buffer prior to binding with RNA as described previously (section 2.8.8.4.2). The protein elutions were pooled before RNeasy Midi preparation, followed by DNase treatment and cleaned using the RNeasy MinElute Cleanup kit. The concentration of RNA, including ribosomal RNA (rRNA), in each of the samples was estimated

using the Nanodrop spectrophotometer and a dilution was prepared (5 ng/μl) to assay on an Agilent Bioanalyzer.

The RNA profiles for the Ni-NTA column experiments shown in panel A and panel B (Fig. 5.1) have sharp, normal distributions, with a short range of RNA sizes in the samples. The actual concentrations were 6.04 ng/μl and 1.97 ng/μl for RNA that bound to RsmA (panel A) and RsmAR44A (panel B) respectively, with minimal rRNA contamination. The RNA extracted from the Ni-NTA magnetic beads purification has a wider range of nucleotide sizes than the RNA extracted from the Ni-NTA column preparation (Fig. 5.1 Panels C and D). The sample concentrations are higher at 10.51 ng/μl for RsmA (panel C) and 5.15 ng/μl for RsmAR44A (panel D). However, this also corresponds to an increase in rRNA contamination of 4.3 - 4.5 %. After the RNA was purified the final concentration was equivalent to the Ni-NTA agarose column experiments. Although data just for RsmA is shown here, the experiment was also performed with RsmN with the same result.

The final RNA quantities obtained after removal of any DNA present was very low, approximately 50 - 100 ng. With the absolute minimum concentration required for deep-sequencing being 500 ng, multiple scale-up experiments would be needed.



**Figure 6.1: Agilent bioanalyzer traces for RNA samples extracted from RsmA bound to a Ni-NTA column and magnetic beads.**

1  $\mu$ l of sample was used in each well of the Nano-RNA chip. Panel A) RsmA and panel B) RsmAR44A were purified from the Ni-NTA columns and panel C and D are of RsmA (C) and RsmAR44 (D) purified using Ni-NTA magnetic beads. Panels A and B both exhibit normal distributions with a narrow variance about the mean, indicating the RNA populations are of similar sizes. Panels C and D exhibit normal distributions with a wide variance about the mean, indicating the RNA populations are of a wider variety of sizes when purified using magnetic beads compared to the Ni-NTA column.

However, they would all be required to have RNA and protein taken from the same sample source, making this method impractical.

Scaling up the experiment would present difficulties in obtaining higher quantities of RNA. More RNA (150  $\mu$ g) was loaded onto both the column and the magnet beads without any further success. It would be expected that the limiting factor in this experiment is that target RNAs (for example, RsmZ and RsmY for RsmA) have a low abundance in the total RNAs extracted from PAO1. However, the low protein concentrations also contribute. A different approach had to be made to obtain a higher RNA concentration, but also an understanding of enrichment or depletion factors needed to be included.

Another consideration would be ensuring removal of the CAP protein prior to loading the column with the total RNA sample. As this co-purified with RsmA when grown in *E. coli* some RNAs in the bound eluent sample might be due to the interaction with CAP instead of RsmA. CAP was not seen on any purification gels of RsmN.

#### 6.2.2.2 RNA extraction from RsmA and RsmN overexpressed in PAO1

As the previous protein-RNA binding experiments were limited by the practical amount of final RNA that could be extracted after binding, a new method was designed to lower the number of experimental steps in order to minimise RNA loss. Therefore the plasmids pRsmA and pRsmN were separately transformed into electrocompetent PAO1 (Lausanne strain). These plasmids are pME6032-based, a *lacI<sup>Q</sup>-P<sub>tac</sub>*, pVS1-p15A shuttle expression vector (Tet<sup>R</sup>). Figure 2.5 in section 2.8.8.4.3 depicts the method utilised for the RNA extraction.

The proteins were purified from *P. aeruginosa* PAO1 using Ni-NTA agarose columns and the enriched RNAs in the subsequent elutions were obtained after phenol:chloroform extractions.

For all the protein-bound RNA and total RNA samples it was necessary to check that there was no DNA present. Therefore using the RNAs as templates and PAO1-L chromosomal DNA as a positive control, PCR reactions were performed using known primers (rsmA1 and rsmA2) and the resulting products examined by agarose gel electrophoresis. As it is important that no DNA is present, DNase digestions were repeated until no PCR products were obtained.

### 6.2.3 RNA Deep-sequencing results

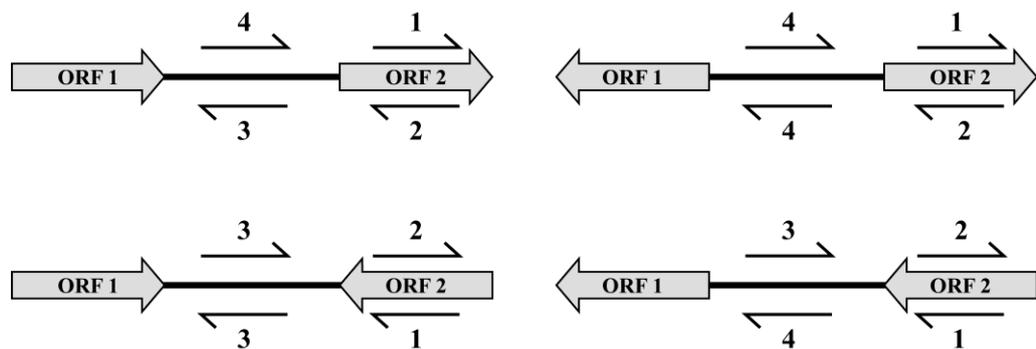
#### 6.2.3.1 RNA transcript identification

Total RNA extracted from the cells and RNA co-purifying with RsmA or RsmN were sequenced at the University of Nottingham Next Generation Sequencing Facility using the SOLiD sequencing system (See Appendix I for sequencing strategy flowchart).

The results were provided as Wiggle files (BioScope™ Software Users Guide), which correspond to tables of the nucleotides from a genomic reference sequence (GenBank accession No. NC\_002516) using a dedicated Perl script was used to identify the genomic context of each significant read sequence and in which every nucleotide has a value corresponding to the number of times that it has been mapped, this value being itself correlated with the abundance of the RNA from which the sequencing reads are derived.

As there isn't an internal standard that can be used to compare the total RNAs with the samples enriched in RNAs that bind RsmN or RsmA to determine their relative abundances, the data in the wiggle files must be normalised first to the average of their values. Then, enrichment factors between RNAs extracted with RsmN or RsmA versus the corresponding total RNAs can be calculated for each nucleotide in the genome. For practical purposes this factor is multiplied by 100, so that it will be greater than this number if there has been an enrichment of a particular nucleotide, or smaller if there has been depletion. To avoid division by zero errors, the arbitrary value of 9999 is used instead for undetermined enrichment factors (*i.e.*, every time that a nucleotide produced

reads in the enriched but not in the corresponding total RNA sample). The BioScope™ program (Applied Biosystems) also uses the genomic position of the nucleotide and the strand from which the reading originated to obtain additional information about the genomic context. Every nucleotide in a genome can be contextually positioned with respect to known upstream and downstream genes allowing the description of a topology for RNA reads spanning over intergenic regions, over known genes, or over a combination of these (Fig. 6.2). Genes identified by this analysis indicate possible targets of RsmN and RsmA.



**Figure 6.2: Interpretation of RNA genetic arrangements**

All the possible genetic arrangements of an RNA and its flanking genes can be classified into 4 groups where 1= RNA where ORF is the target, 2= antisense RNA is the target, 3= potential non-coding RNA and 4: ORF 5'UTR is the potential target. Continuous transcripts of contiguous combinations (e.g. 4-1, 3-2 or 1-3) are also possible, in which case the function overlapping the flanking gene is likely to prevail.

The transcripts were combined into two different types of data sets, semi-condensed and condensed. The semi-condensed data condenses the identified nucleotides into transcripts by recognising nucleotides that are next to each other. The condensed data combines the transcripts with others identified on the same strand within 200 nucleotides.

A minimum and arbitrary threshold enrichment factor of 200 was used to filter the data set, selecting only transcripts that had been enriched by at least two fold (proportionally 2 times more of a specific RNA in the protein-bound than in the total RNA samples). Where transcript reads covered more than one target gene, both PA numbers are indicated. Genes identified as significant and thus potential targets were annotated automatically. This entire analysis was done by computer following a dedicated algorithm (topologies explained in Appendix II), however, each significant result obtained by this computational method was subsequently validated separately manually by visual comparison against the Pseudomonas Genome Database ([www.pseudomonas.com](http://www.pseudomonas.com)) to ensure no errors were made and to extract any biologically relevant information.

### 6.2.3.2 RsmN transcript analysis

#### 6.2.3.2.1 RNAs enriched by binding to RsmN

The number of RsmN transcripts identified and those enriched are shown in Table 6.3. The number of individual transcripts identified for RsmN were 1,141 (data set 1) and 2,608 (data set 2). After the transcript data was condensed (neighbouring transcripts within 200 nt amalgamated), this was reduced to 930 (data set 1) and 458 (data set 2). Transcripts which had been enriched were selected with an average of  $\geq 200$ . Special care had to be taken with the condensed data by checking the transcript locations against the gene location as sometimes transcripts not of interest were included or the topology allocated was inaccurate.

**Table 6.3: Quantity of identified transcripts for RsmN.**

Data Set	Semi-condensed		Condensed	
	1	2	1	2
Total	2,608	1,141	930	458
Average $\geq$ 200	1,876	924	706	394
Average $\leq$ 50	64	49	19	20

Selected enriched transcripts from the RsmN experiment are shown in Table 6.4, with a more comprehensive list in Appendix III. Where the open reading frame (ORF) is the target (topology 1), the structural outer membrane protein encoding genes *popD*, *oprF*, *oprM*, *oprH*, *oprG*, *oprI* and *oprL* were identified, as well as transcriptional regulator genes such as *mvaT*, *vfr* and *pqsR*. Genes involved in secretion, twitching motility, flagellar structure and biofilms were also identified. Many of these have previously been identified as RsmA targets including genes required for pyocyanin, LasA and LecA production. The mRNA encoding RsmA appears to be a target of RsmN.

**Table 6.4: RsmN-enriched Target Transcripts.**

N: Negative, P: Positive strands, CDS: coding sequence. The average is the enrichment value multiplied by 100, only averages &gt;200 have been selected.

PA Number	Gene Name	Strand	Topology	Average	Comment
PA5128*	<i>secB</i>	N	1	9999.00	Secretion protein
PA2958.1	<i>rgsA</i>	P	1	5092.83	sRNA Gac-controlled indirectly
PA4726.11	<i>crcZ</i>	P	1	3091.22	Antagonistic RNA for catabolite repression control protein Crc
PA1871	<i>lasA</i>	P	1	2820.06	LasA protease precursor
PA0432	<i>sahH</i>	N	1	1908.25	S-adenosyl-L-homocysteine hydrolase
PA0524*	<i>norB</i>	P	1	1875.14	Nitric-oxide reductase subunit B
PA5040	<i>pilQ</i>	N	1	1340.63	Type 4 fimbrial biogenesis outer membrane protein PilQ precursor
PA1776/PA1777	<i>sigX/oprF</i>	P	1_4	1276.05	ECF sigma factor/Major porin and structural outer membrane porin OprF precursor
PA0766*	<i>mucD</i>	P	1	924.44	Serine protease MucD precursor
PA4428	<i>sspA</i>	N	1	843.76	Stringent starvation protein A
PA0962		N	1_4	803.09	Probable DNA-binding stress protein
PA2830*	<i>htpX</i>	P	1	481.73	Heat shock protein
PA1455*	<i>fliA</i>	P	1	477.21	Sigma factor
PA1098*	<i>fleS</i>	P	1	476.20	Two-component sensor
PA0427*	<i>oprM</i>	P	1	429.97	Major intrinsic multiple antibiotic resistance efflux outer membrane protein OprM precursor
PA4403*	<i>secA</i>	N	1	349.71	Secretion protein
PA1087*	<i>flgL</i>	P	1	341.04	Flagellar hook-associated protein type 3
PA0396*	<i>pilU</i>	P	1	217.35	Twitching motility protein
PA1001/PA1002*	<i>phnA/phnB</i>	P	1	208.60	Anthranilate synthase component I/ Anthranilate synthase component II
PA4315*	<i>mvaT</i>	P	1	206.30	Transcriptional regulator MvaT, P16 subunit
PA1432*	<i>lasI</i>	P	1	203.24	Autoinducer synthesis protein
PA5563	<i>Soj</i>	P	2	9999.00	Chromosome partitioning protein
PA5213*	<i>PI gcvP1</i>	P	2	9999.00	Glycine cleavage system protein
PA5446*	<i>wbpZ</i>	P	2	9999.00	Glycosyltransferase
PA1674*		P	2	5467.37	GTP cyclohydrolase I precursor

PA5474		N	2	2223.73	Probable metalloprotease
PA0654*	<i>sped</i>	N	2	1491.89	S-adenosylmethionine decarboxylase proenzyme
PA1546*	<i>hemN</i>	P	2	736.46	Oxygen-independent coproporphyrinogen III oxidase
PA1002	<i>phnB</i>	N	2	447.51	Anthranilate synthase component II
PA2423/PA2424*		P	3	369.73	Intergenic PA2423-PA2424
PA0652*	<i>Vfr</i>	N	4_1	9680.80	Transcriptional regulator
PA0519*	<i>nirS</i>	N	4_1	7748.71	Nitrite reductase precursor
PA5239*	<i>Rho</i>	N	4_1	4107.35	Transcription termination factor
PA3126	<i>ibpA</i>	N	4_1	4068.14	Heat-shock protein
PA3266*	<i>capB</i>	P	4_1	2982.23	Cold acclimation protein B
PA1178	<i>oprH</i>	P	4_1	2482.81	PhoP/Q and low Mg <sup>2+</sup> inducible outer membrane protein H1 precursor
PA4205	<i>mexG</i>	P	4_1	2465.75	Hypothetical protein
PA2570	<i>lecA</i>	N	4_1	2266.79	intergenic PA2570 - CDS PA2570
PA1544	<i>Anr</i>	N	4_1	2225.67	Transcriptional regulator
PA1003	<i>pqsR (mvrF)</i>	N	4_1	2131.77	Transcriptional regulator
PA1092	<i>fliC</i>	P	4_1	1902.52	Flagellin type B
PA3361	<i>lecB</i>	P	4_1	1879.25	Fucose-binding lectin PA-IIL
PA3351	<i>flgM</i>	P	1_4_1	1579.24	
PA3385	<i>amrZ</i>	P	4_1	1480.38	Alginate and motility regulator Z
PA0905	<i>rsmA</i>	P	4_1	1324.57	Regulator of secondary metabolites
PA4922	<i>Azu</i>	N	4_1	1308.93	Azurin precursor
PA5253	<i>algP</i>	N	4_1	1236.01	Alginate regulatory protein
PA4067	<i>oprG</i>	P	4_1	1174.02	Outer membrane protein OprG precursor
PA3724	<i>lasB</i>	N	4_1	1150.25	Elastase
PA2853	<i>oprI</i>	P	4_1	1002.37	Outer membrane lipoprotein OprI precursor
PA0762-PA0764	<i>algU/mucA/mucB</i>	P	4_1_4	903.70	Sigma factor/Anti-sigma factor / Negative regulator for alginate biosynthesis
PA4778*	<i>cueR</i>	P	4_1	832.40	

PA1770	<i>ppsA</i>	P	4_1	682.85	phosphoenolpyruvate synthase
PA3476-9	<i>rhlR/rhlAB</i>	N	4_1	674.47	Rhamnosyltransferase chain B
PA1454*	<i>fleN</i>	P	4_1	640.11	Flagellar synthesis regulator
PA1985	<i>pqqA</i>	P	4_1	565.96	Pyrroloquinoline quinone biosynthesis protein A
PA1094	<i>fliD</i>	P	4_1	545.02	Flagellar capping protein
PA1094	<i>fliD</i>	P	4_1	544.61	Flagellar capping protein
PA2231	<i>pslA</i>	P	4_1	532.70	
PA0576	<i>rpoD</i>	N	4_1	477.74	
PA5261/PA5262*	<i>algR/algZ</i>	N	1_4_1	383.28	Alginate biosynthesis protein
PA0973/PA0974	<i>oprL</i>	P	4_1	352.79	Peptidoglycan associated lipoprotein OprL precursor /conserved HP
PA0996-1000	<i>pqsABCDE</i>	P	4_1	337.36	
PA3476/PA3477	<i>rhlR/rhlI</i>	N	4_1	326.81	Transcriptional regulator / autoinducer synthesis protein
PA5261/PA5262	<i>algR/algZ</i>	N	4_1	320.02	Alginate biosynthesis regulatory protein
PA2622	<i>cspD</i>	P	4_1	260.18	Cold-shock protein
PA5183_PA5184	<i>rsmN</i>	N	4_1	250.95	RsmN
PA0408	<i>pilG</i>	P	4_1	213.90	Twitching motility protein

Other RNAs included RgsA, a sRNA which has been shown to be indirectly Gac-controlled (González et al., 2008) and CrcZ. The expression of this small RNA is driven by the CbrA/CbrB system in *P. aeruginosa* which is essential for maintenance of the carbon-nitrogen balance and for growth on energetically unfavourable carbon sources (Abdou et al., 2011). The sRNA, CrcZ antagonizes the repressing effects of the catabolite repression control protein Crc, an RNA-binding protein. Overexpression of *crcZ* relieves catabolite repression *in vivo*, whereas a *crcZ* mutation pleiotropically prevents the utilization of several carbon sources (Sonnleitner et al., 2009). The virulence factor regulator Vfr in *P. aeruginosa* is equivalent to CRP (cAMP receptor protein) in *E. coli*. Vfr can partially complement a *crp* mutation and therefore modulates catabolite repression as a receptor for cAMP binding (West et al., 1994). Soh *et al.*, presented evidence that Vfr binds *E. coli* lac promoter and that this binding requires cAMP (Suh et al., 2002). As catabolite repression control is not affected by *vfr* null mutants, Vfr is not required for catabolite repression. Marden *et al.*, (unpublished results) demonstrated that RsmA positively regulates acute virulence by controlling the cAMP-Vfr regulon by specific binding of the 5' untranslated region of the *vfr* transcript. Both *in vivo* and *in vitro* studies indicate a novel mechanism of positive posttranscriptional regulation, whereby RsmA binding promotes *vfr* translation directly, rather than through increased mRNA stability.

A potential non-coding RNA is located in the intergenic region between PA2423 and PA2424, corresponding to two Rho-independent transcription terminators TERM 1768 and 1769.

In topology 2, where an antisense RNA would be the target, an interesting transcript has been identified which is located on the opposite strand to *phnB*, an anthranilate synthase component. Anthranilate is a precursor of PQS (Essar et al., 1990, Gallagher et al., 2002). WbpZ (PA5446) is one of a cluster of genes that code for a glycosyltransferase which is required for O antigen assembly of A and B band lipopolysaccharides (Lam et al., 1999). PA1546 codes for *hemN*, an oxygen-independent coproporphyrinogen III oxidase involved in heme biosynthesis (Filiatrault et al., 2006). HemN is regulated by the dual action of the redox response regulators Dnr and Anr, the latter has also been identified as a target transcript (Rompf et al., 1998).

Of particular note are those transcripts with the topology 4-1, where the ORF is the target as well as the RBS, where RsmN might be acting as a post-transcriptional regulator affecting translation and/or stability of the mRNAs. The sequencing results identified a ncRNA (PA5183-PA5184) which corresponds to RsmN, which is as yet not annotated. Among the potential targets identified as such are the transcriptional regulators *vfr* (QS regulator), *anr* (anaerobic regulator) and *amzR* (alginate and motility regulator). Genes for the production of lectin, elastase and rhamnolipids were found as well as several required for motility, flagellar assembly, alginate biosynthesis and outer membrane proteins. Another potential target identified is *azu*, a precursor to the copper-binding redox protein azurin which has also been identified as being controlled by RsmA. The transcriptional regulator *rhIR* and autoinducer synthesis protein *rhII* and *lasI* genes were also distinguished as present in the enriched RsmN samples (326.81, 3-fold enrichment).

#### 6.2.3.2.2 Other potential RsmN targets

The impact of RsmN on some targets could not be confirmed (Table 6.5), where the RNA transcript was enriched in one sample, but depleted in the duplicate. Such potential targets included the regulatory RNAs, RsmZ and RsmY, which were both identified as potential targets with average enrichment factors of 330.13 and 198.75 respectively. This is because duplicate rather than triplicate samples were used for these experiments, due to cost and time constraints. The cultures were sampled at stationary phase, therefore in order to investigate RNA expression as a function of growth, samples would need to be taken in triplicate at a variety of time points for example, pre-exponential, exponential, late exponential and stationary.

For *phzB2* (PA1900) the enrichment varies between samples, where in one sample the depletion was 139.91 for a transcript length of 32 nt, and in the other an enrichment of 3666.84 for a transcript of 98 nt. Therefore this transcript is likely to be enriched when RsmN is overexpressed, taking into account the transcript length, but also the position of the transcripts in relation to the RNA of interest. This suggests phenazine production is increased when RsmN is overexpressed.

The small RNA PhrS stimulates synthesis of the *P. aeruginosa* alkylquinolone signal PqsR, a key quorum sensing regulator (Sonnleitner et al., 2011). The expression of *phrS* requires the oxygen-responsive regulator Anr, previously identified in these data sets. As PqsR was identified with an average enrichment factor of 2131.77, this would support the hypothesis that PhrS is

also enriched. The PhrS transcripts were of the same length (186 nt), either enriched two-fold (218.00) or depleted two-fold (46.16).

**Table 6.5: Undetermined RsmN targets.**

N: Negative, P: Positive strands. The average is the enrichment value multiplied by 100. Only transcripts which were enriched in one sample and depleted in the other were selected.

PA Number	Gene Name	Strand	Topology	Data Set 1-2		Data Set 3-4		Average
				Factor	Size (nt)	Factor	Size (nt)	
PA0872	<i>phhA</i>	N	4_1	796.35	79	74.46	115	368.43
PA1900	<i>phzB2</i>	P	1	139.91	32	3666.84	98	2798.67
PA3305.1	<i>PhrS</i>	N	1	218.00	186	46.16	186	132.08
PA3623/ PA3622		N	4_1	100.43	136	1509.39	251	1014.25
PA3621.1	<i>RsmZ</i>	N	4_1	118.03	101	530.33	107	330.13
PA0527.1	<i>RsmY</i>	P	4_1	40.61	96	352.10	99	198.75

#### 6.2.3.2.3 RNAs depleted when RsmN is overexpressed

The number of depleted transcripts identified from the semi-condensed data was 64 (data set 1) and 49 (data set 2). This was further reduced to 19 (data set 1) and 20 (data set 2) when the transcripts were condensed (Table 6.3). Table 6.6 contains the transcripts which were depleted with an enrichment factor average of less than 50 (specific RNA abundance decreased by two-fold or more in the RsmN-bound compared with the total RNA samples). The transcripts that met this criterion were checked in the whole data sets of the semi-condensed and condensed data against neighbouring transcripts coding for the same gene. The vast majority of these transcripts coded for a gene which was not depleted. Therefore only those transcripts which were depleted and only present in one data set (\*) are tabulated together with confirmed depletions from both data sets. Twelve depleted transcripts were identified, including a tRNAs, ribosomal proteins, MucP a metalloprotease involved in

alginate regulation (Damron and Yu) and BphO, a heme oxygenase (Wegele et al., 2004).

**Table 6.6: Depleted RsmN Transcripts.**

N: Negative, P: Positive strands. The average is the enrichment value multiplied by 100, only averages  $\leq 50$  have been selected.

Location	Gene	Strand	Topology	Average
PA4420/PA4421	Conserved hypothetical proteins	N	4_1	49.78
PA3743	tRNA (guanine-N1)-methyltransferase	N	1	45.74
PA3161*	Integration host factor beta subunit	N	1	45.20
PA4741*	30S ribosomal protein S15	N	1	44.31
PA4116*	Heme oxygenase, BphO	P	1_4	43.28
PA5285*	Hypothetical protein	N	1	37.82
PA3649*	MucP	N	1	36.77
PA3742/PA3742-PA3743*	50S ribosomal protein L19	N	4_1	36.74
PA0713/PA0713-PA0714*	Hypothetical protein	P	1_4	35.23
PA5285	Hypothetical protein	N	1	35.13
intergenic PA4581.1-PA4582*	tRNA-Arg/conserved hypothetical protein	P	4	32.21
PA0618*	Probable bacteriophage protein	P	1	12.81

#### 6.2.3.2.4 RNAs enriched by binding to RsmA

The number of enriched RNAs for RsmA was 6,775 (data set 1) and 11,078 (data set 2) from the semi-condensed data sets (Table 6.7). The number of transcripts identified for RsmN was much lower with 1,876 (data set 1) and 924 (data set 2) targets compared to those identified for RsmA.

**Table 6.7: Quantity of identified transcripts for RsmA.**

Data Set	Semi-condensed		Condensed	
	1	2	1	2
Total	10,110	13,022	2,853	3,509
Average $\geq 200$	6,775	11,078	2,061	3,129
Average $\leq 50$	1,934	1,441	451	284

Targets identified include the transcriptional regulators *hfq*, *vfr*, *pqsR*, *fleQ* and *anr* (Table 6.8). Appendix IV contains a more comprehensive table of selected enriched transcripts of interest.

The transcriptional regulator *rhlR* and autoinducer synthesis protein *rhlI* and *lasI* genes were identified together with anthranilate synthases (*trp/phn*), *qscR* (QS control repressor) and *pvdQ* (removal of acyl chains from pyoverdine).

Genes involved in secretion, twitching motility, flagellar structure and biofilms were detected, as well as targets for gene corresponding to production of pyocyanin, LasB, LecA and LecB (PA-IIL) and rhamnolipids. Topology 4\_1 targets RNAs include the small regulatory RNAs RsmY (6329.04) and RsmZ (3357.82). Target RNAs also identified in the RsmN data sets include RgsA, a sRNA is indirectly Gac-controlled (González et al., 2008) and *crcZ*, overexpression of which relieves catabolite repression (Sonnleitner et al., 2009, Abdou et al., 2011). Genes of the *mex* multidrug resistance operon *mexA/R* were identified as potential RsmA targets together with the gene coding for the sigma factor RpoD. The cyclic AMP (cAMP) phosphodiesterase gene *cpdA* is a target, the control of which by RsmA has been mentioned previously (Marden et al., unpublished results).

**Table 6.8: RsmA-enriched Target Transcripts.**

N: Negative, P: Positive strands. The average is the enrichment value multiplied by 100, only averages &gt;200 have been selected.

PA number	Gene	Strand	Topology	Average	Comment
PA1003	<i>mvfR (pqsR)</i>	N	1	9999.00	Transcriptional regulator MvfR (PqsR)
PA4969	<i>cpdA</i>	N	1	9999.00	Cyclic AMP (cAMP) Phosphodiesterase, CpdA
PA0928	<i>gacS</i>	P	1	9204.54	Sensor/response regulator hybrid <i>gacS</i>
PA0764	<i>mucB</i>	P	1	4859.75	Negative regulator for alginate biosynthesis MucB
PA2399	<i>pvdD</i>	N	1	4479.61	Pyoverdine synthetase D
PA1001/PA1002	<i>phnAB</i>	P	1	3986.75	Anthranilate synthase component I/II
PA3724	<i>lasB</i>	N	1	3724.90	Elastase LasB
PA2958.1	<i>rgsA</i>	P	1	1853.49	sRNA Gac-controlled indirectly
PA0609	<i>trpE</i>	N	2	9999.00	Anthranilate synthetase component I
PA1871	<i>lasA</i>	N	2	9999.00	LasA protease precursor
PA1003	<i>mvfR (pqsR)</i>	P	2	9441.53	Transcriptional regulator MvfR (PqsR)
PA0928	<i>gacS</i>	P	2	3351.49	Sensor/response regulator hybrid
PA1898	<i>qscR</i>	N	2	2140.92	Quorum-sensing control repressor
PA0291/PA0290	<i>oprE/HP</i>	N	3	8783.46	Intergenic Anaerobically-induced outer membrane porin OprE precursor/HP
PA2424/PA2425		P	3	2464.58	Intergenic PvdL/PvdG
PA2193	<i>hcnA</i>	P	4_1	9392.835	Hydrogen cyanide synthase
PA2385	<i>pvdQ</i>	N	4_1	9999.00	3-oxo-C12-homoserine lactone acylase PvdQ
PA1898	<i>qscR</i>	P	4_1	9046.99	Quorum-sensing control repressor
PA2570	<i>lecA</i>	N	4_1	8479.50	LecA

PA4704	<i>cbpA</i>	P	4_1	8424.55	cAMP-binding protein A
PA3974	<i>ladS</i>	N	4_1	7147.96	Lost Adherence Sensor, LadS
PA0527.1	<i>rsmY</i>	P	4_1_3_2	6329.04	Regulatory RNA RsmY
PA3361	<i>lecB</i>	P	4_1	6023.61	Fucose-binding lectin PA-IIL
PA0652	<i>vfr</i>	N	4_1	4532.75	Transcriptional regulator Vfr
PA0905	<i>rsmA</i>	P	4_1	3650.58	RsmA, regulator of secondary metabolites
PA3621.1	<i>rsmZ</i>	N	4_1	3357.82	Regulatory RNA RsmZ
PA1544	<i>anr</i>	N	4_1	3348.75	Transcriptional regulator Anr
PA4209	<i>phzM</i>	N	4_1	3288.93	Probable phenazine-specific methyltransferase
PA0996-PA1000	<i>pqsABCDE</i>	P	4_1	2918.45	<i>pqsABCDE</i>
PA1092	<i>fliC</i>	P	4_1	2712.15	Flagellin type B
PA4726.11	<i>crcZ</i>	P	4_1	2389.58	Antagonistic RNA for catabolite repression control protein Crc
PA5183/PA5184	<i>rsmN</i>	N	4_1	2136.77	
PA3476	<i>rhII</i>	N	4_1	2117.86	Autoinducer synthesis protein RhII
PA4315	<i>mvaT</i>	P	4_1	2041.82	Transcriptional regulator MvaT, P16 subunit
PA5239	<i>rho</i>	N	4	1921.89	Transcription termination factor Rho
PA1900	<i>phzB2</i>	P	4_1	1813.83	Probable phenazine biosynthesis protein
PA3385	<i>amrZ</i>	P	4_1	1393.14	Alginate and motility regulator Z
PA4526/PA4527	<i>pilB/pilC</i>	P	1_4_1	1324.14	Type 4 fimbrial biogenesis protein PilB/pilin biogenesis protein PilC
PA1430	<i>lasR</i>	P	4_1	1118.05	Transcriptional regulator LasR
PA3724	<i>lasB</i>	N	4	882.50	Elastase LasB
PA4922	<i>azu</i>	N	4_1	446.86	Azurin precursor
PA4944	<i>hfq</i>	N	4_1	471.69	Hfq
PA1432	<i>lasI</i>	P	4_1	316.04	Autoinducer synthesis protein LasI

There were numerous potential asRNA targets including two homoserine kinase genes *thrH* and *thrB* which are involved in threonine biosynthesis (Singh et al., 2004). The overexpression of *thrH* complements a *serB* mutation in *P. aeruginosa* and *E. coli*, the mutants of which are affected in a phosphoserine phosphatase involved in serine biosynthesis. Coding transcripts were found for *thrB*, however a transcript for *thrH* was also found but only in one data set. There are three other genes identified with both coding and antisense targets, *pqsR* (QS regulator), *gacS* (sensor/response regulator) and *qscR* (QS control repressor). The identification of possible asRNA control in the QS-network, especially on these three quorum-sensing regulators, provides a platform for further investigation into asRNA identification and function in *P. aeruginosa*.

Potential targets of transcriptional regulators include *argR* (controls expression of *argF*, ornithine carbamoyltransferase), *mucB* (alginate biosynthesis), *amrZ* (alginate and mobility regulator Z) and the transcription terminator factor Rho. Another notable target with the topology 4-1 is *ladS* (lost adherence sensor, 7147.96). If LadS and subsequently GacA are activated in the signal cascade, the latter increases *rsmZ* transcription, which leads to more RsmA being sequestered. This supports the sequencing data that RsmZ and RsmY transcripts are enriched. The well known RsmA target *hcnA* (hydrogen cyanide synthase) was also identified (Schubert et al., 2007).

#### 6.2.3.2.5 Depleted Transcripts of RsmA

These transcripts were depleted when RsmA was overexpressed, therefore the lower the average value, the greater the depletion. The number of transcripts in

the identified semi-condensed data was 1,934 (data set 1) and 1,441 (data set 2). Further condensing reduced these number to 451 (data set 1) and 284 (data set 2). The normal precautions were followed when interpreting the data.

Transcripts with topology 1, where the ORF is the target, identified genes involved in cell structure, maintenance and twitching with *pcdJ* (pyoverdine side chain peptide synthetase) and *rmlC*, *rmlD* and *rmlA* (biosynthesis of dTDP-L-rhamnose, a precursor of a key cell wall component), *mexA*, PA2018 and PA3676 (cell division efflux transporters) and *pilC* (fimbrial biosynthesis).

Antisense transcripts, topology 2, were identified for the transcriptional regulator LasR and *phnB* (anthranilate synthase component II). Other transcripts identified were *znuC* (Zinc transport protein) and *pilM* (fimbrial biosynthesis protein).

Transcripts identified with the topology 4-1, where the ORF is the target as well as the RBS, include *pmpR* (pqsR-mediated PQS regulator), *phrS* (PqsR synthesis), *crc* (catabolite repression control protein) and for the response regulator GacA. The sequencing data supports the literature that the sRNA CrcZ antagonizes the repressing effects of the catabolite repression control protein Crc, an RNA-binding protein in analogy to RsmA/RsmZ/RsmY (Sonnleitner et al., 2009).

**Table 6.9: Depleted RsmA target transcripts.**N: Negative, P: Positive strands. The average is the enrichment value multiplied by 100, only averages  $\leq 50$  have been selected.

PA Number	Gene	Topology	Strand	Average	Comment
PA2400	<i>pvdJ</i>	N	1	30.17	PvdJ
PA5164	<i>rmlC</i>	P	1	14.44	dTDP-4-dehydrorhamnose 3,5-epimerase rmlC
PA2018		N	1	14.20	Resistance-Nodulation-Cell Division (RND) multidrug efflux transporter
PA4527	<i>pilC</i>	P	1	13.24	Still frame shift type 4 fimbrial biogenesis protein PilC pilC
PA0426	<i>mexB</i>	P	1	10.75	Resistance-Nodulation-Cell Division (RND) multidrug efflux transporter MexB
PA5162/PA5163	<i>rmlD/rmlA</i>	P	1	9.39	dTDP-4-dehydrorhamnose reductase rmlD/glucose-1-phosphate thymidyltransferase rmlA
PA3676		N	1	9.00	Probable Resistance-Nodulation-Cell Division (RND) efflux transporter
PA5500	<i>znuC</i>	N	2	21.76	Zinc transport protein ZnuC
PA1002	<i>phnB</i>	N	2	9.31	Anthranilate synthase component II
PA5044	<i>pilM</i>	P	2	7.15	Type 4 fimbrial biogenesis protein PilM
PA1430	<i>lasR</i>	N	2	6.63	LasR transcriptional regulator
PA1776/PA1777	<i>sigX/oprR</i>	P	1_4_1	36.21	ECF sigma factor SigX/Major porin and structural outer membrane porin OprF precursor
PA3305.1	<i>phrS</i>	N	4_1	26.53	PhrS
PA0964	<i>pmpR</i>	P	4_1	25.56	PqsR-mediated PQS regulator, PmpR
PA3115	<i>fimV</i>	N	4_1	18.41	Motility protein FimV
PA0376	<i>rpoH</i>	P	4_1	17.88	Heat-shock sigma factor rpoH
PA5332	<i>crc</i>	P	4_1	9.77	Catabolite repression control protein
PA2586	<i>gacA</i>	N	4_1	8.90	Response regulator GacA

Enriched transcripts for *gacS* (Topology 1 (9204.54) and 2 (3351.49)) were identified together with a depleted transcript for *gacA* (Topology 4\_1, 8.90). This is an unexpected result as activation of the Gac pathway would increase transcription of RsmZ, reducing free RsmA. A depleted *gacA* transcript and an enriched asRNA for *gacS* were identified, leading to the possibility that asRNA control of GacS could independently control activation of the Gac pathway.

A comparison of selected transcripts from RsmA and RsmN are shown in Table 6.10, the complete table of transcripts of interest is in Appendix V. Complementary to both RsmA and RsmN data sets was the enrichment of transcripts corresponding to *pqsR*, *lasA*, *lasI*, *lecAB*, *vfr*, *lasB*, *rsmA*, *anr*, *phzB2* and *amrZ*. Depletion of the transcript corresponding to *crc*, the catabolite repression control protein, was consistent in both data sets.

Transcripts which were enriched but to a greater degree in the RsmA data set were *mvaT*, *rhlI*, *pqsABCDE*, *rsmY* and *rsmZ*. The only transcript to be enriched in RsmA (1118.05) and depleted in RsmN (38.03) is *lasR*. When RsmA was overproduced in a *lasI-lacZ* fusion, expression of *lasI* was delayed until the bacterial cells reached an OD<sub>600nm</sub> of around 1.0 (Pessi et al., 2001). Therefore repeating the sequencing with RNA samples taken from different time points and hence optical density, could help elucidate the role of RsmA and RsmN in time and density-dependant gene expression.

**Table 6.10: Comparison of selected RsmA and RsmN data.**

N: Negative, P: Positive strands. The average is the enrichment/depletion value multiplied by 100.

PA Number	Gene	Strand	RsmA		RsmN	
			Topology	Average	Topology	Average
PA1003	<i>pqsR</i>	N	1	9999.00	4_1	2131.77
PA1871	<i>lasA</i>	N	2	9999.00	1	2820.06
PA2570	<i>lecA</i>	N	4_1	8479.50	4_1	2277.80
PA0527.1	<i>rsmY</i>	P	4_1_3_2	6329.04	4_1	198.75
PA3361	<i>lecB</i>	P	4_1	6023.61	4_1	1879.25
PA0652	<i>vfr</i>	N	4_1	4532.75	4_1	9680.80
PA1001/PA1002	<i>phnAB</i>	P	1	3986.75	1	208.60
PA3724	<i>lasB</i>	N	1	3724.90	4_1	1150.25
PA0905	<i>rsmA</i>	P	4_1	3650.58	4_1	1324.57
PA3621.1	<i>rsmZ</i>	N	4_1	3357.82	4_1	330.13
PA1544	<i>anr</i>	N	4_1	3348.75	4_1	2225.67
PA0996-PA1000	<i>pqsABCDE</i>	P	4_1	2918.45	4_1	337.3557
PA4726.11	<i>crcZ</i>	P	4_1	2389.58	1	3091.22
PA3476	<i>rhII</i>	N	4_1	2117.86	4_1	326.81
PA4315	<i>mvaT</i>	P	4_1	2041.82	1	206.30
PA1900	<i>phzB2</i>	P	4_1	1813.83	1	2798.67
PA3385	<i>amrZ</i>	P	4_1	1393.14	4_1	1480.38
PA1430	<i>lasR</i>	P	4_1	1118.05	4_1	38.03
PA1432	<i>lasI</i>	P	4_1	316.04	1	203.24
PA5332*	<i>crc</i>	P	4_1	9.77	4_1	119.36

These results are complementary with previous microarray data performed on RsmA with the identification of many genes including those involved in secretion, structure, cell division and twitching (Burrowes et al., 2006, Brennic and Lory, 2009). Both Burrowes *et al.*, and Brennic and Lory performed transcriptional profiling in an *rsmA* mutant compared to the wild type in PAO1 and PAK identifying 506 and 529 genes respectively that displayed significantly altered transcript levels (greater than two fold). Out of 67 genes common to both, only 36 of these were affected by RsmA in the same direction. Discrepancies could be due to the difference in genomic backgrounds and/or the difference in growth stage sampling where Burrowes *et al.*, sampled in the exponential phase OD 0.8 and Brennic and Lory sampled in

the stationary phase OD 6.0. The study by Brencic and Lory also included an identification of 6 mRNAs that co-purified with RsmA using a histidine-tagged RsmA containing plasmid in the wild type and *rsmA* mutant strains, one of which was *hcnA*, hydrogen cyanide synthase. In this thesis, the *hcnA* gene was identified in the RsmA but not the RsmN sequencing data.

### 6.3 CONCLUSIONS

The use of RNA Deep-sequencing has facilitated analysis of targets of the novel RsmA orthologue, RsmN. The sequencing results produced large data sets for both RsmN and RsmA with many transcripts of interest. The number of semi-condensed enriched RsmN transcripts identified were 1,276 (data set 1) and 924 (data set 2) and the number of depleted transcripts was 64 (data set 1) and 49 (data set 2). In comparison there was a greater pool of RsmA transcripts with 6,775 (data set 1) and 11,078 (data set 2) enriched transcripts. The number of depleted transcripts was 1,934 (data set 1) and 1,441 (data set 2) for RsmA.

RsmN enriched transcripts identified numerous target genes including those required for structural outer membrane proteins, transcriptional regulators as well as genes involved in motility, secretion, flagellar structure and biofilms. RsmA, RsmZ and RsmY were all identified as targets together with the small RNAs RgsA (indirectly gac-controlled) and the antagonistic CrcZ. The virulence factor regulator Vfr in *P. aeruginosa* which is equivalent to CRP (cAMP receptor protein) in *E. coli* was also identified.

The identification of many genes involved in virulence factor regulation in the RsmA sequencing results supports the current literature. The comparison of selected transcripts revealed many genes of interest that were present in both the RsmA and RsmN sequencing results (Table 6.10). Enriched transcripts corresponding to *pqsR*, *lasA*, *lasI*, *lecAB*, *vfr*, *lasB*, *rsmA*, *anr*, *phzB2* and *amrZ*, as well as the targets for *mvaT*, *rhII*, *pqsABCDE*, *rsmY* and *rsmZ* with a lower correlation between relative abundances. Depletion of the transcript

corresponding to *crc*, the catabolite repression control protein, was consistent in both data sets.

By conducting the sequencing experiments with RsmN and RsmA in parallel, the reliability of the technique as well as that of the results was tested. A further improvement would be to perform the experiments in triplicate in order to be better able to discriminate any ambiguous results as shown by the number of undetermined RsmN transcripts (contradictory abundances between the duplicate data sets), together with sampling at different time points along the growth curve at different time points and hence optical densities, could help elucidate the role of RsmA and RsmN in time and density-dependant gene expression. Validation of these results would be required by the construction of new transcriptional and translational reporter fusions, and by conducting *in vitro* binding assays. Cloning of selected RNA targets would help identify those which are most abundant, thereby providing a more targeted approach for further study.

## 7 GENERAL CONCLUSIONS

The CsrA homologue RsmA is a small 6.9 kDa RNA-binding protein which acts as a global post-transcriptional regulator in *P. aeruginosa*. Biochemical and structural data indicates that CsrA/RsmA functions as a homodimer (Dubey et al., 2003) and it has been shown that certain residues are required for maintaining structure and functionality (Heeb et al., 2006). RsmA consists of two monomers, each built of five  $\beta$ -sheets followed by an  $\alpha$ -helix. The three central  $\beta$ -strands from each monomer form a hydrophobic core by hydrogen-bonding. The residue arginine 44 has been characterised and shown to be indispensable for RNA binding (Heeb et al., 2006). It has further been established that the first  $\beta$ -sheet of one monomer and the fifth of the other are vital for interaction with RNA (Mercante et al., 2006a).

Crystallographic structures have been elucidated using X-ray diffraction for RsmA from *P. aeruginosa* ((Rife et al., 2005)) and *Y. enterocolitica* 8081 (Heeb et al., 2006). The solution NMR structures have been solved for CsrA from *E. coli* (Gutiérrez et al., 2005)) CsrA from *B. subtilis* (Koharudin et al., Not published) and RsmE from *P. fluorescens* (Schubert et al., 2007).

RsmA acts as a global post transcriptional regulator by binding to target mRNAs, affecting their translation and/or their stability and mediating the resulting changes in gene expression. This function is modulated by small, untranslated RNAs that are able to titrate out the RNA binding proteins away from the target mRNAs, and via this mechanism control translation and mRNA stability. In *P. aeruginosa*, RsmA can act as both a positive and a negative regulator. RsmA negatively regulates the production of hydrogen cyanide,

pyocyanin, LecA (PA-IL) lectin and AHLs, whereas it positively regulates swarming motility, lipase and rhamnolipid production (Heurlier et al., 2004).

Overexpression and purification of RsmA in this study enabled biophysical techniques to be performed. A combination of CD temperature melts and NMR analysis confirmed RsmA has a high degree of stability. The analysis of protein unfolding as temperature increased using NMR was shown to be non-reversible, contradicting the results found using circular dichroism, which indicated the protein to be stable at 80 °C. RsmA is resistant to changes in pH (7.2 - 5.2) and can be denatured by the addition of a chemical denaturant (GdmCl). The existence of RsmA as both monomers and a dimer was confirmed by ESI-MS. The identification of new target residues for tryptophan mutation could enable analysis of the unfolding of the RsmA dimer.

RsmN is a 7.8 kDa protein which shares 34 % identity and 52 % similarity with the 6.9 kDa protein RsmA. RsmN was discovered from *in silico* analysis of an intergenic region common to 4 clones found using genomic bank screening (M. Messina, PhD thesis) where the clones were identified as capable of restoring the swarming-deficient phenotype of an *rsmA* mutant. A possible antisense gene termed *nmsR* was also discovered. Although swarming assays confirmed that RsmN did not complement the RsmA mutation, further exploration was made into RsmN due to its high similarity of sequence and structure to RsmA. Sequence comparison with RsmA revealed some conserved residues, Arg6, Ala54, Pro55 and Glu64, the corresponding residues of which in RsmA are important for maintenance of structure. The solvent-exposed residue Arg62 was also conserved, where previous study has shown the corresponding residue

in RsmA, R44, is required for retention of biological function (Heeb et al., 2006). The RsmN dimer forms a clam-like structure and CD scans confirmed that RsmN has greater alpha helical content and that RsmA has more unstructured polypeptide chain than RsmN.

The use of transcriptional reporter fusions demonstrated RsmN to have little to no regulatory effect on the expression of AHL synthases *lasI* and *rhII*. Mutations of the transcriptional regulators RhIR and LasR had no significant regulatory effect on the expression of the *rsmN* or *nmsR* promoters. The PqsA mutant strain resulted in an increase in *rsmN* expression, therefore *rsmN* is likely to be repressed by the action of the quinolones or the response regulator PqsE. The *pqsA* mutation had no effect on the *nmsR* promoter. Repeating the experiments using a wide range of IPTG concentrations from 0 to 1000  $\mu$ M could help elucidate the effect of RsmN and RsmA at a range of concentrations on the *lasI*, *rhII* and *pqsA* promoter fusions. Western blot analysis could be repeated with multiple RsmA and RsmN dependant strains using both anti-RsmA and anti-RsmN antibodies for RsmN identification and to determine cross reactivity.

No evidence could be found that RsmN acts as an RsmA homologue using the phenotype assays of swarming, glycogen accumulation, elastase, protease or pyocyanin production under the conditions they were performed. RsmN, unlike RsmA, does not have a control on the restriction modification system of *P. aeruginosa*. Study into the surrounding ORFs to *rsmN* which were common to the four identified swarming complementary clones, PA5182-PA5184

hypothetical proteins, could provide insights into this inability to restore the swarming phenotype.

The expression of *rsmN* and *nmsR* under RsmA control was inconclusive due to contradictory results from the conditional mutant strains. The conditional mutant could have a leaky expression of *rsmA* in the absence of IPTG or there is could be concentration-dependent effect of IPTG on both promoters. Further experiments could be performed using a range of concentrations of IPTG. To elucidate the effect that RsmN has on *rsmA* expression, more transcriptional fusions would need to be constructed in the wild type,  $\Delta rsmN$  mutant and conditional *rsmN* mutant strains with an *rsmA* promoter.

According to the experimental results performed under these particular conditions, RetS, LadS and GacA all appear to have a significant effect as activators on both the *rsmN* and *nmsR* promoters. If RsmN is acting as an RsmA homologue, these results would contradicts the results in published for RsmA in the literature (Ventre et al., 2006). Further elucidation could be obtained by the construction of additional transcriptional fusions, for example looking at the effect of RsmA and RsmN on *rsmZ* and *rsmY* expression. Testing the effect of *rsmA*, *rsmN* and *gacS* mutants, as well as double and triple mutants on target gene expression in a *gacS* mutant background, could provide insight to the effect of RsmN in concert with RsmA. Electrophoretic mobility shift assays using *rsmZ* and *rsmY* as a targets of RsmN could also indicate a possible role of RsmN in the Gac network. Obtaining an expression profile of RsmN would be of interest in comparison to RsmA. The knowledge of expression profiles can provide important information as shown by the observation that RsmE levels were highest at the end of growth in *P.*

*fluorescens* CHA0, suggesting that RsmE could play a role in the termination of GacA-controlled gene expression (Reimann et al., 2005).

The use of RNA Deep-sequencing has facilitated the identification of possible targets of RsmN, producing large data sets for both RsmN and RsmA with many transcripts of interest. RsmN enriched transcripts identified numerous target genes including those required for structural outer membrane proteins, transcriptional regulators as well as genes involved in motility, secretion, flagellar structure and biofilms. RsmA, RsmZ and RsmY were all identified as targets together with the small RNAs RgsA (indirectly gac-controlled) and the antagonistic RNA CrcZ (represses catabolite repression control protein Crc). The virulence factor regulator Vfr in *P. aeruginosa* which is equivalent to CRP (cAMP receptor protein) in *E. coli* was also identified. The identification of many genes involved in virulence factor regulation in the RsmA sequencing results supports the current literature.

A comparison of transcripts present in both the RsmA and RsmN sequencing results revealed a good correlation with genes involved in virulence factor regulation. Targets common to both RsmN and RsmA include the transcriptional regulators Vfr, PqsR, MvaT and Anr, regulatory RNAs RsmZ and RsmY together with transcripts corresponding to the *pqsABCDE* operon, LasB, LecA/B, RhII, LasR/I, Crc and CrcZ. asRNAs targets were identified for both RsmA and RsmN.

Improvements to the experiments would include more replicates to discriminate any ambiguous and sampling at different time points along the growth curve to potentially elucidate the role of RsmA and RsmN in time and

density-dependant gene expression. These results could be validated by the construction of new transcriptional and translational reporter fusions, and by conducting *in vitro* binding assays.

The targets found in these studies can be used for further RsmN phenotypes experiments. Binding studies using the sRNAs (or partial sequences) of CrcZ, PhrS and RgsA, could be conducted with RsmN and RsmA using NMR. Isothermal titration microcalorimetry (ITC) and EMSA experiments of RsmN with targets could be used to identify stoichiometry of binding. The effect of temperature, chemical denaturant and pH on RsmN can be elucidated using NMR, as well as folding studies using a CD temperature melt could be performed.

The identification of many gene targets in RsmN which are identical to targets of RsmA provides evidence that RsmN is involved in global-post-transcriptional regulation of gene expression along the sophisticated QS regulatory networks.

## 8 BIBLIOGRAPHY

- ABDOU, L., CHOU, H.-T., HAAS, D. & LU, C.-D. (2011) Promoter Recognition and Activation by the Global Response Regulator CbrB in *Pseudomonas aeruginosa*. *Journal of Bacteriology*, 193, 2784-2792.
- AENDEKERK, S., DIGGLE, S. P., SONG, Z., HØIBY, N., CORNELIS, P., WILLIAMS, P. & CÁMARA, M. (2005) The MexGHI-OpmD multidrug efflux pump controls growth, antibiotic susceptibility and virulence in *Pseudomonas aeruginosa* via 4-quinolone-dependent cell-to-cell communication. *Microbiology*, 151, 1113-1125.
- AHMER, B. M. M. (2004) Cell-to-cell signalling in *Escherichia coli* and *Salmonella enterica*. *Molecular Microbiology*, 52, 933-945.
- AITKEN, A. & LEARMONTH, M. (1996) *Protein determination by UV absorption*, Humana Press Inc.
- ALTIER, C., SUYEMOTO, M. & LAWHON, S. D. (2000) Regulation of *Salmonella enterica* serovar *Typhimurium* invasion genes by *csrA*. *Infection & Immunity*, 68, 6790-6797.
- ANG, S., HORNG, Y. T., SHU, J. C., SOO, P. C., LIU, J. H., YI, W. C., LAI, H. C., LUH, K. T., HO, S. W. & SWIFT, S. (2001) The role of RsmA in the regulation of swarming motility in *Serratia marcescens*. *Journal of Biomedical Science*, 8, 160-169.
- ANTUNES, L. C. M., FERREIRA, R. B. R., BUCKNER, M. M. C. & FINLAY, B. B. (2010) Quorum sensing in bacterial virulence. *Microbiology*, 156, 2271-2282.
- BABITZKE, P. & ROMEO, T. (2007) CsrB sRNA family: sequestration of RNA-binding regulatory proteins. *Current Opinion in Microbiology*, 10, 156-163.
- BAKER, C. S., EORY, L. A., YAKHNIN, H., MERCANTE, J., ROMEO, T. & BABITZKE, P. (2007) CsrA Inhibits Translation Initiation of *Escherichia coli* hfq by Binding to a Single Site Overlapping the Shine-Dalgarno Sequence. *Journal of Bacteriology*, 189, 5472-5481.
- BAKER, C. S., MOROZOV, I., SUZUKI, K., ROMEO, T. & BABITZKE, P. (2002) CsrA regulates glycogen biosynthesis by preventing translation of *glgC* in *Escherichia coli*. *Molecular Microbiology*, 44, 1599-1610.
- BARNARD, F. M., LOUGHLIN, M. F., FAINBERG, H. P., MESSENGER, M. P., USSERY, D. W., WILLIAMS, P. & JENKS, P. J. (2004) Global regulation of virulence and the stress response by CsrA in the highly adapted human gastric pathogen *Helicobacter pylori*. *Molecular Microbiology*, 51, 15-32.
- BASSLER, B. L. (2002) Small talk: Cell-to-cell communication in bacteria. *Cell*, 109, 421-424.
- BECHER, A. & SCHWEIZER, H. P. (2000) Integration-proficient *Pseudomonas aeruginosa* vectors for isolation of single-copy chromosomal *lacZ* and *lux* gene fusions. *Biotechniques*, 29, 948-954.
- BEISEL, C. L. & STORZ, G. Base pairing small RNAs and their roles in global regulatory networks. Blackwell Publishing Ltd.

- BEITER, T., REICH, E., WILLIAMS, R. & SIMON, P. (2009) Antisense transcription: A critical look in both directions. *Cellular and Molecular Life Sciences*, 66, 94-112.
- BEJERANO-SAGIE, M. & XAVIER, K. B. (2007) The role of small RNAs in quorum sensing. *Current Opinion in Microbiology*, 10, 189-198.
- BLUMER, C., HEEB, S., PESSI, G. & HAAS, D. (1999) Global GacA-steered control of cyanide and exoprotease production in *Pseudomonas fluorescens* involves specific ribosome binding sites. *Proceedings of the National Academy of Sciences of the United States of America*, 96, 14073-14078.
- BRADFORD, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, 72, 248-254.
- BRANTL, S. (2007) Regulatory mechanisms employed by cis-encoded antisense RNAs. *Current Opinion in Microbiology*, 10, 102-109.
- BRENCIC, A. & LORY, S. (2009) Determination of the regulon and identification of novel mRNA targets of *Pseudomonas aeruginosa* RsmA. *Molecular Microbiology*, 72, 612-632.
- BRUCE, A. G. & UHLENBECK, O. C. (1978) REACTIONS AT TERMINI OF TRANSFER-RNA WITH T4 RNA LIGASE. *Nucleic Acids Research*, 5, 3665-3677.
- BURROWES, E., BAYSSE, C., ADAMS, C. & O'GARA, F. (2006) Influence of the regulatory protein RsmA on cellular functions in *Pseudomonas aeruginosa* PAO1, as revealed by transcriptome analysis. *Microbiology*, 152, 405-418.
- CABALLERO, A. R., MOREAU, J. M., ENGEL, L. S., MARQUART, M. E., HILL, J. M. & O'CALLAGHAN, R. J. (2001) *Pseudomonas aeruginosa* protease IV enzyme assays and comparison to other pseudomonas proteases. *Analytical Biochemistry*, 290, 330-337.
- CALLEN, B. P., SHEARWIN, K. E. & EGAN, J. B. (2004) Transcriptional Interference between Convergent Promoters Caused by Elongation over the Promoter. *Molecular Cell*, 14, 647-656.
- CAMARA, M., WILLIAMS, P. & HARDMAN, A. (2002) Controlling infection by tuning in and turning down the volume of bacterial small-talk. *Lancet Infectious Diseases*, 2, 667-676.
- CASE, D. A., CHEATHAM, T. E., DARDEN, T., GOHLKE, H., LUO, R., MERZ, K. M., ONUFRIEV, A., SIMMERLING, C., WANG, B. & WOODS, R. J. (2005) The Amber biomolecular simulation programs. Wiley Subscription Services, Inc., A Wiley Company.
- CHANCEY, S. T., WOOD, D. W. & PIERSON, L. S., 3RD (1999) Two-component transcriptional regulation of *N*-acyl-homoserine lactone production in *Pseudomonas aureofaciens*. *Applied & Environmental Microbiology*, 65, 2294-2299.
- CHANG, W., SMALL, D., TOGHROL, F. & BENTLEY, W. (2005) Microarray analysis of *Pseudomonas aeruginosa* reveals induction of pyocin genes in response to hydrogen peroxide. *BMC Genomics*, 6, 115.
- CHATTERJEE, A., CUI, Y., CHAKRABARTY, P. & CHATTERJEE, A. K. (2010) Regulation of Motility in *Erwinia carotovora* subsp. *carotovora*: Quorum-Sensing Signal Controls FlhDC, the Global Regulator of

- Flagellar and Exoprotein Genes, by Modulating the Production of RsmA, an RNA-Binding Protein. *Molecular Plant-Microbe Interactions*, 23, 1316-1323.
- CHATTERJEE, A., CUI, Y., LIU, Y., DUMENYO, C. K. & CHATTERJEE, A. K. (1995) Inactivation of *rsmA* leads to overproduction of extracellular pectinases, cellulases, and proteases in *Erwinia carotovora* subsp. *carotovora* in the absence of the starvation/cell density-sensing signal, *N*-(3-oxohexanoyl)-L-homoserine lactone. *Applied & Environmental Microbiology*, 61, 1959-1967.
- CHEN, Q. & CROSA, J. H. (1996) Antisense RNA, Fur, Iron, and the Regulation of Iron Transport Genes in *Vibrio anguillarum*. *Journal of Biological Chemistry*, 271, 18885-18891.
- CHHABRA, S. R., HARTY, C., HOOL, D. S., DAYKIN, M., WILLIAMS, P., TELFORD, G., PRITCHARD, D. I. & BYCROFT, B. W. (2003) Synthetic analogues of the bacterial signal (quorum sensing) molecule *N*-(3-oxododecanoyl)-L-homoserine lactone as immune modulators. *Journal of Medicinal Chemistry*, 46, 97-104.
- CHUGANI, S. A., WHITELEY, M., LEE, K. M., D'ARGENIO, D., MANOIL, C. & GREENBERG, E. P. (2001) QscR, a modulator of quorum-sensing signal synthesis and virulence in *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 2752-2757.
- CLOONAN, N., FORREST, A. R. R., KOLLE, G., GARDINER, B. B. A., FAULKNER, G. J., BROWN, M. K., TAYLOR, D. F., STEPTOE, A. L., WANI, S., BETHEL, G., ROBERTSON, A. J., PERKINS, A. C., BRUCE, S. J., LEE, C. C., RANADE, S. S., PECKHAM, H. E., MANNING, J. M., MCKERNAN, K. J. & GRIMMOND, S. M. (2008) Stem cell transcriptome profiling via massive-scale mRNA sequencing. *Nature Methods*, 5, 613-619.
- COLEPICOLO, P., CHO, K. W., POINAR, G. O. & HASTINGS, J. W. (1989) Growth and luminescence of the bacterium *Xenorhabdus luminescens* from a human wound. *Applied & Environmental Microbiology*, 55, 2601-2606.
- CRAMPTON, N., BONASS, W. A., KIRKHAM, J., RIVETTI, C. & THOMSON, N. H. (2006) Collision events between RNA polymerases in convergent transcription studied by atomic force microscopy. *Nucleic Acids Research*, 34, 5416-5425.
- CUI, Y., CHATTERJEE, A., LIU, Y., DUMENYO, C. K. & CHATTERJEE, A. K. (1995) Identification of a global repressor gene, *rsmA*, of *Erwinia carotovora* subsp. *carotovora* that controls extracellular enzymes, *N*-(3-oxohexanoyl)-L-homoserine lactone, and pathogenicity in soft-rotting *Erwinia* spp. *Journal of Bacteriology*, 177, 5108-5115.
- DAMRON, F. H. & YU, H. D. (2011) *Pseudomonas aeruginosa* MucD regulates alginate pathway through activation of MucA degradation via MucP proteolytic activity. *Journal of Bacteriology*, JB.01132-10.
- DARFEUILLE, F., UNOSON, C., VOGEL, J. & WAGNER, E. G. H. (2007) An Antisense RNA Inhibits Translation by Competing with Standby Ribosomes. *Molecular Cell*, 26, 381-392.
- DAVID, L., HUBER, W., GRANOVSKAIA, M., TOEDLING, J., PALM, C. J., BOFKIN, L., JONES, T., DAVIS, R. W. & STEINMETZ, L. M.

- (2006) A high-resolution map of transcription in the yeast genome. *Proceedings of the National Academy of Sciences*, 103, 5320-5325.
- DEZIEL, E., GOPALAN, S., TAMPAKAKI, A. P., LEPINE, F., PADFIELD, K. E., SAUCIER, M., XIAO, G. P. & RAHME, L. G. (2005) The contribution of MvfR to *Pseudomonas aeruginosa* pathogenesis and quorum sensing circuitry regulation: multiple quorum sensing-regulated genes are modulated without affecting lasRI, rhlRI or the production of N-acyl-L-homoserine lactones. *Molecular Microbiology*, 55, 998-1014.
- DÉZIEL, E., LÉPINE, F., MILOT, S. & VILLEMUR, R. (2003) *rhlA* is required for the production of a novel biosurfactant promoting swarming motility in *Pseudomonas aeruginosa*: 3-(3-hydroxyalkanoyloxy)alkanoic acids (HAAs), the precursors of rhamnolipids. *Microbiology*, 149, 2005-2013.
- DIGGLE, S. P., LUMJIAKTASE, P., DIPILATO, F., WINZER, K., KUNAKORN, M., BARRETT, D. A., CHHABRA, S. R., CÁMARA, M. & WILLIAMS, P. (2006) Functional genetic analysis reveals a 2-alkyl-4-quinolone signaling system in the human pathogen *Burkholderia pseudomallei* and related bacteria. *Chemistry & Biology*, 13, 701-710.
- DIGGLE, S. P., MATTHIJS, S., WRIGHT, V. J., FLETCHER, M. P., CHHABRA, S. R., LAMONT, I. L., KONG, X., HIDER, R. C., CORNELIS, P., CÁMARA, M. & WILLIAMS, P. (2007) The *Pseudomonas aeruginosa* 4-quinolone signal molecules HHQ and PQS play multifunctional roles in quorum sensing and iron entrapment. *Chemistry & Biology*, 14, 87-96.
- DIGGLE, S. P., WINZER, K., CHHABRA, S. R., WORRALL, K. E., CÁMARA, M. & WILLIAMS, P. (2003) The *Pseudomonas aeruginosa* quinolone signal molecule overcomes the cell density-dependency of the quorum sensing hierarchy, regulates *rhl*-dependent genes at the onset of stationary phase and can be produced in the absence of LasR. *Molecular Microbiology*, 50, 29-43.
- DORNENBURG, J. E., DEVITA, A. M., PALUMBO, M. J. & WADE, J. T. (2010) Widespread antisense transcription in *Escherichia coli*. *mBio*, 1.
- DOTSCH, A., ECKWEILER, D., SCHNIEDERJANS, M., ZIMMERMANN, A., JENSEN, V., SCHARFE, M., GEFFERS, R. & HAUSSLER, S. (2012) The *Pseudomonas aeruginosa* transcriptome in planktonic cultures and static biofilms using RNA sequencing. *Plos One*, 7, e31092.
- DUBEY, A. K., BAKER, C. S., ROMEO, T. & BABITZKE, P. (2005) RNA sequence and secondary structure participate in high-affinity CsrA-RNA interaction. *RNA*, 11, 1579-1587.
- DUBEY, A. K., BAKER, C. S., SUZUKI, K., JONES, A. D., PANDIT, P., ROMEO, T. & BABITZKE, P. (2003) CsrA regulates translation of the *Escherichia coli* carbon starvation gene, *cstA*, by blocking ribosome access to the *cstA* transcript. *Journal of Bacteriology*, 185, 4450-4460.
- DÜHRING, U., AXMANN, I. M., HESS, W. R. & WILDE, A. (2006) An internal antisense RNA regulates expression of the photosynthesis gene *isiA*. *Proceedings of the National Academy of Sciences*, 103, 7054-7058.

- EBERHARD, A. (1972) Inhibition and activation of bacterial luciferase synthesis. *Journal of Bacteriology*, 109, 1101-&.
- ESSAR, D. W., EBERLY, L., HADERO, A. & CRAWFORD, I. P. (1990) Identification and characterization of genes for a second anthranilate synthase in *Pseudomonas aeruginosa*: interchangeability of the two anthranilate synthases and evolutionary implications. *Journal of Bacteriology*, 172, 884-900.
- FAGERLIND, M. G., NILSSON, P., HARLÉN, M., KARLSSON, S., RICE, S. A. & KJELLEBERG, S. (2005) Modeling the effect of acylated homoserine lactone antagonists in *Pseudomonas aeruginosa*. *Biosystems*, 80, 201-213.
- FENN, J. B., MANN, M., MENG, C. K., WONG, S. F. & WHITEHOUSE, C. M. (1990) Electrospray ionization - principles and practice. *Mass Spectrometry Reviews*, 9, 37-70.
- FETTES, P. S., FORSBACH-BIRK, V., LYNCH, D. & MARRE, R. (2001) Overexpression of a *Legionella pneumophila* homologue of the *E. coli* regulator *csrA* affects cell size, flagellation, and pigmentation. *International Journal of Medical Microbiology*, 291, 353-360.
- FILIATRAULT, M. J., PICARDO, K. F., NGAI, H., PASSADOR, L. & IGLEWSKI, B. H. (2006) Identification of *Pseudomonas aeruginosa* Genes Involved in Virulence and Anaerobic Growth. *Infection and Immunity*, 74, 4237-4245.
- FILIATRAULT, M. J., STODGHILL, P. V., BRONSTEIN, P. A., MOLL, S., LINDEBERG, M., GRILLS, G., SCHWEITZER, P., WANG, W., SCHROTH, G. P., LUO, S., KHREBTUKOVA, I., YANG, Y., THANNHAUSER, T., BUTCHER, B. G., CARTINHO, S. & SCHNEIDER, D. J. (2010) Transcriptome Analysis of *Pseudomonas syringae* Identifies New Genes, Noncoding RNAs, and Antisense Activity. *Journal of Bacteriology*, 192, 2359-2372.
- FINCH, R. G., PRITCHARD, D. I., BYCROFT, B. W., WILLIAMS, P. & STEWART, G. S. A. B. (1998) Quorum sensing: A novel target for anti-infective therapy. *Journal of Antimicrobial Chemotherapy*, 42, 569-571.
- FIROVED, A. M. & DERETIC, V. (2003) Microarray Analysis of Global Gene Expression in Muroid *Pseudomonas aeruginosa*. *Journal of Bacteriology*, 185, 1071-1081.
- FLETCHER, M. P., DIGGLE, S. P., CRUSZ, S. A., CHHABRA, S. R., CÁMARA, M. & WILLIAMS, P. (2007) A dual biosensor for 2-alkyl-4-quinolone quorum-sensing signal molecules. *Environmental Microbiology*, 9, 2683-2693.
- FRANK, L. H. & DEMOSS, R. D. (1959) On the biosynthesis of pyocyanine. *Journal of Bacteriology*, 77, 776-782.
- FRASER, G. M. & HUGHES, C. (1999) Swarming motility. *Current Opinion in Microbiology*, 2, 630-635.
- FRÖHLICH, K. S. & VOGEL, J. (2009) Activation of gene expression by small RNA. *Current Opinion in Microbiology*, 12, 674-682.
- FUQUA, C. (2006) The QscR quorum-sensing regulon of *Pseudomonas aeruginosa*: an orphan claims its identity. *Journal of Bacteriology*, 188, 3169-3171.

- GALLAGHER, L. A., MCKNIGHT, S. L., KUZNETSOVA, M. S., PESCI, E. C. & MANOIL, C. (2002) Functions required for extracellular quinolone signaling by *Pseudomonas aeruginosa*. *Journal of Bacteriology*, 184, 6472-6480.
- GAMBELLO, M. J. & IGLEWSKI, B. H. (1991) Cloning and characterization of the *Pseudomonas aeruginosa lasR* gene, a transcriptional activator of elastase expression. *Journal of bacteriology*, 173, 3000-3009.
- GAMBELLO, M. J., KAYE, S. & IGLEWSKI, B. H. (1993) LasR of *Pseudomonas aeruginosa* is a transcriptional activator of the alkaline protease gene (*apr*) and an enhancer of exotoxin-A expression. *Infection and Immunity*, 61, 1180-1184.
- GAMPER, M., GANTER, B., POLITO, M. R. & HAAS, D. (1992) RNA processing modulates the expression of the *arcDABC* operon in *Pseudomonas aeruginosa*. *Journal of Molecular Biology*, 226, 943-957.
- GARCÍA-MAYORAL, M. F., HOLLINGWORTH, D., MASINO, L., DÍAZ-MORENO, I., KELLY, G., GHERZI, R., CHOU, C.-F., CHEN, C.-Y. & RAMOS, A. (2007) The Structure of the C-Terminal KH Domains of KSRP Reveals a Noncanonical Motif Important for mRNA Degradation. *Structure (London, England : 1993)*, 15, 485-498.
- GEORG, J. & HESS, W. R. (2011) cis-Antisense RNA, Another Level of Gene Regulation in Bacteria. *Microbiology and Molecular Biology Reviews*, 75, 286-300.
- GONZÁLEZ, N., HEEB, S., VALVERDE, C., KAY, E., REIMMANN, C., JUNIER, T. & HAAS, D. (2008) Genome-wide search reveals a novel GacA-regulated small RNA in *Pseudomonas* species *BMC Genomics*, 9, 167.
- GOODMAN, A. L., KULASEKARA, B., RIETSCH, A., BOYD, D., SMITH, R. S. & LORY, S. (2004) A signaling network reciprocally regulates genes associated with acute infection and chronic persistence in *Pseudomonas aeruginosa*. *Developmental cell*, 7, 745-754.
- GOODMAN, A. L., MERIGHI, M., HYODO, M., VENTRE, I., FILLOUX, A. & LORY, S. (2009) Direct interaction between sensor kinase proteins mediates acute and chronic disease phenotypes in a bacterial pathogen. *Genes & Development*, 23, 249-259.
- GORE, M. G. (2000) *Spectrophotometry and spectrofluorimetry*, Oxford University Press.
- GRANT, S. G., JESSEE, J., BLOOM, F. R. & HANAHAN, D. (1990) Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proceedings of the National Academy of Sciences of the United States of America*, 87, 4645-4649.
- GREENSTEIN, J. P. (1938) Sulfhydryl groups in proteins I. Egg albumin in solutions of urea, guanidine, and their derivatives. *Journal of Biological Chemistry*, 125, 501-513.
- GREENSTEIN, J. P. (1939) Sulfhydryl groups in proteins - II. Edestin, excelsin, and globin in solutions of guanidine hydrochloride, urea, and their derivatives. *Journal of Biological Chemistry*, 128, 233-240.
- GROUP, R. G. E. R., GROUP, G. S., CONSORTIUM, T. F., KATAYAMA, S., TOMARU, Y., KASUKAWA, T., WAKI, K., NAKANISHI, M., NAKAMURA, M., NISHIDA, H., YAP, C. C., SUZUKI, M., KAWAI,

- J., SUZUKI, H., CARNINCI, P., HAYASHIZAKI, Y., WELLS, C., FRITH, M., RAVASI, T., PANG, K. C., HALLINAN, J., MATTICK, J., HUME, D. A., LIPOVICH, L., BATALOV, S., ENGSTRÅM, P. G., MIZUNO, Y., FAGHIHI, M. A., SANDELIN, A., CHALK, A. M., MOTTAGUI-TABAR, S., LIANG, Z., LENHARD, B. & WAHLESTEDT, C. (2005) Antisense Transcription in the Mammalian Transcriptome. *Science*, 309, 1564-1566.
- GULIG, P. A., DE CRECY-LAGARD, V., WRIGHT, A. C., WALTERS, B., TELONIS-SCOTT, M. & MCINTYRE, L. M. (2010) SOLiD sequencing of four *Vibrio vulnificus* genomes enables comparative genomic analysis and identification of candidate clade-specific virulence genes. *BMC Genomics*, 11.
- GUTIÉRREZ, P., LI, Y., OSBORNE, M. J., POMERANTSEVA, E., LIU, Q. & GEHRING, K. (2005) Solution structure of the carbon storage regulator protein CsrA from *Escherichia coli*. *Journal of Bacteriology*, 187, 3496-3501.
- HANKINS, J. S., DENROCHE, H. & MACKIE, G. A. (2010) Interactions of the RNA-Binding Protein Hfq with *cspA* mRNA, Encoding the Major Cold Shock Protein. *Journal of Bacteriology*, 192, 2482-2490.
- HART, S. R., WATERFIELD, M. D., BURLINGAME, A. L. & CRAMER, R. (2002) Factors governing the solubilization of phosphopeptides retained on ferric NTA IMAC beads and their analysis by MALDI TOFMS. *Journal of the American Society for Mass Spectrometry*, 13, 1042-1051.
- HEEB, S., BLUMER, C. & HAAS, D. (2002) Regulatory RNA as mediator in GacA/RsmA-dependent global control of exoproduct formation in *Pseudomonas fluorescens* CHA0. *Journal of Bacteriology*, 184, 1046-1056.
- HEEB, S., FLETCHER, M. P., CHHABRA, S. R., DIGGLE, S. P., WILLIAMS, P. & CAMARA, M. (2011) Quinolones: from antibiotics to autoinducers. *FEMS Microbiology Reviews*, 35, 247-274.
- HEEB, S., HEURLIER, K., VALVERDE, C., CÁMARA, M., HAAS, D. & WILLIAMS, P. (2004) Post-transcriptional regulation in *Pseudomonas* spp. via the Gac/Rsm regulatory network. IN RAMOS, J.-L. (Ed.) *Pseudomonas*. New York, Kluwer Academic Publishers.
- HEEB, S., KUEHNE, S. A., BYCROFT, M., CRIVII, S., ALLEN, M. D., HAAS, D., CÁMARA, M. & WILLIAMS, P. (2006) Functional analysis of the post-transcriptional regulator RsmA reveals a novel RNA-binding site. *Journal of Molecular Biology*, 355, 1026-1036.
- HENTZER, M., WU, H., ANDERSEN, J. B., RIEDEL, K., RASMUSSEN, T. B., BAGGE, N., KUMAR, N., SCHEMBRI, M. A., SONG, Z., KRISTOFFERSEN, P., MANEFIELD, M., COSTERTON, J. W., MOLIN, S., EBERL, L., STEINBERG, P., KJELLEBERG, S., HOIBY, N. & GIVSKOV, M. (2003) Attenuation of *Pseudomonas aeruginosa* virulence by quorum sensing inhibitors. *EMBO Journal*, 22, 3803-3815.
- HEURLIER, K., WILLIAMS, F., HEEB, S., DORMOND, C., PESSI, G., SINGER, D., CÁMARA, M., WILLIAMS, P. & HAAS, D. (2004) Positive control of swarming and lipase production by the post-transcriptional RsmA/RsmZ system in *Pseudomonas aeruginosa* PAO1. *Journal of Bacteriology*, 186, 2936-2945.

- HIRANO, S. S., OSTERTAG, E. M., SAVAGE, S. A., BAKER, L. S., WILLIS, D. K. & UPPER, C. D. (1997) Contribution of the regulatory gene *lemA* to field fitness of *Pseudomonas syringae* pv. *syringae*. *Applied & Environmental Microbiology*, 63, 4304-4312.
- HRABAK, E. M. & WILLIS, D. K. (1992) The *lemA* gene required for pathogenicity of *Pseudomonas syringae* pv. *syringae* on bean is a member of a family of two-component regulators. *Journal of Bacteriology*, 174, 3011-3020.
- HUANG, H.-H., LIAO, H.-K., CHEN, Y.-J., HWANG, T.-S., LIN, Y.-H. & LIN, C.-H. (2005) Structural characterization of sialic acid synthase by electrospray mass spectrometry--a tetrameric enzyme composed of dimeric dimers. *Journal of the American Society for Mass Spectrometry*, 16, 324-332.
- HUMAIR, B., GONZALEZ, N., MOSSIALOS, D., REIMMANN, C. & HAAS, D. (2009) Temperature-responsive sensing regulates biocontrol factor expression in *Pseudomonas fluorescens* CHA0. *Multidisciplinary Journal of Microbial Ecology*, 3, 955-965.
- HUTTENHOFER, A. & VOGEL, J. (2006) Experimental approaches to identify non-coding RNAs. *Nucleic Acids Research*, 34, 635-646.
- IRNOV, I., SHARMA, C. M., VOGEL, J. & WINKLER, W. C. (2010) Identification of regulatory RNAs in *Bacillus subtilis*. *Nucleic Acids Research*, 38, 6637-6651.
- ITOH, T. & TOMIZAWA, J. (1980) Formation of an RNA primer for initiation of replication of ColE1 DNA by ribonuclease H. *Proceedings of the National Academy of Sciences*, 77, 2450-2454.
- IVERSEN, S. L. & JØRGENSEN, M. H. (1995) Azocasein assay for alkaline protease in complex fermentation broth. *Biotechnology Techniques*, 9, 573-576.
- JONAS, K., EDWARDS, A. N., SIMM, R., ROMEO, T., RÖMLING, U. & MELEFORS, Ö. (2008) The RNA binding protein CsrA controls cyclic di-GMP metabolism by directly regulating the expression of GGDEF proteins. *Molecular Microbiology*, 70, 236-257.
- JUHAS, M., EBERL, L. & TÜMMLER, B. (2005) Quorum sensing: the power of cooperation in the world of *Pseudomonas*. *Environmental Microbiology*, 7, 459-471.
- KAWANO, M., REYNOLDS, A. A., MIRANDA-RIOS, J. & STORZ, G. (2005) Detection of 5'- and 3'-UTR-derived small RNAs and cis-encoded antisense RNAs in *Escherichia coli*. *Nucleic Acids Research*, 33, 1040-1050.
- KAY, E., DUBUIS, C. & HAAS, D. (2005) Three small RNAs jointly ensure secondary metabolism and biocontrol in *Pseudomonas fluorescens* CHA0. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 17136-17141.
- KAY, E., HUMAIR, B., DENERVAUD, V., RIEDEL, K., SPAHR, S., EBERL, L., VALVERDE, C. & HAAS, D. (2006) Two GacA-dependent small RNAs modulate the quorum-sensing response in *Pseudomonas aeruginosa*. *Journal of Bacteriology*, 188, 6026-6033.
- KIM, W., KILLAM, T., SOOD, V. & SURETTE, M. G. (2003) Swarm-cell differentiation in *Salmonella enterica* serovar typhimurium results in

- elevated resistance to multiple antibiotics. *Journal of Bacteriology*, 185, 3111-3117.
- KING, E. O., WARD, M. K. & RANEY, D. E. (1954) Two simple media for the demonstration of pyocyanin and fluorescin. *Journal of Laboratory and Clinical Medicine*, 44, 301-307.
- KLEEREBEZEM, M., QUADRI, L. E. N., KUIPERS, O. P. & DEVOS, W. M. (1997) Quorum sensing by peptide pheromones and two-component signal- transduction systems in Gram-positive bacteria. *Molecular Microbiology*, 24, 895-904.
- KLINGER, K. W. (1983) Micromethods for the spectrophotometric determination of bacterial protease activities. *Journal of Microbiological Methods*, 1, 329-337.
- KÖHLER, T., CURTY, L. K., BARJA, F., VAN DELDEN, C. & PECHÈRE, J.-C. (2000) Swarming of *Pseudomonas aeruginosa* is dependent on cell-to-cell signaling and requires flagella and pili. *Journal of Bacteriology*, 182, 5990-5996.
- KORADI, R., BILLETER, M. & WUTHRICH, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures. *Journal of Molecular Graphics*, 14, 51-&.
- KOZAK, M. (2005) Regulation of translation via mRNA structure in prokaryotes and eukaryotes. *Gene*, 361, 13-37.
- KRICKA, L. J. (2003) Clinical applications of chemiluminescence. *Analytica Chimica Acta*, 500, 279-286.
- LACATENA, R. M. & CESARENI, G. (1981) Base pairing of RNA I with its complementary sequence in the primer precursor inhibits ColE1 replication. *Nature*, 294, 623-626.
- LAM, J. S., ROCCHETTA, H. L. & BURROWS, L. L. (1999) Glycosyltransferases of *Pseudomonas aeruginosa* that assemble the O antigens of A band and B band lipopolysaccharide. *Journal of Endotoxin Research*, 5, 96-101.
- LAPOUGE, K., SCHUBERT, M., ALLAIN, F. H. & HAAS, D. (2008) Gac/Rsm signal transduction pathway of gamma-proteobacteria: from RNA recognition to regulation of social behaviour. *Molecular Microbiology*, 67, 241-253.
- LAPOUGE, K., SINEVA, E., LINDELL, M., STARKE, K., BAKER, C. S., BABITZKE, P. & HAAS, D. (2007) Mechanism of *hcnA* mRNA recognition in the Gac/Rsm signal transduction pathway of *Pseudomonas fluorescens*. *Molecular Microbiology*, 66, 341-356.
- LATIFI, A., WINSON, M. K., FOGLINO, M., BYCROFT, B. W., STEWART, G. S., LAZDUNSKI, A. & WILLIAMS, P. (1995) Multiple homologues of LuxR and LuxI control expression of virulence determinants and secondary metabolites through quorum sensing in *Pseudomonas aeruginosa* PAO1. *Molecular Microbiology*, 17, 333-343.
- LAVILLE, J., VOISARD, C., KEEL, C., MAURHOFER, M., DÉFAGO, G. & HAAS, D. (1992) Global control in *Pseudomonas fluorescens* mediating antibiotic synthesis and suppression of black root rot of tobacco. *Proceedings of the National Academy of Sciences of the United States of America*, 89, 1562-1566.

- LÉPINE, F., DÉZIEL, E., MILOT, S. & RAHME, L. G. (2003) A stable isotope dilution assay for the quantification of the *Pseudomonas* quinolone signal in *Pseudomonas aeruginosa* cultures. *Biochimica et Biophysica Acta*, 1622, 36-41.
- LEWIS, H. A., CHEN, H., EDO, C., BUCKANOVICH, R. J., YANG, Y. Y., MUSUNURU, K., ZHONG, R., DARNELL, R. B. & BURLEY, S. K. (1999) Crystal structures of Nova-1 and Nova-2 K-homology RNA-binding domains. *Structure with Folding & Design*, 7, 191-203.
- LI, L.-L., MALONE, J. E. & IGLEWSKI, B. H. (2007) Regulation of the *Pseudomonas aeruginosa* Quorum-Sensing Regulator VqsR. *Journal of Bacteriology*, 189, 4367-4374.
- LIAW, S. J., LAI, H. C., HO, S. W., LUH, K. T. & WANG, W. B. (2003) Role of RsmA in the regulation of swarming motility and virulence factor expression in *Proteus mirabilis*. *Journal of Medical Microbiology*, 52, 19-28.
- LIU, M. Y., GUI, G., WEI, B., PRESTON, J. F., OAKFORD, L., YUKSEL, U., GIEDROC, D. P. & ROMEO, T. (1997) The RNA molecule CsrB binds to the global regulatory protein CsrA and antagonizes its activity in *Escherichia coli*. *Journal of Biological Chemistry*, 272, 17502-17510.
- LIU, M. Y. & ROMEO, T. (1997) The global regulator CsrA of *Escherichia coli* is a specific mRNA-binding protein. *Journal of Bacteriology*, 179, 4639-4642.
- LOO, J. A. (1997) Studying noncovalent protein complexes by electrospray ionization mass spectrometry. Wiley Subscription Services, Inc., A Wiley Company.
- LU, C., MEYERS, B. C. & GREEN, P. J. (2007) Construction of small RNA cDNA libraries for deep sequencing. *Methods*, 43, 110-117.
- LYCZAK, J. B., CANNON, C. L. & PIER, G. B. (2000) Establishment of *Pseudomonas aeruginosa* infection: lessons from a versatile opportunist. *Microbes and Infection*, 2, 1051-1060.
- MAJDALANI, N., CUNNING, C., SLEDJESKI, D., ELLIOTT, T. & GOTTESMAN, S. (1998) DsrA RNA regulates translation of *rpoS* message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. *Proceedings of the National Academy of Sciences of the United States of America*, 95, 12462-12467.
- MASSÉ, E., ESCORCIA, F. E. & GOTTESMAN, S. (2003) Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*. *Genes & Development*, 17, 2374-2383.
- MERCANTE, J., EDWARDS, A. N., DUBEY, A. K., BABITZKE, P. & ROMEO, T. (2009) Molecular geometry of CsrA (RsmA) binding to RNA and its implications for regulated expression. *Journal of Molecular Biology*, 392, 511-528.
- MERCANTE, J., SUZUKI, K., CHENG, X., BABITZKE, P. & ROMEO, T. (2006a) Comprehensive alanine-scanning mutagenesis of *Escherichia coli* CsrA defines two subdomains of critical functional importance. *J Biol Chem*, 281, 31832-42.
- MERCANTE, J., SUZUKI, K., CHENG, X., BABITZKE, P. & ROMEO, T. (2006b) Comprehensive alanine-scanning mutagenesis of *Escherichia*

- coli* CsrA defines two subdomains of critical functional importance. *Journal of Biological Chemistry*, 281, 31832-31842.
- MILTON, D. L., O'TOOLE, R., HORSTEDT, P. & WOLF-WATZ, H. (1996) Flagellin A is essential for the virulence of *Vibrio anguillarum*. *Journal of Bacteriology*, 178, 1310-1319.
- MIROUX, B. & WALKER, J. E. (1996) Over-production of proteins in *Escherichia coli*: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *Journal of Molecular Biology*, 260, 289-298.
- MOHANTY, B. K., MAPLES, V. F. & KUSHNER, S. R. (2004) The Sm-like protein Hfq regulates polyadenylation dependent mRNA decay in *Escherichia coli*. *Molecular Microbiology*, 54, 905-920.
- MORITA, T., MAKI, K. & AIBA, H. (2005) RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes & Development*, 19, 2176-2186.
- MORTAZAVI, A., WILLIAMS, B. A., MCCUE, K., SCHAEFFER, L. & WOLD, B. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods*, 5, 621-628.
- MULCAHY, H., O'CALLAGHAN, J., O'GRADY, E. P., ADAMS, C. & O'GARA, F. (2006) The posttranscriptional regulator RsmA plays a role in the interaction between *Pseudomonas aeruginosa* and human airway epithelial cells by positively regulating the type III secretion system. *Infection and Immunity*, 74, 3012-3015.
- NAGALAKSHMI, U., WANG, Z., WAERN, K., SHOU, C., RAHA, D., GERSTEIN, M. & SNYDER, M. (2008) The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science*, 320, 1344-1349.
- NEALSON, K. H., PLATT, T. & HASTINGS, J. W. (1970) Cellular control of synthesis and activity of bacterial luminescent system. *Journal of Bacteriology*, 104, 313-&.
- NIEMANN, H. H., SCHUBERT, W.-D. & HEINZ, D. W. (2004) Adhesins and invasins of pathogenic bacteria: a structural view. *Microbes and Infection*, 6, 101-112.
- O'TOOLE, G. A. & KOLTER, R. (1998) Flagellar and twitching motility are necessary for *Pseudomonas aeruginosa* biofilm development. *Molecular Microbiology*, 30, 295-304.
- OCHSNER, U. A., FIECHTER, A. & REISER, J. (1994) Isolation, characterization and expression in *Echerichia coli* of the *Pseudomonas aeruginosa* *rhlAB* genes encoding a rhamnopyranosyltransferase involved in rhamnolipid biosurfactant synthesis. *Journal of Biological Chemistry*, 269, 19787-19795.
- OCHSNER, U. A., WILDERMAN, P. J., VASIL, A. I. & VASIL, M. L. (2002) GeneChip expression analysis of the iron starvation response in *Pseudomonas aeruginosa*: identification of novel pyoverdine biosynthesis genes. *Molecular Microbiology*, 45, 1277-1287.
- OHMAN, D. E., CRYZ, S. J. & IGLEWSKI, B. H. (1980) Isolation and characterization of a *Pseudomonas aeruginosa* PAO mutant that produces altered elastase. *Journal of Bacteriology*, 142, 836-842.

- OKADA, M., SATO, I., CHO, S. J., IWATA, H., NISHIO, T., DUBNAU, D. & SAKAGAMI, Y. (2005) Structure of the *Bacillus subtilis* quorum-sensing peptide pheromone ComX. *Nature Chemical Biology*, 1, 23-24.
- ONDOV, B. D., VARADARAJAN, A., PASSALACQUA, K. D. & BERGMAN, N. H. (2008) Efficient mapping of Applied Biosystems SOLiD sequence data to a reference genome for functional genomic applications. *Bioinformatics*, 24, 2776-2777.
- OPDYKE, J. A., KANG, J.-G. & STORZ, G. (2004) GadY, a Small-RNA Regulator of Acid Response Genes in *Escherichia coli*. *Journal of Bacteriology*, 186, 6698-6705.
- OVERHAGE, J., LEWENZA, S., MARR, A. K. & HANCOCK, R. E. (2007) Identification of genes involved in swarming motility using a *Pseudomonas aeruginosa* PAO1 mini-Tn5-*lux* mutant library. *Journal of Bacteriology*, 189, 2164-2169.
- PALMER, A. C., AHLGREN-BERG, A., EGAN, J. B., DODD, I. B. & SHEARWIN, K. E. (2009) Potent Transcriptional Interference by Pausing of RNA Polymerases over a Downstream Promoter. *Molecular Cell*, 34, 545-555.
- PAPENFORT, K. & VOGEL, J. (2009) Multiple target regulation by small noncoding RNAs rewires gene expression at the post-transcriptional level. *Research in Microbiology*, 160, 278-287.
- PARAMESWARAN, P., JALILI, R., TAO, L., SHOKRALLA, S., GHARIZADEH, B., RONAGHI, M. & FIRE, A. Z. (2007) A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing. *Nucleic Acids Research*, 35, e130.
- PASSADOR, L., COOK, J. M., GAMBELLO, M. J., RUST, L. & IGLEWSKI, B. H. (1993) Expression of *Pseudomonas aeruginosa* virulence genes requires cell-to-cell communication. *Science*, 260, 1127-1130.
- PASSNER, J. M., SCHULTZ, S. C. & STEITZ, T. A. (2000) Modeling the cAMP-induced Allosteric Transition Using the Crystal Structure of CAP-cAMP at 2.1 Å Resolution. *Journal of Molecular Biology*, 304, 847-859.
- PEARSON, J. P., GRAY, K. M., PASSADOR, L., TUCKER, K. D., EBERHARD, A., IGLEWSKI, B. H. & GREENBERG, E. P. (1994) Structure of the autoinducer required for expression of *Pseudomonas aeruginosa* virulence genes. *Proceedings of the National Academy of Sciences of the United States of America*, 91, 197-201.
- PESCI, E. C., MILBANK, J. B. J., PEARSON, J. P., MCKNIGHT, S., KENDE, A. S., GREENBERG, E. P. & IGLEWSKI, B. H. (1999) Quinolone signaling in the cell-to-cell communication system of *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences of the United States of America*, 96, 11229-11234.
- PESCI, E. C., PEARSON, J. P., SEED, P. C. & IGLEWSKI, B. H. (1997) Regulation of *las* and *rhl* quorum sensing in *Pseudomonas aeruginosa*. *Journal of Bacteriology*, 179, 3127-3132.
- PESSI, G. & HAAS, D. (2000) Transcriptional control of the hydrogen cyanide biosynthetic genes *hcnABC* by the anaerobic regulator ANR and the quorum-sensing regulators LasR and RhIR in *Pseudomonas aeruginosa*. *Journal of Bacteriology*, 182, 6940-6949.

- PESSI, G. & HAAS, D. (2001) Dual control of hydrogen cyanide biosynthesis by the global activator GacA in *Pseudomonas aeruginosa* PAO1. *FEMS Microbiology Letters*, 200, 73-78.
- PESSI, G., WILLIAMS, F., HINDLE, Z., HEURLIER, K., HOLDEN, M. T. G., CÁMARA, M., HAAS, D. & WILLIAMS, P. (2001) The global posttranscriptional regulator RsmA modulates production of virulence determinants and *N*-acylhomoserine lactones in *Pseudomonas aeruginosa*. *Journal of Bacteriology*, 183, 6676-6683.
- PFEIFFER, V., PAPENFORT, K., LUCCHINI, S., HINTON, J. C. D. & VOGEL, J. (2009) Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nature Structural & Molecular Biology*, 16, 840-846.
- POSTIS, V. L. G., DEACON, S. E., ROACH, P. C. J., WRIGHT, G. S. A., XIA, X., INGRAM, J. C., HADDEN, J. M., HENDERSON, P. J. F., PHILLIPS, S. E. V., MCPHERSON, M. J. & BALDWIN, S. A. (2008) A high-throughput assay of membrane protein stability. *Molecular Membrane Biology*, 25, 617-624.
- PRENTKI, P. & KRISCH, H. M. (1984) *In vitro* insertional mutagenesis with a selectable DNA fragment. *Gene*, 29, 303-313.
- RAJAGOPAL, L., VO, A., SILVESTRONI, A. & RUBENS, C. E. (2006) Regulation of cytotoxin expression by converging eukaryotic-type and two-component signalling mechanisms in *Streptococcus agalactiae*. *Molecular Microbiology*, 62, 941-957.
- RAMPIONI, G., PUSTELNY, C., FLETCHER, M. P., WRIGHT, V. J., BRUCE, M., RUMBAUGH, K. P., HEEB, S., CÁMARA, M. & WILLIAMS, P. (2010) Transcriptomic analysis reveals a global alkyl-quinolone-independent regulatory role for PqsE in facilitating the environmental adaptation of *Pseudomonas aeruginosa* to plant and animal hosts. *Environmental Microbiology*, In press.
- RASHID, M. H. & KORNBERG, A. (2000) Inorganic polyphosphate is needed for swimming, swarming, and twitching motilities of *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences of the United States of America*, 97, 4885-4890.
- RATHER, P. N. (2005) Swarmer cell differentiation in *Proteus mirabilis*. *Environmental Microbiology*, 7, 1065-1073.
- RECORDS, A. R. & GROSS, D. C. (2010) Sensor Kinases RetS and LadS Regulate *Pseudomonas syringae* Type VI Secretion and Virulence Factors. *Journal of Bacteriology*, 192, 3584-3596.
- REIMMANN, C., BEYELER, M., LATIFI, A., WINTELER, H., FOGLINO, M., LAZDUNSKI, A. & HAAS, D. (1997) The global activator GacA of *Pseudomonas aeruginosa* PAO positively controls the production of the autoinducer *N*-butyryl-homoserine lactone and the formation of the virulence factors pyocyanin, cyanide, and lipase. *Molecular Microbiology*, 24, 309-319.
- REIMMANN, C., VALVERDE, C., KAY, E. & HAAS, D. (2005) Posttranscriptional repression of GacS/GacA-controlled genes by the RNA-binding protein RsmE acting together with RsmA in the biocontrol strain *Pseudomonas fluorescens* CHA0. *Journal of Bacteriology*, 187, 276-285.

- RIFE, C., SCHWARZENBACHER, R., MCMULLAN, D., ABDUBEK, P., AMBING, E., AXELROD, H., BIORAC, T., CANAVES, J. M., CHIU, H.-J., DEACON, A. M., DIDONATO, M., ELSLIGER, M.-A., GODZIK, A., GRITTINI, C. A., GRZECHNIK, S. K., HALE, J., HAMPTON, E., HAN, G. W., HAUGEN, J., HORNSBY, M., JAROSZEWSKI, L., KLOCK, H. E., KOESEMA, E., KREUSCH, A., KUHN, P., LESLEY, S. A., MILLER, M. D., MOY, K., NIGOGHOSSIAN, E., PAULSEN, J., QUIJANO, K., REYES, R., SIMS, E., SPRAGGON, G., STEVENS, R. C., VAN DEN BEDEM, H., VELASQUEZ, J., VINCENT, J., WHITE, A., WOLF, G., XU, Q., HODGSON, K. O., WOOLEY, J. & WILSON, I. A. (2005) Crystal structure of the global regulatory protein CsrA from *Pseudomonas putida* at 2.05 Å resolution reveals a new fold. *Proteins: Structure, Function, and Bioinformatics*, [Epub ahead of print] <http://dx.doi.org/10.1002/prot.20502>.
- RODRIGUE, A., QUENTIN, Y., LAZDUNSKI, A., MEJEAN, V. & FOGLINO, M. (2000) Two-component systems in *Pseudomonas aeruginosa*: why so many? *Trends in Microbiology*, 8, 498-504.
- ROMEO, T. (1998) Global regulation by the small RNA-binding protein CsrA and the non-coding RNA molecule CsrB. *Molecular Microbiology*, 29, 1321-1330.
- ROMEO, T., GONG, M., LIU, M. Y. & BRUN-ZINKERNAGEL, A. M. (1993) Identification and molecular characterization of *csrA*, a pleiotropic gene from *Escherichia coli* that affects glycogen biosynthesis, gluconeogenesis, cell size, and surface properties. *Journal of Bacteriology*, 175, 4744-4755.
- ROMPF, A., HUNGERER, C., HOFFMANN, T., LINDENMEYER, M., RÖMLING, U., GROß, U., DOSS, M. O., ARAI, H., IGARASHI, Y. & JAHN, D. (1998) Regulation of *Pseudomonas aeruginosa* hemF and hemN by the dual action of the redox response regulators Anr and Dnr. *Molecular Microbiology*, 29, 985-997.
- SAIKI, R. K., SCHARF, S., FALOONA, F., MULLIS, K. B., HORN, G. T., ERLICH, H. A. & ARNHEIM, N. (1985) Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle-cell anemia. *Science*, 230, 1350-1354.
- SALINAS, P. C., WALDBESER, L. S. & CROSA, J. H. (1993) Regulation of the expression of bacterial iron transport genes: possible role of an antisense RNA as a repressor. *Gene*, 123, 33-38.
- SAMBROOK, J., FRITSCH, E. F. & MANIATIS, T. (1989) *Molecular cloning: a laboratory manual, 2nd ed.*, Cold Spring Harbor, N.Y.
- SAMBROOK, J. & RUSSELL, D. W. (2001) *Molecular cloning: a laboratory manual, 3rd ed.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- SCHAGGER, H. (2006) Tricine-SDS-PAGE. *Nature Protocols*, 1, 16-22.
- SCHUBERT, M., LAPOUGE, K., DUSS, O., OBERSTRASS, F. C., JELESAROV, I., HAAS, D. & ALLAIN, F. H. (2007) Molecular basis of messenger RNA recognition by the specific bacterial repressing clamp RsmA/CsrA. *Nature structural & molecular biology*, 14, 807-813.

- SCHUSTER, M., LOSTROH, C. P., OGI, T. & GREENBERG, E. P. (2003) Identification, timing, and signal specificity of *Pseudomonas aeruginosa* quorum-controlled genes: a transcriptome analysis. *Journal of Bacteriology*, 185, 2066-2079.
- SEED, P. C., PASSADOR, L. & IGLEWSKI, B. H. (1995) Activation of the *Pseudomonas aeruginosa lasI* gene by LasR and the *Pseudomonas* autoinducer PAI: an autoinduction regulatory hierarchy. *Journal of Bacteriology*, 177, 654-659.
- SELINGER, D. W., CHEUNG, K. J., MEI, R., JOHANSSON, E. M., RICHMOND, C. S., BLATTNER, F. R., LOCKHART, D. J. & CHURCH, G. M. (2000) RNA expression analysis using a 30 base pair resolution *Escherichia coli* genome array. *Nature Biotechnology*, 18, 1262-1268.
- SHARMA, C. M., HOFFMANN, S., DARFEUILLE, F., REIGNIER, J., FINDEISZ, S., SITTKA, A., CHABAS, S., REICHE, K., HACKERMULLER, J., REINHARDT, R., STADLER, P. F. & VOGEL, J. (2010) The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature*, 464, 250-255.
- SIMON, R., PRIEFER, U. & PUHLER, A. (1983) A broad host range mobilization system for *in vivo* genetic-engineering: transposon mutagenesis in Gram-negative bacteria. *Biotechnology*, 1, 784-791.
- SINGH, S. K., YANG, K., KARTHIKEYAN, S., HUYNH, T., ZHANG, X., PHILLIPS, M. A. & ZHANG, H. (2004) The thrH Gene Product of *Pseudomonas aeruginosa* Is a Dual Activity Enzyme with a Novel Phosphoserine:Homoserine Phosphotransferase Activity. *Journal of Biological Chemistry*, 279, 13166-13173.
- SITTKA, A., LUCCHINI, S., PAPPENFORTH, K., SHARMA, C. M., ROLLE, K., BINNEWIES, T. T., HINTON, J. C. & VOGEL, J. (2008) Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. *PLoS Genetics*, 4, e1000163-e1000163.
- SMITH, R. D. & LIGHT-WAHL, K. J. (1993) The observation of non-covalent interactions in solution by electrospray ionization mass spectrometry: Promise, pitfalls and prognosis. *Biological Mass Spectrometry*, 22, 493-501.
- SNEPPEN, K., DODD, I. B., SHEARWIN, K. E., PALMER, A. C., SCHUBERT, R. A., CALLEN, B. P. & EGAN, J. B. (2005) A Mathematical Model for Transcriptional Interference by RNA Polymerase Traffic in *Escherichia coli*. *Journal of Molecular Biology*, 346, 399-409.
- SONNLEITNER, E., ABDOU, L. & HAAS, D. (2009) Small RNA as global regulator of carbon catabolite repression in *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences*, 106, 21866-21871.
- SONNLEITNER, E., GONZALEZ, N., SORGER-DOMENIGG, T., HEEB, S., RICHTER, A. S., BACKOFEN, R., WILLIAMS, P., HÜTTENHOFER, A., HAAS, D. & BLÄSI, U. (2011) The small RNA PhrS stimulates synthesis of the *Pseudomonas aeruginosa* quinolone signal. *Molecular Microbiology*, 80, 868-885.
- SOPER, T., MANDIN, P., MAJDALANI, N., GOTTESMAN, S. & WOODSON, S. A. (2010) Positive regulation by small RNAs and the

- role of Hfq. *Proceedings of the National Academy of Sciences*, 107, 9602-9607.
- STAZIC, D., LINDELL, D. & STEGLICH, C. (2011) Antisense RNA protects mRNA from RNase E degradation by RNA-RNA duplex formation during phage infection. *Nucleic Acids Research*.
- STOCK, A. M., ROBINSON, V. L. & GOUDREAU, P. N. (2000) Two-component signal transduction. *Annual Review of Biochemistry*, 69, 183-215.
- STORK, M., DI LORENZO, M., WELCH, T. J. & CROSA, J. H. (2007) Transcription Termination within the Iron Transport-Biosynthesis Operon of *Vibrio anguillarum* Requires an Antisense RNA. *Journal of Bacteriology*, 189, 3479-3488.
- STOVER, C. K., PHAM, X. Q., ERWIN, A. L., MIZOGUCHI, S. D., WARRENER, P., HICKEY, M. J., BRINKMAN, F. S. L., HUFNAGLE, W. O., KOWALIK, D. J., LAGROU, M., GARBER, R. L., GOLTRY, L., TOLENTINO, E., WESTBROCK-WADMAN, S., YUAN, Y., BRODY, L. L., COULTER, S. N., FOLGER, K. R., KAS, A., LARBIG, K., LIM, R., SMITH, K., SPENCER, D., WONG, G. K. S., WU, Z., PAULSEN, I. T., REIZER, J., SAIER, M. H., HANCOCK, R. E. W., LORY, S. & OLSON, M. V. (2000) Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. *Nature*, 406, 959-964.
- SUH, S.-J., RUNYEN-JANECKY, L. J., MALENIK, T. C., HAGER, P., MACGREGOR, C. H., ZIELINSKI-MOZNY, N. A., PHIBBS, P. V. & WEST, S. E. H. (2002) Effect of vfr mutation on global gene expression and catabolite repression control of *Pseudomonas aeruginosa*. *Microbiology*, 148, 1561-1569.
- SWIFT, S., LYNCH, M. J., FISH, L., KIRKE, D. F., TOMAS, J. M., STEWART, G. S. A. B. & WILLIAMS, P. (1999) Quorum sensing-dependent regulation and blockade of exoprotease production in *Aeromonas hydrophila*. *Infection and Immunity*, 67, 5192-5199.
- TOMASZ, A. (1965) Control of Competent State in Pneumococcus by A Hormone-Like Cell Product - An Example for A New Type of Regulatory Mechanism in Bacteria. *Nature*, 208, 155-&.
- VALVERDE, C., HEEB, S., KEEL, C. & HAAS, D. (2003) RsmY, a small regulatory RNA, is required in concert with RsmZ for GacA-dependent expression of biocontrol traits in *Pseudomonas fluorescens* CHA0. *Molecular Microbiology*, 50, 1361-1379.
- VAN DELDEN, C. & IGLEWSKI, B. H. (1998) Cell-to-cell signaling and *Pseudomonas aeruginosa* infections. *Emerging Infectious Diseases*, 4, 551-560.
- VAN ELDERE, J. (2003) Multicentre surveillance of *Pseudomonas aeruginosa* susceptibility patterns in nosocomial infections. *The Journal of Antimicrobial Chemotherapy*, 51, 347-352.
- VENTRE, I., GOODMAN, A. L., VALLET-GELY, I., VASSEUR, P., SOSCIA, C., MOLIN, S., BLEVES, S., LAZDUNSKI, A., LORY, S. & FILLOUX, A. (2006) Multiple sensors control reciprocal expression of *Pseudomonas aeruginosa* regulatory RNA and virulence genes. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 171-176.

- VIEIRA, J. & MESSING, J. (1991) New pUC-derived cloning vectors with different selectable markers and DNA replication origins. *Gene*, 100, 189-194.
- VITALI, P., ROYO, H., SEITZ, H., BACHELLERIE, J. P., HUTTENHOFER, A. & CAVAILLE, J. (2003) Identification of 13 novel human modification guide RNAs. *Nucleic Acids Research*, 31, 6543-6551.
- VOGEL, J. & LUISI, B. F. (2011) Hfq and its constellation of RNA. *Nature Reviews Microbiology*, 9, 578-589.
- WAGNER, E. G. H. & SIMONS, R. W. (1994) Antisense RNA Control in Bacteria, Phages, and Plasmids. *Annual Review of Microbiology*, 48, 713-742.
- WAGNER, V., BUSHNELL, D., PASSADOR, L., BROOKS, A. & IGLEWSKI, B. (2003) Microarray analysis of *Pseudomonas aeruginosa* quorum-sensing regulons: effects of growth phase and environment. *Journal of Bacteriology*, 185, 2080 - 2095.
- WALDBESER, L. S., CHEN, Q. & CROSA, J. H. (1995) Antisense RNA regulation of the fatB iron transport protein gene in *Vibrio anguillarum*. *Molecular Microbiology*, 17, 747-756.
- WALDBESER, L. S., TOLMASKY, M. E., ACTIS, L. A. & CROSA, J. H. (1993) Mechanisms for negative regulation by iron of the fatA outer membrane protein gene expression in *Vibrio anguillarum* 775. *Journal of Biological Chemistry*, 268, 10433-10439.
- WANG, X., DUBEY, A. K., SUZUKI, K., BAKER, C. S., BABITZKE, P. & ROMEO, T. (2005) CsrA post-transcriptionally represses *pgaABCD*, responsible for synthesis of a biofilm polysaccharide adhesin of *Escherichia coli*. *Molecular Microbiology*, 56, 1648-1663.
- WANG, Z., GERSTEIN, M. & SNYDER, M. (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, 10, 57-63.
- WASSARMAN, K. M., REPOILA, F., ROSENOW, C., STORZ, G. & GOTTESMAN, S. (2001) Identification of novel small RNAs using comparative genomics and microarrays. *Genes & Development*, 15, 1637-1651.
- WEGELE, R., TASLER, R., ZENG, Y., RIVERA, M. & FRANKENBERG-DINKEL, N. (2004) The Heme Oxygenase(s)-Phytochrome System of *Pseudomonas aeruginosa*. *Journal of Biological Chemistry*, 279, 45791-45802.
- WEI, B., SHIN, S., LAPORTE, D., WOLFE, A. J. & ROMEO, T. (2000) Global regulatory mutations in *csrA* and *rpoS* cause severe central carbon stress in *Escherichia coli* in the presence of acetate. *Journal of Bacteriology*, 182, 1632-1640.
- WEI, B. L., BRUN-ZINKERNAGEL, A.-M., SIMECKA, J. W., PRÜSS, B. M., BABITZKE, P. & ROMEO, T. (2001) Positive regulation of motility and *flhDC* expression by the RNA-binding protein CsrA of *Escherichia coli*. *Molecular Microbiology*, 40, 245-256.
- WEILBACHER, T., SUZUKI, K., DUBEY, A. K., WANG, X., GUDAPATY, S., MOROZOV, I., BAKER, C. S., GEORGELLIS, D., BABITZKE, P. & ROMEO, T. (2003) A novel sRNA component of the carbon storage regulatory system of *Escherichia coli*. *Molecular Microbiology*, 48, 657-670.

- WEST, S. E. H., KAYE, S. A., HAMOOD, A. N. & IGLEWSKI, B. H. (1994) Characterization of *Pseudomonas aeruginosa* mutants that are deficient in exotoxin-A synthesis and are altered in expression of regA, a positive regulator of exotoxin-A. *Infection and Immunity*, 62, 897-903.
- WINSON, M. K., CAMARA, M., LATIFI, A., FOGLINO, M., CHHABRA, S. R., DAYKIN, M., BALLY, M., CHAPON, V., SALMOND, G. P. C., BYCROFT, B. W., LAZDUNSKI, A., STEWART, G. S. A. B. & WILLIAMS, P. (1995) Multiple N-acyl-L-homoserine lactone signal molecules regulate production of virulence determinants and secondary metabolites in *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences of the United States of America*, 92, 9427-9431.
- WINZER, K., FALCONER, C., GARBER, N. C., DIGGLE, S. P., CAMARA, M. & WILLIAMS, P. (2000) The *Pseudomonas aeruginosa* lectins PA-II and PA-III are controlled by quorum sensing and by RpoS. *Journal of Bacteriology*, 182, 6401-6411.
- WONG, S. M. & AKERLEY, B. J. (2005) Environmental and genetic regulation of the phosphorylcholine epitope of *Haemophilus influenzae* lipooligosaccharide. *Molecular Microbiology*, 55, 724-738.
- WORKENTINE, M. L., CHANG, L., CERI, H. & TURNER, R. J. (2009) The GacS–GacA two-component regulatory system of *Pseudomonas fluorescens*: a bacterial two-hybrid analysis. Blackwell Publishing Ltd.
- XIAO, G. P., DEZIEL, E., HE, J. X., LEPINE, F., LESIC, B., CASTONGUAY, M. H., MILOT, S., TAMPAKAKI, A. P., STACHEL, S. E. & RAHME, L. G. (2006) MvfR, a key *Pseudomonas aeruginosa* pathogenicity LTTR-class regulatory protein, has dual ligands. *Molecular Microbiology*, 62, 1689-1699.
- XU, Z., WEI, W., GAGNEUR, J., PEROCCHI, F., CLAUDER-MUNSTER, S., CAMBLONG, J., GUFFANTI, E., STUTZ, F., HUBER, W. & STEINMETZ, L. M. (2009) Bidirectional promoters generate pervasive transcription in yeast. *Nature*, 457, 1033-1037.
- YANG, H., LIU, M. Y. & ROMEO, T. (1996) Coordinate genetic regulation of glycogen catabolism and biosynthesis in *Escherichia coli* via the CsrA gene product. *Journal of Bacteriology*, 178, 1012-1017.
- ZHANG, A., WASSARMAN, K. M., ROSENOW, C., TJADEN, B. C., STORZ, G. & GOTTESMAN, S. (2003) Global analysis of small RNA and mRNA targets of Hfq. *Molecular Microbiology*, 50, 1111-1124.
- ZHANG, Y., LIU, X. S., LIU, Q.-R. & WEI, L. (2006) Genome-wide in silico identification and analysis of cis natural antisense transcripts (cis-NATs) in ten species. *Nucleic Acids Research*, 34, 3465-3475.
- ZHOU, Y., ZHANG, X. & EBRIGHT, R. H. (1993) Identification of the activating region of catabolite gene activator protein (CAP): isolation and characterization of mutants of CAP specifically defective in transcription activation. *Proceedings of the National Academy of Sciences*, 90, 6081-6085.
- ZUKER, M. (1989) On finding all suboptimal foldings of an RNA molecule. *Science*, 244, 48-52.

## 9 ANNEX

### 9.1 SINGLE-TAILED T TEST

A t-test compares two independent sample means. Assume  $\mu_1 - \mu_2$  follows a t distribution, where the assumption is the underlying distribution of the means is approximately normal, but for small populations ( $n < 20$ ), the distribution follows the Student t-distribution. We make the following hypotheses;

Null Hypothesis:  $H_0$ : There is no statistically significant difference between the mean levels of the two populations  $\mu_1 = \mu_2$

For a 1-tailed test looking at whether one distribution is significantly higher than the other, the hypothesis,  $H_1$ , is the mean level of first population is significantly greater than the mean level of the second population  $\mu_1 > \mu_2$  or  $\mu_1 - \mu_2 > 0$ .

Determine the degrees of freedom (DoF) =  $(n_1 + n_2) - 2$ , where n is the population size (or replicates) in sample 1.

n = population size

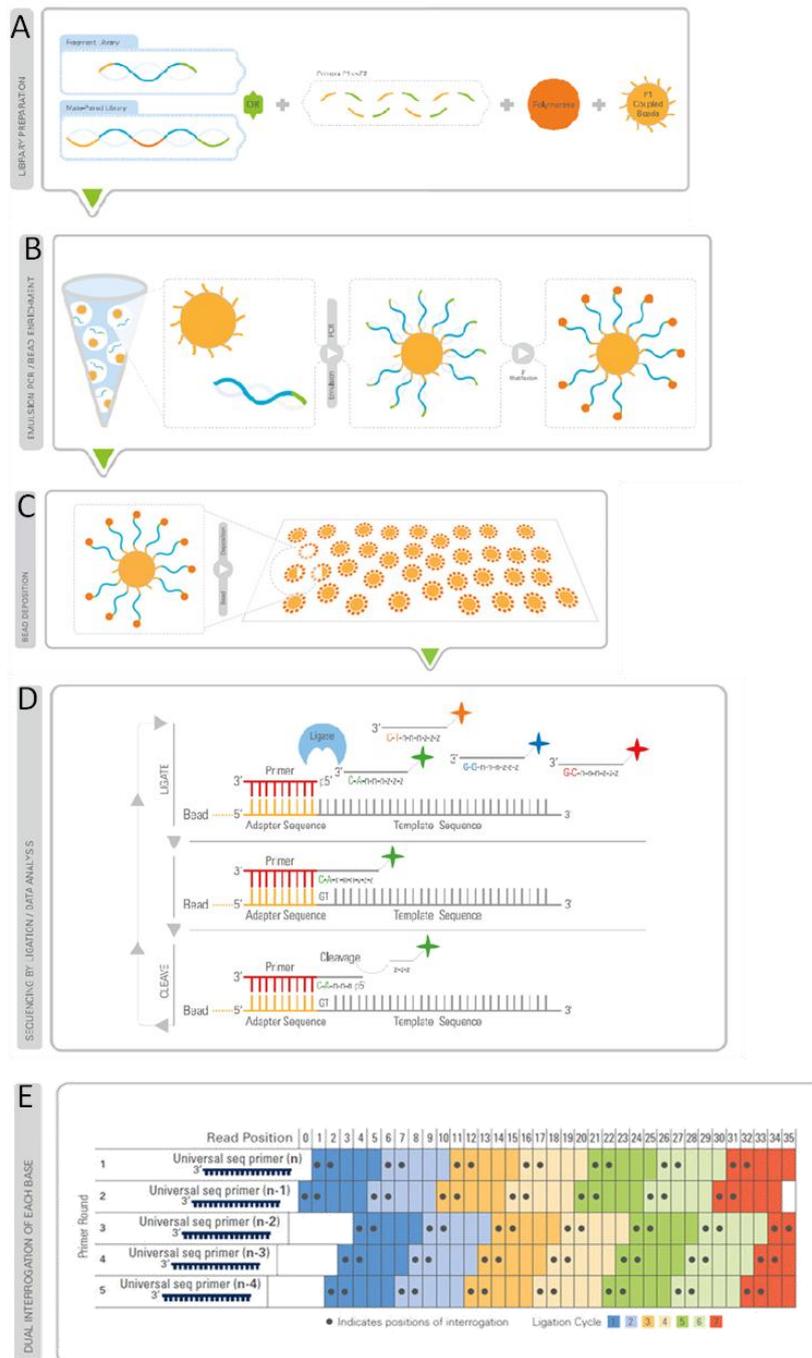
$\mu$  = population mean

$s^2$  = population variance

$$t = \frac{\mu_1 - \mu_2}{\sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}}$$

The resultant t value is determined to be significant or not by comparison to the critical value of the t-distribution corresponding to the degrees of freedom at a chosen percentile. The critical values used in this thesis correspond to a single-tailed distribution to 5 %.

## 9.2 APPENDIX I



Overview of SOLiD Sequencing system (Applied Biosciences)

Sequencing fragment library was prepared (A) for SOLiD™ System. There are two choices of library, sequence-fragment and mate-pair depending on the application to be performed and the information required, in this case, sequencing fragments. The clonal bead populations are prepared (B) in

microreactors containing template, PCR reaction components, beads, and primers. After PCR and template denaturation, bead enrichment is performed to separate beads with extended templates from undesired beads. The template on the selected beads undergoes a 3' modification to allow covalent attachment to the slide.

The 3' modified beads are deposited onto a glass slide (C). Primers hybridize to the P1 adapter sequence on the templated beads (D) and a set of four fluorescently labeled di-base probes compete for ligation to the sequencing primer. Specificity of the di-base probe is achieved by interrogating every 1st and 2nd base in each ligation reaction. Multiple cycles of ligation, detection and cleavage are performed with the number of cycles determining the eventual read length. Following a series of ligation cycles, the extension product is removed and the template is reset with a primer complementary to the n-1 position for a second round of ligation cycles.

Five rounds of primer reset are completed for each sequence tag (E). Through the primer reset process, virtually every base is interrogated in two independent ligation reactions by two different primers. For example, the base at read position 5 is assayed by primer number 2 in ligation cycle 2 and by primer number 3 in ligation cycle 1.

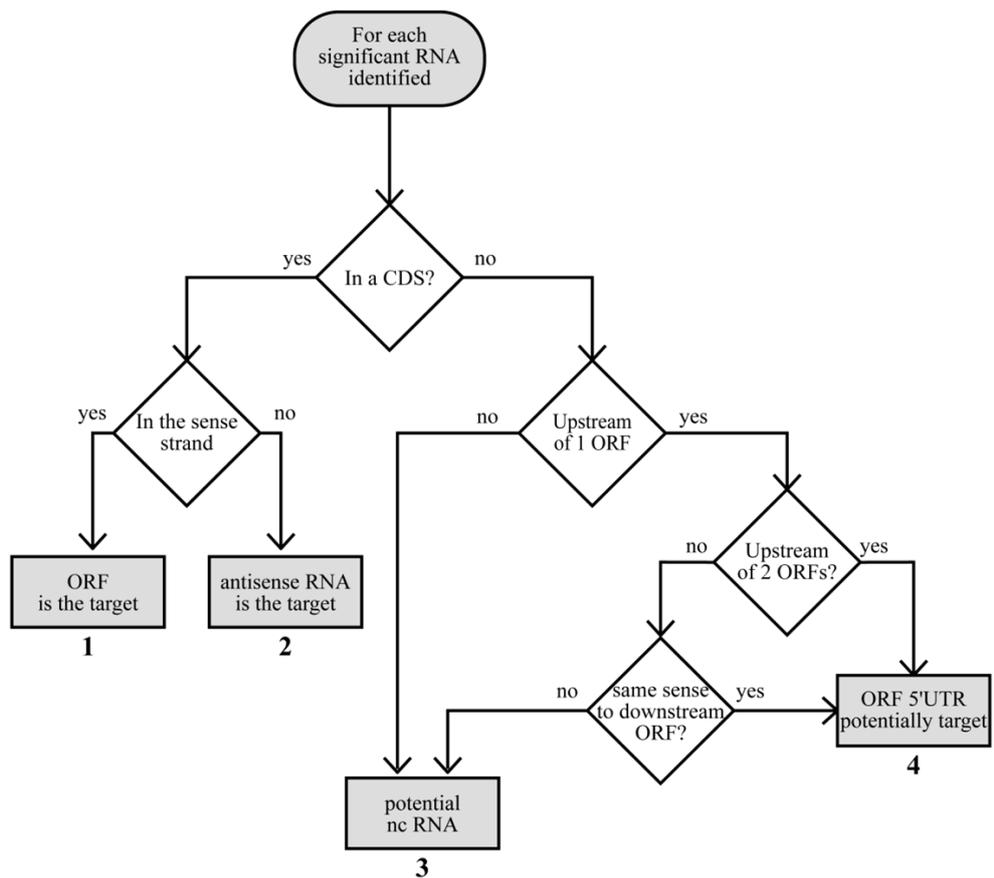
For more information:

<http://www.appliedbiosystems.com/absite/us/en/home/applications-technologies/solid-next-generation-sequencing/next-generation-systems.html>

### 9.3 APPENDIX II

#### RNA classification for Deep-Seq transcripts

Result interpretation for RNA transcript classification according to a logic flowchart diagram into 1 of 4 groups according to their most likely function. The group location or topology of the transcript is identified by applying the RNA of interest to the flowchart, if it is located in a CDS. If the RNA is allocated to topology 1: the ORF is the target, topology 2: the antisense RNA is the target, topology 3: the transcript corresponds to a ncRNA and topology 4: the ORF 5'UTR is the potential target.



## 9.4 APPENDIX III

### Deep-sequencing RsmN enriched target transcripts

N: Negative, P: Positive strands, CDS: Coding Sequence. The average is the enrichment value multiplied by 100, only averages >200 have been selected.

PA Number	Gene Name	Strand	Topology	Average	Comment
PA5128*	<i>secB</i>	N	1	9999.00	Secretion protein
PA2958.1	<i>rgsA</i>	P	1	5092.83	sRNA Gac-controlled indirectly
PA0362	<i>fdx1</i>	N	1	4848.60	Ferredoxin [4Fe-4S]
PA1754	<i>cysB</i>	P	1	4010.36	Transcriptional regulator (sulphur metabolism)
PA0044	<i>exoT</i>	P	1	4009.47	Exoenzyme T
PA1709	<i>popD</i>	P	1	3511.87	Translocator outer membrane protein PopD precursor
PA4726.11	<i>crcZ</i>	P	1	3091.22	Antagonistic RNA for catabolite repression control protein Crc
PA1871	<i>lasA</i>	P	1	2820.06	LasA protease precursor
PA0432	<i>sahH</i>	N	1	1908.25	S-adenosyl-L-homocysteine hydrolase
PA0524*	<i>norB</i>	P	1	1875.14	Nitric-oxide reductase subunit B
PA1708	<i>popB</i>	P	1	1512.03	
PA5040	<i>pilQ</i>	N	1	1340.63	Type 4 fimbrial biogenesis outer membrane protein PilQ precursor
PA1776/PA1777	<i>sigX/oprF</i>	P	1_4	1276.05	ECF sigma factor/Major porin and structural outer membrane porin OprF precursor
PA1151*	<i>imm2</i>	P	1	1131.09	Pyocin S2 immunity protein
PA0766*	<i>mucD</i>	P	1	924.44	Serine protease MucD precursor
PA4428	<i>sspA</i>	N	1	843.76	Stringent starvation protein A
PA0962		N	1_4	803.09	Probable dna-binding stress protein
PA0969*	<i>tolQ</i>	P	1	565.88	TolQ protein

---

PA4175*	<i>Piv</i>	P	1	521.06	Protease IV
PA2830*	<i>htpX</i>	P	1	481.73	Heat shock protein
PA1455*	<i>fliA</i>	P	1	477.21	Sigma factor
PA1098*	<i>fleS</i>	P	1	476.20	Two-component sensor
PA0427*	<i>oprM</i>	P	1	429.97	Major intrinsic multiple antibiotic resistance efflux outer membrane protein OprM precursor
PA3104*	<i>xcpP</i>	P	1	408.37	Secretion protein
PA3813/PA3814	<i>iscU/iscS</i>	N	1	384.87	Probable iron-binding protein/L-cysteine desulfurase
PA5565*	<i>gidA</i>	N	1	368.55	Glucose-inhibited division protein A
PA4403*	<i>secA</i>	N	1	349.71	Secretion protein
PA1087*	<i>flgL</i>	P	1	341.04	Flagellar hook-associated protein type 3
PA0396*	<i>pilU</i>	P	1	217.35	Twitching motility protein
PA1001/PA1002*	<i>phnA/phnB</i>	P	1	208.60	Anthranilate synthase component I/ Anthranilate synthase component II
PA4315*	<i>mvaT</i>	P	1	206.30	Transcriptional regulator MvaT, P16 subunit
PA1432*	<i>lasI</i>	P	1	203.24	Autoinducer synthesis protein
PA5563	<i>Soj</i>	P	2	9999.00	Chromosome partitioning protein
PA5213*	<i>P1 gcvP1</i>	P	2	9999.00	Glycine cleavage system protein
PA5446*	<i>wbpZ</i>	P	2	9999.00	Glycosyltransferase
PA1674*		P	2	5467.37	GTP cyclohydrolase I precursor
PA5474		N	2	2223.73	Probable metalloprotease
PA0654*	<i>sped</i>	N	2	1491.89	S-adenosylmethionine decarboxylase proenzyme
PA1546*	<i>hemN</i>	P	2	736.46	Oxygen-independent coproporphyrinogen III oxidase
PA1002	<i>phnB</i>	N	2	447.51	Anthranilate synthase component II

---

PA2423/PA2424*		P	3	369.73	Intergenic PA2423-PA2424
PA0652*	<i>Vfr</i>	N	4_1	9680.80	Transcriptional regulator
PA0519*	<i>nirS</i>	N	4_1	7748.71	Nitrite reductase precursor
PA5239*	<i>Rho</i>	N	4_1	4107.35	Transcription termination factor
PA3126	<i>ibpA</i>	N	4_1	4068.14	Heat-shock protein
PA0355*	<i>pfpI</i>	P	4_1	3716.53	Protease
PA3266*	<i>capB</i>	P	4_1	2982.23	Cold acclimation protein B
PA1178	<i>oprH</i>	P	4_1	2482.81	PhoP/Q and low Mg <sup>2+</sup> inducible outer membrane protein H1 precursor
PA4205	<i>mexG</i>	P	4_1	2465.75	Hypothetical protein
PA2570	<i>lecA</i>	N	4_1	2266.79	intergenic PA2570 - CDS PA2570
PA1544	<i>Anr</i>	N	4_1	2225.67	Transcriptional regulator
PA0852*	<i>cbpD</i>	N	4_1	2152.75	Chitin-binding protein CbpD precursor
PA1003	<i>pqsR (mvrF)</i>	N	4_1	2131.77	Transcriptional regulator
PA5427	<i>adhA</i>	P	4_1	1979.30	Alcohol dehydrogenase
PA1092	<i>fliC</i>	P	4_1	1902.52	Flagellin type B
PA3361	<i>lecB</i>	P	4_1	1879.25	Fucose-binding lectin PA-IIL
PA4385/PA4386		N	4_1	1699.74	GroEL protein groEL / groES
PA3351	<i>flgM</i>	P	1_4_1	1579.24	
PA3385	<i>amrZ</i>	P	4_1	1480.38	Alginate and motility regulator Z
PA0905	<i>rsmA</i>	P	4_1	1324.57	Regulator of secondary metabolites
PA4922	<i>Azu</i>	N	4_1	1308.93	Azurin precursor
PA5253	<i>algP</i>	N	4_1	1236.01	Alginate regulatory protein
PA4067	<i>oprG</i>	P	4_1	1174.02	Outer membrane protein OprG precursor
PA5170-3	<i>arcDABC</i>	P	4_1	1153.77	Arginine/ornithine antiporter

PA3724	<i>lasB</i>	N	4_1	1150.25	Elastase
PA2853	<i>oprI</i>	P	4_1	1002.37	Outer membrane lipoprotein OprI precursor
PA3326	<i>clpP2</i>	N	4_1	1002.24	
PA0762-PA0764	<i>algU/mucA/mucB</i>	P	4_1_4	903.70	Sigma factor/Anti-sigma factor / Negative regulator for alginate biosynthesis
PA4778*	<i>cueR</i>	P	4_1	832.40	
PA1770	<i>ppsA</i>	P	4_1	682.85	phosphoenolpyruvate synthase
PA3476-9	<i>rhlR/rhlAB</i>	N	4_1	674.47	Rhamnosyltransferase chain B
PA1454*	<i>fleN</i>	P	4_1	640.11	Flagellar synthesis regulator
PA1985	<i>pqqA</i>	P	4_1	565.96	Pyrroloquinoline quinone biosynthesis protein A
PA1094	<i>fliD</i>	P	4_1	545.02	Flagellar capping protein
PA1094	<i>fliD</i>	P	4_1	544.61	Flagellar capping protein
PA2231	<i>pslA</i>	P	4_1	532.70	
PA0576	<i>rpoD</i>	N	4_1	477.74	
PA0888	<i>aotJ</i>	P	4_1	441.12	Arginine/ornithine binding protein
PA5261/PA5262*	<i>algR/algZ</i>	N	1_4_1	383.28	Alginate biosynthesis protein
PA0973/PA0974	<i>oprL</i>	P	4_1	352.79	Peptidoglycan associated lipoprotein OprL precursor /conserved HP
PA0996-1000	<i>pqsABCDE</i>	P	4_1	337.36	
PA3476/PA3477	<i>rhlR/rhlI</i>	N	4_1	326.81	Transcriptional regulator / autoinducer synthesis protein
PA5261/PA5262	<i>algR/algZ</i>	N	4_1	320.02	Alginate biosynthesis regulatory protein
PA2622	<i>cspD</i>	P	4_1	260.18	Cold-shock protein
PA5183_PA5184	<i>rsmN</i>	N	4_1	250.95	RsmN
PA0408	<i>pilG</i>	P	4_1	213.90	Twitching motility protein

## 9.5 APPENDIX IV

### Deep-sequencing RsmA enriched target transcripts

N: Negative, P: Positive strands, CDS: Coding Sequence. The average is the enrichment value multiplied by 100, only averages >200 have been selected.

PA number	Gene	Strand	Topology	Average	Comment
PA1003	<i>mvfR (pqsR)</i>	N	1	9999.00	Transcriptional regulator MvfR (PqsR)
PA4969	<i>cpdA</i>	N	1	9999.00	Cyclic AMP (cAMP) Phosphodiesterase, CpdA
PA3820	<i>secF</i>	N	1	9926.12	Secretion protein
PA3821	<i>secD</i>	N	1	9220.76	Secretion protein
PA0928	<i>gacS</i>	P	1	9204.54	Sensor/response regulator hybrid gacS
PA0893	<i>argR</i>	P	1	6103.00	Transcriptional regulator ArgR
PA0425	<i>mexA</i>	P	1	5835.52	Resistance-Nodulation-Cell Division (RND) multidrug efflux membrane fusion protein MexA precursor
PA0764	<i>mucB</i>	P	1	4859.75	Negative regulator for alginate biosynthesis MucB
PA2399	<i>pvdD</i>	N	1	4479.61	Pyoverdine synthetase D
PA1001/PA1002	<i>phnAB</i>	P	1	3986.75	Anthranilate synthase component I/II
PA3724	<i>lasB</i>	N	1	3724.90	Elastase LasB
PA2958.1	<i>rgsA</i>	P	1	1853.49	sRNA Gac-controlled indirectly
PA0609	<i>trpE</i>	N	2	9999.00	Anthranilate synthetase component I
PA5495	<i>thrB</i>	N	2	9999.00	Homoserine kinase
PA1757	<i>thrH</i>	N	2	9999.00	Homoserine kinase
PA1871	<i>lasA</i>	N	2	9999.00	LasA protease precursor
PA5128	<i>secB</i>	P	2	9999.00	Secretion protein SecB
PA1003	<i>mvfR (pqsR)</i>	P	2	9441.53	Transcriptional regulator MvfR (PqsR)

PA0928	<i>gacS</i>	P	2	3351.49	Sensor/response regulator hybrid
PA1898	<i>qscR</i>	N	2	2140.92	Quorum-sensing control repressor
PA0291/PA0290	<i>oprE/HP</i>	N	3	8783.46	Intergenic Anaerobically-induced outer membrane porin OprE precursor/HP
PA2424/PA2425		P	3	2464.58	Intergenic PvdL/PvdG
PA2193	<i>hcnA</i>	P	4_1	9392.835	Hydrogen cyanide synthase
PA2385	<i>pvdQ</i>	N	4_1	9999.00	3-oxo-C12-homoserine lactone acylase PvdQ
PA0424	<i>mexR</i>	N	4_1	9355.57	Multidrug resistance operon repressor MexR
PA1898	<i>qscR</i>	P	4_1	9046.99	Quorum-sensing control repressor
PA2396	<i>pvdF</i>	N	4_1	8857.95	Pyoverdine synthetase F
PA1727	<i>mucR</i>	N	4_1	8846.57	MucR
PA2570	<i>lecA</i>	N	4_1	8479.50	LecA
PA4704	<i>cbpA</i>	P	4_1	8424.55	cAMP-binding protein A
PA0517- PA0519	<i>nirCMS</i>	N	4_1	8080.31	Heme d1 biosynthesis protein NirC (probable c-type cytochrome precursor )/cytochrome c-551 precursor/nitrite reductase precursor nirS
PA3974	<i>ladS</i>	N	4_1	7147.96	Lost Adherence Sensor, LadS
PA0527.1	<i>rsmY</i>	P	4_1_3_2	6329.04	Regulatory RNA RsmY
PA3361	<i>lecB</i>	P	4_1	6023.61	Fucose-binding lectin PA-IIL
PA4778	<i>cueR</i>	P	4_1	5713.72	CueR
PA1985	<i>pqqA</i>	P	4_1	5301.95	Pyroloquinoline quinone biosynthesis protein A
PA0652	<i>vfr</i>	N	4_1	4532.75	Transcriptional regulator Vfr
PA1004	<i>nadA</i>	P	4_1	4406.13	Quinolinate synthetase A
PA1178	<i>oprH</i>	P	4_1	4087.01	PhoP/Q and low Mg <sup>2+</sup> inducible outer membrane protein H1 precursor
PA0905	<i>rsmA</i>	P	4_1	3650.58	RsmA, regulator of secondary metabolites

PA3621.1	<i>rsmZ</i>	N	4_1	3357.82	Regulatory RNA RsmZ
PA1544	<i>anr</i>	N	4_1	3348.75	Transcriptional regulator Anr
PA0576	<i>rpoD</i>	N	4_1	3328.50	Sigma factor RpoD
PA4209	<i>phzM</i>	N	4_1	3288.93	Probable phenazine-specific methyltransferase
PA5253	<i>algP</i>	N	4_1	3183.45	Alginate regulatory protein AlgP
PA5261/PA5262	<i>algR/alaZ</i>	N	4_1	2965.92	Alginate biosynthesis regulatory protein AlgR /alginate biosynthesis protein AlgZ/FimS
PA0996-PA1000	<i>pqsABCDE</i>	P	4_1	2918.45	Probable coenzyme A ligase pqsABCDE
PA1546	<i>hemN</i>	N	4_1	2717.87	Oxygen-independent coproporphyrinogen III oxidase
PA1092	<i>fliC</i>	P	4_1	2712.15	Flagellin type B
PA5040-PA5044	<i>pilMNOPQ</i>	N	4_1	2512.01	Type 4 fimbrial biogenesis outer membrane protein PilQ precursor
PA4726.11	<i>crcZ</i>	P	4_1	2389.58	Antagonistic RNA for catabolite repression control protein Crc
PA0432	<i>sahH</i>	N	4_1	2208.05	S-adenosyl-L-homocysteine hydrolase
PA5183/PA5184	<i>rsmN</i>	N	4_1	2136.77	
PA3476	<i>rhII</i>	N	4_1	2117.86	Autoinducer synthesis protein RhII
PA4315	<i>mvaT</i>	P	4_1	2041.82	Transcriptional regulator MvaT, P16 subunit
PA5239	<i>rho</i>	N	4	1921.89	Transcription termination factor Rho
PA0362	<i>fdx1</i>	N	4_1	1844.03	Ferredoxin [4Fe-4S]
PA4403	<i>secA</i>	N	4_1	1824.03	Secretion protein
PA1900	<i>phzB2</i>	P	4_1	1813.83	Probable phenazine biosynthesis protein
PA3351	<i>flgM</i>	P	4_1	1698.07	FlgM
PA3385	<i>amrZ</i>	P	4_1	1393.14	Alginate and motility regulator Z
PA4526/PA4527	<i>pilB/pilC</i>	P	1_4_1	1324.14	Type 4 fimbrial biogenesis protein PilB/pilin biogenesis protein PilC
PA1430	<i>lasR</i>	P	4_1	1118.05	Transcriptional regulator LasR

---

PA3724	<i>lasB</i>	N	4	882.50	Elastase LasB
PA4922	<i>azu</i>	N	4_1	446.86	Azurin precursor
PA4944	<i>hfq</i>	N	4_1	323.11	Hfq
PA1432	<i>lasI</i>	P	4_1	316.04	Autoinducer synthesis protein LasI
PA5495	<i>thrB</i>	P	4_1	1646.25	Homoserine Kinase

---

## 9.6 APPENDIX V

Deep-sequencing RsmA and RsmN transcript comparison table for genes of interest.

N: Negative, P: Positive strands, CDS: Coding sequence, \*appears only in one data set. The average is the enrichment/depleted value multiplied by 100.

PA Number	Gene	Strand	RsmA		RsmN	
			Topology	Average	Topology	Average
PA1003	<i>mvfR (pqsR)</i>	N	1	9999.00	4_1	2131.77
PA1871	<i>lasA</i>	N	2	9999.00	1	2820.06
PA5128	<i>secB</i>	P	2	9999.00	1	9999.00
PA2570	<i>lecA</i>	N	4_1	8479.50	4_1	2277.80
PA0517-PA0519	<i>nirCMS</i>	N	4_1	8080.31	4_1	7748.71
PA0527.1	<i>rsmY</i>	P	4_1_3_2	6329.04	4_1	198.75
PA3361	<i>lecB</i>	P	4_1	6023.61	4_1	1879.25
PA4778	<i>cueR</i>	P	4_1	5713.72	4_1	832.40
PA1985	<i>pqqA</i>	P	4_1	5301.95	4_1	565.96
PA0764	<i>mucB</i>	P	1	4859.75	4_1_4	903.70
PA0652	<i>vfr</i>	N	4_1	4532.75	4_1	9680.80
PA1178	<i>oprH</i>	P	4_1	4087.01	4_1	2482.81
PA1001/PA1002	<i>phnAB</i>	P	1	3986.75	1	208.60
PA3724	<i>lasB</i>	N	1	3724.90	4_1	1150.25
PA0905	<i>rsmA</i>	P	4_1	3650.58	4_1	1324.57
PA3621.1	<i>rsmZ</i>	N	4_1	3357.82	4_1	330.13
PA1544	<i>anr</i>	N	4_1	3348.75	4_1	2225.67
PA0576	<i>rpoD</i>	N	4_1	3328.50	4_1	477.74
PA5253	<i>algP</i>	N	4_1	3183.45	4_1	1236.01
PA5261/PA5262	<i>algR/alaZ</i>	N	4_1	2965.92	1_4_1	383.28
PA0996-PA1000	<i>pqsABCDE</i>	P	4_1	2918.45	4_1	337.3557
PA1092	<i>fliC</i>	P	4_1	2712.15	4_1	1916.18
PA5040-PA5044	<i>pilMNOPQ</i>	N	4_1	2512.01	1	1340.63
PA2424/PA2425		P	3	2464.58	3	369.73
PA4726.11	<i>crcZ</i>	P	4_1	2389.58	1	3091.22
PA0432	<i>sahH</i>	N	4_1	2208.05	1	1908.25
PA3476	<i>rhII</i>	N	4_1	2117.86	4_1	326.81
PA4315	<i>mvaT</i>	P	4_1	2041.82	1	206.30
PA5239	<i>rho</i>	N	4	1921.89	4_1	4107.35
PA2958.1	<i>rgsA</i>	P	1	1853.49	1	5092.83
PA0362	<i>fdxI</i>	N	4_1	1844.03	1	4848.60
PA4403	<i>secA</i>	N	4_1	1824.03	1	349.71
PA1900	<i>phzB2</i>	P	4_1	1813.83	1	2798.67
PA3351	<i>flgM</i>	P	4_1	1698.07	1_4_1	1579.24
PA3385	<i>amrZ</i>	P	4_1	1393.14	4_1	1480.38
PA1430	<i>lasR</i>	P	4_1	1118.05	4_1	38.03
PA3724	<i>lasB</i>	N	4	882.50	4_1	1150.25

PA4922	<i>azu</i>	N	4_1	446.86	4_1	1308.93
PA4944	<i>hfq</i>	N	4_1	323.11	4_1	471.69
PA1432	<i>lasI</i>	P	4_1	316.04	1	203.24
PA1776/PA1777	<i>sigX/oprR</i>	P	1_4_1	36.21	1_4	1276.05
PA3305.1	<i>phrS</i>	N	4_1	26.53	1	132.08
PA0376	<i>rpoH</i>	P	4_1	17.88	1_4	165.00
PA5332	<i>crc</i>	P	4_1	9.77	4_1	119.36*
PA1002	<i>phnB</i>	N	2	9.31	2	447.51
PA1430	<i>lasR</i>	N	2	6.63	4_1	38.03