Imperial College London

Department of Medicine

Division of Infectious Diseases

Functional analysis of the DNA-binding domain of the *Staphylococcus aureus* master virulence transcription factor AgrA

Sophie Nicod

Submitted for the degree of Doctor of Philosophy

September 2014

ABSTRACT

The LytTR domain is a DNA-binding domain that is present in 5% of all DNA-binding transcription factors (TFs). The LytTR family of DNA-binding TFs is poorly characterised but crucially important as LytTR domain-containing TFs (LDCTFs) are disproportionately associated with virulence genes regulation. Although some members of the family have been extensively studied, how LDCTFs activate transcription remains to be determined.

The accessory gene regulator A (AgrA) in *Staphylococcus aureus* (*S. aureus*) is the master TF responsible for the activation of transcription at the *agr* operon, a key regulatory locus involved in the switch between early colonisation and the spread of infection. AgrA is an LDCTF and the only one for which the crystal structure of the LytTR domain in complex with DNA is available.

Here an experimental system was developed to study AgrA activity *in vivo* in order to try to unravel the mechanism of transcription activation by LDCTFs using AgrA as a model.

A systematic alanine scanning mutagenesis of the LytTR domain of AgrA was performed and the mutants were characterised *in vivo* and *in vitro*.

Overall, the results identify amino acid (aa) residues in the LytTR domain of AgrA that are important for DNA-binding and transcription activation and provide novel molecular and mechanistic insights on how AgrA and other LDCTFs activate transcription. More specifically, a key residue for transcription activation by AgrA, Y229, was identified. This residue is important for transcription activation but is not involved in the ability to bend DNA. These findings challenge the current assumption that the DNA bending activity of AgrA is responsible for its ability to activate transcription. We therefore propose a new model for transcription activation by AgrA based on our findings.

DECLARATION OF AUTHORSHIP

I confirm that this thesis entitled "Functional analysis of the DNA-binding domain of the *Staphylococcus aureus* master virulence transcription factor AgrA" is written entirely by me and that the research to which it refers to is my own. I confirm that any ideas or quotations from the work of others have always been clearly referenced and that all main sources of help have been acknowledged.

Sophie Nicod

COPYRIGHT DECLARATION

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work

ACKNOWLEDGMENTS

First of all, I would like to thank my supervisor Ramesh. Thanks for always believing in me and pushing me to accomplish more. But most of all, I want to thank you for your genuine excitement and interest for my project. It was nice to know there was at least one other person in the world that would be as excited as I was about a good or interesting result.

Then I would like to thank all the Wigs, past and present for their help and support during my PhD and for being such amazing friends. A few words for: Ellen who was a joy to work with and helped me a lot during the past few years, Dan for always being supportive, especially when I was writing up, Aisha who has been a very good friend for the past year, Carol for all her help, especially at the start of my PhD and Andrey, for being Drey…

I am grateful to my friends in the CMBI, past and present: Nathalie, Erin, Michaella and Kim who have always been very supportive and fun to be around but most of all, amazing friends.

And thank you to the whole CMBI3/4, it has been a joy to be around you guys and I could not have imagined a better place to do my PhD. I will remember all of our social events, Thirsty Thursdays, Friday cakes but mostly just the friendly day to day atmosphere.

I also want to thank my friends back home, Mélanie and Valerie for their friendship and support.

A big thank you to my parents who have always supported me. I would not be where I am without you. Thank you for the way you raised me: to believe that if you work hard everything is possible.

I am grateful to my sisters who are always there when I need them and my grandmothers for always checking up on me.

Last but not least, I want to thank my boyfriend Olivier. Thank you for your help and support during my PhD. Thank you for taking care of me when I was writing up, making sure I was ok and doing everything possible to help me. Thank you also for your help with my thesis and for always being there for me for the past 8 years.

ABREVIATIONS

µ	Micro, 10^{-6}
µCi	Microcurie
µg	Microgram
µM	Micromolar
^{32}P	Radio-labelled Phosphate
A	Adenine
aa	Amino acid
agr	Accessory gene regulator operon
agr IR	Accessory gene regulator operon intergenic region
AIP	Autoinducing peptide
aMD	Accelerated Molecular dynamics
BACTH	Bacterial adenylate cyclase two-hybrid
bp	Base pair(s)
C	Cytosine
CA	Catalytic and ATP binding
CBD	Chitin-binding domain
Ci	Curie
DHp	Dimerisation and histidine phosphotransfer
dNTP	Deoxy nucleotide triphosphate
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EMSA	Electophoretic mobility shift assay
Eσ	RNAp holoenzyme
g	Gram
G	Guanine
GFP	Green fluorescent protein
GFP-Fluo	GFP fluorescence units
h	Hour(s)
HK	Histidine protein kinase
HTH	Helix-turn-helix
kb	Kilobase
kDa	Kilodalton
LDCTF	LytTR domain-containing transcription factor
M	Molar
Mab	Monoclonal antibody
MD	Molecular dynamics
mg	Milligram
min	Minute(s)
ml	Millilitres
mM	Millimolar
MRSA	Methycilin resistant *S. aureus*

n	Nano, 10^{-9}
ng	Nanogram
nm	Nanometre
NTP	Nucleoside triphosphate
OD	Optical density
P	Promoter
p	Pico,10^{-12}
REC	Receiver
RNA	Ribonucleic acid
RNAp	DNA-dependent RNA-polymerase(s)
RPc	Closed promoter complex(es)
rpm	Revolutions per minute
RPo	Open promoter complex(es)
RR	Response regulator
RT	Room temperature
S. aureus	*Staphylococcus aureus*
SD	Shine-Dalgarno sequence
SDS	Sodium dodecyl sulphate
sec	Seconds
spp.	Species
T	Thymine
TAD	Transcription activation deficient
TBS	Tris-buffered saline
TCS	Two-component systems
TEMED	N',N',N',N'-Tetramethylethylenediamine
TF	Transcription factor
Tm	Melting temperature
Tris	Tris(hydroxymethyl)aminomethane
U	Uridine
v/v	Volume:Volume
VISA	Vancomycin resistant *S. aureus*
VRSA	Vancomycin intermediate *S. aureus*
w/v	Weight:Volume
WT	Wild-type
x g	Relative centrifugal force
α	RNAp alpha subunit
β	RNAp beta subunit
β'	RNAp beta prime subunit
δ	RNAp delta subunit
σ	Sigma factor
ω	RNAp omega subunit

Amino Acid	3-Letter Code	1-Letter Code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartate	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamate	Glu	E
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

CONTENTS

ABSTRACT 2

DECLARATION OF AUTHORSHIP 3

COPYRIGHT DECLARATION 3

ACKNOWLEDGMENTS 4

ABREVIATIONS 5

CONTENTS 8

LIST OF FIGURES 17

LIST OF TABLES 22

CHAPTER ONE 23

1. Introduction 23

1.1. Transcription 23

1.1.1. The promoter 25

1.1.2. RNAp 27

1.1.3. The bacterial RNAp 28

1.1.4. The sigma factors (σ) 29

1.2. Transcription activation at σ^{70} promoters 31

1.2.1. Activation by promoter remodelling 31

1.2.2. Activation by contacting the RNAp 32

1.2.2.1. Class I activation 33

1.2.2.2. Class II activation 33

1.2.2.3. Class III activation 34

1.2.2.4. Activation by pre-recruitment 35

1.3. Two-component systems (TCS) 36

1.3.1. Histidine Kinases (HKs) 37

1.3.2. Response Regulators (RRs) 39

1.3.2.1. The LytTR family of DNA-binding RRs 41
1.4. *Staphylococcus aureus* 43
1.4.1. Historical background 44
1.4.2. *S. aureus* and antibiotic resistance 44
1.4.3. Pathogenicity of *S. aureus* 45
1.4.4. TCS in *S. aureus* 46
1.4.4.1. The AgrCA TCS 48
1.4.4.2. The LytSR TCS 48
1.4.4.3. The GraSR TCS 48
1.4.4.4. The SaeSR TCS 49
1.4.4.5. The DesKR TCS 49
1.4.4.6. The ArlSR TCS 49
1.4.4.7. The SrrBA TCS 49
1.4.4.8. The PhoPR TCS 50
1.4.4.9. The AirSR TCS 50
1.4.4.10. The VraSR TCS 51
1.4.4.11. The KdpDE TCS 51
1.4.4.12. The HssSR TCS 51
1.4.4.13. The NreBC TCS 52
1.4.4.14. The BceSR TCS 52
1.4.4.15. The WalKR TCS 52
1.4.4.16. Summary of *S. aureus* TCS 53
1.5. The accessory gene regulator (*agr*) operon 53
1.5.1. *agr* interference 55
1.5.2. AgrB 56
1.5.3. AgrD 57

1.5.4. AgrC....59
1.5.5. AgrA....62
1.5.5.1. The AgrA CheY-like receptor domain....62
1.5.5.2. The AgrA LytTR domain....64
1.5.5.3. AgrA and transcription activation at the P2 and P3 promoters....64
1.5.5.4. AgrA regulation outside of the *agr* operon....67
1.5.5.5. Interactions of AgrA with AgrC....67
1.5.6. RNAIII....68
1.5.7. The regulation of the *agr* operon....70
1.5.7.1. SarA....71
1.5.7.2. SarR....72
1.5.7.3. SarU....72
1.5.7.4. SarZ....72
1.5.7.5. MgrA....72
1.5.7.6. CcpA....73
1.5.7.7. SarX....73
1.5.7.8. SarT....73
1.5.7.9. RsR....73
1.5.7.10. σ^B....73
1.5.7.11. CodY....73
1.6. Objectives....75
CHAPTER TWO....76
2. Materials and methods....76
2.1. General Methods....76
2.1.1. Composition of growth media....76
2.1.2. GFP growth curves using a multiwall plate reader....77

2.1.3. Blood agar plate heamolysis assay 77
2.1.4. General lab reagents and solutions 77
2.1.5. Bacterial strains and plasmids 79
2.2. General DNA methods 80
2.2.1. Purification of chromosomal DNA 80
2.2.2. Purification of plasmid DNA from *E. coli* (XL1 blue or DC10B) 81
2.2.3. Purification of plasmid DNA from *S. aureus* (RN4220) 82
2.2.4. Restriction analysis of DNA and site specific DNA digestion 82
2.2.5. Agarose gel electrophoresis 82
2.2.6. Gel extraction and PCR purification of DNA 83
2.2.7. Polymerase chain reaction (PCR) 83
2.2.8. Site directed mutagenesis (SDM) 84
2.2.9. Sequencing 85
2.2.10. pGEM-T easy vector ligation 85
2.2.11. Vector ligation 86
2.2.12. Chemically Competent cells preparation (*E. coli* cells) 86
2.2.13. Heat shock transformation 86
2.2.14. Electro-competent cells preparation (*S. aureus* cells) 87
2.2.15. Electroporation 87
2.3. General protein methods 88
2.3.1. Protein Expression 88
2.3.1.1. The pTYB2 and the pET systems 88
2.3.1.2. Expression of [6His]-AgrA in pET28b^{+} 89
2.3.1.3. Expression of AgrC-HK in pET28b^{+} 89

2.3.1.4. Expression of AgrA , $AgrA_{L171A}$, $AgrA_{E181A}$, $AgrA_{H200A}$, $AgrA_{Y229A}$ and $AgrA_{R233A}$ in pSN-*agrA,* pSN-*agrA* $_{L171A}$, pSN-*agrA*$_{E181A}$, pSN-*agrA*$_{H200A}$, pSN-*agrA*$_{Y229A}$ *and* pSN-*agrA*$_{R233A}$ 89

2.3.2. Expression of the S. aureus core RNA polymerase (RNAp) 90

2.3.3. SDS-Polyacrylamide gel electrophoresis (SDS-PAGE) 90

2.3.4. Protein purification 92

2.3.4.1. Nickel affinity chromatography 92

2.3.4.2. Immunoaffinity chromatography 93

2.3.4.3. Intein mediated purification with an affinity chitin-binding tag (IMPACT) 94

2.3.5. Determination of protein concentration by Bradford microassay 94

2.3.6. Polyclonal antibody production 95

2.3.7. Affinity purification of antibody serum 95

2.3.8. Western blotting 96

2.4. General biochemical methods 97

2.4.1. Probe labelling with ^{32}P 97

2.4.2. Electrophoretic gel mobility shift assay (EMSA) with acetyl phosphate activation 97

2.4.3. Electrophoretic gel mobility shift assay (EMSA) with AgrC-HK activation 98

2.4.4. Short RNA synthesis assay with AgrC-HK activation 98

2.4.5. Full length *in vitro* transcription assay with acetyl phosphate activation 99

2.4.6. DNA bending assays 99

2.4.7. Acetyl Phosphate phosphorylation of AgrA 100

2.4.8. AgrC-HK phosphorylation of AgrA 100

2.5. *In silico* method 100

2.5.1. Molecular Dynamics (MD) Simulations (performed by Dr. Robert Weinzierl) 100

CHAPTER THREE 102

3. Development of an experimental system to study AgrA activity *in vivo*.....102

3.1. Introduction 102

3.1.1. The P1, P2 and P3 promoters 102

3.1.2. The *S. aureus* haemolysins.....104

3.1.2.1. α-haemolysin.....104

3.1.2.2. β-haemolysin.....104

3.1.2.3. γ-haemolysin 104

3.1.2.4. δ-haemolysin.....105

3.1.3. The L54a phage and pCL55 105

3.2. Objectives.....107

3.3. Polyclonal AgrA antibody production and purification 107

3.4. Characterisation of the SH1000^{-} *agr* operon-deficient strain 109

3.5. Development of a method to measure transcription activation at the P3 promoter by AgrA *in vivo* 111

3.6. Complementation of the SH1000$^{-}_{agr\ \text{IR P3-GFP}}$ *agrA*-deficient strain.....113

3.7. Conclusion.....117

CHAPTER FOUR 121

4. Construction, screening and characterisation of a library of single alanine mutants in the AgrA LytTR domain 121

4.1. Introduction 121

4.2. Objectives.....130

4.3. Selection of the amino acid residues to mutate in the AgrA LytTR domain.....130

4.4. Construction of a single alanine mutant library in the AgrA LytTR domain 131

4.5. Scanning of the alanine mutant library in the SH1000$^{-}_{agr\ \text{IR P3-GFP strain}}$ 131

4.6. *In vivo* characterisation of a selection of single alanine mutants in the AgrA LytTR domain.....134

4.7. Synonymous and non-synonymous mutations at selected residues in the AgrA LytTR domain 140

4.8. Conclusion 144

CHAPTER FIVE 147

5. Further characterisation of the four putative TAD mutants 147

5.1. Introduction 147

5.1.1. AgrA 147

5.1.2. Activation of AgrA 148

5.1.3. The *S. aureus* RNA polymerase 151

5.1.4. Sigma factor (σ) 151

5.1.5. DNA templates 151

5.2. Objectives 152

5.3. Tooling up for the *in vitro* analysis 152

5.3.1. Expression and purification of AgrA, $AgrA_{E181A}$, $AgrA_{L171A}$, $AgrA_{H200A}$, $AgrA_{Y229A}$ and $AgrA_{R233A}$ 153

5.3.1.1. Purification AgrA 153

5.3.1.2. Purification $AgrA_{L171A}$ 154

5.3.1.3. Purification $AgrA_{E181A}$ 154

5.3.1.4. Purification $AgrAH_{200A}$ 155

5.3.1.5. Purification $AgrA_{R233A}$ 156

5.3.2. Expression and purification of AgrC-HK 156

5.3.3. Expression and purification of the *S. aureus* RNA polymerase (RNAp) 157

5.4. Electrophoretic mobility shift assay (EMSA) 157

5.4.1. EMSA with acetyl phosphate activation 158

5.4.2. EMSA with AgrC-HK activation 159

5.5. *In vitro* transcription assay 161

5.5.1. Short RNA synthesis assay with AgrC-HK activation.....161

5.5.2. Full-length *in vitro* transcription assay with acetyl phosphate activation.....162

5.6. *In vivo* analysis: double mutants.....163

5.6.1. *In vitro* phosphorylation of AgrA.....164

5.6.2. Acetyl Phosphate phosphorylation of AgrA.....165

5.6.3. AgrC-HK phosphorylation of AgrA.....165

5.6.4. Position of L171A in the context of the predicted structure of the full-length AgrA.....166

5.7. *In silico* simulation of the effect of the Y229A mutation (performed by Dr. Robert Weinzierl).....167

5.8. DNA bending assay.....169

5.9. Conclusion.....171

CHAPTER SIX.....175

6. Final conclusions and overview.....175

6.1. Introduction.....175

6.2. Summary of findings.....177

6.3. Future directions.....187

APPENDICES.....190

7. Appendix I.....190

7.1. Vector maps.....190

7.1.1. pGEM-T Easy – used for cloning into pTYB2 and pCN34*itet*.....190

7.1.2. pCN34 – used to clone P2-RBS-AgrA-TT and P1-AgrA-TT.....190

7.1.3. pCN44 – used as PCR template for the TT.....190

7.1.4. pCN34*itet* – used to clone RBS-AgrA-TT.....191

7.1.5. pTYB2 – used to clone AgrA.....191

7.1.6. pET28b^{+} -used to clone AgrC-HK.....192

7.2. Protein ladders 192

8. Appendix II 193

8.1. Oligonucleotides and vectors constructed 193

8.1.1. Oligonucleotides for cloning and *in vitro* analysis 193

8.1.2. Vectors constructed 193

8.1.3. SDM pSN-P2-*agrA* primers 194

REFERENCES 197

PUBLICATION 221

Nicod SS, Weinzierl RO, Burchell L, Escalera-Maurer A, James EH, Wigneshweraraj S (2014) Systematic mutational analysis of the LytTR DNA binding domain of Staphylococcus aureus virulence gene transcription factor AgrA. *Nucleic acids research*

LIST OF FIGURES

Figure 1: The bacterial transcription cycle. 23

Figure 2: DNA sequence of -10 and -35 promoter elements. 26

Figure 3: DNA element contributing to the promoter recognition by the RNAp. 27

Figure 4: Composition of the RNAp in the three domains of life. 28

Figure 5: Crystal structure of *the Escherichia coli (E. coli)* RNAp. 29

Figure 6: Domain organisation of σ^{70} and its promoter recognition regions. 30

Figure 7: Activation by promoter remodelling. 32

Figure 8: Class I activation. 33

Figure 9: Class II activation. 33

Figure 10: Class III activation Class II-like. 34

Figure 11: Class III activation Class I-like. 35

Figure 12: Activation by pre-recruitment. 35

Figure 13: Schematic of a TCS. 37

Figure 14: Schematic of signal transduction in TCS. 38

Figure 15: Distribution of the different effector domains of response regulators. 40

Figure 16: Repartition of the different types of DNA-binding domains in response regulators (RRs) from *S. aureus* NCTC 8325. 47

Figure 17: Schematic representation of the organisation of the *agr* operon and the auto-catalytic circuit. 55

Figure 18: Conserved and variable regions in *agrB, D, C* and *A*. 56

Figure 19: Old and new model of the transmembrane topology of AgrB. 57

Figure 20: Model of the autoinducing peptide (AIP) synthetic pathway. 58

Figure 21: Structure of the AIP from the 4 different *agr* groups. 58

Figure 22: Topology of AgrC1. 60

Figure 23: Domain architecture of the AgrC dimer. 61

Figure 24: Crystal structure of the AgrC CA domain. 61

Figure 25: Crystal structure of the inactive CheY. 63

Figure 26: Crystal structure of the AgrA LytTR domain. 64

Figure 27: AgrA binding sites and footprint at the P2 and P3 promoters. 65

Figure 28: Conservation of the -10 and -35 promoter elements of the P2 and P3 promoters. 65
Figure 29: Model of oxidation sensing via AgrA. 67
Figure 30: Model of the interaction between the AgrC Ca-DHp dimer and a dimer of AgrA receiver domain. 68
Figure 31: Secondary structure of RNAIII. 69
Figure 32: AgrA binding sites and footprint, SarA and SarR binding sites at the P2 and P3 promoters 72
Figure 33: Schematic of the P1, P2 and P3 promoters of the *agr* operon. 102
Figure 34: The P1 and the P2 promoter. 103
Figure 35: pCL55 map. 106
Figure 36: Sequence of the *agr* IR P3-GFP of the pCL55 $_{agr\ \text{IR P3-GFP}}$ construct. 106
Figure 37: Nickel affinity chromatography of [6His]-AgrA. 107
Figure 38: Western blot of rabbits 369 and 370. 108
Figure 39: Western blot comparing the affinity of the serum and the polyclonal antibody against AgrA for AgrA. 108
Figure 40: Salt bridges between H174, E226 and E163 on the AgrA LytTR domain crystal structure. 109
Figure 41: Growth curve of SH1000 and SH1000^{-}. 110
Figure 42: Western blot of SH1000 and SH1000^{-}. 110
Figure 43: Blood agar plate of SH1000 and SH1000^{-}. 111
Figure 44: Schematic representation of the SH1000$^{-}_{agr\ \text{IR P3-GFP}}$ strain. 112
Figure 45: GFP growth curve of SH1000$_{agr\ \text{IR P3-GFP}}$ and SH1000$^{-}_{agr\ \text{IR P3-GFP}}$. 112
Figure 46: Bar chart of SH1000$_{agr\ \text{IR P3-GFP}}$ (blue) and SH1000$^{-}_{agr\ \text{IR P3-GFP}}$ (red). 113
Figure 47: Bar chart of the different SH1000$^{-}_{agr\ \text{IR P3-GFP}}$ complementations. 114
Figure 48: Bar chart of the different SH1000$^{-}_{agr\ \text{IR P3-GFP}}$ complementations compared to the pSN-P2-*agrA* complementation. 115
Figure 49: Blood agar plate complementation SH1000$^{-}_{agr\ \text{IR P3-GFP}}$. 116
Figure 50: Western blot of SH1000$^{-}_{agr\ \text{IR P3-GFP}}$ complementation, without HIgG. 117
Figure 51: Western blot of SH1000$^{-}_{agr\ \text{IR P3-GFP}}$ complementation, with HIgG. 117
Figure 52: Conservation of the -35 and -10 of the *agr* promoters. 119
Figure 53: Secondary structure of the AgrA LytTR domain with amino acid (aa) sequence. 121

Figure 54: Crystal structure of the AgrA LytTR domain with the three residues directly interacting with the DNA. 122
Figure 55: Crystal structure and amino acid (aa) sequence of the AgrA LytTR domain with all the residues making contact with the DNA. 123
Figure 56: Crystal Structure of the AgrA LytTR domain with the salt bridges interactions. 124
Figure 57: C-terminal *agrA* variants. 125
Figure 58: Savirin binding site. 127
Figure 59: Clustal alignment LytTR domain containing transcription factors (LDCTFs). 127
Figure 60: Secondary structure and amino acid sequence of the AgrA LytTR domain with FFRCHNSFVVN motif. 128
Figure 61: Clustal alignment of AgrA with VirR from *Clostridum perfringens*. 129
Figure 62: List of residues targeted by Site Directed Mutagenesis 130
Figure 63: SDM library construction. 131
Figure 64: Systematic mutational analysis of the AgrA LytTR domain. 132
Figure 65: Position of the mutants of category 1 on the structure. 133
Figure 66: Position of the mutants of category 2 on the structure. 133
Figure 67: Cartoon of the *agr* operon feedback loop. 135
Figure 68: GFP growth curve of pSN-*itet*-*agr.* 136
Figure 69: Comparison of the GFP fluorescence of the 24 putative TAD mutants in pSN-P2-*agrA* and in pSN-*itet*-*agrA*. 137
Figure 70: GFP activity vs AgrA expression for the 24 putative TAD mutants. 138
Figure 71: AgrA crystal structure with eleven residues not detectably expressed when mutated to alanine. 138
Figure 72: Multiple Sequence alignment of the amino acid (aa) sequence of the AgrA LytTR domain of 29 representative *Staphylococci* strains with the four TAD mutants. 139
Figure 73: Multiple Sequence alignment of the amino acid (aa) sequence of the LytTR domain of 43 representative LDCTFs with the four TAD mutants. 140
Figure 74: AgrA crystal structure with the four residues excluded. 141
Figure 75: Synonymous and non-synonymous mutations. 141
Figure 76: Multiple Sequence alignment the amino acid (aa) sequence of the LytTR domain of forty-three representative LDCTFs with the nine residues mutated highlighted. 143
Figure 77: Activity of the residues in the FFRCHNSFVVN motif. 145

Figure 78: Principle of the IMPACT^{TM} kit. 148
Figure 79: AgrC topology. 149
Figure 80: Phosphorylation of a RR by Acetyl phosphate. 150
Figure 81: DNA templates for *in vitro* analysis. 152
Figure 82: WT AgrA purification. 153
Figure 83: $AgrA_{L171A}$ purification. 154
Figure 84: $AgrA_{E181A}$ purification. 154
Figure 85: $AgrA_{H200A}$ purification. 155
Figure 86: $AgrA_{R233A}$ purification. 156
Figure 87: AgrC-HK purification. 156
Figure 88: RNAp purification. 157
Figure 89: Crystal structure of the AgrA LytTR domain with the 4 TAD mutants. 158
Figure 90: EMSA with acetyl phosphate activation. 158
Figure 91: EMSA with AgrC-HK activation. 159
Figure 92: Short RNA synthesis assay with AgrC-HK activation. 162
Figure 93: Full-length transcription assay with acetyl phosphate activation. 163
Figure 94: Substitutions and double mutants at positions 171 and 229. 163
Figure 95: Phosphorylation of AgrA with γ^{32}P-acetyl phosphate. 165
Figure 96: Phosphorylation of AgrA with AgrC-HK. 166
Figure 97: Phyre 2 predicted full-length structure of AgrA. 167
Figure 98: Molecular dynamics analysis of AgrA LytTR domain with the Y229A mutation. 168
Figure 99: Molecular dynamics analysis of AgrA LytTR domain with the Y229A or I219A mutations. 169
Figure 100: Schematic of the DNA bending vector pAM1847. 170
Figure 101: DNA bending assay AgrA and $AgrA_{Y229A}$. 170
Figure 102: AgrA binding sites and footprint at the P2 and P3 promoter. 174
Figure 103: Schematic representation of the organisation of the *agr* operon and the auto-catalytic circuit. 176
Figure 104: AgrA binding sites and footprint at the P2 and P3 promoters DNA sequence of the *agr* IR. 176
Figure 105: Salt bridges interactions in the AgrA LytTR domain with H174, E163 and E226 in red. 178

Figure 106: Crystal structure of the AgrA LytTR domain with H200 and the three residues directly interacting with the DNA. 180
Figure 107: Residues in the FhRh[RK][SNG]hhVN motif identified in the screen. 181
Figure 108: Crystal structure of the AgrA LytTR domain in complex with its DNA-binding sequence and the P2 promoter. 183
Figure 109: Model of transcription by AgrA. 184
Figure 110: Schematic of the principle of the BACTH system. 185
Figure 111: BACTH interaction between the AgrA-LytTR domain and σ4. 185
Figure 112: Vector map of the pGEM-T Easy vector from Promega. 190
Figure 113: Vector map of the pCN34 vector. Adapted from (Charpentier et al, 2004). 190
Figure 114: Vector map of the pCN44 vector. Adapted from (Charpentier et al, 2004). 190
Figure 115: Vector map of the pCN34itet. (Corrigan et al, 2011). 191
Figure 116: Vector map pTYB2 from New England Biolabs. 191
Figure 117: Vector map for pET28b^{+} from Novagen. 192
Figure 118: Protein Ladders. 192

LIST OF TABLES

Table 1: Examples of LDCTFs and the function they regulate. 41

Table 2: Distribution of LDCTFs among bacteria. 43

Table 3: The 16 TCS identified in NCTC 8325. 47

Table 4: List of processes regulated by more than one two component system (TCS). 53

Table 5: Transcriptional regulators involved in the regulation of the *agr* operon. 71

Table 6: Composition of growth media. 76

Table 7: Summary of the antibiotics used in this work. 77

Table 8: Bacterial strains used in this work. 79

Table 9: Bacterial plasmids used in this work. 80

Table 10: Composition of the SDS resolving and stacking gels used in this study. 91

Table 11: Salt bridges interactions. 124

Table 12: Examples of RRs successfully activated by acetyl phosphate. 150

Table 13: List of oligonucleotides used for cloning, EMSA and short RNA synthesis assay in this study. 193

Table 14: List of vectors constructed in this study. 193

Table 15: SDM primer Table 1. 194

Table 16: SDM primer Table 2. 195

Table 17: SDM primer Table 3. 196

CHAPTER ONE

1. Introduction

1.1. Transcription

The immense diversity of bacterial life is probably not the result of differences in DNA sequences of genes as those are mostly conserved among bacteria. Therefore these differences must stem from differences in the amount and the timing of the expression of these genes, their transcription (Haugen et al, 2008). Transcription is the first step of gene expression, the first opportunity for regulation and is mediated by the RNA polymerase (RNAp). The RNAp is a large multi-subunit enzyme responsible for gene transcription in all organisms from bacteria to eukaryotes. Transcription is the synthesis of an RNA copy complementary to the DNA template by the RNAp using nucleoside tri-phosphates (NTP(s)) and consists of three main steps: initiation, elongation and termination (See Figure 1) (Borukhov & Nudler, 2008). Regulation can take place at every step of the transcription cycle but transcription initiation is widely accepted as being the most frequently regulated step (Haugen et al, 2008).

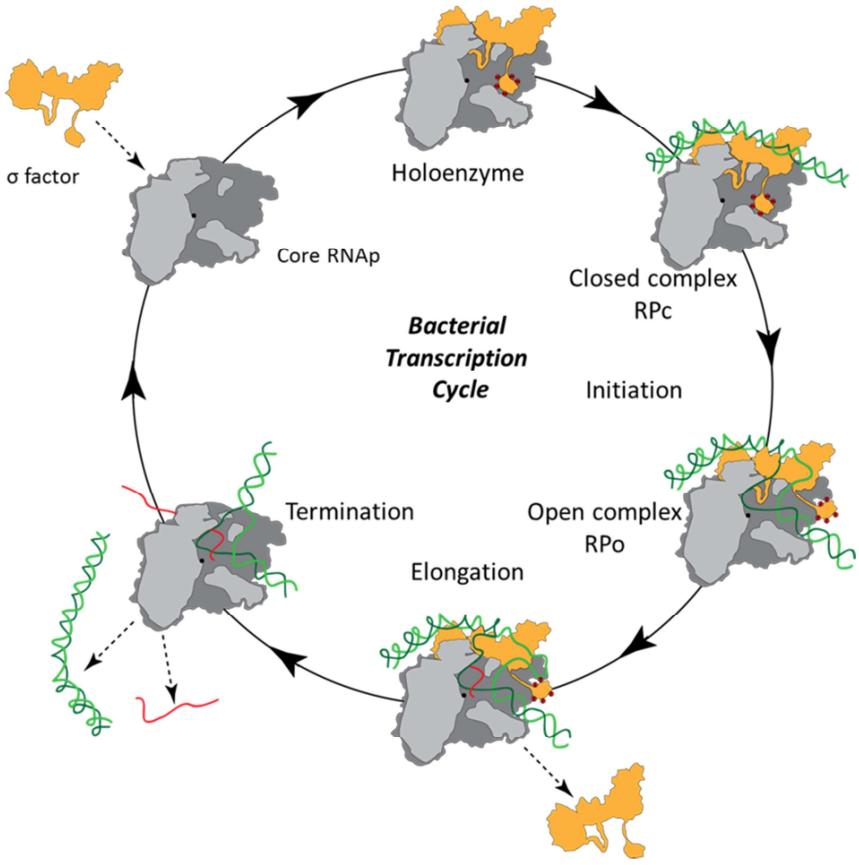


Figure 1: The bacterial transcription cycle.

The core RNAp is coloured in grey, the σ factor in orange, the DNA in green and the RNA in red. The transcription cycle starts with the association of the core RNAp with the σ factor to form the holoenzyme. Then the holoenzyme locates and binds to the promoter DNA to form the closed complex (RPc). The complex goes through a series of intermediate states until it reaches the open complex (RPo) conformation where the DNA is melted open. The initiation starts with the binding of the first NTP at the +1 site and thus the start of the synthesis of the complementary RNA. Then the holoenzyme escapes the promoter and the σ factor dissociates from the core RNAp, leading to the formation of the elongation complex. The RNA synthesis is continued until the transcriptional stop site is reached and transcription is stopped. Upon transcription termination, the newly synthesised RNA molecule and the DNA are released and the core RNAp is free to start a new transcription cycle. For more details, refer to the main text. Figure by Ellen James, adapted from (Haugen et al, 2008).

An entire thesis could be written about the transcription cycle in order to describe the molecular mechanisms taking place at each stage in detail. Here I will give a brief description of the events happening, focusing on transcription initiation as it is the step the most relevant to this thesis. The different stages are depicted in Figure 1.

Transcription initiation starts with the association of the core RNAp consisting of 2α subunits, the β subunit, the β' subunit and the ω subunit with the σ factor to create the holoenzyme. The σ factor is required for the recognition of the promoter DNA.

Then the holoenzyme locates the promoter DNA sequence and specifically binds to it, in a conformation referred to as closed complex (RPc) because the DNA is still double stranded, hence closed.

The complex then isomerises and goes through a series of intermediate states until it reaches the open complex (RPo) conformation where the DNA is melted open from positions -11 to +2 (See section 1.1.1) and is therefore accessible to serve as template for the synthesis of the RNA molecule.

During the isomerisation, a series of conformational changes take places in the DNA and the RNAp: the down-stream DNA moves into the main channel and unwinds, the template strand (T) is placed in the active site of the RNAp with the +1 in position for base pairing with the first NTP and σ region 1.1 is displaced from the main channel by the incoming DNA (See sections 1.1.3 and 1.1.4).

The initiation phase of transcription starts with the binding of the first NTP at the +1 site and continues with the binding of the second NTP and the formation of a phosphodiester bond between the two NTPs in the active site. The first transcripts synthesised are usually short and abortive and happen while the RNAp is still firmly anchored to the promoter sequence and pulls the downstream DNA into itself, a mechanism called scrunching that generates free energy. After a few rounds of abortive initiation, the free energy accumulated is enough to overcome the strength of promoter binding and the RNAp therefore escapes the promoter which leads to the dissociation of the σ factor from the core and the formation of the

elongation complex. Elongation consists of the repetitive formation of phosphodiester bonds as the RNA-DNA hybrid moves across the catalytic cleft in the RNAp due to the translocation of the RNAp along the DNA, one base pair (bp) forward at a time.

Upon termination, the synthesised RNA strand dissociates from the RNA-DNA hybrid and the RNAp in response to specific signals and the RNAp is free to start a new cycle of transcription. (Borukhov & Nudler, 2008; Haugen et al, 2008; Imashimizu et al, 2014; Saecker et al, 2011).

1.1.1. The promoter

The promoter is the DNA sequence located upstream of a gene responsible for the regulation of the expression of this gene. It is also the sequence that is recognised by the RNAp and brings the RNAp towards the gene to drive its expression. The promoter consists of several specific conserved motifs that interact with specific subunits of the RNAp, mainly the σ factor. Therefore, there are two main types of promoters, the σ^{70} promoters and the σ^{54} promoters that present a different organisation (See section 1.1.4). This section will focus on σ^{70} promoters (Haugen et al, 2008).

The first promoter elements to have been identified are the -10 (5'-TATAAT-3', from -12 to-7) and -35 (5'-TTGACA-3', from -35 to -30) (See Figure 2) (Maniatis et al, 1975; Pribnow, 1975). The names of these elements stem from their position on the DNA in respect to the transcriptional start site. The transcriptional start site corresponds to position +1 and the base immediately upstream to position -1 and so forth. The -10 and -35 promoter elements are separated by a spacer. The optimal size for that spacer is 17 bp and promoters having a smaller or bigger spacer have reduced activity (Aoyama et al, 1983; Siebenlist, 1979). The -35 and -10 promoter elements are conserved amongst bacterial promoters but some positions within these elements can vary without affecting the efficiency of the promoter too much whilst others cannot. At the -10 promoter element, the TAA motif from position -10 to -8 is variable whereas the T at position -12, the A at position -11 and the T at position -7 are very well conserved. Any mutation at the three later positions results in a dramatically reduced promoter activity. Positions -11 and -7 are the most highly conserved positions of the -10 element (See Figure 2) (Moyle et al, 1991; Siebenlist et al, 1980). Recently, the reason for this high level of conservation was identified. Upon recognition of the -10 promoter element by region 2 of the σ factor, the DNA is melted open and the A at position -11 and the T at position -7 are flipped out from the DNA base stack and buried into protein pockets on the σ factor. The A at position -11 is buried in a hydrophobic pocket which can perfectly

accommodate the shape of an A but not a different base. Similarly, the T at position -7 is buried in a hydrophilic pocket. Although this pocket is bigger, it can only accommodate a pyrimidine base (T or C) and not a purine base (A or G) (Feklistov & Darst, 2011; Zhang et al, 2012).

At the -35 promoter element, the only position that can vary without affecting the promoter activity too much and that is less conserved is the A at position -30. Any mutation at any of the other positions reduces the activity of the promoter (See Figure 2) (Moyle et al, 1991; Siebenlist et al, 1980).

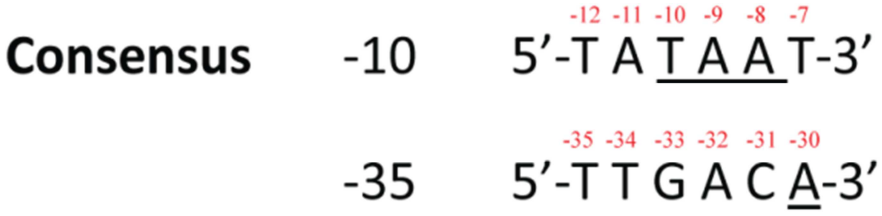


Figure 2: DNA sequence of -10 and -35 promoter elements.

-35 refers to the -35 element (5'-TTGACA-3') and -10 refers to the -10 element (5'-TATAAT-3'). The red numbers above the sequence correspond to the position of each base with regards to the transcriptional start site. The nucleotides that can vary from the consensus sequence without affecting promoter activity drastically and are thus less conserved are underlined in black.

However, some promoters have a very poorly conserved -35 promoter element but retain a high activity. This is due to another promoter element, called the extended -10. The extended -10 consists of two bases, a T at position -15 and a G at position -14. Not all promoters have an extended -10 promoter but those that do have a greater activity. The presence of an extended -10 at a promoter can compensate for a poorly conserved -35 promoter element and even abolish the need for it. In a similar fashion, the presence of an extended -10 promoter element can compensate for a suboptimal spacer length (Burr et al, 2000; Keilty & Rosenberg, 1987; Mitchell et al, 2003).

The last promoter element to be identified was the discriminator region. The optimal discriminator consists of three Gs at position -6, -5 and -4 and is required for binding of the RNAp to the promoter. It also helps the RNAp to recognise the promoter in the absence of the -35 promoter element (Feklistov et al, 2006; Haugen et al, 2006).

All of the promoter elements listed above are involved in the interaction with the σ factor (See Figure 3). However, the σ factor is not the only subunit of the RNAp involved in the interaction with the promoter DNA. The UP element consist of an extremely A+T rich region, alternating As and Ts, stretching from position -59 to -38 that interact with the C-

terminal domain of the α subunits of the RNAp. When present, the UP element leads to an increase in transcriptional activity (See Figure 3) (Gourse et al, 2000; Ross et al, 1993).

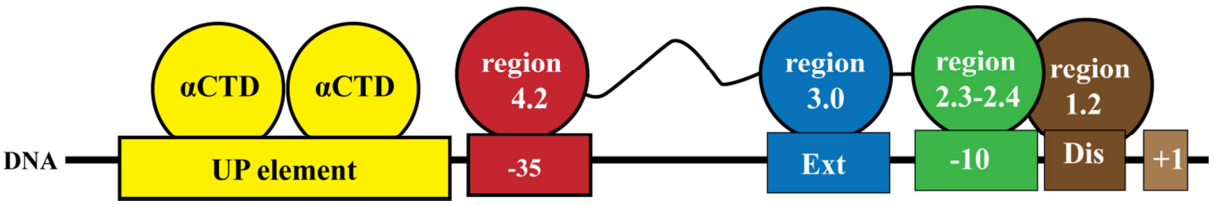


Figure 3: DNA element contributing to the promoter recognition by the RNAp.

The UP element are the binding site for α-CTD (α-C-terminal domain) at the promoter and consist of a very AT rich stretch of DNA from position -59 to -38. -35 refers to the -35 element (5'-TTGACA-3') recognised by the σ factor. Ext refers to the extended -10 element, T at -15 and G at -14 recognised by the σ factor. -10 refers to the -10 element (5'-TATAAT-3') recognised by the σ factor. Dis refers to the discriminator region, three Gs from -6 to -4 recognised by the σ factor. +1 refers to the transcriptional start site.

1.1.2. RNAp

RNA polymerases are separated into two main families: the single subunit and the multi-subunit RNAp. Single subunit RNAp are utilised by bacteriophages, mitochondria and chloroplasts. The most studied single subunit RNAp is probably the T7 RNAp. Although the structure of single subunit RNAp is very different from the multi-subunit ones and resembles that of a DNA polymerase, their biochemical properties are similar. Therefore, the small size of single subunit RNAp (~100 kDa) has made them attractive for biochemical and structural studies (Cheetham & Steitz, 2000; Cramer, 2002a).

Multi-subunit RNAp share a conserved core of 5 subunits and many regions that diverge in term of sequence but are structurally conserved (See Figure 4). The bacterial RNAp consists of the 5 core subunits: β' (~150 kDa), β (~150 kDa), two α (40 kDa) and ω (10 kDa) for a total mass of about 400 kDa (See Figure 5). Bacteria also possess a σ factor and some have an additional subunit, the δ subunit (~20 kDa) that is found in Gram positive bacteria that have a low-G+C content. Eukaryotes possess three RNAp, eRNAp I, II and III that contain 10 common subunits and 4, 2 and 5 additional subunits, respectively. Archaea contain only one RNAp which is composed of 12 subunits, 11 of which are similar to the subunits of eRNAp II (Cramer, 2002b; Lane & Darst, 2010).

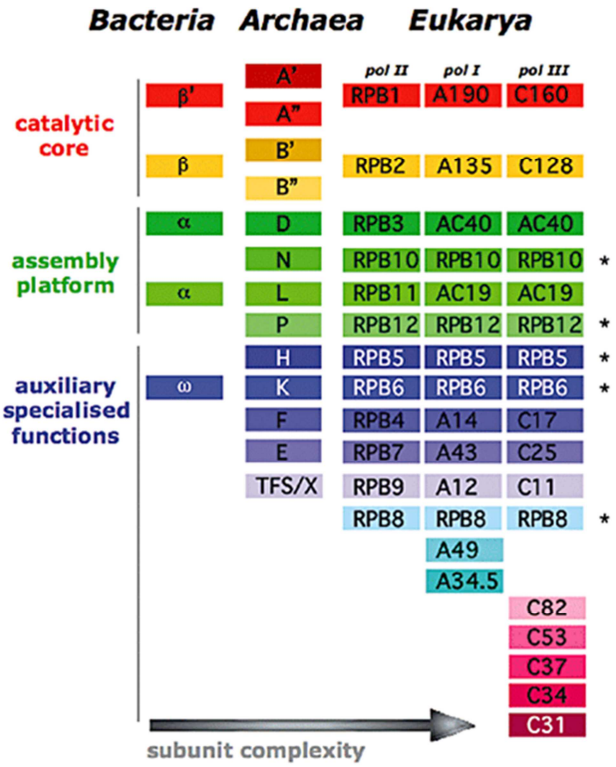


Figure 4: Composition of the RNAp in the three domains of life.

The subunits are organised according to their function and homologous subunits are colour coded. The direction of the arrow indicates the increase in complexity. The asterisks mark the subunits that are conserved in all three eRNAP (Werner, 2007).

1.1.3. The bacterial RNAp

The active site of the RNAp is formed by the β and β' subunits and drives the ability to bind to the template DNA and the RNA product during transcription (Korzheva et al, 2000). The α subunits consist of two independent domains connected by a flexible linker. The N-terminal part of α (αNTD) dimerises and is involved in the assembly of the β and β' subunits (Ebright & Busby, 1995). The C-terminal part (αCTD) has a DNA-binding ability and is important for promoter recognition by recognizing the UP element (See section 1.1.1) (Blatter et al, 1994; Ross et al, 1993). The αCTD interacts with region 4 of the σ factor in addition (Chen et al, 2003; Ross et al, 2003). The ω subunit is a chaperone and helps the folding of the β' subunit (Minakhin et al, 2001) (See Figure 5). The δ subunit is involved in promoter recognition in *Staphylococcus aureus* (*S. aureus*) and drives the RNAp towards strongly expressed promoters, mainly the key virulence genes (Weiss et al, 2014).

The RNAp binding to the promoter region involves firstly the binding of the αCTD to the UP element and the binding of σ to the -35 promoter element, the extended -10 promoter element and the discriminator region (Barne et al, 1997; Feklistov et al, 2006; Gardella et al, 1989; Haugen et al, 2006; Ross et al, 1993). The interaction between the -10 promoter element and

the σ factor takes place at a later stage, upon strand separation. Moreover, the fact that the bases at positions -11 and -7 are inserted into pockets on the σ factor (See section 1.1.1) enables to use binding energy to drive DNA unwinding and accounts for the ability of the σ factor to facilitate promoter unwinding (Feklistov & Darst, 2011; Zhang et al, 2012).

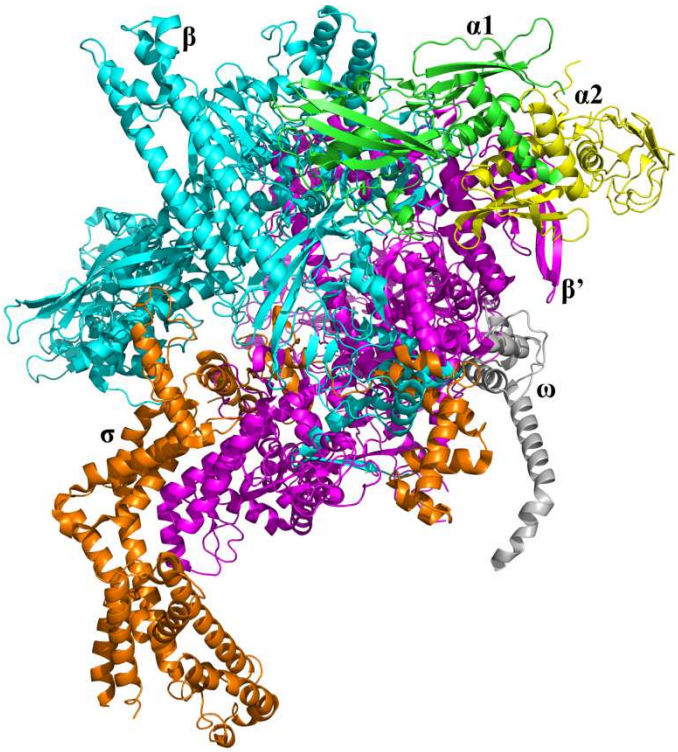


Figure 5: Crystal structure of *the Escherichia coli (E. coli)* RNAp.

Ribbon representation of the *E. coli* RNAp. The α1 subunit is coloured in green, the α2 subunit is coloured in yellow, the β subunit is coloured in cyan, the β' subunit is coloured in magenta, the ω subunit is coloured in grey and the σ factor in orange. All the subunits are labelled in black. Pymol, PDB 4LK0.

1.1.4. The sigma factors (σ)

σ factors are responsible for the ability of the holoenzyme to locate and bind to promoter sequences and are thus an essential component of the transcription machinery.

The σ factors are separated into two classes: the σ^{70} class containing the majority of σ factors that are mostly structurally related to the housekeeping σ^{70} factor from *E. coli* and the σ^{54} class. σ^{54} is the only member of this class and has a different structure. σ^{54} is involved in the regulation of genes involved in various functions (nitrogen metabolism, carbon metabolism, phage and heat shock, etc…) and recognises different promoter elements than the other family. σ^{54} recognises promoters that have conserved -24 and -12 promoter elements. These promoters are unable to transition from RPc to RPo without the help of an activator ATPase that hydrolyses ATP (Buck et al, 2000; Studholme & Buck, 2000; Wigneshweraraj et al, 2008).

The σ^{70} family can be divided into 4 groups, based on phylogenetic similarities: the essential primary σ factors (σ^{70}), the closely related σ factors that are not essential for growth, the σ factors responding to specific signals and the extracytoplasmic function (ECF) σ factors that are highly divergent from σ^{70} and represent the largest group (Gruber & Gross, 2003).

The σ^{70} family recognise the conserved -35 and -10 promoter elements, the extended -10 and the discriminator region. The σ^{70} factor consists of 4 domains: σ1, σ2, σ3 and σ4 that are subdivided into smaller sub-domains. Region 1.2 is responsible for binding to the discriminator region, regions 2.3 and 2.4 for binding to the -10 promoter element, region 3.0 for binding to the extended -10 and region 4.2 for binding to the -35 promoter element (See Figure 6 and section 1.1.1) (Barne et al, 1997; Feklistov et al, 2006; Gardella et al, 1989; Gruber & Gross, 2003; Haugen et al, 2006; Siegele et al, 1989).

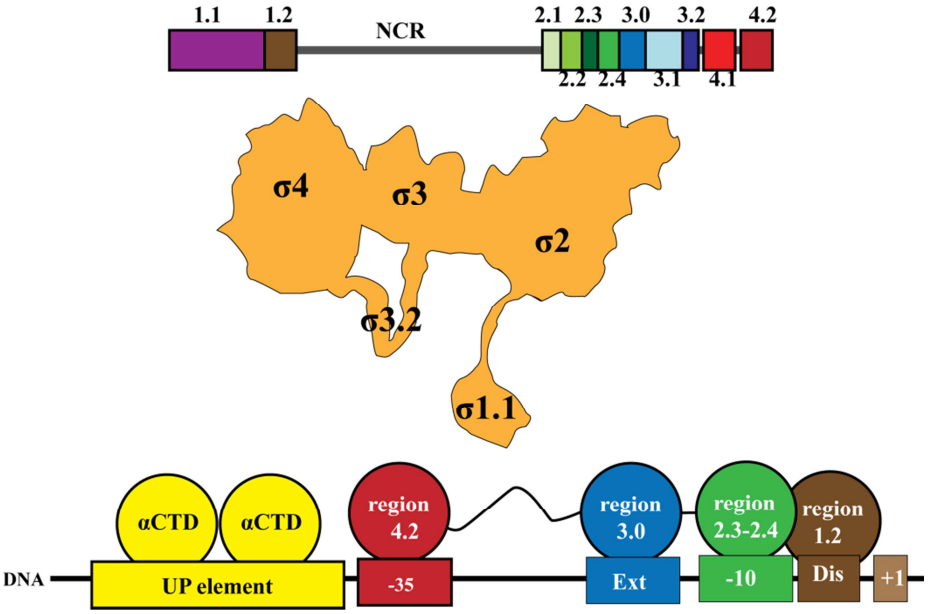


Figure 6: Domain organisation of σ^{70} and its promoter recognition regions.

The top figure represents the domain organisation of σ^{70}. NCR means non-coding region. The subdomains are colour coded. The middle figure is a cartoon representation of σ^{70}. The bottom figure is a cartoon representation the promoter elements and the σ^{70} subdomains that recognise them. The σ^{70} subdomains are colour coded. The UP element are the binding site for α-CTD (α-C-terminal domain) at the promoter and consist of a very AT rich stretch of DNA from position -59 to -38. -35 refers to the -35 element (5'-TTGACA-3') recognised by the σ factor. Ext refers to the extended -10 element, T at -15 and G at -14 recognised by the σ factor. -10 refers to the -10 element (5'-TATAAT-3') recognised by the σ factor. Dis refers to the discriminator region, three Gs from -6 to -4 recognised by the σ factor. +1 refers to the transcriptional start site. Adapted from (Haugen et al, 2008).

S. aureus contains 4 different sigma factors: the equivalent of σ^{70}, called σ^{A} regulating housekeeping genes that is essential (Deora & Misra, 1996), σ^{B} involved in the environmental stress response (Chan et al, 1998), σ^{S} an apparent ECF σ factor involved in stress responses (starvation, growth under elevated temperature and lysis) (Shaw et al, 2008)

and σ^H that is only activated under certain conditions and enables *S. aureus* to become naturally competent for transformation by DNA (Morikawa et al, 2012).

1.2. Transcription activation at σ^{70} promoters

Transcription activation is an additional step that allows a tighter regulation. Some promoters are not optimal by design. They either have weak conservation of their characteristics -10 and -35 promoter elements or a spacer length longer or shorter than the optimal 17 nt. Transcription from such promoters will be weak unless other factors come into place to help them achieve a higher level of transcription. These factors will mainly act to help the recruitment of the RNAp, stabilise the RPc or help the transition towards the RPo. They are called transcription factors (TFs) and act at to regulate transcription either by activating it or repressing it, in response to an environmental signal. They act either at the RNAp in which case they do not necessarily have a DNA-binding ability, or are DNA-binding TFs and bind to the promoter region of the gene they regulate. Some of the DNA-binding TFs are response regulators (RRs) (See section 1.3.2), like the *S. aureus* AgrA that regulates transcription in response to quorum sensing (See section 1.5.5.3). Transcription activation has been extensively characterised in *E. coli* and backed-up by crystal structures of promoter DNA/ TF complexes (Lee et al, 2012). This led to the definition of major mechanisms of activation that will be listed here.

1.2.1. Activation by promoter remodelling

DNA-binding TFs can bind to the promoter DNA and alter its conformation to render it more attractive for the RNAp and thus increase the transcription efficiency. They can bind between the -10 and -35 promoter elements and rectify a non-optimal spacer by twisting the DNA. A non-optimal spacer length results in the σ factor region 2.4 being misaligned in respected to the -10 promoter element after binding to the -35 and UP promoter elements and thus defective transcription (See section 1.1.1). The bend induced by the TF will act to align σ region 2.4 with the -10 promoter element and thereby activate transcription (See Figure 7). Examples of such TFs include the MerR family of transcriptional regulators (*E. coli* MerR, *E. coli* SoxR, *Streptomyces* TipA, *Bacillus subtilis* BltR, *Bacillus subtilis* BmrR) and the cII protein from the bacteriophage λ (Brown et al, 2003; Heldwein & Brennan, 2001; Jain et al, 2005; Lee et al, 2012).

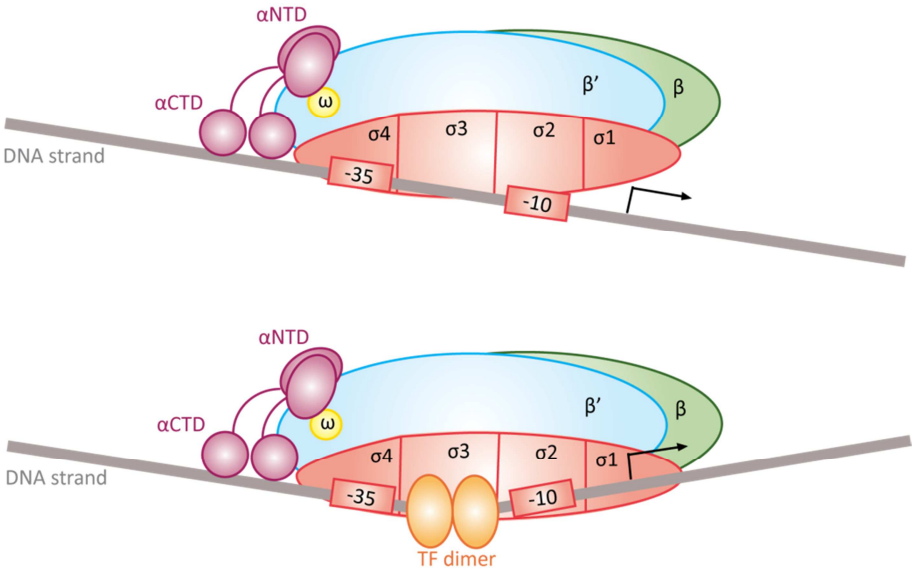


Figure 7: Activation by promoter remodelling.

The TF dimer binds between the -10 and -35 promoter elements and induces a conformational change that allows the RNAp to interact with the promoter better. The RNAp subunits are colour coded and labelled in black.-35 refers to the -35 element (5'-TTGACA-3') and -10 to the -10 element (5'-TATAAT-3'). αCTD refers to the α-C-terminal domain and αNTD to the α-N-terminal domain. Adapted from (Lee et al, 2012).

1.2.2. Activation by contacting the RNAp

In cases where the -10 or -35 promoter elements have a weak conservation compared to the consensus sequence, TFs can bind to the promoter DNA upstream or overlapping the -35 promoter element and directly contact the RNAp when it binds to the promoter region. In such cases the interaction between the TF and the RNAp will strengthen the defective interaction between the RNAp and the promoter. Thus the final RPo will be similar to that of an independent promoter. It appears that such TFs act either by helping recruiting the RNAp to the promoter site or by stabilising the contacts between the RNAp and the promoter DNA to allow optimal transcription. In this case, TFs can interact with different subunits of the RNAp depending on where the TF binds in regard to the -35 promoter element. TFs that bind far upstream of the -35 promoter element will be able to interact with the α-CTD subunit. Because the α-CTD is connected to the α-NTD by a flexible linker, it has a considerable freedom of movement. Therefore, even though the α-CTD usually interacts with TFs with binding site located at -61.5, it has been shown to interact with TFs located at -72.5, -82.5 or even -93.5. TFs that have a binding site overlapping the -35 promoter element will be able to interact with region 4 of the σ factor (not all of them do), the α-CTD and the α-NTD. These types of activation have been separated into classes, according to the number of contacting regions between the TF and the RNAp (Lee et al, 2012).

1.2.2.1. Class I activation

Class I is used to describe the promoters were the TF binding site or sites, in the case of a dimeric TF, are located upstream of the -35 promoter element. In this case, the TFs will be able to interact each with the α-CTD. One α-CTD will bind upstream of the TF biding site and contact the TF further upstream and the other α-CTD will maintain its interaction with region 4 of the σ factor and interact with the downstream TF on the other face (See Figure 8). Examples of such TFs in *E. coli* include the catabolite activator protein (CAP) and the fumarate and nitrate reductase regulatory protein (FNR) (Lawson et al, 2004; Lee et al, 2012; Wing et al, 1995).

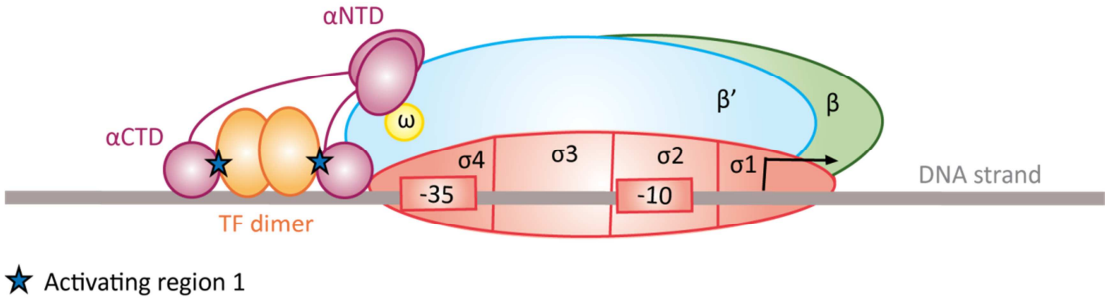


Figure 8: Class I activation.

The TF dimer binds upstream of the RNAp and contacts the two α-C-terminal domains (αCTD).The RNAp subunits are colour coded and labelled in black.-35 refers to the -35 element (5'-TTGACA-3') and -10 to the -10 element (5'-TATAAT-3'). αNTD refers to the α-N-terminal domain. Adapted from (Lee et al, 2012).

1.2.2.2. Class II activation

Class II refers to promoters where the TF binding site overlaps the -35 promoter element. In this case, the localisation of the TF enables it to make contacts with three different subunits of the RNAp. Because the TF binds to the upstream part of -35 promoter element the interaction between region 4 of the σ factor and α-CTD is prevented and the α-CTD subunits will bind further upstream. The TF will be able to contact region 4 of the σ factor, the α-CTD subunit and the α-NTD subunit (See Figure 9). Examples of such TFs in the *E. coli* include the CAP and FNR TFs (Belyaeva et al, 1996; Busby & Ebright, 1997; Lee et al, 2012; Niu et al, 1996; Wing et al, 2000).

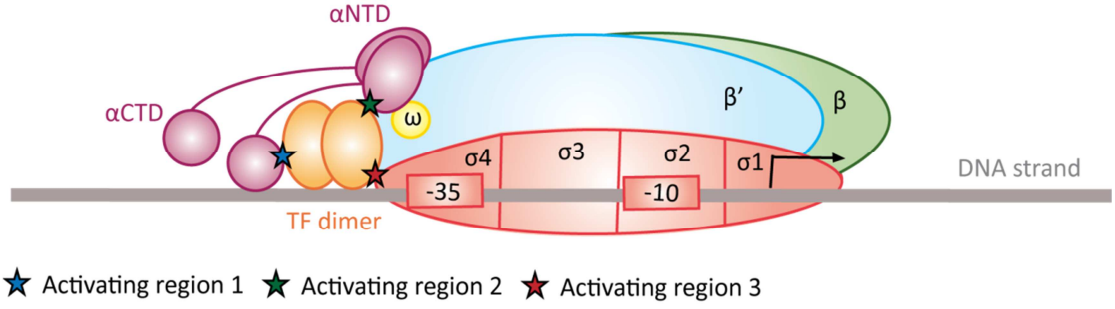


Figure 9: Class II activation.

The TF dimer binds close to the RNAp and can interact with three subunits of the RNAP: one α-C-terminal domain (αCTD), the α-N-terminal domain (αNTD) and region 4 of the σ factor. The RNAp subunits are colour coded and labelled in black.-35 refers to the -35 element (5'-TTGACA-3') and -10 to the -10 element (5'-TATAAT-3'). Adapted from (Lee et al, 2012).

1.2.2.3. Class III activation

Class III is used to describe the cases where two TF dimers make contact with the RNAp. In this situation, if one of the TF binding site overlaps the -35 promoter element and the other binding site is further upstream, the situation will be similar to a Class II activation but with additional contacts. Because the first TF dimer binds to the upstream part of -35 promoter element the interaction between region 4 of the σ factor and α-CTP is prevented and the α-CTP subunits will bind further upstream. The downstream TF dimer will be able to contact region 4 of the σ factor, the α-CTP subunit and the α-NTP subunit and the upstream TF dimer will contact the other the α-CTP subunit (See Figure 10). The CAP protein from *E. coli* is an example of such TFs (Belyaeva et al, 1998; Joung et al, 1993; Lee et al, 2012; Murakami et al, 1997).

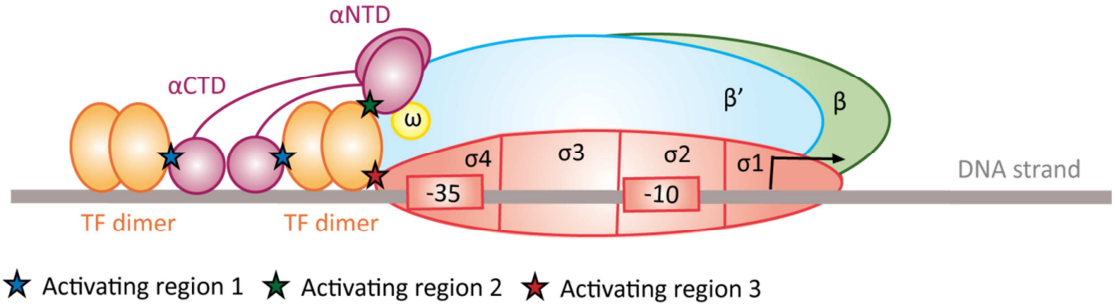


Figure 10: Class III activation Class II-like.

One TF dimer binds close to the RNAp and can interact with three subunits of the RNAP: one α-C-terminal domain (αCTD), the α-N-terminal domain (αNTD) and region 4 of the σ factor. The other TF dimer binds upstream of the αCTD and can interact with the most upstream αCTD. The RNAp subunits are colour coded and labelled in black.-35 refers to the -35 element (5'-TTGACA-3') and -10 to the -10 element (5'-TATAAT-3'). Adapted from (Lee et al, 2012).

If both TFs binding sites are upstream of the -35 promoter element the situation will be similar to a Class I activation. The TF dimers will be able to interact each with an α-CTD. One α-CTD will bind in between both TFs biding site and contact the TF dimer further upstream and the other α-CTD will maintain its interaction with region 4 of the σ factor and interact with the downstream TF dimer on the other face (See Figure 11). The CAP protein from *E. coli* is an example of such TFs (Beatty et al, 2003; Joung et al, 1993; Lee et al, 2012; Tebbutt et al, 2002).

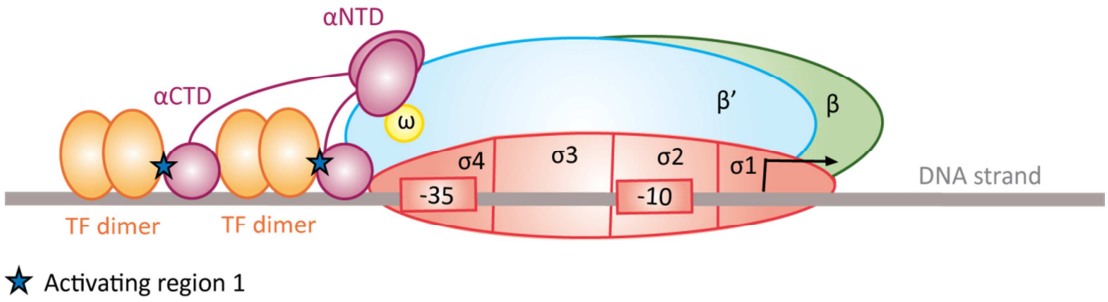


Figure 11: Class III activation Class I-like.

One TF dimer binds upstream of the RNAp and can interact with the most upstream α-C-terminal domain (αCTD). The other TF dimer binds in-between both αCTD and interacts with the downstream αCTD. The α-N-terminal domain (αNTD). The RNAp subunits are colour coded and labelled in black.-35 refers to the -35 element (5'-TTGACA-3') and -10 to the -10 element (5'-TATAAT-3'). αNTD refers to the α-N-terminal domain. Adapted from (Lee et al, 2012).

1.2.2.4. Activation by pre-recruitment

Some TFs contact the RNAp before it has reached the promoter and act to change its promoter preferences in order to drive it to the promoter they activate. This is called pre-recruitment and modifies the competition between promoters. The TF binds to the RNAp and then guides it to the promoter it regulates by binding to its binding site upstream of the -35 promoter element (See Figure 12). Examples of such TFs in *E. coli* include MarA and SoxS. The binding sites for MarA and SoxS are far more frequent than the proteins themselves. Therefore the action of binding to the RNAp first extends the recognition sequence and directs MarA and SoxS towards the binding sites located upstream of a promoter (Griffith et al, 2002; Griffith & Wolf, 2004; Lee et al, 2012; Martin et al, 2002).

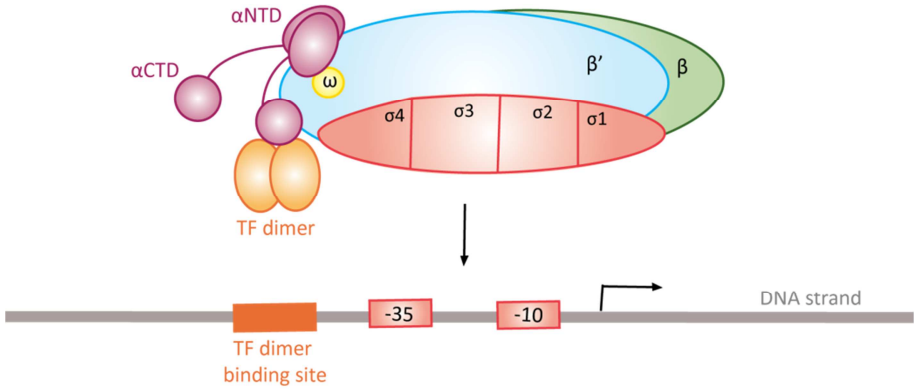


Figure 12: Activation by pre-recruitment.

The TF dimer binds to the RNAp before it is bound to the promoter DNA and directs it towards a specific promoter. The RNAp subunits are colour coded and labelled in black.-35 refers to the -35 element (5'-TTGACA-3') and -10 to the -10 element (5'-TATAAT-3'). αCTD refers to the α-C-terminal domain and αNTD to the α-N-terminal domain. Adapted from (Lee et al, 2012).

1.3. Two-component systems (TCS)

Two-component systems (TCS) are often the mediators of the ability of bacteria to adapt efficiently to changes in their environment and are involved in most aspects of bacterial regulation. TCS act at all the levels of adaptation used by bacteria: at the level of individual genes and proteins, at the level of global regulons, at the whole-cell level and at the population level. TCS could therefore be considered as essential prerequisites for the pathogenicity of bacteria. TCS have also been identified in eukaryotic organisms (plant, fungi, yeast, protozoa) but are much less wide-spread in these organisms than they are in bacteria. They are not present in animals, which makes them good targets for antibacterial development. The number of TCS identified keeps increasing as more and more new bacterial genomes are being sequenced. There seems to be a correlation between the number of TCS present in a genome and the size of that said genome. Bacteria that live in a very versatile environment and encounter a variety of different conditions will typically encode more TCS than bacteria living in a uniform habitat. Thus, bacteria that have adapted to a specific niche such as obligate intracellular bacteria or extracellular pathogens will have a very small number of TCS with the extreme examples of Mycoplasma spp. that have none and *Myxococcus xantus* that contains over a hundred (Beier & Gross, 2006; Gao et al, 2007; Kiil et al, 2005). However, most bacteria encode over a dozen of TCS that are involved in the regulation of a wide variety of mechanisms. These mechanisms range from the fundamental processes that are metabolism and motility to much more complex behaviours such as virulence and development (Galperin, 2010; Gao et al, 2007; Gao & Stock, 2009).

TCS allow the coupling of environmental stimuli with an adaptive response. The signal is transduced via a phosphotransfer between the sensor histidine kinase (HK) that is usually membrane bound and a cytosolic RR (See Figure 13). Typically, the HK αrylates at a conserved histidine residue in response to an external stimulus, taking a γ-phosphoryl group from ATP. This autophosphorylation is thus ATP-dependent. This high-energy phosphoryl group is then transferred to a conserved exposed aspartate residue in the RR. The phosphorylation of the RR induces conformational changes which transition the RR from the inactive state to the active state and affect the affinity of the effector domain for its target. The activated RR is then able to elicit the cellular response. The lifetime of the RR phosphorylation is controlled via dephosphorylation either by the kinase or by itself. The autophosphorylation of the HK, the phosphotransfer to the RR and the dephosphorylation of the RR all require Mg^{2+} cations and can be regulated by environmental signals. The level of

activation of the output response depends on the level of phosphorylation of the RR and not its absolute expression level. Therefore the system is limited by the HK to RR ratio. Additional proteins can act to regulate the HK or the RR and many TCS consist of more than two proteins. These additional proteins add to the complexity of the network and provide alternative targets for a more precise regulation (Casino et al, 2010; Gao & Stock, 2009; Perry et al, 2011; Stock et al, 2000).

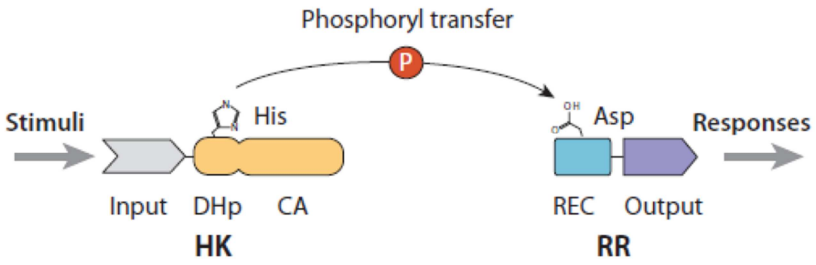


Figure 13: Schematic of a TCS.

Activation of a response regulator (RR) by phosphotransfer from a histidine kinase (HK). The input domain of the HK detects the stimuli and autophosphorylates at a conserved histidine (His) residue in the dimerisation and phosphotransfer (DHp) domain. The phosphate is then transferred to a conserved aspartate residue (Asp) in the receiver domain (REC) of the RR thereby activating it. Once the RR is activated, its output domain is ready to elicit its specific responses. Adapted from (Gao & Stock, 2009).

1.3.1. Histidine Kinases (HKs)

HKs are modular proteins consisting of a variable input (sensing) domain (varying in function of the signal that the HK is responsible for sensing) attached to a conserved catalytic core by a linker domain. HKs are able to sense a wide array of stimuli such as small molecules, light, pressure, cell-envelope stress, redox potential and electrochemical gradients. These signals are either perceived by the HK directly through its sensory domain or by auxiliary signal transduction proteins. There are three main kinds of sensory domains: extra-cytoplasmic domains allowing the perception of extracellular stimuli, membrane-spanning domains in order to sense membrane-associated stimuli and cytoplasmic sensory domains responsible for sensing internal stimuli.

The cytoplasmic catalytic core consists of two domains: a well-conserved C-terminal catalytic and ATP binding (CA) domain carrying the catalytic activity of transferring phosphoryl group from ATP to the conserved histidine residue and a less-conserved dimerisation and histidine phosphotransfer (DHp) domain containing that conserved histidine residue (See Figure 14). In addition to their kinase activity, many HKs also exhibit a phosphatase activity towards their cognate RR. This phosphatase activity enables a rapid

silencing of the signalling pathway and is under the control of the DHp domain. It is also thought to act in order to supress non-specific phosphorylation. The DHp domain is responsible for the dimerisation. This dimerisation is necessary for the HK to be functional. It is thought that the detection of the stimulus changes the interactions within the HK dimer and that this perturbation is relayed to the core domain. The linker domain of the HK is thought to be essential for this signal transmission. Upon detection of the signal, the interface between the DHp and CA domain is modified allowing the CA domain to move, either to align the ATP molecule with the conserved histidine to allow autophosphorylation of the kinase or to present the aspartate of the RR for phosphorylation or dephosphorylation of the latter (Gao & Stock, 2009; Perry et al, 2011). The CA domain of the HK can host one ATP molecule. In the case of autophosphorylation, the conserved histidine in the DHp domain attacks the γ-phosphate of the ATP. This process can happen in *trans* where the histidine from one monomer attacks the γ-phosphate of the ATP hosted by the CA domain from the other monomer or in *cis* with the DHp and CA domain from the same monomer. Some HKs autophosphorylate in *trans* while others do so in *cis* (Casino et al, 2010).

The abundance of paralogous TCS in the bacterial cell raises the question of cross-phosphorylation between non-cognate HKs and RRs. In most cases, there is a strong kinetic preference for cognate HK-RR phosphorylation. However, in some cases, one HK can activate multiple RRs (one-to-many) or multiple HKs can activate the same RR (many-to-one) (Casino et al, 2010; Gao & Stock, 2009; Perry et al, 2011).

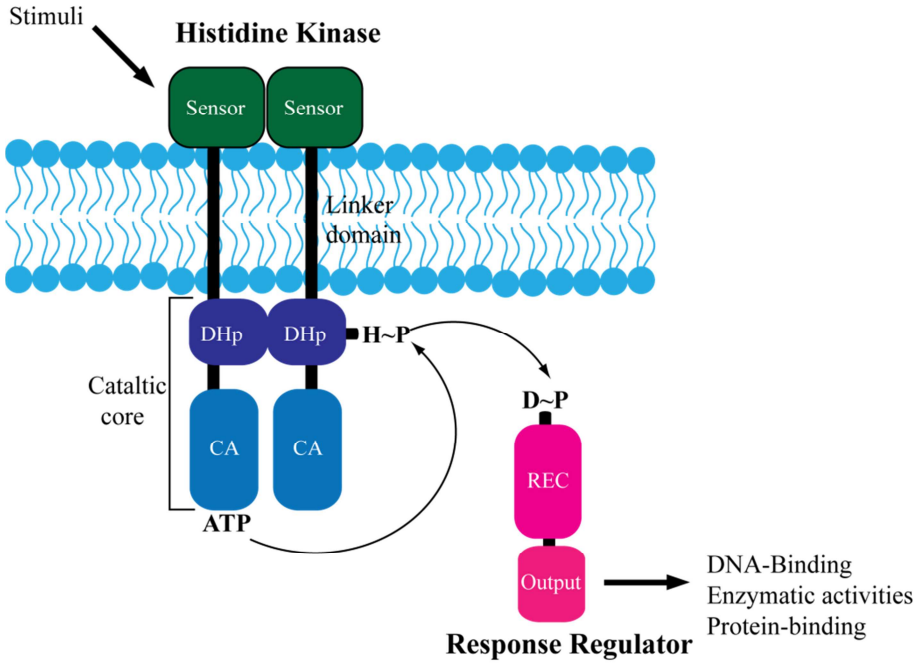


Figure 14: Schematic of signal transduction in TCS.

The sensor domain of the Histidine kinase detects the stimuli and autophosphorylates at a conserved histidine (H) residue in the dimerisation and phosphotransfer (DHp) domain by attacking the ATP hosted in the catalytic and ATP binding (CA) domain. The phosphate is then transferred to a conserved aspartate residue (D) in the receiver domain (REC) of the response regulator thereby activating it. Once the RR is activated, its output domain is ready to elicit its specific response (DNA-binding, Enzymatic activities or Protein-binding). Adapted from (Perry et al, 2011).

1.3.2. Response Regulators (RRs)

Most RRs contain two domains: an N-terminal conserved regulatory or receiver (REC) domain linked to a variable effector domain (See Figure 14). However, about 17% of RRs consist of a stand-alone REC domain. The REC domain is involved in the catalysis of phosphoryl transfer from the HK and regulates the activity of the effector domain in a phosphorylation-dependent manner (Gao & Stock, 2009; Perry et al, 2011). The REC domain of the RR binds to its cognate HK by clinging to a helix in the DHp domain, under the conserved histidine. This arrangement places the phosphor-acceptor aspartate residue of the RR close to the phosphorylated histidine in the HK. Thus the partner specificity is conferred by the cytosolic part of the DHp domain of the HK containing the interacting helix and the part of the REC domain of the RR that interacts (Casino et al, 2010).

The REC domain exists in equilibrium between the active and inactive form. Un-phosphorylated, the REC domain mainly exists in the inactive state and phosphorylation shifts the equilibrium towards the active state. Phosphorylation takes place at a conserved aspartate residue and generates a high-energy acyl phosphate. This acyl phosphate provides the energy that drives a conformational change in the REC domain. The REC domain catalyses phosphoryl transfer and autodephosphorylation. Therefore, the phosphotransfer from the HK to the RR is the result of the contribution of both proteins. The catalytic activity of the RR explains the ability of most RRs to autophosphorylate using high-energy small phosphodonors such as acetyl phosphate. However, these types of phosphotransfer have been shown to be about 100 fold slower than with the cognate HK. The level of phosphorylation of the RR is determined by the rates of phosphorylation and dephosphorylation which are influenced by several factors: acyl phosphates are unstable, RRs catalyse autodephosphorylation and in many cases the HK phosphatase activity dephosphorylates the RR. RRs commonly dimerise or oligomerise upon phosphorylation.

REC domains usually have on average 26% sequence identity and contain a few highly conserved residues. These residues consists of two aspartate or glutamate residues in the active site, the phosphor-acceptor aspartate which is responsible for coordinating an essential Mg^{2+} and a conserved lysine that forms a salt bridge with the phosphate in the active state.

The binding of the Mg^{2+} to the RR drives conformational changes that are then stabilised once the RR is phosphorylated.

The effector domain is responsible for the output response. The wide variety of effector domains is responsible for the diversity of processes that are regulated by TCS. The effector domains of RRs are mainly DNA-binding, RNA-binding, enzymatic, or protein binding. However, the majority of RRs (63%) contain a DNA-binding effector domain. DNA-binding RRs act as TFs and modulate the transcription of their target genes by binding to their promoter region. They can function as activators or as repressors; regulating a single operon or 100 genes throughout the genome. Most DNA-binding RRs positively regulate their own expression and the expression of their cognate HK. The DNA-binding domain of RRs usually consists of some version of the classical helix-turn helix domain. These consist of winged-helix domain (30% of all RRs, the OmpR/PhoB family), a four-helix helix-turn-helix (H-T-H) domain (17% of all RRs, NarL/FixJ family) and an AAA+ATPase domain fused to a factor of inversion (Fis)-type H-T-H domain (10% of all RRs, NtrC/DctD family). But in 3% of all RRs, the DNA-binding domain consists of an unusual domain, the LytTR domain which consists mainly of β-sheets (See Figure 15) (Gao & Stock, 2009; Perry et al, 2011).

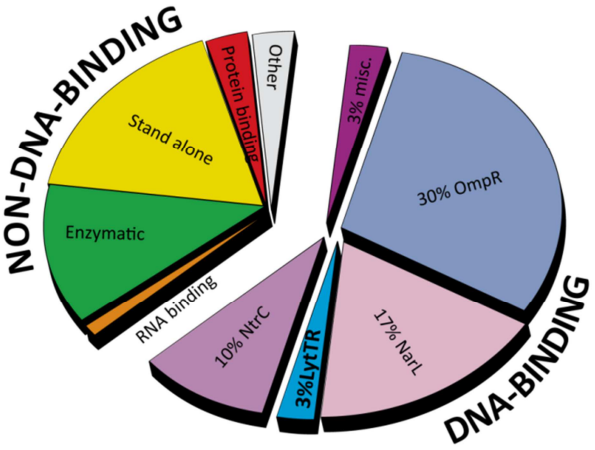


Figure 15: Distribution of the different effector domains of response regulators.

Pie chart with the different effector domains of response regulators (RRs). The effector domains are separated into two groups: the DNA-binding ones and the non-DNA-binding ones. The percentage of all RRs having a specific domain is indicated for the DNA-binding ones. Adapted from (Gao & Stock, 2009).

The OmpR family of winged-helix DNA-binding domains contacts the DNA through α-helices. One recognition helix interacts with the major groove of the DNA and the flanking loops interact with the minor groove. In the NarL and NtrC family, binding to the DNA is also mediated by α-helices (Stock et al, 2000).

1.3.2.1. The LytTR family of DNA-binding RRs

The LytTR family of DNA-binding RRs was first identified by Nikolskaya and Galperin in 2002. They analysed the sequences of bacterial genomes and identified a new kind of DNA-binding domain that was present in RRs and displayed several conserved motifs. This novel domain displayed no significant structural or sequence similarity to the different types of H-T-H DNA-binding domains and was predicted to consist mainly of β-sheets with only one or two α-helices. The RRs containing this LytTR DNA-binding domain all contained a CheY-like receiver domain. Interestingly, LytTR domain-containing transcription factors (LDCTFs) that had been previously characterised were found to be mostly involved in the regulation of genes involved in virulence. The virulence genes regulated by these LDCTFs are either involved in virulence towards the host such as toxins or extracellular polysaccharides or virulence towards other bacteria such as bacteriocins (See Table 1) (Nikolskaya & Galperin, 2002).

TFs	Organism	Regulated processes	Reference
AlgR	*Pseudomonas aeruginosa*	Alginate biosynthesis	(Deretic et al, 1989)
MrkE	*Klebsiella pneumoniae*	Type 3 fimbriae expression	(Allen et al, 1991)
LytR	*S. aureus*	Autolysis	(Brunskill & Bayles, 1996a)
VirR	*Clostridium perfringens*	Perfringolysin O production	(Shimizu et al, 1994)
AgrA	*S. aureus*	Toxins production	(Recsei et al, 1986)
ComE	*Streptococcus pneumoniae*	Competence	(Pestova et al, 1996)
CdtR	*Clostridium difficile*	Toxins production	(Carter et al, 2007)
SilA	*Streptococcus* group A	Toxins production	(Hidalgo-Grass et al, 2002)
FasA	*Streptococcus* group A	Toxins production	(Malke & Steiner, 2004)
StxR	*Lactobacillus sakei*	bacteriocins production	(Vaughan et al, 2003)
CbaR	*Carnobacterium piscicola*	bacteriocins production	(Quadri et al, 1997)
EntR	*Enterococcus faecium*	bacteriocins production	(O'Keeffe et al, 1999)
FsrA	*Enterococcus faecalis*	bacteriocins production	(Qin et al, 2001)
PlnC	*Lactobacillus plantarum*	bacteriocins production	(Diep et al, 1996)
PlnD	*Lactobacillus plantarum*	bacteriocins production	(Diep et al, 1996)
PlsR	*Lactobacillus plantarum*	bacteriocins production	(Stephens et al, 1998)
SppR	*Lactobacillus sakei*	bacteriocins production	(Huhne et al, 1996)
SapR	*Lactobacillus sakei*	bacteriocins production	(Axelsson & Holck, 1995)
BlpR	*Streptococcus peumoniae*	bacteriocins production	(Lange et al, 1999)

Table 1: Examples of LDCTFs and the function they regulate.

This table contains examples of transcription factors (TFs) containing a LytTR domain. The name of is TF is listed, the bacterium it comes from, the main function of the gene(s) it regulates as well as a reference paper. Adapted from (Nikolskaya & Galperin, 2002).

The bioinformatical analysis of the sequence of LDCTFs identified a motif that was extremely conserved, the FhRhH[RK][SNQ]hhVN motif where h represents a hydrophobic amino acid (aa) and a cluster of positively charged residues (K and R) at the C-terminal end of the domain that was hypothesised to play a role in DNA-binding (Nikolskaya & Galperin, 2002). The analysis of the genes regulated by LDCTFs established that these RRs bind to imperfect direct repeats and that their sequence pattern was [TA][AC][CA]GTTN[AG][TG] separated by a 12-13 bp spacer (de Saizieu et al, 2000; Diep et al, 1996; Nikolskaya & Galperin, 2002; Risoen et al, 1998; Ween et al, 1999).

LDCTFs represent only 3% of all RRs and yet they are disproportionally involved in the regulation of virulence genes which raised a lot of interest for this new family of DNA-binding RRs.

The number of LDCTFs appears to correlate with genome size, with large genomes typically encoding more LDCTFs than small ones. Thus, although LDCTFs are mostly associated with virulence genes regulation, free-living bacteria that tend to have larger genomes encode more LDCTFs (Nikolskaya & Galperin, 2002). These bacteria are known to encode more TCS in general as they have to adapt to a variety of different environments. Almost all Firmicutes encode LDCTFs with the exception of *Mycoplasma* spp. that do not possess any TCS (Beier & Gross, 2006; Kiil et al, 2005). A census of all LDCTFs in bacterial genomes conducted in 2007 identified 200 bacteria containing LDCTFs out of 413 genomes screened. A summary of this census can be found in Table 2. LDCTFs are only found in bacteria and are not present in archaea or eukaryotes (Nikolskaya & Galperin, 2002; Nikolskaya & Galperin, 2007).

Organism	**Number of LDCTFs**	**Frequency of LDCTFs**
Acidobacteria (2)	7-10	100%
Actinobacteria (39)	1-2, mostly 1	51%
Bacteroidetes (11)	3-23, avg 9	82%
Cholobi/green sulphur bacteria(5)	0	0%
Chlamydiae (7)	0	0%
Chloroflexi (7)	0	0%
Cyanobacteria (15)	0	0%
Deinococcus/Thermus (3)	0	0%
Firmicutes (70) *	1-10, mostly 2 or 3	91% *
α-proteobacteria (58) **	1-8, mostly 1 or 2	53% **
β-proteobacteria (38)	1 or 2, one has 4	66%
γ-proteobacteria (92)	1-7, mostly 1 or 2	69%
δ-proteobacteria (17)	1-3	53%

ε-proteobacteria (13)	1	15%
Spirochaetes (7)	1	14%
Thermotogae (6)	1	50%

Table 2: Distribution of LDCTFs among bacteria.

The first column contains the name of the Phylum and in bracket the number of member from that phylum. The second column contains the range of LDCTFs contained in that Phylum and if the range varies a lot, the number of LDCTFs contained in most member of the Phylum or an average (avg). The third column contains the percentage of members of the phylum that contain LDCTFs. * = excluding Mycoplasma spp. that do not contain LDCTFs. ** = excluding Rickettsia spp. that do not contain LDCTFs. Adapted from the updated version of supplementary figure S3 from (Nikolskaya & Galperin, 2002). This table was updated in 2007 with the additional bacterial genomes sequenced at that time (Nikolskaya & Galperin, 2007).

In 2008, the first crystal structure of the LytTR domain of the most extensively characterised LDCTF, AgrA from *S. aureus*, was published and confirmed the structure predictions. The LytTR domain of the RR AgrA (PDB 3BS1) consists mostly of β-sheets (See section 1.5.5.2). Interestingly, the structure enabled the discovery of a new method of contacting the DNA. Unlike H-T-H-containing TFs that bind to the DNA via α-helices, the AgrA LytTR domain binds DNA via three loops that separate the β-sheets (Sidote et al, 2008). The LytTR domain of AgrA will be discussed in more details in sections 1.5.5.2 and 4.1

To this date, the structure of two more LDCTFs have been deposited in the Protein Data Bank (PDB): a putative methyl-accepting/DNA RR from Bacillus cereus in 2008 by Osipiuk et al. but not published (PDB 3D6W) and ComE (See Table 1) from *Streptococcus pneumoniae* (PDB 4CBV) in 2014 (Boudes et al, 2014).

1.4. *Staphylococcus aureus*

S. aureus is a Gram positive bacterium and a major human pathogen. It is present in the nasal cavity, anterior nares and skin of the carriers but the anterior nares are thought to be the primary reservoir. 30%-60% of the population will carry *S. aureus* intermittently and ~20% of the population are persistent carriers thus "almost always" carriers. It is still unclear what exact factors determine whether someone is likely to be a carrier or not but it appears to be linked with the nasal epithelium cells. Interestingly, *S. aureus* has a greater affinity for the nasal epithelium cells from carriers as opposed to non-carriers and for patients with eczema as opposed to patients without eczema and thus adheres better to the cells of these patients. *S. aureus* is a human commensal, however it is also a frequent source of community-acquired and hospital-acquired infections. These two lifestyles are linked as nasal carriage is an important risk factor for the development of *S. aureus* surgical wound infection. *S. aureus* is an important nosocomial pathogen and the second most common cause of bacteraemia in the hospital settings. A *S. aureus* carrier is more likely to be infected than a non-carrier and in

80% of cases, the patients were infected by their own strain of *S. aureus*. Despite an increased susceptibility, the rate of mortality from *S. aureus* bacteraemia is higher in non-carriers than in carriers (Kluytmans et al, 1997; Peacock et al, 2001; Wertheim et al, 2005). Consequences of *S. aureus* infections are diverse and range from mild skin or soft tissue infections to life-threatening conditions such as endocarditis, bacteraemia, sepsis and toxic shock syndrome (Lowy, 1998).

1.4.1. Historical background

S. aureus was first identified in 1880 by a Scottish surgeon, Sir Alexander Ogston. He observed Gram positive micrococci forming masses that looked like bunches of grapes (that he later called *Staphylococci*, from the Greek staphyle meaning bunch of grapes) in the pus from an acute abscess. He observed next that the micrococci were always present in such abscesses in abundance. After trying different media to cultivate the micrococci he succeeded using newly-laid eggs. When he injected the micrococci cultivated in an egg or directly from the pus of abscesses into healthy animals they developed similar abscesses. He therefore demonstrated that the micrococci were causing the observed abscesses and validated Koch's postulates (Ogston, 1880). Four years later, a German physician, Friedrich Julius Rosenbach noticed that there were two different kinds of *Staphylococci* exhibiting a different pigmentation. He called the white pigmented micrococci *Staphylococcus pyogenes albus* (later renamed *Staphylococcus epidermidis* as it colonises the skin) and the yellow pigmented ones *Staphylococcus pyogenes aureus.* The *pyogenes* part of the name was later abandoned and the bacteria referred to as *Staphylococcus aureus* (Rosenbach, 1884).

1.4.2. *S. aureus* and antibiotic resistance

Resistance to antibiotics has always been a problem with *S. aureus*. Resistance was first encountered in the mid-1940s when penicillin-resistant strains started to spread in hospitals shortly after the introduction of the drug. The acquisition of a plasmid encoding a β-lactamase able to hydrolyse the β-lactam ring of penicillin was responsible for the resistance. The introduction of methicillin solved the problem until the first methicillin resistant *S. aureus* (MRSA) were reported in hospitals almost immediately after methicillin had been introduced in 1960. The acquisition of a single genetic element by the bacteria enabled them to be resistant to most commonly used antibiotics, β-lactams such as penicillins, cephalosporins and carbapenems. This large mobile genetic element, the staphylococcal cassette chromosome *mec* (SCC*mec*) encodes the *mecA* gene responsible for the methicillin resistance by encoding an additional penicillin binding protein (PBP2A) with reduced affinity

for β-lactam antibiotics. MRSA strains started to spread and had become endemic by the mid-1980s in hospitals worldwide and the leading-cause of hospital-acquired infections. These strains were however strictly hospital associated (HA-MRSA) and thus targeting people that were already vulnerable. The early 1990s saw the apparition of a new type of MRSA, the community acquired MRSA (CA-MRSA). CA-MRSA infections were reported in patients with no link to the healthcare settings and that were otherwise healthy. These strains were hyper virulent, lead to worse clinical outcomes than HA-MRSA and thus a higher mortality rate. The rapid spread of the CA-MRSA in the community worldwide highlighted *S. aureus* as a global health problem and raised serious concern (Chambers & Deleo, 2009; Thurlow et al, 2012).

The HA-MRSA and CA-MRSA pandemic lead to an increase in the use of vancomycin that was the only efficient antibiotic against MRSA. This triggered the appearance of the first vancomycin-intermediate *S. aureus* strains (VISA) in 1997, resistant to low doses of vancomycin. VISA strains have a thicker cell wall that absorbs the vancomycin and enables enough synthesis of new peptidoglycan for survival. These strains were unfit and therefore had a low endemic potential. However, the vancomycin-resistant *S. aureus* (VRSA) strains that have appeared in 2002 are more worrying as vancomycin is the preferred antibiotic for MRSA treatment and cases of VRSA have been reported in the USA, in India, in Iran and more recently in Europe and in Latin America. These strains acquired the vancomycin resistance gene cluster *vanA* from *Enterococcus faecalis* and have mostly emerged independently from poly-microbial infections treated with vancomycin. They appear to be less contagious than MRSA as they are not known to be transmitted by close contact. (Chambers & Deleo, 2009; Friaes et al, 2014; Lindsay, 2010; Rossi et al, 2014).

1.4.3. Pathogenicity of *S. aureus*

The pathogenic success of *S. aureus* can be attributed to a large array of virulence factors that are expressed in a growth-phase and cell density-dependent manner. *S. aureus* received its name because of its golden colour. This pigmentation is mediated by the staphyloxanthin which is also involved in virulence. It protects the bacterium against oxidative stress thus against neutrophil killing (Liu et al, 2005).

The colonisation of the host is the first step of the infection process and involves adherence to the host tissues. In *S. aureus,* this step is mediated by a number adhesins called microbial surface components recognising adhesive matrix molecules (MSCRAMMs). A few examples of MSCRAMMs include protein A, which displays anti-phagocytic properties due to its

ability to bind to the Fc chain of immunoglobulins, fibronectin-binding proteins that allow adhesion to the host fibronectin and elastin and are essential for the invasion and clumping factor A involved in the resistance to phagocytic uptake (Lowy, 1998).

Once the infection is established, the next step is to spread the infection through host tissue destruction. This step involves a plethora of secreted virulence factors such as the cytotoxins haemolysins (α, β, γ and δ), toxic shock syndrome toxin 1 (TSST-1), Panton-Valentine leukocidin (PVL), proteases, lipases and phenol-soluble modulins (PSMs). The TSST-1 binds to T cells and causes an expansion of clonal T-cells and results in a massive release of cytokines that mediate the toxic shock syndrome. Haemolysins mediate the lysis of red blood cells through pore formation for α-haemolysin which is critical for virulence. β and δ haemolysins act synergistically to prevent phagocytosis. The proteases and lipases are involved in host tissue destruction. PSMs are involved in neutrophil and red blood cells killing, biofilm formation and dispersal. PVL is a bi-component toxin, composed of two secreted proteins S (slow) and F (fast) and is able to lyse white blood cells (Dinges et al, 2000; Giese et al, 2011; Lowy, 1998; Peschel & Otto, 2013)

1.4.4. TCS in *S. aureus*

There are 16 or 17 TCS clearly identified as such in *S. aureus* (Kuroda et al, 2001). The strains that have the SCCmec cassette (Ito et al, 2001) have 17 TCS as a homolog of the KdpD/E TCS is encoded on this cassette. The strains that do not have the cassette encode 16 TCS (Hanssen & Ericson Sollid, 2006). The 16 TCS identified in the NCTC8325 strain are listed in Table 3. For the sake of clarity, only one function was listed in this table. The different functions of each TCS will be detailed in the text below and the different DNA-binding domains they contain listed in Figure 16. This section is meant as an overview of the main functions assumed by TCS in *S. aureus* and does not provide a complete description of all the functions of each TCS.

The nomenclature for the same TCS in S. aureus vary in the literature, therefore to facilitate comprehension, the TCS will be referred to using the most recent name and the previous names will be listed as well. For the TCS that have not been characterised in *S. aureus* but are homologous to a TCS that has been characterised in another bacteria, the function of the characterised TCS will be described (See Table 3 and Figure 16).

HK name	HK number	RR name	RR number	Function
AgrC	SAOUHSC_02264	AgrA	SAOUHSC_02265	Virulence genes regulation
unassigned	SAOUHSC_00185	unassigned	SAOUHSC_00184	Unknown
LytS	SAOUHSC_00230	LytR	SAOUHSC_00231	Autolysis
GraS	SAOUHSC_00666	GraR	SAOUHSC_00665	Antibiotic resistance
SaeS	SAOUHSC_00714	SaeR	SAOUHSC_00715	Virulence genes regulation
DesK	SAOUHSC_01313	DesR	SAOUHSC_01314	Unknown in *S. aureus*
ArlS	SAOUHSC_01419	ArlR	SAOUHSC_01420	Biofilm formation
SrrB	SAOUHSC_01585	SrrA	SAOUHSC_01586	Virulence genes regulation
PhoP	SAOUHSC_01799	PhoR	SAOUHSC_01800	Unknown in *S. aureus*
AirS	SAOUHSC_01981	AirR	SAOUHSC_01980	Nitrate respiration
VraS	SAOUHSC_02099	VraR	SAOUHSC_02098	Cell wall biosynthesis
KpdD	SAOUHSC_02314	KdpE	SAOUHSC_02315	Virulence genes regulation
HssS	SAOUHSC_02644	HssR	SAOUHSC_02643	Heme sensing
NreB	SAOUHSC_02676	NreC	SAOUHSC_0267C	Unknown in *S. aureus*
BceS	SAOUHSC_02956	BceR	SAOUHSC_02955	Antibiotic resistance
WalK	SAOUHSC_00021	WalR	SAOUHSC_00020	Cell wall biosynthesis

Table 3: The 16 TCS identified in NCTC 8325.

This table lists the two component systems (TCS) identified in the N315 strain and the gene they correspond to in the NCTC 8325 strain. The table contains the name of the histidine kinase (HK) and the corresponding response regulator (RR) (when assigned) and their accession number as well as their function, when available. More details can be found for each TCS in the text below. This list was compiled based on the TCS reported in the N317 strain (Kawada-Matsuo et al, 2011; Kuroda et al, 2001) and in the MW2 strain (Matsuo et al, 2010). The KEGG pathways database (Kanehisa et al, 2014), the AureusDB (http://aureusdb.biologie.uni-greifswald.de/index.shtml) and the BioCyc database (Caspi et al, 2014) were used to identified the TCS that were not annotated in the NCTC 8325 genome.

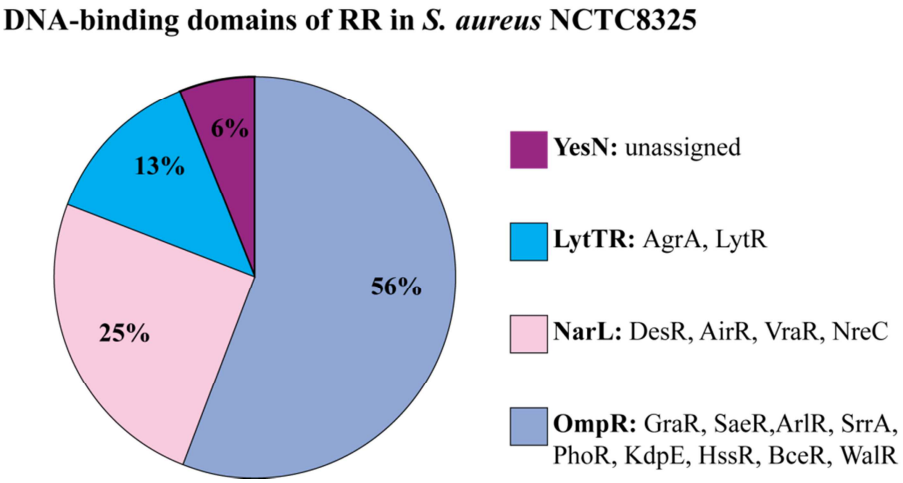


Figure 16: Repartition of the different types of DNA-binding domains in response regulators (RRs) from *S. aureus* NCTC 8325.

Pie chart representing the proportion of each type of DNA-binding domain. The DNA-binding domain types are colour coded and the percentage of each is indicated. The DNA-binding domains of each of the 16 RRs from *S. aureus* NCTC 8325 have been identified using the P2CS database (Barakat et al, 2011).

1.4.4.1. The AgrCA TCS

The AgrCA TCS is a key component in the ability of *S. aureus* to regulate the expression of colonisation and virulence factors in response to environmental cues (Novick, 2003). It is responsible for the upregulation of secreted virulence factors and the downregulation of the synthesis of cell surface-associated virulence factors that takes place during the late exponential phase (Peng et al, 1988; Recsei et al, 1986). The AgrA RR directly regulates phenol soluble modulins (PSMs) and the *bsaA* gene that is essential for survival under oxidative stress (See section 1.5.5.4) (Queck et al, 2008; Sun et al, 2012b). However, most of the genes regulated by AgrCA (over 200) are regulated via RNAIII. RNAIII is the effector RNA of the *agr* operon that is activated by AgrA at the transcriptional level (Cheung et al, 2011; Ji et al, 1995; Morfeldt et al, 1995; Morfeldt et al, 1996; Novick et al, 1995; Novick et al, 1993). The AgrCA TCS also regulates biofilm formation in *S. aureus* (Vuong et al, 2000; Yarwood et al, 2004). Therefore the AgrCA TCS is involved in the regulation of virulence genes, biofilm formation and resistance to oxidative stress as well as a number of other processes.

This section is meant as an overview of the role of AgrCA in *S. aureus*. This TCS will be described in more detail later in this chapter (See section 1.5).

1.4.4.2. The LytSR TCS

The LytSR TCS is essential for normal biofilm development and is important to control cell lysis during biofilm formation. The LytR RR regulates genes involved in cell death and lysis such as peptidoglycan (murein) hydrolases and thereby represses cell wall hydrolytic activity (Brunskill & Bayles, 1996b; Sharma-Kuinkel et al, 2009). Therefore, LytSR mutants display increased autolysis activity (Brunskill & Bayles, 1996a). In addition, LytSR is involved in the adaptive response to antimicrobial peptides perturbing the cell membrane most-likely by detecting subtle changes in the transmembrane potential. LytSR mutants have been shown to have increased susceptibility to cationic antimicrobial peptides *in vitro* (Yang et al, 2013).

1.4.4.3. The GraSR TCS

The GraSR (ApsSR) TCS is actually a three-component system consisting of the GraS, GraR and GraX proteins. It is required for growth of *S. aureus* at high temperatures, for resistance to oxidative stress and for cationic antimicrobial peptide sensing and resistance (Falord et al, 2012; Falord et al, 2011). The involvement of the GraSRX in resistance against antimicrobial peptides makes it an important system for the virulence of *S. aureus in vivo* (Li et al, 2007). GraSRX also controls the stress response and cell wall metabolism signal transduction

pathways. The GraSRX is subject to regulation by the *agr* system thereby linking *agr* to antimicrobial agent susceptibility. It is therefore strongly expressed during the exponential phase and repressed by *agr* during the stationary phase (Matsuo et al, 2011).

1.4.4.4. The SaeSR TCS

SaeRS is an important regulator of virulence factors and a key element of *S. aureus* virulence regulation (Novick & Jiang, 2003; Rogasch et al, 2006). It is required in order for *S. aureus* to cause infections through its regulation of genes involved in adhesion and invasion and affects bacterial survival during an infection (Liang et al, 2006). SaeRS has been shown to be critical for the pathogenesis of *S. aureus* in a murine infection model. It is responsible for the upregulation of virulence gene transcription. It is also linked with other TCS: SaeRS regulates the ArlSR TCS which negatively regulates secreted virulence genes and acts synergistically with the *agr*. It lies downstream of *agr* in the exoproteins activation pathway (Novick & Jiang, 2003; Nygaard et al, 2010)

1.4.4.5. The DesKR TCS

The DesKR TCS has not been investigated in *S. aureus* but has been characterised in another Gram positive bacterium, *Bacillus subtilis*. In *B. subtilis*, DesKR is responsible for the control of gene expression at low temperatures, mainly membrane lipids. It is involved in the adjustment of membrane lipid composition at cold temperature. The RR DesR activates a lipid desaturase which is responsible for the change in membrane composition under cold temperature in order to regulate membrane fluidity (Aguilar et al, 2001).

1.4.4.6. The ArlSR TCS

The ArlSR TCS is involved in autolysis and is also a repressor of biofilm development. The disruption of the *arlS* gene increases the activity of peptidoglycan hydrolases in the cell wall leading to autolysis (Fournier et al, 2001; Toledo-Arana et al, 2005). ArlSR is also indirectly responsible for the decreased production of virulence factors by downregulating their transcription. Mutations in the ArlSR TCS result in an increased production of secreted proteins. ArlSR regulates *sarA* and *agr* therefore its effect on virulence factors is mainly dependent on them. It stimulates SarA and downregulates RNAIII transcription (Fournier et al, 2001)

1.4.4.7. The SrrBA TCS

The SrrBA TCS acts in the global regulation of *S. aureus* virulence. It is responsible for the regulation of exotoxins and cell surface-associated virulence factors in response to the level

of oxygen in the environment. This regulation is partly *agr* mediated as RNAIII expression is inversely dependent on SrrBA. SrrA regulates *agr* by binding to the P2 and P3 promoters (Pragman et al, 2004; Yarwood et al, 2001). Mutations in SrrBA result in a decrease in virulence of the bacteria in rabbit endocarditis model and knockout strains exhibit a growth defect. The expression of SrrBA is necessary for growth in low oxygen or on a media poor in nutrient (Pragman et al, 2007; Pragman et al, 2004). SrrBA is also a major activator of intracellular adhesins and polysaccharide intracellular adhesins in an anaerobic environment which contributes to the protection of *S. aureus* against non-oxidative defence mechanisms (Ulrich et al, 2007).

1.4.4.8. The PhoPR TCS

The PhoPR TCS has not been characterised in *S. aureus* but in *B. subtilis* it controls the synthesis of the ATPases, alkaline phosphodiesterase and glycerophosphoryl diester phosphodiesterase and lipoproteins during phosphate starvation (Antelmann et al, 2000; Hulett et al, 1994).

1.4.4.9. The AirSR TCS

The AirSR (YhcSR) TCS has been mainly characterised by two different research groups that do not seem to validate each other's findings. Therefore there is controversy as to its precise functions. It was first shown to be essential for growth *in vitro*. Its downregulation results in increased susceptibility to the cell wall synthesis inhibitor phleomycin (Sun et al, 2005). AirSR is also involved in the modulation of the nitrate respiratory pathway under anaerobic conditions. Nitrate induces the transcription of AirSR under anaerobic conditions. In turn, AirSR directly positively regulates the transcription of the nitrate reductase *narG* and the oxygen responsive NreABC TCS. This enables the modulation of the nitrate respiratory pathway of *S. aureus* under anaerobic conditions. The downregulation of AirSR eliminates the enhancing effect of nitrate on bacterial growth under anaerobic conditions and results in a dramatic decrease in transcription of the genes involved in nitrate respiration (Yan et al, 2011). In order to be fully active the AirS needs to be oxidised with a $[2Fe\text{-}2S]^{2+}$cluster. Over oxidation either by prolonged O_2 exposure or contact with H_2O_2 or NO leads to the inactivation of the AirS kinase activity. Mutation in the AirSR TCS impacts the expression of ~355 genes under anaerobic conditions. The mutant strain displays an increased resistance to H_2O_2, vancomycin, norflaxin and ciproflaxin under anaerobic conditions. Therefore the AirSR is a redox-dependent global regulatory system playing an important role in gene regulation using redox active Fe-S cluster under oxygen limited conditions (Sun et al, 2012a).

The deletion of AirSR results in the downregulation of genes involved in cell-wall metabolism, reduced autolysis and reduced viability with vancomycin. AirSR is able to directly bind and regulate genes functioning in cell wall metabolism, protection against phagocytosis and autolysis. The involvement of AirSR in cell wall biosynthesis and turnover is linked to vancomycin resistance (result of a thickened cell wall). However, in this study it was to found to be essential, which could be the result of strain variation (Sun et al, 2013).

1.4.4.10. The VraSR TCS

The VraSR TCS is responsible for the coordination of important steps in cell wall biosynthesis. It is induced by exposure to cell wall affecting antibiotics such as glycopeptides, beta-lactams, bacitracin, vancomycin and knockout strains display increased susceptibility to these antibiotics (Gardete et al, 2006; Kuroda et al, 2003).

1.4.4.11. The KdpDE TCS

The TCS KdpDE upregulates the *cap* operon, responsible for the synthesis of capsular polysaccharides. They are important cell wall components that can interact with the host immune system and allow to resist uptake and killing by phagocytes. The *cap* operon is a major determinant of *S. aureus* virulence (Zhao et al, 2010). KdpDE is also a global regulator of virulence genes and activates the transcription of genes encoding cell wall associated proteins and polysaccharides synthesis genes while repressing the transcription of toxin genes in response to external K^+ concentration. In addition, the transcription level of KdpD is affected by external K^+ concentration. It decreases as the level of external K^+ increases. This indicates that *S. aureus* might modulate its infectious status through sensing specific external K^+ stimuli in different environments. KdpDE is strongly upregulated by *agr*, and might be an important regulator in order to coordinate external K^+ with *agr* during pathogenesis. In the environment, high levels of KdpD help to activate the transcription of cell wall proteins and polysaccharides synthesis genes beneficial for colonisation. In the host, high K^+ leads to a decrease in KdpD transcript and low expression of cell wall proteins but high production of extracellular toxins and enzyme facilitating local invasion. Therefore KdpD plays an important role in the pathogenesis of *S. aureus* (Xue et al, 2011). KdpD acts also as specific c-di-AMP binding protein and receptor for c-di-AMP (Corrigan et al, 2013).

1.4.4.12. The HssSR TCS

The HssSR TCS is the heme sensor system. It is necessary for growth with high concentration of heme. It responds to heme exposure and activates the expression of the

heme-regulated transporters, the efflux pumps. These pumps protect the cells against heme-mediated damage. Heme is an essential source of iron but can be toxic when in excess due to its reactivity. The inactivation of HssSR results in increased virulence in a vertebrate infection model. Without HssSR, *S. aureus* cannot sense and excrete the excess heme and therefore exhibits an increased virulence factor expression and secretion which leads to increased host tissue damages. HssSR is only activated by extracellular heme and specifically recognizes host derived heme. Therefore HssSR enables the detection of the host internal tissues as heme is sensed as a marker of internal host tissues and regulates virulence to avoid inflicting excessive host tissue damage (Torres et al, 2007) HssRS also modulates the susceptibility to the antimicrobial peptide plectasin. Mutations in HssRS make *S. aureus* more resistant to pectasin and eurocin therefore provides susceptibility to these antimicrobial peptides (Thomsen et al, 2010).

1.4.4.13. The NreBC TCS

The NreBC TCS has not been yet characterised in *S. aureus* but has been in *S. carnosus*. It is actually a three component system composed of the NreA, NreB and NreC genes. NreA is a nitrate receptor which controls the phosphorylation level of the HK NreB, an oxygen sensor, in a nitrate dependent manner. In turn, NreB acts to control NreC. NreA interacts with NreB but not with NreC. There is therefore nitrate/oxygen co-sensing performed by the NreABC system (Nilkens et al, 2014). Deletion mutants have been shown to be severely impaired in nitrate and nitrite reduction and to have growth defects. NreABC is responsible for enhancing the transcription of genes involved in nitrate reduction and transport and a nitrite reductase. It is also involved in regulating transcription in the absence of oxygen (Fedtke et al, 2002).

1.4.4.14. The BceSR TCS

The BceRS (BraSR) TCS regulates two transporter that determine susceptibility to bacitracin, an antimicrobial peptide. It might therefore provide a mechanism of resistance to antimicrobial agents targeting cell wall biosynthesis in the membrane (Hiron et al, 2011; Kawada-Matsuo et al, 2011).

1.4.4.15. The WalKR TCS

The WalKR (YycFG, VicKR) is the only TCS that is universally accepted as being essential for survival (Martin et al, 1999). It is the master regulator of cell wall formation. WalKR activates the transcription of genes involved in cell wall degradation, controls autolytic activities by positively regulating the synthesis of autolysins (cell wall hydrolases) and

positively regulates biofilm formation. Without WalKR, the peptidoglycan biosynthesis and turnover is significantly decreased, the growth is therefore slower and the cells are unable to form biofilms. WalkKR mutants are also unable to lyse and are hypersensitive to macrolide and lincosamide antibiotics and vancomycin due to a thickened cell wall (Dubrac et al, 2007; Shoji et al, 2011).

1.4.4.16. Summary of *S. aureus* TCS

The *S. aureus* TCS regulate a wide range of functions essential for bacterial survival and adaptation to changing environments. Some functions are regulated by different TCS in order to enable a better modulation and to respond to different environment cues. This section has provided a summary of the functions that are regulated by more than one TCS in *S. aureus*. This illustrates that almost all the function essential for bacterial life are tightly regulated by one or more TCS (See Table 4).

Function	TCS involved
Cell wall metabolism	VraSR, WalKR, KdpDE
Biofilm formation	ArlSR, AgrCA, WalKr, LytSR
Autolysis	LytSR, ArlSR, WalKR
Antibiotic resistance	VraSR, GraSR, ArlSR, SrrRS, BceRS, AgrCA, LytSR, HssSR, WalKR
Virulence factors regulation	ArlSR, GraSR, SaeRS, VraSR, SrrRS, KdpDE, AgrCA
Redox response	SrrAB, AgrCA
Regulated by *agr*	GraSR, ArlSR, SrrRS, SaeRS
Involved in *agr* regulation	ArlSR,SrrAB
Anaerobic respiration	SrrAB, AirSR

Table 4: List of processes regulated by more than one two component system (TCS).

The functions regulated by more than one TCS are listed and the name of the TCS involved are listed.

1.5. The accessory gene regulator (*agr*) operon

It is essential for bacteria to exchange information in order to act as a population. This mechanism, called quorum-sensing is universal in the microbial world. In *S. aureus*, the ability to switch gene program once a certain threshold of population has been reached and transition from a colonisation state to the invasion mode is mediated by quorum sensing via the *agr* operon. The *agr* operon is the main quorum-sensing system in *S. aureus* (Singh & Ray, 2014).

The *agr* operon consists of two transcriptional units, RNAII and RNAIII, transcribed by two divergent promoters, P2 and P3, respectively. RNAII encodes the quorum sensing system of the operon consisting of the *agr* B, D, C and A genes (See Figure 17). The *agrD* gene

encodes a pro-peptide which is processed and secreted by the transmembrane AgrB, with the help of the SpsB peptidase, thus releasing the finalised autoinducing peptide (AIP). The AIP is the quorum signalling molecule of the operon. AgrC and AgrA are members of a TCS in which AgrC is the transmembrane receptor HK which senses the cell density-dependent accumulation of the AIP. Once a certain threshold of AIP concentration has been reached, the AIP binds to AgrC. Upon binding the AIP, AgrC autophosphorylates in an ATP-dependent manner. The autophosphorylation of AgrC leads to the activation of the LytTR domain-containing DNA-binding RR AgrA via phosphotransfer. The activated AgrA is then able to bind to direct tandem repeats at the P2 and P3 promoter and thus activate transcription of these two promoters. AgrA therefore stimulates its own synthesis in an autocatalytic circuit as mentioned in section 1.4.4.1 (Novick & Geisinger, 2008). RNAIII is the pleiotropic effector RNA molecule of the operon and is responsible for the downregulation of cell surface-associated virulence factors and the upregulation of the secreted virulence factors (main examples listed in section 1.4.3). RNAIII is involved in the regulation of over 200 genes and thus has a tremendous metabolic burden. This is probably the reason for the very high frequency of spontaneous mutations inactivating the *agr* operon. The C-terminal part of RNAIII encodes a toxin, δ-haemolysin (Cheung et al, 2011; Novick & Geisinger, 2008).

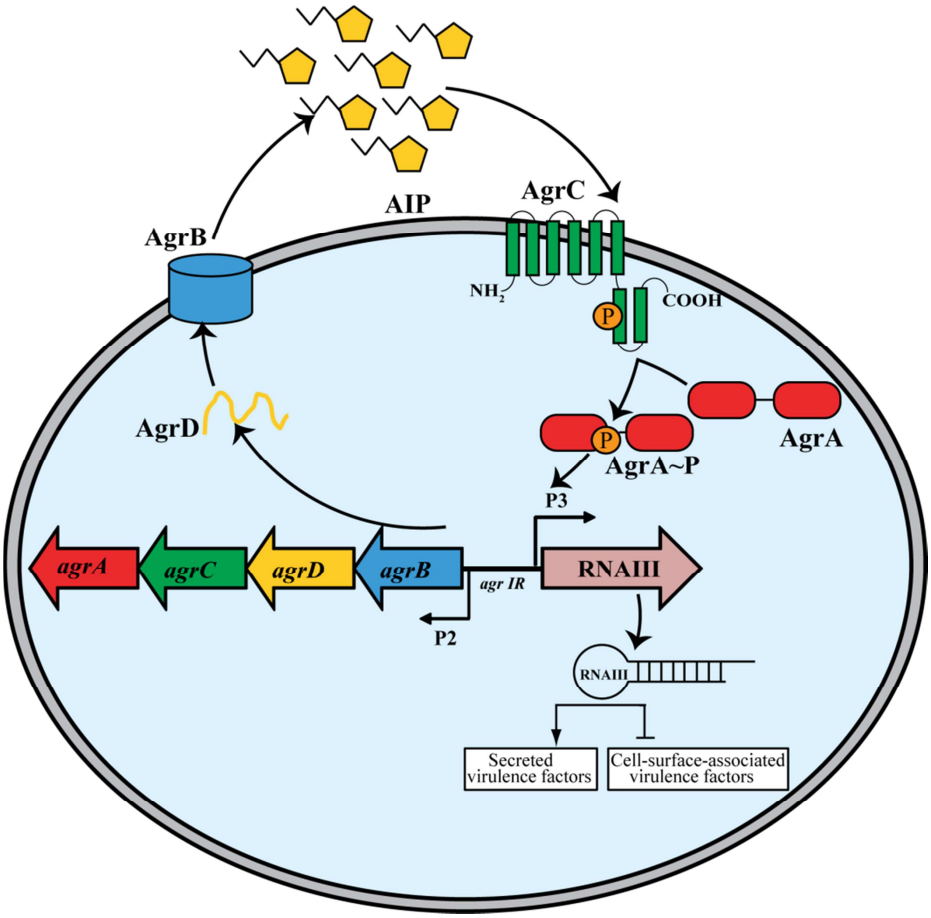


Figure 17: Schematic representation of the organisation of the *agr* operon and the auto-catalytic circuit.

agrD encodes a propeptide that is processed and secreted by the transmembrane AgrB with the help of SpsB releasing the finalised AIP. AgrC senses the cell density-dependent accumulation of the AIP and autophosphorylates upon binding the AIP. The phosphorylated AgrC is then able to activate AgrA via phosphotransfer. The phosphorylated AgrA is then able to bind to specific direct repeats at the *agr* IR and activate transcription at the P2 and P3 promoter. The P3 promoter encodes RNAIII, the pleiotropic effector RNA of the operon. RNAIII is responsible for the upregulation of secreted virulence factors and the downregulation of cell surface-associated virulence factors. By activating the P2 promoter, AgrA creates an autocatalytic circuit as it stimulates its own transcription.

1.5.1. *agr* interference

There are 4 *agr* specific groups in *S. aureus* resulting from allelic variation in the *agr* B, D and C genes which correspond to 4 different strain pathotypes. These 3 genes have evolved in parallel in order to remain compatible. The *agr* operon of each group is mutually inhibited by the AIP from the other groups in a mechanism where the foreign AIP "blocks" the AgrC receptor thus rendering it unavailable for its cognate AIP. This mechanism is a mode of bacterial interference through inhibition of the synthesis of virulence factors. The sequence of AgrB is variable except for the first 34 aa. AgrC is conserved except for the N-terminus responsible for the interaction with the AIP. AgrD is variable except for its C-terminal charged tail and essential cysteine residue involved in the thiolactone ring formation.

Therefore, only the regions involved in the specificity of AgrD processing and for ligand-receptor interactions are variable (See Figure 18). The different groups exhibit differences in the timing and the magnitude of the *agr* response (Dowell et al, 2001; Dufour et al, 2002; Geisinger et al, 2012; Ji et al, 1997; Lyon et al, 2002).

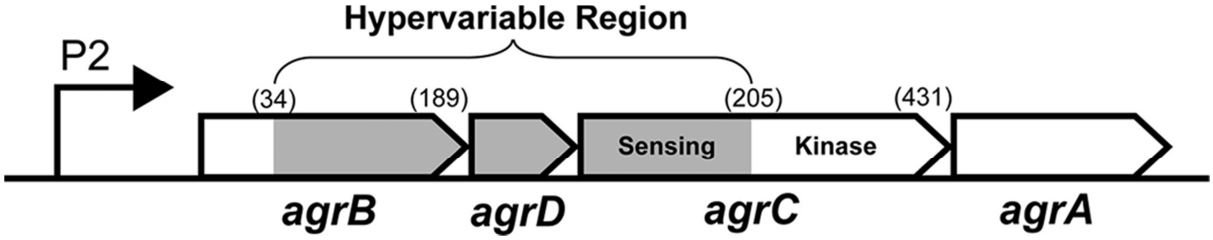


Figure 18: Conserved and variable regions in *agrB, D, C* and *A*.

Schematic of the *agr* operon. The conserved regions are displayed in white and the variable ones in grey. The number above the genes indicates the number of the residues at the border between the conserved and variable areas (Thoendel et al, 2011).

1.5.2. AgrB

AgrB is a cysteine endopeptidase transmembrane protein responsible for the proteolytic processing of AgrD. AgrB consists of 6 transmembrane segments, 4 hydrophobic ones and 2 hydrophilic ones. AgrB is an endopeptidase and two residues (cysteine 84 and histidine 77) form the catalytic centre responsible for the cleavage of AgrD at its C-terminal processing site. The interaction between AgrB and the AgrD propeptide is group specific, i.e. the AgrB from group 1 is responsible for the processing of the AgrD propetide from that same group. The region responsible for this group specificity differs between different groups. The first transmembrane α-helix and extracellular loop are essential for group specific interaction between AgrB1 and AgrD1 (See Figure 19**A**). In AgrB2, the two hydrophilic regions are required for the specific processing of AgrD2 (See Figure 19**A**). Recently, a random mutagenesis study enabled the identification of a number of residues in AgrB that are important for the processing of AgrD (See Figure 19**B**). This study also reported that the model of AgrB topology that had been used until then was wrong and that the C-terminal and N-terminal end of AgrB were in fact located inside the cell and not outside as was previously thought (See Figure 19**B**) (Qiu et al, 2005; Thoendel & Horswill, 2013; Zhang et al, 2002; Zhang & Ji, 2004).

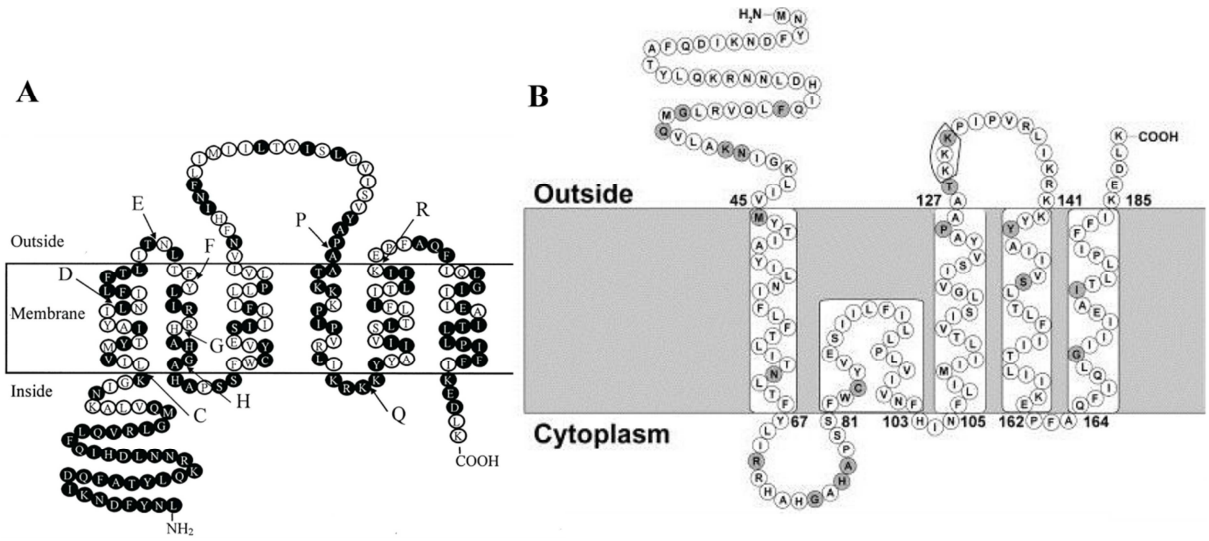


Figure 19: Old and new model of the transmembrane topology of AgrB.

A: Old model of AgrB1 where the N and C-terminus are intracellular. The aa residues that are identical between AgrB1 and AgrB2 are highlighted in black (Zhang & Ji, 2004).

B: New model of AgrB1 where the N and C-terminus are extracellular. The aa that lead to decrease autoinducing peptide (AIP) production when mutated are shaded in grey. The numbers indicated correspond to the residues numbers (Thoendel & Horswill, 2013).

1.5.3. AgrD

AgrD is a membrane protein and is anchored in the membrane via its N-terminal region. This membrane localisation is essential in order for AgrD to be processed into a mature AIP. AgrD is composed of three different parts: the 24 N-terminal residues are part of an amphipathic leader, the central 8 residues constitute the AIP and the 14 C-terminal residues that are predominantly charged. The C-terminal charged tail of AgrB contains 9 residues that are essential for AIP production and the endopeptidase activity of AgrB (Kavanaugh et al, 2007; Thoendel & Horswill, 2009; Zhang et al, 2004).

The processing of AgrD into the finalised AIP consists of five steps for which both AgrB and AgrD are required (See Figure 20). Firstly, AgrD is localised to the cytoplasmic membrane via its N-terminal leader. Secondly, the C-terminal tail of AgrD is removed by AgrB using its endopeptidase activity. Thirdly, a conserved cysteine residue in AgrD catalyses the formation of the thiolactone ring through thioester exchange and the release from AgrB. Fourthly, the AIP precursor is exported through the membrane by an unknown mechanism probably involving AgrB. Finally, a signal peptidase, SpsB removes the N-terminal leader (Kavanaugh et al, 2007; Qiu et al, 2005; Zhang et al, 2002; Zhang et al, 2004)

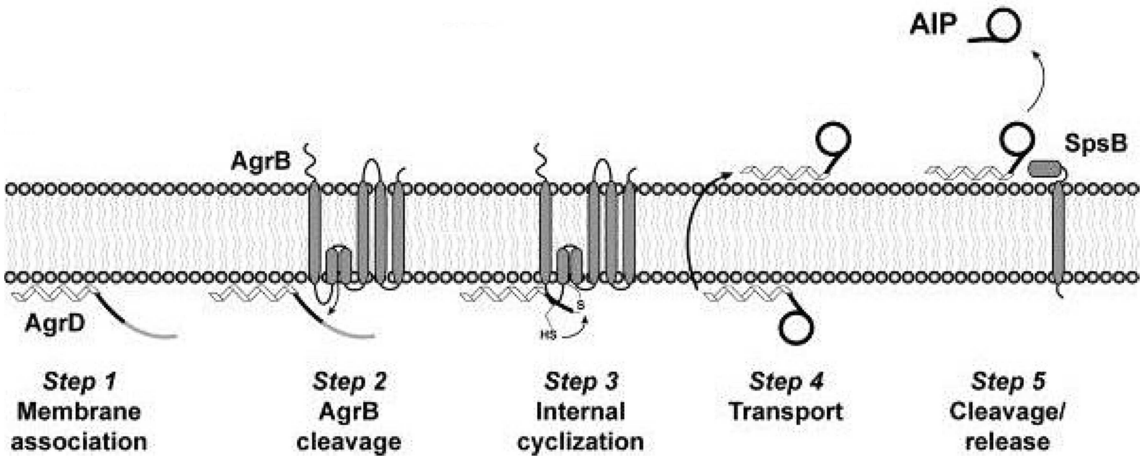


Figure 20: Model of the autoinducing peptide (AIP) synthetic pathway.

Step 1: AgrD localises to the cytoplasmic membrane via its N-terminal leader. Step 2: The C-terminal tail of AgrD is removed by the endopeptidase activity of AgrB. Step 3: Formation of the thiolactone ring through thioester exchange and release from AgrB. Step 4: The AIP precursor is exporter through the membrane via an unknown mechanism. Step 5: SpsB removes the N-terminal leader of AIP precursor (Thoendel & Horswill, 2013).

The processed product of AgrD is a thiolactone peptide pheromone responsible for the activation of the *agr* operon processed from the 46 aa AgrD precursor and AgrC is its receptor. The AIP varies from seven to nine aa depending on the compatibility group and all groups contain a conserved cysteine residue 5 aa from the C-terminus (See Figure 21). The sulfhydryl group of that cysteine is involved in a thiolactone ring with the α-carboxyl group of the peptide and is required for the interaction with AgrC. The ring and hydrophobic tail are necessary and sufficient for activation of AgrC by the AIP and for inhibition (Dufour et al, 2002; Ji et al, 1997; Ji et al, 1995; Lyon et al, 2002; Mayville et al, 1999; Wright et al, 2004; Zhang et al, 2002).

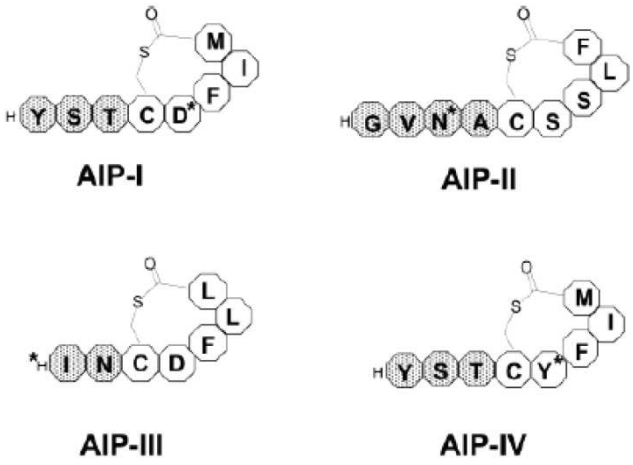


Figure 21: Structure of the AIP from the 4 different *agr* groups.

The sulphur atom of the cysteine and the carbonyl from the C-terminal involved in the thioester linkage are shown. Residues in the ring are shown in white while the ones from the tail are shaded in grey. The residues

that are critical for receptor activation are marked with an asterisk. For AIP-III the C-terminal asterisk illustrates that adding additional aa prevents receptor activation (Lyon et al, 2002).

1.5.4. AgrC

AgrC is the sensor transmembrane receptor HK of the *agr* operon and a member of the family of peptide-inducible HKs. AgrC consists of an N-terminal sensor domain containing 6 membrane spanning α-helixes (aa 1-200 in AgrC1), a linker region (aa 201-221 in AgrC1), a cytoplasmic domain consisting of the DHp domain (aa 222-290 in AgrC1) containing the conserved histidine constituting the autophosphorylation site (H239 in AgrC1) and a C-terminal CA domain (aa 293-430 in AgrC1) responsible for accepting an ATP molecule (See Figure 22). The DHp and the CA domain are required for autophosphorylation but the DHp domain is sufficient for phosphotransfer (Gao & Stock, 2009; Thoendel et al, 2011; Wang et al, 2014).

The interaction of AgrC with the AIP regulates the catalytic activity of AgrC and the autophosphorylation of AgrC is only triggered by the AIP. The ligand-binding site is located in the third extracellular loop of the protein, in the N-terminal part of the sensory domain and consists of a hydrophobic pocket. When the sensory domain is missing, the truncated AgrC has the ability to autophosphorylate spontaneously in the absence of any signal. Therefore the sensory domain must act to prevent spontaneous autophosphorylation. The AIP makes two types of interactions with AgrC: a hydrophobic interaction involving residues in the C-terminal hydrophobic tail of the AIP and sequence specific contacts that trigger the activation or inactivation of the receptor (Lina et al, 1998; Wright et al, 2004).

Mutations that lead to constitutively active AgrC have been identified and target residues located either in the last transmembrane helix or in the DHp domain, close to the histidine responsible for autophosphorylation. The constitutive mutants containing mutations in the DHp domain were not sensitive to inhibition with the AIP from another group and were therefore "locked" in the active state. On the contrary, constitutively active mutants containing a mutation in the sensory domain were inhibited by certain AIP from other groups that were able to reverse the constitutive activity of AgrC. A residue in last extracellular loop, I171, was identified as the determinant of inhibition as the mutation I171K this residue reduced the sensitivity of AgrC to inhibition by AIPs from other groups. Instead, the AgrC bearing this mutation exhibited a broadened specificity and could be activated by the AIP from several other groups (See Figure 22). The signal recognition mechanism appears to be

confined to the N- terminus of the protein, the membrane associated and extracellular components that are more variable. This suggests that these regions have evolved in order to respond to different ligands while the C-terminal part of the protein has been conserved in order to preserve the catalytic activity (Geisinger et al, 2009; Thoendel et al, 2011).

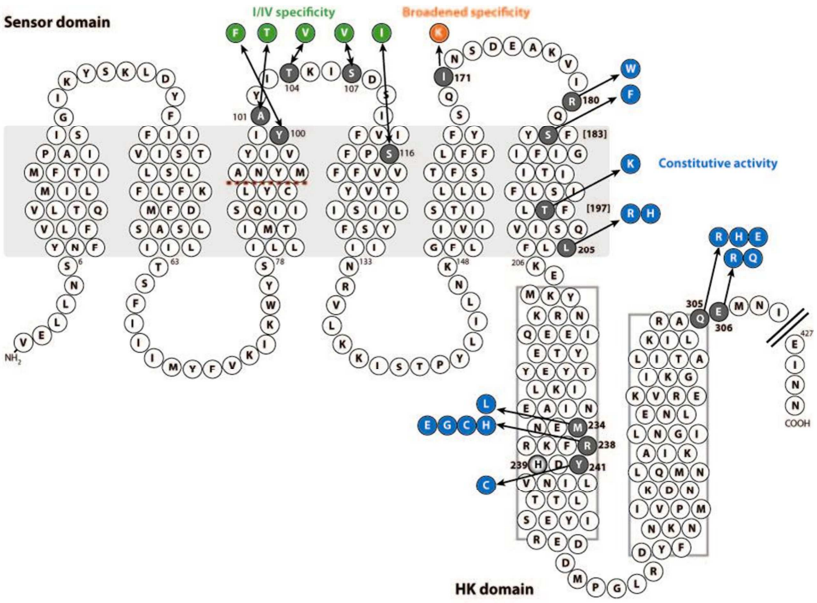


Figure 22: Topology of AgrC1.

The mutations coloured in green represent the differences between AgrC1 and AgrC4 in the sensor domain. The residue involved in broadened specificity, I171 is coloured in red and the mutations leading to a constitutive activity of AgrC are coloured in blue. The numbers correspond to the number of the residues (Novick & Geisinger, 2008).

AgrC is an obligate dimer, like all HKs and dimerisation happens before any interaction with the AIP (See Figure 23). Mutant receptor studies have demonstrated that the activation of one monomer by the AIP triggers *trans*-autophosphorylation and that one sensory domain is sufficient for activation. Similarly, a mutant dimer containing a constitutively active monomer and a WT monomer displayed full AIP-independent activity as observed with a dimer of two constitutively active AgrC mutants. Therefore, the activation of any monomer triggers the symmetric activation of both kinases domains and the autophosphorylation activity of AgrC happens in *trans* (George Cisar et al, 2009).

Very recently, the mechanism of activation/inhibition by the AIP was uncovered. It was shown that the sensory and the kinase domain are connected via a helical linker. The binding of the AIP to the sensor domain of AgrC triggers a twisting of the linker in different directions depending whether the AIP is the cognate AIP (counter clockwise rotation) or an inhibitor (clockwise rotation). Furthermore, it was demonstrated that AgrC lacks the

phosphatase activity present in some HKs and that AgrA alone is responsible for its dephosphorylation. This implicated that the AIP cannot modulate the rate of AgrA dephosphorylation. (Wang et al, 2014)

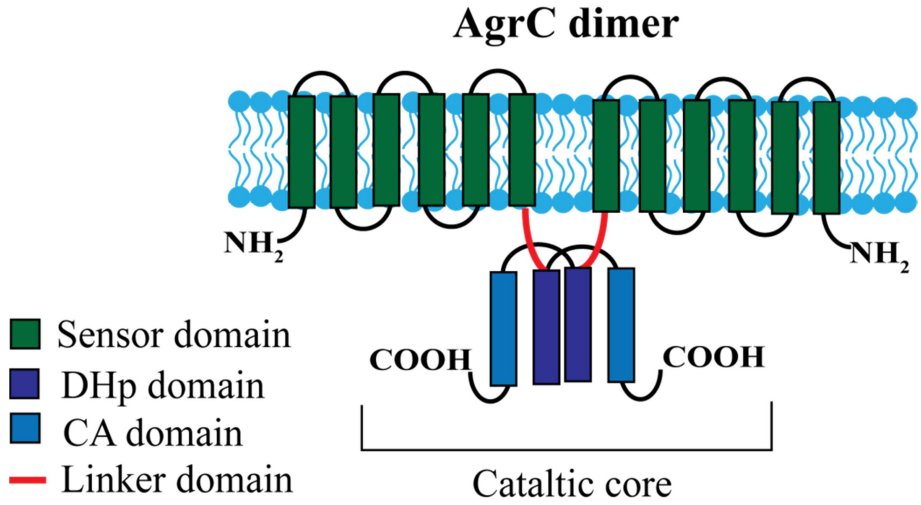


Figure 23: Domain architecture of the AgrC dimer.

Adapted from Wang et al, 2014.

Very recently, the first crystal structure of the AgrC CA domain in the presence of ATP was reported (aa 278-430) and consists of a 6 stranded β-sheet stacked against 4 α-helices (See Figure 24). The dimerisation helix and the catalytic histidine residue (H239) could not be modelled as a result of poor electron density (Srivastava et al, 2014).

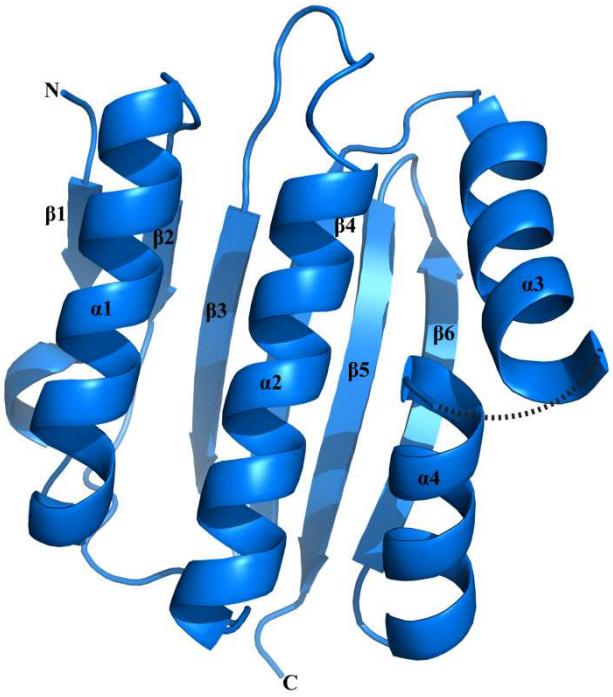


Figure 24: Crystal structure of the AgrC CA domain.

Ribbon representation of the CA domain of AgrC crystallised in presence of ATP depicted in blue. The 6 β-sheets, the 4 α-helix and the N- and C-terminus are labelled in black. The black dotted line represents the residues that could not be modelled onto the structure, aa 389-393. Pymol, PDB 4BXI.

The structure of the CA domain of AgrC is essentially similar to those of kinases from the same family, apart from a few key differences. In AgrC the β1-sheet is part of a β-sheet with the 5 other strands whereas is other related kinases structure this sheet forms a linker with the DHp domain. This β-sheet also assumes the role of covering the nucleotide binding pocket and is called the ATP-lid. In AgrC it is shorter than in other structures. In the AgrC structure, α-helices α3 and α4 are at the entrance of the active site in a conformation that appears to block it. This conformation is different than that observed with structures of related kinases bound to ATP (Srivastava et al, 2014).

In the same study, the timing of the autophosphorylation of the AgrC CA domain was investigated by means of a time-course experiment. The results revealed that AgrC CA domain achieves autophosphorylation *in vitro* after 5 min and is saturated within 20 min at 25°C. It was observed that the phosphorylated AgrC CA domain is stable for 1h at 25°C (Srivastava et al, 2014).

1.5.5. AgrA

AgrA is the RR of the *agr* operon and is responsible for activating transcription at both the P2 and the P3 promoter when activated. The AgrC HK is responsible for the activation of AgrA. AgrA consists of an N-terminal CheY-like receptor domain (aa 1-136) and a C-terminal LytTR DNA-binding domain (137-238) (Novick & Geisinger, 2008).

AgrA functions as a dimer and the dimerisation of AgrA is activation dependent. The incubation of AgrA with acetyl phosphate converts the AgrA monomers into dimers (Srivastava et al, 2014).

1.5.5.1. The AgrA CheY-like receptor domain

The structure of the AgrA CheY-like receiver domain has not been resolved but can be predicted based on the crystal structure of the CheY RR. Similarly, the changes taking place in the AgrA receptor domain upon phosphorylation have not been characterised. CheY consists of a stand-alone receiver domain from *S*almonella typhimurium and *E. coli*. CheY contains five parallel β-sheets surrounded by five α-helices (See Figure 25). Five residues in CheY have been identified as important for phosphorylation: D12, D13, D57, T87 and K109. Residues D12, D13 and D57 have carboxylated side chains that form a conserved acidic pocket in CheY (D8, D9 and D59 in AgrA) (Lukat et al, 1990). This acidic pocket has been

shown to be important for the binding Mg^{2+} ions that are essential for the phosphorylation of CheY (Lukat et al, 1991; Lukat et al, 1990). In addition, D57 is the phosphor-acceptor site of CheY (Sanders et al, 1989). Residues D12 and D13 and D57 bind to the Mg^{2+} ions and residues T87 and K109 are important for an event taking place after the phosphorylation event, potentially a conformational change essential to reach the active state or for the stabilisation of the active state. Another residue, Y106 is thought to be involved in signal transduction via a rearrangement of its side chain (Appleby & Bourret, 1998; Lee et al, 2001; Lukat et al, 1991; Stock et al, 1993; Zhu et al, 1996)

Several crystal structures have been resolved for CheY: with and without Mg^{2+} bound (See Figure 25) and activated. The coordination of the metal cation induces conformational changes in the side chains of aspartate 57 and lysine 109 residues which result in the abolition of the salt bridge interaction present between these two residues in the structure without magnesium. This liberates the aspartate 57 for the phosphoryl transfer reaction. In the activated structure (with beryllium fluoride), the beryllium fluoride is bound to D57 and forms a hydrogen bond with T87 and a salt bridge with K109. The hydroxyl group of T87 forms a hydrogen bond with an active site acceptor (probably activate D57) and the aromatic side chain of Y106 is buried in a hydrophobic pocket. The activated structure revealed that only modest changes were taking place upon phosphorylation of CheY and were mainly localised to one region known to be flexible in inactive CheY, the α4 helix and the β4-sheet (See Figure 25) (Lee et al, 2001; Stock et al, 1993).

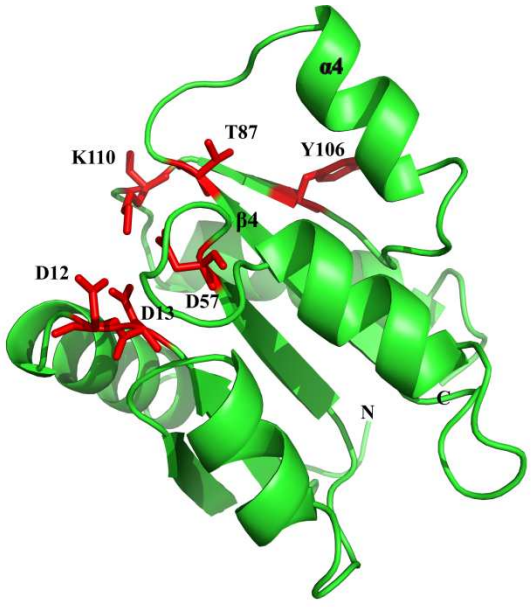


Figure 25: Crystal structure of the inactive CheY.

Ribbon representation of the inactive CheY from *Salmonella typhimurium* CheY is depicted in green. The 5 residues potentially involved in phosphorylation, D12, D13, D57, K110 and the residue involved in signal

transduction Y106 are coloured in red and labelled in black. The N- and C-terminus and the α4 region and β4 region where most of changes upon phosphorylation take place are labelled. Pymol, PDB 2CHF.

1.5.5.2. The AgrA LytTR domain

The crystal structure of the isolated AgrA LytTR domain bound to a 15 bp DNA fragment containing the 9-bp consensus binding sequence (based on the P2 binding sequence) was reported in 2008 and consists of a 10 stranded β-fold (See Figure 26). A number of residues involved in contacting the DNA and maintaining the structural integrity of the protein were identified. The binding of AgrA to the promoter DNA induced a substantial bend in the DNA (38°) as it conforms to the AgrA surface (Sidote et al, 2008). The LytTR domain of AgrA will be discussed in more detail in section 4.1.



Figure 26: Crystal structure of the AgrA LytTR domain.

Ribbon representation of the AgrA LytTR domain-DNA complex crystal structure (the LytTR domain and DNA are coloured in cyan and grey, respectively). The N-terminus and C-terminus are indicated. Pymol, PDB 3BS1.

1.5.5.3. AgrA and transcription activation at the P2 and P3 promoters

AgrA binds to a pair of direct repeats located in the *agr* intergenic region (*agr*IR) separated by 12 bp at the P2 and the P3 promoter and activates the transcription of both promoters. The activation of AgrA with acetyl phosphate, a small phosphodonor, enhances its binding to both promoters demonstrating that AgrA needs to be activated in order to activate transcription. AgrA has a higher affinity for the P2 direct repeats due to a 2 bp difference in the sequence of the repeats. At position 1 and 2 in the downstream repeat, the P2 promoter has an AC and the P3 promoter a CT. Mutation of the CT motif in the P3 repeat to AC lead to a higher affinity of AgrA for the mutated P3 than P2. The area protected by AgrA upon binding (footprint) overlaps the -35 promoter element at both the P2 and the P3 promoter (See Figure 27) (Koenig et al, 2004). The consensus binding sequence for LDCTFs has been identified as [TA][AC][CA]GTTN[AG][TG] (Nikolskaya & Galperin, 2002).

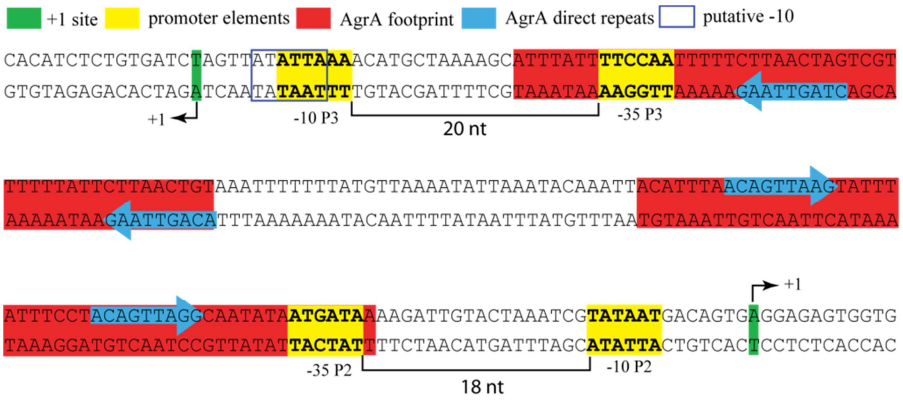


Figure 27: AgrA binding sites and footprint at the P2 and P3 promoters.

DNA sequence of the *agr* IR. -35 refers to the -35 element (5'-TTGACA-3') and -10 refers to the -10 element (5'-TATAAT-3').The nucleotide length of the spacer between the -10 and -35 promoter elements is indicated for the P2 and the P3 promoter. The published -10 and -35 promoter elements are highlighted in yellow and the putative -10 of the P3 promoter is boxed in blue. Adapted from (Koenig et al, 2004).

The position of the -10 and -35 promoter elements of the P2 and the P3 promoter were inferred based on the 5' end of the RNAII and RNAII transcripts, respectively. The 5' end of RNAII and RNAIII were mapped using primer extension and nuclease protection assay (Janzon et al, 1989; Novick et al, 1993). Thus the published position of the -10 and -35 promoter element for both P2 and P3 promoters are indicated in yellow in Figure 27. Position -11 of the P3 promoter differs from the consensus -10 sequence although it is one of the mostly highly conserved position of this promoter element (See sections 1.1.1 and 1.1.3) (Feklistov & Darst, 2011; Zhang et al, 2012). However, the study of the sequence adjacent to the -10 of the P3 promoter revealed that if the -10 (5'-TTTAAT-3') was shifted by two nt downstream, the resulting new -10 element (5'- TAATAT -3') would have position -11 conserved (See Figure 27 and Figure 28). Thus it is possible that the published position for the -10 of the P3 promoter is incorrect.

Consensus	-10	-12-11-10-9-8-7 5'-T A T A A T-3'		**P3 published**	-10	-12-11-10-9-8-7 5'-T T T A A T-3'	5/6	-15-14 TG
	-35	-35-34-33-32-31-30 5'-T T G A C A-3'			-35	-35-34-33-32-31-30 5'-T T G G A A-3'	4/6	
P2	-10	-12-11-10-9-8-7 5'-T A T A A T-3'	6/6	**P3 putative**	-10	-12-11-10-9-8-7 5'-T A A T A T-3'	4/6	-15-14 TG
	-35	-35-34-33-32-31-30 5'-A T G A T A-3'	4/6		-35	-35-34-33-32-31-30 5'-T T G G A A-3'	4/6	

Figure 28: Conservation of the -10 and -35 promoter elements of the P2 and P3 promoters.

Conservation of the -10 and -35 promoter elements of the P2 and P3 promoter (the published one and the putative one illustrated in Figure 27) compared to the consensus promoter elements. -35 refers to the -35 element (5'-TTGACA-3') and -10 refers to the -10 element (5'-TATAAT-3'). The red numbers above the sequence correspond to the position of each base with regards to the transcriptional start site. The nucleotides that can vary from the consensus sequence without affecting promoter activity drastically and are thus less

conserved are underlined in black. The nucleotides of the P2 or P3 promoter that differ from the consensus sequence are highlighted in yellow. The extended -10 TG motif is indicated for the P3 promoter.

The spacer length at the P2 and the P3 promoter differs from the optimal spacer length of 17 nucleotides (nt) (Aoyama et al, 1983). More significantly at the P3 promoter where the spacer is 20 nt long and slightly at the P2 promoter where the spacer is 18 nt long. Thus the P3 promoter appears to be a weaker promoter than the P2. In addition, the -10 promoter element of the P3 promoter has a weaker conservation (5 out of 6) compared to the consensus than the -10 element of the P2 promoter (6 out of 6) (See Figure 28). However, P3 has an extended -10 TG motif at position -15 and -14 (respectively) which has been shown to compensate for the sub-optimally placed -35 promoter elements (Burr et al, 2000). Indeed the extended -10 promoter element was shown to be essential for optimal promoter activity (Reynolds & Wigneshweraraj, 2011). *In vivo* and *in vitro* studies have shown that shortening the P3 promoter to the optimal length of 17 nt significantly increases its activity and reduces the effect of the activation by AgrA (Morfeldt et al, 1996; Reynolds & Wigneshweraraj, 2011). These findings and the fact that AgrA significantly bends the DNA led to the hypothesis that the bending activity of AgrA is responsible for its activation function. The binding of AgrA would shorten the promoter spacer length to the optimal length and thereby increase transcription by facilitating open and closed complex formation (RPo and RPc) (Reyes et al, 2011; Sidote et al, 2008).

The RNA polymerase (RNAp) can bind to the P2 and the P3 promoter equally well in the absence of AgrA and initiate transcription from both promoters. However, open complexes (RPo) formed at the P3 promoter are less stable than the P2 ones and the level of basal transcription (without AgrA) is higher at the P2 promoter (~20-fold) than at the P3 promoter. However, the increase of transcription observed in presence of AgrA is significant at the P3 promoter (~15 fold) but marginal at the P2 promoter (~2-fold). AgrA was also shown to increase promoter occupancy by the RNAp at both P2 and P3 promoters (Reynolds & Wigneshweraraj, 2011). These findings are in concordance with a model where AgrA activates transcription via bending the DNA which would have a more significant at the P3 promoter because of its longer spacer length. Moreover, the increased promoter occupancy observed in presence of AgrA suggests that AgrA could also be involved in the recruitment of the RNAp. This theory is reinforced by the fact that the AgrA footprint at the P2 and the P3 promoter significantly overlaps with the -35 promoter element of both promoters. Therefore the *agr* promoters resemble a class-II-like promoter (Reynolds & Wigneshweraraj, 2011). In class II promoters, the TF facilitates the binding of the RNAp to the promoter to

increase RPc formation by recruiting the RNAp and is some cases facilitate the conversion from RPc to RPo (See Figure 9) (Lee et al, 2012).

1.5.5.4. AgrA regulation outside of the *agr* operon

AgrA directly upregulates the expression of phenol-soluble modulins PSMαs and PSMβs by directly binding to their promoter region. There are 7 PSM peptides in *S. aureus*, four are part of the PSMα operon, three of the PSMβ operon and the last is δ-haemolysin which is encoded by RNAIII. Both PSM operons are under RNAIII-independent activation via AgrA. The AgrA foototprint overlaps with the -10 and -35 promoter elements of the PSMα promoter (Queck et al, 2008).

AgrA is also responsible for the repression on the *bsaA* gene, an essential gene for survival under oxidative stress. Upon oxidative stress, a disulphide bond is formed between two cysteine residues in the AgrA LytTR domain (C199 and C228) which leads to the dissociation of AgrA from the *bsaA* promoter and in turn the activation of that promoter. When one of the cysteine is mutated into a serine (C199S) the disulphide bond is not formed and AgrA inhibits the *bsaA* gene even under oxidative stress leading to an increase susceptibility to oxidative stress (See Figure 29). This function illustrates that there is a link between the quorum sensing function of the *agr* operon and oxidation sensing and that link is AgrA. It is also the first case of AgrA acting as a repressor (Sun et al, 2012b).

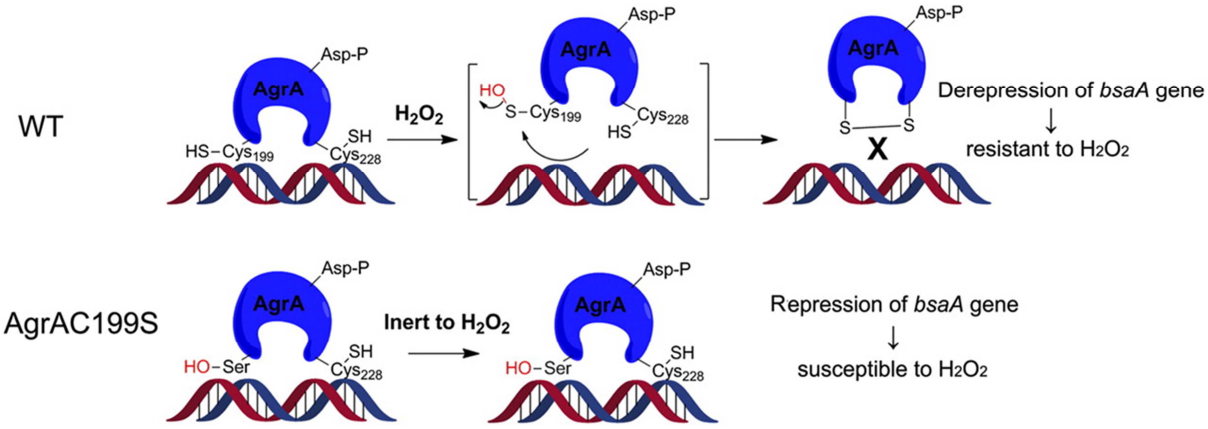


Figure 29: Model of oxidation sensing via AgrA.

Under WT conditions, the oxidation of AgrA leads to a disulphide bond formation between C199 and C228 and the dissociation of AgrA from the DNA. With the C199S mutation, the disulphide bond cannot be formed and therefore AgrA stays on the DNA preventing the activation of the *bsaA* promoter. This blockage of the *bsaA* gene renders the mutated strain susceptible to H_20_2 (Sun et al, 2012b).

1.5.5.5. Interactions of AgrA with AgrC

In vitro analysis of the phospho-transfer between AgrA and AgrC using different recombinant proteins lacking one of the domains revealed that both the DHp and the CA

domain of AgrC are required for efficient phosphotransfer to AgrA and that the AgrA LytTR domain is not required for the phosphotransfer. The phosphotransfer reaction revealed to be very fast with 50% of phosphoryl group transferred from AgrC to AgrA within a min. These experiments were performed in presence of a 3-fold excess of AgrA. Phosphotransfer between AgrC and two non-cognate RRs proved to be poor supporting the fact that AgrC only phosphorylates AgrA (Srivastava et al, 2014).

The crystal structure of the AgrC CA domain was used to construct a mechanistic model of the interaction between this domain and the receiver domain of AgrA (See Figure 30). The AgrC catalytic domain dimer was modelled based on the dimeric structure of a related kinase, HK853 from *Thermotoga maritima* and the AgrA receiver domain based on the CheB RR from *Salmonella enterica*. The AgrA receiver domains were then modelled onto the predicted AgrC-CA dimer by *in silico* docking. In this model, the AgrA receiver domains are stacked on the dimerisation helices of AgrC (Srivastava et al, 2014).

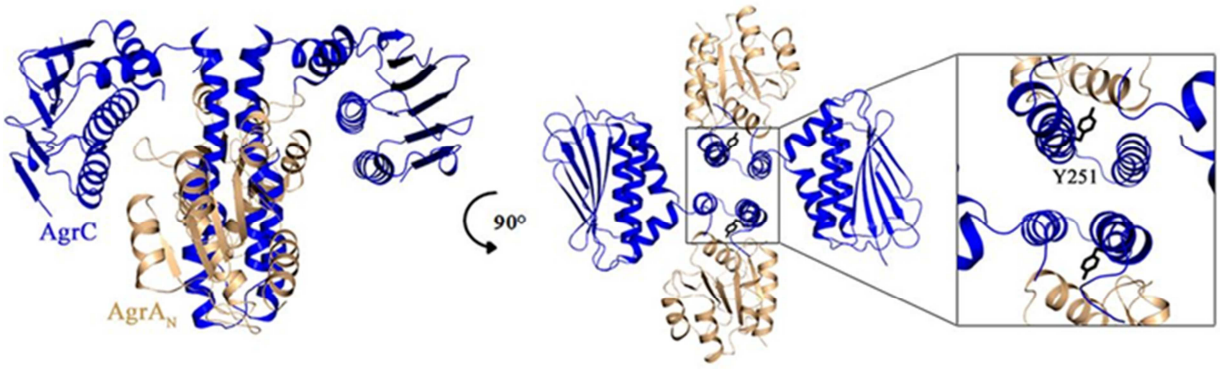


Figure 30: Model of the interaction between the AgrC Ca-DHp dimer and a dimer of AgrA receiver domain.

Two different perspectives are shown as well as a zoom on the direct interacting region (Srivastava et al, 2014).

The complex model revealed that the cognate/non-cognate discrimination is probably driven by specific sequence and structural features at the interface between the HK and the RR. Several residues involved in specifics contact were identified in the AgrA receiver domain and in AgrC. The interacting segment in the AgrA receiver domain contains residues located on both sides of the putative phosphorylation site D59. Further experiments to characterise the interaction between AgrA and AgrC revealed that their interaction is influenced by the presence of ATP (Srivastava et al, 2014).

1.5.6. RNAIII

RNAIII is a 514 nt RNA and the effector of the *agr* operon (Novick et al, 1993). It has a complex secondary structure that is well conserved among staphylococci although the

sequence is not, and consists of 14 hairpin structures connected by unpaired nucleotides (See Figure 31) (Benito et al, 2000). RNAIII contains three sites of long-range interaction where the distant nucleotides come together and form helices. They divide RNAIII into three domains. The 5' domain contains the first 31 nt and hairpin 1. The central domain contains nt 32-382, hairpins 2-11. It encodes δ-haemolysin (See Figure 31) (Janzon et al, 1989), contains most of the regions responsible for the upregulation of α-haemolysin (in green in Figure 31) and contains hairpin 7 which is involved in the negative post-translational regulation of cell surface-associated virulence factors and Rot. The 3' domain contains nt 383 to 514, hairpins 12-14 and is mainly involved in the translational inhibition of cell surface-associated virulence factors and the *spa* gene (protein A) (Benito et al, 2000; Thoendel et al, 2011).

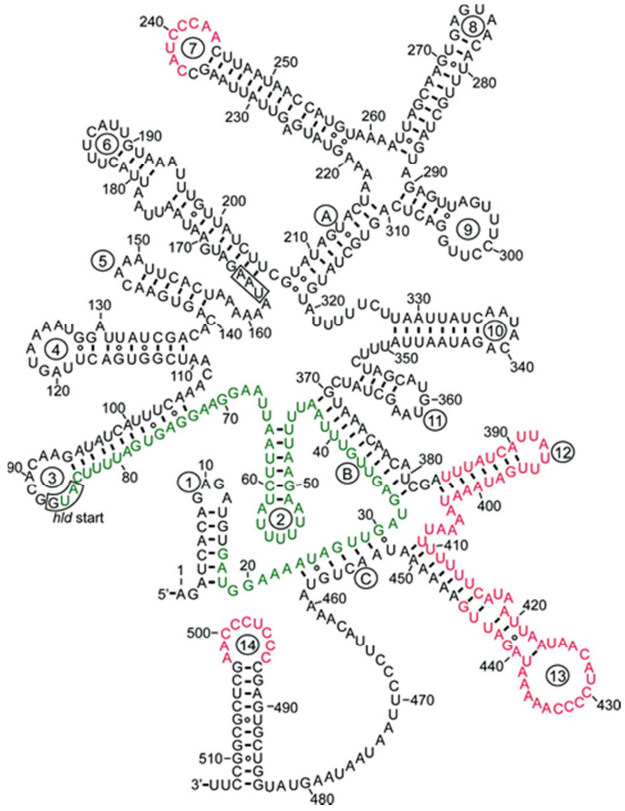


Figure 31: Secondary structure of RNAIII.

The stemloops are numbered and the letters indicate long-range interactions that establish domains. The start and end codon of hld are boxed and labelled. The green letters represent the nucleotides that basepair with and activated *hla* translation. The nucleotides involved in translation inhibition are coloured in red (Thoendel et al, 2011).

The mechanism of the translation of δ-haemolysin is unclear. Its Shine-Dalgarno (SD) sequence is located at the base of hairpin 3 which suggest that the ribosomes might not recognize it in the secondary structure. However, ribosomes have been shown to be able to bind to native RNAIII (Benito et al, 2000). The δ-haemolysin SD and start codon are also

part of the region involved in the upregulation of α-haemolysin via base pairing (Balaban & Novick, 1995).

The half-life of RNAIII is greater than 45 min, which is considered long (Benito et al, 2000; Huntzinger et al, 2005). The expression of secreted virulence factors is upregulated at the post-transcriptional level by countering a translation-blocking secondary structure in their mRNA leader that hides the SD from the ribosomes (Morfeldt et al, 1995). The repression of the cell surface-associated virulence factors (such as protein A) occurs also at the translational level but by a different mechanism involving the pairing of RNAIII with the translation-initiation regions of their transcript (Novick et al, 1993). This base pairing hides the SD sequence from the ribosomes and therefore prevents translation (Thoendel et al, 2011). This post-transcriptional repression is irreversible as it also involves a cleavage event, by the endoribonuclease RNAseIII that binds to RNAIII (an endoribonuclease that cleaves double-stranded RNAs (Thoendel et al, 2011)) (Geisinger et al, 2006). The cleavage by RNAseIII results in the destabilisation of the mRNA and sometimes the removal of the SD (Huntzinger et al, 2005). The same mechanism of repression is used by RNAIII to inhibit the major transcription factor Rot, a fibrinogen binding protein and coagulase (Boisset et al, 2007; Chevalier et al, 2010; Geisinger et al, 2006). The repression of Rot by RNAIII is responsible for most of the RNAIII downregulated and upregulated genes as Rot upregulates the translation of cell-associated virulence factors and downregulates the translation of secreted virulence factors (Geisinger et al, 2006; Said-Salim et al, 2003). The targets of Rot downregulation include the haemolysins α, β and γ, extracellular proteases, urease, lipase and enterotoxin B. The genes upregulated by Rot include protein A, coagulases and clumping factor B. There is a significant overlap between *agr* regulated and Rot regulated genes (Said-Salim et al, 2003). It should be noted that these effects are not entirely due to Rot as the overlap of targets is not perfect and RNAIII has direct targets as well, as described above. For a list of TCS regulated by RNAIII refer to section 1.4.4.16.

1.5.7. The regulation of the *agr* operon

The regulation of the *agr* operon is a very complex process and involves large number of regulators. It is not in the scope of this section to discuss all of them in detail. Instead I will give a brief overview of the main factors acting to regulate *agr*.

The regulation of *agr* includes RRs and transcription factors (TFs) (See Table 5) and environmental cues such as pH, and glucose (via CcpA). (Regassa et al, 1992; Seidl et al, 2006).

Regulator	Effect on *agr*
SarA	+
SarR	-
SarU	+
SarZ	+
MgrA	+
CcpA	+
ArlSR	+
SarX	-
SarT	-
Rsr	-
SrrAB	-
σ^B	-
CodY	-

Table 5: Transcriptional regulators involved in the regulation of the *agr* operon.

List of the transcriptional regulators involved in the regulation of the *agr* operon and whether they activate (+) or inhibit (-) it.

For more information about ArlSR and SrrBA, refer to sections 1.4.4.6 and 1.4.4.7.

1.5.7.1. SarA

The Staphylococcal accessory regulator (SarA) is part of a large family of TFs (the SarA family) involved in the regulation of virulence genes in *S. aureus* and was the first member characterised (Cheung et al, 1992). SarA is a global regulator of transcription and is involved in the regulation many genes (over 100) involved in a variety of processes. SarA activates transcription at the *agr* operon through direct binding to the *agr* IR (Rechtin et al, 1999). SarA also represses cell surface-associated virulence factors while upregulating secreted virulence factors, like *agr*. Some of this regulation is carried out via *agr* but SarA has also been shown to regulate certain virulence factors directly by binding to their promoter region. Many studies have been conducted to determine the overlap and differences between the *agr*-mediated effects of SarA from its direct effects (Dunman et al, 2001). Interestingly, the SarA binding site at the *agr* IR overlaps with the AgrA footprint at the P2 promoter (See Figure 32) (Koenig et al, 2004; Sterba et al, 2003). Furthermore, it has been shown that SarA is required for maximum transcription from the P2 promoter (but not the P3 promoter). SarA bends the DNA upon binding, and is therefore thought to activate transcription in a mechanism similar to the one proposed for AgrA (Reyes et al, 2011).

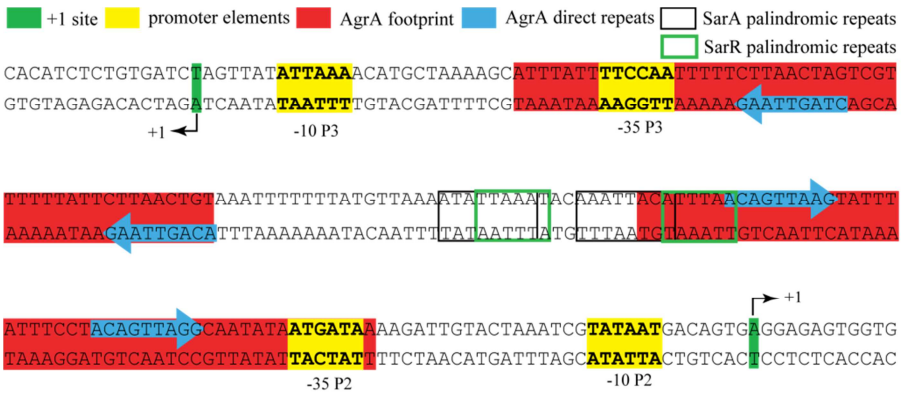


Figure 32: AgrA binding sites and footprint, SarA and SarR binding sites at the P2 and P3 promoters

DNA sequence of the *agr* IR. -35 refers to the -35 element (5'-TTGACA-3') and -10 refers to the -10 element (5'-TATAAT-3'). The published -10 and -35 promoter elements are highlighted in yellow, the SarA palindromic repeats are boxed in black and the SarR palindromic repeats are boxed in green. Adapted from (Koenig et al, 2004).

1.5.7.2. SarR

SarR is a member of the SarA family and the negative regulator of SarA. SarR regulates SarA by binding to its promoter region (Manna & Cheung, 2001). SarR binds to the *agr* IR and its binding sites are located in the vicinity of the SarA binding sites (See Figure 32) (Manna & Cheung, 2006b). SarR is a repressor of the P2 promoter and competes for binding at the *agr* IR. SarR binds more avidly than SarA and thus can displace SarA and repress the P2 promoter (but not the P3) in presence of SarA and AgrA. As both AgrA and SarA bend DNA but not SarR it is though that the inhibition by SarR involves preventing the DNA from bending (Reyes et al, 2011).

1.5.7.3. SarU

SarU is a SarA homolog and a positive regulator of the *agr* operon (Manna & Cheung, 2003).

1.5.7.4. SarZ

SarZ is a SarA paralog and is activated by MgrA. SarZ is involved in virulence gene regulation and is an activator of the *agr* operon (Ballal et al, 2009; Kaito et al, 2006).

1.5.7.5. MgrA

MgrA is also called Rat and is a TF related to the SarA family and a member of the MarR family. It is involved in the regulation of cell wall turnover, autolysis and the upregulation of secreted toxins and proteases and is a global regulator of virulence genes expression. MgrA also positively regulates the TCS LytSR and ArlSR (See section 1.4.4). MgrA positively a*gr* by binding to the *agr* region and also directly upregulates α-haemolysin by binding to its promoter (Ingavale et al, 2005; Ingavale et al, 2003; Luong et al, 2003).

1.5.7.6. CcpA

CcpA, the carbon catabolite protein is a TF that activates *agr* in the presence of glucose. CcpA regulates the carbon metabolism but is also a major regulator of virulence gene expression via its effect on *agr* (Seidl et al, 2006).

1.5.7.7. SarX

SarX is a TF member of the SarA family. SarX is activated by MgrA. SarX is a negative regulator of *agr* and therefore indirectly act on the genes regulated by a*gr*. SarX regulates *agr* by binding to the *agr* IR (Manna & Cheung, 2006a).

1.5.7.8. SarT

SarT is a SarA homolog that is repressed by SarA and *agr*. SarT represses α-haemolysin and *agr*, most-likely by repressing SarU which is an activator of *agr* (Manna & Cheung, 2003; Schmidt et al, 2001).

1.5.7.9. RsR

RsR is a repressor of the *agr* operon. The mechanism of repression of the *agr* by RsR is unclear as it does not bind to the *agr* IR (Tamber et al, 2010).

1.5.7.10. σ^B

σ^B is an alternative σ factor involved in extracellular stress response. σ^B is also responsible for the activation of surface proteins, pigment production and the downregulation of secreted toxins and proteases. This regulation has been shown to be *agr* mediated as σ^B is a repressor of the *agr* operon and activates SarA. This regulation was uncovered because *S. aureus* strains from the NCTC 8325 exhibited an increased *agr* activity that was shown to be due to a mutation inactivating σ^B. The exact mechanism of σ^B regulation of *agr* is unclear (Bischoff et al, 2001; Kullik et al, 1998).

1.5.7.11. CodY

CodY is a TF involved in the adaptation to nutrient stress and nutrient limitation. It is also a regulator of virulence factors in response to nutrient availability and represses biofilm formation. CodY is a sensor of intracellular concentrations of isoleucine, leucine, valine and GTP. CodY negatively regulates *agr* when the nutrients levels are high. When the nutrients are scarce, CodY does not affect *agr*. CodY binds only weakly to the *agr* IR and the mechanism of repression by CodY at the *agr* operon is unknown. CodY regulation appears to affect differently the P2 and the P3 promoter. It has been shown that in a CodY mutant *agr* is

overexpressed and that this overexpression is dependent on phosphorylated AgrA and AgrC and results in an increased production of AIP. SarA is not involved in the CodY dependent regulation of *agr* (Majerczyk et al, 2008; Roux et al, 2014; Thoendel et al, 2011).

1.6. Objectives

LytTR domain-containing transcription factors (LDCTFs) are involved in the transcriptional regulation of virulence genes in a number of important human pathogens. However, the mechanism of this regulation remains unclear. The most extensively studied LDCTF AgrA, a key player of *S. aureus* virulence, has been investigated for nearly 30 years yet the precise mechanism of transcription activation via AgrA remains cryptic. The chief objective of this thesis is to elucidate the individual contributions of aa residues in the LytTR domain of LDCTFs to transcription activation using the *S. aureus* AgrA as a model. In this regard, the specific aims include:

I. The development of an experimental system to study AgrA *in vivo.*
II. The construction, screening and characterisation of a library of single alanine mutants in the AgrA LytTR domain.
III. Further characterisation of putative transcription activation deficient (TAD) mutants identified using *in vivo* and *in vitro* techniques.

CHAPTER TWO

2. Materials and methods

2.1. General Methods

2.1.1. Composition of growth media

All of the bacterial growth media were autoclaved for sterilisation at 120°C for 25 minutes (min). Table 6 contains the different media, when they were used and their composition.

Media	Composition (1L)	Use
Lysogeny Broth (Luria-Bertani, LB) (Sigma)	10 g Tryptone, 5 g Yeast Extract, 5 g NaCl	*E. coli* growth media
Lysogeny Agar (Luria-Bertani Agar, LA) (Sigma)	10 g Tryptone, 5 g Yeast extract, 5 g Sodium chloride, 10 g Agar	*E. coli* growth media
Tryptic Soy Broth (TSB) (BD)	17 g Bacto™ Tryptone (Pancreatic Digest of Casein), 3 g BactoSoytone (Peptic Digest of Soybean Meal), 2.5 g Glucose (= Dextrose), 5 g Sodium Chloride, 2.5 g Dipotassium Hydrogen Phosphate	*S. aureus* growth media
Tryptic Soy Agar (TSA) (BD)	15 g Bacto™ Tryptone (Pancreatic Digest of Casein), 5 g BactoSoytone (Papaic Digest of Soybean Meal), 5 g Sodium Chloride, 15 g Agar	*S. aureus* growth media
Super Optimal Broth with Catabolite repression (SOC) (BD)	20 g Tryptone, 5 g Yeast Extract, 0.5 g Sodium Chloride, 2.4 g Magnesium Sulfate (anhydrous), 186 mg Potassium Chloride, 20 mM glucose	Heat-shock transformation
Brain Heart Infusion Broth with Sucrose (BHI + sucrose) (Oxoid)	12.5 g Brain infusion solids, 5 g Beef heart infusion solids, 10 g Proteose peptone, 2 g Glucose, 5 g Sodium chloride, 2.5 g Disodium phosphate, 0.5 M Sucrose	Electroporation
Columbia Blood Agar (Fluka)	23 g Special nutrient substrate, 5 g Sodium Chloride, 1 g Starch, 15 g Agar, 5% sheep blood (Sigma)	Blood Agar plate haemolysis assay

Table 6: Composition of growth media.

The name, composition for 1L and use is listed for each media.

When appropriate, the sterile growth media were supplemented with antibiotics in order to select for the bacterial strain or plasmid of interest. The antibiotics were filter sterilised before use and are listed in Table 7. All agar plates and liquid cultures were incubated at 37°C, and the liquid cultures were then shaken at 180 rpm in a shaking incubator, unless stated otherwise. *E. coli* cultures were grown in LB or LA and *S. aureus* cultures in TSB or TSA, unless otherwise stated.

Antibiotic	Stock concentration	Final Concentration	Storage	Abbreviation
Ampicillin	100 mg/ml, in ethanol	100 µg/ml	-20°C	Amp^R
Kanamycin	50 mg/ml, in H_2O or 90 mg/ml, in H_2O	50 µg/ml or 90 µg/ml	-20°C	Kan^R
Chloramphenicol	7.5 mg/ml, in ethanol	7.5 µg/ml	-20°C	Cm^R
Anhydrotetracycline	200 µg/ml, in ethanol	200 ng/ml	-20°C	Atet

Table 7: Summary of the antibiotics used in this work.

The name, stock concentration, final concentration, storage condition and abbreviation are listed for each antibiotic. Kanamycin was used at 50 µg/ml to select for *E. coli* and at 90 µg/ml to select for *S. aureus*.

2.1.2. GFP growth curves using a multiwall plate reader

Seed cultures were grown at 37 °C, shaken at 700 rpm for 17 hours (h) in a THERMOstar (BMG Labtech) plate incubator by directly inoculating a colony from a freshly transformed TSA plate into 200 µl of TSB medium supplemented with the appropriate antibiotics in a 96-well black microtitre plates with clear bottoms (Corning). The experimental growth curves were also performed in 96-well microtitre plates in a POLARstar Omega multiwell plate reader (BMG Labtech). The seed cultures were diluted 1 : 50 in a final volume of 200 µl of fresh TSB medium supplemented with the appropriate antibiotics and incubated at 37 °C, shaken at 700 rpm for 17 h. Simultaneous growth ($OD_{600\ nm}$) and GFP fluorescence measurements (with excitation and emission filters of 485 and 520 nm respectively) were performed. At least three biological replicates (each defined as a single colony) were performed for each growth curve.

2.1.3. Blood agar plate heamolysis assay

The *S. aureus* strains to be tested were streaked out on fresh TSA plates complemented with the appropriate antibiotics and grown overnight. TSB seed cultures were inoculated and grown for 17 h, and then 100 µl aliquots were taken and pelleted. The cell pellet was resuspended in 20 µl TSB and doted onto Columbia Agar containing 5% sheep's blood and left to grow for 17 h at 37°C and then 24 h at 4°C prior to image capture of the plates using a Gel Doc^{TM} EZ system (BIO-RAD).

2.1.4. General lab reagents and solutions

Acrylamide proto gel (Solution I) 30% (w/v) acrylamide, 0.8% (w/v)

(National diagnostics)	bisacrylamide stock solution (37.5: 1)
Acrylamide Solution II	1.5 M Tris-HCl, pH 8.8, 0.3% (w/v) SDS
Acrylamide Solution III	0.5 M Tris-HCl, pH 6.8, 0.4% (w/v) SDS
DNA loading dye (6x) (Fermentas)	10 mM Tris-HCl (pH 7.6), 0.03% (w/v) bromophenol blue, 0.03% (w/v) xylene, 60% (v/v) glycerol, 60 mM EDTA.
Protein Kinase A buffer (PKA) (2x)	80 mM Tris-HCl (pH 6.4), 400 mM NaCl, 50 mM $MgCl_2$, 4 mM DTT, 50% (v/v) glycerol
Phosphate Buffered Saline (PBS) (10x)	80 g NaCl, 2 g KCl, 11.5 g Na_2HPO_4•7 H_20, 2 g KH_2PO_4 in 1 l dH_2O, pH 7.4
Tris Magnesium buffer (TM) (10x)	10 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$
Transblot buffer (20x)	24.2 g Tris, 150 g glycine, dH_2O to 1 L.
Transblot buffer (1x)	20ml 20x transblot buffer, 40 ml methanol, dH_2O to 400 ml.
Tris Borate EDTA (TBE) (10x) (Lonza)	0.89 M Tris-borate, 0.02 M EDTA, pH 8.3
Tris Buffered Saline (TBS) (10x)	50 mM Tris, 150 mM NaCl, pH 7.6
TBS Tween 20 (TBSTT) (10x)	TBS with 0.1% (v/v) Tween 20
Tris Glycine (TG) (10x) (National Diagnostics)	0.25 M Tris base and 1.92 M glycine
Tris Glycine SDS (TGS) (10x) (w/v)(National Diagnostics)	0.25 M Tris base, 1.92 M glycine, 1% SDS

Tris-HCl stock (1M)	121 g Tris base in 800 ml water, adjust to desired pH with concentrated HCl. (Mix and make up to 1L)
Urea gel system concentrate (National Diagnostics)	237.5 g acrylamide, 12.5 g methylene bis-acrylamide, 7.5 M urea, dH_2O to 1 L.
Urea gel system diluent (National Diagnostics)	7.5 M urea.

2.1.5. Bacterial strains and plasmids

A list of all the bacterial strains used in this work can be found in Table 8, all the plasmids used in this work can be found in Table 9 and all the plasmids constructed in this work can be found in Table 14 (See section 8.1.2).

Strain	Genotype	Reference
E. coli XL1 blue	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F´ proAB lacIqZΔM15 Tn10 (Tetr)]	Stratagene
E. coli BL21 (DE3)	F^-ompT gal dcm lon hsd $S_B(r_B^-m_B^-)$ λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])	Stratagene
E. coli DC10B	Δdcm dam^+ Δhsd RMS endA1 recA1	(Monk et al, 2012)
E. coli ER2566	fhuA2 lacZ::T7 gene1 [lon] ompT gal sulA11 R(mcr-73::miniTn10--TetS)2 [dcm] R(zgb-210::Tn10--TetS) endA1 Δ(mcrC-mrr)114::IS10	New England Biolabs
S. aureus SH1000	Functional *rsbU* derivative of 8325-4 *rsbU+*	(Horsburgh et al, 2002)
S. aureus $SH1000^-$	*agr*-deficient SH1000 because of mutation H174L in *agrA*	(Tsompanidou et al, 2011)
S. aureus SH1001	SH1000 *agr*::*tet*	(Horsburgh et al, 2002)
S. aureus RN4220	MNNG mutagenized derivative of RN450, accepts foreign DNA	(Peng et al, 1988)
S. aureus $SH1000_{agr\ IR\ P3\text{-}GFP}$	$pCL55_{agr\ IR\ P3\text{-}GFP}$ integrated in SH1000	(James et al, 2013)
S. aureus $SH1000^-_{agr\ IR\ P3\text{-}GFP}$	$pCL55_{agr\ IR\ P3\text{-}GFP}$ integrated in $SH1000^-$	This study
S. aureus $SH1001_{agr\ IR\ P3\text{-}GFP}$	$pCL55_{agr\ IR\ P3\text{-}GFP}$ integrated in SH1001	(James et al, 2013)

Table 8: Bacterial strains used in this work.

The name, genotype and reference are listed for each strain.

Vector	Resistance	Use	Inducer	Reference
pJR28-[6His]*agrA*	Kan^R	Protein Expression	IPTG	(Reynolds & Wigneshweraraj, 2011)
pJR28-[6His]*rpoD*	Kan^R	Protein Expression	IPTG	(Reynolds & Wigneshweraraj, 2011)
pGEM-T Easy	Amp^R	Cloning	N/A	Promega
pJR_{P2+P3}	Amp^R	Promoter template for *in vitro* transcription	N/A	(Reynolds & Wigneshweraraj, 2011)
pCN34	Amp^R in *E. coli*, Kan^R in *S. aureus*	Expression in *S. aureus*	N/A	(Charpentier et al, 2004)
pCN34*itet*	Amp^R in *E. coli*, Kan^R in *S. aureus*	Expression in *S. aureus*	Atet	(Corrigan et al, 2011)
pCN44	Amp^R in *E. coli*, Erm^R in *S. aureus*	Cloning of TT	N/A	(Charpentier et al, 2004)
pCL55 *agr* IR P3-GFP	Amp^R in *E. coli*, Cm^R in *S. aureus*	Cloning of P2-*agr* IR	N/A	(James et al, 2013)
pTYB2	Amp^R	Protein Expression	IPTG	New England Biolabs
pAM1847	Amp^R	DNA bending vector	N/A	(Reyes et al, 2011)
$pET28b^+$	Kan^R	Protein expression	IPTG	Novagen

Table 9: Bacterial plasmids used in this work.

The name, the antibiotic resistance, the use, the inducer and the reference are listed for each plasmid.

2.2. General DNA methods

2.2.1. Purification of chromosomal DNA

The chromosomal DNA from *S. aureus* SH1000 (See Table 8) was purified using the Wizard Genomic DNA Purification kit (Promega) and used for cloning. 5 ml of *S. aureus* SH1000 overnight culture were pelleted, resuspended in 50 µl of TSM (50 mM Tris-HCl pH 7.5, 0.5 M Sucrose, 10 mM $MgCl_2$ – filter sterilized) and transferred to a microcentrifuge tube. Then 2.5 µl of lysostaphin (2 mg/ml in 0.5 M Tris pH 8.0) was added to digest the cell wall and the resuspended cells were incubated at 37°C for 30 min to digest the cell wall. 600µl of Nuclei Lysis Solution (from the kit) were then added and the sample was incubated at 80°C for 5 min. The sample was then mixed by pipetting up and down to break the cells. If the sample did not look homogenous, it was incubated at 80°C for 5 more min. Then the sample was allowed to cool to room temperature (RT) before the addition of 3 µl of RNase Solution (from the kit, 4 mg/ml). Thereafter the sample was gently mixed and incubated at 37°C for 30

min then cooled to RT. 200 µl of Protein Precipitation Solution (from the kit) were added then the sample was vortexed and incubated on ice for 5 min. The sample was then spun down at 14,000 x g for 5 min. The supernatant was then transferred to a clean tube containing 600 µl of RT isopropanol and mixed. The sample was then spun down at 14,000 x g again for 5 min, the supernatant was decanted and the position of the pellet marked outside of the tube. The pellet was washed with 600 µl of RT 70% ethanol pipetted slowly over the pellet. The sample was then spun down at 14,000 x g for 5 min. The ethanol was aspirated and the pellet air-dried for 15 min (or until all the ethanol was gone and the pellet was invisible).
The DNA was rehydrated in 25 µl or 50 µl of rehydratation solution (from the kit) for 1 hour at RT or overnight at 4°C and resuspended by pipetting up and down. The concentration of the sample was measured using a NanoVue Plus (GE Healthcare) spectrophotometer and the sample was stored at -20°C.

2.2.2. Purification of plasmid DNA from *E. coli* (XL1 blue or DC10B)

Plasmid DNA was purified from *E. coli* XL1 blue cloning strains (See Table 8) in order to transform competent cells or to perform restriction analysis. Plasmid DNA was purified from *E. coli* DC10B in order to transform the *S. aureus* destination strain by electroporation. The plasmid DNA was isolated using the principle and some buffers of the QIAGEN's QIAprep miniprep system. Bacterial culture of 5 ml were grown overnight to saturation, pelleted and the supernatant was discarded. The cell pellet was then resuspended in 250 µl of P1 buffer (50 mM Tris-HCl pH 8, 10 mM EDTA, 100 µg/ml RNase A (from the kit) and mixed by inverting the tube. 250 µl of P2 buffer (200 mM NaOH, 1% SDS (w/v)) were then added to lyse the cells and the sample was mixed by inverting the tube. 350 µl of N3 buffer (from the kit) were then added to neutralise and the sample was mixed by inverting the tube. The cell debris were then pelleted by centrifugation at 18,000 x g for 10 min in a microcentrifuge (subsequent steps required 1 min centrifugation at 18,000 x g). The supernatant was then transferred to an anion exchange column (NBS Biologicals) and the DNA bound to the membrane following centrifugation. The DNA was washed using PE buffer (10 mM Tris-HCl pH 7.5, 80% ethanol) and eluted with 25 to 50 µl ddH_20. The concentration of the purified DNA was measured using a NanoVue Plus (GE Healthcare) spectrophotometer. The sample was analysed using agarose gel electrophoresis (See 2.2.5) then stored at -20°C. The pJR_{P2+P3} and pAM1847 plasmids were purified using the Quiagen Maxi Prep plasmid purification kit as per manufacturer's instructions which follow similar steps as above with a larger volume of cell culture (500 ml).

2.2.3. Purification of plasmid DNA from *S. aureus* (RN4220)

Plasmid DNA from *S. aureus* (RN4220) (See Table 8) was purified in order to transform the *S. aureus* destination strain by electroporation. The plasmid purification was performed exactly as for *E. coli* cells except for the first few steps. 5 ml of *S. aureus* RN4220 overnight culture was pelleted, resuspended in 50 µl of TSM (50 mM Tris-HCl pH 7.5, 0.5 M sucrose, 10 mM $MgCl_2$ – filter sterilized) and transferred to a microcentrifuge tube. Then 2.5µl of lysostaphin (2 mg/ml in 0.5 M Tris pH 8.0) were added to digest the cell wall and the resuspended cells were incubated at 37°C for 30 min. Then, as for *E. coli* cells, 250 µl of P1 buffer were added. The rest of the protocol was conducted as for *E. coli* cells but with adjusted volume of P2 and N3 buffer, 300 µl and 360 µl, respectively.

2.2.4. Restriction analysis of DNA and site specific DNA digestion

Restriction analysis and site specific DNA digestion of inserts or vectors was carried out in 10 µl reaction volumes and typically contained 1 µl of a 10x reaction buffer (optimal for the restriction enzyme(s)), 7 µl of plasmid DNA or PCR product (as prepared in section 2.2.2 or in section 2.2.7) and 1 unit of restriction enzyme(s). If required, Bovine Serum Albumin (NEB) was added at a final concentration of 10 mg /ml. The reactions prepared as above were then incubated at 37°C for 2 h, heat inactivated according to manufacturer's instructions when needed and then analysed by agarose gel electrophoresis (section 2.2.5). Digested plasmid DNA to be used for ligations was dephosphorylated at the 5' end to prevent self-ligation through the addition of 2.5 µl of FastAP Thermosensitive Alkaline Phosphatase (AP, Fermentas) and 2.5 µl of x10 AP buffer (Fermentas) to the 10 µl digestion reaction. The reaction was incubated at 37°C for 15 min, the AP enzyme was deactivated at 75°C for 5 min then the digestions purified (section 2.2.6).

2.2.5. Agarose gel electrophoresis

Agarose gels were prepared at either 1% agarose (w/v, Sigma), to visualise DNA fragments >200 bp, or 2% (w/v) for DNA fragments <200 bp, in 1x TBE buffer (Lonza). The agarose and 1x TBE solution was heated in a microwave until all the agarose was dissolved and being left to cool. Then a 1:10,000 dilution of *SYBR*® Safe DNA gel stain (Invitrogen) was added and the solution poured into a pre-levelled gel casting tray with a gel comb added to create the wells. 6x DNA loading dye (Fermentas) was added to the samples at 1x final concentration and then the samples were loaded onto the gel alongside a 1kb or 100 bp DNA marker (O'Generuler™ 1 kb or 100 bp, Fermentas), depending on expected DNA size range,

to determine DNA length. The gel was run at 110 V in 1x TBE for approximately 40 min before being visualised using a Safe Imager Blue Light Transilluminator (Invitrogen). Gel images were captured using a Gel Doc^{TM} EZ system (BIO-RAD).

2.2.6. Gel extraction and PCR purification of DNA

To extract and purify DNA from agarose gels the QIA quick gel extraction kit (QIAGEN) was used. The DNA was purified according to the manufacturer's instructions. This first involved cutting the DNA band of interest out from the agarose gel using a sterile blade and the Safe Imager Blue Light Transilluminator to visualise it. The gel band was then placed into a clean a microcentrifuge tube and dissolved at 50°C in 300 µl QG binding buffer (from the kit) per 100 mg of gel. Isopropanol was added to the gel fragment at a 1:1 (w/v), to the weight of the gel slice, and the solution was bound to the column (NBS Biologicals) using a microcentrifuge at 14,000 x g for 1 min. The column was then washed through PE buffer and then the DNA eluted in 25 or 50 µl of ddH_2O. The concentration of the sample was measured using a NanoVue Plus (GE Healthcare) spectrophotometer and the sample was stored at -20°C.

The products of the PCR reactions were purified as the method above using the QIAquick PCR purification kit (QIAGEN) except that the PCR reaction was re-suspended in 5 volumes binding buffer PB (from the kit) and then bound to the column. Then the protocol above was followed as described.

2.2.7. Polymerase chain reaction (PCR)

The genes of interests were amplified from chromosomal DNA (section 2.2.1) or from plasmid DNA (section 2.2.2) using Herculase II Fusion DNA polymerase (Agilent) and the relevant oligonucleotides (primers). The primers were synthesised by Sigma and can be found in Table 13 (See section 8.1.1). Typically, a standard PCR reaction of 50 µl total volume contained 50-100 ng template DNA, 0.25 µM of each primer, 10 mM dATP, dGTP, dCTP and dTTP, 10 µl 5x reaction buffer and 5 units of DNA polymerase mixed altogether. The PCR reaction was conducted in 0.2 ml thin-walled PCR tubes in a GenePro thermocycler (BIOER) according to the following protocol: 95ºC for 2 min, then thirty repetitions of 95ºC for 20 sec, 54ºC (or Tm- 5ºC) for 30 sec, 72ºC for 30 sec/kb and then a last step at 72ºC for 3 min. The annealing temperature of 54 ºC was standardly used instead of varying the melting temperature (Tm) according to each primer set but was modified according to the primer Tm when needed.

For the amplification of staphylococcal DNA the standard PCR protocol was modified in order to facilitated the binding of the primers to very AT rich sequence. The staphylococcal PCR protocol consisted of 95ºC for 2 min, 95ºC for 20 sec, 45ºC for 30 sec, 72ºC for 30 sec/kb then thirty repetitions of 95ºC for 20 sec, 54ºC (or Tm- 5ºC) for 30 sec, 72ºC for 30 sec/kb and then a last step at 72ºC for 3 min.

Because Herculase is a high fidelity enzyme, it does not have a polyadenylating activity. Therefore, The PCR products destined to be ligated into the pGEM-T easy vector (Promega) were polyadenylated using the GoTaq DNA polymerase (Promega) which polyadenylates linear DNA ends. Typically, 7 µl of PCR purified product was mixed with 5 units of Go Taq polymerase, 2 µl 5x Go Taq Buffer and 0.2 mM dATP in a 10 µl reaction and incubated at 72 ºC for 20 min.

PCR was also used to confirm the presence of cloned genes in the appropriate vector within individual colonies (termed colony PCR). This was performed by resuspending a single colony in 10 µl LB and using 8 µl of this suspension in a 50 µl PCR reaction (as described above). The remaining 2µl were spotted onto a LB agar plate with appropriate antibiotics in order to maintain the tested colony for future use. For colony PCR the GoTaq DNA polymerase was used with the following protocol : 95ºC for 15 min, then thirty repetitions of 95ºC for 30 sec, 54ºC (or Tm- 5ºC) for 30 sec, 72ºC for 1 min/kb and then a last step at 72ºC for 5 min. Similarly as for standard PCR, with Staphylococcal colonies, the protocol was modified to include 95ºC for 30 sec, 45ºC for 30 sec, 72ºC for 1 min/kb before the thirty repetitions.

2.2.8. Site directed mutagenesis (SDM)

Site directed mutagenesis (SDM) was conducted to construct the library of single alanine mutants using the pSN-P2-*agrA* plasmid as template and the high fidelity PfuUltra Fusion II HS DNA polymerase (Agilent) and the relevant complementary oligonucleotides (See Table 15, Table 16 and Table 17). The 50 µl reaction contained ~100 ng template DNA, 0.1 µM of each primer, 10 mM dATP, 10 mM dGTP, 10 mM dCTP and 10 mM dTTP, 5 µl 10x PfuUltra reaction buffer and 2.5 unit of PfuUltra DNA polymerase mixed. The PCR reaction was performed according to the following protocol: 95ºC for 50 sec, 95ºC for 50 sec, 45ºC for 1 min, 72ºC for 2 min and 30 sec then fifteen repetitions of 95ºC for 50 sec, Tm- 5ºC for 1 min, 68ºC for 2 min and 30 sec and then a last step at 68ºC for 7 min. The negative control reaction included everything but the primers and was treated exactly as the others samples

were. A separate negative control was performed for each different annealing temperature. Two separate PCR reactions were performed for each primer pair.

After the PCR reaction was completed, the pSN-P2-*agrA* template plasmid was degraded using the *Dpn*I restriction enzyme. This enzyme specifically digests methylated DNA with the target sequence 5'-GmGATC-3'. After the PCR reaction, only the remaining template DNA is methylated and therefore will be specifically degraded. 6 µl of 10 x buffer 4 (New England biolabs), 3 µl ddH_2O, and 20 units *Dpn*I were added to each PCR reaction. The samples were incubated for 2-3 h at 37ºC. The two PCR reactions for each primer pair were then pooled together before being purified using the QIA quick PCR purification kit (section 2.2.6) and eluted in 30 µl ddH_2O. 15 µl of purified SDM product were used to transform *E. coli* XL1 blue competent cells (See section 2.2.12) the rest was stored at -20°C. The next day, the transformation plates were checked for colonies. If the transformation plates for the primer pair had a satisfactory amount of colonies compared to the corresponding negative control (at least 2-fold more), two colonies per primer pair were selected and used to inoculate liquid culture for plasmid purification (See section 2.2.2). If the ratio was not satisfactory, the SDM was repeated with a different annealing temperature. The purified plasmids were then sent to sequencing using the appropriate primer to verify that the mutation had been successfully incorporated.

2.2.9. Sequencing

The samples for sequencing were sent to GATC Biotech and sequenced using Sanger ABI 3730xl. The reactions sent to sequencing contained 25 pmoles of the appropriate primer and 500 ng of DNA in a total volume of 10 µl. All sequencing results were analysed using BioEdit sequence Alignment Editor (Hall, 1999).

2.2.10. pGEM-T easy vector ligation

The use of the pGEM-T easy vector simplifies molecular cloning in the case of problem with the restriction digest of the insert, especially in case of blunt end cutting. The pGEM-T easy was used to clone into the pTYB2 and the pCN34-*itet* vectors (See Table 9).

The ligation of DNA into the linear pGEM-T easy vector requires purified DNA with polyadenylated ends. The polyadenylated PCR products required for direct ligation into the linear pGEM-T easy vector with poly-thymine tails using T4 DNA ligase (Promega) were prepared as in section 2.2.7. A typical 10 µl ligation reaction consisted of 1 µl pGEM-T easy vector, 3 µl insert DNA, 5 µl x2 ligase buffer and 1 unit T4 DNA ligase. The reaction was

then incubated at 4°C overnight and then transformed into competent *E. coli* XL1 blue (See section 2.2.12).

2.2.11. Vector ligation

Ligations were setup using either insert DNA digested from pGEM-T easy plasmids (section 2.2.10) or insert DNA digested from PCR reactions (section 2.2.4) for ligation into the corresponding restriction enzyme digested plasmid DNA (section 2.2.4). A typical 10 µl ligation reaction consisted of 50-100 ng of plasmid DNA, 3-fold excess (150-300 ng) of insert DNA, 5 units of T4 DNA ligase (Fermentas) and 1 µl of 10x T4 DNA ligase buffer. The reaction was then incubated at RT for 1 hour and the enzyme heat inactivated at 65°C for 10 min before transformation into competent *E. coli* XL1 blue (See section 2.2.12).

2.2.12. Chemically Competent cells preparation (*E. coli* cells)

Chemically competent cells were prepared using the $CaCl_2$ method and used with *E. coli* XL1 and DC10B strains. The *E. coli* strain of interest was streaked out onto an LA plate and incubated at 37°C overnight. A seed culture was prepared by inoculating 5 ml of LB with a single bacterial colony and incubating at 37°C overnight. 200 ml of LB was inoculated with 200 µl of seed culture and cells were grown at 37°C shaking until mid-exponential phase, $OD_{600\ nm}$ = 0.4-0.45. The culture was then centrifuged at 3200 x g for 10 min at 4°C. The cell pellet was resuspended in 40 ml of pre-chilled 100 mM $CaCl_2$ and 15% (v/v) glycerol solution. Cells were then incubated on ice, in a 4°C room overnight. The cells were centrifuged again at 3200 x g for 10 min at 4°C. The pellet was then resuspended in 6 ml of pre-chilled 100 mM $CaCl_2$ and 15% (v/v) glycerol solution. Cells were either used immediately or aliquoted into 50 µl volumes, snap frozen in liquid nitrogen and stored at -80°C.

2.2.13. Heat shock transformation

To transform purified *E. coli* plasmid DNA (section 2.2.2) or ligation product (section 2.2.10 or 2.2.11) into competent *E. coli* strains (section 2.2.12) aliquots of the appropriate cells were taken from -80°C and thawed on ice for 5 min. ~100 ng of plasmid DNA was added to each aliquot and incubated on ice for 30 min. The cells were then heat shocked by placing the cell aliquot into a 42°C water bath for 1 min and then promptly placing the aliquot back on ice for 2 min. 950 µl of SOC media were added to the cells under sterile conditions and the cells placed in a shaking incubator at 37°C for 1 hour to recover. The cells were then pelleted at 7,000 x g for 2 min and the supernatant removed. The pellet was resuspended in 50 µl of the

original SOC media and spread onto a LB agar plate containing the appropriate antibiotics. The LB agar plate was incubated at 37°C overnight to allow colony formation.

2.2.14. Electro-competent cells preparation (*S. aureus* cells)

Electro-competent cells were prepared using the sucrose method to prepare all *S. aureus* strains. The *S. aureus* strain of interest was streaked out onto a TSA plate and incubated at 37°C overnight. A seed culture was prepared by inoculating 4 ml of TSB with a single bacterial colony and incubating at 37°C overnight. 200 ml of TSB were inoculated with 200 µl of seed culture and cells were grown at 37°C shaking for 3 h. The culture was then centrifuged at 3200 x g for 10 min at 4°C. The cell pellet was washed twice with 200 ml of 0.5 M sterile sucrose solution and once with 100 ml of 0.5 M sterile sucrose solution before being resuspended in 2 ml of 0.5 M sterile sucrose solution. The cells were either used immediately or aliquoted into 110 µl volumes and stored at -80°C.

2.2.15. Electroporation

Electroporation was used to transform all *S. aureus* strains. RN4220 can be transformed with any kind of *E. coli* plasmid DNA as it is a restriction deficient strain. All other *S. aureus* destination strains can be transformed either with plasmid DNA purified from RN4220 (section 2.2.3) or from *E. coli* DC10B (section 2.2.2) as it a strain deficient in dcm methylation and is therefore not recognised as *E. coli* DNA. All other *E. coli* plasmid DNA will be degraded.

To transform electro-competent cells (section 2.2.14) aliquots of the appropriate cells were taken from -80°C and thawed on ice for 10 min. 20µl of 600-1000 ng/µl of appropriate plasmid DNA were added to a 0.025 µm membrane filter (Millipore) and dialysed against sterile ddH_2O for 15 min in a Petri dish. 15 µl of dialysed plasmid were then transferred to an aliquot of S. aureus thawed competent cells. The mixture was then transferred into a 0.1 cm cuvette (VWR) and electroporated using a Gene Pulser (BIO-RAD) with the following settings: 100 Ω, 2.5 kV, 25 µF. Immediately after electroporation, the cells were transferred into a microcentrifuge tube containing 0.9 ml sterile BHI 0.5 M sucrose medium. The cells were placed in a shaking incubator at 37°C for 1 hour to recover. The cells were then pelleted at 7,000 x g for 2 min and the supernatant removed. The pellet was re-suspended in 50 µl of the original BHI 0.5 M sucrose medium and spread onto a TSA plate containing the appropriate antibiotics. The TSA plate was incubated at 37°C overnight to allow colony formation.

2.3. General protein methods

2.3.1. Protein Expression

The proteins overexpressed in this work were expressed using two similar systems (the pTYB2 and the pET system) and three different expression conditions. The three different expression conditions will be therefore listed in this section as well as the principle of the expression systems.

The *S. aureus* core RNA polymerase was purified from *S. aureus* SH100 cells without use of an overexpression system. The expression method is described below as well.

2.3.1.1. The pTYB2 and the pET systems

The pTYB2 and the pET system are two similar expression systems. In both cases, the expression vector contains a T7 promoter located upstream of the multiple cloning site (MCS), and as well as a copy of the *lacI* repressor and a *lac* operator sequence just downstream of the T7 promoter. This T7 promoter region is specifically recognised by the bacteriophage T7 DNA dependent RNA polymerase (T7 RNAp). The two expression strains used to propagate these vectors (*E. coli* BL21 (DE3) and *E. coli* ER2566, Table 8) contain a chromosomal copy of the T7 RNAp under the control of a *lac* promoter. The BL21 (DE3) strain is a DE3 λ prophage lysogen carrying the T7 RNAp under the control of the *lacUV5* promoter and the *lacI* repressor. The ER2566 strain contains the T7 RNAp inserted in the *lacZ* gene and therefore also under the control of the *lac* promoter. The gene of interest cloned in the MCS of any of the two vectors will therefore be under the control of the T7 promoter and the T7 RNAp. The expression of the gene of interest and of the chromosomal T7 RNAp is inhibited by the *lac* I repressor which occupies the lac operator sequence, therefore blocking the *lac* promoter. The addition of the inducer isopropyl β-D-1-thiogalactopyranoside (IPTG), the molecular mimic of allolactose, outcompetes and displaces the lac repressor therefore releasing both *lac* promoters and allowing the production of T7 RNAp. The T7 RNAp is then able to bind to the T7 promoter of the pET or pTYB2 vector and transcribe the target gene encoded in the MCS. Unlike allolactose, IPTG is non-hydrolysable and therefore remains at a constant level. However, in the absence of the inducer, a basal level of transcription is still present due to the dynamic equilibrium between occupied and unoccupied operator sites by the repressor protein. Therefore a low level of target protein is expressed in the absence of the inducer. The pET28b^{+} vector (See Table 9) encodes a six-histidine tag (6His tag) at the target protein N-terminus for nickel affinity

chromatography (See section 2.3.4.1). This system was used to express the *S. aureus* proteins AgrA and AgrC-HK. The pTYB2 vector encodes a self-cleavable intein-tag with a chitin-binding domain at the target protein C-terminus for Intein Mediated Purification with an Affinity Chitin-binding Tag (IMPACT) (See section 2.3.4.3). This system was used to express the *S. aureus* AgrA, $AgrA_{L171A}$, $AgrA_{E181A}$, $AgrA_{H200A}$, $AgrA_{Y229A}$ and AgrAR233A.

2.3.1.2. Expression of [6His]-AgrA in pET28b+

The [6His]-AgrA protein is under the control of the pET system (See section 2.3.1.1). The pJR28-[6His]*agrA* plasmid (See Table 9) was transformed, as detailed in section 2.2.13, into *E. coli* BL21 (DE3) (See Table 8). The transformation plate was used to inoculate 1 l of LB media with the corresponding antibiotics in a 2 l flask. The culture was grown in a shaking incubator at 37°C without induction until the culture reached $OD_{600\ nm}$ = 1. The cells were then pelleted at 4,000 x g for 20 min and the cell pellet stored at -20°C overnight, ready for protein purification. 1 ml samples were taken after inoculation and before the cells were pelleted. These samples were centrifuged at 17,000 x g and the pellet was resuspended in 100 µl of ddH2O for future SDS-PAGE analysis (See section 2.3.3).

2.3.1.3. Expression of AgrC-HK in pET28b+

AgrC-HK is under the control of the pET system (See section 2.3.1.1). The pSN-*agrC-HK* (See Table 14) was transformed, as detailed in section 2.2.13, into *E. coli* BL21 (DE3) (Table 8). The transformation plate was used to inoculate 1 l of LB media with the corresponding antibiotics in a 2 l flask. The culture was grown in a shaking incubator at 37°C to $OD_{600\ nm}$ = 0.4-0.45, then transferred into a shaking incubator 25°C for 30 min before the addition of the inducer (1 mM IPTG for pET vectors). The cultures were then grown for a further 3 h at 25°C shaking. Cells were then pelleted at 4,000 x g for 20 min and the cell pellet stored at -20°C overnight, ready for protein purification.1 ml samples were taken before and after induction. These samples were centrifuged at 17,000 x g and the pellet was resuspended in 100 µl of ddH2O for future SDS-PAGE analysis (See section 2.3.3).

2.3.1.4. Expression of AgrA , $AgrA_{L171A}$, $AgrA_{E181A}$, $AgrA_{H200A}$, $AgrA_{Y229A}$ and $AgrA_{R233A}$ in pSN-*agrA*, pSN-*agrA* $_{L171A}$, pSN-*agrA*$_{E181A}$, pSN-*agrA*$_{H200A}$, pSN-*agrA*$_{Y229A}$ *and* pSN-*agrA*$_{R233A}$

AgrA is under the control of the pTYB2 system. The pSN-*agrA,* pSN-*agrA* $_{L171A}$,pSN-*agrA* $_{E181A}$,pSN-*agrA* $_{H200A}$,pSN-*agrA* $_{Y229A}$ *and* pSN-*agrA* $_{R233A}$ (See Table 14) were transformed, section 2.2.13, into *E. coli* ER2566 (Table 8). The transformation plates were used to

inoculate 5 ml seed cultures of LB media with the corresponding antibiotics that were grown at 37°C overnight. 1ml of each seed culture was used to inoculate 100ml of LB media with the corresponding antibiotics in 1 l flasks. The cultures were grown in a shaking incubator at 37°C to $OD_{600\ nm}$ = 0.6, then transferred in a shaking incubator at 16°C for 30 min before the addition of the inducer (0.25mM IPTG). The cultures were then grown overnight at 16°C shaking. Cells were then pelleted at 4,000 x g for 20 min and the cell pellets were used for protein purification immediately.1 ml samples were taken before and after induction. These samples were centrifuged at 17,000 x g and the pellet was resuspended in 100 µl of ddH2O for future SDS-PAGE analysis (See section 2.3.3).

2.3.2. Expression of the S. aureus core RNA polymerase (RNAp)

The *S. aureus* SH1000 strain (See Table 8) was streaked out onto a TSA plate and incubated at 37°C overnight. A 100 ml TSB starter culture was inoculated with the plate and incubated at 37°C overnight. 6 2 l flasks containing 1 l of TSB were each inoculated with 10 ml of starter culture and grown in a shaking incubator at 37°C to $OD_{600\ nm}$ = 0.5. Cells were pelleted by centrifugation at 4°C at 5,000 x g for 20 min and the cell pellet re-suspended in 100 ml of RNAp lysis buffer A (10mM Tris-HCl pH 7.5, 1 M KCl, 15 mM $MgCl_2$, 5 mM EDTA and 10 mM 2-mercaptoethanol) before re-centrifugation and the pellet was stored at -20°C overnight.

2.3.3. SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)

Sodium dodecyl sulphate (SDS) polyacrylamide gel electrophoresis (PAGE) was used to resolve protein for visualisation by coomassie staining or western blot. SDS-PAGE gels were prepared at an appropriate percentage according to the molecular weight of the protein of interest (the acrylamide percentage is inversely proportional to protein size). Solutions were prepared as described in Table 10 with the addition of APS and TEMED immediately prior to pouring. The resolving solution was quickly poured into the 1 mm thick space between two Mini-PROTEAN glass plates (BIO-RAD) to 1 cm below the well level. The rest was filled with isopropanol to level the gel solution. Once the gel was set, the isopropanol was decanted and the top of the gel was washed with ddH_2O. Residual ddH_2O was removed by absorption with Whatmann paper. The stacking gel was then prepared and poured on top of the resolving gel and either a 10 or 15 well comb inserted. Once set, the gels were either used immediately or wrapped in wet paper towel and clingfilm and stored at 4°C for future use.

	Resolving gel 10%	Resolving gel 12.5%	Stacking gel 4%
Acrylamide solution I	3.25 ml	4.1 ml	0.6 ml
Acrylamide solution II	2.5 ml	2.5 ml	-
Acrylamide solution III	-	-	1 ml
H_2O	4.2 ml	3.4 ml	2.4 ml
10% (w/v) APS	100 µl	100 µl	40 µl
TEMED (Sigma)	10 µl	10 µl	4 µl

Table 10: Composition of the SDS resolving and stacking gels used in this study.

The composition of 10% and 12.5% resolving gel and 4% stacking gel is listed.

The gels were assembled in a Mini-PROTEAN Tetra Electrophoresis gel tank (BIO-RAD) with 1x Tris Glycine SDS (TGS, National Diagnostics) buffer and pre-warmed by running at 200 V. Protein samples were prepared by mixing 20 µl or 40 µl of protein sample with 20 µl or 40 µl of 2x Laemmli buffer (4% (w/v) SDS, 20% (v/v) glycerol, 10% (v/v) 2-mercaptoethanol, 0.004% (w/v) bromophenol blue, 0.125 M Tris-HCl, pH 6.8, Sigma), incubated at 95°C for 5 min and then briefly centrifuged. Samples from the IMPACT purification were mixed with a loading buffer that did not contain any thiol, as they would have released the intein tag from the protein. The 3x SDS Sample buffer was used (187.5 mM Tris-HCl pH 6.8, 6% (w/v) SDS, 30% glycerol and 0.03% (w/v) bromophenol blue) with the IMPACT samples. The samples were then loaded into the pre-warmed gel alongside 8 µl of Page Ruler Prestained protein ladder (Fermentas) to help determine protein size and the gel run at 100 V for 10 min and then 200 V for approximately 40 min or until the dye front reached the bottom of the gel. The gel was then removed from the glass plates, washed with dH_2O for 5 min and then stained with 20 ml of Bio-safe coomassie stain (BIO-RAD) for 30 min. The gel was then destained in dH_2O through several washing steps. The GelDoc EZ imager was used to capture the gel images. For western blotting (See section 2.3.8), the gel was removed from the glass plates and equilibrated in 1x Transblot buffer.

2.3.4. Protein purification

2.3.4.1. Nickel affinity chromatography

Proteins with a 6His tag ([6His]-AgrA (section 2.3.1.2) and AgrC-HK (section 2.3.1.3)) were purified using Nickel affinity chromatography. Nickel has a strong affinity for histidines. The 6His-tagged proteins will therefore bind to the nickel column allowing their separation from the other proteins present in the whole cell lysate. 6His-tagged proteins will remain bound to the column until the addition of imidazole. Imidazole will act as a molecular mimic of the histidine imidazole ring side-chain and at a high concentration displace the 6His-tagged proteins from the nickel column resulting in the pure protein being eluted.

Cell pellets containing expressed target proteins stored at -20°C were thawed on ice for 20 min and re-suspended in 30 ml Lysis buffer (100 ml nickel buffer A + 1 protease inhibitor tablet (Roche)). 500 mg of lysozyme (Sigma) were sprinkled on top of the cell suspension before the cells were sonicated for 12 min (30 sec on-off cycle) at 40% amplitude using a Vibra cell sonicator (SONICS) until lysed. The lysed sample was spun in a high speed centrifuge tube at 4°C at 32,000 x g for 30 min to separate the soluble from the insoluble protein fraction. The supernatant was removed and stored on ice for loading onto a HiTrap^{TM} Chelating HP nickel column (GE Healthcare) and a 40 µl sample was taken for SDS-PAGE analysis (section 2.3.3). The cell pellet was resuspended in 30 ml nickel buffer A (25 mM NaH_2PO_4 pH 7, 500 mM NaCl, 5% (v/v) glycerol) and a 40 µl sample was taken for SDS-PAGE analysis (section 2.3.3). The Fast Protein Liquid Chromatography (FLPC) system (ÄKTA prime™, GE healthcare) was prepared by washing the system through with pre-cooled nickel buffer A and B (nickel buffer A + 1 M imidazole). The nickel column was loaded with 0.1 M $NiSO_4$ and equilibrated with 3 column volumes of nickel buffer A, B and then A. The nickel column was attached to the FPLC system, with caution taken to reduce air blockages, and the sample loaded into the column at 0.5 ml per min. The column was then washed at 2 ml per min with 2 column volumes of buffer A followed by 3% buffer B and then buffer A again. A 40 ml gradient was then started at 1 ml per min, where the buffer composition was slowly increased from 0% (v/v) buffer B to 100% (v/v) buffer B to elute and collect column bound proteins in 1 ml fractions. Samples were taken throughout the purification steps in order to check the presence of the protein and to determine which elution fraction to dialyze through SDS-PAGE electrophoresis (See section 2.3.3). [6His] AgrA was dialysed in antibody storage buffer (10 mM Tris-HCl, pH 8.0, 50 mM NaCl, 0.1 mM EDTA, 10% (v/v) glycerol) and AgrC-HK was dialysed in HK storage buffer (10 mM HEPES(4-(2-

hydroxyethyl)-1-piperazineethanesulfonic acid)-KOH pH 8, 50 mM KCl, 0.1 mM EDTA, 1 mM DTT and 10% (v/v) glycerol) overnight at 4°C using Spectra/Por 3 dialysis tubing (Spectrum) and the protein concentration measured (See section 2.3.5). Protein samples were aliquoted into microcentrifuge tubes and stored at -80°C.

2.3.4.2. Immunoaffinity chromatography

Immunoaffinity chromatography was used for the extraction of *S. aureus* SH1000 core RNAp (section 2.3.2) using the polyol responsive 8RB13 affinity resin (Neoclone) (Bergendahl et al, 2003). The cell pellet was thawed on ice and re-suspended in 100 ml of RNAp lysis buffer B (10 mM Tris-HCl, pH 7.5, 50 mM KCl, 15 mM $MgCl_2$, 5 mM EDTA and 10 mM 2-mercaptoethanol) containing 1 tablet of RNase inhibitors (Roche) and 200 µl of 2 mg/ml lysostaphin. The cells were lysed by passing the cell suspension three times through a French pressure cell press (SIM-AMINCO) pre-chilled to 4°C and pre-washed with RNAp buffer A, at 11,000 psi. The soluble fraction was collected through centrifugation at 4°C at 32,000 x g for 30 min. All further steps were carried out at 4°C. The nucleic acid and RNAp were precipitated by the addition of 0.6% (v/v) polyethyleneimine (PEI) (Sigma) from a 10% stock (pH7.9 using HCl), gradually added and mixed over 15 min before a further 30 min mixing. The precipitate was pelleted in a high speed centrifuge tube at 32,000 x g for 20 min and the pellet thoroughly re-suspended in 10 ml RNAp wash buffer (10 mM Tris-HCl pH 7.9, 0.1 mM EDTA, 300 mM NH_4Cl, 5% (v/v) glycerol and 0.1 mM DTT). The suspension was re-centrifuged at 32,000 x g for 20 min and re-suspended in 10 ml RNAp elution buffer (RNAp wash buffer adjusted to pH 7 + 1 M NH_4Cl), before further centrifugation at 32,000 x g for 20 min. The supernatant was collected and gently mixed in a glass beaker with 0.4 g $(NH_4)SO_4$ (Sigma)/ml supernatant, gradually added over 15 min, followed by a further 30 min mixing. The precipitate was pelleted at 32,000 x g for 10 min and the pellet stored at -20°C overnight.

The $(NH_4)SO_4$ pellet was thawed on ice and re-suspended in 10 ml TE buffer (50 mM Tris-HCl pH 7.9 and 0.1 mM EDTA) and spun at 5000 rpm for 10 min at 4°C to remove insoluble materials. The supernatant was then passed three times through a prewashed column (BIO-RAD), with 10 ml TE buffer, packed with 1 ml 8RB13 polyol responsive antibody resin (Neoclone) to maximise RNAp contact time. The column was then washed with 10 ml TE buffer plus 0.15 M NaCl followed by 10 ml TE buffer plus 0.5 M NaCl. The remaining bound protein, including RNAp, were eluted with 5 ml TE buffer plus 0.75 M NaCl and 40% propylene glycol (Sigma), with 1 ml elution fractions collected. 20 µl samples taken

throughout the purification were analysed through SDS PAGE analysis (See section 2.3.3) and fractions containing >95% homogeneity of *S. aureus* core RNAp were concentrated through centrifugation at 5,000 x g using Amicon Ultra-4 spin columns (Millipore), with a 5 kDa cut-off, before being dialysed in standard storage buffer and stored at -80°C, as in section 2.3.4.1.

2.3.4.3. Intein mediated purification with an affinity chitin-binding tag (IMPACT)

AgrA, $AgrA_{L171A}$, $AgrA_{E181A}$, $AgrA_{H200A}$, $AgrA_{Y229A}$ and $AgrA_{R233A}$ (See section 2.3.1.4) were purified using the IMPACT^{TM} kit (New England Biolabs) according to the manufacturer's instructions. Briefly, the cells were lysed in 6 ml column buffer (20 mM Tris-HCl pH 9, 1 M NaCl and 1 mM EDTA) through sonication for 12 min (30 sec on-off cycle) at 40% amplitude and centrifuged in a high speed centrifuge tube at 4°C at 32,000 x g for 30 min to remove cellular debris. The pellet was resuspended in 6 ml column buffer and the supernatant was then loaded on a 10 ml gravity flow column (BIO-RAD) packed with 2 ml Chitin Resin (New England Biolabs) prewashed with 12 ml column buffer. The column was washed with 40 ml of column buffer and the protein was cleaved from the intein tag after incubation for 16 h at 4°C in 6 ml of cleavage buffer (column buffer + 200 mM DTT). 40 µl samples taken throughout the purification were analysed through SDS PAGE analysis (section 2.3.3) The protein was concentrated using Amicon Ultracel-10K (Millipore) and dialysed in low glycerol storage buffer (10 mM Tris-HCl pH 8, 50 mM NaCl, 0.1 mM EDTA, 1 mM DTT and 20% (v/v) glycerol) as in section 2.3.4.1. The proteins were aliquoted and stored at -20°C.

2.3.5. Determination of protein concentration by Bradford microassay

The Bradford assay is a colorimetric protein assay that is based on the absorbance shift of the Coomassie Brilliant Blue G-250 dye. The dye exists in three different forms, a blue one (anionic), a green one (neutral) and a red one (acidic). Under acidic conditions, the dye is therefore mainly in its red form. However, when the dye binds to proteins it is converted to a stable un-protonated blue form. Therefore, the amount of blue dye is a measure for the protein concentration and can be estimated by use of an absorbance reading. The absorbance of the bound blue dye is at $OD_{595\ nm}$ so the increased absorbance at $OD_{595\ nm}$ is proportional to the bound dye and thus to the amount of protein in the sample.

The concentration of the purified proteins in the different storage buffers was determined using the Bradford reagent dye (Quick start Bradford 1x Dye reagent, BIO-RAD) which contains the Coomassie Brilliant Blue G-250 dye. The protein concentration was calculated using a standard curve constructed with the absorbance of known concentration of protein bovine serum albumin (BSA) standards. Typically, the standard curve contained serial dilutions of BSA standards ranging from 0.05 mg/ml to 2 mg/ml.

5 µl of protein sample and of each BSA sample were pipetted in clear flat-bottomed 96-well plate (Sterillin) and 250 µl of Bradford reagent were added to each sample. The 96-well plate was incubated at RT for 15 min and then the absorbance of the protein samples at 595 nm was determined using the POLAR star Omega plate reader.

2.3.6. Polyclonal antibody production

Purified [6His]-AgrA (section 2.3.4.1) was sent to Eurogentec for rabbit immunisation. Two different rabbit were immunised with the protein. Bleeds were taken pre-immunisation, 21 days after immunisation and after termination (final bleed). Anti-peptide ELISA analysis was conducted for the final bleed prior to shipping.

2.3.7. Affinity purification of antibody serum

50 µg untagged AgrA (section 2.3.4.3) was resolved by SDS-PAGE and transferred to a polyvinylidenedifluoride (PVDF) membrane (BIO-RAD) as per the western blotting protocol (refer to section 2.3.8). The membrane was then stained with Ponceau S solution (Sigma) for 5 min and rinsed with dH_2O to visualise the protein. The section of the membrane containing the AgrA protein was cut and transferred to a 50 ml falcon tube. The membrane was blocked with 25 ml of 1x PBS with 5% fat-free milk powder for 1 hour and then incubated with 25 ml of 1x PBS with 5% fat-free milk powder and 5 ml of final bleed antibody serum at 4°C overnight. The membrane was rinsed twice in 1x PBS for a few sec each time and then the antibody was eluted with 4 ml of 100 mM glycine pH 2.7 for exactly 2 min. The eluate was transferred into a 15 ml falcon tube containing 300 µl 1.5 M Tris-HCl, pH 8.8 to neutralise the acid and was then dialysed against 500 ml PBS twice over 8 h. The antibody was then aliquoted and frozen at -20°C.

2.3.8. Western blotting

The S. aureus samples for Western blot were prepared differently than standard samples for SDS-PAGE analysis (See section 2.3.3). 1 ml or 200 µl of S. aureus culture at $OD_{600\ nm}$ = ~2 were spun down and resuspended in 500 µl or 100 µl of ddH_2O. 25 µl or 5 µl of 2 mg/ml lysostaphin were added to the sample and it was incubated 30 min at 37°C. Then 500 µl or 100 µl of 2x Laemmli buffer were added to the sample and the sample was boiled for 5 min at 95°C and briefly spun down. The samples were then resolved using the appropriate percentage SDS-PAGE (See section 2.3.3). The gel was then equilibrated in 1x transblot buffer for 10 min. At the same time, a gel-sized sheet of PVDF membrane was soaked in methanol for 10 s, in water for 5 min and then equilibrated in 1x transblot buffer for 10 mins. Meanwhile, eight gel-sized sheets of Whatmann paper were soaked in 1x transblot buffer. The blot was assembled onto a Trans-Blot Turbo transfer system (Bio-RAD) in the form of a sandwich. The "bread" on either side was constituted by the 4 Whatmann paper pieces, the "meat" was represented by the PVDF membrane, above the terminal 4 Whatmann paper pieces and the "salad" was represented by the gel, above the membrane. Any bubbles were then removed by rolling over the blot with a stripette. The proteins in the gel were transferred onto the membrane by running at 10 V for 1 hour. The membrane was then blocked in 25 ml 1x TBS with 5% fat-free milk powder and 37.5 µl of 5 mg/ml human immunoglobulin G (hIgG, Sigma) overnight at 4°C. The membrane was washed twice for 10 min each in TBSTT and for a further 10 min in TBS. The membrane was then incubated for 1 hour in 20 ml 1x TBS with 5% fat-free milk powder, 30 µl of 5 mg/ml hIgG and 80 µl the primary antibody, the agrA polyclonal antibody. The membrane was then washed twice for 10 min each in TBSTT and for a further 10 min in TBS. The membrane was then incubated for 1 hour in 20 ml 1x TBS with 5% fat-free milk powder, 30 µl of 5 mg/ml hIgG and 6.7 µl of the secondary antibody, the anti-rabbit–horseradish peroxidase-conjugated antibody (Dako). After the incubation the membrane was washed 5 times for 10 min each in TBSTT. The blot was then placed with the protein face down on premixed ECL prime western blotting detection reagent (GE Healthcare, 500 µl of reagent A and 500 µl of reagent B) and incubated for 5 min. The blot was then briefly dried with tissue paper and placed between the two transparent sheets of a X-ray film cassette. The blot was either exposed on CL-XPOsure Film (Thermo Scientific) for the appropriate amount of time (10 s to 30 min usually) and developed using a CURIX 60 developer (AGFA) or visualised using a LAS-3000 imaging system (Fujifilm).

2.4. General biochemical methods

2.4.1. Probe labelling with ^{32}P

Typically, 20 pmoles of single stranded oligonucleotides or double stranded DNA (PCR products or restricted fragment) were labelled with 35 µCi γ-^{32}P-ATP (500 Ci/mmol-PerkinElmer) using a T4 ligase (Promega) to form ^{32}P labelled DNA promoter probes for electrophoretic mobility shift assays (section 2.4.2 and 2.4.3) or DNA bending assay (section 2.4.6). T4 DNA ligase catalyses the joining of the 5'-phosphate and 3'-hydroxyl groups of adjacent nucleotides and therefore is able to transfer the ^{32}P labelled γ-phosphate from γ-^{32}P-ATP to the free hydroxyl group of DNA oligonucleotides. Typically, 20 µl labelling reactions contained 20 pmoles of DNA, 35 µCi γ-^{32}P-ATP, 2 µl x10 T4 DNA ligase buffer A (300 mM Tris-HClpH7.8, 100 mM $MgCl_2$, 100 mM DTT and 10 mM ATP, Promega) and 10 units of T4 kinase . Reactions were incubated at 37°C for 30 min before the kinase was denatured at 70°C for 10 min. For oligonucleotides, duplex DNA was formed in 40 µl reactions with the 20 µl of ^{32}P labelled single stranded oligonucleotides, 40 pmoles of the complementary single stranded unlabelled oligonucleotide and 4 µl x10 TM buffer (10 mM Tris-HCl pH 10 and 1 mM $MgCl_2$) made up to 40 µl in distilled water to give a 500 nM final DNA concentration. Reactions were heated in a heat block at 95°C for 5 min before the slow cooling of the reactions through placing the heat block at RT overnight. The DNA duplex was then stored at -20°C.

2.4.2. Electrophoretic gel mobility shift assay (EMSA) with acetyl phosphate activation

Phosphorylation of AgrA was carried out by pre-incubating 3 µM AgrA with 50 mM acetyl phosphate (Sigma) and 5 mM $MgCl_2$ for 1 h at 37°C. 10 µl binding reactions were set up in reaction buffer (10 mM HEPES pH 7.6, 50 mMKCl, 1 mM EDTA, 2 mM DTT, 0.5% (v/v) TritonX-100 and 5% (v/v) glycerol) using 1 µM final concentration of phosphorylated AgrA and 10 nM of a 214 base pairs long γ-32P-labelled DNA fragment representing the *agr* operon intergenic region including the P2 and P3 core promoter sequences up to position +11 of the P2 promoter and +15 of the P3 promoter (See Figure 81). The phosphorylated AgrA was incubated in reaction buffer for 10 min at 37°C before adding the DNA probe. The reactions were incubated at 37°C for another 10 min and stopped with native gel loading dye (reaction buffer + 50% (v/v) glycerol and 0.05% (w/v) bromophenolblue) and resolved on a 4-20% (w/v) gradient native polyacrylamide gel (BIO-RAD). The gel was run in 1x TG buffer (National diagnostics) using the Mini-PROTEAN Tetra Electrophoresis system (BIO-

RAD) and pre-warmed by running at 100 V for 30 min before loading the samples. The gel was run at 100 V for 2 hour and 20 min before drying for 1 hour on Whatmann paper (Whatmann) using a gel dryer (Scie-Plas-GD4534) and pump (KNK lab).The dried gel was visualized and quantified using a GE Typhoon FLA 2000 Phosphor Imager and the Image Quant TL software, respectively.

2.4.3. Electrophoretic gel mobility shift assay (EMSA) with AgrC-HK activation

These were conducted exactly as in the previous section 2.4.2 but without the pre-incubation step. Instead, 10 µl reactions were set up in reaction buffer with 200 nM AgrA, 800 nM AgrC-HK and 0.5 mM dATP. The rest was exactly as in the previous section.

2.4.4. Short RNA synthesis assay with AgrC-HK activation

Short RNA synthesis assays measure the transcriptional output of RNAp promoter complexes. By providing the RNAp with only a short dinucleotide primer complementary to the promoter template strand transcriptional start site and also only the next complimentary nucleotide in radiolabeled form only short RNA synthesis can occur.

10 µl binding reactions were setup. 50 nM *S. aureus* core RNAp and 200 nM *S. aureus* σ^A were incubated for 5 min at 37°C in presence of 0.25 µl RNasin® Plus RNase Inhibitor (Promega) in reaction buffer. Simultaneously, 200 nM of AgrA was incubated for 5 min at 37°C in presence of 800 nM AgrC-HK, 500 nM dATP and 10 nM P3-IR promoter probe (See Figure 81). After the 5 min incubation, both reactions were mixed and incubated for another 1 min at 37°C before the addition of the elongation mix (0.5 mM GpA dinucleotide primer, 3 µCi α-^{32}P-UTP (800 Ci/mmol-PerkinElmer) and 100 µg/ml heparin) to allow dinucleotide primed short RNA synthesis to occur. Reactions were incubated at 37°C for a further 10 min before the addition of 2 µl of transcription stop dye (3% (w/v) xylene cyanol, 3% (w/v) bromophenol blue and 20 mM EDTA in deionized formamide). Reactions were loaded and resolved on a 20% (w/v) urea denaturing gel (20 ml system concentrate, 2.5 ml system dilutent, 2.5 ml 10x TBE (Lonza), 150 µl 10% APS, 20 µl TEMED) using a Mini-PROTEAN Tetra Electrophoresis system. The gel was pre-warmed by running at 200 V for 30 min before loading the samples. The gel was run at 200 V for 50 min. The wet gel was visualized and quantified using a GE Typhoon FLA 2000 Phosphor Imager and the Image Quant TL software, respectively.

2.4.5. Full length *in vitro* transcription assay with acetyl phosphate activation

Full length transcription assays were performed to measure the formation and transcriptional output of RNAp promoter complexes formed on supercoiled templates over time.

Phosphorylation of AgrA was carried out by pre-incubating 3 µM AgrA with 50 mM acetyl phosphate and 5 mM $MgCl_2$ for 1 h at 37°C. 10 µl microliter binding reactions were set up in transcription buffer (40 mM Tris–acetate (pH 7.9), 100 mM NaCl, 20 mM $MgCl_2$ and 0.2 mM DTT). 50 nM *S. aureus* core RNAp and 200 nM *S. aureus* σ^A, 1 µM AgrA, 10 nM pJR_{P2+P3} (See Table 9) and 0.25 µl RNasin® Plus RNase Inhibitor were incubated at 37°C for 5 min separately and then mixed and incubated for another 5 min. RNA synthesis was initiated by adding an elongation mix containing 0.5 mM ATP, CTP and GTP; 0.25 mM UTP; 0.75 µCi of α-^{32}P UTP; 1 µl of 100 µg/ml heparin per reaction was added to the elongation mix for single-round transcription reactions. The reaction was incubated for a further 10 min at 37 °C. The reactions were stopped with transcription stop dye and resolved on a 10% (w/v) urea-denaturing polyacrylamide gel (28 ml system concentrate, 35 ml system dilutent, 7 ml 10x TBE, 560µl 10% APS, 28 µl TEMED). The gel was cast and run using the Sequi-Gen Sequencing cell (BIO-RAD) in 1x TBE. The gel was pre-run at 40 W for 1 hour or until the gel running apparatus had reached the optimal running temperature of 80°C. The samples were then loaded and the gel run at 40 W for 3h until the bands were sufficiently resolved, based on the loading dye position. (on a 10% urea gel, bromophenol blue runs like a 10 nucleotide product and xylene cyanole like a 55 nucleotides product).The gel was then dried for 3 h on Whatmann paper using a gel dryer and pump.The dried gel was visualized and quantified using a GE Typhoon FLA 2000 Phosphor Imager and the Image Quant TL software, respectively.

2.4.6. DNA bending assays

The DNA bending plasmid pAM1847 containing the AgrA binding site (Reyes et al, 2011) was digested with the *EcoR*I, *Hind*III, *BstN*I, *EcoR*V, *Nhe*I and *BamH*I restriction enzymes (section 2.2.4). The fragments were purified (section 2.2.6) and individually labeled with γ-^{32}P ATP (section 2.4.1). 7 ng of the labeled fragments were then incubated with 1 µM of phosphorylated AgrA in reaction buffer with 0.3 µg of sonicated calf thymus (Sigma) at RT for 20 min. The reactions were stopped with native loading dye and resolved on a 4.5% (w/v) native polyacrylamide gel (3 ml Solution I, 2 ml 10x TG, 15 ml H_2O, 200 µl APS, 20 µl TEMED). The gel was run in 1x TG buffer (National diagnostics) using the Mini-PROTEAN Tetra Electrophoresis system (BIO-RAD) and pre-warmed by running at 100 V for 30 min

before loading the samples. The gel was run at 100 V for 1 hour and 40 min before drying for 30 min on Whatmann paper using a gel dryer and pump. The dried gel was visualized and quantified using a GE Typhoon FLA 2000 Phosphor Imager and the Image Quant TL software, respectively.

2.4.7. Acetyl Phosphate phosphorylation of AgrA

γ-^{32}P acetyl phosphate was synthesized with 0.3 unit of acetate kinase (Sigma) and 10 µCi γ-^{32}P-ATP in 7.5 µl of AKP buffer (25 mM Tris–HCl pH 7.4, 60 mM KOAc, 10 mM $MgCl_2$; final pH 7.6) in a 10 µl reaction for 20 min at RT (Quon et al, 1996). 1.46 µM AgrA was then added to the γ-^{32}P acetyl phosphate and the reaction was incubated at 37°C for 1 hour before stopping with the addition of 2x Laemmli buffer. The samples were loaded on a 12.5% SDS-PAGE gel and the gel was run at 200 V for 50 min (section 2.3.3). The gel was dried for 1 hour on Whatmann paper using a gel dryer and pump. The dried gel was visualized and quantified using a GE Typhoon FLA 2000 PhosphorImager and the Image Quant TL software, respectively.

2.4.8. AgrC-HK phosphorylation of AgrA

Two separate 10 µl reactions were set up. The first one contained 5µM AgrC-HK, 5 µCi γ-^{32}P-ATP, 2 mM DTT, 5 mM $MgCl_2$, 500 nM dATP and 5 mM DTT in HK storage buffer and the second one 10 µM AgrA and 5 mM $MgCl_2$ in HK storage buffer. Both reactions were incubated for 1 hour at 30°C then 5 µl were taken from each reaction and incubated together in a new reaction tube for 1 hour at 30°C. The reactions were stopped with the addition of Laemmli buffer and loaded on a 10% SDS-PAGE gel. The gel was run at 200 V for 50 min (section 2.3.3) and dried for 1 hour on Whatmann paper using a gel dryer and pump. The dried gel was visualized and quantified using a GE Typhoon FLA 2000 Phosphor Imager and the Image Quant TL software, respectively.

2.5. *In silico* method

2.5.1. Molecular Dynamics (MD) Simulations (performed by Dr. Robert Weinzierl)

The protonation states of residues present in the LytTR DNA-binding domain in PDB 4G4K (Sidote et al, 2008) were computed using H++ (Gordon et al, 2005). The resulting atomic coordinates were embedded in a TIP3P (Jorgensen, 1983) cubic water box extending for a minimum of 15 Å beyond the protein structure. The assembly was charge-neutralized, adjusted to 150 mM NaCl and parameterized with the AMBER99SB (Ponder & Case, 2003)

force field. Simulations were run on Nvidia GPUs certified to carry out accurately reproducible calculations using the AMBER 12 package (Salomon-Ferrer, 2013). The parameterized structures were energetically minimized by rapid descent and repeated annealing minimization before equilibration at 300 Kelvin and one atmosphere of pressure (NPT ensemble). For all MD simulations, the bond length between hydrogen and heavy atoms was fixed using the SHAKE algorithm (Ryckaert, 1977). The time-step was 2 femtosec and the Particle Mesh Ewald (Essmann, 1995) cut-off distance was set to 8 Å. The simulation engine was pmemd.cuda, a CUDA-accelerated MD production engine based on a hybrid single-/double precision (SPFP) algorithm. Conventional MD simulations were initially carried out for 50 nanosec (ns) to obtain representative values for the total potential and dihedral energy values for adjustment in the subsequent dual-boost accelerated MD (aMD) production runs. For boosting the dihedral potential, an energy contribution of 3.5 kcal/mol/residue per degree of freedom was assumed. A boost factor (α) factor of 0.2 was used for adjustment of the dihedral and potential energy in all aMD simulations. All aMD simulations lasted 100 ns, but are likely – due to the acceleration - to represent protein motions occurring in the tens to hundreds of microsec range in real time (Pierce et al, 2012). Trajectories were processed, visualized and analyzed using the Visual Molecular Dynamics (VMD) suite (Humphrey et al, 1996).

CHAPTER THREE

3. Development of an experimental system to study AgrA activity *in vivo*

3.1. Introduction

3.1.1. The P1, P2 and P3 promoters

In the early days of study of the *agr* operon, two promoters able to drive the expression of AgrA were identified, P1 and P2 (See Figure 33 and Figure 34). P1 is a weak constitutive promoter located about 100 bp upstream of *agrA,* in the 3'end of *agrC* (Peng et al, 1988), P2 is a strong promoter located about 2kb upstream of *agrA* in the *agr* IR, more active in the post-exponential phase of growth responsible for the transcription of RNAII (Janzon et al, 1989; Morfeldt et al, 1988).The P3 promoter is a strong promoter also located in in the *agr* IR but on the opposite strand from the P2 and P1 promoters and responsible for the transcription of RNAIII, the effector RNA of the *agr* operon involved in the temporal regulation of virulence factors (See Figure 33 and section 1.5 for more details). P2 and P3 are divergent promoters (Janzon & Arvidson, 1990; Morfeldt et al, 1988; Novick et al, 1993). Both the P2 and the P3 promoters are activated by AgrA (Janzon et al, 1989).

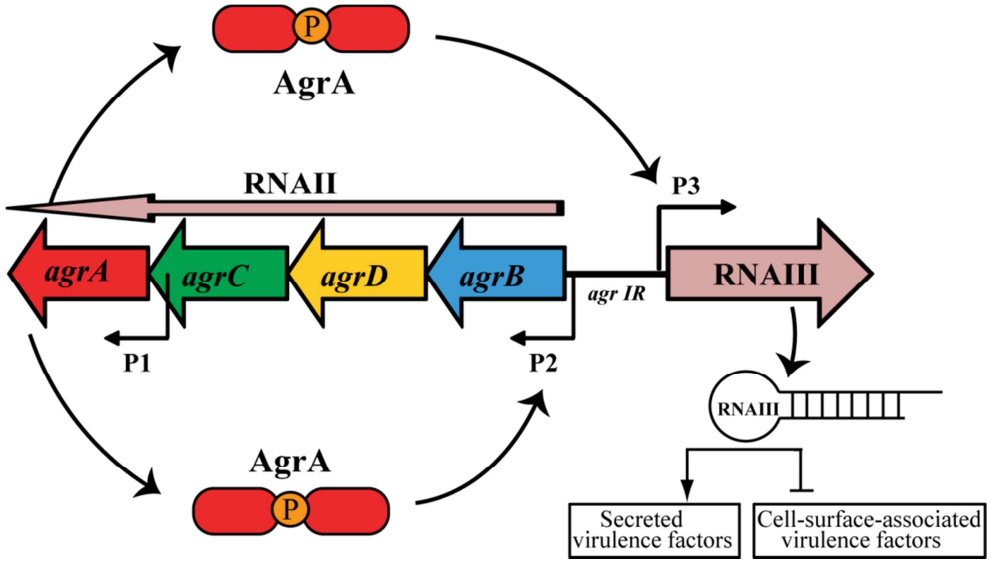


Figure 33: Schematic of the P1, P2 and P3 promoters of the *agr* operon.

The P2 and the P3 promoter are activated by the phosphorylated AgrA. P2 encodes the RNAII transcript consisting of the *argB*, *agrD*, *agrC* and *agrA* genes. P3 encodes RNAIII, the pleiotropic effector RNA of the operon responsible for the upregulation of the secreted virulence factors and the downregulation of the cell surface-associated virulence factors.

As mentioned in section 1.5.5.3, the -10 and -35 promoter elements of the P2 and P3 promoters have been mapped based on the start of the RNAII and RNAII transcripts (Janzon

et al, 1989; Novick et al, 1993). However, the transcriptional start site of the P1 promoter has not been determined and therefore the precise position of the P1 -10 and -35 promoter elements is also unknown.

P1 driven AgrA was shown to weakly complement *agrA* activity in an *agr* deficient strain (Peng et al, 1988) whereas P2 driven AgrA was able to fully complement under similar conditions (Morfeldt et al, 1988). Both promoters were mentioned in the literature until the mid-nineties when it became widely accepted that P2 was indeed the primary *agrA* promoter (Novick et al, 1995) and articles mentioning P1 became scarce. However, P1-*agrA* constructs (See Figure 34) were used to complement AgrA activity until much later, with more or less success depending on the *S. aureus* strain used (Adhikari et al, 2007; Traber & Novick, 2006; Tsompanidou et al, 2011). It can be noted that none of these complementation restored the full AgrA activity. However, the best results were obtained in the strains of the NCTC 8325 lineage bearing a mutation in *rbsU*, encoding a phosphatase responsible for the activation of the *S. aureus* alternative sigma factor σ^B (Kullik et al, 1998). This is interesting as σ^B is thought to antagonise *agr* activity (Novick, 2003). Indeed, SH1000, the *rbsU* repaired strain has been shown to display important differences in growth patterns and exoproteins production, the latter resulting from a decreased *agr* activity, compared to other strains of the NCTC 8325 lineage (Horsburgh et al, 2002; Novick, 2003).

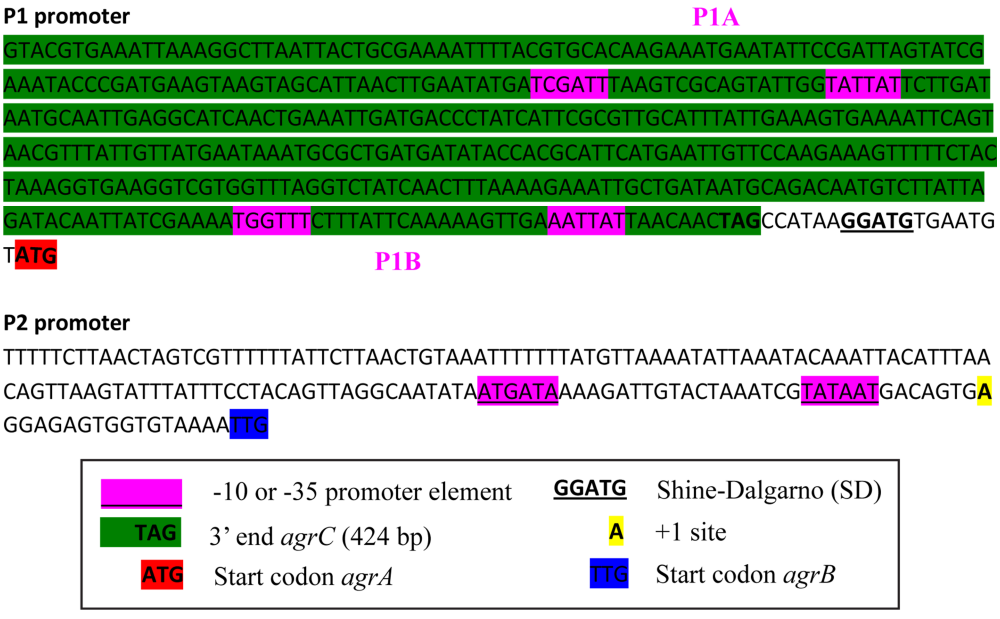


Figure 34: The P1 and the P2 promoter.

DNA sequence of the P1 promoter used to complement for AgrA activity by Peng et al. and the P2 promoter. -35 refers to the -35 element (5'-TTGACA-3') and -10 refers to the -10 element (5'-TATAAT-3'). The P1 promoter -10 and -35 promoter elements have not been identified. Therefore, the two putative -10 and -35 promoter elements present in the sequence are highlighted in magenta and labelled in magenta but not underlined. The

sequence of the P2 promoter contains the *agr* intergenic region, starting immediately after the -35 promoter element of the P3 promoter.

3.1.2. The *S. aureus* haemolysins

The haemolytic activity of *S. aureus* on blood plates was first observed over a century ago (Julianelle, 1922). It is primarily mediated by the ability of the staphylococcal haemolysins α, β, γ and δ to lyse red blood cells (Coleman et al, 1986). The vast majority of *S. aureus* strains contain all the corresponding genes, *hla*, *hlb*, *hlg* and *hld*, respectively (Vandenesch et al, 2012). However, *hlb* is a preferential site for bacteriophage insertion. Therefore, the large majority of human clinical isolates are β-haemolysin-negative (van Wamel et al, 2006). The activity of α and δ haemolysins on blood agar plates can be used as a surrogate for *agr* function (Adhikari et al, 2007) and all four haemolysins are under the control of the *agr* operon (Bronner et al, 2000; Recsei et al, 1986)

3.1.2.1. α-haemolysin

α-haemolysin (α-toxin) has been extensively studied for about 100 years. α-haemolysin monomers are secreted and integrate into the target cell membrane to form a cylindrical heptamer leading to the formation of a pore. This pore allows the influx/efflux of ions and small molecules leading to osmotic swelling and eventually cell death (Dinges et al, 2000).

3.1.2.2. β-haemolysin

β-haemolysin (Sphyngomyelinase C) is also called the "hot-cold" haemolysin because its haemolytic activity is enhances by incubation below 10°C following incubation at 37°C (Bigger, 1933). It is a sphyngomyelinase and hydrolyses sphingomyelin, a plasma membrane lipid (Doery et al, 1963). Sphingomyelin is rich in cholesterol and sphyngomyelinase treatment of the lipid bilayer leads to aggregation of cholesterol. The exact mechanism leading to cell death remains unclear but it is thought that it might be due to destabilisation of the plasma membrane (Ira & Johnston, 2008; Vandenesch et al, 2012)

3.1.2.3. γ-haemolysin

γ-haemolysin is a bi-component toxin consisting of two non-associated secreted proteins, termed S (slow) and F (fast) (Woodin, 1959). γ-haemolysin cannot be identified on blood agar plates as its activity is inhibited by agar (Prevost et al, 1995). The γ-haemolysin proteins bind sequentially to the host cell surface as monomer then oligomerise into an octamer consisting of four S elements alternating with four F elements. The mature octamer forms a pore across the membrane leading to host cell lysis (Vandenesch et al, 2012).

3.1.2.4. δ-haemolysin

The δ-haemolysin (δ-toxin) open reading frame is located at the 5' end of RNAIII and is therefore closely linked with the *agr* operon (Janzon et al, 1989). It is secreted and forms an α-helix with hydrophobic and hydrophilic domains on opposite sides (Freer & Birkbeck, 1982). There are three hypotheses explaining its haemolytic activity. δ-haemolysin could bind to cell surface and aggregate to form a pore, bind to the cell surface and destabilise the plasma membrane or act as a detergent and solubilize the membrane (Verdon et al, 2009).

3.1.3. The L54a phage and pCL55

The *S. aureus* structural lipase gene (glycerol ester hydrolase: *geh*) is negatively regulated by lysogenic conversion by the prophage L54a (Lee & Iandolo, 1985). This is a temperate prophage which is capable of both lysogenic and lytic life style (Duval-Iflah, 1972). While in a lysogenic lifestyle, it inactivates the extracellular lipase of its host by integrating on the *S. aureus* chromosome thus disrupting the lipase structural gene (Lee & Iandolo, 1986b). Curing of the phage restores the lipase activity of the strain (Lee & Iandolo, 1985). The integrative recombination of L54a occurs between the specific viral attachment site (*att*P) on the bacteriophage genome and the *att*B site on the host chromosome. This leads to the creation of two attachment sites, *att*L and *att*R generated on the left and right side of the insertion. The recombination is not only site specific but also orientation specific. The four *att* sites share an 18 bp identical core region. The *att*B site is localised near the 3'end of the *geh* open reading frame (Lee & Iandolo, 1986a). Thus the lysogens will display a lipase negative activity as the insertion disrupts aa 635 to 644 of the 690 aa enzyme. This activity can be tested on egg yolk agar plates (Lee & Iandolo, 1985). The integration of the phage requires the viral integrase (*int*) gene whereas the excision necessitates the integrase and excisionase (*xis)* viral genes (Ye & Lee, 1989).

pCL55 is a single copy integration vector based on the recombination system of the prophage L54a. It is capable of autonomous replication in *E. coli* but not in *S. aureus* as it carries a single origin of replication solely compatible with *E. coli*. This plasmid contains both the *att*P and the *int* sites enabling its integration on the chromosome of the host at the *geh* locus in an orientation specific manner (See Figure 35). Integration can be easily tested as correct integration will produce a lipase negative phenotype. Because the plasmid does not contain the *xis* gene necessary for excision it can be stably maintained even without selective pressure. Therefore it provides a straight forward method to integrate a single copy of the desired gene on the chromosome of *S. aureus* (Lee et al, 1991).

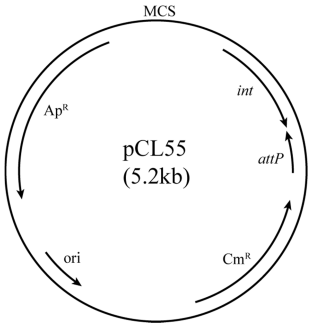


Figure 35: pCL55 map.

Ap^R represent the ampicillin resistance gene. *int* represents the integrase gene. *att*P represents the viral attachment site. Cm^R represent the chloramphenicol resistance gene. ori represents the origin of replication. Adapted from (Lee et al, 1991).

The pCL55 plasmid was used to construct a GFP transcriptional fusion in the laboratory by Ellen James. The P3 promoter and the *agr* intergenic region (IR) up to, but not including, the -35 promoter of the divergent P2 promoter were cloned upstream of the gene encoding GFP with an optimal Shine-Dalgarno (SD) sequence thus creating the pCL55 $_{agr\ \text{IR P3-GFP}}$ (See Figure 36). This construct enabled to quantify the level of activation of the P3 promoter by measuring the GFP fluorescence emitted (James et al, 2013).

TATCATTATATTGCCTAACTGTAGGAAATAAATACTTAACTGTTAAATGTAATTTGTATTTAATATTTTAACATA

AAAAAATTTACAGTTAAGAATAAAAAACGACTAGTTAAGAAAAA**TTGGAA**AATAAATGCTTTTAGCATGT**TTTAA**

-35$_{P3}$ -10$_{P3}$

+1

TATAACTAGATCACAGAGATGTGATCCT**AAGGAGG**AAAAAAAATG

SD GFP

Figure 36: Sequence of the *agr* IR P3-GFP of the pCL55 $_{agr\ \text{IR P3-GFP}}$ construct.

The *agr* intergenic region, starting immediately after the -35 promoter element of the P2 promoter and stopping upstream of RNAIII of the pCL55 $_{agr\ \text{IR P3-GFP}}$ construct. -35 refers to the -35 element (5'-TTGACA-3') and -10 refers to the -10 element (5'-TATAAT-3').The bold boxes highlighted in red represent the -35 and -10 promoter element of the P3 promoter, the +1 site is highlighted in yellow, the Shine-Dalgarno sequence is bold highlighted and the start codon of the GFP reporter is highlighted in green.

The stability of the integrated construct was tested and after 24 h of growth in TSB without antibiotic selection, and all the colonies tested had retained the construct. To test whether the presence of a second copy of the *agr* IR affected the expression of the *agr* operon, mRNA level of *agrA* (P2 transcription), *hld* (P3 transcription) were compared relative to *gyrB* (constitutively expressed gyrase) by qRT PCR. No obvious differences in the level of mRNA of *agrA* and *hld* were detected between the strain bearing the reporter and an empty strain. Blood agar haemolysis was used to illustrate that the reporter strain was not compromised for β-haemolysis. In addition, the growth rates of the reporter strain in TSB were very similar to that of the empty strain (James et al, 2013).

Therefore it was concluded that the presence of the GFP reporter does not affect the strain in any way and that construct can be used to portray the regulatory events taking place at the native *agr* operon.

3.2. Objectives

The overall objective of this chapter was to develop tools to study transcription activation by AgrA *in vivo*. In this regard, the specific aim included:

I. Purify AgrA to send for antibody production then purification of the polyclonal antibody obtained
II. Characterise the SH1000^{-} *agr*-deficient strain
III. Development of a method to measure transcription activation at the P3 promoter by AgrA *in vivo*
IV. Complement the SH1000^{-} *agr*-deficient strain

3.3. Polyclonal AgrA antibody production and purification

AgrA was expressed under the control of the pET expression system, as in section 2.3.1.1 in the *E. coli* BL21 (DE3) expression strain (See Table 8) without induction, as in section 2.3.1.2. The vector pET28b^{+} encoding an N-terminal 6His tag was used. The [6His]-AgrA was then purified by a single step nickel affinity chromatography, as in section 2.3.4.1. The SDS-PAGE gel shown in Figure 37 contains the samples taken during the purification and illustrates that the [6His]-AgrA was eluted from the nickel column at high concentration (See Figure 37, lanes 13-14). The most concentrated elution fractions, lanes 13 to 15 were collected and dialysed into antibody storage buffer. The total [6His]-AgrA concentration obtained was of 1.4 mg and was sufficient to send for antibody production (Eurogentec, see section 2.3.6).

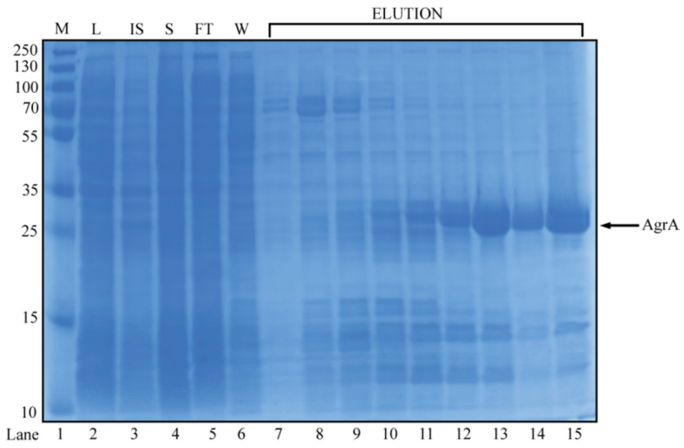


Figure 37: Nickel affinity chromatography of [6His]-AgrA.

Image of a 12.5% (w/v) SDS-PAGE gel from the nickel affinity purification of [6His]-AgrA. M = Marker, L = whole-cell lysate, IS = insoluble fration, S = soluble fraction, FT = flow-through and W = low imidazole wash. The molecular weight of the marker is shown in kDa and the migratory position of [6His]-AgrA is indicated with an arrow.

The polyclonal serum of both of the immunised rabbits was tested by Western Blot (section 2.3.8) in order to select the best candidate for purification. (See Figure 38)

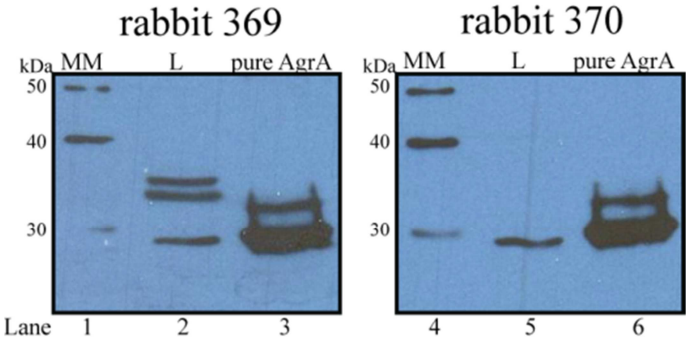


Figure 38: Western blot of rabbits 369 and 370.

Western blot analysis comparing the affinity of the serum from each rabbit for AgrA. MM = magic marker and L = whole cell lysate. The lane number and kDa of the magic marker bands are indicated.

The serum fron rabbit 370 recognised AgrA in whole cell lysate with high affinity and only one band specific for AgrA was detected (See Figure 38, lane 5). Therefore, an aliquot of serum from rabbit 370 was purified against pure untagged AgrA, as described in section 2.3.7. The purified polyclonal antibody against AgrA was tested by Western Blot (as in section 2.3.8) and compared to the serum (See Figure 39). The purified antibody displayed a cleaner signal than the serum.

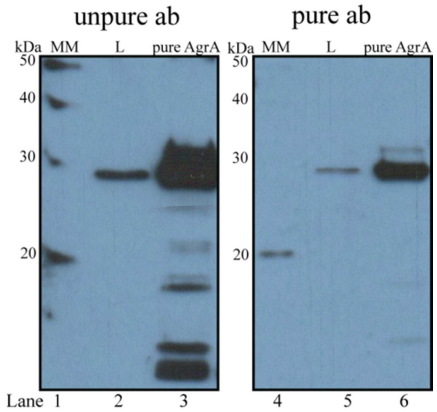


Figure 39: Western blot comparing the affinity of the serum and the polyclonal antibody against AgrA for AgrA.

Western blot analysis comparing the specificity of the serum and the polyclonal antibody against AgrA for AgrA. MM = magic marker and L = whole cell lysate. The lane number and kDa of the magic marker bands are indicated.

3.4. Characterisation of the SH1000$^-$ *agr* operon-deficient strain

Initially we acquired a mutant *S. aureus* strain displaying an *agr*–defective phenotype that we were kindly given by Jan Maarten van Dijl from the University of Groningen, Netherlands. This strain, called SH1000$^-$, had been phenotypically characterised as lacking haemolytic activity and successfully complemented with a plasmid-borne ArgA, suggesting a mutation in the *agrA* gene (Tsompanidou et al, 2011). However, the spontaneous mutation in the *agr* operon responsible for the phenotype had not been identified. Thus, the *agrC* and *agrA* genes of this strain were sequenced, as they are the mutational hotspots of the operon, and the mutation identified by Ellen James. The point mutation at aa H174 in *agrA* resulted in a non-synonymous change from a histidine to a leucine residue. This position had been previously identified by Sidote et al (Sidote et al, 2008) as important for maintaining the overall structure of the protein through salt bridges interactions with two glutamate residues at position 226 and 163 (See Figure 40).

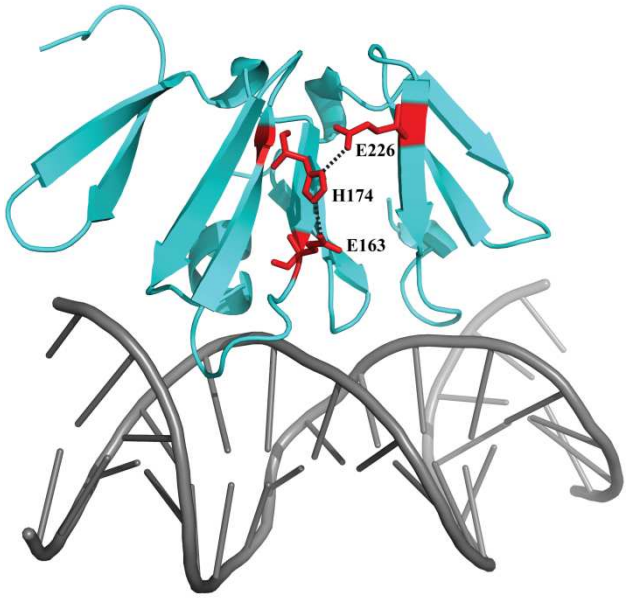


Figure 40: Salt bridges between H174, E226 and E163 on the AgrA LytTR domain crystal structure.

Ribbon representation of the AgrA LytTR domain-DNA complex crystal structure (the LytTR domain and DNA are coloured in cyan and grey, respectively). H174, E163 and E226 are highlighted in red, labelled in black and the salt bridge interactions indicated by dotted lines. Pymol, PDB 3BS1.

The growth rate of the mutant strain was tested and compared to that of the WT SH1000 strain by growth curve in a multiwell plate reader, as in section 2.1.2 (See Figure 41).

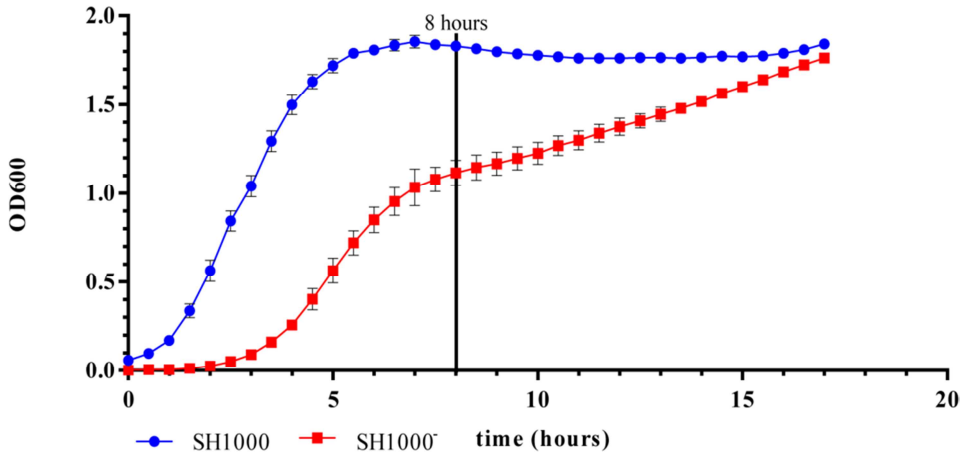


Figure 41: Growth curve of SH1000 and SH1000^{-}.

Graph showing growth curves ($OD_{600\ nm}$) of the *S. aureus* SH1000 (blue) and SH1000^{-} (red) strains grown in TSB. The black line highlights the 8 h' time point corresponding to the beginning of the stationary phase for the SH1000 strain. Data obtained from three biological replicates.

The SH1000^{-} strain displayed a different growth pattern consisting of a longer lag phase and exponential phase as the cells did not seem to reach stationary phase but kept on growing linearly. This growth pattern seems to indicate a deficiency in quorum sensing consistent with a defective *agrA*.

In order to determine whether the mutant AgrA was stably produced, whole-cell extract from the SH1000^{-} strain were tested by Western blotting using the polyclonal antibody raised against AgrA and compared with the WT (WT) strain SH1000 (See Figure 42 and section 3.3).

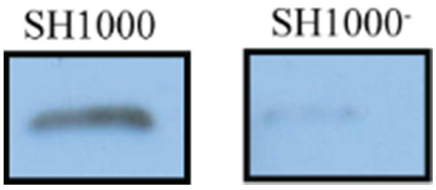


Figure 42: Western blot of SH1000 and SH1000^{-}.

Western blot analysis comparing the AgrA protein level between the SH1000 and SH1000^{-} strains in whole-cell lysates using a polyclonal antibody against AgrA (for full image, see Figure 51).

The SH1000^{-} mutant $AgrA_{H174L}$ (Figure 42, box 2) was barely detectable when compared with WT AgrA from SH1000 (Figure 42, box 1) that was expressed 5 fold higher.

To assess whether the $AgrA_{H174L}$ mutant retained the ability to activate transcription at the P3 promoter, the SH1000^{-} strain and the WT SH1000 strain were doted on blood agar plates (See Figure 43 and section 2.1.3).

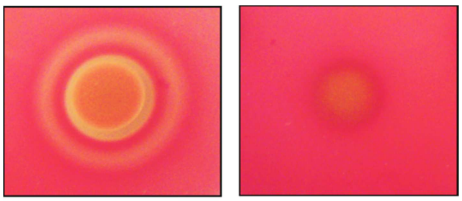


Figure 43: Blood agar plate of SH1000 and SH1000$^-$.

Sheep blood agar haemolysis assay comparing the haemolytic activity of the SH1000$^-$ and SH1000 strains. Data obtained from three biological replicates, representative image shown.

The production of haemolysins responsible for the characteristic β-haemolysis of blood, visible on blood agar plates as a clear halo around the colony in strain SH1000 (See Figure 43, box 1), is abolished in the absence of AgrA. Therefore and as expected, the SH1000$^-$ strain (See Figure 43, box 2) produced no such halo when plated on blood agar meaning that the mutant $AgrA_{H174L}$ is unable to activate transcription at the P3 promoter.

After characterising the SH1000$^-$ strain we can indeed conclude that it is an *agrA* deficient strain. Moreover, we can hypothesise that the salt bridges interactions between residues H174, E226 and E163 in the AgrA LytTR domain are important for maintaining the activity of the protein.

3.5. Development of a method to measure transcription activation at the P3 promoter by AgrA *in vivo*

As shown in the previous section (section 3.4), the ability of AgrA to activate transcription at the P3 promoter can be assessed *in vivo* using blood agar plates. However, this does not target the P3 promoter itself but rather the ability of RNAIII to activate transcription of the haemolysins, adding an extra regulatory step. In order to look at expression of the P3 promoter itself directly, a P3-GFP reporter approach was used. The $pCL55_{agr\ IR\ P3\text{-}GFP}$ reporter has been developed and tested previously in the laboratory (See Table 9 and James *et al*, 2013). This system enables to quantify the level of activation of the P3 promoter by AgrA through GFP readings. Thus, the $pCL55_{agr\ IR\ P3\text{-}GFP}$ was transduced in the SH1000 and SH1000$^-$ strains by Ellen James (See Figure 44).

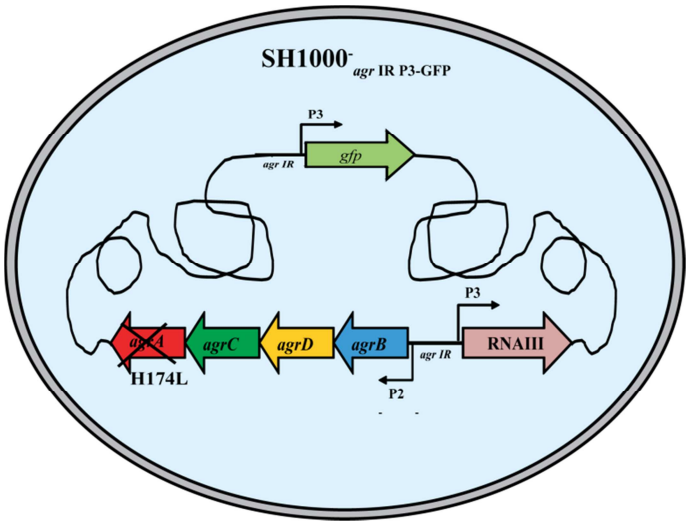


Figure 44: Schematic representation of the SH1000⁻ *agr* IR P3-GFP strain.

Schematic representation of the SH1000⁻ *agr* IR P3-GFP strain (with defective *agrA*$_{H174L}$ on its chromosome) containing the pCL55 *agr* IR P3-GFP reporter inserted ectopically on the chromosome at the *geh* locus. There are two copies of the P3 promoter, one at the *geh* locus and one at the *agr* operon. The *agrA* gene of the SH1000⁻ *agr* IR P3-GFP strain bears the H174L mutation.

Both strains were then used in growth curves in triplicates using a multiwell plate reader (See section 2.1.2) for 17 h, and the $OD_{600\ nm}$ and the GFP fluorescence units (GFP-Fluo) were measured every 30 min. To take into account the differences in growth rates, the GFP readings were normalised according to the $OD_{600\ nm}$ (See Figure 45).

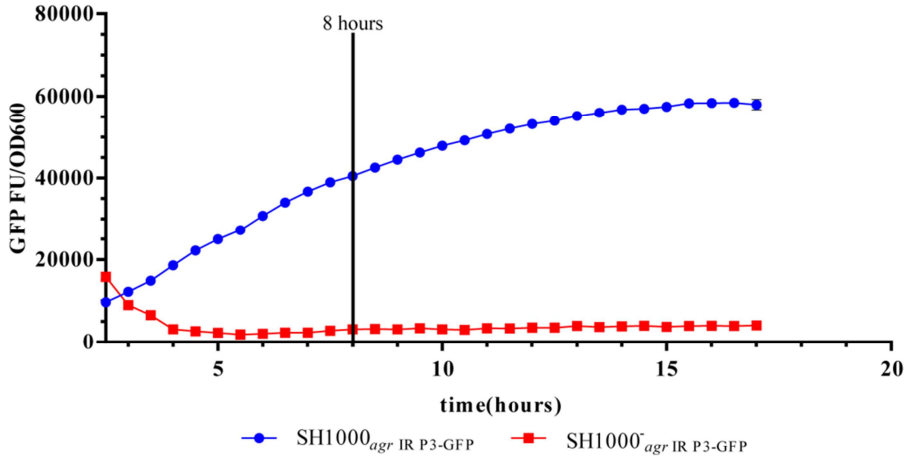


Figure 45: GFP growth curve of SH1000 *agr* IR P3-GFP and SH1000⁻ *agr* IR P3-GFP.

Graph showing GFP expression [as GFP fluorescence units (GFP-Fluo)] as a function of growth ($OD_{600\ nm}$) over time for the *S. aureus* SH1000 *agr* IR P3-GFP (blue) and SH1000⁻ *agr* IR P3-GFP (red) strains grown in TSB. The black line highlights the 8 h' time point corresponding to the early stationary phase for the SH1000 *agr* IR P3-GFP strain. Data obtained from three biological replicates.

As expected from the results obtained in section 3.4, the SH1000 *agr* IR P3-GFP strain displayed increasing GFP fluorescence over time as AgrA was able to activate transcription at the P3 promoter leading to the production of GFP. On the contrary, SH1000⁻ *agr* IR P3-GFP displayed a

constant low level of fluorescence illustrating that $AgrA_{H174L}$ is unable to activate transcription at the P3 promoter. The 8 h' time point, highlighted in Figure 45, was taken as a representative time point as it is the beginning of stationary phase (See Figure 41) and provides a simplified representation of the results. From here onwards, the GFP growth curves will be represented as bar charts displaying the 8 h' time point (See Figure 46).

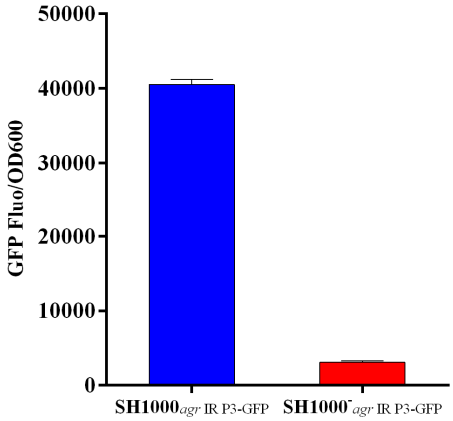


Figure 46: Bar chart of SH1000$_{agr\ IR\ P3\text{-}GFP}$ (blue) and SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ (red).

Graph showing GFP expression [as GFP fluorescence units (GFP-Fluo)] as a function of growth ($OD_{600\ nm}$) for the *S. aureus* SH1000$_{agr\ IR\ P3\text{-}GFP}$ and SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ strains grown in TSB for 8 h. Data obtained from three biological replicates.

3.6. Complementation of the SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ *agrA*-deficient strain

In section 3.5, the pCL55$_{agr\ IR\ P3\text{-}GFP}$ reporter was used successfully to illustrate the differences in P3 transcription between a strain containing a functional AgrA (SH1000 $_{agr\ IR\ P3\text{-}GFP}$) and one with a defective one (SH1000^{-} $_{agr\ IR\ P3\text{-}GFP}$). However, this was a clear cut result which does not give much indication about the sensitivity of the system and its ability to detect subtle variations in transcription activation. Therefore, the SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ was complemented using two different constructs in the high copy number shuttle vector, pCN34 (Charpentier et al, 2004). As mentioned in section 3.1.1, Tsompanidou et al attempted to complement the SH1000^{-} strain but only achieved a weak complementation because they were using a P1 driven a*grA*. We decided to try to complement using two different constructs: pSN:*agrC832-1261-agrA* containing *agrA* upstream of the 3' end of *agrC* (424bp) including the P1 promoter (See Figure 34) and pSN-P2-*agrA*, containing *agrA* under the control of the P2 promoter, the *agr* intergenic region (IR) and a strong Shine-Dalgarno sequence. The pSN-P2-$agrA_{H174L}$ was also constructed to assess whether the overexpression of the mutant $AgrA_{H174L}$ could lead to some activity. The SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ strain was transformed with all three constructs and their ability to complement the strain was assessed

by GFP growth curves (See Figure 47). The SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ and SH1000$_{agr\ IR\ P3\text{-}GFP}$ strains were also transformed with the empty pCN34 vector, pSN-P2-empty as a control.

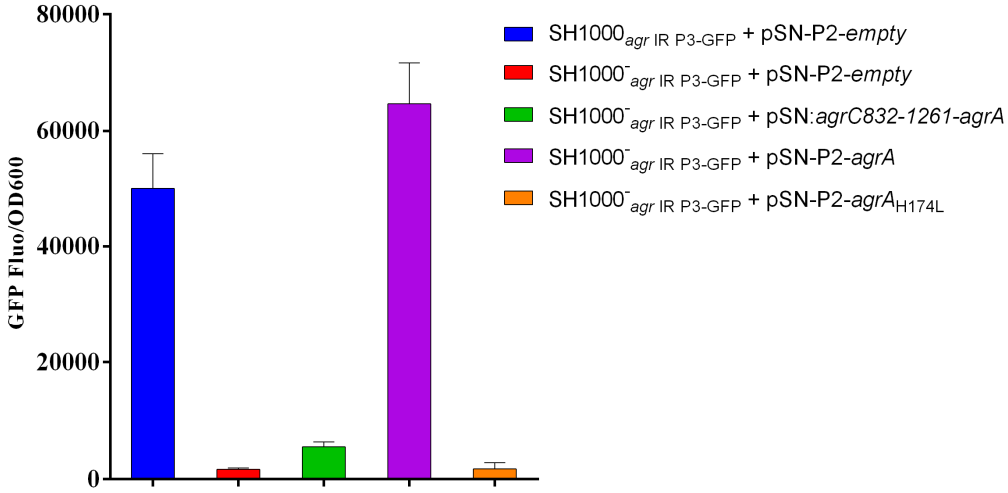


Figure 47: Bar chart of the different SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ complementations.

Graph showing GFP expression [as GFP fluorescence units (GFP-Fluo)] as a function of growth ($OD_{600\ nm}$) for the *S. aureus* strains SH1000$_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*empty* (blue), SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*empty* (red), SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN:*agrC832-1261-agrA* (green), SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*agrA* (purple)and SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*agrA*$_{H174L}$ (orange) grown in TSB for 8 h. Data obtained from three biological replicates.

The complementation with the pSN:*agrC832-1261-agrA* was weak, displaying only a 3-fold increase compared to the negative control and over 10-fold less transcription than the pSN-P2-*agrA* construct. This is not surprising as P1 driven AgrA was never shown to fully complement (see section 3.1.1). However, as expected, the pSN-P2-*agrA* construct was able to fully complement, displaying an even higher level of fluorescence than the WT strain SH1000$_{agr\ IR\ P3\text{-}GFP}$ with the pSN-P2-empty. This is due to the fact that AgrA is overexpressed on a high copy number plasmid, hence giving a higher signal because there are 25-30 additional *agrA* copies in the cell. But it is also an indication that in the SH1000$_{agr\ IR\ P3\text{-}GFP}$ strain, the system is not saturated; more AgrA can lead to more transcription. In addition, it indicates that the chromosomal copy of *agrA*$_{H174L}$ present in the SH1000^{-} strain does not act as a sink for the phosphates transferred by AgrC to activate AgrA. The pSN-P2-*agrA*$_{H174L}$ complementation displays more or less the same signal as the SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ strain with the pSN-P2-empty. This illustrates that even when overexpressed, AgrA$_{H174L}$ remains non-functional. For a clearer comparison, the activity of each strain was calculated compared to the SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ with pSN-P2-*agrA* set at 100% activity (See Figure 48). From now on, the GFP growth curves will be presented in comparison to the pSN-P2-*agrA* complementation, "WT" set at 100%.

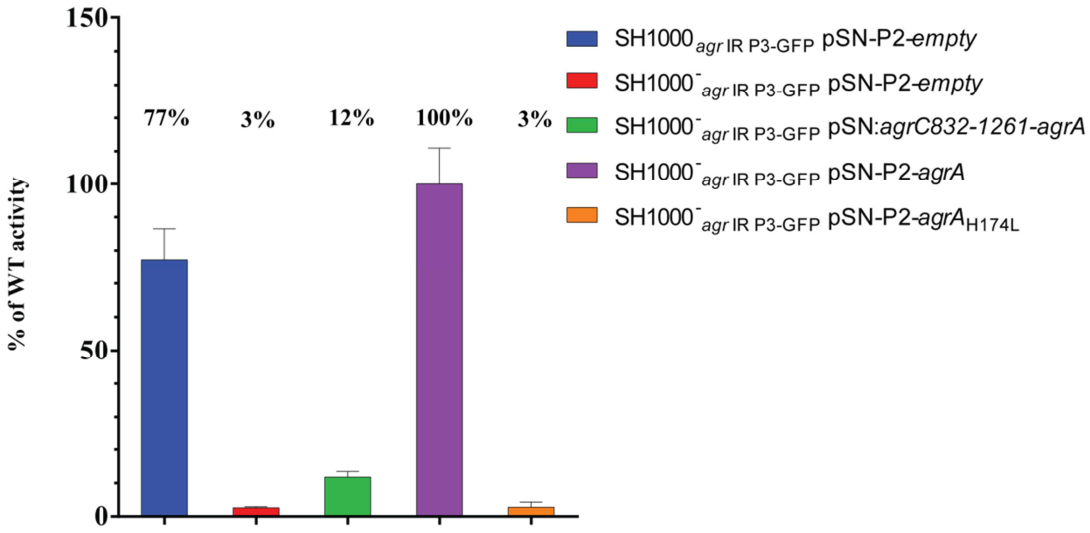


Figure 48: Bar chart of the different SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ complementations compared to the pSN-P2-*agrA* complementation.

Graph showing GFP expression [as GFP fluorescence units (GFP-Fluo)] as a function of growth ($OD_{600\ nm}$) for the *S. aureus* strains SH1000$_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*empty* (blue), SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*empty* (red), SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN:*agrC832-1261-agrA* (green) and SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*agrA*$_{H174L}$ (orange) compared to SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*agrA* (purple) set at 100%, grown in TSB for 8 h. Data obtained from three biological replicates.

The results obtained with the GFP growth curves were confirmed using blood agar plates and by western blot (See Figure 49, Figure 50 and Figure 51).

The result of the blood agar plate experiment was consistent with that of the GFP growth curves (See Figure 49). The SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ strain with the pSN-P2-empty or with the pSN-P2-*agrA*$_{H174L}$ produced no sign of haemolysis. The SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ strain with the pSN:*agrC832-1261-agrA* showed signs of weak haemolysis, which was only visible after incubation at 4°C, indicating that β-haemolysin was most likely responsible for the faint halo surrounding the colony (See section 3.1.2). The SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ with pSN-P2-*agrA* and SH1000$_{agr\ IR\ P3\text{-}GFP}$ strain with the pSN-P2-empty strains were both highly haemolytic, as expected.



1. SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN:*agrC832-1261-agrA*
2. SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*agrA*
3. SH1000$_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*empty*
4. SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*empty*
5. SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*agrA*$_{H174L}$

Figure 49: Blood agar plate complementation SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$.

Sheep blood agar haemolysis assay comparing the haemolytic activity of the for the *S. aureus* strains SH1000$_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*empty*, SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*empty*, SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN:*agrC832-1261-agrA*, SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*agrA*$_{H174L}$ and SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*agrA* . Data obtained from three biological replicates, representative image shown.

The western blot was performed using whole-cell lysate of each strain and without adding human immunoglobulin (HIgG). HIgG is routinely added when performing western blots (See 2.3.8) in *S. aureus* to prevent the staphylococcal immunoglobulin binding proteins Spa (protein A) and Sbi from binding to the antibodies and giving background signal. However, this step is not necessary when working with strains producing functional AgrA as both Spa and Sbi are negatively regulated by the *agr* operon (see section 1.5.6). Thus, in the SH1000$_{agr\ IR\ P3\text{-}GFP}$ strain with the pSN-P2-empty and the SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ with pSN-P2-*agrA* strain that are both producing a functional AgrA there was no sign of IgG binding proteins and only one band specific for AgrA was detected (see Figure 50, lanes 3 and 4). However, in the SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ strain with the pSN-P2-empty or with the pSN-P2-*agrA*$_{H174L}$ and in the SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ strain with the pSN:*agrC832-1261-agrA* , no specific band for AgrA could be distinguished. Instead there was a variety of bands of different sizes, ranging from 40 to 25 kDa, most likely corresponding to degradation products from Spa and Sbi (See Figure 50, lanes 3, 5 and 6). The band pattern observed for these three strains was similar to that of RN4220, which also contains a dysfunctional AgrA and cannot inhibit the IgG binding proteins (See Figure 50, lane 7 and section 1.5.6). This is yet another confirmation that neither the pSN-P2-*agrA*$_{H174L}$ or the pSN:*agrC832-1261-agrA* were not able to fully complement AgrA activity.

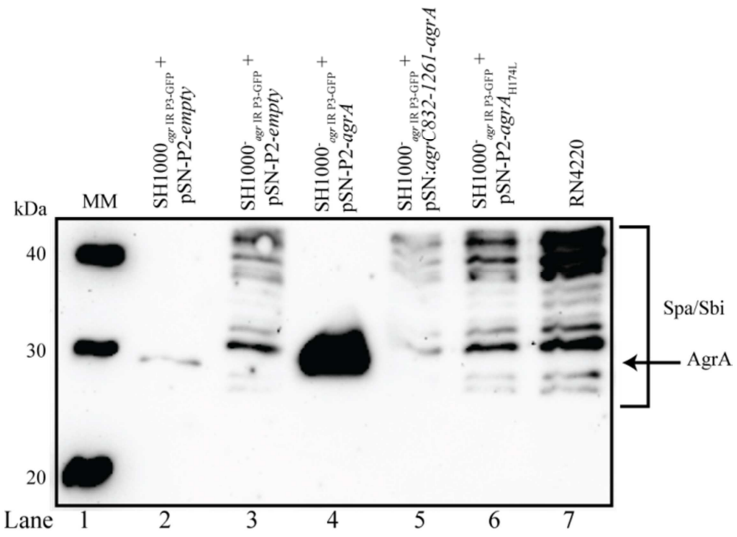


Figure 50: Western blot of SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ complementation, without HIgG.

Western blot analysis comparing the AgrA protein level in the *S. aureus* strains SH1000$_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*empty*, SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*empty*, SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN:*agrC832-1261-agrA*, SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*agrA*$_{H174L}$, SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*agrA* and RN4220. MM = magic marker. The lane number and kDA of the magic marker bands are indicated.

Another western blot was then performed using whole-cell lysate of each strain and in presence of HIgG and therefore no contaminating band was detected for any strains (See Figure 51). A faint band corresponding to AgrA$_{H174L}$ was detected for SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ strain with the pSN-P2-empty (See Figure 51, lane 2), a bigger band was detected in SH1000$_{agr\ IR\ P3\text{-}GFP}$ strain with the pSN-P2-empty (See Figure 51, lane 4). In SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ with pSN-P2-*agrA*, a very strong band was detected (See Figure 51, lane 1) and no protein was detectable in the agr operon deletion strain SH1001 $_{agr\ IR\ P3\text{-}GFP}$ (See Figure 51, lane 3).

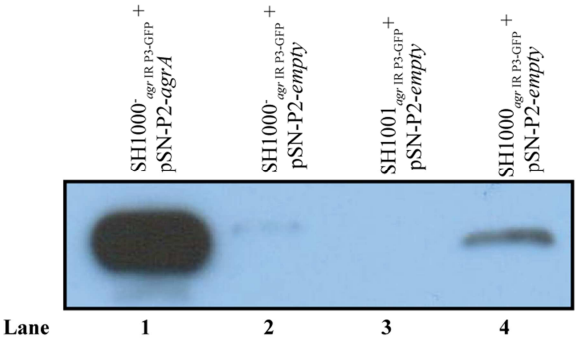


Figure 51: Western blot of SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ complementation, with HIgG.

Western blot analysis comparing the AgrA protein level in the *S. aureus* strains SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*agrA*, SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*empty*, SH10001$_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*empty* and SH1000$_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*empty*. The lane number are indicated.

3.7. Conclusion

The main objective of this section was to set up an *in vivo* experimental system to study the activation of transcription by AgrA.

A polyclonal rabbit antibody specific for AgrA was synthesised and purified. It is highly specific for AgrA and highly sensitive, opening a number of downstream applications to detect AgrA efficiently in whole cell lysates.

The choice of a suitable *S. aureus* strain to study AgrA is not an easy task. Many strains contain mutations in AgrA or have an altered *agr* regulon due to mutations in genes involved in the regulation of the operon. SH1000 (See Table 8) was our strain of choice as it is *rbsU* cured and contains a functional *agr* operon. We characterised the SH1000^{-} (See Table 8) strain, and have illustrated that it seems to be very close to its parental strain, SH1000, but contains a mutation in AgrA that renders the protein non-functional. This strain can be easily complemented with a plasmid borne AgrA which makes it an ideal strain for our *in vivo* set-up. Furthermore, we have identified a new naturally occurring mutation that leads to inactivation of AgrA. Such mutations are frequently observed in the laboratory and clinical settings and will be discussed further in the next chapter (See section 4.1)

The GFP-reporter system described in this chapter was shown to be suitable to measure transcription activation by AgrA at the P3 promoter in the SH1000$^{-}_{agr\ \text{IR P3-GFP}}$ strain and the results were confirmed by other standard methods. The reporter system has proved to be a better system to look at P3 activation by AgrA as it gives a direct read-out and is highly quantitative, enabling the detection of subtle differences. It is also more sensitive than the traditional methods used such as blood agar plates.

The SH1000$^{-}_{agr\ \text{IR P3-GFP}}$ strain was successfully complemented with pSN-P2-*agrA*. The complemented strain was able to activated the P3 promoter at a level even higher than the SH1000$_{agr\ \text{IR P3-GFP}}$ with pSN-P2-*empty* strain, displayed an haemolytic activity on blood agar plates slightly higher than SH1000$_{agr\ \text{IR P3-GFP}}$ with pSN-P2-*empty*, was able to inhibit the production of the staphylococcal immunoglobulin binding proteins Spa and Sbi and was produced at a much higher level than SH1000$_{agr\ \text{IR P3-GFP}}$ with pSN-P2-*empty* as detected by western blotting experiments. In fact, there is a much bigger difference in protein level between the complemented strain and the WT SH1000$_{agr\ \text{IR P3-GFP}}$ with pSN-P2-*empty* strain than the one observed with the GFP fluorescence or the blood agar plates. This is due to the fact that the protein level does not directly correlate with the activity. AgrA needs to be activated by AgrC in order to activate transcription. Therefore, AgrC is the limiting factor in this equation and AgrC is produced at WT level, not overexpressed. Thus, most of the AgrA protein detected by western blot is probably inactive because it cannot be phosphorylated by AgrC because the AgrC/AgrA ratio is skewed.

It was also shown that the use of a P1 driven AgrA is a very weak way to complement for AgrA activity as it produces a very weak activation of the P3 promoter, hence low haemolytic activity, is barely detectable by western blot and unable to inhibit the production of the staphylococcal immunoglobulin binding proteins. Both of the putative P1 promoters (P1A further upstream and P1B just upstream of a*grA*, See Figure 34) have a poor conservation of the -35 promoter element which could be somewhat complemented by the fact that they both have an extended -10 promoter element. The P1A -10 promoter element has a good conservation (5/6) with a change at position -9 which is not too bad in term of activity. The P1B -10 promoter element has a poor conservation (4/6) with a change at position -12 which often leads to a decreased activity (See Figure 52 and sections 1.1.1 and 1.1.3) (Moyle et al, 1991). The P1A is the most likely P1 promoter, based on this analysis. Nevertheless, both putative P1 promoters are weaker than the P2 and P3 promoter in terms of conservation of the -10 and -35 promoter elements which correlates with the observed phenotype. Thus, this proves furthermore that the P2 promoter is the "true" AgrA promoter and therefore should be used to fully complement AgrA activity. However, we cannot exclude that the massive difference observed between the P1 and P2 driven AgrA is partially be due to the SD. The pSN:*agrC832-1261-agrA* plasmid contains the normal SD located in the 3' end of *agrC* immediately upstream of *agrA*. There are a few putative SD in that region; hence we did not know which one was the one used. Therefore, in the pSN-P2-*agrA* plasmid we did not use the natural AgrA SD but a SD optimised for *S. aureus*. However, the result showing a weak activation of the P1 driven AgrA is consistent with the belief that P1 is responsible for the production of a basal low level of AgrA (Novick, 2003).

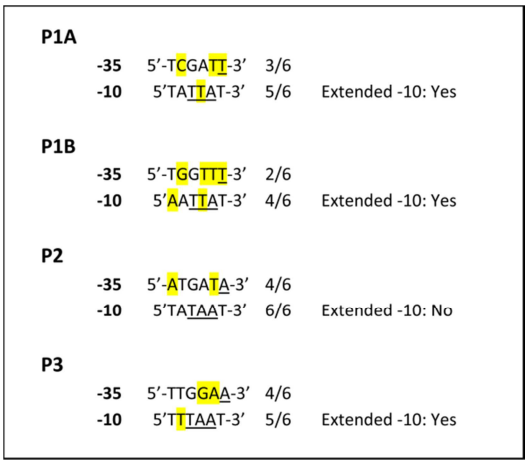


Figure 52: Conservation of the -35 and -10 of the *agr* promoters.

Sequence of the -10 and -35 promoter elements of the putative P1 promoters (P1A and P1B), the P2 and P3 promoters. -35 refers to the -35 element (5'-TTGACA-3') and -10 refers to the -10 element (5'-TATAAT-3'). Extended -10 refers to the extended -10 element, T at -15 and G at -14. The nucleotides that can vary from the

consensus sequence without affecting promoter activity drastically and are thus less conserved are underlined in black. The nucleotides that differ from the consensus promoter elements are highlighted in yellow. The presence or absence of an extended -10 motif is indicated.

Overall, this chapter describes how an *in vivo* system to study transcription activation by AgrA was successfully set-up, combining all the tools developed in this section. This system will be very useful to help unravelling the mechanism of transcription activation by AgrA.

CHAPTER FOUR

4. Construction, screening and characterisation of a library of single alanine mutants in the AgrA LytTR domain

4.1. Introduction

Not much is known about the AgrA LytTR domain responsible for the DNA-binding activity of AgrA in terms of individual aa contributions. When Sidote et al reported the crystal structure of the isolated AgrA LytTR domain in 2008, the knowledge of the field was significantly increased (Sidote et al, 2008). The structure is that of the isolated AgrA LytTR domain in complex with a pentadecamer DNA duplex containing the 9 bp consensus binding sequence for AgrA based on the P2 AgrA binding site.

The AgrA LytTR domain presents an unusual fold consisting mainly of beta strands. It adopts a 10-stranded β-fold which consists of two duplicated 5-stranded β motifs (β1-β5, β6-β10) and contains a short two-turn α-helix (See Figure 53).

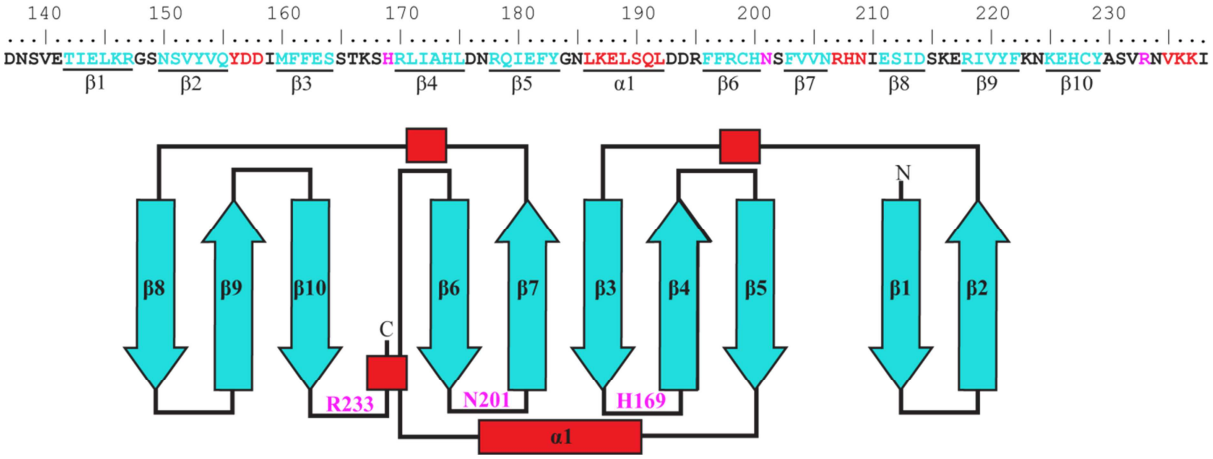


Figure 53: Secondary structure of the AgrA LytTR domain with amino acid (aa) sequence.

The 10 β-strands are coloured in cyan and numbered from 1 to 10. The α-helix is coloured in red and labelled α1 and the three turns are represented by red boxes in the schematic and highlighted in red in the aa sequence. The three residues involved in direct interaction with the DNA are highlighted in magenta. Adapted from (Sidote et al, 2008).

The crystal structure reveals that the AgrA LytTR domain binds to the DNA via three loops that insert into two subsequent major grooves and the minor groove separating them. Three residues located in these loops make direct contact with the DNA and two of them are essential in order for the protein to be able to bind DNA (See Figure 53 and Figure 54, in magenta).

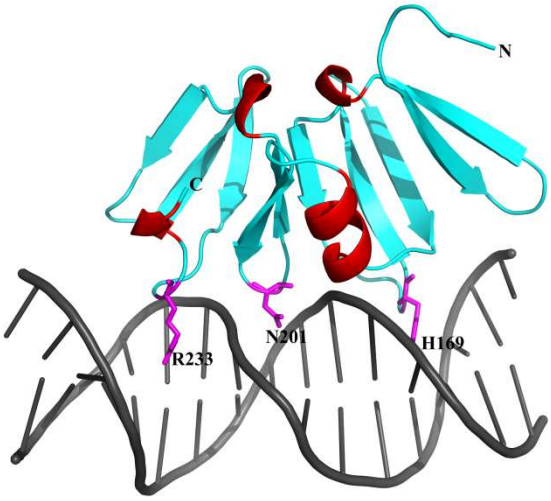


Figure 54: Crystal structure of the AgrA LytTR domain with the three residues directly interacting with the DNA.

Ribbon representation of the AgrA LytTR domain-DNA complex crystal structure. The 10 β-strands are coloured in cyan, the α-helix and the three turns are coloured in red and the DNA in grey. The three residues making direct contact with the DNA are coloured in magenta, displayed as sticks and labelled in black. The N- and C-terminal ends of the structure are indicated in black. Pymol, PDB 3BS1.

The first essential residue is a histidine located at position 169 (H169), in the loop between β3 and β4 (see Figure 53, in magenta), and inserts in a major groove of the DNA (see Figure 54, in magenta). The second essential residue is an arginine located at position 233 (R233A), in the loop between β10 and the C-terminal end of the protein (See Figure 53, in magenta), and inserts in a major groove of the DNA (See Figure 54, in magenta). If H169 or R233 are mutated into an alanine, the protein is no longer able to interact with the DNA. Hence H169 and R233 are essential for the DNA-binding activity of the protein. The third residue making direct contact with the DNA is an asparagine located at position 201 (N201), in the loop between β6 and β7 (See Figure 53, in magenta) and inserting in a minor groove of the DNA (See Figure 54). If this residue is mutated into an alanine, the affinity of protein for the DNA is drastically reduced. Therefore N201 is thought to stabilise the interaction between AgrA and the DNA backbone.

Fourteen other residues located on the face of AgrA proximal to the DNA make indirect contact with the DNA (See Figure 55, in green).

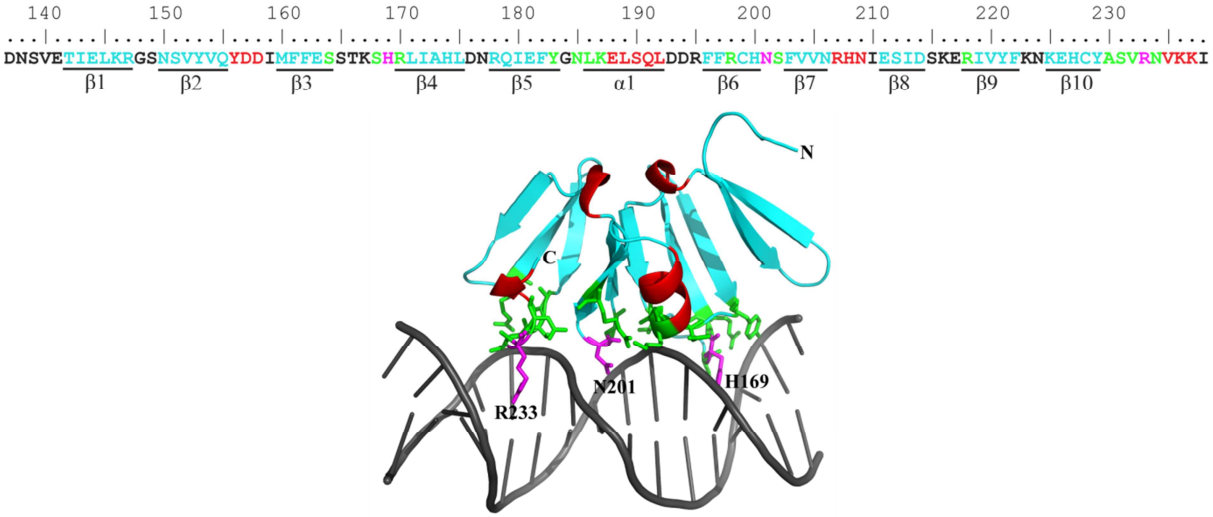


Figure 55: Crystal structure and amino acid (aa) sequence of the AgrA LytTR domain with all the residues making contact with the DNA.

Ribbon representation of the AgrA LytTR domain-DNA complex crystal structure. The 10 β-strands are coloured in cyan, the α-helix and the three turns are coloured in red and the DNA in grey. The three residues making direct contact with the DNA are coloured in magenta, displayed as sticks and labelled in black. The fourteen residues making indirect contact with the DNA are coloured in green and displayed as sticks. The N- and C-terminal ends of the structure are indicated in black. Pymol, PDB 3BS1.

As mentioned in section 1.5.5, upon binding by AgrA, the DNA bends to conform to the surface of the protein, creating a bend of approximately 38°. This bend is thought to be important for the transcription activation function of AgrA.

AgrA is predominantly negatively charged with the majority of the positively charged residues located along the DNA-binding surface of the protein.

The structural integrity of the protein is maintained through eight salt bridges interactions involving ten different aa residues (See Figure 56 and Table 11). One residue, R218, is involved in four different salt bridges interactions. The histidine responsible for the *agrA* minus phenotype of the SH1000^{-} strain, H174, is involved in two salt bridges interactions with E163 and E226 (See Figure 40 and section 3.4) (Sidote et al, 2008).

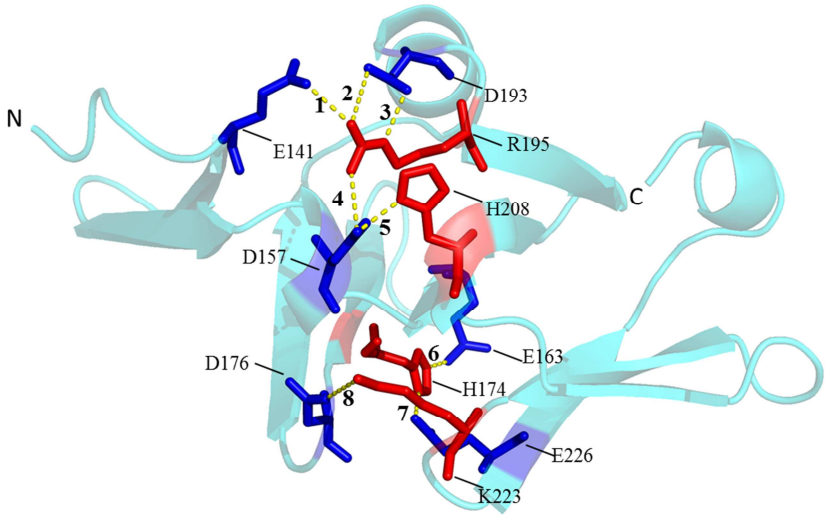


Figure 56: Crystal Structure of the AgrA LytTR domain with the salt bridges interactions.

Ribbon representation of the AgrA LytTR domain. The LytTR domain is coloured in cyan. The ten residues involved in salt bridges interactions are displayed as sticks and coloured according to their charge. The positively charged residues are coloured in red and the negatively charged residues in blue. The eight salt bridges interactions are represented by yellow dotted lines and numbered from 1 to 8. Pymol, PDB 3BS1. Adapted from (Sidote et al, 2008).

	pos. partner	neg. partner	stabilises interaction between: pos./neg.
1	R195	E141	end loop between α1 and β6/beginnning β1
2	R195	D193	end loop between α1 and β6/start loop between α1 and β6
3	R195	D193	end loop between α1 and β6/start loop between α1 and β6
4	R195	D157	end loop between α1 and β6/ loop between β2 and β3
5	H208	D157	loop between β7 and β8/ loop between β2 and β3
6	H174	E163	β4/β3
7	H174	E226	β4/β10
8	K223	D176	loop between β9 and β10/loop between β4 and β5

Table 11: Salt bridges interactions.

List of the salt bridges interactions maintaining the structural integrity of the AgrA LytTR domain. The table contains the name of each of the residues involved (the positive (pos.) partner and the negative (neg.) partner) and the two regions stabilised by their interaction.

Two cysteine residues, C199 and C228 are involved in resistance to oxidative stress. Under reducing conditions, AgrA binds to the promoter of the *bsaA* gene (encoding the glutathione peroxidase essential for survival under oxidative stress) thereby repressing it. Under oxidative stress, a disulphide bond is created between C199 and C228 leading to the dissociation of AgrA from the b*saA* promoter. Therefore the *bsaA* gene is no longer repressed enabling the bacteria to adapt its behaviour to the oxidative stress. In any of the two cysteine residues is mutated into a serine, the disulphide bond cannot be formed and the *bsaA* gene gene is repressed by AgrA under any conditions. This leads to an increase susceptibility to oxidative stress (See section 1.5.5.4). Furthermore, if C228 is mutated into a serine, the mutated $AgrA_{C228S}$ lacks haemolytic activity which is indicative of an *agrA*- phenotype (See section

3.1.2). Thus this residue is thought to be essential for the correct folding of the protein. To the contrary, if C199 is mutated into a serine, the $AgrA_{C199S}$ mutant displays a hyperactive haemolytic activity compared to WT AgrA (Sun et al, 2012b).

Spontaneous mutations in the *agr* operon resulting in a non-haemolytic phenotype occur very frequently in a laboratory setting but also in clinical isolates. The historical *S. aureus* laboratory strain RN4220 (See Table 8) is a nitrosoguanidine-induced mutant able to accept foreign non-staphylococcal DNA and an example of such mutations. RN4220 presents a non-haemolytic phenotype which is due to a frameshift mutation at the C-terminal end of the AgrA LytTR domain. An additional adenine is added to the stretch of seven adenines at the C-terminal end of AgrA (aa 236-238, See Figure 57) creating a frameshift resulting in a protein with three additional aa. This *agrA* 8A is defective and creates a delay in RNAIII production which in turns leads to failure to translate α and δ haemolysins (Traber & Novick, 2006). The *agrA* 8A was also observed in clinical isolated along other naturally occurring mutations targeting the same adenine stretch (Traber et al, 2008). This adenine stretch is conserved in all published *S. aureus* genomes except for the strains Mu3 and Mu50. These strains contain two additional adenines which results in a non-synonymous mutation changing isoleucine 238 into a lysine in *agrA* 9A (See Figure 57). This AgrA is however fully functional (Traber & Novick, 2006). Another spontaneous mutation observed in clinical isolates targeting the same adenine stretch presents a much more drastic effect. If an adenine is removed, the resulting frameshift lengthens the protein by 21 aa (See Figure 57) and the *agrA* 6A is completely inactive (Traber & Novick, 2006; Traber et al, 2008). Similarly, another mutation observed in clinical isolates consists of an *agrA* shortened by one aa residue. Lysine 238 is mutated into a stop codon and the resulting *agrA* is non-haemolytic (Shopsin et al, 2010).The C-terminal end of AgrA therefore seems to be a hotspot for mutations conferring an *agr*- phenotype.

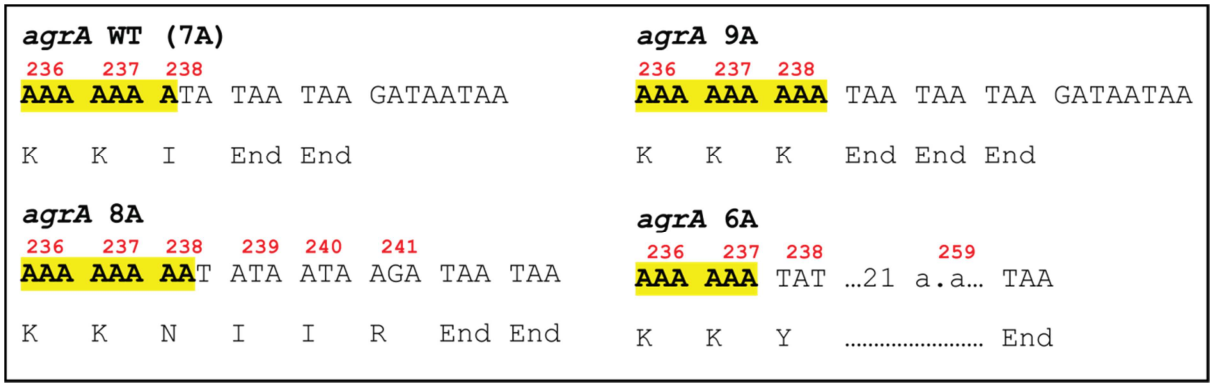


Figure 57: C-terminal *agrA* variants.

C-terminal end of the DNA sequence of the two non-haemolytic variants of *agrA* (8A and 6A) compared to the WT (7A) and the 9A variant. The stretch of As containing the mutations is highlighted in yellow. The amino acid are indicated and numbered in red. Adapted from (Traber & Novick, 2006).

A number of other mutations present in clinical isolates have been reported as *agrA* is the mutational hotspot of the operon, together with the 3' end of a*grC*. These mutations consist of frameshift insertions/deletions, nonsense mutations and non-synonymous changes resulting in a non-haemolytic phenotype. Most of them are located into the AgrA LytTR domain (Shopsin et al, 2010). Only the five non-synonymous mutations identified will be mentioned here, as they are indicative of aa residues important for the function of the AgrA protein. They consist of V140A, V140K, L171F, L186I and R218H. Each of these single aa mutations is sufficient to render the protein inactive and therefore suggests that these aa are essential for the normal function of the protein. Note that residues L186 and R218 are involved in indirect interaction with the DNA (See Figure 55, in green).

The C-terminal loop of AgrA contains a hydrophobic cleft to which fragment compounds can bind leading to inhibition the DNA-binding ability of the protein. This region contains the aa residues S231, V232, R233, N234, K236, K237 and I238 (Leonard et al, 2012). This loop is important for the DNA-binding activity of the protein as residue R233 is essential for the DNA-binding activity of the protein and residues S231, V232 and N234 make indirect contact with the DNA (See Figure 55). In addition, mutations in that region result in partially deficient or complete lack of activity of the protein (See Figure 57). That region had been identified as a putative DNA-binding region because of its hydrophobic nature in the first report about the LytTR domain (See section 1.3.2.1) (Nikolskaya & Galperin, 2002).

Very recently, a small molecule inhibitor able to selectively target AgrA in *S. aureus* was identified. This inhibitor, called savirin, inhibits the activation of the P3 promoter and the production of RNAIII by AgrA. It also inhibits the DNA-binding activity of the AgrA LytTR domain.

The docking region for savirin was identified as the region between R218 on β9 and Y229 on β10 (See Figure 53 and Figure 58). Residues E217 and Y229 are within interaction distance (Sully et al, 2014).

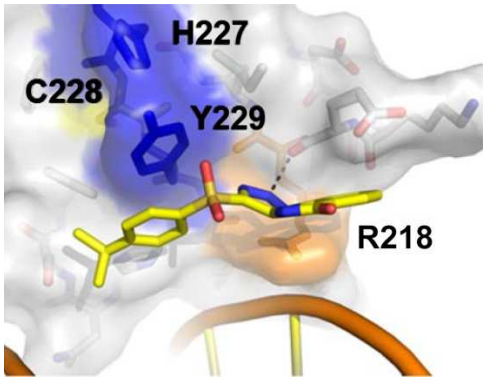


Figure 58: Savirin binding site.

In silico docking of savirin to the AgrA LytTR domain. Savirin is represented by yellow sticks. The residues in the close vicinity are annotated and displayed as sticks (Sully et al, 2014).

The savirin docking region contains C228, involved is disulphide bond and important for AgrA folding (See above) and R218 which is involved in indirect interaction with the DNA (See Figure 55) and important for the activity of the protein (see above). This region therefore seems to be important for the activity of AgrA.

The most conserved motif found in LytTR domain containing transcription factor is the FhRhH[RK][SNQ]hhVN motif (h corresponds to hydrophobic residues, see Figure 59 and section 1.3.2.1) (Nikolskaya & Galperin, 2002) spanning from aa residues 196 to 206 and corresponding to FFRCHNSFVVN in AgrA (See Figure 60). This motif contains the aa N201 involved in direct interaction with the DNA (See Figure 53 and Figure 54), the aa R198 and S202 making indirect contact with the DNA (See Figure 53 and Figure 54) and C199 involved in disulphide bond under oxidative stress (see above).

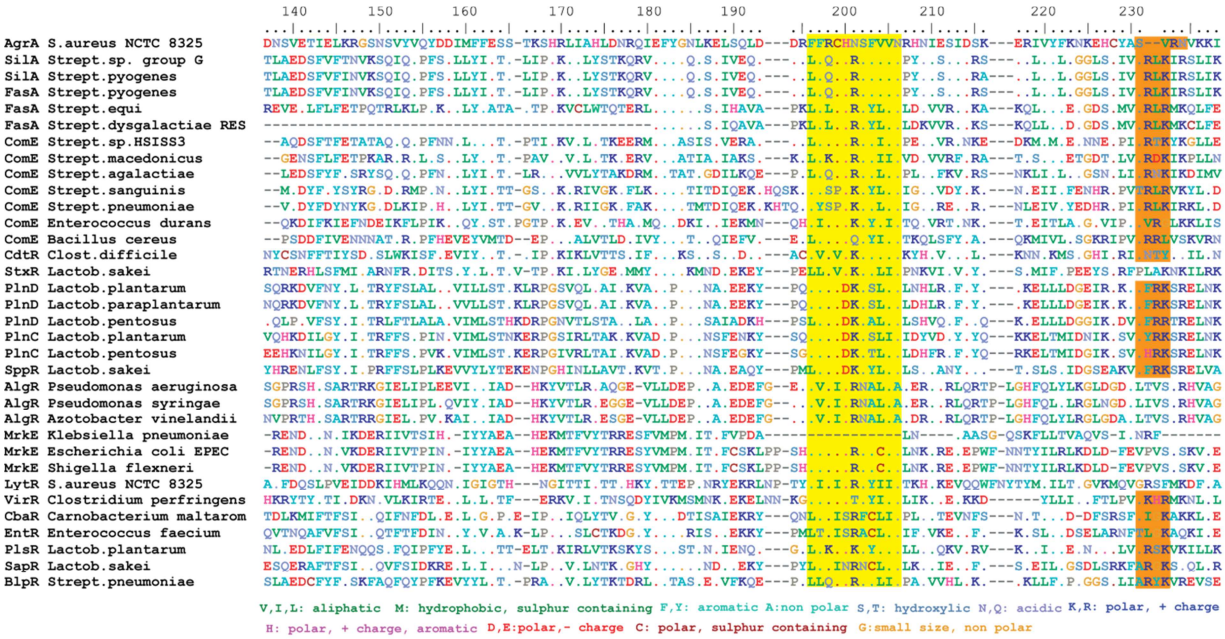


Figure 59: Clustal alignment LytTR domain containing transcription factors (LDCTFs).

Multiple sequence alignment of the amino acid (aa) sequence of the LytTR domain of forty-three representative LDCTFs. Aa residues with similar properties have similar colour (See legend). Conserved aa are represented by a dot. The FhRhH[RK][SNQ]hhVN motif is highlighted in yellow and the S[RK][RK][RKHY] motif in orange.

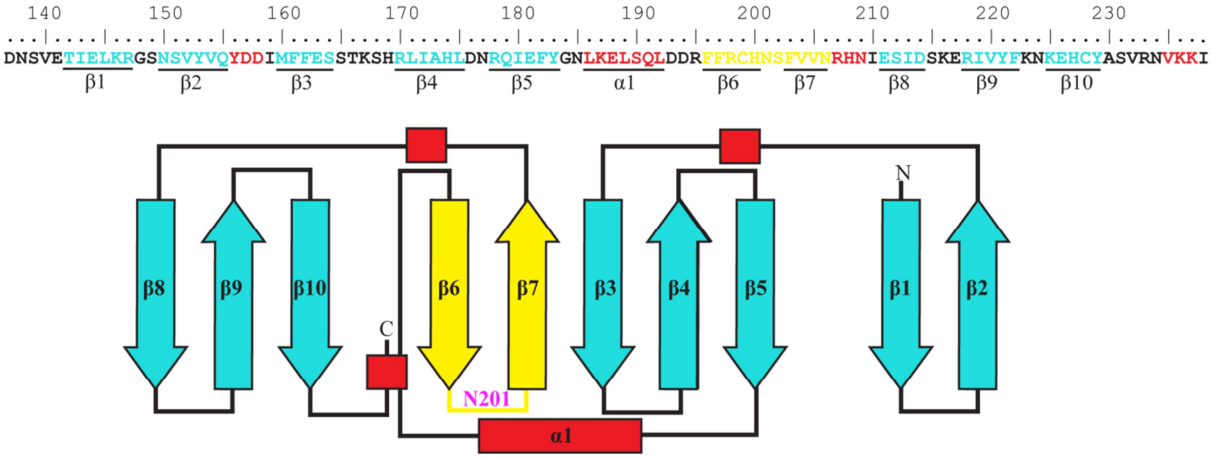


Figure 60: Secondary structure and amino acid sequence of the AgrA LytTR domain with FFRCHNSFVVN motif.

The 10 β-strands are coloured in cyan and numbered from 1 to 10. The α-helix is coloured in red and labelled α1 and the three turns are represented by red boxes in the cartoon and highlighted in red in the aa sequence. The FFRCHNSFVVN motif is highlighted in green and the residue N201 making direct contact with the DNA is highlighted in magenta in the schematic. Adapted from (Sidote et al, 2008).

In addition, this motif (FFRCHTSYIVN in the LDCTF VirR from *Clostridium perfringens*) has been shown to be involved in the DNA-binding activity of VirR and to be essential for the activity of the protein. Residues R186, H188 and S190 from VirR (corresponding to R198, H200 and S202 from AgrA, See Figure 61) have been shown to be essential for the activity of the protein and R186 and S190 to be involved in DNA-binding. Residues L179, F184, N194 and L195 (corresponding to D193, F196, N206 and R207 in AgrA, See Figure 61) are also essential for the activity of the VirR protein. Residues E176, F185 and I198 (corresponding to S190, F197 and I210 in AgrA, See Figure 61) are important for the activity of the VirR protein but are not essential (McGowan et al, 2002). The wide conservation of this motif among LDCTFs suggests that the corresponding residues in *S. aureus* AgrA could assume a similar function.

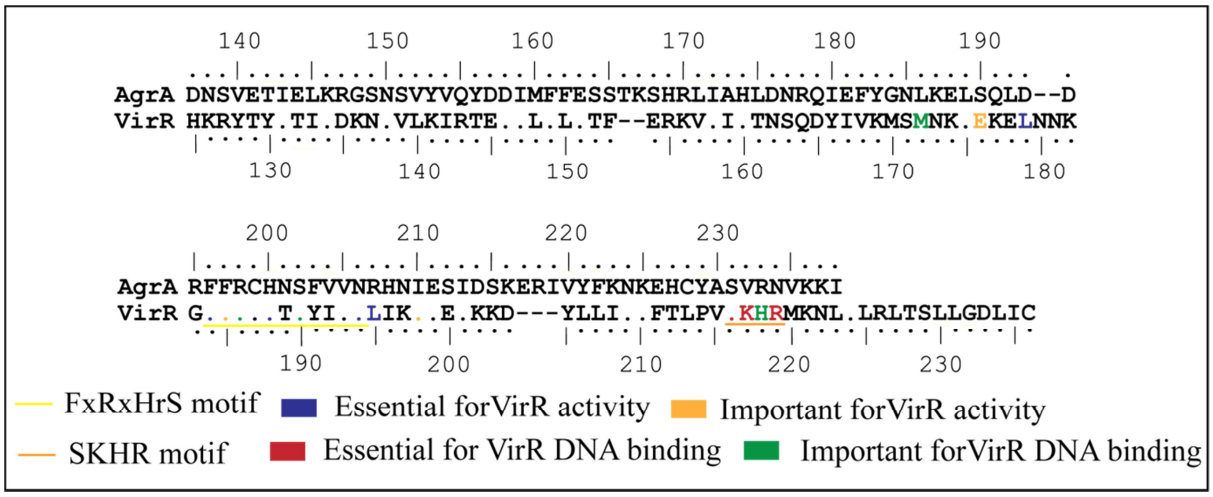


Figure 61: Clustal alignment of AgrA with VirR from *Clostridum perfringens*.

Amino acid (aa) sequence alignment of the LytTR domain of AgrA and VirR. The conserved aa are indicated by a dot. The number corresponding to each aa is indicated above for AgrA and below for VirR.

Another motif conserved among most LDCTFs, S[RK][RK][RKHY] (See Figure 59) was characterised in VirR. The SKHR motif from VirR (aa 216-219) corresponds to the SVRN motif in AgrA (aa 231-234, see Figure 61). Each residue of the motif is essential for the activity of the VirR protein. Residue M172 (L186 in AgrA, see Figure 61) was also identified as being required for the biological activity of VirR. Residues M172, and H218 are involved the in the DNA-binding activity of the protein and residues S216, K217 and R219 are essential for the DNA-binding activity of VirR (See Figure 61) (McGowan et al, 2003). This motif seems to assume a similar function in AgrA. R233A is essential for the DNA-binding activity of the protein and S231, V232 and N234 make indirect contact with the DNA (See Figure 54 and Figure 55) and all are part of the hydrophobic region important for the DNA-binding activity of the protein (See above).

All the residues involved in DNA-binding in VirR have already been identified as being involved in DNA-binding in AgrA as well. Therefore the DNA-binding mode could very well be conserved among LDCTFs.

In summary, the knowledge of the AgrA LytTR domain individual aa contributions is limited to DNA-binding and structural (salt bridges) roles and nothing is known about potential aa involved in transcription activation.

4.2. Objectives

The objective of this section was to try to elucidate the individual contribution of the aa residues in AgrA LytTR domain to the transcription activation function of the protein. To this aim the specific section involved:

I. Selection of the aa residues to mutate in the AgrA LytTR domain
II. Construction of a single alanine mutant library in the AgrA LytTR domain
III. Scanning of the alanine mutant library in the SH1000$^{-}_{agr\ \text{IR P3-GFP}}$ strain
IV. *In vivo* characterisation of a selection of single alanine mutants in the AgrA LytTR domain
V. Synonymous and non-synonymous mutations at selected residues in the AgrA LytTR domain

4.3. Selection of the amino acid residues to mutate in the AgrA LytTR domain

The AgrA LytTR contains 101 aa residues, 29 of which have an assigned function (See Figure 62). The remaining 70% of the aa of the AgrA LytTR domain have no assigned function. In order to try to elucidate the function of the residues with unknown function and hoping to unravel aa involved in the transcription activation function of AgrA we decided to target all these residues for alanine mutagenesis. A few residues with known function were selected as well. One residue essential for binding the DNA (R233),three residues involved in salt bridges interactions (H218, K223, H174 [mutated into a leucine, like in the SH1000^{-} strain, See section 3.6]), three residues making indirect contact with the DNA (R198, S202, R218) and the two cysteine residues involved in disulphide bond formation (changed into serine residues). This led to a library of 81 mutants to construct by Site Directed Mutagenesis (SDM) (See Figure 62).

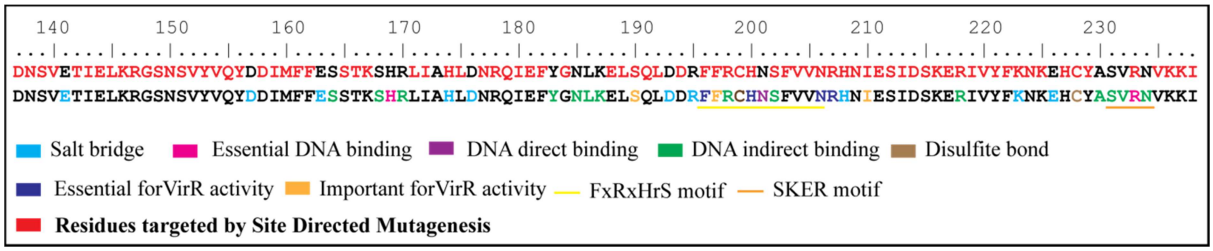


Figure 62: List of residues targeted by Site Directed Mutagenesis

Amino acids (aa) sequence of the AgrA LytTR domain. In the first line, the residues that were targeted for SDM that are coloured in red. In the second line, the residues with a known function are coloured according to that function (See key).

4.4. Construction of a single alanine mutant library in the AgrA LytTR domain

The alanine library was constructed by Site Directed Mutagenesis (SDM) using the pSN-P2-*agrA* plasmid as a template, as described in Section 2.2.8 and in Figure 63 with the primers listed in Table 15, Table 16 and Table 17.

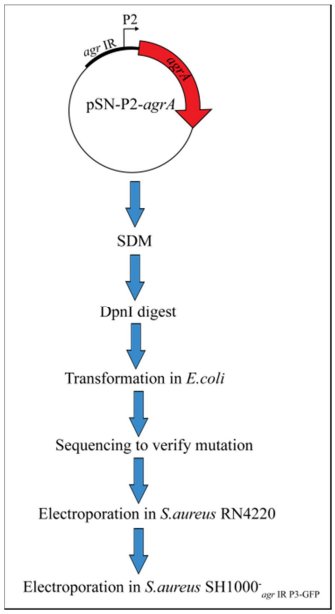


Figure 63: SDM library construction.

Flow-chart describing how the *S. aureus* single alanine mutant library was constructed. In the pSN-P2-*agrA* plasmid, *agrA* is under the control of the P2 promoter and the *agr* intergenic region (IR).

4.5. Scanning of the alanine mutant library in the SH1000^{-} *agr* IR P3-GFP strain

The SH1000^{-} *agr* IR P3-GFP strain was transformed with each of the 81 mutants and their ability to complement for AgrA activity was tested by GFP growth curve using a plate reader (See Section 2.1.2 and Section 3.6). As described in Section 3.6, the GFP activity for each mutant is calculated according to the growth (GFP-Fluo/$OD_{600\ nm}$) after 8 h of growth and is displayed compared to the WT complementation with pSN-P2-*agrA* set at 100%. Three controls were included: SH1000^{-} *agr* IR P3-GFP + pSN-P2-*agrA*$_{H174L}$ (H174L), SH1000^{-} *agr* IR P3-GFP + pSN-P2-*empty* (SH^{-}) and SH1000^{-} *agr* IR P3-GFP + pSN-P2-*agrA* (WT).

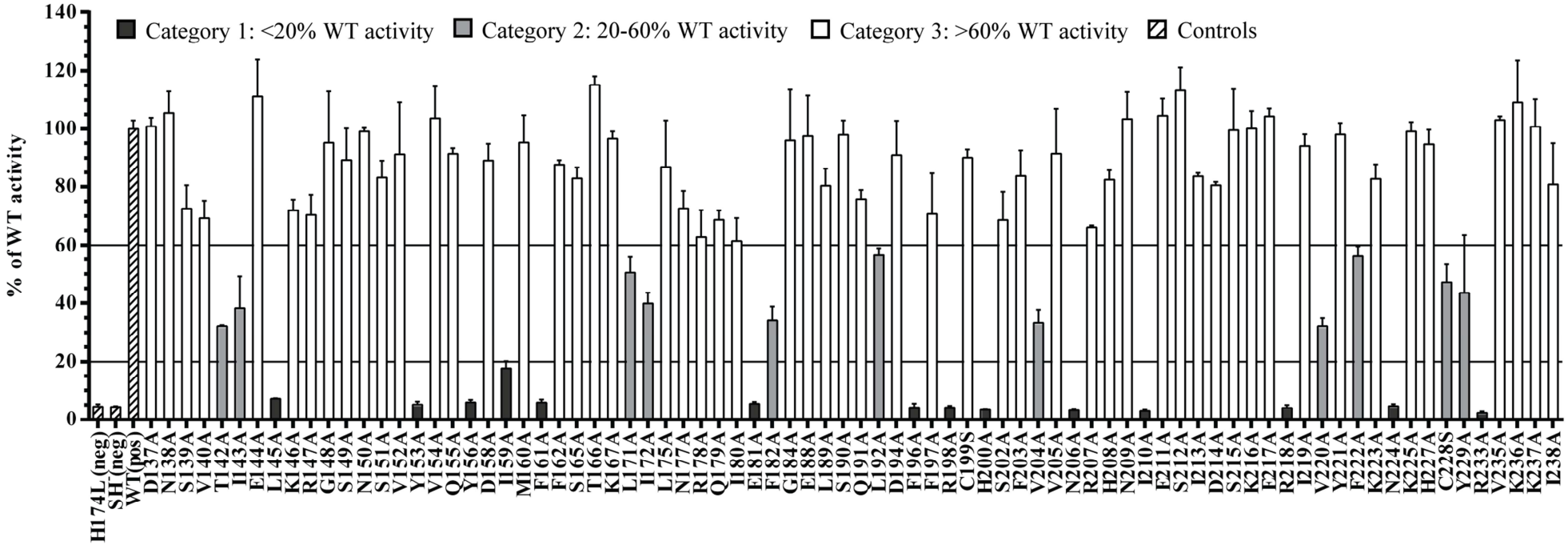


Figure 64: Systematic mutational analysis of the AgrA LytTR domain.

Graph showing GFP expression [as GFP fluorescence units (GFP-Fluo)] as a function of growth ($OD_{600\ nm}$) relative to the SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*agrA* for each of the 81 mutant after 8 h of growth in TSB. AgrA mutants displaying more than 60% activity compared to the WT AgrA are shown in white, the AgrA mutants displaying between 20 and 60% activity compared to the WT AgrA are shown in light grey, the AgrA mutants displaying less than 20% activity compared to the WT AgrA are shown in dark grey and the controls (SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*agrA*$_{H174L}$, SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*empty* and SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*agrA*) are shown in white with black stripes. Data obtained from at least three biological replicate.

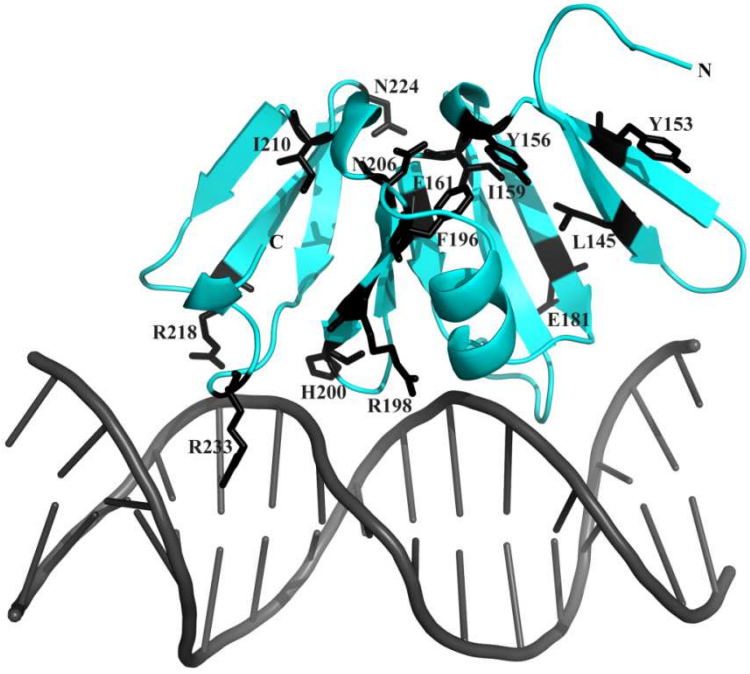


Figure 65: Position of the mutants of category 1 on the structure.

Ribbon representation of the AgrA LytTR domain-DNA complex crystal structure. The AgrA LytTR domain is coloured in cyan and the DNA in grey. The fourteen residues that display less than 20% activity compared to the WT when mutated to alanine are displayed as sticks, coloured in black and labelled in black. Pymol, PDB 3BS1.

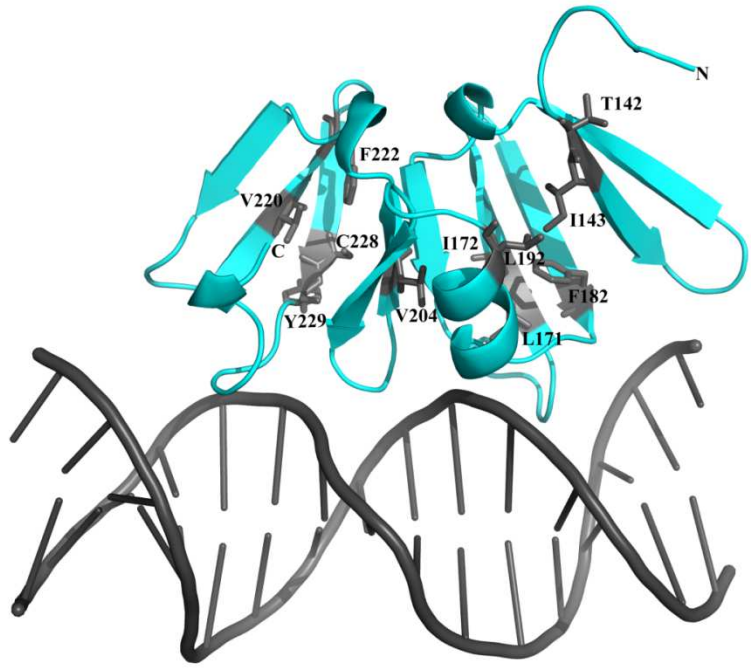


Figure 66: Position of the mutants of category 2 on the structure.

Ribbon representation of the AgrA LytTR domain-DNA complex crystal structure. The AgrA LytTR domain is coloured in cyan and the DNA in grey. The eleven residues that display less between 20 and 60% activity

compared to the WT when mutated to alanine are displayed as sticks, coloured in grey and labelled in black. Pymol, PDB 3BS1.

The mutants displayed a wide range of activities and were separated into three activity categories. The first activity category contained the fourteen mutants the most drastically affected and displaying less than 20% activity compared to the WT (in dark grey in Figure 64). The second activity category contained the eleven mutants displaying between 20 and 60% activity compared to the WT (in light grey in Figure 64). The position on the structure of the AgrA LytTR domain of the residues belonging to the first and second category is illustrated in Figure 65 and Figure 66, respectively. The third activity category contained the fifty-five mutants displaying more than 50% activity compared to the WT (in white in Figure 64). The third category was not investigated further as the attention was focused on the two remaining groups. The aim of the screen was to identify putative transcription activation defective (TAD) mutants and it seemed unlikely that such mutants would retain more than 60% activity *in vivo*.

However, a few residues are worth mentioning as they have been discussed above. For example, the C199S mutation has been described as hyperactively haemolytic which does not seem to be the case here as it displays 90% activity compared to the WT. The two residues involved in salt bridges interactions H208 and K223 (See Figure 56) do not appear to be essential to maintain the structural integrity of the protein as H2008A and K233A both display 82% activity compared to the WT. The C-terminal end residues of AgrA V235,K236, K237 and I238 part of the hydrophobic cleft do not seem to be affected by the change to alanine as they display 103%, 110%, 100% and 80% activity, respectively. Maybe replacing these residues with hydrophilic residues would have a more drastic effect. So even if this region has been reported to be a hotspot for mutations inactivating AgrA (See Figure 57), this is not reflected in this screen.

The mutation V140A was observed in clinical isolates and resulted in an *agrA*- phenotype. However, in this screen it displayed 70% activity.

4.6. *In vivo* characterisation of a selection of single alanine mutants in the AgrA LytTR domain

The mutants for the first and the second category (See Figure 64) were selected for further analysis aiming to identify potential TAD mutants. C228 was excluded as it has a known function in oxidation sensing (See section 4.1). It can be mentioned that this residue was thought to be essential for the folding of AgrA as the C228S mutation was observed to be

non-haemolytic. However, in this screen this mutant displays 47% activity which means it would still be partially haemolytic.

In order to exclude the possibility that the TAD property is caused by effects on protein stability by the mutation in question, we aimed to determine whether the 25 putative TAD AgrA mutants were produced at level similar to WT AgrA under the assay conditions. To investigate this, the 25 putative TAD mutants were transferred into the plasmid pSN-*itet-agrA* (See Table 14), so that expression of AgrA could be induced with anhydrotetracycline and thus be independent of the P2 promoter. This was necessary because in pSN-P2-*agrA* the native AgrA-dependent P2 promoter drives *agrA* transcription and therefore a putative TAD mutant would fail to or poorly activate its own transcription (See Figure 67 and section 1.5).

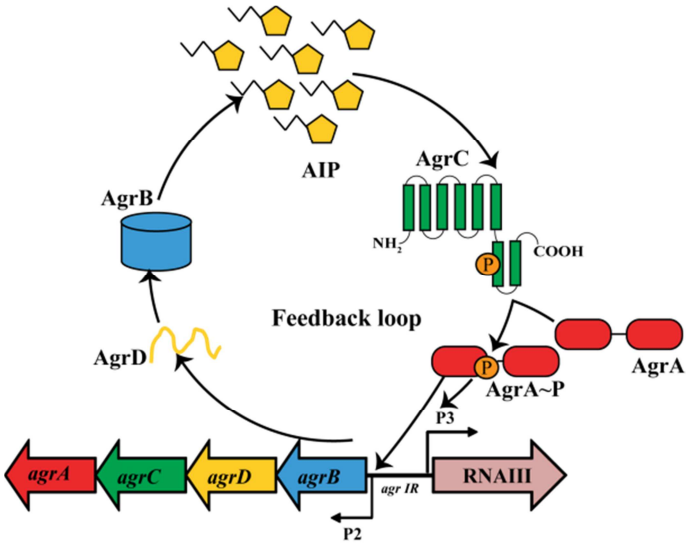


Figure 67: Cartoon of the *agr* operon feedback loop.

The *agr* operon feedback loop is illustrated by black arrows forming a circle. *agrD* encodes a propeptide that is processed and secreted by the transmembrane AgrB with the help of SpsB releasing the finalised AIP. AgrC senses the cell density-dependent accumulation of the AIP and autophosphorylates upon binding the AIP. The phosphorylated AgrC is then able to activate AgrA via phosphotransfer. The phosphorylated AgrA is then able to bind to specific direct repeats at the *agr* IR and activate transcription at the P2 and P3 promoter. By activating the P2 promoter, AgrA creates an autocatalytic circuit as it stimulates its own transcription. For more details, see section 1.5.

pSN-*itet-agrA* contained the same plasmid backbone (pCN34) as pSN-P2-*agrA* but with an inducible tetracycline promoter in place of the P2 promoter and *agr* IR (pCN34*itet*, see Table 9). Initially we tested if P3 activity in SH1000^{-} $_{agr\ IR\ P3\text{-}GFP}$ could be restored with pSN-*itet-agrA*. As shown in Figure 68, P3 activity was restored to comparable levels in SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ containing pSN-*itet-agrA*, both in the presence and absence of anhydrotetracycline, but, as expected, not in SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ containing pSN-*itet-agrA*$_{H174L}$ even in the presence of anhydrotetracycline. It seems that leaky expression of AgrA from pSN-*itet-agrA* is sufficient

to restore full P3 activity in SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ containing pSN-*itet-agrA*, therefore all downstream experiments involving pSN-*itet-agrA* were conducted in the absence of anhydrotetracycline.

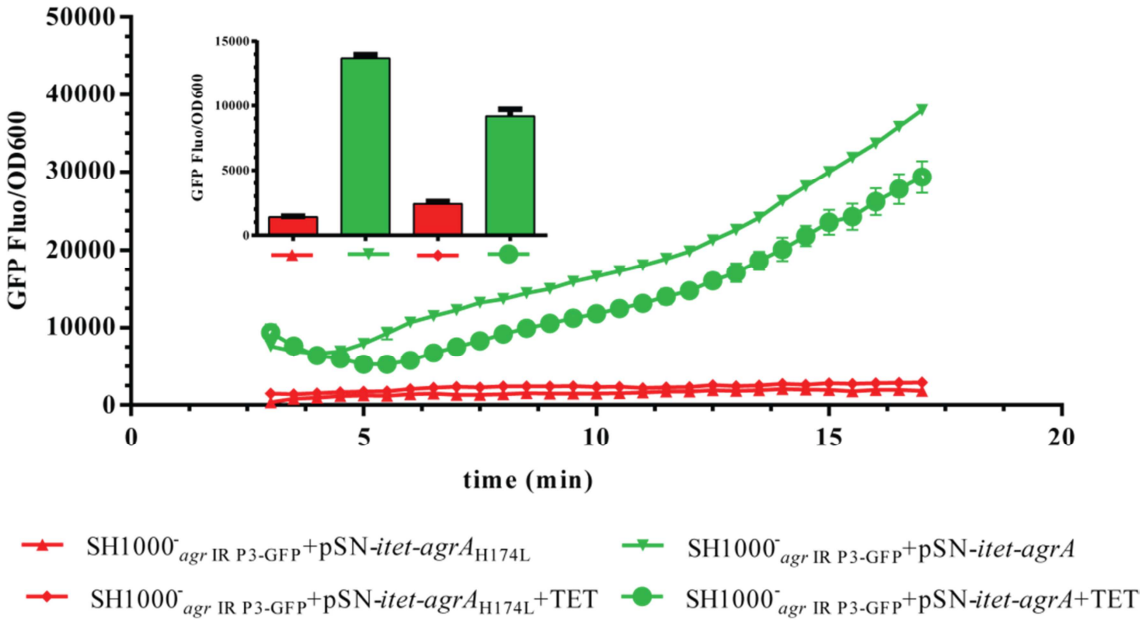


Figure 68: GFP growth curve of pSN-*itet-agr.*

Graphs showing GFP expression [as GFP fluorescence units (GFP-Fluo)] as a function of growth ($OD_{600\ nm}$) over time for SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-*itet-agrA* and SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-*itet-agrA*$_{H174L}$ strains grown in TSB with and without the addition of anhydrotetracycline. The bar chart in the insert represents the GFP expression of the same samples (colour coded accordingly) at the 8 h' time point. Data obtained from at least three biological replicates.

The GFP fluorescence for each mutant was measured as well and compared to the values obtained in SN-P2-*agrA*.The 16 h' time point was chosen instead of the 8 h' time point in order to get enough protein to be detectable by western blot. No significant differences were observed between the two plasmids and the two different time points for most of the mutants (See Figure 69).

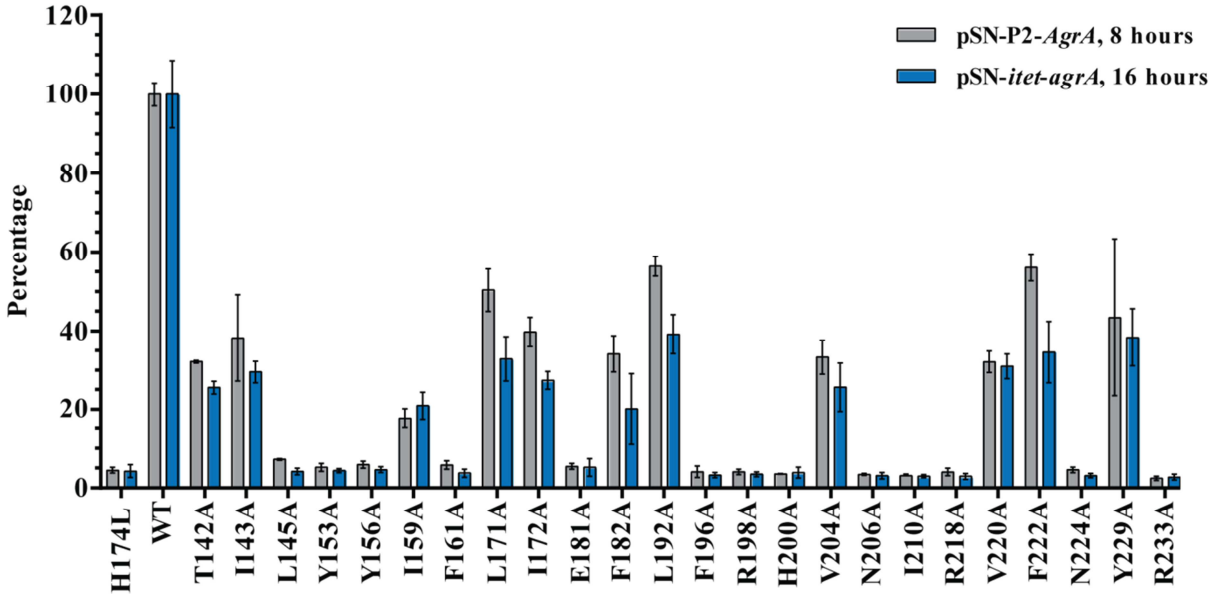


Figure 69: Comparison of the GFP fluorescence of the 24 putative TAD mutants in pSN-P2-*agrA* and in pSN-*itet-agrA*.

Graph showing GFP expression [as GFP fluorescence units (GFP-Fluo)] as a function of growth ($OD_{600\ nm}$) relative to the $SH1000^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-*itet-agrA* for the mutants in pSN-*itet-agrA* or relative to the $SH1000^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*agrA* for the mutants in pSN-P2-*agrA* for the 24 single alanine mutants displaying less than 60% WT activity after 8 h (pSN-P2-*agrA*) or 16 h (pSN-*itet-agrA*) of growth in TSB. Data obtained from at least three biological replicates.

The variation that was observed for most of the mutants was within the range of the error bars except for five mutants which all displayed a higher activity in pSN-P2-*agrA* : T142A (32.2% vs 25.7%), L171A (50.4% vs 32.9%), I172A (39.8% vs 27.4%), L192 (56.4% vs 39%) and F222A (56% vs 34.6%). These differences could be due to the feedback loop system of the P2 promoter which would mean that none of these five mutants are likely to be TAD mutants as they are apparently able to activate their own transcription. Alternatively, these differences could be the result of another TF binding at the agr IR to activate transcription at the P2 and that is not binding to the *itet* promoter.

Nevertheless, the protein expression level of the 24 selected putative TAD mutants was tested by western blot after growth for 16 h in TSB (See section 2.3.8 and Figure 70).

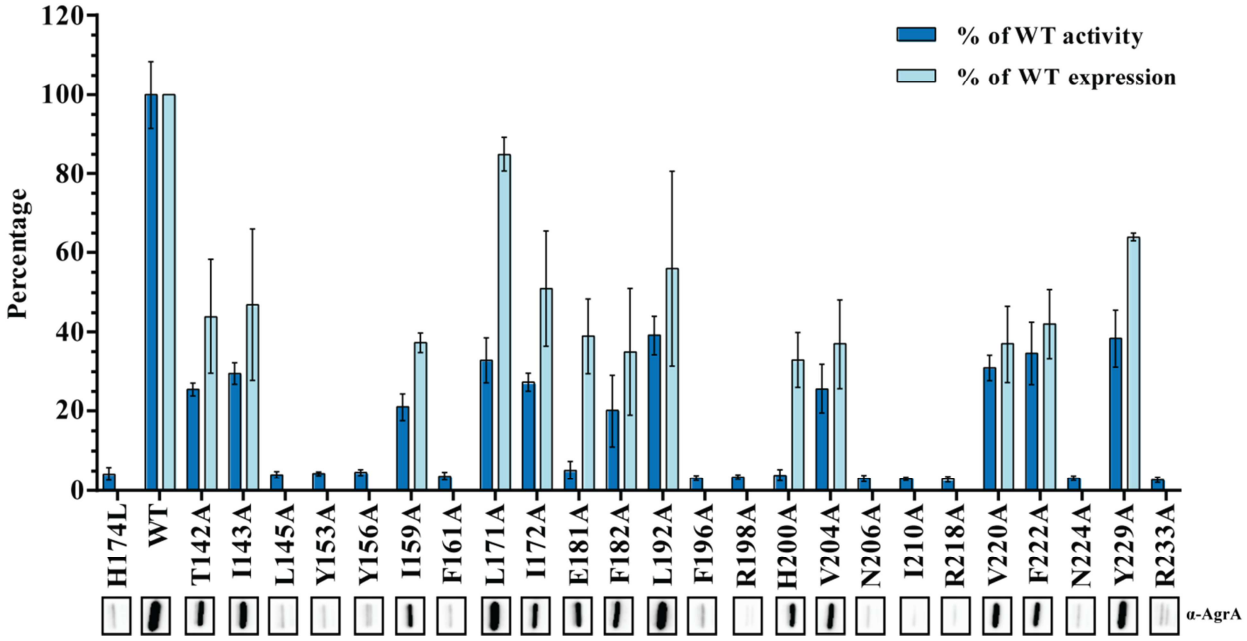


Figure 70: GFP activity vs AgrA expression for the 24 putative TAD mutants.

Graph showing GFP expression [as GFP fluorescence units (GFP-Fluo)] as a function of growth ($OD_{600\ nm}$) relative to the $SH1000^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-*itet-agrA* for each of the 24 single alanine mutant displaying less than 60% WT activity after 16 h of growth in TSB. A section of the Western blot image indicating AgrA protein levels in whole cell lysates for each mutant is shown. For the 13 mutants that are detectably expressed, the AgrA protein levels in whole cell lysates relative to the WT AgrA levels is shown on the graph. Data were obtained from at least three biological replicates.

Eleven of the twenty-four mutants are not detectably expressed, which suggests that that alanine substitution at positions L145, Y153, Y156, F161, F196, R198, N206, I210, N224 and R233 in the LytTR domain of AgrA could impair the gross structural stability of the protein under the assay conditions. Consistent with this view, side-chains of Y156, F161, F196, N206, I210 and N224 are either fully or partly buried within the hydrophobic core of AgrA (See Figure 71, colored in blue and labelled in red).

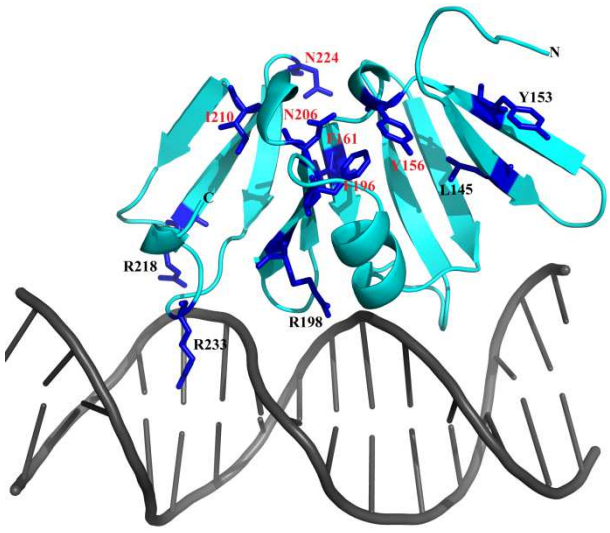


Figure 71: AgrA crystal structure with eleven residues not detectably expressed when mutated to alanine.

Ribbon representation of the AgrA LytTR domain-DNA complex crystal structure. The AgrA LytTR domain is coloured in cyan and the DNA in grey. The eleven residues that are not detectably expressed when mutated to alanine are displayed as sticks and coloured in blue. They are labelled in black but the eight residues that are buried in the hydrophobic core are labelled in red. Pymol, PDB 3BS1.

For 4 (L171A, E181A, H200A and Y229A) of the 13 AgrA mutants which are detectably expressed, the ability to activate transcription did not correlate with the expression levels (both relative to WT AgrA) under the assay conditions. In other words, even though the L171A, E181A, H200A and Y229A mutants are expressed to levels at least ≥ 35-50% of that of WT AgrA, a relatively reduced level of P3 activity (by ≥60% loss of activity) is seen in cells containing these mutants. Thus, the results suggest that aa residues L171, E181, H200 and Y229 in the LytTR domain could potentially play a significant role in transcription activation by AgrA. This conclusion is further substantiated by the fact that L171, E181, H200 (100% identity) and Y229 (100% similarity) are fully conserved in AgrA found in all staphylococci (See Figure 72). Moreover, three of these residues, L171, E181 and H200 are well conserved among LDCTFs as well (See Figure 73).

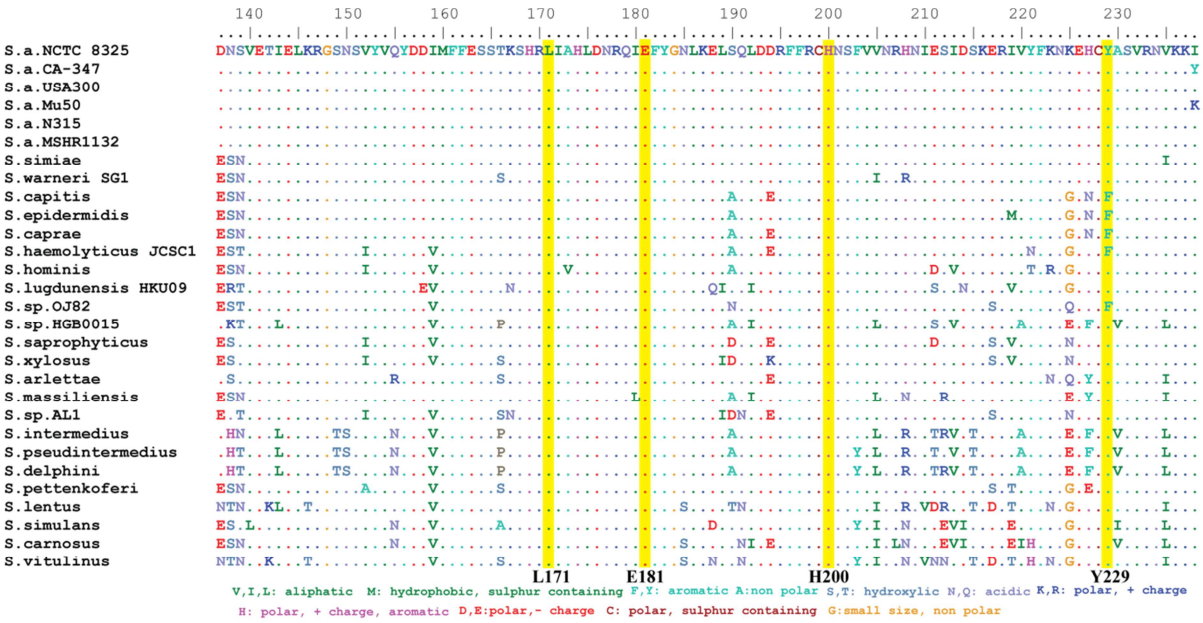


Figure 72: Multiple Sequence alignment of the amino acid (aa) sequence of the AgrA LytTR domain of 29 representative *Staphylococci* strains with the four TAD mutants.

Aa residues with similar properties have similar colour (See legend). Conserved aa are represented by a dot. The four putative transcription activation deficient (TAD) mutants L171, E181, H200 and Y229 are highlighted in yellow and labelled.

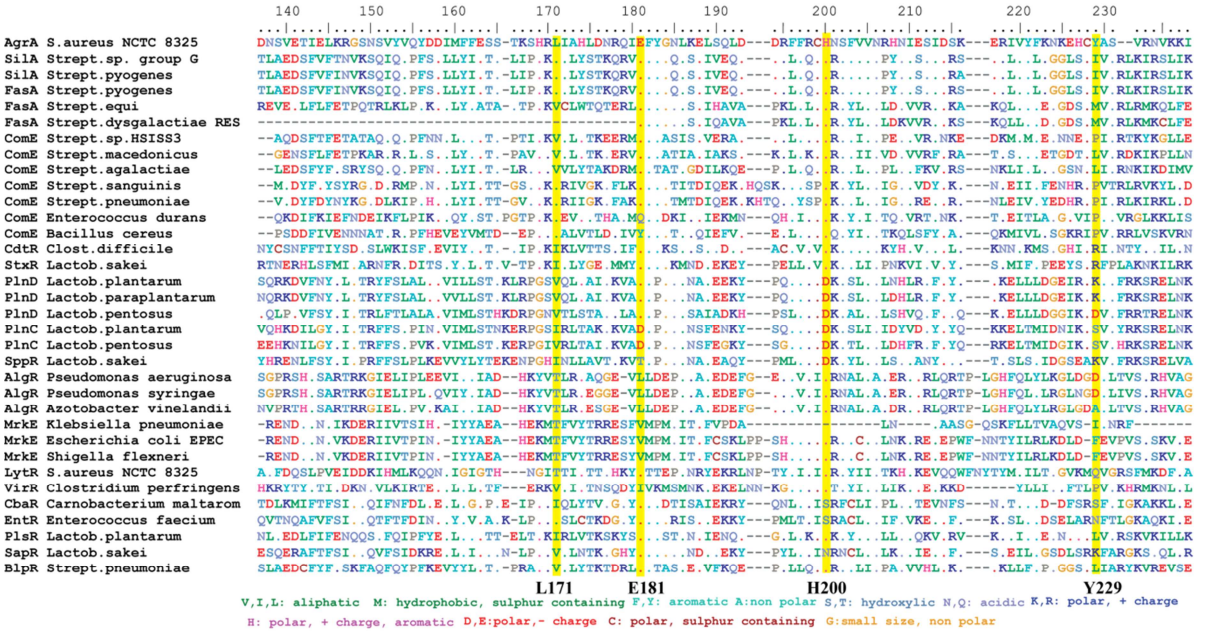


Figure 73: Multiple Sequence alignment of the amino acid (aa) sequence of the LytTR domain of 43 representative LDCTFs with the four TAD mutants.

Conserved residues are represented by a dot. The aa residues displaying similar colours have similar properties (See legend). The four putative transcription activation deficient (TAD) mutants L171, E181, H200 and Y229 are highlighted in yellow and labelled in black.

4.7. Synonymous and non-synonymous mutations at selected residues in the AgrA LytTR domain

In parallel to the *in vivo* characterisation in section 4.6, a number of synonymous and non-synonymous mutations in selected mutants from the first and second category were investigated. This mutagenesis was conducted in order to try to determine the aa requirements at each of these positions. The eleven residues that were not detectably expressed were excluded as well as five residues that seem to be buried in the structure, I159, V204, V220 and F222 (See Figure 74, in orange).

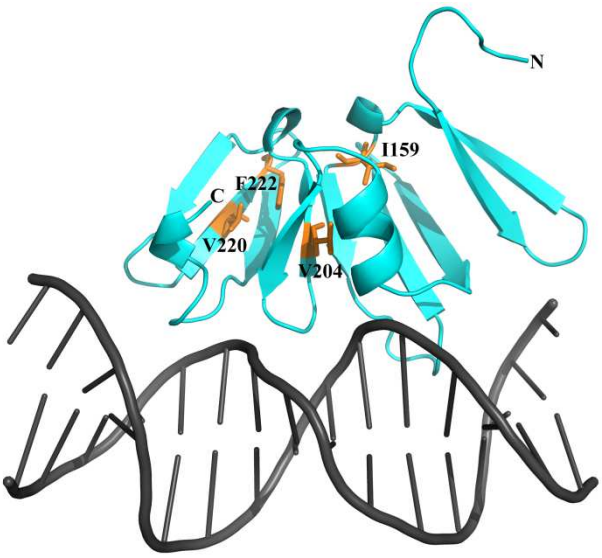


Figure 74: AgrA crystal structure with the four residues excluded.

Ribbon representation of the AgrA LytTR domain-DNA complex crystal structure. The AgrA LytTR domain is coloured in cyan and the DNA in grey. The four residues that were excluded are displayed as sticks, coloured in orange and labelled in black. Pymol, PDB 3BS1.

The mutations were created by SDM as described in section 4.4 and screened as described in section 4.5. Nine of the twenty-four TAD mutants were targeted: T142, I143, L171, I172, E181, F182, L192, H200 and Y229 (See Figure 75).

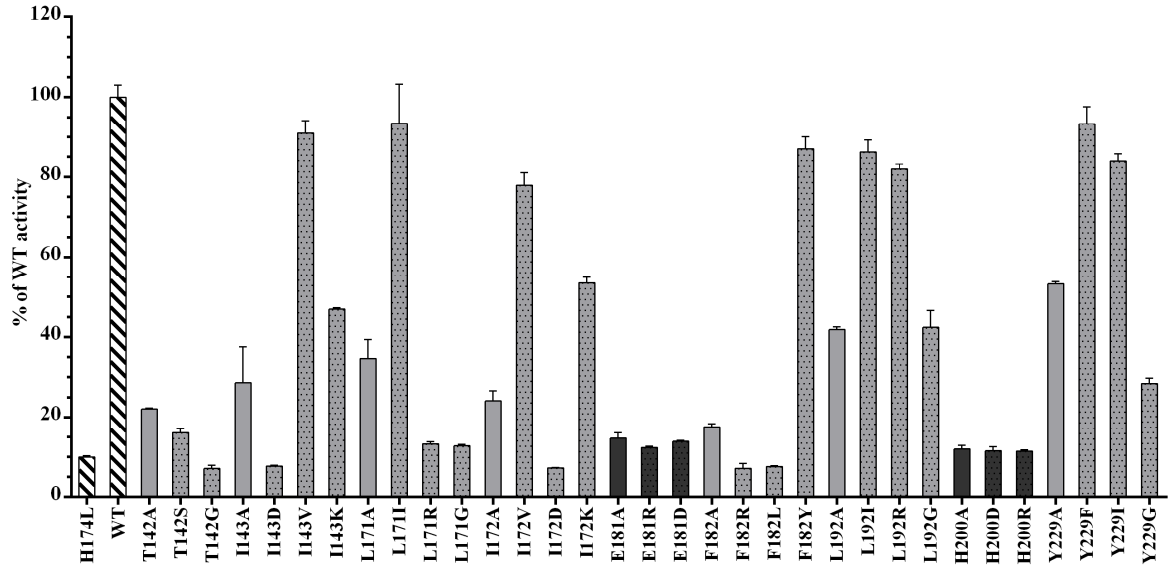


Figure 75: Synonymous and non-synonymous mutations.

Graph showing GFP expression [as GFP fluorescence units (GFP-Fluo)] as a function of growth ($OD_{600\ nm}$) relative to the $SH1000^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*agrA* for each of the 33 single aa mutants after 8 h of growth in TSB. The alanine mutants are colour coded according to their activity category (See Figure 64).The mutants displaying less than 20% activity compared to the WT are coloured in black and the ones displaying between 20% and 60% compared to the WT in grey. The WT and H174L mutants are coloured in white with black stripes. The different mutants at each position are shown in the same colour, the alanine mutant without any

pattern and the other mutants with a dotted pattern. Data were obtained from at least three biological replicates.

Position 142 does not seem to tolerate any other aa than a threonine. As expected, the substitution with a non-polar glycine (7%) was even more disruptive than the alanine substitution. But the substitution with a serine, differing from a threonine only by a hydrogen group instead of a methyl group, was expected to produce a higher activity than the alanine substitution. However, it was not the case and both the alanine and the serine substitution had a very similar activity, 22% and 16%, respectively (See Figure 75). This position is poorly conserved among LDCTFs (See Figure 76).

The requirements at position 143 and 172 appear to be very similar. The most disruptive substitution is with the negatively charged aspartate (8% and 7%, respectively). A positively charged aa, lysine, is much more tolerated giving about 50% activity (47% and 53%, respectively), better than with the alanine substitution (29% and 24%, respectively). The valine substitution, very similar to the WT isoleucine, gives almost full activity (91%) at position 143 and 78% activity at position 172 (See Figure 75). This suggests that a bulky, aliphatic and hydrophobic residue is needed at both positions but an isoleucine is required in order to get full activity at position 172. However, looking at the conservation of position 143 among LDCTFs, the only substitutions that are observed are with a leucine and a phenylalanine (See Figure 76) suggesting that these are the only accepted substitutions. Interestingly, although position 172 is conserved with 100% identity among *Staphylococci* it is poorly conserved among LDCTFs suggesting that this position is important only in the context of AgrA (See Figure 76).

Position 171 does not tolerate small residues like a glycine (13%) nor a positively charged arginine (13%) which gives a lower activity than the alanine substitution (35%). However, an isoleucine, very similar to the WT leucine gives almost full activity (See Figure 75). This substitution can be observed in other LDCTFs as well as valine and threonine substitutions (See Figure 76).This suggests that an aliphatic and hydrophobic, preferably bulky residue is needed at that position

Position 192 is fairly similar to position L171 but more tolerant. The small residue glycine has about 40% activity, similar to the alanine substitution (43%). With a positively charged arginine (82%) or an isoleucine (86%), very similar to the WT leucine, the activity is much higher but still lower than the WT (See Figure 75).This suggests that a positively charged or a

bulky, aliphatic and hydrophobic residue is tolerated at that position. This position is poorly conserved among LDCTFs (See Figure 76).

Position 181 seems to absolutely require a glutamate. All other substitutions lead to an activity of 15% or less, even with a negatively charged aspartate (See Figure 75). This position is very well conserved among other LDCTFs (See Figure 76) supporting this view.

At position 182, an aromatic reside is required for the activity of the protein. Substitutions with an arginine (7%) or the bulky leucine (8%) give lower activity than an alanine (18%). Substitutions with another aromatic residue, tyrosine gives 87% activity (See Figure 75). This position is very well conserved among other LDCTFs (See Figure 76).

Position 200 does not tolerate any other aa than a histidine. Other polar aa, positively charged arginine or negatively charged aspartate or an alanine all give about 13% activity (See Figure 75). This position is very well conserved among other LDCTFs (See Figure 76) suggesting that a histidine is required for full activity.

Position 229 tolerates some variation. An alanine substitution gives 50% activity, a bit less with the smaller glycine, 28%. The bulky, aliphatic and hydrophobic residue isoleucine has a high activity (83%). Aromatic residue phenylalanine has almost WT activity (93%, See Figure 75). The phenylalanine substitution is observed in other *Staphylococci* (See Figure 72). This position can tolerate another bulky aa but an aromatic residue is needed at that position for full activity. This position is poorly conserved among other LDCTFs (See Figure 76)

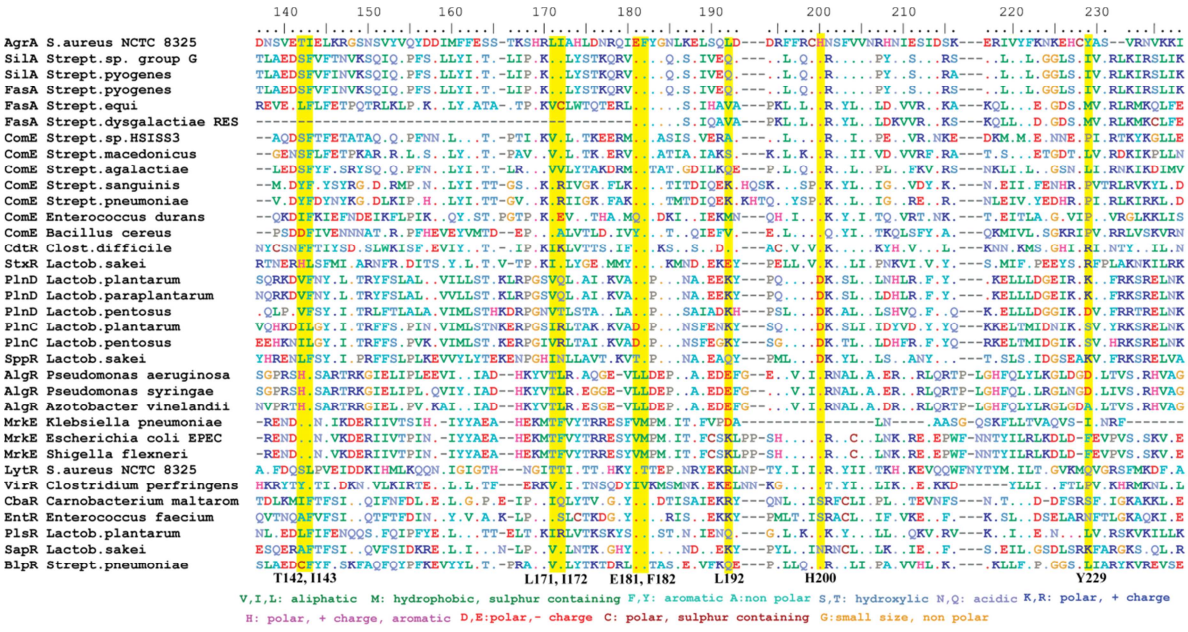


Figure 76: Multiple Sequence alignment the amino acid (aa) sequence of the LytTR domain of forty-three representative LDCTFs with the nine residues mutated highlighted.

The conserved residues are represented by a dot. The aa residues displaying similar colours have similar properties (See legend). The nine residues T142, I143, L171, I172, E181, F182, L192, H200 and Y229 are highlighted in yellow and labelled in black.

4.8. Conclusion

The aim of this chapter was to clarify the individual aa contribution in the LytTR domain to the activity of the AgrA protein. Overall, the aim was reached. Essential and important residues in the LytTR domain were identified by means of a systematic alanine screen. The screen enabled an unbiased approach to identify potentially important residues in the AgrA LytTR domain.

The residues were categorised according to their activity in the screen and a subset of residues were further characterised *in vivo*. This *in vivo* approach consisting of a western blot analysis enabled the identification of 11 residues potentially essential for the protein to assume the correct structural fold. These residues display a very low activity in the screen (~4%) and the proteins are not stably detected by western blot. These residues have never before been implicated in a structural role in AgrA. However, residues F196 and N206 have been reported as being essential for the activity of the VirR protein and are in the FxRxHrS motif (See Figure 62) which is involved in DNA-binding in VirR. One could therefore hypothesise that these residues are likely involved in the DNA-binding activity of the AgrA protein.

Residue I210 has been reported as being important for the activity of the VirR protein (See Figure 62) and it appears to be essential.

R233A is recognised as being essential for the DNA-binding function of the protein and R198 and R218 make indirect contact with the DNA (see section 4.1). R198 is also in the the FxRxHrS motif (See Figure 62). As mentioned in section 4.1, the FhRhH[RK][SNQ]hhVN motif is the most conserved motif among LDCTFs (See Figure 59) and important for the function of the *Clostridum perfringens* regulator VirR. All the residues of the motif except N201 were included in screen (See Figure 77).

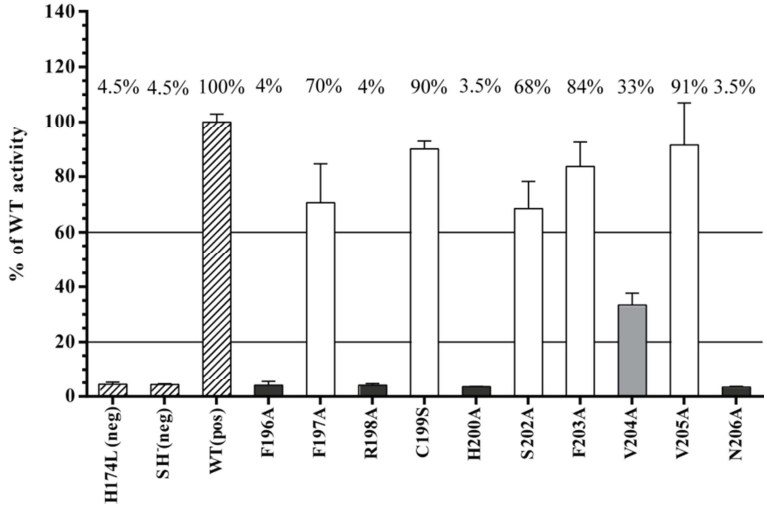


Figure 77: Activity of the residues in the FFRCHNSFVVN motif.

Graph showing GFP expression [as GFP fluorescence units (GFP-Fluo)] as a function of growth ($OD_{600\ nm}$) relative to the SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*agrA* for each mutant in the FFRCHNSFVVN motif after 8 h of growth in TSB. AgrA mutants displaying more than 60% activity compared to the WT AgrA are shown in white, the AgrA mutants displaying between 20 and 60% activity compared to the WT AgrA are shown in light grey, the AgrA mutants displaying less than 20% activity compared to the WT AgrA are shown in dark grey and the controls (SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*agrA*$_{H174L}$, SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*empty* and SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*agrA*) are shown in white with black stripes. Data obtained from at least three biological replicates.

The result clearly shows that four of the most conserved residues (F196, R198, H200 and N206, in dark grey in Figure 77) appear to be essential for the activity of AgrA. Residue V204 seems to be important for the function of AgrA as it displays a low activity as well, 33% (in light grey in Figure 77). These five residues will be further characterised and discussed further below. However, residue S202 which is essential in VirR and makes indirect contact with the DNA in AgrA does not appear to be essential as it retains 68% activity. Another essential residue in VirR is R207 which does not seem to be essential in AgrA either as it displays still 66% activity (See Figure 64). Even if not all the residues of the motif are important in AgrA, this motif to be globally important for the activity of the AgrA.

The *in vivo* characterisation of the selected 25 putative TAD mutants resulted in reducing the list of potential TAD mutants to only 4 candidates. Because the next step of the characterisation of the TAD mutants involves *in vitro* analysis with the need to purify each individual mutant protein it was essential to refine the number of candidates as much as possible in order to have a feasible number of candidates. Therefore, our criteria for the selection were very stringent. It cannot therefore be excluded that potential TAD mutants were missed because of that. However, because of the stringency of our criteria, we can be

absolutely sure that none of the 4 TAD mutants are involved in maintaining the structural integrity of the protein.

A number of synonymous and non-synonymous mutations were investigated in order to elucidate the aa requirements at each of the targeted positions. These substitutions allowed a further characterisation of the 4 TAD mutants. H200 and E181 are very conserved position among LDCTFs and no other residue appears to be tolerated at these positions. L171 can be substituted with a hydrophobic and aliphatic residue, preferably bulky. Y229 can be substituted with a hydrophobic bulky residue, preferably aromatic.

These 4 TAD mutants will be further characterised *in vitro* in the next chapter (See chapter 0).

CHAPTER FIVE

5. Further characterisation of the four putative TAD mutants

5.1. Introduction

In the previous chapter (chapter 0) four putative transcription activation deficient (TAD) mutants were identified. They were shown to have a reduced activity *in vivo* compared to the WT AgrA protein which was not due to any kind of structural instability. However, further characterisation of these mutants is necessary in order to unravel the mechanism responsible for their decreased activity. An *in vitro* approach was selected as we were mainly interested in looking at transcription activation. A fully native *in vitro* system to study transcription in *S. aureus* has been previously successfully developed in the lab (Reynolds & Wigneshweraraj, 2011) and is readily available. Furthermore, such an *in vitro* analysis enables the isolation of each mechanistic step in order to identify which one is responsible for the observed phenotype. We can therefore look at transcription activation directly without having any other mechanism taking place at the same time. However, it is important to keep in mind that an *in vitro* setting is artificial and can lead to artefactual effects. This should not be the case here as the mutants we have selected have proved to be deficient for transcription activation *in vivo*.

The use of an *in vitro* approach implies the purification of all the individual necessary components for the reaction in order to recreate the system artificially. Therefore, all the proteins components needed to be purified and a suitable DNA template to be designed.

5.1.1. AgrA

AgrA had been successfully purified in the lab with an N-terminal 6His tag (See section 3.3 and 2.3.4.1) (Reynolds & Wigneshweraraj, 2011). However, in order to have a protein as native as possible, and to exclude potential detrimental effects of the tag, we decided to adopt a tag-less approach. Therefore, the intein mediated purification with an affinity chitin-binding tag (IMPACT^{TM}) method was selected. We used the pTYB2 plasmid, containing an N-terminal intein tag. The method makes use of the inducible self-cleavage activity of protein splicing elements called inteins to separate the target protein from the affinity tag. The splice junction of the intein is modified in order to allow controllable peptide bond cleavage. The cleavage of the intein is activated by thiol reagents such as DTT. The intein tag also contains a chitin-binding domain (CBD) that will enable it to bind to chitin. Therefore, the purification can be achieved with a column packed with chitin slurry by means of on column cleavage

with thiol reagents that will result in the target protein being released from the intein tag (See Figure 78). This enables the purification of a native recombinant protein without the use of proteases (Chong et al, 1997; Chong et al, 1998).

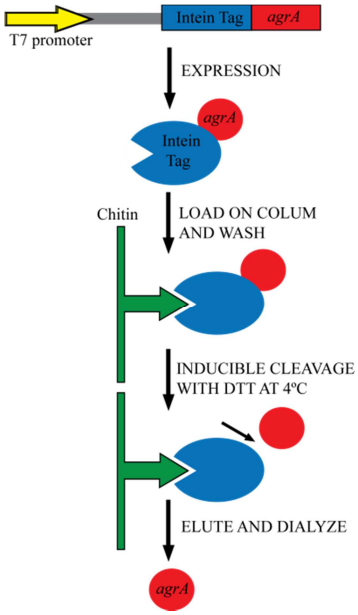


Figure 78: Principle of the IMPACT™ kit.

Schematic representation of the mechanism of the Intein Mediated Purification with an Affinity Chitin-binding Tag (IMPACT). For more details refer to section 2.3.4.3. Adapted from (Chong et al, 1997).

5.1.2. Activation of AgrA

In the cell, AgrA needs to be phosphorylated by AgrC in order to be in the active state. Therefore, we needed to find a way to activate AgrA in a similar fashion in order to produce active AgrA. Two different approaches were undertaken. The first approach involved creating a system as native as possible via the introduction of the natural sensor HK responsible for the phosphorylation of AgrA, AgrC. However, AgrC is by definition a transmembrane protein therefore most likely poorly soluble. Therefore, in order to maximise the chances of successful purification, we decided to attempt using only the cytoplasmic part of the AgrC protein, the HK domain, from aa residues 206-430 (AgrC-HK, See Figure 79) to phosphorylate AgrA. This approach has already been successfully conducted for a large number of HKs and their cognate RRs in *Caulobacter crescentus* (Skerker et al, 2005).

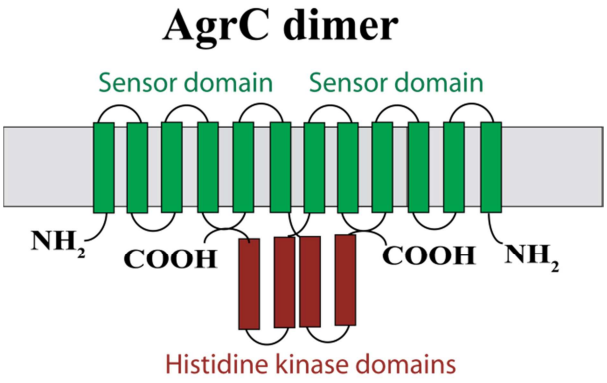


Figure 79: AgrC topology.

Schematic representation of the domain organisation of the AgrC dimer. The transmembrane sensor domain of each monomer (aa 1-200) is coloured in green, the linker (aa 201-221) in black and the Histidine kinase domain (aa 222-430) in brown.

The second approach, widely used for *in vitro* activation of RRs (See Table 12), involved using a small phophodonor, acetyl phosphate, to activate AgrA. Acetyl phosphate is a high energy form of phosphate. It has the ability to store more energy than almost any other phosphorylated compound, even more than ATP. It has been shown to have the ability to donate phosphoryl group to many RRs partners in two-component signalling systems (TCS) *in vitro* thereby activating the RR in the absence of its cognate HK (See Table 12) (Wolfe, 2005). Acetyl phosphate has been previously successfully used to activate AgrA, in presence of $MgCl_2$ (Koenig et al, 2004; Queck et al, 2008; Reyes et al, 2011). The presence of $MgCl_2$ has been shown to be essential in order for the CheY RR to be phosphorylated by acetyl phosphate *in vitro* (Lukat et al, 1992; McCleary & Stock, 1994)

RR	HK	Organism(s)	Present *In Vitro*
AlgR	AlgZ	*Pseudomonas aeruginosa*	+
BvgA	BvgS	*Bordetella pertussis*	+
CitB	CitA	*Klebsiella pneumoniae*	+
CheY	CheA	*Escherichia coli*	+
CheY1, Y2	CheA	*Sinorhizobium meliloti*	+
CheY1, Y3–6	CheA	*Rhodobacter sphaeroides*	+
CheV2		*Helicobacter pylori*	+
ComA	ComP	*Bacillus subtilis*	+
CpxR	CpxA	*Shigella sonnei, E. coli*	+
CsrR	CsrS	*Streptococcus pyogenes*	+
CyaC	CyaC	*Spirulina platensis*	+
DcuR	DcuS	*E. coli*	+
FixJ	FixL	*Sinorhizobium meliloti*	+
KdpE	KdpD	*Clostridium acetobutylicum*	+
KdpE	KdpD	*E. coli*	+
MprA	MprB	*Mycobacterium tuberculosis*	+
NarL	NarX	*E. coli*	+
NodW	NodV	*Bradyrhizobium japonicum*	+
NRI (NtrC)	NRII (NtrB)	*E. coli*	+
OmpR	EnvZ	*E. coli*	+
PhoB	PhoR	*E. coli*	+
PhoP*	PhoQ	*Salmonella enterica*	+
PrrA	PrrB	*R. sphaeroides*	+
RcpA	CphA	*Calothrix* sp. strain PCC7601	+
RcpB	CphB	*Calothrix* sp. strain PCC7601	+
RegA	RdeA	*Dictyostelium discoideum*	+
RssB	none	*E. coli, S. enterica*	+
SirA	BarA	*S. enterica*	
SSK1	SLN1-YPD1	*Saccharomyces cerevisiae*	+
VanR	VanS	*Enterococcus faecium*	+

Table 12: Examples of RRs successfully activated by acetyl phosphate.

The table lists the name of each response regulator (RR), the name of the corresponding histidine kinase (HK), the organism(s) they are found in and whether the RR could be activated with acetyl phosphate *in vitro*. Adapted from (Wolfe, 2005).

Acetyl phosphate is an intermediate in the conversion pathway of acetyl coenzyme A (Acetyl-CoA) to acetate. Acetyl-CoA can be converted into acetyl phosphate by the phosphotransacetylase (PTA) enzyme whereas acetyl phosphate can be converted into acetyl-CoA by the PTA. Similarly, the acetyl phosphate can be converted into acetate by the acetate kinase (AK) whereas the acetate can be converted into acetyl phosphate by the AK. The conversion of acetate into acetyl phosphate by the AK releases a phosphoryl group that can be acquired by a RR (See Figure 80) (Wolfe, 2005).

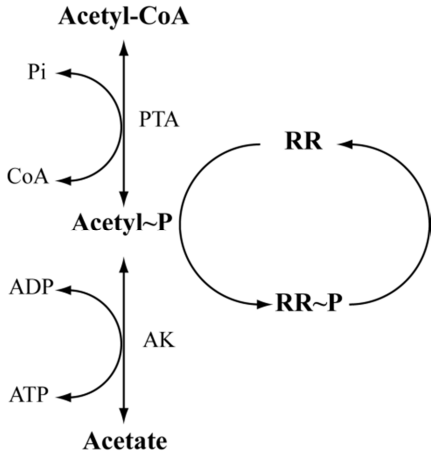


Figure 80: Phosphorylation of a RR by Acetyl phosphate.

Schematic of acetyl phosphate in the conversion pathway of acetyl-CoA to acetate and phosphorylation of a response regulator (RR) by acetyl phosphate. Acetate is converted into acetyl phosphate (Acetyl~P) by the acetate kinase (AK) using ATP and releasing ADP. The same enzyme can convert Acetyl~P into acetate, thereby converting ADP in ATP. The conversion of Acetyl~P into acetyl coenzyme A (Acetyl-CoA) by the phosphotransacetylase (PTA) uses coenzyme A (CoA) and releases inorganic phosphate (Pi) that can be acquired by the RR thereby phosphorylating it (RR~P). The PTA can also convert Acetyl-CoA into Acetyl~P using Pi and releasing CoA. Based on (Wolfe, 2005).

5.1.3. The *S. aureus* RNA polymerase

The native *S. aureus* core RNAp was purified with a method previously optimised in the laboratory for *S. aureus* (Reynolds & Wigneshweraraj, 2011) (See section 2.3.4.2). This method consists of an immunoaffinity chromatography using a polyol-responsive antibody in order to reduce the need for harsh conditions and preserve core RNAp activity. The cross-reactive polyol-responsive monoclonal antibody (Mab) 8RB13 can be used to purify core RNAp from a number of bacteria (*E. coli, Bacillus subtilis, Pseudomonas aeruginosa*, etc…) that contain the epitope recognised by the Mab. 8RB13 specifically detects an epitope in the flap tip helix of the β-flap domain. The epitope, corresponding to aa sequence 5'-PEEKLLRAIFGEKAS-3', is conserved in many bacteria, hence the wide recognition spectrum of the Mab (Bergendahl et al, 2003). The use of a polyol responsive antibody is ideal for large proteins such as the core RNAp as the elution is gentle and only requires a polyol containing buffer and salt. This method also minimises the chances of contamination with sigma (σ) factors.

5.1.4. Sigma factor (σ)

The *S. aureus* σ factor σ_A, was cloned in the pET28b$^+$ vector to create an N-terminal 6His tag , expressed and purified by Jonathan Reynolds (Reynolds & Wigneshweraraj, 2011).

5.1.5. DNA templates

Three different DNA templates (See Figure 81) were used for the *in vitro* analysis. One linear 216 bp DNA template containing the P2 and the P3 promoter, as well as the *agr* IR was used for the EMSAs. For the short RNA synthesis assay, a 96 bp linear probe containing the P3 promoter only was used. The third template consisted of the same 216 bp fragment inserted in the circular plasmid pJR_{P2+P3} (Reynolds & Wigneshweraraj, 2011) and was used for the *in vitro* transcription assay.

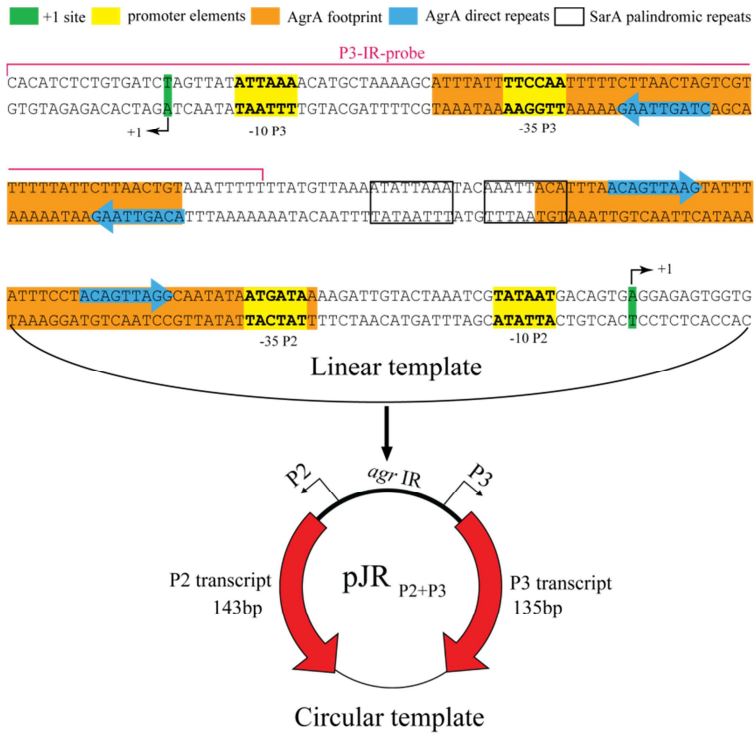


Figure 81: DNA templates for *in vitro* analysis.

DNA sequence of the promoter templates used for the *in vitro* analysis. The linear template was used for the EMSAs (See section 5.4). The circular template was used for the *in vitro* transcription assay (See section 5.5.2). For this template, the red arrows represent the transcripts originating from the P2 and the P3 promoter, respectively. The P3-IR-probe is indicated in pink and was used for the short RNA synthesis assay (See section 5.5.1).

5.2. Objectives

The aim of this chapter was to further characterise the four putative TAD mutants in order to determine whether they are indeed *bona fide* transcription activation deficient mutants. This further characterisation consisted mainly of an *in vitro* analysis. This *in vitro* analysis was then complemented with further *in vivo* and *in silico* characterisation. The specific sections involved:

I. Tooling up for the *in vitro* analysis
II. Electrophoretic mobility shift assay (EMSA)
III. *In vitro* transcription assay
IV. *In vivo* analysis: double mutants
V. *In vitro* phosphorylation of AgrA
VI. *In silico* simulation of the effect of the Y229Amutation (performed by Dr. Robert Weinzierl)
VII. DNA-bending assay

5.3. Tooling up for the *in vitro* analysis

The first step of the *in vitro* analysis involved purifying and preparing all the components necessary for the *in vitro* analysis which is here referred to as “tooling up”.

5.3.1. Expression and purification of AgrA, $AgrA_{E181A}$, $AgrA_{L171A}$, $AgrA_{H200A}$, $AgrA_{Y229A}$ and $AgrA_{R233A}$

The WT *agrA* gene and the TAD mutants (*agrA*$_{L171A}$, *agrA*$_{E181A}$, *agrA*$_{H200A}$ and *agrA*$_{Y229A}$) were cloned into the pTYB2 plasmid and expressed in the *E. coli* expression strain ER2566 at 37°C and induced with IPTG at mid-log phase, see section 2.3.1.4. Another *agrA* mutant was selected, R233A, which is unable to bind DNA (See section 4.1), to use as a negative control in the *in vitro* assays. The five AgrA proteins were lysed and purified using the IMPACT^{TM} kit, as in section 2.3.4.3. The purification was performed at 4°C and the cleavage reaction was performed at 4°C O/N before the protein was eluted. A sample for the chitin resin was taken and the elution fraction was then reloaded (except for $AgrA_{R233A}$) into the column and incubated for another hour before being eluted again. Another sample from the chitin resin was taken and the elution fraction was concentrated to about 1/12 of the initial elution volume and dialysed O/N at 4°C in a low glycerol (20%) storage buffer.

5.3.1.1. Purification AgrA

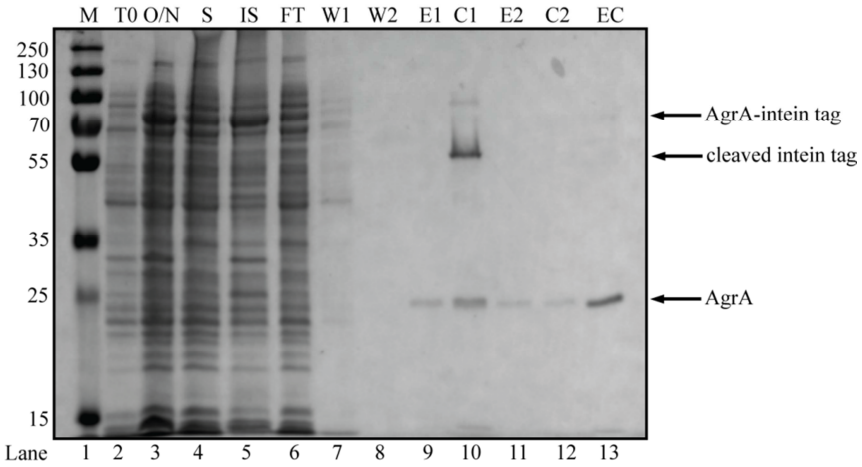


Figure 82: WT AgrA purification.

Image of a 12.5% (w/v) SDS-PAGE gel showing the IMPACT^{TM} purification of AgrA. M = Marker, T0 = before induction, O/N = after induction and expression O/N, S = Soluble fraction, IS = Insoluble fraction, FT = Flow-through, W = Wash, E = Elution, C = Chitin resin, EC = Concentrated elution. The lane numbers and molecular weight of the marker bands in kDa are indicated.

The purification was successful and the native untagged AgrA was successfully eluted from the column (Figure 82, lane 9) with minimal contamination. After the first elution, there was still some untagged protein stuck in the chitin slurry (Figure 82, lane 10). The second elution (Figure 82, lane 11) did not appear to contain more untagged protein than the first one but the chitin resin appeared to contain less protein (Figure 82, lane 12). After dialysis, the

concentration of protein obtained was approximatly 10-20 µM, which was enough for *in vitro* analysis.

5.3.1.2. Purification $AgrA_{L171A}$

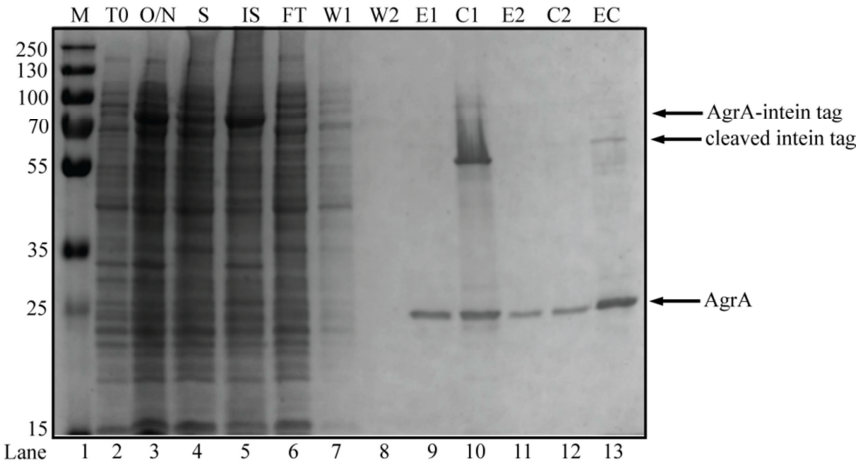


Figure 83: $AgrA_{L171A}$ purification.

Image of a 12.5% (w/v) SDS-PAGE gel showing the IMPACT™ purification of $AgrA_{L171A}$. M = Marker, T0 = before induction, O/N = after induction and expression O/N, S = Soluble fraction, IS = Insoluble fraction, FT = Flow-through, W = Wash, E = Elution, C = Chitin resin, E C = Concentrated elution. The lane numbers and molecular weight of the marker bands in kDa are indicated.

The purification was successful and the native untagged $AgrA_{L171A}$ was successfully eluted from the column (Figure 83, lane 9, 11) with minimal contamination. However, after concentration of the elution fraction, a faint contaminating band, probably corresponding to the cleaved intein tag was observed (Figure 83, lane 13). After dialysis, the concentration of protein obtained was between 10-20 µM.

5.3.1.3. Purification $AgrA_{E181A}$

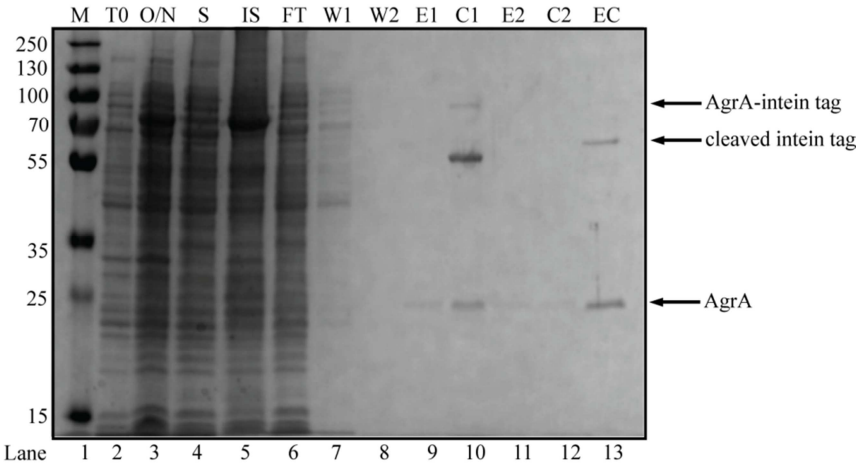


Figure 84: $AgrA_{E181A}$ purification.

Image of a 12.5% (w/v) SDS-PAGE gel showing the IMPACT™ purification of $AgrA_{E181A}$. M = Marker, T0 = before induction, O/N = after induction and expression O/N, S = Soluble fraction, IS = Insoluble fraction, FT = Flow-through, W =Wash, E =Elution, C = Chitin resin, EC = Concentrated elution. The lane numbers and molecular weight of the marker bands in kDa are indicated.

The purification was successful and the native untagged $AgrA_{E181A}$ was successfully eluted from the column but the band was quite faint (Figure 84, lane 9, 11) with minimal contamination. However, after concentration of the elution fraction, a faint contaminating band, probably corresponding to the cleaved intein tag was observed (Figure 84, lane 13). After dialysis, the concentration of protein obtained was ~5 μM.

5.3.1.4. Purification $AgrAH_{200A}$

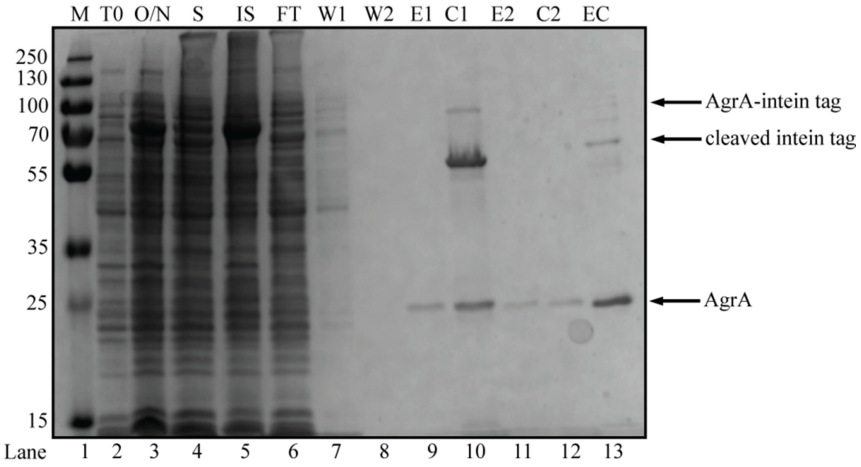


Figure 85: $AgrA_{H200A}$ purification.

Image of a 12.5% (w/v) SDS-PAGE gel showing the IMPACT™ purification of $AgrA_{H200A}$. M = Marker, T0 = before induction, O/N = after induction and expression O/N, S = Soluble fraction, I S = Insoluble fraction, FT = Flow-through, W = Wash, E = Elution, C = Chitin resin, EC = Concentrated elution. The lane numbers and molecular weight of the marker bands in kDa are indicated.

The purification was successful and the native untagged $AgrA_{H200A}$ was successfully eluted from the column (Figure 85, lane 9, 11) with minimal contamination. However, after concentration of the elution fraction, a faint contaminating band, probably corresponding to the cleaved intein tag was observed (Figure 85, lane 13). After dialysis, the concentration of protein obtained was between 10-20 μM.

5.3.1.5. Purification $AgrA_{R233A}$

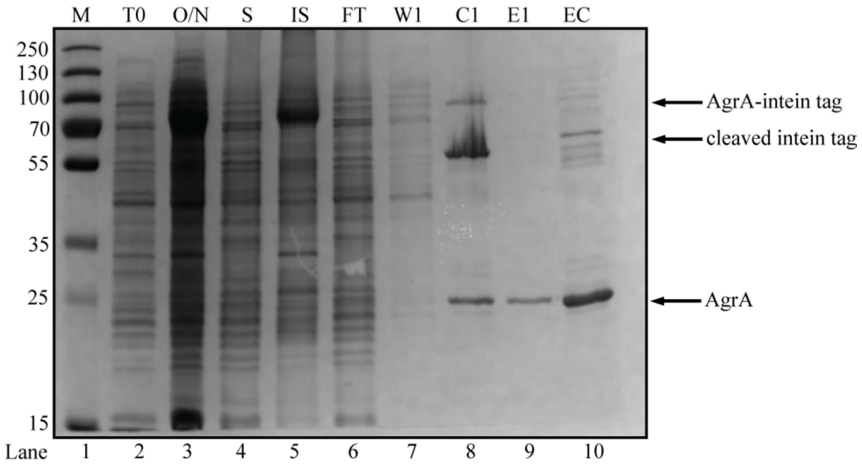


Figure 86: $AgrA_{R233A}$ purification.

Image of a 12.5% (w/v) SDS-PAGE gel showing the IMPACT™ purification of $AgrA_{R233A}$. M = Marker, T0 = before induction, O/N = after induction and expression O/N, S = Soluble fraction, IS = Insoluble fraction, FT = Flow-through, W = Wash, E = Elution, C = Chitin resin, EC = Concentrated elution. The lane numbers and molecular weight of the marker bands in kDa are indicated.

The purification was successful and the native untagged $AgrA_{R233A}$ was successfully eluted from the column (Figure 86, lane 9, 11) with minimal contamination. However, after concentration of the elution fraction, a few faint contaminating bands, probably corresponding to the cleaved intein tag were observed (Figure 86, lane 13). After dialysis, the concentration of protein obtained was between 10-20 μM.

5.3.2. Expression and purification of AgrC-HK

The histidine kinase domain of AgrC (AgrC-HK, aa 206-430) was cloned into the $pET28b^+$ vector to encode an N-terminal 6His tag, expressed in the *E. coli* expression strain BL21 at 37°C and induced with IPTG at mid-log phase, see section 2.3.1.3. The [6His]-AgrC-HK was purified with magnetic nickel beads following the methodology in section 2.3.4.1.

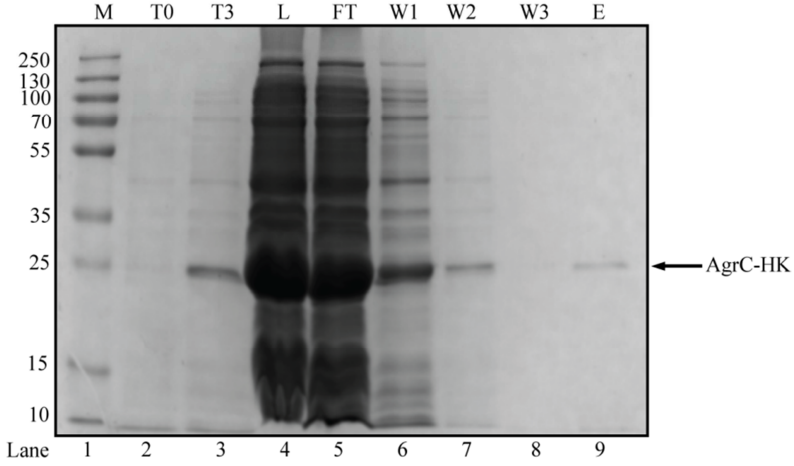


Figure 87: AgrC-HK purification.

Image of a 12.5% (w/v) SDS-PAGE gel showing the magne-His purification of [6His]-AgrC-HK. M = Marker, T0 = before induction, O/N = after induction and expression O/N, L = Lysate FT = Flow-through, W = Wash, E = Elution. The lane numbers and molecular weight of the marker bands in kDa are indicated.

The elution fraction (See Figure 87, lane 9) contained soluble [6His]-AgrC-HK and was dialysed O/N in a storage buffer with low glycerol (10%). After dialysis, the concentration of protein obtained was ~10 μM.

5.3.3. Expression and purification of the *S. aureus* RNA polymerase (RNAp)

The Mab 8RB13 was used to purify *S. aureus* SH1000 core RNAp extracted from exponentially growing cells (See section 2.3.2). The core RNAp was purified by cell lysis and extraction using Polymin P and ammonium sulphate precipitation steps, as in section 2.3.4.2. The extracted fraction was then purified by immunoaffinity chromatography using a column pre-packed with 8RB13 Mab resin.

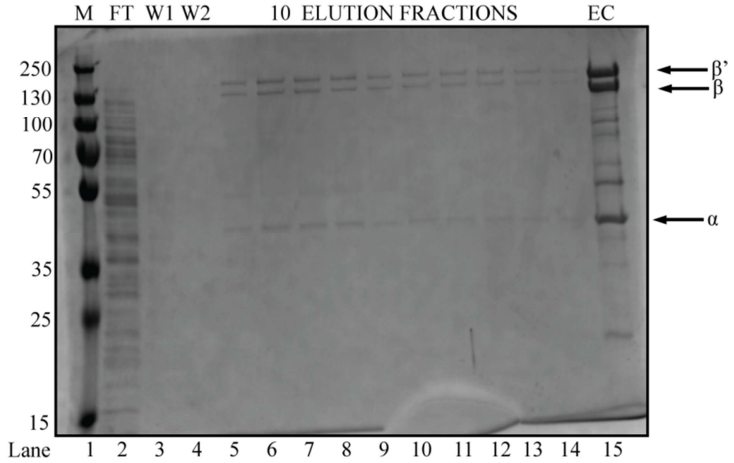


Figure 88: RNAp purification.

Image of a 10% (w/v) SDS-PAGE gel showing the immunoaffinity chromatography of *S. aureus* core RNAp. M = Marker, FT = Flow-through, W = Wash, EC = Concentrated elution. The lane numbers and molecular weight of the marker bands in kDa are indicated.

The *S. aureus* RNAp was successfully purified to a high purity level (Figure 88, lanes 6-14). The elution fractions were pooled together, concentrated (Figure 88, lanes 15) and dialysed O/N in storage buffer.

5.4. Electrophoretic mobility shift assay (EMSA)

In the crystal structure of the AgrA LytTR domain-DNA complex,the 4 putative TAD mutants, aa residues L171, E181, H200 and Y229 are located proximal to the DNA (See Figure 89).

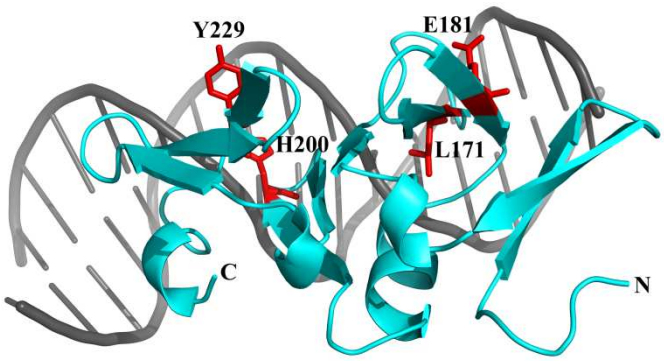


Figure 89: Crystal structure of the AgrA LytTR domain with the 4 TAD mutants.

Ribbon representation of the AgrA LytTR domain-DNA complex crystal structure (the LytTR domain and DNA are coloured in cyan and grey, respectively). The 4 transcription activation deficient (TAD) mutants are coloured in red and labelled in black. Pymol, PDB 3BS1.

We initially wanted to determine, using purified proteins, if alanine substitutions at L171, E181, H200 and Y229 impaired the DNA-binding activity of the mutant AgrA protein, and thereby contributed to the observed TAD phenotype. To test the DNA-binding activity of the L171A, E181A, H200A and Y229A AgrA mutants, we conducted electrophoretic mobility shift based protein-DNA-binding assays (EMSAs) with a 214 bp long DNA probe containing the intergenic region of the *agr* operon (See Figure 81) and with the purified AgrA mutants. Since AgrA requires phosphorylation by AgrC for specific binding to DNA, *in vitro* phosphorylation of AgrA was achieved using the small phosphodonor acetyl phosphate and with AgrC-HK in presence of ATP (See section 5.1.2). The acetyl phosphate activation involved the pre-incubation of AgrA in the presence of acetyl phosphate and $MgCl_2$ for 1h at 37°C (See section 2.4.2). For the activation with AgrC-HK, purified AgrC-HK was incubated for 10 min at RT with AgrA in the presence of dATP (See section 2.4.3). The $AgrA_{R233A}$ mutant served as the negative control in the protein-DNA-binding assays.

5.4.1. EMSA with acetyl phosphate activation

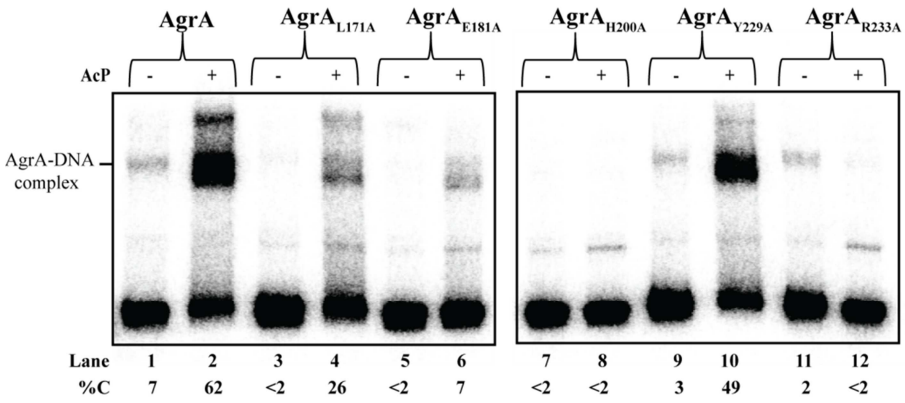


Figure 90: EMSA with acetyl phosphate activation.

Autoradiograph image of a 4-20% (w/v) native polyacrylamide gel comparing the ability of the 4 TAD mutants $AgrA_{L171A}$ (lanes 3,4), $AgrA_{E181A}$ (lanes 5, 6), $AgrA_{H200A}$ (lanes 7, 8), $AgrA_{Y229A}$ (lanes 9, 10) with that of WT AgrA (lane 1, 2) to bind to a DNA probe representing the intergenic region of the *agr* operon in absence (lane 1, 3, 5, 7, 9, 11) and presence (lane 2, 4, 6, 8, 10, 12) of acetyl phosphate. Lanes 11 and 12 contain the DNA-binding $AgrA_{R233A}$ mutant, which is used here as a negative control. The percentage of DNA bound by AgrA (%C) in each reaction is given at the bottom of the gel and was calculated from three independent experiments.

As shown in Figure 90, and as expected, phosphorylation with acetyl phosphate markedly increased the binding of WT AgrA to the DNA probe (Figure 90 lanes 1 and 2) and the $AgrA_{R233A}$ mutant did not detectably bind to the DNA probe in the presence of acetyl phosphate compared to WT AgrA (Figure 90 lanes 11 and 12), thus showing the specific binding of AgrA to the DNA probe in the presence of acetyl phosphate. Of the four putative TAD mutants, only $AgrA_{Y229A}$ bound to the DNA probe to a level comparable to the WT AgrA (49% vs 62%, respectively), whilst $AgrA_{L171A}$, $AgrA_{E181A}$ and $AgrA_{H200A}$ bound to the DNA probe with different degrees of reduced efficiency compared to WT AgrA (Figure 90, lanes 3-8). Overall, the results indicate that an alanine substitution at aa residue Y229 in the LytTR domain of AgrA, whilst only moderately affecting DNA-binding, significantly impairs the ability of AgrA to activate transcription. Further, since the $AgrA_{Y229A}$ mutant only binds DNA in a phosphorylation-dependent manner, we can exclude the possibility that the TAD property of this mutant is due to defects associated with phosphorylation.

5.4.2. EMSA with AgrC-HK activation

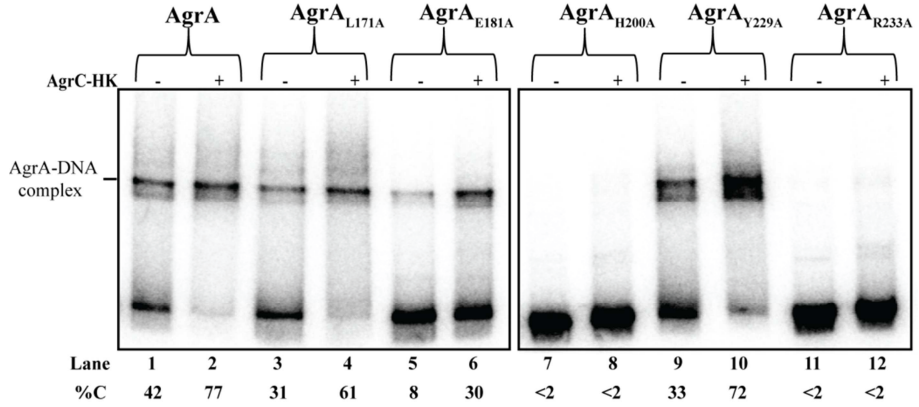


Figure 91: EMSA with AgrC-HK activation.

Autoradiograph image of a 4-20% (w/v) native polyacrylamide gel comparing the ability of the 4 TAD mutants $AgrA_{L171A}$ (lanes 3,4), $AgrA_{E181A}$ (lanes 5, 6), $AgrA_{H200A}$ (lanes 7, 8), $AgrA_{Y229A}$ (lanes 9, 10) with that of WT AgrA (lane 1, 2) to bind to a DNA probe representing the intergenic region of the *agr* operon in absence (lane 1, 3, 5, 7, 9, 11) and presence (lane 2, 4, 6, 8, 10, 12) of AgrC-HK. Lanes 11 and 12 contain the DNA-binding $AgrA_{R233A}$ mutant, which is used here as a negative control. The percentage of DNA bound by AgrA (%C) in each reaction is given at the bottom of the gel and was calculated from three independent experiments.

Overall, the results obtained for the activation with AgrC-HK (See Figure 91) are mostly consistent with those obtained with acetyl phosphate activation (See Figure 90). However,

there is a major difference. With the AgrC-HK activation, the unactivated AgrA samples are able to bind to the DNA probe efficiently (Figure 91, lanes 1, 3, 5, 9) which is not the case with the acetyl phosphate activation (Figure 90, lanes 1, 3, 5, 9). This is probably due to the difference in treatment of the unactivated samples between the two methods. In the case of the acetyl phosphate, the unactivated sample consists of AgrA pre-incubated with $MgCl_2$ for 1h at 37°C without acetyl phosphate. The AgrA protein is highly unstable *in vitro* and the treatment at 37°C for 1h most likely destabilizes the protein. However, the activation state renders the protein more stable. Therefore, the ability of the protein to bind DNA also depends on its ability to be activated by phosphorylation. This explains why there is no DNA-binding activity observed for the unactivated samples with the acetyl phosphate activation.

In the case of the activation with AgrC-HK, the unactivated samples are incubated for 10 min at RT in the absence of AgrC. Because this step is much shorter and at a lower temperature, it does not destabilise the protein hence why the untreated samples retain their ability to bind DNA.

Taken together, the results of the two different EMSAs allow drawing a number of observations. One of the TAD mutants, $AgrA_{H200A}$, is clearly unable to bind DNA under any conditions (Figure 90 and Figure 91, lanes 7 and 8), explaining the very low activity observed *in vivo*. In addition, it behaves exactly like the negative control $AgrA_{R233A}$. Residue R233 is essential in order for AgrA to be able to bind DNA and therefore it appears that H200 could also be an essential residue for the DNA-binding activity of AgrA. This is also supported by the fact that H200 is proximal to the DNA in the crystal structure of the AgrA LytTR domain (See Figure 89).

$AgrA_{E181A}$ binds to DNA with a reduced efficiency (>50%) compared to the WT in any conditions (Figure 90 and Figure 91, lanes 5 and 6). Therefore, the reduced activity of this mutant *in vivo* can be attributed to a defect in DNA-binding. This suggests that this residue is involved in the DNA-binding activity of the AgrA protein, as the position of this residue on the structure suggested (See Figure 89).

$AgrA_{L171A}$ appears to have a more complicated behavior. When activated with acetyl phosphate, it binds to DNA with a reduced efficiency compared to the WT (~40%, See Figure 90, lanes 3 and 4). However, when unactivated or activated with AgrC-HK (See Figure 91, lanes 3 and 4, its DNA-binding ability is much higher, 74% and 79% of the WT, respectively. Therefore, it appears that the decreased ability of $AgrA_{L171A}$ to bind DNA when activated

with acetyl phosphate could be due to a defect in the ability to be activated by phosphorylation.

$AgrA_{Y229A}$ binds to DNA with a slightly reduced efficiency compared to the WT (Figure 90 and Figure 91, lanes 9 and 10) but that is not enough to explain the decreased activity observed *in vivo*. $AgrA_{Y229A}$ is therefore still a putativeTAD mutant.

5.5. *In vitro* transcription assay

To ascertain that $AgrA_{Y229A}$ is a *bona fide* TAD mutant, we tested the ability of the $AgrA_{Y229A}$ to activate transcription from the *agr* operon P3 promoter *in vitro*. The $AgrA_{L171A}$ mutant was also included in the *in vitro* transcription assays as it displayed the best DNA-binding activity (~50% WT activity; Figure 90 and Figure 91, compare lanes 3 and 4) compared to $AgrA_{E181A}$ and $AgrA_{H200A}$ mutants.

5.5.1. Short RNA synthesis assay with AgrC-HK activation

As the *in vitro* activation of AgrA by the HK domain of its cognate two component partner AgrC has never been used for *in vitro* transcription before, the WT AgrA was first tested along with the relevant controls (See section 2.4.4). As shown in Figure 92, the results revealed that this method of activation was flawed in respect to *in vitro* transcription. An increase of 4-fold in transcription is observed under activation conditions (AgrA, AgrC-HK and ATP; Figure 92, lane 5) suggesting that AgrA is activated and activates transcription more than when unactivated (Figure 92, lane 4). However, an equivalent 4-fold increase in transcription is observed in presence of AgrC-HK and dATP but without AgrA (Figure 92, lane 3). This result has no biological relevance and must be due to an artefact of the *in vitro* system. AgrC is a transmembrane protein, therefore it is highly unlikely that it would ever come in contact with the DNA, let alone be involved in transcription activation. Moreover, this would render AgrA useless as it acts as the intermediate between AgrC and the DNA. We therefore moved on to using acetyl phosphate activation only, as it has been successfully used to activate transcription i*n vitro* previously.

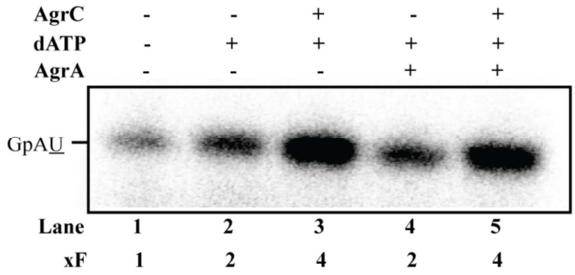


Figure 92: Short RNA synthesis assay with AgrC-HK activation.

Autoradiograph image of a 20% (w/v) urea-denaturing-PAGE gel showing the synthesis of the short RNA product GpAU from the P3 promoter at 25°C, (indicated by a black bar, underlined nucleotide in ^{32}P-labelled. Reactions are conducted in the absence of AgrA, AgrC-HK and dATP (lane 1), in the absence of AgrC-HK and dATP but in presence of AgrA (lane 2), in the presence of AgrC-HK and dATP but without AgrA (lane 3) , in the absence of AgrC-HK but with AgrA and dATP (lane 4), in the presence of AgrC-HK, AgrA and dATP (lane 5). The amount of short RNA product synthesised is expressed as fold increase in short RNA from P3 (xF) relative to the basal level of P3 transcription without any activation (lane 1) with data from three independent experiments.

5.5.2. Full-length *in vitro* transcription assay with acetyl phosphate activation

In order to allow the monitoring of transcription initiation in presence of AgrA, we decided to use full length *in vitro* transcription assay instead of short RNA synthesis assay (See section 2.4.5). Moreover, the effect of AgrA activation of transcription is clearer with this type of assays and this was therefore a better way to test whether $AgrA_{Y229A}$ is deficient for transcription activation.

As shown in Figure 93, phosphorylated $AgrA_{Y229A}$ failed to activate transcription from the P3 promoter to a level comparable to that seen in reactions containing WT AgrA (Figure 93, compare lanes 2 and 3 with lanes 6 and 7). $AgrA_{Y229A}$ is therefore a true TAD mutant.

In contrast, phosphorylated $AgrA_{L171A}$ activated transcription from the P3 promoter to a level almost comparable to that of WT AgrA (Figure 93, compare lanes 2 and 3 with lanes 4 and 5), even though this mutant binds DNA approximately 50-60% less efficiently than the $AgrA_{Y229A}$ mutant or WT AgrA when activated with acetyl phosphate (See Figure 90). This observation supports the theory that $AgrA_{L171A}$ is not truly deficient in DNA-binding but more likely deficient for phosphorylation (See section 5.4.2). Therefore one could imagine a model where $AgrA_{L171A}$ binds DNA less efficiently because of its phosphorylation defect, but will be able to activate transcription almost as well as the WT when it does bind. Therefore, if the system is saturated and there is an excess of AgrA, there might be enough $AgrA_{L171A}$ able to bind to DNA and activate transcription so the amount of transcript produced with $AgrA_{L171A}$ will be very similar to that produced with the WT. This is apparently the case here. As expected, the DNA-binding defective mutant AgrAR233A is unable to activate transcription at the P3 promoter (See Figure 93, lanes 8 and 9).

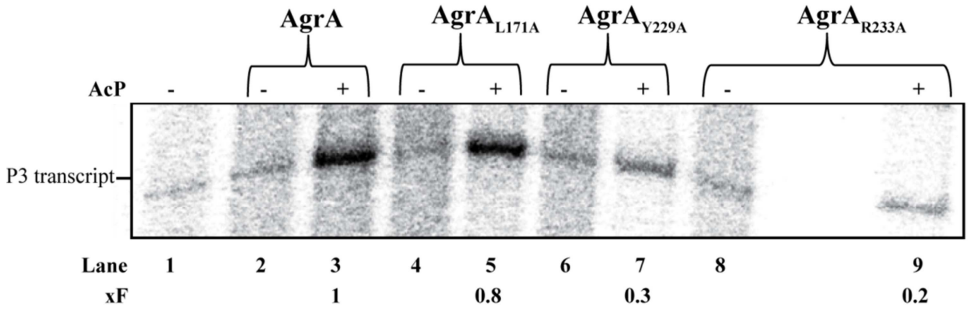


Figure 93: Full-length transcription assay with acetyl phosphate activation.

Autoradiograph image of a 10% (v/v) denaturing urea gel showing the synthesis of the P3 transcript by WT AgrA (lane 2, 3), AgrA$_{L171A}$ (lane 4, 5), AgrA$_{Y229A}$ (lane 6, 7) and AgrA$_{R233A}$ (lane 8, 9) in the absence (lane 2, 4, 6, 8) and presence (lane 3, 5, 7, 9) of acetyl phosphate. The fold decrease in transcription (xF) compared to the P3 transcript in the presence of activated AgrA (lane 3) in each reaction is given at the bottom of the gel with data obtained from three independent experiments.

5.6. *In vivo* analysis: double mutants

In order to attempt to clarify the role of residues L171 and Y229, a series of double mutants were constructed and tested *in vivo* (as in sections 4.4 and 4.5). The double mutants AgrA$_{L171A/Y229A}$, AgrA$_{L171A/Y229F}$, AgrA$_{L171I/Y229A}$ and AgrA$_{L171I/Y229F}$ were tested alongside the relevant single mutants (as in sections 4.7 and 4.5, See Figure 94)

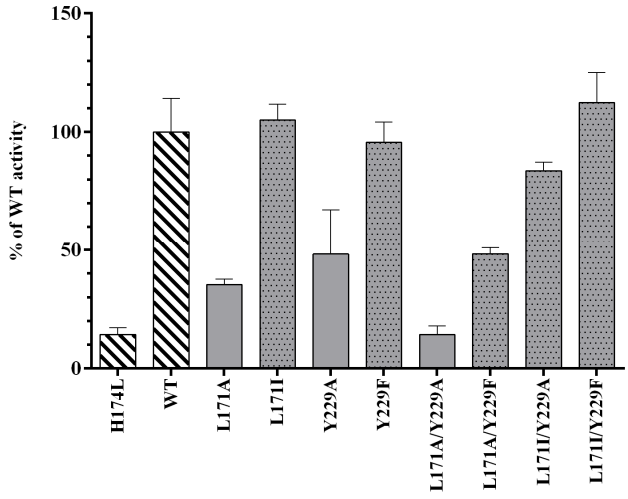


Figure 94: Substitutions and double mutants at positions 171 and 229.

Graph showing GFP expression [as GFP fluorescence units (GFP-Fluo)] as a function of growth ($OD_{600\ nm}$) relative to the SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*agrA* strain for each of the mutants in SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ after 8 h of growth in TSB. The tested mutants are coloured in grey, the alanine mutants without any pattern and the other mutants with a dotted pattern. The controls, WT and H174L are coloured in white with black stripes. Data were obtained from at least three biological replicates.

The mutants AgrA$_{L171I}$ and AgrA$_{Y229F}$ have been tested previously (See section 4.7) but gave a slightly higher activity in this screen, as they are both equivalent to the WT in terms of activity. As discussed previously, substitution of L171 with an hydrophobic, aliphatic and bulky isoleucine does not affect the ability of AgrA to activate transcription. Similarly,

substitution of Y229 by the aromatic phenylalanine, differing from a tyrosine only by a hydroxyl group, did not significantly affect the ability of the $AgrA_{Y229I}$ to activate transcription (See Figure 94).

The double mutant $AgrA_{L171A/Y229A}$ has a very low activity (14%), similar to that of the negative control $AgrA_{H174L}$. This illustrates that the two mutations have a cumulative effect which suggests that these two residues have different functions, as is expected from our analysis so far (See Figure 94).

The double mutant $AgrA_{L171A/Y229F}$ displays 48% activity, which is higher than the single mutant $AgrA_{L171A}$ which displays 35% activity. This suggests that the Y229F mutation has a beneficial effect as it can partially complement for the L171A mutation. If Y229 is indeed important for transcription activation it might be possible that $AgrA_{Y229F}$ is able to activate transcription better than the WT AgrA and therefore complement for the defect in transcription activation caused by the L171A mutation (See Figure 94).

The double mutant $AgrA_{L171I/Y229A}$ displays 83% activity. This is much higher than the activity of the $AgrA_{Y229A}$ mutant (48%). This suggests that the L171I mutation, similar to the Y229F mutation, has a beneficial effect. The presence of this secondary mutation almost doubles the transcriptional activity of the $AgrA_{Y229A}$. If L171A is indeed deficient for phosphorylation as suspected, L171F could be a hyper-phosphorylation mutant or a constitutively active AgrA. Such mechanism could possibly complement the transcription activation detrimental effect caused by the Y229A mutation (See Figure 94).

The double mutant $AgrA_{L171I/Y229F}$ displays 112% activity. This is higher than the WT range of activity. This illustrates that, as expect from the result obtained with the other mutants, the L171I and Y229F mutations are both beneficial and have a cumulative effect, as observed with the double mutant $AgrA_{L171A/Y229A}$ (See Figure 94).

5.6.1. *In vitro* phosphorylation of AgrA

After section 5.6, we strongly suspected that residue L171 is somehow involved in the ability of AgrA to phosphorylate. In order to test that, we decided to develop an assay to look at the ability of AgrA to phosphorylate in order to be able to compare the level of phosphorylation of $AgrA_{L171A}$ with the WT AgrA. As discussed in section 5.1.2, we developed two different methods to activate AgrA with acetyl phosphate and with AgrC-HK. The activation via phosphorylation with AgrC-HK proved to be problematic in the context of *in vitro* transcription but enabled the successful activation of AgrA for the EMSA (See section 5.4.2).

5.6.2. Acetyl Phosphate phosphorylation of AgrA

In order to be able to visualise AgrA activated with acetyl phosphate we used radiolabelled acetyl phosphate (Quon et al, 1996). This method was first tested with the WT AgrA. Potassium acetate (KOAc) was incubated in presence of the acetate kinase and γ^{32}P-ATP in order to create γ^{32}P-acetyl phosphate. The negative control reaction involved KOAc incubated in presence of γ^{32}P-ATP but without the acetate kinase which should result in no γ^{32}P-acetyl phosphate being synthesised. The γ^{32}P-acetyl phosphate or negative reaction was then incubated for 1h at 37°C in presence of AgrA (See section 2.4.7), similar to the activation performed in sections 5.4.1 and 5.5.2. The results, shown in Figure 95, were disappointing. There was only a low level of phosphorylation of AgrA by γ^{32}P-acetyl phosphate (Figure 95, lane 2) which was barely detectable. This might be due to the fact that a low concentration of AgrA was used (1.5 μM final) and can probably be improved with a higher concentration of protein. However, the main problem was that AgrA was able to self-phosphorylate in presence of γ^{32}P-ATP. This did not really make sense therefore we decided not to use this method to test the ability of $AgrA_{L171A}$ to be phosphorylated.

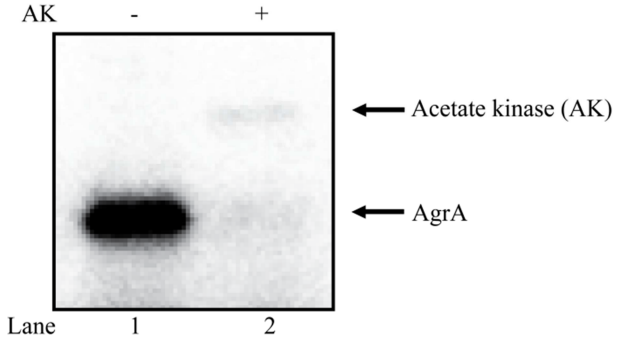


Figure 95: Phosphorylation of AgrA with γ^{32}P-acetyl phosphate.

Autoradiograph of a 4.5% (w/v) native-PAGE gel showing the ability of AgrA to self-phosphorylate in presence of γ^{32}P-ATP and KOAc (lane 1) or γ^{32}P-acetyl phosphate (lane 2). The migratory positions of the acetate kinase and of AgrA are indicated with arrows.

5.6.3. AgrC-HK phosphorylation of AgrA

We then tried to use the ability of AgrC-HK to phosphorylate AgrA in presence of γ^{32}P-ATP to visualise the phosphorylated AgrA. The method was first tested with the WT AgrA. Two separate reactions were incubated for 1h at 30°C. The first reaction contained AgrC-HK, γ^{32}P-ATP and $MgCl_2$. The second reaction contained AgrA and $MgCl_2$. The two reactions were then mixed and incubated for 1h at 30°C (See section 2.4.8). As shown in Figure 96, AgrC-HK was able to autophosphorylate in presence of γ^{32}P-ATP (Figure 96, lane 1) and transfer phosphate to AgrA (Figure 96, lane 2). A Coomassie stain confirmed that the faint

band observed above AgrC-HK was indeed AgrA (Figure 96, lane 2 and 3). Moreover, the band observed for AgrC-HK in presence of AgrA was fainter than the one obtained with AgrC-HK alone (Figure 96, lane 1 and 2) as expected if AgrC-HK is able to phosphorylate AgrA.

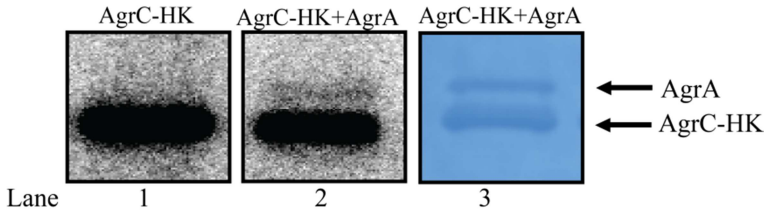


Figure 96: Phosphorylation of AgrA with AgrC-HK.

Lane 1 and 2 are autoradiograph of a 10% SDS-PAGE gel. Lane 1 contains the AgrC histidine kinase domain (AgrC-HK) and lane 2 AgrC-HK and AgrA. Lane 3 is a Coomassie staining of lane 2. The migratory positions of AgrA and AgrC-HK are indicated with arrows.

This result was promising but the ability of AgrAL171A to be phosphorylated by AgrC-HK could unfortunately not be tested as there was not enough time left at this point.

5.6.4. Position of L171A in the context of the predicted structure of the full-length AgrA

As the time restrictions did not allow us to investigate a potential role for residue L171 in phosphorylation via a phosphorylation assay, we decided to take a closer look at its position in the structure relative to the CheY-like phosphorylation domain. However, as the crystal structure available for AgrA does not include the CheY-like domain, we resolved to use a structure prediction software. The **P**rotein **H**omology/Analog**Y** **R**ecognition **E**ngine 2 (Phyre 2) (Kelley & Sternberg, 2009) was used to obtain the predicted structure. The predicted full-length structure of AgrA obtained was modelled on the crystal structure of the full-length LDCTF ComE from *Streptococcus peumoniae* (PDB 4CBV) (Boudes et al, 2014) with 100% confidence and a coverage of 98% (234/238 residues). This predicted structure was used to analyse the position of residue L171 relative to the CheY-like domain and the residues potentially important for the phosphorylation of AgrA (See Figure 97).

The precise site of phosphorylation in AgrA is not known but can be predicted based on the information available on CheY from *Salmonella typhimurium* and *E.coli* as AgrA contains a CheY-like phosphorylation domain. Five residues in CheY have been identified as important for phosphorylation: D12, D13, D57, T87, Y106 and K109. The corresponding residues in

AgrA are D8, D9, D59, T88, F107 and K110. D57, D59 in AgrA is the phosphorylation site (See Figure 97). Refer to section 1.5.5.1 for details.

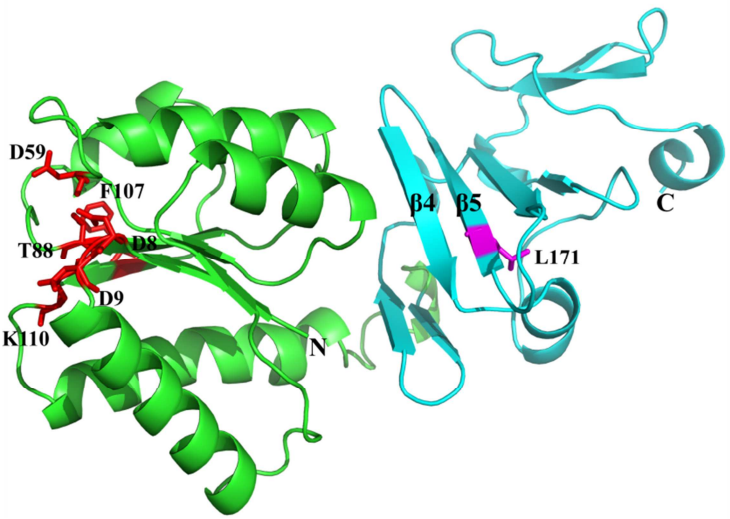


Figure 97: Phyre 2 predicted full-length structure of AgrA.

Ribbon representation of the predicted crystal structure of the full-length AgrA. The CheY-like domain is depicted in green and the LytTR domain in blue. The 6 residues potentially involved in phosphorylation, D8, D9, D59, T88, F107 and K110 are coloured in red and labelled in black. The L171 residue is coloured in magenta and labelled in black. The two β-sheets of the LytTR domain that are the closest to the CheY-like domain are labelled in black β4 and β5.

The predicted structure of AgrA revealed that, while L171A is not located close to the cluster of residues important for activation in CheY (See Figure 97, residues coloured in red), it is located in close proximity to the CheY-like domain. The β4-sheet is the region of the LytTR domain that is the closest to the CheY-like domain and the β5-sheet containing L171 is next in line. Therefore, if L171 is somehow involved in maintaining the spatial localization of β4 and β5, its mutation to alanine might allow these two β-sheets to move away from the rest of the LytTR domain. This movement could induce a steric clash with the CheY-like domain and prevent it from adopting an active conformation.

5.7. *In silico* simulation of the effect of the Y229A mutation (performed by Dr. Robert Weinzierl)

To further interrogate the interactions made by aa residue Y229 in AgrA a collaboration with Dr. Robert Weinzierl, in the Department of Life Sciences at Imperial College was established. Dr Weinzierl carried out a series of fully atomistic accelerated molecular dynamics (aMD) stimulations. aMD is a powerful technique to explore the conformational space available to macromolecules (Markwick & McCammon, 2011) and is thus a particularly effective method for investigating structural changes induced by mutations.

Simulation of the WT AgrA LytTR domain revealed that the core of the structure is very stable and deviates only within the expected range (root-mean square deviation < 2.4 Å) from the original crystal structure during 100 nanosec of aMD simulation. In contrast, simulation of AgrA LytTR domain containing the *in silico* Y229-A substitution under identical conditions shows a substantial destabilization of the carboxyl-terminal domain (See Figure 98 and Figure 99).

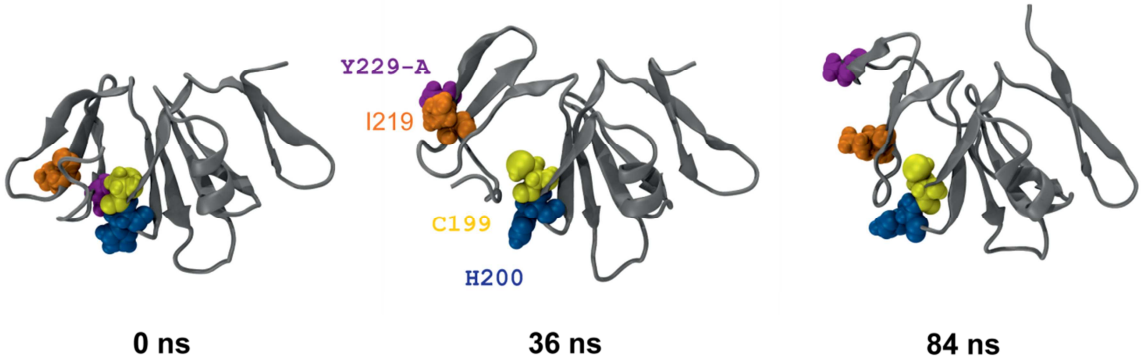


Figure 98: Molecular dynamics analysis of AgrA LytTR domain with the Y229A mutation.

Snapshots of the aMD simulation of Y229-A at different time points. The polypeptide chain of the AgrA LytTR domain is shown as a grey ribbon. The side chains of aa of interest are represented as van der Waals spheres (C199 yellow, H200 blue, I219 orange and the Y229-A *in silico* substitution in purple). At the beginning of the simulation (t = 0 ns), Y229-A is in close contact (as Y229 would be in the WT structure) with C199 and H200. This association dissociates during the course of the simulation, resulting in the carboxyl-terminal three β-sheets to dissociate from the remainder of the LytTR domain (t = 36). At a later stage Y229-A and I219 dissociate from each other (t = 84).

In the WT structure Y229 makes close contact with C199, H200 and I219 based on a dense network of van der Waals-contacts, hydrogen-bonding and hydrophobic interactions. These contacts are severely disrupted in Y229-A, resulting in the rapid dissociation of the three C-terminal β-strands from the main β-sheet transversing the LytTR domain (See Figure 98 and Figure 99, 36 ns). At a later stage of the simulation it becomes also evident that Y229-A further destabilises the two carboxyl-terminal β-strands and causes a reversible unfolding of this sub-structure due to the absence of interaction between I219 and Y229 (See Figure 98 and Figure 99, 84 ns). Taking into account the structural relevance of the interaction between I219 and Y229, it appeared initially puzzling that the $AgrA_{I219A}$ substitution displayed no discernible defect in activating transcription from the P3 promoter (See Figure 64 and section 4.5).

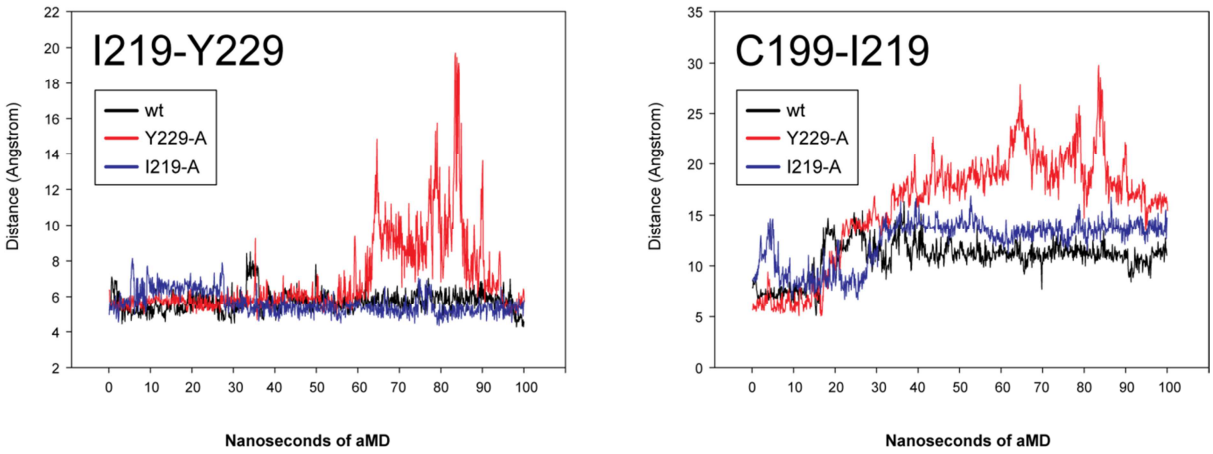


Figure 99: Molecular dynamics analysis of AgrA LytTR domain with the Y229A or I219A mutations.

Quantitation of the aMD results shown in Figure 98: The center-of-mass distances between I219 and Y229-A (left panel) and I219-A (right panel) are plotted against aMD simulation time for the WT (wt) structure (black), Y229-A (red) and I219-A (blue). The distance between I219 and Y229-A increases substantially at one stage of the simulation (from 55 ns onwards), but appears to reform in the final stages of the simulation. In contrast, the WT and I219-A simulations maintain a high degree of stability of this interaction. The distance between C199 and I219 serves as a measure of the dissociation of the three carboxyl-terminal β-sheets.

Simulations of the *in silico* I219-A substitution revealed, however, that in this situation Y229 forms an alternative interaction network involving the adjacent residues E217 and R218. It is therefore evident, from both theoretical as well as experimental observations, that I219 plays a structurally redundant role (which also explains the evolutionary variability in this position (See Figure 72)). Overall the aMD simulations show that an alanine substitution at Y229 is predicted to have a substantial impact on the carboxyl-terminal part of the LytTR domain of AgrA. The mutation causes a 'localised' structural destabilisation that may prevent AgrA from adopting a conformation required for transcription activation.

5.8. DNA bending assay

A previous study by Reyes et al reported that AgrA induced bending of DNA is the main driving force for activation of transcription at the P2 and P3 promoters (Reyes et al, 2011). However, based on the results above (See Figure 94, Figure 98 and Figure 99), it is unlikely that Y299 is involved in the DNA bending activity of AgrA. Therefore, to directly rule out this possibility, we carried out the DNA bending assay previously used by Reyes et al to compare the DNA bending activity of the $AgrA_{Y229A}$ mutant with that of the WT AgrA in the presence of acetyl phosphate. This assay uses the pAM1847 plasmid containing the AgrA tandem binding site upstream of the P2 promoter region cloned between the *Sac*I and *Bgl*I sites (See Figure 100). Digestion of the plasmid with *EcoR*I, *Hind*III, *BstN*I, *EcoR*V, *Nhe*I,

and *BamH*I results in DNA fragments of identical length and composition in which only the position of the AgrA binding site is permuted with respect to the 5' terminus of the fragments (See Figure 100 and section 2.4.6).

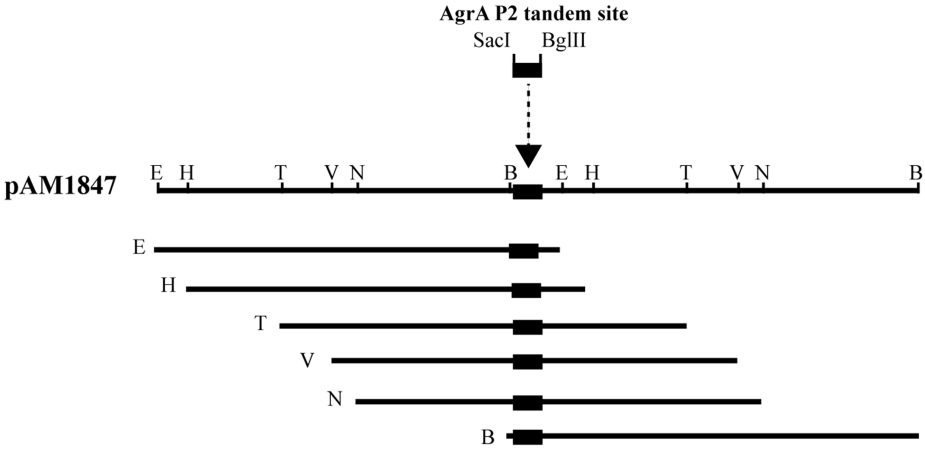


Figure 100: Schematic of the DNA bending vector pAM1847.

Schematic (based on Reyes et al) of the DNA bending vector pAM1847 containing the AgrA P2 tandem binding site (represented by a black rectangle) cloned into the *Sac*I-*Bgl*I site. The *EcoR*I-*Sac*I fragment on the 5' side of the AgrA P2 binding site is identical to the *Bgl*I-*BamH*I fragment on the 3' side of the AgrA P2 binding site. Hence, digestion of the recombinant vector with any of the six restriction enzymes *EcoR*I (E), *Hind*III (H), *BstN*I (N), *EcoR*V (V), *Nhe*I (N) or *BamH*I (B) produces DNA fragments of identical length but with a different position of the AgrA P2 binding site with respect to the 5' and 3' end of the fragment.

Therefore, AgrA-mediated DNA bending can be monitored by non-denaturing polyacrylamide electrophoretic analysis of the AgrA-DNA complexes, as the mobility of these will be strongly dependent on the position of the bend in the DNA molecule as a result of AgrA binding. As shown in Figure 101, no detectable differences were seen in the pattern of the mobility of the WT and mutant AgrA-DNA complexes, thus indicating that $AgrA_{Y229A}$ and WT AgrA bend the DNA equally well and that conserved aa Y229 in the LytTR domain of AgrA is not a major determinant of the DNA bending activity of AgrA.

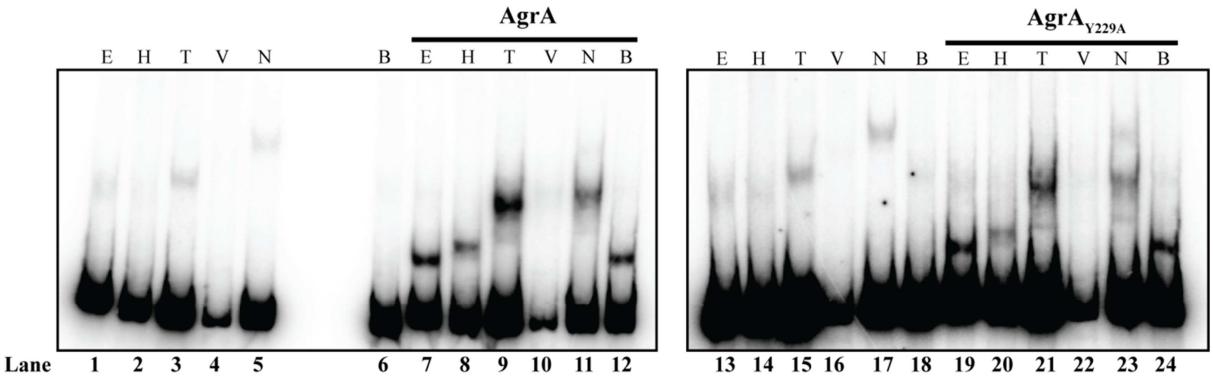


Figure 101: DNA bending assay AgrA and $AgrA_{Y229A}$.
The autoradiograph image of a 4.5% (v/v) non-denaturing polyacrylamide gel shows the mobilities of phosphorylated $AgrA_{Y229A}$ and WT AgrA-DNA complexes bound to each of the DNA fragments generated upon digestion of pAM1847 with

*EcoR*I (E), *Hind*III (H), *BstN*I (N), *EcoR*V (V), *Nhe*I (N) or *BamH*I (B). Lanes 1-6 and 13-18 contain no protein. Data from at least two independent experiments.

5.9. Conclusion

In this section, AgrA was successfully purified using a tag-less method and a new method to phosphorylate AgrA using AgrC-HK was developed. Although this mode of activation proved to be unsuccessful in the case of *in vitro* transcription, it was successfully used to visualise the level of phosphorylation of AgrA by AgrC-HK in presence of γ^{32}P-ATP.

But mostly, our analysis has provided several novel insights into AgrA function in *S. aureus* and other LytTR containing transcription factors:

Alanine substitutions at aa residues E181 or H200 significantly impair the DNA-binding activity of AgrA. Alternatively, alanine substitutions at these aa residues could also indirectly affect the ability of mutant AgrA to become phosphorylated, however, this is unlikely for H200 given its proximity to DNA in the context of the structure of the AgrA LytTR domain-DNA complex and the fact that even when activated via AgrC it cannot bind DNA at all (See Figure 89). However, E181A binds to DNA a little bit better when activated with AgrC. Therefore it could, like L171A, be deficient for activation via phosphorylation. Nevertheless, even when activated with AgrC it still binds DNA 60% less efficiently than the WT and thus is most likely deficient for DNA-binding as well.

The most conserved sequence motif within the LytTR domain is FhRhHRS (where 'h' indicates a hydrophobic aa (See section 1.3.2.1) (Nikolskaya & Galperin, 2002). In AgrA the FhRhHRS corresponds to FFRCHNS (residues 196-202). In the crystal structure of the AgrA LytTR domain-DNA complex only aa residues R198, N201 and S202 were shown to be involved in interaction with the DNA and alanine substitution at N201 reduced the ability of the AgrA LytTR domain to bind DNA (Sidote et al, 2008). These observations clearly imply a role for aa residue H200 within the highly conserved FFRCHNS motif in the binding of AgrA to DNA, and, based on the proximity of H200 to the DNA in the AgrA LytTR domain-DNA complex, base specific contacts between H200 and the DNA cannot be excluded. In further support of this view, alanine substitution of H188 (equivalent residue of H200 in AgrA) in the LytTR domain of *Clostridium perfringens* RR-TF VirR, which regulates virulence and toxin gene expression, confers a loss of activity phenotype and residues R186 and S190 in the FhRhHRS motif of VirR has been shown to be involved in DNA-binding (McGowan et al, 2002) (See section 4.1).To summarize, we have identified two novel residues important for the DNA-binding activity of AgrA, E181 and H200. The mutation of H200 into alanine completely abolishes the ability of AgrA to bind DNA, similar to the effect

of the R233A mutation. Therefore H200 appears to be essential for the DNA-binding activity of AgrA.

The phenylalanine substitution at aa residues L171 in AgrA has been previously reported to confer the non-haemolytic phenotype (indicative of *agr* operon dysfunction) in a nosocomial methicillin-resistant *S. aureus* isolate (Shopsin et al, 2008) (See section 4.1). Intriguingly, our results reveal that even though the $AgrA_{L171A}$ mutant displays a reduced (by ~50%) DNA-binding activity compared to WT AgrA when activated with acetyl phosphate and slightly reduced (80%) when activated with AgrC, its ability to activate transcription is only moderately affected (~20% reduction compared to WT AgrA). Thus, it is possible that an aromatic side chain at position L171 in the LytTR domain is more deleterious for AgrA activity than the presence of alanine. We have hypothesised that the defect in DNA-binding observed for the L171A mutant is due to a defect in phosphorylation. Therefore, if this is the case, maybe the L171F mutation completely abolishes the ability of AgrA to be activated resulting in non-haemolytic phenotype.

Double mutant studies have revealed that the L171I mutation has a beneficial effect as it can partially complement the Y229A mutation. Therefore this mutation probably confers a higher phosphorylation or renders AgrA constitutively active. All these arguments go towards a role for position L171 in activation by phosphorylation. We have shown, using a model of the AgrA full-length structure, that residue L171 is not located close to the determinants of phosphorylation in the AgrA CheY-like domain. However, we have shown that it is located in the second β-sheet the closest to the LytTR domain and have therefore concluded that a disruption of a putative contact between L171A and the β3-sheet could result in more freedom of movement for the β4-sheet containing L171 and the β5-sheet and a steric clash with the CheY-like domain. Interestingly, the β5-sheet contains the E181 residue which also appears to show a defect in phosphorylation.

Alternatively, L171 could be important in order for AgrA to assume the active conformation. The activation by phosphorylation results in changes in the CheY-like domain but the LytTR domain must be slightly modified as well as its ability to bind DNA is increased when active. Therefore, it is possible that residue L171 is an important determinant for reaching the active conformation and that when mutated to alanine AgrA cannot be fully active. Similarly, the L171I mutation could facilitate the transition to the active state or further stabilise it and the E181 residue could also be involved.

Importantly, the results identify the highly conserved aa residue Y229 in the LytTR domain of AgrA as a major determinant for transcription activation by AgrA. When mutated into an

alanine, a drop of 70% for transcription activation *in vitro* is observed but this residue is not involved in the DNA bending activity of AgrA. Interestingly, in the structure of the AgrA LytTR domain DNA-complex, Y229 is adjacent to aa residues critical for AgrA folding (C228) and DNA interaction (R218), thus indicating that Y229 is part of a functional hotspot in AgrA (See section 4.1). Consistent with this notion, a recent study reported that a novel antibacterial compound, called savarin, binds to the LytTR domain of AgrA in a region proximal to aa residues C228, R218 and Y229 and thereby abrogates AgrA function (Sully et al, 2014) (See section 4.1). The proximity of both AgrA binding sites to the P2 and P3 core promoter elements (See Figure 102) suggests that transcription activation at both promoters could involve direct protein-protein interaction involving aa residue Y229 in AgrA and the RNAp and thus could occur via a simple 'recruitment' mechanism, whereby AgrA could facilitate the binding of the RNAp to the promoter to yield a transcriptionally-proficient promoter complex or via a class I or class II mechanism (See section 1.2.2.1 and 1.2.2.2). Moreover, double mutants studies have shown that the Y229F mutation has a beneficial effect as it can partially complement the L171A mutation. Therefore if Y229 is involved in recruiting or contacting the RNAp maybe the Y229F mutant can recruit the RNAp better or have a more stable interaction resulting in more transcription. These theories will be further developed in the next chapter (See chapter 0).

A previous study by Reyes et al reported that transcription from P2 and P3 *agr* operon promoters is differentially regulated, with the former (P2) dependent on AgrA and SarA and the later (P3) dependent only on AgrA for maximum promoter activity (Reyes et al, 2011) and the SarA binding site overlaps with the AgrA footprint at the P2 promoter (See Figure 102). Therefore, it is possible that Y229 in AgrA is a determinant for interaction with SarA for the activation and that the $SH1000^{-}_{\text{agr IR P3-GFP}}$ reporter strain used in this study indirectly indicates P2 activity (which drives its own transcription – see section 1.5). However, this is unlikely because $AgrA_{Y229A}$ mutant displays the same level of activity in the context of pSN-P2-*agrA* and pSN-*itet-agrA* in $SH1000^{-}_{\text{agr IR P3-GFP}}$ strain and because the Y229A is deficient for transcription activation *in vitro* compared to the WT in the absence of SarA.

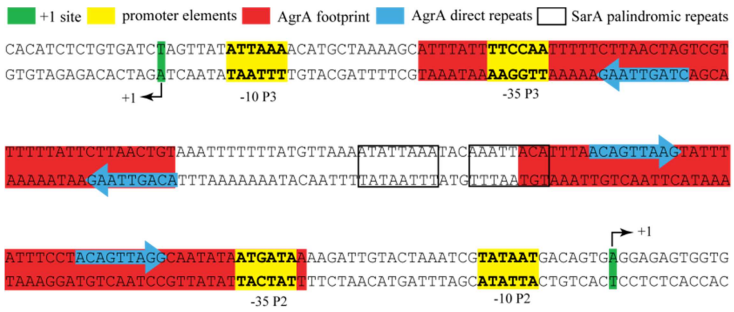


Figure 102: AgrA binding sites and footprint at the P2 and P3 promoter.

The nucleotide sequence of the *S. aureus agr* operon intergenic region with the different regions of relevance indicated. -35 refers to the -35 element (5'-TTGACA-3') and -10 refers to the -10 element (5'-TATAAT-3'). The published -10 and -35 promoter elements are highlighted in yellow and the SarA palindromic repeats boxed in black. Adapted from (Koenig et al, 2004).

The aMD analysis reveals that an alanine substitution at Y229 could cause a 'localised' structural destabilisation that may prevent AgrA from adopting a conformation required for making interactions with the RNAp and/or other co-activators for transcription activation from the P2 and P3 promoters. However, as with SarA, the defect in transcription activation observed for Y229A compared to the WT is unlikely to be due to an inability to interact with another factor binding to the *agr* IR as the defect is also seen *in vitro* when no other factors are presents. In fact, the only external factor that is present in the *in vitro* system is the RNAp therefore the most likely explanation for the Y229A defect in transcription is the lack of an interaction with the RNAp that takes place with the WT AgrA.

The main results obtained in this chapter will be further discussed and developed in the next chapter, chapter 0.

CHAPTER SIX

6. Final conclusions and overview

6.1. Introduction

The LytTR domain is an unusual DNA-binding domain that is present in ~3% of all RRs. Unlike other DNA-binding domains, the LytTR domain consists mainly of β-sheets (See Figure 26) and binds to the DNA via 3 loops that connect these β-sheets. Interestingly, LytTR domain-containing transcription factors (LDCTFs) are disproportionally associated with the regulation of virulence genes (Nikolskaya & Galperin, 2002) (See section1.3.2.1).

The best characterised LDCTF is AgrA from the important human pathogen *Staphylococcus aureus* (*S. aureus*) and the only one for which the DNA-binding mechanism has been demonstrated. AgrA is a key regulator of virulence gene expression in *S. aureus* and is responsible for the downregulation of cell surface-associated virulence factors and the upregulation of secreted virulence factors that takes place in a growth phase and cell density-dependent manner thanks to a two component quorum sensing system (TCS). The tight regulation of virulence factors operated by the *agr* operon allows *S. aureus* to switch its behaviour from a colonisation state to a more invasive state once a certain threshold of population has been reached and contributes to the success of *S. aureus* as a pathogen and its ability to adapt rapidly to changing environments (Novick & Geisinger, 2008).

The *agr* operon consists of two transcriptional units: RNAII and RNAIII encoded by two divergent promoters: P2 and P3, respectively (See Figure 103 and section 1.5).

RNAII consists of the quorum sensing system, genes *agr*B, D, C and A. AgrD is the precursor of the autoinducing peptide (AIP) which is processed and secreted by the transmembrane endopeptidase AgrB with the help of another peptidase, SpsB releasing the finalised AIP. AgrC and AgrA are members of a TCS in which AgrC is the transmembrane receptor histidine kinase which senses the growth phase and cell density-dependent accumulation of the AIP. Upon binding the AIP, AgrC autophosphorylates and is then able to activate the DNA-binding RR AgrA via phosphotransfer. The activated AgrA is then able to bind to specific sequences within the *agr* intergenic region (IR) and activate transcription at the P2 and P3 promoters. Like many other RRs, AgrA therefore stimulates its transcription in a feed-back loop mechanism.

RNAIII is the effector RNA molecule of the agr operon and is responsible for the regulation of over 200 genes (Novick & Geisinger, 2008).

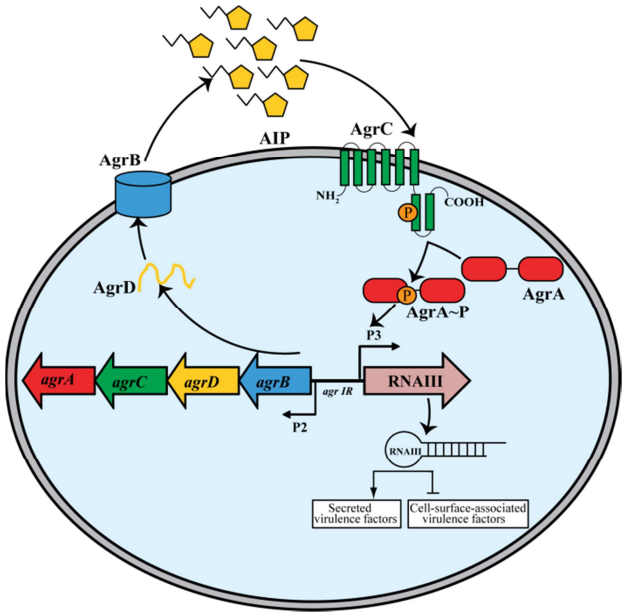


Figure 103: Schematic representation of the organisation of the *agr* operon and the auto-catalytic circuit.

AgrD encodes a propeptide that is processed and secreted by the transmembrane AgrB with the help of SpsB releasing the finalised AIP. AgrC senses the cell density-dependent accumulation of the AIP and autophosphorylates upon binding the AIP. The phosphorylated AgrC is then able to activate AgrA via phosphotransfer. The phosphorylated AgrA is then able to bind to specific direct repeats at the *agr* IR and activate transcription at the P2 and P3 promoter. The P3 promoter encodes RNAIII, the pleiotropic effector RNA of the operon. RNAIII is responsible for the upregulation of secreted virulence factors and the downregulation of cell surface-associated virulence factors. By activating the P2 promoter, AgrA creates an autocatalytic circuit as it stimulates its own transcription.

The P2 and P3 promoters present a non-optimal spacer length of 18 and 20 nt, respectively and AgrA was shown to induce a significant bend in the promoter DNA upon binding (See Figure 104) (Sidote et al, 2008). The artificial shortening of the P3 spacer *in vivo* was shown to significantly increase its activity and *in vitro* to result in a level of transcription equivalent as when activated by AgrA (Morfeldt et al, 1995; Reynolds & Wigneshweraraj, 2011). Together, these findings have led to the theory that the bend in the DNA induced by AgrA is responsible for its ability to activate transcription by a typical remodelling of the promoter type of activation (See Figure 7)

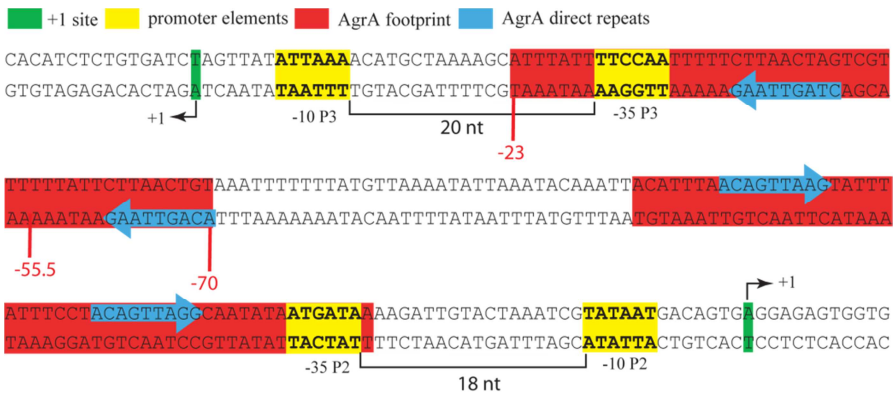


Figure 104: AgrA binding sites and footprint at the P2 and P3 promoters DNA sequence of the *agr* IR.

The nucleotide sequence of the *S. aureus agr* operon intergenic region with the different regions of relevance indicated. -35 refers to the -35 element (5'-TTGACA-3') and -10 refers to the -10 element (5'-TATAAT-3'). The published -10 and -35 promoter elements are highlighted in yellow. The nucleotide length of the spacer between the -10 and -35 promoter elements is indicated for the P2 and the P3 promoter. The end and start of the AgrA footprint and the centre of the AgA direct repeats at the P3 promoter are indicated in red. Adapted from (Koenig et al, 2004).

The *agr* operon has been extensively characterised since its discovery nearly 30 years ago and a number of TFs acting to regulate its expression have been identified. Nevertheless, the exact mechanism of transcription activation by AgrA remains elusive.

6.2. Summary of findings

The objective of this study was to try to clarify the mechanism of transcription activation by LCDTFs at their target promoter using AgrA as a model. The first step involved the characterisation of the SH1000$^-$ strain, previously identified as a possible *agrA*- strain due to a spontaneous mutation. This strain was shown to be *agrA* deficient due to the mutation H174L in the AgrA LytTR domain. The SH1000$^-$ strain was successfully complemented using a plasmid-borne AgrA, pSN-P2-*agrA* therefore demonstrating that the H174L mutation was causing the *agrA*- phenotype. Residue H174 was shown to be involved in two salt bridges interactions with residues E163 and E226 in the crystal structure of the AgrA LytTR domain (Sidote et al, 2008). These salt bridges stabilise the interactions between the β-sheets β4 and β3 (H174-E163) and between the β-sheets β4 and β10 (H174 and E226). As illustrated in Figure 105, these two salt bridges interactions are the only ones taking place on this face of the protein, the only ones linking the β-sheet β10 to the rest of the protein and the only ones connecting the β-sheets β4 and β3. Therefore, this observation and the fact that the $AgrA_{H174L}$ was barely detectable by Western blot using a polyclonal antibody again AgrA led us to conclude that these salt bridges interactions are essential in order for AgrA to adopt the correct structural fold and that residue H174 is thus essential for AgrA function. This information is novel as this residue has never been identified as being essential for AgrA function before. The fact that this mutation was the result of a spontaneous mutation indicates that this position is sometime targeted by bacteria in order to inactivate the *agr* response.

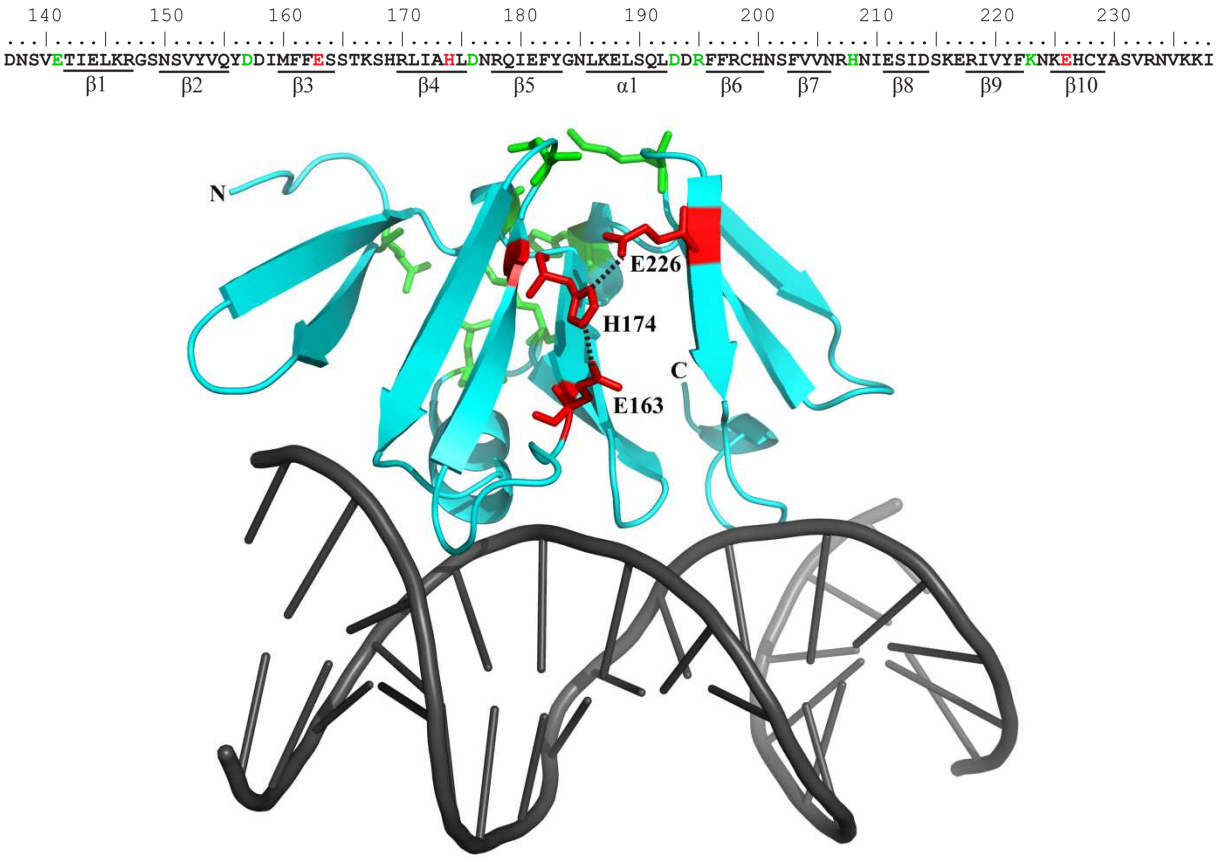


Figure 105: Salt bridges interactions in the AgrA LytTR domain with H174, E163 and E226 in red.

Sequence in amino acid (aa) of the AgrA LytTR domain and ribbon representation of the AgrA LytTR domain-DNA complex (the LytTR domain and the DNA are coloured in cyan and grey, respectively). H174, E163 and E226 are coloured in red. They are labelled in black and the salt bridges interactions are represented by a black dotted line on the structure. The other 7 residues involved in salt bridges interactions: R195, E141, D193, D157, H208, K223 and D176 are coloured in green. The N- and C-terminal ends of the protein are labelled in black on the structure. The different β-sheets and the α-helix are labelled and numbered in the sequence. Pymol, PDB 3BS1.

In order to try to understand the mechanism of transcription action by AgrA by elucidating the individual contributions of the aa in the AgrA LytTR domain to transcription activation, a systematic mutational analysis of the AgrA LytTR domain was performed. A library of 81 single aa mutants (78 single alanine) in pSN-P2-*agrA* targeting 80% of the residues of that domain was therefore constructed by site directed mutagenesis.

A P3-GFP reporter system that had been previously developed in the lab (James et al, 2013) was transduced into the SH1000^{-} strain creating the SH1000$^{-}_{agr\ \mathrm{IR\ P3\text{-}GFP}}$. The SH1000$^{-}_{agr\ \mathrm{IR\ P3\text{-}GFP}}$ was transformed with each of the library mutant plasmid and the whole library was tested by growth curves using a multiwall plate reader measuring GFP fluorescence units (GFP-Fluo) and $OD_{600\ nm}$. This system enabled the quantification of the ability of each library mutant AgrA to activate transcription at the P3 promoter by measuring the GFP-Fluo emitted by the SH1000$^{-}_{agr\ \mathrm{IR\ P3\text{-}GFP}}$ strain complemented with each mutant plasmid and comparing it to

the WT AgrA complementation level (See Figure 64). The screening of the library resulted in a wide range of transcriptional activities and demonstrated that this method is highly sensitive and quantitative and enables to discriminate between subtle differences in activity. We have therefore developed an efficient and simple method to quantify AgrA transcriptional activity at the P3 promoter *in vivo*.

The screen enabled the identification of 25 putative transcription activation defective (TAD) mutants that displayed a reduced ability to activate the P3 promoter by at least 40% compared to WT AgrA. The structural stability of these mutants was then tested by western blot in order to exclude the mutants whose decrease in activity resulted from structural instability.

11 of the putative TAD mutants were not detectably expressed (similar to H174L level) and were therefore concluded to be structurally impaired. These residues are probably important in order to maintain the structural fold of the protein. This is a new finding as none of these residues have been identified as important for the activity of the AgrA protein before. However, 2 of these residues: R198 and R218 have been shown to make contact with the DNA. Interestingly, the essential DNA-binding residue R233 was also barely detectable suggesting that this residue is also involved in maintaining the structural fold of the AgrA protein.

For 9 of the remaining putative TAD mutants, the decrease in transcriptional activity appeared to correlate with an impaired stability. These mutants were therefore excluded as well.

After this stringent selection, 4 putative TAD mutants remained: L171A, E181A, H200A and Y229A. They were purified alongside the DNA-binding defective control mutant R233A using a tag-less method and characterised *in vitro* by EMSA. The EMSAs revealed that all mutants but Y229A had a reduced DNA-binding ability compared to the WT. E181A was severely impaired in its DNA-binding ability with a 90% reduction compared to the WT with acetyl phosphate activation. The H200A was completely unable to bind DNA, as the R233A mutant. We have therefore identified two residues involved in DNA-binding in AgrA: E181 that is important and H200 that is essential for the DNA-binding activity of AgrA. Therefore both of these residues are most likely involved in the interaction between AgrA and the DNA. Synonymous mutations performed at both these positions revealed that no other aa is tolerated at these positions and that these residues at very conserved among LDCTFs which highlights their importance.

The LytTR domain binds to DNA though three loops that insert into three subsequent grooves in the DNA (two minor and one major). Two residues, located on two of these loops, H169 and R233 have been shown to be essential in order for the protein to be able to bind to DNA. However, the residue located on the third loop inserting in the minor groove, N201, was shown to make direct contact with the DNA and to be important in order for AgrA to bind DNA, but not essential. H200 is located in this third loop, immediately before N201 and appears to be essential. Therefore, it is conceivable that H200 is the main residue responsible for the interaction between the AgrA LytTR domain and the DNA in the third loop, the third "finger" essential to correctly position the LytTR domain onto the DNA (See Figure 106).

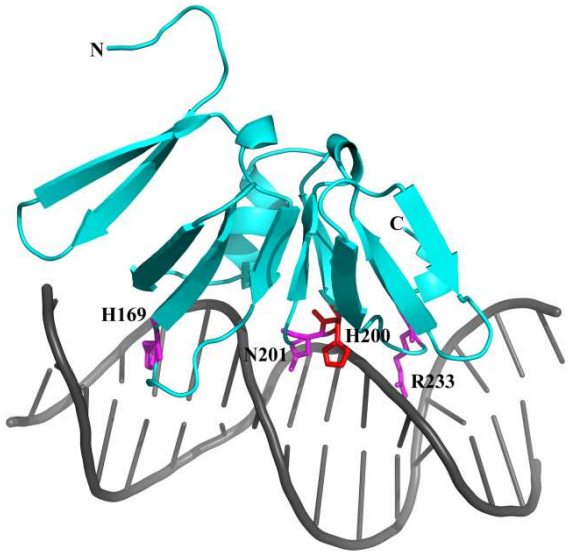


Figure 106: Crystal structure of the AgrA LytTR domain with H200 and the three residues directly interacting with the DNA.

Ribbon representation of the AgrA LytTR domain-DNA complex crystal structure. The LytTR domain is coloured in cyan and the DNA in grey. H200 is coloured in red and labelled in the black. The three residues making direct contact with the DNA are coloured in magenta, displayed as sticks and labelled in black. The N- and C-terminal ends of the structure are indicated in black. Pymol, PDB 3BS1.

The FhRh[RK][SNG]hhVN motif (See Figure 107) was identified as the most conserved motif among LDCTFs (Nikolskaya & Galperin, 2002) and as involved in DNA-binding in VirR (McGowan et al, 2002) is also very important in *S. aureus* and involved in DNA-binding. The essential DNA-binding residue identified, H200 is part of that motif and two residues that were shown to make contact with the DNA, R198 and R218 as well. The N206 residue was identified as resulting in structural instability in the screen (See Figure 107). Therefore, judging by the high degree of conservation of this motif among LDCTFs (See Figure 59), it highly probable that the involvement of this motif in DNA-binding is conserved among LDCTFs. This would implicate that H200 could also be essential for the DNA-binding activity of other LDCTFs.

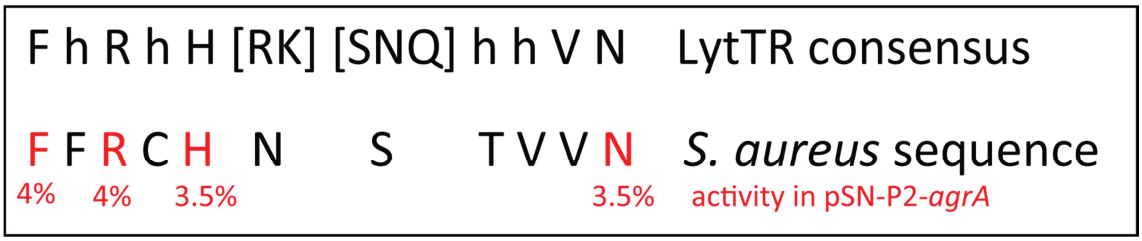


Figure 107: Residues in the FhRh[RK][SNG]hhVN motif identified in the screen.

The residues displaying 4% activity or less in the screen and therefore thought to be important in *S. aureus* are coloured in red. Their percentage of activity in pSN-P2-*agrA* is indicated in red as well.

The last remaining TAD mutant, Y229A was used in an *in vitro* transcription assay to assess whether it was able to activate transcription at the P3 promoter. The L171A mutant was tested as well as it had displayed the best DNA-binding activity after Y229A and was suspected to be deficient for activation via phosphorylation.

The Y229A mutant was unable to activate transcription at the P3 promoter to a level equivalent to the WT AgrA. It was able to activate transcription i*n vitro* to ~30% of the WT activity which correlates with the *in vivo* observed activation that was 40-50% of that of the WT. Y299A is therefore the only true TAD mutant that was identified by our screen. Interestingly, another mutation at that position, Y229F, appears to have the opposite effect and to have a higher ability to activate transcription than the WT AgrA. When combined with the L171A mutation, the Y229F mutant can compensate for this mutation and the level of transcription activation is higher than in the single L171A mutant. Importantly, this result identifies the highly conserved Y229A residue in the AgrA LytTR domain as a key determinant for maximal transcription activation by AgrA.

Surprisingly, the L171A was able to activate transcription *in vitro* to about 80% of the WT ability (See Figure 93). This corresponds to the 20% drop in the ability to bind DNA observed when the L171A mutant was activated with AgrC and further strengthens the case for L171A being deficient for phosphorylation. The L171A mutant displayed a reduced ability to bind DNA when activated by acetyl phosphate that appeared to be due to a deficient activation. Therefore we hypothesised that this position might be important for activation via phosphorylation in the LytTR domain. Upon phosphorylation, a RR undergoes structural changes and adopts the active state. These structural changes are mainly localised in the receiver domain of the RR (Lee et al, 2001) but must affect the DNA-binding domain as well as the ability of the RR to bind DNA is dependent on its activation. Therefore it is conceivable that position L171 is important in order for AgrA to reach the active state and

that its mutation to alanine prevents AgrA from being fully active. Therefore the DNA-ability of the L171A is reduced as it is not fully active.

The L171I mutant appears to be more transcriptionally active than the WT AgrA as it can partially compensate for the Y229A mutation in a double mutant. Therefore having an isoleucine at position L171 appears to be more advantageous than a leucine. Following the theory that position L171 is important in order to reach the active conformation, the L171I mutant could be hyper-phosphorylated or a constitutively active mutant. However, these theories are purely speculative now and further work needs to be done in order to confirm or infirm them.

Having identified a true TAD mutant we decided to try to understand the nature of the inability of the Y229A mutant to activate transcription to WT level

An in silico analysis of that mutant was performed by Robert Weinzierl using molecular dynamics demonstrated the importance of that residue to connect the C-terminal part of the protein to the rest. It predicted that the Y229A mutation creates a local structural destabilisation which probably explains the reduced protein stability observed compared to the WT AgrA by western blot.

The ability of AgrA to activate transcription at the P3 and P2 promoter is thought to be the result of its ability to bend DNA. We therefore tested the ability of the Y229A mutant to bend DNA at the P2 promoter using an *in vitro* DNA bending assay and showed that the Y229A while unable to properly activate transcription at the P3 promoter retained its DNA bending ability (See Figure 101). This result brings to question the current belief that the transcription activation function of AgrA is solely due to its ability to bend the promoter DNA. Therefore, we propose a new model for transcription activation via AgrA that takes into account our new findings. As the AgrA footprint at the P2 and P3 promoter overlap the -35 promoter element, an interaction between AgrA and the RNAp is conceivable. Therefore activation by AgrA would implicate remodelling of the promoter by DNA bending (See section 1.2.1) and direct interaction via Class II-like mechanisms. In this case, AgrA could potentially interact with σ4, α-CTD or α-NTD (See section 1.2.2.2). Moreover, we have observed that the Y229F mutant has a higher ability to activate transcription than the WT AgrA. Therefore, If Y229A does indeed interact with the RNAp, the Y229F probably has a better ability to interact than the WT AgrA.

The crystal structure of the AgrA-LytTR domain/DNA complex indicates that the N-terminal part of the LytTR domain is directed towards the -35 promoter element (Sidote et al, 2008) (See Figure 108).

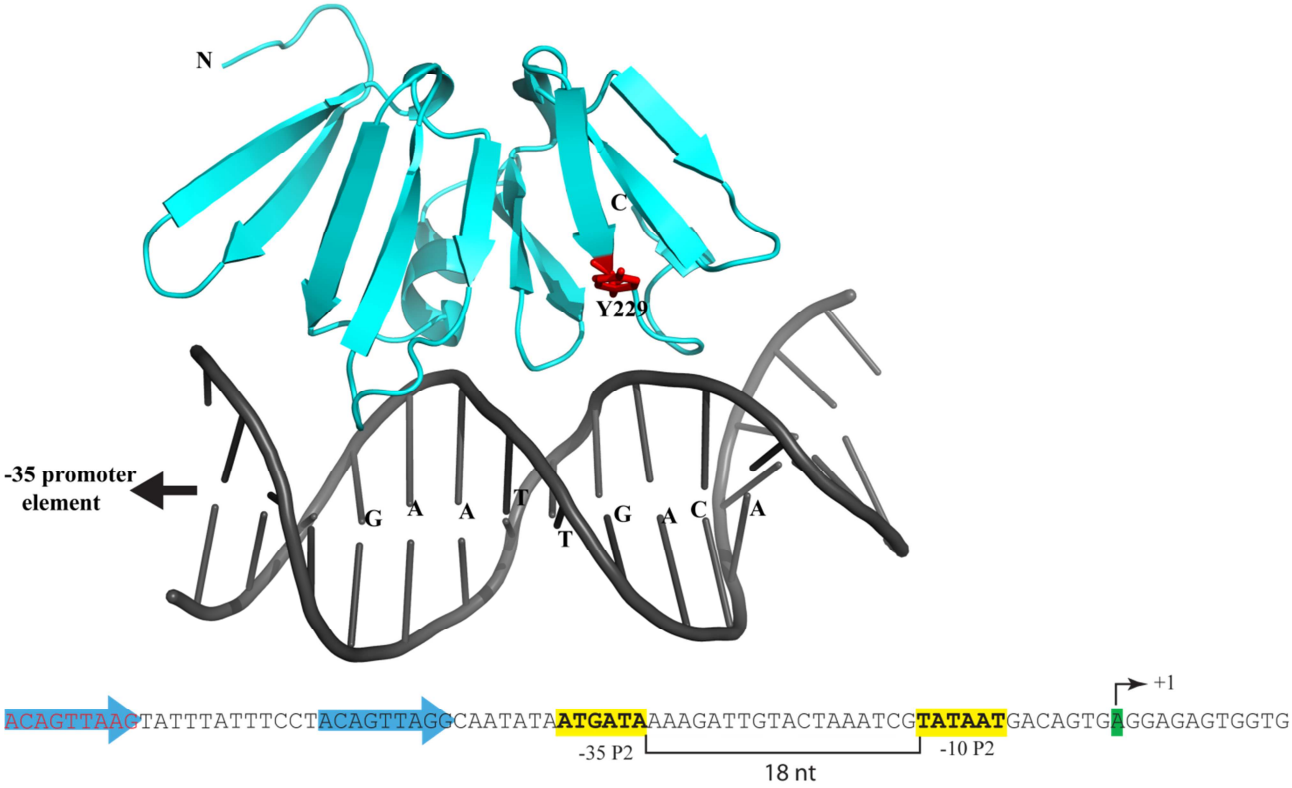


Figure 108: Crystal structure of the AgrA LytTR domain in complex with its DNA-binding sequence and the P2 promoter.

Ribbon representation of the AgrA LytTR domain-DNA complex (the LytTR domain and the DNA are coloured in cyan and grey, respectively) and DNA sequence of the P2 promoter. Y229A is highlighted in red and labelled. The consensus AgrA binding sequence is annotated on the corresponding base in the DNA on the structure and the direction of the -35 promoter element is indicated with an arrow. The sequence of the P2 promoter with the AgrA binding sites (the blue arrows) is displayed under the structure. The binding site corresponding to the structure binding site is coloured in red. -35 refers to the -35 element (5'-TTGACA-3', in yellow) and -10 refers to the -10 element (5'-TATAAT-3', in yellow). The +1 site is highlighted in green and the spacer length between the -10 and the -35 promoter elements is indicated in black. Pymol, PDB 3BS1.

As Y229A is located in the C-terminal part of the protein, it seems unlikely that this residue would be involved in an interaction with σ4. Therefore the best candidate to interact with AgrA would be the α-CTD subunit. The AgrA footprint at the P3 promoter extends from position -23 to position -70 and the centre of the direct repeats is at -55.5 (See Figure 104). Therefore, the α-CTD subunit could only bind after position -70. It has been demonstrated that activation via interaction with the α-CTD subunit can happen until position -90 as the linker joining it to the α-NTD subunit is very flexible (Gaston et al, 1990; Lee et al, 2012; Ushida & Aiba, 1990). Therefore, our model of activation by AgrA is represented in Figure 109.

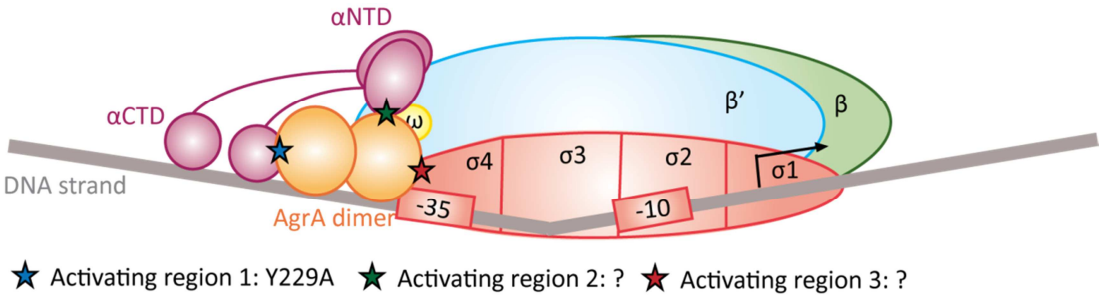


Figure 109: Model of transcription by AgrA.

AgrA bends the DNA upon binding and Y229A interacts with the α-C-terminal domain (α-CTD) by a Class II mechanism. AgrA could therefore also interact with the α-N-terminal domain (α-NTD) and region 4 of the σ factor (σ4).

AgrA makes a contact with the RNAp α-CTD subunit via residue Y229A. Such interaction could take place before the AgrA dimer binds the promoter DNA and therefore act as a pre-recruitment mechanism (See Figure 12) or it could stabilise the interaction between the RNAp and the promoter DNA (See Figure 9). In this model, residues located in the N-terminal part of the AgrA LytTR domain could possible interact with σ4 or the α-NTD subunit. In addition, the δ subunit of the *S. aureus* RNAp was recently shown to be involved somehow in promoter recognition as it appears to direct the RNAp towards strongly expressed promoters. Therefore the δ subunit could also be considered as a candidate region for interaction.

We have evidence to support that AgrA does indeed interact with region 4 of σ. Preliminary results of a bacterial adenylate cyclase two-hybrid (BACTH) screen have identified a weak interaction between the AgrA LytTR domain and σ4 (See Figure 111). The BACTH system enables the detection of the interaction of two proteins of interest fused to the complementary fragments of the adenylate cyclase from *Bordetella pertussis* T25 and T18 *in vivo* in an *E. coli* strain where the adenylate cyclase gene has been deleted, BTH101. When the two domains T25 and T18 are fused to two interacting proteins, the T25 and T18 fragments can interact and the adenylate cyclase is functional and can produced cyclic AMP (cAMP). The cAMP produced interacts with the catabolite activator protein (CAP) and forms a cAMP/CAP complex that can activate the transcription of many *E. coli* genes. In this case, the complex will activate transcription of the *lac* operon involved in lactose catabolism and containing the *lacZ* gene encoding the β-galactosidase enzyme (See Figure 110). The amount of β-galactosidase enzyme produced is proportional to the ability of both of the target protein to interact with each other and can be quantified in Miller Units. When the two domains T25 and T18 are physically separated by two non-interacting proteins, there is no production of

cAMP, no activation of the transcription of *lacZ* and therefore no β-galactosidase production (Karimova et al, 1998).

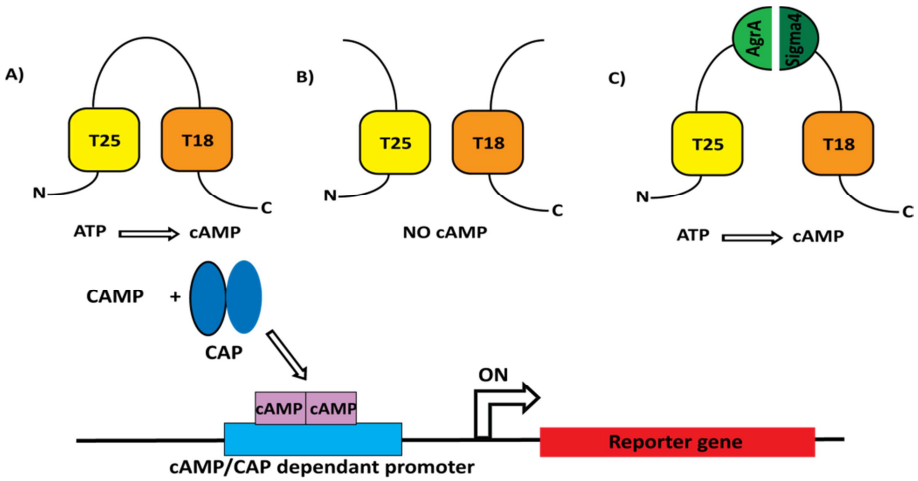


Figure 110: Schematic of the principle of the BACTH system.

A) The T25 and the T18 complementary fragments of the adenylate cyclase from *Bordetella pertussis* are physically linked which results in a functional adenylate cyclase, cAMP production and activation of the *lacZ* reporter gene. B) The T25 and the T18 complementary fragments are not physically linked, the adenylate cyclase is non-functional and there is not cAMP produced hence no transcription activation. C) AgrA and σ4 are fused to the T25 and the T18 complementary fragments of the adenylate cyclase and if they interact, the physical link between T25 and T18 will be restored, the adenylate cyclase will be functional, cAMP will be produced and the *lacZ* reporter gene activated. Adapted from (Karimova et al, 1998).

In this case, we tested the ability of several subunits of the RNAp (α-CTD, α-NTD, β-flap, β-flap-tip and σ4) to interact with the AgrA LytTR domain and the AgrA CheY-like domain. A weak interaction was detected between σ4 and the AgrA LytTR domain but not the AgrA CheY-like domain. None of the other subunits seemed to interact.

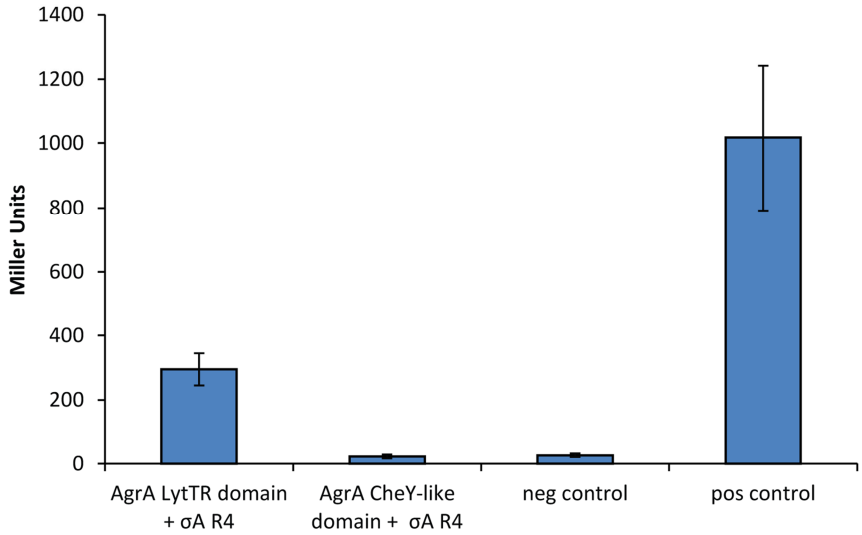


Figure 111: BACTH interaction between the AgrA-LytTR domain and σ4.

The interaction between the AgrA-LytTR domain and σ4 is over 10-fold above the negative control. The CheY-like domain does not interact with σ4 as it displays the same level of activity as the negative control. The interaction is measured in Miller Units.

The interaction between the AgrA LytTR domain and σ4 was weak and difficult to detect. This might be because this system does not involve DNA and AgrA is not activated. The interaction between AgrA and the RNAp is probably mediated by the presence of the promoter DNA and requires AgrA to be active. Thus, the fact that we did not detect an interaction between AgrA and the other subunits we tested, mainly α-CTD, does not mean that they do not interact. We therefore conclude that the BACTH is not a suitable system to investigate the putative interaction between the AgrA LytTR domain and the RNAp. Nevertheless, this system identified that the AgrA LytTR domain does interact with another RNAp subunit than the α-CTD.

To summarise, our putative model of activation by AgrA involves the bending of the promoter DNA in order to reduce the spacer length between the -10 and -35 promoter, direct contact with the α-CTD of the RNAp via residue Y229 in the AgrA LytTR domain, interaction of unknown residue(s) probably in the N-terminal part of the AgrA LytTR domain with σ region 4 and putatively a third interaction of unknown residue(s) probably in the N-terminal part of the AgrA LytTR domain could take place with the α-NTD subunit. Examples of TFs that bind directly to the RNAp and bend the DNA include: CAP, MerR and Fis.

The CAP TF from *E. coli* binds to its target promoters (>100) in presence of the allosteric effector cyclic AMP and enhances the ability of the RNAp to initiate transcription from these promoters. The binding of CAP to the promoter DNA induces a ~ 80° bend in the DNA. CAP is also involved in direct interaction with the RNAp. There are Class I promoters where CAP interacts with the α-CTD subunit and this interaction facilitates the binding of the RNAp to the promoter DNA and thereby stimulates transcription. At Class II promoters, CAP interacts with the α-CTD subunit and recruits the RNAp to the promoter. The interaction of CAP with the α-NTD subunit and σ4 facilitate the isomerisation from closed complex (RPc) to open complex (RPo) thereby activating transcription (Belyaeva et al, 1996; Lawson et al, 2004).

MerR is a metal ion responsive TF in *E. coli* and activates transcription as a result of binding to mercuric (Hg) ion. MerR binds to its target promoters and recruits the RNAp by interaction with σ4. However a ~33° bend in the promoter DNA is induced upon MerR binding which results in the -10 and -35 being misaligned in regards to the σ factor. The binding of Hg to MerR results in an allosteric underwinding of the DNA which realigns the -10 and -35 promoter element so that the σ factor can contact them (Ansari et al, 1992; Hobman, 2007).

FIS (Factor for Inversion Stimulation) is a global regulator of gene expression in *E. coli*. Fis induces a 40-90° bend in the promoter DNA upon binding. The DNA bending induced by FIS is not sufficient for transcription activation on its own. FIS activates transcription via an interaction with the α-CTD subunit (Bokal et al, 1997; Finkel & Johnson, 1993; Gille et al, 1991; Thompson & Landy, 1988).

The principal limitation of this study is the fact that we might have missed other important residues for transcription activation. At first we excluded residues involved in salt bridges interactions and DNA-binding, then in the screen we excluded the mutants that were structurally impaired, had reduced stability or a reduced ability to bind to DNA. We cannot exclude the fact that some of the residues we excluded are involved in transcription activation as one residue can be involved in different functions. For example, we exclude residues R218 and R198 because they appeared to be structurally impaired but these residues are involved in contacting the DNA. However, the stringency of the criteria we used was necessary in order to be certain that the drop in activity we were observing was due to a defect in transcription activation.

In this study, we have successfully identified a key residue for transcription activation in the AgrA LytTR domain, Y229, which is not deficient for DNA bending and thus challenges the current model that AgrA activates transcription by its DNA bending activity. We have also identified two new residues in the AgrA LytTR domain that are involved in contacting the DNA, E181 and H200 and discovered that H200 is essential for the DNA-binding activity of AgrA.

6.3. Future directions

In order to try to illustrate the interaction between the RNAp, we will use a co-immuno-precipitation approach to identify proteins interacting with AgrA. This approach implicates the use of protein A beads, and the polyclonal antibody against AgrA. Protein A specifically binds to the Fc chain of antibodies. First the protein A beads will be coated with polyclonal AgrA antibody then incubated with *S. aureus* whole cell lysate in order for the antibody to bind to AgrA. A wash step will remove un-specifically bound and excess protein. Then the protein A beads coated with the AgrA/anti-AgrA antibody complex will be incubated in *S. aureus* whole cell lysate again in order for putative proteins interacting with AgrA to bind to AgrA. The complexes will be released from the beads after a washing step by incubation with Laemmly buffer. The samples will then be analysed by SDS-PAGE electrophoresis to determine the nature of the proteins bound to AgrA (Firestone & Winguth, 1990).

Alternatively, b-benzoyl-p-phenylalanine (Bpa) crosslinking could be used to try to illustrate binding of the RNAp to AgrA. Bpa is a photoreactive artificial aa that can be introduced in a protein at a specific site during translation. The chosen insertion sites need to be replaced with TAG stop codons. The incorporation of Bpa takes place in a strain carrying a Bpa-specific suppressor tRNA and aminoacyl-tRNA synthetase (available in the lab) that will incorporate the Bpa aa in place of the TAG stop codon. Then the protein containing the artificial Bpa can be purified and used in a crosslinking experiment. Bpa can crosslink to aliphatic side chains of aa when exposed to long-wave UV light (Dorman & Prestwich, 1994; Forne et al, 2012; Liu & Schultz, 2010). Therefore, we could incorporate the Bpa aa close to residue Y229A in the AgrA LytTR domain and perform a crosslinking experiment with AgrA in presence of the RNAp and see if they interact.

A third method could be used to confirm that AgrA activation involves contact with the α-CTD subunit of the RNAp using a reconstituted RNAp with a truncated α subunit lacking the CTD (aa 235-329) (Igarashi et al, 1991; Igarashi & Ishihama, 1991). This mutant RNAp could be used in an *in vitro* transcription assay, as described in section 5.5, to test whether the ability of WT AgrA mutant to activate transcription at the P3 promoter is dependent on the presence of the α-CTD. The $AgrA_{Y229A}$ could also be tested alongside the WT AgrA and if the Y229A is truly involved in interaction with the α-CTD that activates transcription, they should display a similar level of transcription activation.

Fourthly, it would be interesting to investigate the putative role of the δ subunit in transcription activation by AgrA at the P3 promoter. An *rpoE* (δ) transposon mutant is available in the Nebraska transposon library (NARSA) and was used to determine that δ is involved in promoter recognition in *S. aureus* (Weiss et al, 2014). We could acquire this strain and transduce the mutation to our $SH1000^{-}_{agr\ IR\ P3\text{-}GFP}$ strain. Then we would be able to test the ability of pSN-P2-agrA to activate transcription in the absence of the δ subunit and the whole mutant library could be screened in this background.

Finally, we will purify the Y229F mutant that displayed a higher ability to activate transcription than the WT AgrA and test it in an *in vitro* transcription assay to confirm the result observed *in vivo*.

To investigate whether the $AgrA_{L171A}$ mutant is truly phosphorylation deficient, a phosphorylation assay with the HK domain of AgrC, AgrC-HK will be performed, as in section 2.4.8 and 5.6.3. The ability of $AgrA_{L171A}$ to be activated by AgrC-HK will be compared to that of the WT AgrA.

Alternatively, the effect of the L171A could be test in an AgrA constitutively active mutant using the SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ strain. Several constitutively active CheY mutants have been described (Da Re et al, 2002) but one, the double mutant CheY D13K Y106W has been shown to adopt a conformation similar to that of the activated CheY (Dyer et al, 2004). The corresponding AgrA double mutant (D9K F107W) could be constructed by site directed mutagenesis (SDM) as described in section 4.4. To check that CheY mutants are truly constitutively active, they are usually tested in presence of the D57A mutation (the phosphorylation site in CheY corresponding to D59 in AgrA) that cannot be phosphorylated *in vitro*. Therefore we could use this method to test whether the D9K F107W mutations result in a constitutively active AgrA by adding the D59A mutation by SDM and testing the ability of the mutant to activate transcription at the P3 promoter in the SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ strain. Once we have confirmed whether the AgrA$_{D9K\ F107W}$ is constitutively active, we will be able to test the effect of adding the L171A mutation to the AgrA$_{D9K\ F107W}$ mutant using the SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ strain.

Alternatively, a simpler method would be to test whether the AgrA$_{D9K\ F107W}$ mutant is able to activate the *agr* response in serum to test whether it is constitutively active AgrA mutant. The *agr* response is blocked in serum because the autoinducing peptide, AIP, cannot reach AgrC. This blockage can be reversed using a constitutively active AgrC and therefore should also be reverted using a constitutively active AgrA (James et al, 2013). Therefore we could test the ability of the AgrA$_{D9K\ F107W}$ mutant to activate transcription at the P3 promoter in presence of serum using the SH1000$^{-}_{agr\ IR\ P3\text{-}GFP}$ strain.

Finally, we could assess whether the AgrA$_{L171I}$ mutant is constitutively active using the exact same methods as for the AgrA$_{D13KY106W}$ mutant.

APPENDICES

7. Appendix I

7.1. Vector maps

7.1.1. pGEM-T Easy – used for cloning into pTYB2 and pCN34*itet*

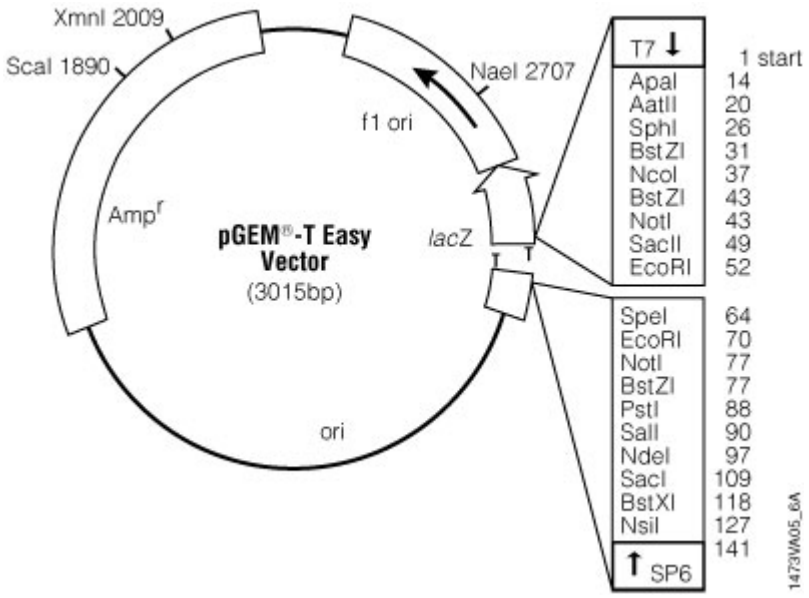


Figure 112: Vector map of the pGEM-T Easy vector from Promega.

7.1.2. pCN34 – used to clone P2-RBS-AgrA-TT and P1-AgrA-TT

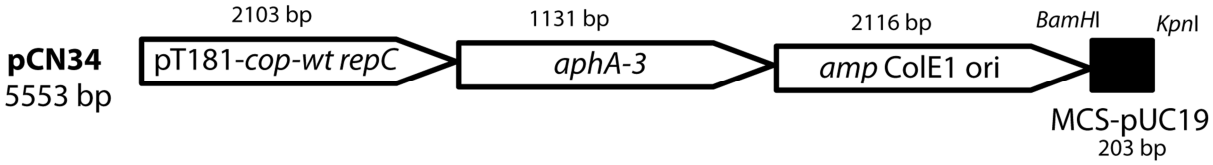


Figure 113: Vector map of the pCN34 vector. Adapted from (Charpentier et al, 2004).

7.1.3. pCN44 – used as PCR template for the TT

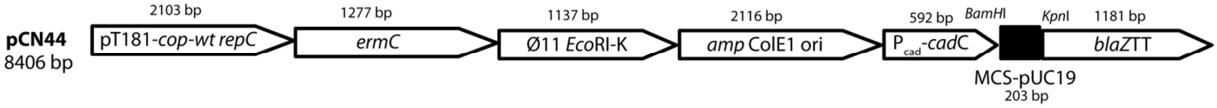


Figure 114: Vector map of the pCN44 vector. Adapted from (Charpentier et al, 2004).

7.1.4. pCN34*itet* – used to clone RBS-AgrA-TT

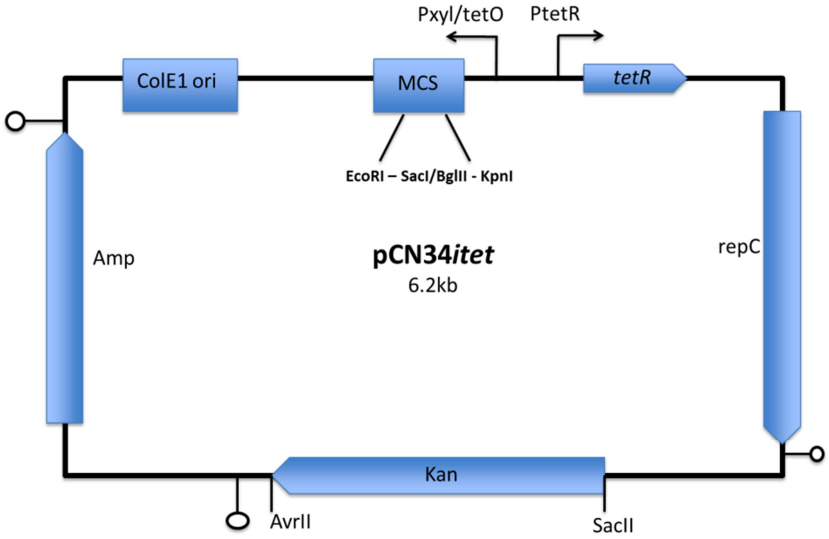


Figure 115: Vector map of the pCN34itet. (Corrigan et al, 2011).

7.1.5. pTYB2 – used to clone AgrA

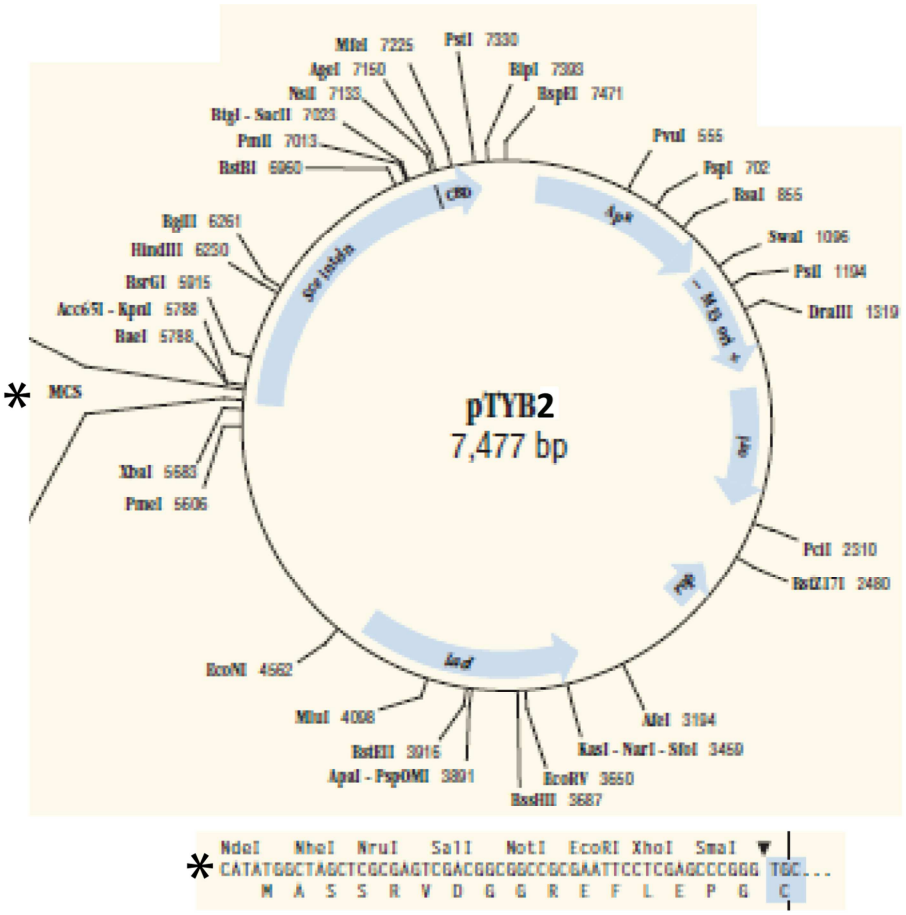


Figure 116: Vector map pTYB2 from New England Biolabs.

7.1.6. pET28b+ -used to clone AgrC-HK

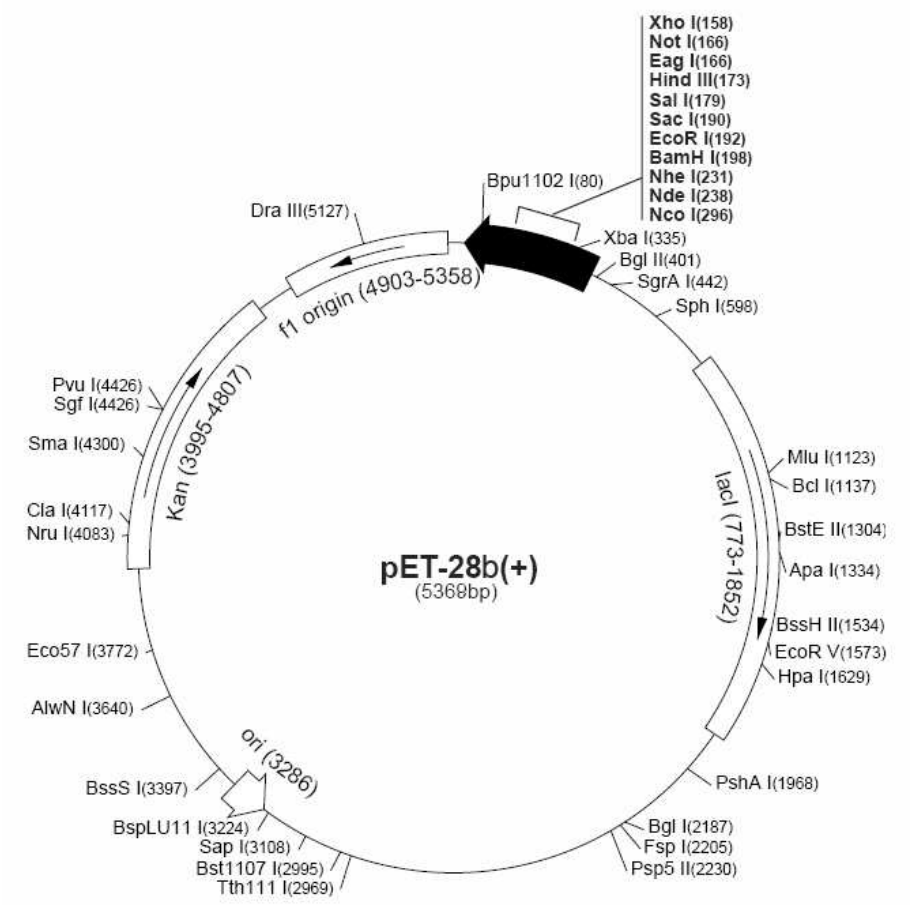


Figure 117: Vector map for pET28b+ from Novagen.

7.2. Protein ladders

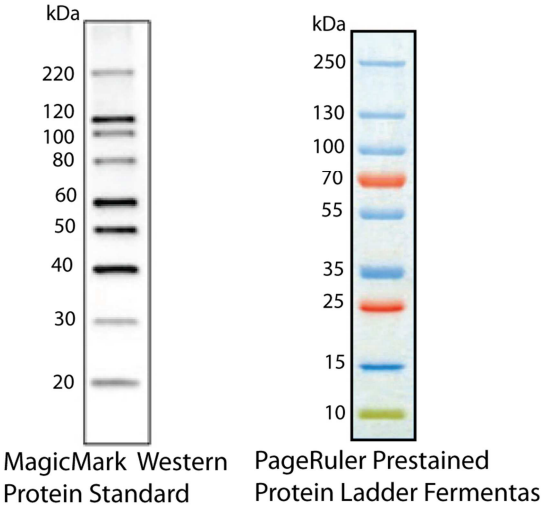


Figure 118: Protein Ladders.

PageRuler was used for SDS-PAGE gels and MagicMark for Western blot.

8. Appendix II

8.1. Oligonucleotides and vectors constructed

8.1.1. Oligonucleotides for cloning and *in vitro* analysis

Primer	Use	Sequence (5'-3')	Name vector created
P2-RBS F (for P2-AgrA)	cloning P2-RBS-AgrA-TT into pCN34; fuse to AgrA. *BamH*I site.	CGCGGATCCCCGGAATTCTTTTTCTTAACTAGTCG	pSN-P2-*agrA*
P2-RBS R (for P2-AgrA)	cloning P2-RBS-AgrA-TT into pCN34; fuse to AgrA. Contains RBS.	CGCAAATGAAAATTTTCATTTTTTTTCCTCCTTAGGATCACC	pSN-P2-*agrA*
P2-AgrA F (for P2-AgrA)	cloning P2-RBS-AgrA-TT into pCN34; fuse to P2-RBS	ATGAAAATTTTCATTTGCGAAGACGATCCAAAAC	pSN-P2-*agrA*
P2-AgrA R (for P2-AgrA)	cloning P2-RBS-AgrA-TT into pCN34; fuse to P2-RBS	GCATTTAGAATATTATATTTTTTTAACGTTTCTCACCG	pSN-P2-*agrA*
P2-TT F (for P2-AgrA)	cloning P2-RBS-AgrA-TT into pCN34; fuse to P2-RBS-AgrA.	CGTTAAAAAAATATAATATTCTAAATGCATAATAAATACTG	pSN-P2-*agrA*
P2-TT R (for P2-AgrA)	cloning P2-RBS-AgrA-TT into pCN34.fuse to P2-RBS-AgrA. *Kpn*I site.	CGGGGTACCTGTCACTTTGCTTGATATATGAGAATTATTTAACC	pSN-P2-*agrA*
P1-AgrA F (for P1-AgrA)	cloning P1-AgrA-TT into pCN34; fuse to TT. *BamH*I site.	CGCGGATCCGTACGTGAAATTAAAGGCTTAATTACTGC	pSN:*agrC832-1261-agrA*
P1-AgrA R (for P1-AgrA)	cloning P1-AgrA-TT into pCN34; fuse to TT.	GCATTTAGAATATTATATTTTTTTAACGTTTCTCACCG	pSN:*agrC832-1261-agrA*
P1-TT R (for P1-AgrA)	cloning P1-AgrA-TT into pCN34; fuse P1-AgrA.	CGTTAAAAAAATATAATATTCTAAATGCATAATAAATACTG	pSN:*agrC832-1261-agrA*
P1-TT F (for P1-AgrA)	cloning P1-AgrA-TT into pCN34; fuse P1-AgrA. KpnI site.	CGGGGTACCTGTCACTTTGCTTGATATATGAGAATTATTTAACC	pSN:*agrC832-1261-agrA*
AgrA pCN34*itet* F	cloning RBS-AgrA-TT into pCN34*itet*. Contains RBS. *KpN*I site.	CGGGGTACCGATCCTAAGGAGGAAA	pSN-*itet-agrA*
AgrA pCN34*itet* R	cloning RBS-AgrA-TT into pCN34*itet*. *EcoR*I site.	CCGGAATTCTGTCACTTTGCTTGATATATGAGAATTATTTAACC	pSN-*itet-agrA*
AgrA pTYB2 F	cloning AgrA into pTYB2. *Nde*I site.	GAAGCATATGAAAATTTTCATTTGCGAAGACG	pSN-*agrA*
AgrA pTYB2 R	cloning AgrA into pTYB2. *Sma*I site.	GAAGCCCGGGTATTTTTTTAACGTTTCTCACCGATGC	pSN-*agrA*
AgrC-HK pET28 F	cloning AgrC-HK into pET28b+. *Nde*I site.	GGAATTCCATATGATATTAATGTTTACAATACCAGC	pSN-*agrC-HK*
AgrC-HK pET28 R	cloning AgrC-HK into pET28b+. *Hind*III site.	GGGAAGCTTGTTGTTAATAATTTCAACTTTTTGAATAAAG	pSN-*agrC-HK*
P2-IR-P3 F	EMSA probe	ACCACTCTCCTCACTGTCATTATACG	
P2-IR-P3 R	EMSA probe	ACATCTCTGTGATCTAGTTATATTAAAACATGC	
P3-IR non-template	Short RNA synthesis probe AAAATTTACAGTTAAGAATAAAAAAACGACTAGTTAAGAAAAATTGGAAAATAAATGTTTTCATGTTTTAATATAACTAGATCACAGAGATGTG		
P3-IR template	Short RNA synthesis probe CACATCTCTGTGATCTAGTTATATTAAAACATGAAAACATTTATTTTCCAATTTTTCTTAACTAGTCGTTTTTTTATTCTTAACTGTAAATTTT		

Table 13: List of oligonucleotides used for cloning, EMSA and short RNA synthesis assay in this study.

Primers AgrA pCN34*itet* F and R were used to subclone the AgrA mutants from pSN-P2-*agrA* to pSN-*itet-agrA*. Primers AgrA pTYB2 F and R were used to subclone the AgrA mutants from pSN-P2-*agrA* to pSN *agrA.*

8.1.2. Vectors constructed

Name	Insert	Use	Reference
pSN-P2-*agrA*	1480 bp *BamH*I-*Kpn*I fragment carrying P2-RBS-AgrA-TT in pCN34	complementation SH1000^{-}, SDM template	this study
pSN:*agrC832-1261-agrA*	1227 bp *BamH*I-*Kpn*I fragment carrying P1-AgrA-TT in pCN34	complementation SH1000^{-}	this study
pSN-*itet-agrA*	1059 bp *Ecor*I-*Nde*I fragment carrying RBS-AgrA-TT in pCN34*itet*	complementation SH1000^{-}, Western blot	this study
pSN-*agrA*	714 bp *Nde*I-*Sma*I fragment carrying AgrA in pTYB2	IMPACT purification	this study
pSN-*agrC-HK*	698 bp *Nde*I-*Hind*III fragment carrying AgrC-HK in pEt28b+	AgrC purification	this study

Table 14: List of vectors constructed in this study.

8.1.3. SDM pSN-P2-*agrA* primers

Primer	Sequence (5'-3')	Primer	Sequence (5'-3')
D137A F	CGCTTACAATTGTTGTCTAAAGCTAATAGCGTTGAAACG	D158A F	GTGTATGTTCAATATGATGCTATTATGTTTTTTGAATC
D137A R	CGTTTCAACGCTATTAGCTTTAGACAACAATTGTAAGCG	D158A R	GATTCAAAAAACATAATAGCATCATATTGAACATACAC
N138A F	CGCTTACAATTGTTGTCTAAAGATGCTAGCGTTGAAACGATTG	I159A F	GTATGTTCAATATGATGATGCTATGTTTTTTGAATCATC
N138A R	CAATCGTTTCAACGCTAGCATCTTTAGACAACAATTGTAAGCG	I159A R	GATGATTCAAAAAACATAGCATCATCATATTGAACATAC
S139A F	GTTGTCTAAAGATAATGCCGTTGAAACGATTGAATTAAAACG	M160A F	GTTCAATATGATGATATTGCGTTTTTTGAATCATCAAC
S139A R	CGTTTTAATTCAATCGTTTCAACGGCATTATCTTTAGACAAC	M160A R	GTTGATGATTCAAAAAACGCAATATCATCATATTGAAC
V140A F	GTTGTCTAAAGATAATAGCGCTGAAACGATTGAATTAAAACG	F161A F	CAATATGATGATATTATGGCTTTTGAATCATCAACAAAATC
V140A R	CGTTTTAATTCAATCGTTTCAGCGCTATTATCTTTAGACAAC	F161A R	GATTTTGTTGATGATTCAAAAGCCATAATATCATCATATTG
T142A F	CTAAAGATAATAGCGTTGAAGCGATTGAATTAAAACGTGGC	F162A R	GAGATTTTGTTGATGATTCAGCAAACATAATATCATC
T142A R	GCCACGTTTTAATTCAATCGCTTCAACGCTATTATCTTTAG	F162A F	GATGATATTATGTTTGCTGAATCATCAACAAAATCTC
T142G F	CTAAAGATAATAGCGTTGAAGGGATTGAATTAAAACGTGGC	S165A F	GATATTATGTTTTTTGAATCAGCAACAAAATCTCACAGAC
T142G R	GCCACGTTTTAATTCAATCCCTTCAACGCTATTATCTTTAG	S165A R	GTCTGTGAGATTTTGTTGCTGATTCAAAAAACATAATATC
T142S F	CTAAAGATAATAGCGTTGAATCGATTGAATTAAAACGTGGC	T166A R	TGAGTCTGTGAGATTTTGCTGATGATTCAAAAAAC
T142S R	GCCACGTTTTAATTCAATCGATTCAACGCTATTATCTTTAG	T166A F	GTTTTTTGAATCATCAGCAAAATCTCACAGACTCA
I143A F	GATAATAGCGTTGAAACGGCTGAATTAAAACGTGGC	K167A F	GTTTTTTGAATCATCAACAGCATCTCACAGACTCATTGCC
I143A R	GCCACGTTTTAATTCAGCCGTTTCAACGCTATTATC	K167A R	GGCAATGAGTCTGTGAGATGCTGTTGATGATTCAAAAAAC
I143D F	GATAATAGCGTTGAAACGGATGAATTAAAACGTGGC	L171A F	CAACAAAATCTCACAGAGCCATTGCCCATTTAGATAACCG
I143D R	GCCACGTTTTAATTCATCCGTTTCAACGCTATTATC	L171A R	CGGTTATCTAAATGGGCAATGGCTCTGTGAGATTTTGTTG
I143K F	GATAATAGCGTTGAAACGAAAGAATTAAAACGTGGC	L171G F	CAACAAAATCTCACAGAGGCATTGCCCATTTAGATAACCG
I143K R	GCCACGTTTTAATTCTTTCGTTTCAACGCTATTATC	L171G R	CGGTTATCTAAATGGGCAATGCCTCTGTGAGATTTTGTTG
I143V F	GATAATAGCGTTGAAACGGTTGAATTAAAACGTGGC	L171I F	CAACAAAATCTCACAGAATCATTGCCCATTTAGATAACCG
I143V R	GCCACGTTTTAATTCAACCGTTTCAACGCTATTATC	L171I R	CGGTTATCTAAATGGGCAATGATTCTGTGAGATTTTGTTG
E144A F	GCGTTGAAACGATTGCATTAAAACGTGGCAGTAATTC	L171R F	CAACAAAATCTCACAGACGCATTGCCCATTTAGATAACCG
E144A R	GAATTACTGCCACGTTTTAATGCAATCGTTTCAACGC	L171R R	CGGTTATCTAAATGGGCAATGCGTCTGTGAGATTTTGTTG
L145A F	GCGTTGAAACGATTGAAGCAAAACGTGGCAGTAATTC	I172A F	CAAAATCTCACAGACTCGCTGCCCATTTAGATAACCG
L145A R	GAATTACTGCCACGTTTTGCTTCAATCGTTTCAACGC	I172A R	CGGTTATCTAAATGGGCAGCGAGTCTGTGAGATTTTG
K146A R	ACTGAATTACTGCCACGTGCTAATTCAATCGTTTCAACGC	I172D F	CAAAATCTCACAGACTCGATGCCCATTTAGATAACCG
K146A F	GCGTTGAAACGATTGAATTAGCACGTGGCAGTAATTCAGT	I172D R	CGGTTATCTAAATGGGCATCGAGTCTGTGAGATTTTG
R147A F	CGTTGAAACGATTGAATTAAAAGCTGGCAGTAATTCAGTGTATGTTC	I172K F	CAAAATCTCACAGACTCAAAGCCCATTTAGATAACCG
R147A R	GAACATACACTGAATTACTGCCAGCTTTTAATTCAATCGTTTCAACG	I172K R	CGGTTATCTAAATGGGCTTTGAGTCTGTGAGATTTTG
G148A F	CGATTGAATTAAAACGTGCCAGTAATTCAGTGTATG	I172V F	CAAAATCTCACAGACTCGTTGCCCATTTAGATAACCG
G148A R	CATACACTGAATTACTGGCACGTTTTAATTCAATCG	I172V R	CGGTTATCTAAATGGGCAACGAGTCTGTGAGATTTTG
S149A R	CATATTGAACATACACTGAATTAGCGCCACGTTTTAATTCAATCG	H174L F	CTCACAGACTCATTGCCCTTTTAGATAACCGTCAAATTG
S149A F	CGATTGAATTAAAACGTGGCGCTAATTCAGTGTATGTTCAATATG	H174L R	CAATTTGACGGTTATCTAAAAGGGCAATGAGTCTGTGAG
N150A F	CGATTGAATTAAAACGTGGCAGTGCTTCAGTGTATGTTCAATATGATG	L175A F	CACAGACTCATTGCCCATGCAGATAACCGTCAAATTG
N150A R	CATCATATTGAACATACACTGAAGCACTGCCACGTTTTAATTCAATCG	L175A F	CTCACAGACTCATTGCCCATGCAGATAACCGTCAAATTG
S151A F	CGATTGAATTAAAACGTGGCAGTAATGCAGTGTATGTTCAATATG	L175A R	CAATTTGACGGTTATCTGCATGGGCAATGAGTCTGTG
S151A R	CATATTGAACATACACTGCATTACTGCCACGTTTTAATTCAATCG	L175A R	CAATTTGACGGTTATCTGCATGGGCAATGAGTCTGTGAG
V152A F	CGTGGCAGTAATTCAGCGTATGTTCAATATGATG	N177A F	CTCATTGCCCATTTAGATGCCCGTCAAATTGAATTTTATGG
V152A R	CATCATATTGAACATACGCTGAATTACTGCCACG	N177A R	CCATAAAATTCAATTTGACGGGCATCTAAATGGGCAATGAG
Y153A F	CGTGGCAGTAATTCAGTGGCTGTTCAATATGATGATATTATG	R178A F	GCCCATTTAGATAACGCTCAAATTGAATTTTATGG
Y153A R	CATAATATCATCATATTGAACAGCCACTGAATTACTGCCACG	R178A R	CCATAAAATTCAATTTGAGCGTTATCTAAATGGGC
V154A F	GGCAGTAATTCAGTGTATGCTCAATATGATGATATTATG	Q179A F	GCCCATTTAGATAACCGTGCAATTGAATTTTATGG
V154A R	CATAATATCATCATATTGAGCATACACTGAATTACTGCC	Q179A R	CCATAAAATTCAATTGCACGGTTATCTAAATGGGC
Q155A F	GGCAGTAATTCAGTGTATGTTGCATATGATGATATTATG	I180A F	GCCCATTTAGATAACCGTCAAGCTGAATTTTATGGTAATTTAAAAG
Q155A R	CATAATATCATCATATGCAACATACACTGAATTACTGCC	I180A R	CTTTTAAATTACCATAAAATTCAGCTTGACGGTTATCTAAATGGGC
Y156A F	GGCAGTAATTCAGTGTATGTTCAAGCTGATGATATTATGTTTTTTG	E181A F	GCCCATTTAGATAACCGTCAAATTGCATTTTATGGTAATTTAAAAG
Y156A R	CAAAAAACATAATATCATCAGCTTGAACATACACTGAATTACTGCC	E181A R	CTTTTAAATTACCATAAAATGCAATTTGACGGTTATCTAAATGGGC

Table 15: SDM primer Table 1.

Primers for pSN-P2-*agrA* SDM, D137A-E181A.

Primer	Sequence (5'-3')	Primer	Sequence (5'-3')
E181D F	GCCCATTTAGATAACCGTCAAATTGATTTTTATGGTAATTTAAAAG	F203A F	CGTTTCTTTAGATGTCATAATAGCGCTGTCGTCAATCGCC
E181D R	CTTTTAAATTACCATAAAAATCAATTTGACGGTTATCTAAATGGGC	F203A R	GGCGATTGACGACAGCGCTATTATGACATCTAAAGAAACG
E181R F	GCCCATTTAGATAACCGTCAAATTCGATTTTATGGTAATTTAAAAG	V204A F	GATGTCATAATAGCTTTGCCGTCAATCGCCATAATATTG
E181R R	CTTTTAAATTACCATAAAATCGAATTTGACGGTTATCTAAATGGGC	V204A R	CAATATTATGGCGATTGACGGCAAAGCTATTATGACATC
F182A F	CATTTAGATAACCGTCAAATTGAAGCTTATGGTAATTTAAAAGAACTG	V205A F	GATGTCATAATAGCTTTGTCGCCAATCGCCATAATATTG
F182A R	CAGTTCTTTTAAATTACCATAAGCTTCAATTTGACGGTTATCTAAATG	V205A R	CAATATTATGGCGATTGGCGACAAAGCTATTATGACATC
F182L F	CCGTCAAATTGAACTGTATGGTAATTTAAAAG	N206A R	CTATAGATTCAATATTATGGCGAGCGACGACAAAGCTATTATGACATC
F182L R	CTTTTAAATTACCATACAGTTCAATTTGACGG	N206A F	GATGTCATAATAGCTTTGTCGTCGCTCGCCATAATATTGAATCTATAG
F182R R	CTTTTAAATTACCATAACGTTCAATTTGACGG	R207A F	GCTTTGTCGTCAATGCCCATAATATTGAATCTATAGATTCG
F182R F	CCGTCAAATTGAACGTTATGGTAATTTAAAAG	R207A R	CGAATCTATAGATTCAATATTATGGGCATTGACGACAAAGC
F182Y R	CTTTTAAATTACCATAGTATTCAATTTGACGG	H208A F	GCTTTGTCGTCAATCGCGCTAATATTGAATCTATAGATTCG
F182Y F	CCGTCAAATTGAATACTATGGTAATTTAAAAG	H208A R	CGAATCTATAGATTCAATATTAGCGCGATTGACGACAAAGC
G184A F	CGTCAAATTGAATTTTATGCTAATTTAAAAGAACTGAG	N209A F	GCTTTGTCGTCAATCGCCATGCTATTGAATCTATAGATTCG
G184A R	CTCAGTTCTTTTAAATTAGCATAAAATTCAATTTGACG	N209A R	CGAATCTATAGATTCAATAGCATGGCGATTGACGACAAAGC
E188A F	GAATTTTATGGTAATTTAAAAGCACTGAGTCAATTAGATGATCG	I210A F	GCTTTGTCGTCAATCGCCATAATGCTGAATCTATAGATTCG
E188A R	CGATCATCTAATTGACTCAGTGCTTTTAAATTACCATAAAATTC	I210A R	CGAATCTATAGATTCAGCATTATGGCGATTGACGACAAAGC
L189A F	GGTAATTTAAAAGAAGCGAGTCAATTAGATGATCG	E211A F	CGTCAATCGCCATAATATTGCATCTATAGATTCGAAAGAGCG
L189A R	CGATCATCTAATTGACTCGCTTCTTTTAAATTACC	E211A R	CGCTCTTTCGAATCTATAGATGCAATATTATGGCGATTGACG
S190A F	GGTAATTTAAAAGAACTGGCTCAATTAGATGATCG	S212A F	CGTCAATCGCCATAATATTGAAGCTATAGATTCGAAAGAGCG
S190A R	CGATCATCTAATTGAGCCAGTTCTTTTAAATTACC	S212A R	CGCTCTTTCGAATCTATAGCTTCAATATTATGGCGATTGACG
Q191A F	GGTAATTTAAAAGAACTGAGTGCATTAGATGATCGTTTC	I213A F	CGCCATAATATTGAATCTGCAGATTCGAAAGAGCG
Q191A R	GAAACGATCATCTAATGCACTCAGTTCTTTTAAATTACC	I213A R	CGCTCTTTCGAATCTGCAGATTCAATATTATGGCG
L192A R	CTAAAGAAACGATCATCTGCTTGACTCAGTTC	D214A F	GCCATAATATTGAATCTATAGCTTCGAAAGAGCGAATTGTC
L192A F	GAACTGAGTCAAGCAGATGATCGTTTCTTTAG	D214A R	GACAATTCGCTCTTTCGAAGCTATAGATTCAATATTATGGC
L192G F	GAACTGAGTCAAGGAGATGATCGTTTCTTTAG	S215A F	CCATAATATTGAATCTATAGATGCGAAAGAGCGAATTGTC
L192G R	CTAAAGAAACGATCATCTCCTTGACTCAGTTC	S215A R	GACAATTCGCTCTTTCGCATCTATAGATTCAATATTATGG
L192I F	GAACTGAGTCAAATAGATGATCGTTTCTTTAG	K216A F	GAATCTATAGATTCGGCAGAGCGAATTGTCTATTTTAAAAATAAAG
L192I R	CTAAAGAAACGATCATCTATTTGACTCAGTTC	K216A R	CTTTATTTTTAAAATAGACAATTCGCTCTGCCGAATCTATAGATTC
L192R F	GAACTGAGTCAACGAGATGATCGTTTCTTTAG	E217A F	GAATCTATAGATTCGAAAGCGCGAATTGTCTATTTTAAAAATAAAG
L192R R	CTAAAGAAACGATCATCTCGTTGACTCAGTTC	E217A R	CTTTATTTTTAAAATAGACAATTCGCGCTTTCGAATCTATAGATTC
D194A F	GAACTGAGTCAATTAGATGCTCGTTTCTTTAGATG	R218A F	GAATCTATAGATTCGAAAGAGGCAATTGTCTATTTTAAAAATAAAG
D194A R	CATCTAAAGAAACGAGCATCTAATTGACTCAGTTC	R218A R	CTTTATTTTTAAAATAGACAATTGCCTCTTTCGAATCTATAGATTC
F196A F	GAGTCAATTAGATGATCGTGCCTTTAGATGTCATAATAGC	I219A F	CTATAGATTCGAAAGAGCGAGCTGTCTATTTTAAAAATAAAG
F196A R	GCTATTATGACATCTAAAGGCACGATCATCTAATTGACTC	I219A R	CTTTATTTTTAAAATAGACAGCTCGCTCTTTCGAATCTATAG
F197A F	GTCAATTAGATGATCGTTTCGCTAGATGTCATAATAGC	V220A F	GATTCGAAAGAGCGAATTGCCTATTTTAAAAATAAAGAAC
F197A R	GCTATTATGACATCTAGCGAAACGATCATCTAATTGAC	V220A R	GTTCTTTATTTTTAAAATAGGCAATTCGCTCTTTCGAATC
R198A F	GATGATCGTTTCTTTGCATGTCATAATAGCTTTGTCG	Y221A F	CGAAAGAGCGAATTGTCGCTTTTAAAAATAAAGAACACTGC
R198A R	CGACAAAGCTATTATGACATGCAAAGAAACGATCATC	Y221A R	GCAGTGTTCTTTATTTTTAAAAGCGACAATTCGCTCTTTCG
C199S F	GATCGTTTCTTTAGATCTCATAATAGCTTTGTCGTCAATCG	F222A F	CGAAAGAGCGAATTGTCTATGCTAAAAATAAAGAACACTGC
C199S R	CGATTGACGACAAAGCTATTATGAGATCTAAAGAAACGATC	F222A R	GCAGTGTTCTTTATTTTTAGCATAGACAATTCGCTCTTTCG
H200A F	CGTTTCTTTAGATGTGCTAATAGCTTTGTCGTCAATCG	K223A R	GCAGTGTTCTTTATTTGCAAAATAGACAATTC
H200A R	CGATTGACGACAAAGCTATTAGCACATCTAAAGAAACG	K223A F	GAATTGTCTATTTTGCAAATAAAGAACACTGC
H200D F	CGTTTCTTTAGATGTGATAATAGCTTTGTCGTCAATCG	N224A F	CGAATTGTCTATTTTAAAGCTAAAGAAACACTGCTATGC
H200D R	CGATTGACGACAAAGCTATTATCACATCTAAAGAAACG	N224A R	GCATAGCAGTGTTCTTTAGCTTTAAAATAGACAATTCG
H200R F	CGTTTCTTTAGATGTCGTAATAGCTTTGTCGTCAATCG	K225A F	GTCTATTTTAAAAATGCAGAACACTGCTATGCATCG
H200R R	CGATTGACGACAAAGCTATTACGACATCTAAAGAAACG	K225A R	CGATGCATAGCAGTGTTCTGCATTTTTAAAATAGAC
S202A F	CGTTTCTTTAGATGTCATAATGCCTTTGTCGTCAATCGCC	H227A F	CTATTTTAAAAATAAAGAAGCCTGCTATGCATCGGTG
S202A R	GGCGATTGACGACAAAGGCATTATGACATCTAAAGAAACG	H227A R	CACCGATGCATAGCAGGCTTCTTTATTTTTAAAATAG

Table 16: SDM primer Table 2.

Primers for pSN-P2-*agrA* SDM, E181D-H227A.

Primer	Sequence (5'-3')
C228S F	CTATTTTAAAAATAAAGAACACTCCTATGCATCGGTGAGAAAC
C228S R	GTTTCTCACCGATGCATAGGAGTGTTCTTTATTTTTAAAATAG
Y229A F	CTATTTTAAAAATAAAGAACACTGCGCTGCATCGGTGAGAAACG
Y229A R	CGTTTCTCACCGATGCAGCGCAGTGTTCTTTATTTTTAAAATAG
Y229F F	CTATTTTAAAAATAAAGAACACTGCTTTGCATCGGTGAGAAACG
Y229F R	CGTTTCTCACCGATGCAAAGCAGTGTTCTTTATTTTTAAAATAG
Y229G F	CTATTTTAAAAATAAAGAACACTGCGGTGCATCGGTGAGAAACG
Y229G R	CGTTTCTCACCGATGCACCGCAGTGTTCTTTATTTTTAAAATAG
Y229I F	CTATTTTAAAAATAAAGAACACTGCATTGCATCGGTGAGAAACG
Y229I R	CGTTTCTCACCGATGCAATGCAGTGTTCTTTATTTTTAAAATAG
R233A F	CTGCTATGCATCGGTGGCAAACGTTAAAAAAATAGG
R233A R	CCTATTTTTTTAACGTTTGCCACCGATGCATAGCAG
V235A F	GCATCGGTGAGAAACGCTAAAAAAATATAATATTCTAAATGC
V235A R	GCATTTAGAATATTATATTTTTTTAGCGTTTCTCACCGATGC
K236A F	GCATCGGTGAGAAACGTTGCAAAAATATAATATTCTAAATGC
K236A R	GCATTTAGAATATTATATTTTTGCAACGTTTCTCACCGATGC
K237A F	GCATCGGTGAGAAACGTTAAAGCAATATAATATTCTAAATGC
K237A R	GCATTTAGAATATTATATTGCTTTAACGTTTCTCACCGATGC
I238A F	CGGTGAGAAACGTTAAAAAAGCATAATATTCTAAATGC
I238A R	GCATTTAGAATATTATGCTTTTTTAACGTTTCTCACCG

Table 17: SDM primer Table 3.

Primers for pSN-P2-*agrA* SDM, C228S-I238A.

REFERENCES

Adhikari RP, Arvidson S, Novick RP (2007) A nonsense mutation in agrA accounts for the defect in agr expression and the avirulence of Staphylococcus aureus 8325-4 traP::kan. *Infection and immunity* **75:** 4534-4540

Aguilar PS, Hernandez-Arriaga AM, Cybulski LE, Erazo AC, de Mendoza D (2001) Molecular basis of thermosensing: a two-component signal transduction thermometer in Bacillus subtilis. *The EMBO journal* **20:** 1681-1691

Allen BL, Gerlach GF, Clegg S (1991) Nucleotide sequence and functions of mrk determinants necessary for expression of type 3 fimbriae in Klebsiella pneumoniae. *Journal of bacteriology* **173:** 916-920

Ansari AZ, Chael ML, O'Halloran TV (1992) Allosteric underwinding of DNA is a critical step in positive control of transcription by Hg-MerR. *Nature* **355:** 87-89

Antelmann H, Scharf C, Hecker M (2000) Phosphate starvation-inducible proteins of Bacillus subtilis: proteomics and transcriptional analysis. *Journal of bacteriology* **182:** 4478-4490

Aoyama T, Takanami M, Ohtsuka E, Taniyama Y, Marumoto R, Sato H, Ikehara M (1983) Essential structure of E. coli promoter: effect of spacer length between the two consensus sequences on promoter function. *Nucleic acids research* **11:** 5855-5864

Appleby JL, Bourret RB (1998) Proposed signal transduction role for conserved CheY residue Thr87, a member of the response regulator active-site quintet. *Journal of bacteriology* **180:** 3563-3569

Axelsson L, Holck A (1995) The genes involved in production of and immunity to sakacin A, a bacteriocin from Lactobacillus sake Lb706. *Journal of bacteriology* **177:** 2125-2137

Balaban N, Novick RP (1995) Translation of RNAIII, the Staphylococcus aureus agr regulatory RNA molecule, can be activated by a 3'-end deletion. *FEMS Microbiology Letters* **133:** 155-161

Ballal A, Ray B, Manna AC (2009) sarZ, a sarA family gene, is transcriptionally activated by MgrA and is involved in the regulation of genes encoding exoproteins in Staphylococcus aureus. *Journal of bacteriology* **191:** 1656-1665

Barakat M, Ortet P, Whitworth DE (2011) P2CS: a database of prokaryotic two-component systems. *Nucleic acids research* **39:** D771-776

Barne KA, Bown JA, Busby SJ, Minchin SD (1997) Region 2.5 of the Escherichia coli RNA polymerase sigma70 subunit is responsible for the recognition of the 'extended-10' motif at promoters. *The EMBO journal* **16:** 4034-4040

Beatty CM, Browning DF, Busby SJ, Wolfe AJ (2003) Cyclic AMP receptor protein-dependent activation of the Escherichia coli acsP2 promoter by a synergistic class III mechanism. *Journal of bacteriology* **185:** 5148-5157

Beier D, Gross R (2006) Regulation of bacterial virulence by two-component systems. *Current opinion in microbiology* **9:** 143-152

Belyaeva TA, Bown JA, Fujita N, Ishihama A, Busby SJ (1996) Location of the C-terminal domain of the RNA polymerase alpha subunit in different open complexes at the Escherichia coli galactose operon regulatory region. *Nucleic acids research* **24:** 2242-2251

Belyaeva TA, Rhodius VA, Webster CL, Busby SJ (1998) Transcription activation at promoters carrying tandem DNA sites for the Escherichia coli cyclic AMP receptor protein: organisation of the RNA polymerase alpha subunits. *Journal of molecular biology* **277:** 789-804

Benito Y, Kolb FA, Romby P, Lina G, Etienne J, Vandenesch F (2000) Probing the structure of RNAIII, the Staphylococcus aureus agr regulatory RNA, and identification of the RNA domain involved in repression of protein A expression. *Rna* **6:** 668-679

Bergendahl V, Thompson NE, Foley KM, Olson BM, Burgess RR (2003) A cross-reactive polyol-responsive monoclonal antibody useful for isolation of core RNA polymerase from many bacterial species. *Protein expression and purification* **31:** 155-160

Bigger JW (1933) The production of staphylococcal elemolysin with observations on its mode of action. *The Journal of Pathology and Bacteriology* **36:** 87-114

Bischoff M, Entenza JM, Giachino P (2001) Influence of a functional sigB operon on the global regulators sar and agr in Staphylococcus aureus. *Journal of bacteriology* **183:** 5171-5179

Blatter EE, Ross W, Tang H, Gourse RL, Ebright RH (1994) Domain organization of RNA polymerase alpha subunit: C-terminal 85 amino acids constitute a domain capable of dimerization and DNA-binding. *Cell* **78:** 889-896

Boisset S, Geissmann T, Huntzinger E, Fechter P, Bendridi N, Possedko M, Chevalier C, Helfer AC, Benito Y, Jacquier A, Gaspin C, Vandenesch F, Romby P (2007) Staphylococcus aureus RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes & development* **21:** 1353-1366

Bokal AJ, Ross W, Gaal T, Johnson RC, Gourse RL (1997) Molecular anatomy of a transcription activation patch: FIS-RNA polymerase interactions at the Escherichia coli rrnB P1 promoter. *The EMBO journal* **16:** 154-162

Borukhov S, Nudler E (2008) RNA polymerase: the vehicle of transcription. *Trends in microbiology* **16:** 126-134

Boudes M, Sanchez D, Graille M, van Tilbeurgh H, Durand D, Quevillon-Cheruel S (2014) Structural insights into the dimerization of the response regulator ComE from Streptococcus pneumoniae. *Nucleic acids research* **42:** 5302-5313

Bronner S, Stoessel P, Gravet A, Monteil H, Prevost G (2000) Variable expressions of Staphylococcus aureus bicomponent leucotoxins semiquantified by competitive reverse transcription-PCR. *Applied and environmental microbiology* **66:** 3931-3938

Brown NL, Stoyanov JV, Kidd SP, Hobman JL (2003) The MerR family of transcriptional regulators. *FEMS microbiology reviews* **27:** 145-163

Brunskill EW, Bayles KW (1996a) Identification and molecular characterization of a putative regulatory locus that affects autolysis in Staphylococcus aureus. *Journal of bacteriology* **178:** 611-618

Brunskill EW, Bayles KW (1996b) Identification of LytSR-regulated genes from Staphylococcus aureus. *Journal of bacteriology* **178:** 5810-5812

Buck M, Gallegos MT, Studholme DJ, Guo Y, Gralla JD (2000) The bacterial enhancer-dependent sigma(54) (sigma(N)) transcription factor. *Journal of bacteriology* **182:** 4129-4136

Burr T, Mitchell J, Kolb A, Minchin S, Busby S (2000) DNA sequence elements located immediately upstream of the -10 hexamer in Escherichia coli promoters: a systematic study. *Nucleic acids research* **28:** 1864-1870

Busby S, Ebright RH (1997) Transcription activation at class II CAP-dependent promoters. *Molecular microbiology* **23:** 853-859

Carter GP, Lyras D, Allen DL, Mackin KE, Howarth PM, O'Connor JR, Rood JI (2007) Binary toxin production in Clostridium difficile is regulated by CdtR, a LytTR family response regulator. *Journal of bacteriology* **189:** 7290-7301

Casino P, Rubio V, Marina A (2010) The mechanism of signal transduction by two-component systems. *Current opinion in structural biology* **20:** 763-771

Caspi R, Altman T, Billington R, Dreher K, Foerster H, Fulcher CA, Holland TA, Keseler IM, Kothari A, Kubo A, Krummenacker M, Latendresse M, Mueller LA, Ong Q, Paley S, Subhraveti P, Weaver DS, Weerasinghe D, Zhang P, Karp PD (2014) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic acids research* **42:** D459-471

Chambers HF, Deleo FR (2009) Waves of resistance: Staphylococcus aureus in the antibiotic era. *Nature reviews Microbiology* **7:** 629-641

Chan PF, Foster SJ, Ingham E, Clements MO (1998) The Staphylococcus aureus alternative sigma factor sigmaB controls the environmental stress response but not starvation survival or pathogenicity in a mouse abscess model. *Journal of bacteriology* **180:** 6082-6089

Charpentier E, Anton AI, Barry P, Alfonso B, Fang Y, Novick RP (2004) Novel cassette-based shuttle vector system for gram-positive bacteria. *Applied and environmental microbiology* **70:** 6076-6085

Cheetham GM, Steitz TA (2000) Insights into transcription: structure and function of single-subunit DNA-dependent RNA polymerases. *Current opinion in structural biology* **10:** 117-123

Chen H, Tang H, Ebright RH (2003) Functional Interaction between RNA Polymerase α Subunit C-Terminal Domain and σ70 in UP-Element- and Activator-Dependent Transcription. *Molecular cell* **11:** 1621-1633

Cheung AL, Koomey JM, Butler CA, Projan SJ, Fischetti VA (1992) Regulation of exoprotein expression in Staphylococcus aureus by a locus (sar) distinct from agr. *Proceedings of the National Academy of Sciences of the United States of America* **89:** 6462-6466

Cheung GY, Wang R, Khan BA, Sturdevant DE, Otto M (2011) Role of the accessory gene regulator agr in community-associated methicillin-resistant Staphylococcus aureus pathogenesis. *Infection and immunity* **79:** 1927-1935

Chevalier C, Boisset S, Romilly C, Masquida B, Fechter P, Geissmann T, Vandenesch F, Romby P (2010) Staphylococcus aureus RNAIII binds to two distant regions of coa mRNA to arrest translation and promote mRNA degradation. *PLoS pathogens* **6:** e1000809

Chong S, Mersha FB, Comb DG, Scott ME, Landry D, Vence LM, Perler FB, Benner J, Kucera RB, Hirvonen CA, Pelletier JJ, Paulus H, Xu MQ (1997) Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. *Gene* **192:** 271-281

Chong S, Montello GE, Zhang A, Cantor EJ, Liao W, Xu MQ, Benner J (1998) Utilizing the C-terminal cleavage activity of a protein splicing element to purify recombinant proteins in a single chromatographic step. *Nucleic acids research* **26:** 5109-5115

Coleman DC, Arbuthnott JP, Pomeroy HM, Birkbeck TH (1986) Cloning and expression in Escherichia coli and Staphylococcus aureus of the beta-lysin determinant from Staphylococcus aureus: evidence that bacteriophage conversion of beta-lysin activity is caused by insertional inactivation of the beta-lysin determinant. *Microbial pathogenesis* **1:** 549-564

Corrigan RM, Abbott JC, Burhenne H, Kaever V, Grundling A (2011) c-di-AMP is a new second messenger in Staphylococcus aureus with a role in controlling cell size and envelope stress. *PLoS pathogens* **7:** e1002217

Corrigan RM, Campeotto I, Jeganathan T, Roelofs KG, Lee VT, Grundling A (2013) Systematic identification of conserved bacterial c-di-AMP receptor proteins. *Proceedings of the National Academy of Sciences of the United States of America* **110:** 9084-9089

Cramer P (2002a) Common structural features of nucleic acid polymerases. *BioEssays : news and reviews in molecular, cellular and developmental biology* **24:** 724-729

Cramer P (2002b) Multisubunit RNA polymerases. *Current opinion in structural biology* **12:** 89-97

Da Re S, Tolstykh T, Wolanin PM, Stock JB (2002) Genetic analysis of response regulator activation in bacterial chemotaxis suggests an intermolecular mechanism. *Protein science : a publication of the Protein Society* **11:** 2644-2654

de Saizieu A, Gardes C, Flint N, Wagner C, Kamber M, Mitchell TJ, Keck W, Amrein KE, Lange R (2000) Microarray-based identification of a novel Streptococcus pneumoniae regulon controlled by an autoinduced peptide. *Journal of bacteriology* **182:** 4696-4703

Deora R, Misra TK (1996) Characterization of the primary sigma factor of Staphylococcus aureus. *The Journal of biological chemistry* **271:** 21828-21834

Deretic V, Dikshit R, Konyecsni WM, Chakrabarty AM, Misra TK (1989) The algR gene, which regulates mucoidy in Pseudomonas aeruginosa, belongs to a class of environmentally responsive genes. *Journal of bacteriology* **171:** 1278-1283

Diep DB, Havarstein LS, Nes IF (1996) Characterization of the locus responsible for the bacteriocin production in Lactobacillus plantarum C11. *Journal of bacteriology* **178:** 4472-4483

Dinges MM, Orwin PM, Schlievert PM (2000) Exotoxins of Staphylococcus aureus. *Clinical microbiology reviews* **13:** 16-34, table of contents

Doery HM, Magnusson BJ, Cheyne IM, Sulasekharam J (1963) A phospholipase in staphylococcal toxin which hydrolyses sphingomyelin. *Nature* **198:** 1091-1092

Dorman G, Prestwich GD (1994) Benzophenone photophores in biochemistry. *Biochemistry* **33:** 5661-5673

Dowell PM, Affas Z, Reynolds C, Holden MT, Wood SJ, Saint S, Cockayne A, Hill PJ, Dodd CE, Bycroft BW, Chan WC, Williams P (2001) Structure, activity and evolution of the group I thiolactone peptide quorum-sensing system of Staphylococcus aureus. *Molecular microbiology* **41:** 503-512

Dubrac S, Boneca IG, Poupel O, Msadek T (2007) New insights into the WalK/WalR (YycG/YycF) essential signal transduction pathway reveal a major role in controlling cell wall metabolism and biofilm formation in Staphylococcus aureus. *Journal of bacteriology* **189:** 8257-8269

Dufour P, Jarraud S, Vandenesch F, Greenland T, Novick RP, Bes M, Etienne J, Lina G (2002) High genetic variability of the agr locus in Staphylococcus species. *Journal of bacteriology* **184:** 1180-1186

Dunman PM, Murphy E, Haney S, Palacios D, Tucker-Kellogg G, Wu S, Brown EL, Zagursky RJ, Shlaes D, Projan SJ (2001) Transcription profiling-based identification of Staphylococcus aureus genes regulated by the agr and/or sarA loci. *Journal of bacteriology* **183:** 7341-7353

Duval-Iflah Y (1972) Lysogenic conversion of the lipase in Staphylococcus pyogenes group III strains. *Canadian Journal of Microbiology* **18:** 1491-1497

Dyer CM, Quillin ML, Campos A, Lu J, McEvoy MM, Hausrath AC, Westbrook EM, Matsumura P, Matthews BW, Dahlquist FW (2004) Structure of the constitutively active double mutant CheYD13K Y106W alone and in complex with a FliM peptide. *Journal of molecular biology* **342:** 1325-1335

Ebright RH, Busby S (1995) The Escherichia coli RNA polymerase α subunit: structure and function. *Current Opinion in Genetics & Development* **5:** 197-203

Essmann UP, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G.; (1995) A Smooth Particle Mesh Ewald Method. *Journal of Chemical Physics* **103:** 8577-8593

Falord M, Karimova G, Hiron A, Msadek T (2012) GraXSR proteins interact with the VraFG ABC transporter to form a five-component system required for cationic antimicrobial peptide sensing and resistance in Staphylococcus aureus. *Antimicrobial agents and chemotherapy* **56:** 1047-1058

Falord M, Mader U, Hiron A, Debarbouille M, Msadek T (2011) Investigation of the Staphylococcus aureus GraSR regulon reveals novel links to virulence, stress response and cell wall signal transduction pathways. *PloS one* **6:** e21323

Fedtke I, Kamps A, Krismer B, Gotz F (2002) The nitrate reductase and nitrite reductase operons and the narT gene of Staphylococcus carnosus are positively controlled by the novel two-component system NreBC. *Journal of bacteriology* **184:** 6624-6634

Feklistov A, Barinova N, Sevostyanova A, Heyduk E, Bass I, Vvedenskaya I, Kuznedelov K, Merkiene E, Stavrovskaya E, Klimasauskas S, Nikiforov V, Heyduk T, Severinov K, Kulbachinskiy A (2006) A basal promoter element recognized by free RNA polymerase sigma subunit determines promoter recognition by RNA polymerase holoenzyme. *Molecular cell* **23:** 97-107

Feklistov A, Darst SA (2011) Structural basis for promoter-10 element recognition by the bacterial RNA polymerase sigma subunit. *Cell* **147:** 1257-1269

Finkel SE, Johnson RC (1993) The Fis protein: it's not just for DNA inversion anymore. *Molecular microbiology* **7:** 1023

Firestone GL, Winguth SD (1990) [52]Immunoprecipitation of proteins. In *Methods in Enzymology*, Murray PD (ed), Vol. Volume 182, pp 688-700. Academic Press

Forne I, Ludwigsen J, Imhof A, Becker PB, Mueller-Planitz F (2012) Probing the conformation of the ISWI ATPase domain with genetically encoded photoreactive crosslinkers and mass spectrometry. *Molecular & cellular proteomics : MCP* **11:** M111 012088

Fournier B, Klier A, Rapoport G (2001) The two-component system ArlS-ArlR is a regulator of virulence gene expression in Staphylococcus aureus. *Molecular microbiology* **41:** 247-261

Freer JH, Birkbeck TH (1982) Possible conformation of delta-lysin, a membrane-damaging peptide of Staphylococcus aureus. *Journal of theoretical biology* **94:** 535-540

Friaes A, Resina C, Manuel V, Lito L, Ramirez M, Melo-Cristino J (2014) Epidemiological survey of the first case of vancomycin-resistant Staphylococcus aureus infection in Europe. *Epidemiology and infection***:** 1-4

Galperin MY (2010) Diversity of structure and function of response regulator output domains. *Current opinion in microbiology* **13:** 150-159

Gao R, Mack TR, Stock AM (2007) Bacterial response regulators: versatile regulatory strategies from common domains. *Trends in biochemical sciences* **32:** 225-234

Gao R, Stock AM (2009) Biological insights from structures of two-component proteins. *Annual review of microbiology* **63:** 133-154

Gardella T, Moyle H, Susskind MM (1989) A mutant Escherichia coli σ70 subunit of RNA polymerase with altered promoter specificity. *Journal of molecular biology* **206:** 579-590

Gardete S, Wu SW, Gill S, Tomasz A (2006) Role of VraSR in antibiotic resistance and antibiotic-induced stress response in Staphylococcus aureus. *Antimicrobial agents and chemotherapy* **50:** 3424-3434

Gaston K, Bell A, Kolb A, Buc H, Busby S (1990) Stringent spacing requirements for transcription activation by CRP. *Cell* **62:** 733-743

Geisinger E, Adhikari RP, Jin R, Ross HF, Novick RP (2006) Inhibition of rot translation by RNAIII, a key feature of agr function. *Molecular microbiology* **61:** 1038-1048

Geisinger E, Chen J, Novick RP (2012) Allele-dependent differences in quorum-sensing dynamics result in variant expression of virulence genes in Staphylococcus aureus. *Journal of bacteriology* **194:** 2854-2864

Geisinger E, Muir TW, Novick RP (2009) agr receptor mutants reveal distinct modes of inhibition by staphylococcal autoinducing peptides. *Proceedings of the National Academy of Sciences of the United States of America* **106:** 1216-1221

George Cisar EA, Geisinger E, Muir TW, Novick RP (2009) Symmetric signalling within asymmetric dimers of the Staphylococcus aureus receptor histidine kinase AgrC. *Molecular microbiology* **74:** 44-57

Giese B, Glowinski F, Paprotka K, Dittmann S, Steiner T, Sinha B, Fraunholz MJ (2011) Expression of delta-toxin by Staphylococcus aureus mediates escape from phago-endosomes of human epithelial and endothelial cells in the presence of beta-toxin. *Cellular microbiology* **13:** 316-329

Gille H, Egan JB, Roth A, Messer W (1991) The FIS protein binds and bends the origin of chromosomal DNA replication, oriC, of Escherichia coli. *Nucleic acids research* **19:** 4167-4172

Gordon JC, Myers JB, Folta T, Shoja V, Heath LS, Onufriev A (2005) H++: a server for estimating pKas and adding missing hydrogens to macromolecules. *Nucleic acids research* **33:** W368-371

Gourse RL, Ross W, Gaal T (2000) UPs and downs in bacterial transcription initiation: the role of the alpha subunit of RNA polymerase in promoter recognition. *Molecular microbiology* **37:** 687-695

Griffith KL, Shah IM, Myers TE, O'Neill MC, Wolf RE, Jr. (2002) Evidence for "pre-recruitment" as a new mechanism of transcription activation in Escherichia coli: the large excess of SoxS binding sites per cell relative to the number of SoxS molecules per cell. *Biochemical and biophysical research communications* **291:** 979-986

Griffith KL, Wolf RE, Jr. (2004) Genetic evidence for pre-recruitment as the mechanism of transcription activation by SoxS of Escherichia coli: the dominance of DNA-binding mutations of SoxS. *Journal of molecular biology* **344:** 1-10

Gruber TM, Gross CA (2003) Multiple sigma subunits and the partitioning of bacterial transcription space. *Annual review of microbiology* **57:** 441-466

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41:** 95-98

Hanssen A-M, Ericson Sollid JU (2006) SCCmec in staphylococci: genes on the move. *FEMS Immunology & Medical Microbiology* **46:** 8-20

Haugen SP, Berkmen MB, Ross W, Gaal T, Ward C, Gourse RL (2006) rRNA promoter regulation by nonoptimal binding of sigma region 1.2: an additional recognition element for RNA polymerase. *Cell* **125:** 1069-1082

Haugen SP, Ross W, Gourse RL (2008) Advances in bacterial promoter recognition and its control by factors that do not bind DNA. *Nature reviews Microbiology* **6:** 507-519

Heldwein EE, Brennan RG (2001) Crystal structure of the transcription activator BmrR bound to DNA and a drug. *Nature* **409:** 378-382

Hidalgo-Grass C, Ravins M, Dan-Goor M, Jaffe J, Moses AE, Hanski E (2002) A locus of group A Streptococcus involved in invasive disease and DNA transfer. *Molecular microbiology* **46:** 87-99

Hiron A, Falord M, Valle J, Debarbouille M, Msadek T (2011) Bacitracin and nisin resistance in Staphylococcus aureus: a novel pathway involving the BraS/BraR two-component system (SA2417/SA2418) and both the BraD/BraE and VraD/VraE ABC transporters. *Molecular microbiology* **81:** 602-622

Hobman JL (2007) MerR family transcription activators: similar designs, different specificities. *Molecular microbiology* **63:** 1275-1278

Horsburgh MJ, Aish JL, White IJ, Shaw L, Lithgow JK, Foster SJ (2002) sigmaB modulates virulence determinant expression and stress resistance: characterization of a functional rsbU strain derived from Staphylococcus aureus 8325-4. *Journal of bacteriology* **184:** 5457-5467

Huhne K, Axelsson L, Holck A, Krockel L (1996) Analysis of the sakacin P gene cluster from Lactobacillus sake Lb674 and its expression in sakacin-negative Lb. sake strains. *Microbiology* **142 (Pt 6):** 1437-1448

Hulett FM, Lee J, Shi L, Sun G, Chesnut R, Sharkova E, Duggan MF, Kapp N (1994) Sequential action of two-component genetic switches regulates the PHO regulon in Bacillus subtilis. *Journal of bacteriology* **176:** 1348-1358

Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. *Journal of molecular graphics* **14:** 33-38, 27-38

Huntzinger E, Boisset S, Saveanu C, Benito Y, Geissmann T, Namane A, Lina G, Etienne J, Ehresmann B, Ehresmann C, Jacquier A, Vandenesch F, Romby P (2005) Staphylococcus aureus RNAIII and the endoribonuclease III coordinately regulate spa gene expression. *The EMBO journal* **24:** 824-835

Igarashi K, Hanamura A, Makino K, Aiba H, Aiba H, Mizuno T, Nakata A, Ishihama A (1991) Functional map of the alpha subunit of Escherichia coli RNA polymerase: two modes of transcription activation by positive factors. *Proceedings of the National Academy of Sciences of the United States of America* **88:** 8958-8962

Igarashi K, Ishihama A (1991) Bipartite functional map of the E. coli RNA polymerase alpha subunit: involvement of the C-terminal region in transcription activation by cAMP-CRP. *Cell* **65:** 1015-1022

Imashimizu M, Shimamoto N, Oshima T, Kashlev M (2014) Transcription elongation: Heterogeneous tracking of RNA polymerase and its biological implications. *Transcription* **5**

Ingavale S, van Wamel W, Luong TT, Lee CY, Cheung AL (2005) Rat/MgrA, a regulator of autolysis, is a regulator of virulence genes in Staphylococcus aureus. *Infection and immunity* **73:** 1423-1431

Ingavale SS, Van Wamel W, Cheung AL (2003) Characterization of RAT, an autolysis regulator in Staphylococcus aureus. *Molecular microbiology* **48:** 1451-1466

Ira, Johnston LJ (2008) Sphingomyelinase generation of ceramide promotes clustering of nanoscale domains in supported bilayer membranes. *Biochimica et biophysica acta* **1778:** 185-197

Ito T, Katayama Y, Asada K, Mori N, Tsutsumimoto K, Tiensasitorn C, Hiramatsu K (2001) Structural comparison of three types of staphylococcal cassette chromosome mec integrated in the chromosome in methicillin-resistant Staphylococcus aureus. *Antimicrobial agents and chemotherapy* **45:** 1323-1336

Jain D, Kim Y, Maxwell KL, Beasley S, Zhang R, Gussin GN, Edwards AM, Darst SA (2005) Crystal structure of bacteriophage lambda cII and its DNA complex. *Molecular cell* **19:** 259-269

James EH, Edwards AM, Wigneshweraraj S (2013) Transcriptional downregulation of agr expression in Staphylococcus aureus during growth in human serum can be overcome by constitutively active mutant forms of the sensor kinase AgrC. *FEMS microbiology letters* **349:** 153-162

Janzon L, Arvidson S (1990) The role of the delta-lysin gene (hld) in the regulation of virulence genes by the accessory gene regulator (agr) in Staphylococcus aureus. *The EMBO journal* **9:** 1391-1399

Janzon L, Lofdahl S, Arvidson S (1989) Identification and nucleotide sequence of the delta-lysin gene, hld, adjacent to the accessory gene regulator (agr) of Staphylococcus aureus. *Molecular & general genetics : MGG* **219:** 480-485

Ji G, Beavis R, Novick RP (1997) Bacterial interference caused by autoinducing peptide variants. *Science* **276:** 2027-2030

Ji G, Beavis RC, Novick RP (1995) Cell density control of staphylococcal virulence mediated by an octapeptide pheromone. *Proceedings of the National Academy of Sciences of the United States of America* **92:** 12055-12059

Jorgensen WLC, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. (1983) Comparison of Simple Potential Functions for Simulating Liquid Water. *Journal of Chemical Physics* **79:** 926-935

Joung JK, Le LU, Hochschild A (1993) Synergistic activation of transcription by Escherichia coli cAMP receptor protein. *Proceedings of the National Academy of Sciences of the United States of America* **90:** 3083-3087

Julianelle LA (1922) Studies of Hemolytic Staphylococci: Hemolytic Avtivity-Biochemical Reactions-Serologic Reactions. *The Journal of Infectious Diseases* **31:** 256-284

Kaito C, Morishita D, Matsumoto Y, Kurokawa K, Sekimizu K (2006) Novel DNA-binding protein SarZ contributes to virulence in Staphylococcus aureus. *Molecular microbiology* **62:** 1601-1617

Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M (2014) Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic acids research* **42:** D199-205

Karimova G, Pidoux J, Ullmann A, Ladant D (1998) A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proceedings of the National Academy of Sciences of the United States of America* **95:** 5752-5756

Kavanaugh JS, Thoendel M, Horswill AR (2007) A role for type I signal peptidase in Staphylococcus aureus quorum sensing. *Molecular microbiology* **65:** 780-798

Kawada-Matsuo M, Yoshida Y, Nakamura N, Komatsuzawa H (2011) Role of two-component systems in the resistance of Staphylococcus aureus to antibacterial agents. *Virulence* **2:** 427-430

Keilty S, Rosenberg M (1987) Constitutive function of a positively regulated promoter reveals new sequences essential for activity. *The Journal of biological chemistry* **262:** 6389-6395

Kelley LA, Sternberg MJ (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nature protocols* **4:** 363-371

Kiil K, Ferchaud JB, David C, Binnewies TT, Wu H, Sicheritz-Ponten T, Willenbrock H, Ussery DW (2005) Genome update: distribution of two-component transduction systems in 250 bacterial genomes. *Microbiology* **151:** 3447-3452

Kluytmans J, van Belkum A, Verbrugh H (1997) Nasal carriage of Staphylococcus aureus: epidemiology, underlying mechanisms, and associated risks. *Clinical microbiology reviews* **10:** 505-520

Koenig RL, Ray JL, Maleki SJ, Smeltzer MS, Hurlburt BK (2004) Staphylococcus aureus AgrA binding to the RNAIII-agr regulatory region. *Journal of bacteriology* **186:** 7549-7555

Korzheva N, Mustaev A, Kozlov M, Malhotra A, Nikiforov V, Goldfarb A, Darst SA (2000) A structural model of transcription elongation. *Science* **289:** 619-625

Kullik I, Giachino P, Fuchs T (1998) Deletion of the alternative sigma factor sigmaB in Staphylococcus aureus reveals its function as a global regulator of virulence genes. *Journal of bacteriology* **180:** 4814-4820

Kuroda M, Kuroda H, Oshima T, Takeuchi F, Mori H, Hiramatsu K (2003) Two-component system VraSR positively modulates the regulation of cell-wall biosynthesis pathway in Staphylococcus aureus. *Molecular microbiology* **49:** 807-821

Kuroda M, Ohta T, Uchiyama I, Baba T, Yuzawa H, Kobayashi I, Cui L, Oguchi A, Aoki K, Nagai Y, Lian J, Ito T, Kanamori M, Matsumaru H, Maruyama A, Murakami H, Hosoyama A, Mizutani-Ui Y, Takahashi NK, Sawano T, Inoue R, Kaito C, Sekimizu K, Hirakawa H, Kuhara S, Goto S, Yabuzaki J, Kanehisa M, Yamashita A, Oshima K, Furuya K, Yoshino C, Shiba T, Hattori M, Ogasawara N, Hayashi H, Hiramatsu K (2001) Whole genome sequencing of meticillin-resistant Staphylococcus aureus. *Lancet* **357:** 1225-1240

Lane WJ, Darst SA (2010) Molecular evolution of multisubunit RNA polymerases: sequence analysis. *Journal of molecular biology* **395:** 671-685

Lange R, Wagner C, de Saizieu A, Flint N, Molnos J, Stieger M, Caspers P, Kamber M, Keck W, Amrein KE (1999) Domain organization and molecular characterization of 13 two-component systems identified by genome sequencing of Streptococcus pneumoniae. *Gene* **237:** 223-234

Lawson CL, Swigon D, Murakami KS, Darst SA, Berman HM, Ebright RH (2004) Catabolite activator protein: DNA-binding and transcription activation. *Current opinion in structural biology* **14:** 10-20

Lee CY, Buranen SL, Ye ZH (1991) Construction of single-copy integration vectors for Staphylococcus aureus. *Gene* **103:** 101-105

Lee CY, Iandolo JJ (1985) Mechanism of bacteriophage conversion of lipase activity in Staphylococcus aureus. *Journal of bacteriology* **164:** 288-293

Lee CY, Iandolo JJ (1986a) Integration of staphylococcal phage L54a occurs by site-specific recombination: structural analysis of the attachment sites. *Proceedings of the National Academy of Sciences of the United States of America* **83:** 5474-5478

Lee CY, Iandolo JJ (1986b) Lysogenic conversion of staphylococcal lipase is caused by insertion of the bacteriophage L54a genome into the lipase structural gene. *Journal of bacteriology* **166:** 385-391

Lee DJ, Minchin SD, Busby SJ (2012) Activating transcription in bacteria. *Annual review of microbiology* **66:** 125-152

Lee SY, Cho HS, Pelton JG, Yan D, Berry EA, Wemmer DE (2001) Crystal structure of activated CheY. Comparison with other activated receiver domains. *The Journal of biological chemistry* **276:** 16425-16431

Leonard PG, Bezar IF, Sidote DJ, Stock AM (2012) Identification of a hydrophobic cleft in the LytTR domain of AgrA as a locus for small molecule interactions that inhibit DNA-binding. *Biochemistry* **51:** 10035-10043

Li M, Cha DJ, Lai Y, Villaruz AE, Sturdevant DE, Otto M (2007) The antimicrobial peptide-sensing system aps of Staphylococcus aureus. *Molecular microbiology* **66:** 1136-1147

Liang X, Yu C, Sun J, Liu H, Landwehr C, Holmes D, Ji Y (2006) Inactivation of a two-component signal transduction system, SaeRS, eliminates adherence and attenuates virulence of Staphylococcus aureus. *Infection and immunity* **74:** 4655-4665

Lina G, Jarraud S, Ji G, Greenland T, Pedraza A, Etienne J, Novick RP, Vandenesch F (1998) Transmembrane topology and histidine protein kinase activity of AgrC, the agr signal receptor in Staphylococcus aureus. *Molecular microbiology* **28:** 655-662

Lindsay JA (2010) Genomic variation and evolution of Staphylococcus aureus. *International journal of medical microbiology : IJMM* **300:** 98-103

Liu CC, Schultz PG (2010) Adding new chemistries to the genetic code. *Annual review of biochemistry* **79:** 413-444

Liu GY, Essex A, Buchanan JT, Datta V, Hoffman HM, Bastian JF, Fierer J, Nizet V (2005) Staphylococcus aureus golden pigment impairs neutrophil killing and promotes virulence through its antioxidant activity. *The Journal of experimental medicine* **202:** 209-215

Lowy FD (1998) Staphylococcus aureus infections. *The New England journal of medicine* **339:** 520-532

Lukat GS, Lee BH, Mottonen JM, Stock AM, Stock JB (1991) Roles of the highly conserved aspartate and lysine residues in the response regulator of bacterial chemotaxis. *The Journal of biological chemistry* **266:** 8348-8354

Lukat GS, McCleary WR, Stock AM, Stock JB (1992) Phosphorylation of bacterial response regulator proteins by low molecular weight phospho-donors. *Proceedings of the National Academy of Sciences of the United States of America* **89:** 718-722

Lukat GS, Stock AM, Stock JB (1990) Divalent metal ion binding to the CheY protein and its significance to phosphotransfer in bacterial chemotaxis. *Biochemistry* **29:** 5436-5442

Luong TT, Newell SW, Lee CY (2003) Mgr, a novel global regulator in Staphylococcus aureus. *Journal of bacteriology* **185:** 3703-3710

Lyon GJ, Wright JS, Muir TW, Novick RP (2002) Key determinants of receptor activation in the agr autoinducing peptides of Staphylococcus aureus. *Biochemistry* **41:** 10095-10104

Majerczyk CD, Sadykov MR, Luong TT, Lee C, Somerville GA, Sonenshein AL (2008) Staphylococcus aureus CodY negatively regulates virulence gene expression. *Journal of bacteriology* **190:** 2257-2265

Malke H, Steiner K (2004) Control of streptokinase gene expression in group A & C streptococci by two-component regulators. *The Indian journal of medical research* **119 Suppl:** 48-56

Maniatis T, Ptashne M, Backman K, Kield D, Flashman S, Jeffrey A, Maurer R (1975) Recognition sequences of repressor and polymerase in the operators of bacteriophage lambda. *Cell* **5:** 109-113

Manna A, Cheung AL (2001) Characterization of sarR, a modulator of sar expression in Staphylococcus aureus. *Infection and immunity* **69:** 885-896

Manna AC, Cheung AL (2003) sarU, a sarA homolog, is repressed by SarT and regulates virulence genes in Staphylococcus aureus. *Infection and immunity* **71:** 343-353

Manna AC, Cheung AL (2006a) Expression of SarX, a negative regulator of agr and exoprotein synthesis, is activated by MgrA in Staphylococcus aureus. *Journal of bacteriology* **188:** 4288-4299

Manna AC, Cheung AL (2006b) Transcriptional regulation of the agr locus and the identification of DNA-binding residues of the global regulatory protein SarR in Staphylococcus aureus. *Molecular microbiology* **60:** 1289-1301

Markwick PR, McCammon JA (2011) Studying functional dynamics in bio-molecules using accelerated molecular dynamics. *Physical chemistry chemical physics : PCCP* **13:** 20053-20065

Martin PK, Li T, Sun D, Biek DP, Schmid MB (1999) Role in cell permeability of an essential two-component system in Staphylococcus aureus. *Journal of bacteriology* **181:** 3666-3673

Martin RG, Gillette WK, Martin NI, Rosner JL (2002) Complex formation between activator and RNA polymerase as the basis for transcriptional activation by MarA and SoxS in Escherichia coli. *Molecular microbiology* **43:** 355-370

Matsuo M, Kato F, Oogai Y, Kawai T, Sugai M, Komatsuzawa H (2010) Distinct two-component systems in methicillin-resistant Staphylococcus aureus can change the susceptibility to antimicrobial agents. *The Journal of antimicrobial chemotherapy* **65:** 1536-1537

Matsuo M, Oogai Y, Kato F, Sugai M, Komatsuzawa H (2011) Growth-phase dependence of susceptibility to antimicrobial peptides in Staphylococcus aureus. *Microbiology* **157:** 1786-1797

Mayville P, Ji G, Beavis R, Yang H, Goger M, Novick RP, Muir TW (1999) Structure-activity analysis of synthetic autoinducing thiolactone peptides from Staphylococcus aureus responsible for virulence. *Proceedings of the National Academy of Sciences of the United States of America* **96:** 1218-1223

McCleary WR, Stock JB (1994) Acetyl phosphate and the activation of two-component response regulators. *The Journal of biological chemistry* **269:** 31567-31572

McGowan S, Lucet IS, Cheung JK, Awad MM, Whisstock JC, Rood JI (2002) The FxRxHrS motif: a conserved region essential for DNA-binding of the VirR response regulator from Clostridium perfringens. *Journal of molecular biology* **322:** 997-1011

McGowan S, O'Connor JR, Cheung JK, Rood JI (2003) The SKHR motif is required for biological function of the VirR response regulator from Clostridium perfringens. *Journal of bacteriology* **185:** 6205-6208

Minakhin L, Bhagat S, Brunning A, Campbell EA, Darst SA, Ebright RH, Severinov K (2001) Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. *Proceedings of the National Academy of Sciences of the United States of America* **98:** 892-897

Mitchell JE, Zheng D, Busby SJ, Minchin SD (2003) Identification and analysis of 'extended -10' promoters in Escherichia coli. *Nucleic acids research* **31:** 4689-4695

Monk IR, Shah IM, Xu M, Tan MW, Foster TJ (2012) Transforming the untransformable: application of direct transformation to manipulate genetically Staphylococcus aureus and Staphylococcus epidermidis. *mBio* **3**

Morfeldt E, Janzon L, Arvidson S, Lofdahl S (1988) Cloning of a chromosomal locus (exp) which regulates the expression of several exoprotein genes in Staphylococcus aureus. *Molecular & general genetics : MGG* **211:** 435-440

Morfeldt E, Taylor D, von Gabain A, Arvidson S (1995) Activation of alpha-toxin translation in Staphylococcus aureus by the trans-encoded antisense RNA, RNAIII. *The EMBO journal* **14:** 4569-4577

Morfeldt E, Tegmark K, Arvidson S (1996) Transcriptional control of the agr-dependent virulence gene regulator, RNAIII, in Staphylococcus aureus. *Molecular microbiology* **21:** 1227-1237

Morikawa K, Takemura AJ, Inose Y, Tsai M, Nguyen Thi le T, Ohta T, Msadek T (2012) Expression of a cryptic secondary sigma factor gene unveils natural competence for DNA transformation in Staphylococcus aureus. *PLoS pathogens* **8:** e1003003

Moyle H, Waldburger C, Susskind MM (1991) Hierarchies of base pair preferences in the P22 ant promoter. *Journal of bacteriology* **173:** 1944-1950

Murakami K, Owens JT, Belyaeva TA, Meares CF, Busby SJ, Ishihama A (1997) Positioning of two alpha subunit carboxy-terminal domains of RNA polymerase at promoters by two transcription factors. *Proceedings of the National Academy of Sciences of the United States of America* **94:** 11274-11278

Nikolskaya AN, Galperin MY (2002) A novel type of conserved DNA-binding domain in the transcriptional regulators of the AlgR/AgrA/LytR family. *Nucleic acids research* **30:** 2453-2459

Nikolskaya AN, Galperin MY. (2007) A census of LytTR domain proteins encoded in bacterial genomes. In http://www.ncbi.nlm.nih.gov/Complete_Genomes/LytTR.html (ed.).

Nilkens S, Koch-Singenstreu M, Niemann V, Gotz F, Stehle T, Unden G (2014) Nitrate/oxygen co-sensing by an NreA/NreB sensor complex of Staphylococcus carnosus. *Molecular microbiology* **91:** 381-393

Niu W, Kim Y, Tau G, Heyduk T, Ebright RH (1996) Transcription activation at class II CAP-dependent promoters: two interactions between CAP and RNA polymerase. *Cell* **87:** 1123-1134

Novick RP (2003) Autoinduction and signal transduction in the regulation of staphylococcal virulence. *Molecular microbiology* **48:** 1429-1449

Novick RP, Geisinger E (2008) Quorum sensing in staphylococci. *Annual review of genetics* **42:** 541-564

Novick RP, Jiang D (2003) The staphylococcal saeRS system coordinates environmental signals with agr quorum sensing. *Microbiology* **149:** 2709-2717

Novick RP, Projan SJ, Kornblum J, Ross HF, Ji G, Kreiswirth B, Vandenesch F, Moghazeh S (1995) The agr P2 operon: an autocatalytic sensory transduction system in Staphylococcus aureus. *Molecular & general genetics : MGG* **248:** 446-458

Novick RP, Ross HF, Projan SJ, Kornblum J, Kreiswirth B, Moghazeh S (1993) Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. *The EMBO journal* **12:** 3967-3975

Nygaard TK, Pallister KB, Ruzevich P, Griffith S, Vuong C, Voyich JM (2010) SaeR binds a consensus sequence within virulence gene promoters to advance USA300 pathogenesis. *The Journal of infectious diseases* **201:** 241-254

O'Keeffe T, Hill C, Ross RP (1999) Characterization and heterologous expression of the genes encoding enterocin a production, immunity, and regulation in Enterococcus faecium DPC1146. *Applied and environmental microbiology* **65:** 1506-1515

Ogston A (1880) Classics in infectious diseases. "On abscesses". Alexander Ogston (1844-1929). *Reviews of infectious diseases* **6:** 122-128

Peacock SJ, de Silva I, Lowy FD (2001) What determines nasal carriage of Staphylococcus aureus? *Trends in microbiology* **9:** 605-610

Peng HL, Novick RP, Kreiswirth B, Kornblum J, Schlievert P (1988) Cloning, characterization, and sequencing of an accessory gene regulator (agr) in Staphylococcus aureus. *Journal of bacteriology* **170:** 4365-4372

Perry J, Koteva K, Wright G (2011) Receptor domains of two-component signal transduction systems. *Molecular bioSystems* **7:** 1388-1398

Peschel A, Otto M (2013) Phenol-soluble modulins and staphylococcal infection. *Nature reviews Microbiology* **11:** 667-673

Pestova EV, Havarstein LS, Morrison DA (1996) Regulation of competence for genetic transformation in Streptococcus pneumoniae by an auto-induced peptide pheromone and a two-component regulatory system. *Molecular microbiology* **21:** 853-862

Pierce LC, Salomon-Ferrer R, Augusto FdOC, McCammon JA, Walker RC (2012) Routine Access to Millisecond Time Scale Events with Accelerated Molecular Dynamics. *Journal of chemical theory and computation* **8:** 2997-3002

Ponder JW, Case DA (2003) Force fields for protein simulations. *Advances in protein chemistry* **66:** 27-85

Pragman AA, Ji Y, Schlievert PM (2007) Repression of Staphylococcus aureus SrrAB using inducible antisense srrA alters growth and virulence factor transcript levels. *Biochemistry* **46:** 314-321

Pragman AA, Yarwood JM, Tripp TJ, Schlievert PM (2004) Characterization of virulence factor regulation by SrrAB, a two-component system in Staphylococcus aureus. *Journal of bacteriology* **186:** 2430-2438

Prevost G, Couppie P, Prevost P, Gayet S, Petiau P, Cribier B, Monteil H, Piemont Y (1995) Epidemiological data on Staphylococcus aureus strains producing synergohymenotropic toxins. *Journal of medical microbiology* **42:** 237-245

Pribnow D (1975) Nucleotide sequence of an RNA polymerase binding site at an early T7 promoter. *Proceedings of the National Academy of Sciences of the United States of America* **72:** 784-788

Qin X, Singh KV, Weinstock GM, Murray BE (2001) Characterization of fsr, a regulator controlling expression of gelatinase and serine protease in Enterococcus faecalis OG1RF. *Journal of bacteriology* **183:** 3372-3382

Qiu R, Pei W, Zhang L, Lin J, Ji G (2005) Identification of the putative staphylococcal AgrB catalytic residues involving the proteolytic cleavage of AgrD to generate autoinducing peptide. *The Journal of biological chemistry* **280:** 16695-16704

Quadri LE, Kleerebezem M, Kuipers OP, de Vos WM, Roy KL, Vederas JC, Stiles ME (1997) Characterization of a locus from Carnobacterium piscicola LV17B involved in bacteriocin

production and immunity: evidence for global inducer-mediated transcriptional regulation. *Journal of bacteriology* **179:** 6163-6171

Queck SY, Jameson-Lee M, Villaruz AE, Bach TH, Khan BA, Sturdevant DE, Ricklefs SM, Li M, Otto M (2008) RNAIII-independent target gene control by the agr quorum-sensing system: insight into the evolution of virulence regulation in Staphylococcus aureus. *Molecular cell* **32:** 150-158

Quon KC, Marczynski GT, Shapiro L (1996) Cell cycle control by an essential bacterial two-component signal transduction protein. *Cell* **84:** 83-93

Rechtin TM, Gillaspy AF, Schumacher MA, Brennan RG, Smeltzer MS, Hurlburt BK (1999) Characterization of the SarA virulence gene regulator of Staphylococcus aureus. *Molecular microbiology* **33:** 307-316

Recsei P, Kreiswirth B, O'Reilly M, Schlievert P, Gruss A, Novick RP (1986) Regulation of exoprotein gene expression in Staphylococcus aureus by agar. *Molecular & general genetics : MGG* **202:** 58-61

Regassa LB, Novick RP, Betley MJ (1992) Glucose and nonmaintained pH decrease expression of the accessory gene regulator (agr) in Staphylococcus aureus. *Infection and immunity* **60:** 3381-3388

Reyes D, Andrey DO, Monod A, Kelley WL, Zhang G, Cheung AL (2011) Coordinated regulation by AgrA, SarA, and SarR to control agr expression in Staphylococcus aureus. *Journal of bacteriology* **193:** 6020-6031

Reynolds J, Wigneshweraraj S (2011) Molecular insights into the control of transcription initiation at the Staphylococcus aureus agr operon. *Journal of molecular biology* **412:** 862-881

Risoen PA, Havarstein LS, Diep DB, Nes IF (1998) Identification of the DNA-binding sites for two response regulators involved in control of bacteriocin synthesis in Lactobacillus plantarum C11. *Molecular & general genetics : MGG* **259:** 224-232

Rogasch K, Ruhmling V, Pane-Farre J, Hoper D, Weinberg C, Fuchs S, Schmudde M, Broker BM, Wolz C, Hecker M, Engelmann S (2006) Influence of the two-component system SaeRS on global gene expression in two different Staphylococcus aureus strains. *Journal of bacteriology* **188:** 7742-7758

Rosenbach FJ (1884) *Mikro-organismen bei den Wund-Infections-Krankheiten des Menschen*, Wiesbaden, Germany.

Ross W, Gosink KK, Salomon J, Igarashi K, Zou C, Ishihama A, Severinov K, Gourse RL (1993) A third recognition element in bacterial promoters: DNA-binding by the alpha subunit of RNA polymerase. *Science* **262:** 1407-1413

Ross W, Schneider DA, Paul BJ, Mertens A, Gourse RL (2003) An intersubunit contact stimulating transcription initiation by E coli RNA polymerase: interaction of the alpha C-terminal domain and sigma region 4. *Genes & development* **17:** 1293-1307

Rossi F, Diaz L, Wollam A, Panesso D, Zhou Y, Rincon S, Narechania A, Xing G, Di Gioia TS, Doi A, Tran TT, Reyes J, Munita JM, Carvajal LP, Hernandez-Roldan A, Brandao D, van der Heijden IM, Murray BE, Planet PJ, Weinstock GM, Arias CA (2014) Transferable vancomycin resistance in a community-associated MRSA lineage. *The New England journal of medicine* **370:** 1524-1531

Roux A, Todd DA, Velazquez JV, Cech NB, Sonenshein AL (2014) CodY-mediated regulation of the Staphylococcus aureus Agr system integrates nutritional and population density signals. *Journal of bacteriology* **196:** 1184-1196

Ryckaert JPC, G.; Berendsen, H. J. C. (1977) Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. *Journal of Computational Physics* **23:** 327-341

Saecker RM, Record MT, Jr., Dehaseth PL (2011) Mechanism of bacterial transcription initiation: RNA polymerase - promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis. *Journal of molecular biology* **412:** 754-771

Said-Salim B, Dunman PM, McAleese FM, Macapagal D, Murphy E, McNamara PJ, Arvidson S, Foster TJ, Projan SJ, Kreiswirth BN (2003) Global regulation of Staphylococcus aureus genes by Rot. *Journal of bacteriology* **185:** 610-619

Salomon-Ferrer RC, D. A.; Walker, R. C. (2013) An overview of the Amber biomolecular simulation package *WIREs Computational Molecular Science* **3:** 195-210

Sanders DA, Gillece-Castro BL, Stock AM, Burlingame AL, Koshland DE, Jr. (1989) Identification of the site of phosphorylation of the chemotaxis response regulator protein, CheY. *The Journal of biological chemistry* **264:** 21770-21778

Schmidt KA, Manna AC, Gill S, Cheung AL (2001) SarT, a repressor of alpha-hemolysin in Staphylococcus aureus. *Infection and immunity* **69:** 4749-4758

Seidl K, Stucki M, Ruegg M, Goerke C, Wolz C, Harris L, Berger-Bachi B, Bischoff M (2006) Staphylococcus aureus CcpA affects virulence determinant production and antibiotic resistance. *Antimicrobial agents and chemotherapy* **50:** 1183-1194

Sharma-Kuinkel BK, Mann EE, Ahn JS, Kuechenmeister LJ, Dunman PM, Bayles KW (2009) The Staphylococcus aureus LytSR two-component regulatory system affects biofilm formation. *Journal of bacteriology* **191:** 4767-4775

Shaw LN, Lindholm C, Prajsnar TK, Miller HK, Brown MC, Golonka E, Stewart GC, Tarkowski A, Potempa J (2008) Identification and characterization of sigma, a novel component of the Staphylococcus aureus stress and virulence responses. *PloS one* **3:** e3844

Shimizu T, Ba-Thein W, Tamaki M, Hayashi H (1994) The virR gene, a member of a class of two-component response regulators, regulates the production of perfringolysin O, collagenase, and hemagglutinin in Clostridium perfringens. *Journal of bacteriology* **176:** 1616-1623

Shoji M, Cui L, Iizuka R, Komoto A, Neoh HM, Watanabe Y, Hishinuma T, Hiramatsu K (2011) walK and clpP mutations confer reduced vancomycin susceptibility in Staphylococcus aureus. *Antimicrobial agents and chemotherapy* **55:** 3870-3881

Shopsin B, Drlica-Wagner A, Mathema B, Adhikari RP, Kreiswirth BN, Novick RP (2008) Prevalence of agr dysfunction among colonizing Staphylococcus aureus strains. *The Journal of infectious diseases* **198:** 1171-1174

Shopsin B, Eaton C, Wasserman GA, Mathema B, Adhikari RP, Agolory S, Altman DR, Holzman RS, Kreiswirth BN, Novick RP (2010) Mutations in agr do not persist in natural populations of methicillin-resistant Staphylococcus aureus. *The Journal of infectious diseases* **202:** 1593-1599

Sidote DJ, Barbieri CM, Wu T, Stock AM (2008) Structure of the Staphylococcus aureus AgrA LytTR domain bound to DNA reveals a beta fold with an unusual mode of binding. *Structure* **16:** 727-735

Siebenlist U (1979) Nucleotide sequence of the three major early promoters of bacteriophage T7. *Nucleic acids research* **6:** 1895-1907

Siebenlist U, Simpson RB, Gilbert W (1980) E. coli RNA polymerase interacts homologously with two different promoters. *Cell* **20:** 269-281

Siegele DA, Hu JC, Walter WA, Gross CA (1989) Altered promoter recognition by mutant forms of the sigma 70 subunit of Escherichia coli RNA polymerase. *Journal of molecular biology* **206:** 591-603

Singh R, Ray P (2014) Quorum sensing-mediated regulation of staphylococcal virulence and antibiotic resistance. *Future microbiology* **9:** 669-681

Skerker JM, Prasol MS, Perchuk BS, Biondi EG, Laub MT (2005) Two-component signal transduction pathways regulating growth and cell cycle progression in a bacterium: a system-level analysis. *PLoS biology* **3:** e334

Srivastava SK, Rajasree K, Fasim A, Arakere G, Gopal B (2014) Influence of the AgrC-AgrA complex on the response time of Staphylococcus aureus quorum sensing. *Journal of bacteriology* **196:** 2876-2888

Stephens SK, Floriano B, Cathcart DP, Bayley SA, Witt VF, Jimenez-Diaz R, Warner PJ, Ruiz-Barba JL (1998) Molecular analysis of the locus responsible for production of plantaricin S, a

two-peptide bacteriocin produced by Lactobacillus plantarum LPCO10. *Applied and environmental microbiology* **64:** 1871-1877

Sterba KM, Mackintosh SG, Blevins JS, Hurlburt BK, Smeltzer MS (2003) Characterization of Staphylococcus aureus SarA binding sites. *Journal of bacteriology* **185:** 4410-4417

Stock AM, Martinez-Hackert E, Rasmussen BF, West AH, Stock JB, Ringe D, Petsko GA (1993) Structure of the Mg(2+)-bound form of CheY and mechanism of phosphoryl transfer in bacterial chemotaxis. *Biochemistry* **32:** 13375-13380

Stock AM, Robinson VL, Goudreau PN (2000) Two-component signal transduction. *Annual review of biochemistry* **69:** 183-215

Studholme DJ, Buck M (2000) The biology of enhancer-dependent transcriptional regulation in bacteria: insights from genome sequences. *FEMS microbiology letters* **186:** 1-9

Sully EK, Malachowa N, Elmore BO, Alexander SM, Femling JK, Gray BM, DeLeo FR, Otto M, Cheung AL, Edwards BS, Sklar LA, Horswill AR, Hall PR, Gresham HD (2014) Selective chemical inhibition of agr quorum sensing in Staphylococcus aureus promotes host defense with minimal impact on resistance. *PLoS pathogens* **10:** e1004174

Sun F, Ji Q, Jones MB, Deng X, Liang H, Frank B, Telser J, Peterson SN, Bae T, He C (2012a) AirSR, a [2Fe-2S] cluster-containing two-component system, mediates global oxygen sensing and redox signaling in Staphylococcus aureus. *Journal of the American Chemical Society* **134:** 305-314

Sun F, Liang H, Kong X, Xie S, Cho H, Deng X, Ji Q, Zhang H, Alvarez S, Hicks LM, Bae T, Luo C, Jiang H, He C (2012b) Quorum-sensing agr mediates bacterial oxidation response via an intramolecular disulfide redox switch in the response regulator AgrA. *Proceedings of the National Academy of Sciences of the United States of America* **109:** 9095-9100

Sun H, Yang Y, Xue T, Sun B (2013) Modulation of cell wall synthesis and susceptibility to vancomycin by the two-component system AirSR in Staphylococcus aureus NCTC8325. *BMC microbiology* **13:** 286

Sun J, Zheng L, Landwehr C, Yang J, Ji Y (2005) Identification of a novel essential two-component signal transduction system, YhcSR, in Staphylococcus aureus. *Journal of bacteriology* **187:** 7876-7880

Tamber S, Reyes D, Donegan NP, Schwartzman JD, Cheung AL, Memmi G (2010) The staphylococcus-specific gene rsr represses agr and virulence in Staphylococcus aureus. *Infection and immunity* **78:** 4384-4391

Tebbutt J, Rhodius VA, Webster CL, Busby SJ (2002) Architectural requirements for optimal activation by tandem CRP molecules at a class I CRP-dependent promoter. *FEMS microbiology letters* **210:** 55-60

Thoendel M, Horswill AR (2009) Identification of Staphylococcus aureus AgrD residues required for autoinducing peptide biosynthesis. *The Journal of biological chemistry* **284:** 21828-21838

Thoendel M, Horswill AR (2013) Random mutagenesis and topology analysis of the autoinducing peptide biosynthesis proteins in Staphylococcus aureus. *Molecular microbiology* **87:** 318-337

Thoendel M, Kavanaugh JS, Flack CE, Horswill AR (2011) Peptide signaling in the staphylococci. *Chemical reviews* **111:** 117-151

Thompson JF, Landy A (1988) Empirical estimation of protein-induced DNA bending angles: applications to lambda site-specific recombination complexes. *Nucleic acids research* **16:** 9687-9705

Thomsen LE, Gottlieb CT, Gottschalk S, Wodskou TT, Kristensen HH, Gram L, Ingmer H (2010) The heme sensing response regulator HssR in Staphylococcus aureus but not the homologous RR23 in Listeria monocytogenes modulates susceptibility to the antimicrobial peptide plectasin. *BMC microbiology* **10:** 307

Thurlow LR, Joshi GS, Richardson AR (2012) Virulence strategies of the dominant USA300 lineage of community-associated methicillin-resistant Staphylococcus aureus (CA-MRSA). *FEMS immunology and medical microbiology* **65:** 5-22

Toledo-Arana A, Merino N, Vergara-Irigaray M, Debarbouille M, Penades JR, Lasa I (2005) Staphylococcus aureus develops an alternative, ica-independent biofilm in the absence of the arlRS two-component system. *Journal of bacteriology* **187:** 5318-5329

Torres VJ, Stauff DL, Pishchany G, Bezbradica JS, Gordy LE, Iturregui J, Anderson KL, Dunman PM, Joyce S, Skaar EP (2007) A Staphylococcus aureus regulatory system that responds to host heme and modulates virulence. *Cell host & microbe* **1:** 109-119

Traber K, Novick R (2006) A slipped-mispairing mutation in AgrA of laboratory strains and clinical isolates results in delayed activation of agr and failure to translate delta- and alpha-haemolysins. *Molecular microbiology* **59:** 1519-1530

Traber KE, Lee E, Benson S, Corrigan R, Cantera M, Shopsin B, Novick RP (2008) agr function in clinical Staphylococcus aureus isolates. *Microbiology* **154:** 2265-2274

Tsompanidou E, Sibbald MJ, Chlebowicz MA, Dreisbach A, Back JW, van Dijl JM, Buist G, Denham EL (2011) Requirement of the agr locus for colony spreading of Staphylococcus aureus. *Journal of bacteriology* **193:** 1267-1272

Ulrich M, Bastian M, Cramton SE, Ziegler K, Pragman AA, Bragonzi A, Memmi G, Wolz C, Schlievert PM, Cheung A, Doring G (2007) The staphylococcal respiratory response regulator SrrAB induces ica gene transcription and polysaccharide intercellular adhesin expression,

protecting Staphylococcus aureus from neutrophil killing under anaerobic growth conditions. *Molecular microbiology* **65:** 1276-1287

Ushida C, Aiba H (1990) Helical phase dependent action of CRP: effect of the distance between the CRP site and the -35 region on promoter activity. *Nucleic acids research* **18:** 6325-6330

van Wamel WJ, Rooijakkers SH, Ruyken M, van Kessel KP, van Strijp JA (2006) The innate immune modulators staphylococcal complement inhibitor and chemotaxis inhibitory protein of Staphylococcus aureus are located on beta-hemolysin-converting bacteriophages. *Journal of bacteriology* **188:** 1310-1315

Vandenesch F, Lina G, Henry T (2012) Staphylococcus aureus hemolysins, bi-component leukocidins, and cytolytic peptides: a redundant arsenal of membrane-damaging virulence factors? *Frontiers in cellular and infection microbiology* **2:** 12

Vaughan A, Eijsink VG, Van Sinderen D (2003) Functional characterization of a composite bacteriocin locus from malt isolate Lactobacillus sakei 5. *Applied and environmental microbiology* **69:** 7194-7203

Verdon J, Girardin N, Lacombe C, Berjeaud JM, Hechard Y (2009) delta-hemolysin, an update on a membrane-interacting peptide. *Peptides* **30:** 817-823

Vuong C, Saenz HL, Gotz F, Otto M (2000) Impact of the agr quorum-sensing system on adherence to polystyrene in Staphylococcus aureus. *The Journal of infectious diseases* **182:** 1688-1693

Wang B, Zhao A, Novick RP, Muir TW (2014) Activation and inhibition of the receptor histidine kinase AgrC occurs through opposite helical transduction motions. *Molecular cell* **53:** 929-940

Ween O, Gaustad P, Havarstein LS (1999) Identification of DNA-binding sites for ComE, a key regulator of natural competence in Streptococcus pneumoniae. *Molecular microbiology* **33:** 817-827

Weiss A, Ibarra JA, Paoletti J, Carroll RK, Shaw LN (2014) The delta subunit of RNA polymerase guides promoter selectivity and virulence in Staphylococcus aureus. *Infection and immunity* **82:** 1424-1435

Werner F (2007) Structure and function of archaeal RNA polymerases. *Molecular microbiology* **65:** 1395-1404

Wertheim HF, Melles DC, Vos MC, van Leeuwen W, van Belkum A, Verbrugh HA, Nouwen JL (2005) The role of nasal carriage in Staphylococcus aureus infections. *The Lancet Infectious diseases* **5:** 751-762

Wigneshweraraj S, Bose D, Burrows PC, Joly N, Schumacher J, Rappas M, Pape T, Zhang X, Stockley P, Severinov K, Buck M (2008) Modus operandi of the bacterial RNA polymerase containing the sigma54 promoter-specificity factor. *Molecular microbiology* **68:** 538-546

Wing HJ, Green J, Guest JR, Busby SJ (2000) Role of activating region 1 of Escherichia coli FNR protein in transcription activation at class II promoters. *The Journal of biological chemistry* **275:** 29061-29065

Wing HJ, Williams SM, Busby SJ (1995) Spacing requirements for transcription activation by Escherichia coli FNR protein. *Journal of bacteriology* **177:** 6704-6710

Wolfe AJ (2005) The acetate switch. *Microbiology and molecular biology reviews : MMBR* **69:** 12-50

Woodin AM (1959) Fractionation of a leucocidin from Staphylococcus aureus. *The Biochemical journal* **73:** 225-237

Wright JS, 3rd, Lyon GJ, George EA, Muir TW, Novick RP (2004) Hydrophobic interactions drive ligand-receptor recognition for activation and inhibition of staphylococcal quorum sensing. *Proceedings of the National Academy of Sciences of the United States of America* **101:** 16168-16173

Xue T, You Y, Hong D, Sun H, Sun B (2011) The Staphylococcus aureus KdpDE two-component system couples extracellular K+ sensing and Agr signaling to infection programming. *Infection and immunity* **79:** 2154-2167

Yan M, Yu C, Yang J, Ji Y (2011) The essential two-component system YhcSR is involved in regulation of the nitrate respiratory pathway of Staphylococcus aureus. *Journal of bacteriology* **193:** 1799-1805

Yang SJ, Xiong YQ, Yeaman MR, Bayles KW, Abdelhady W, Bayer AS (2013) Role of the LytSR two-component regulatory system in adaptation to cationic antimicrobial peptides in Staphylococcus aureus. *Antimicrobial agents and chemotherapy* **57:** 3875-3882

Yarwood JM, Bartels DJ, Volper EM, Greenberg EP (2004) Quorum sensing in Staphylococcus aureus biofilms. *Journal of bacteriology* **186:** 1838-1850

Yarwood JM, McCormick JK, Schlievert PM (2001) Identification of a novel two-component regulatory system that acts in global regulation of virulence factors of Staphylococcus aureus. *Journal of bacteriology* **183:** 1113-1123

Ye ZH, Lee CY (1989) Nucleotide sequence and genetic characterization of staphylococcal bacteriophage L54a int and xis genes. *Journal of bacteriology* **171:** 4146-4153

Zhang L, Gray L, Novick RP, Ji G (2002) Transmembrane topology of AgrB, the protein involved in the post-translational modification of AgrD in Staphylococcus aureus. *The Journal of biological chemistry* **277:** 34736-34742

Zhang L, Ji G (2004) Identification of a staphylococcal AgrB segment(s) responsible for group-specific processing of AgrD by gene swapping. *Journal of bacteriology* **186:** 6706-6713

Zhang L, Lin J, Ji G (2004) Membrane anchoring of the AgrD N-terminal amphipathic region is required for its processing to produce a quorum-sensing pheromone in Staphylococcus aureus. *The Journal of biological chemistry* **279:** 19448-19456

Zhang Y, Feng Y, Chatterjee S, Tuske S, Ho MX, Arnold E, Ebright RH (2012) Structural basis of transcription initiation. *Science* **338:** 1076-1080

Zhao L, Xue T, Shang F, Sun H, Sun B (2010) Staphylococcus aureus AI-2 quorum sensing associates with the KdpDE two-component system to regulate capsular polysaccharide synthesis and virulence. *Infection and immunity* **78:** 3506-3515

Zhu X, Amsler CD, Volz K, Matsumura P (1996) Tyrosine 106 of CheY plays an important role in chemotaxis signal transduction in Escherichia coli. *Journal of bacteriology* **178:** 4208-4215

PUBLICATION

Published online 28 October 2014 *Nucleic Acids Research, 2014, Vol. 42, No. 20* **12523–12536**
doi: 10.1093/nar/gku1015

Systematic mutational analysis of the LytTR DNA binding domain of *Staphylococcus aureus* virulence gene transcription factor AgrA

Sophie S. Nicod[1], Robert O. J. Weinzierl[2], Lynn Burchell[1], Andres Escalera-Maurer[2], Ellen H. James[1] and Sivaramesh Wigneshweraraj[1,*]

[1]MRC Centre for Molecular Microbiology and Infection, Imperial College London, London, UK and [2]Department of Life Sciences, Imperial College London, London, UK

Received August 26, 2014; Accepted October 07, 2014

Downloaded from http://nar.oxfordjournals.org/ at Imperial College London on March 17, 2015

ABSTRACT

Most DNA-binding bacterial transcription factors contact DNA through a recognition α-helix in their DNA-binding domains. An emerging class of DNA-binding transcription factors, predominantly found in pathogenic bacteria interact with the DNA via a relatively novel type of DNA-binding domain, called the LytTR domain, which mainly comprises β strands. Even though the crystal structure of the LytTR domain of the virulence gene transcription factor AgrA from *Staphylococcus aureus* bound to its cognate DNA sequence is available, the contribution of specifi amino acid residues in the LytTR domain of AgrA to transcription activation remains elusive. Here, for the fi st time, we have systematically investigated the role of amino acid residues in transcription activation in a LytTR domain-containing transcription factor. Our analysis, which involves *in vivo* and *in vitro* analyses and molecular dynamics simulations of *S. aureus* AgrA identifie a highly conserved tyrosine residue, Y229, as a major amino acid determinant for maximal activation of transcription by AgrA and provides novel insights into structure–function relationships in *S. aureus* AgrA.

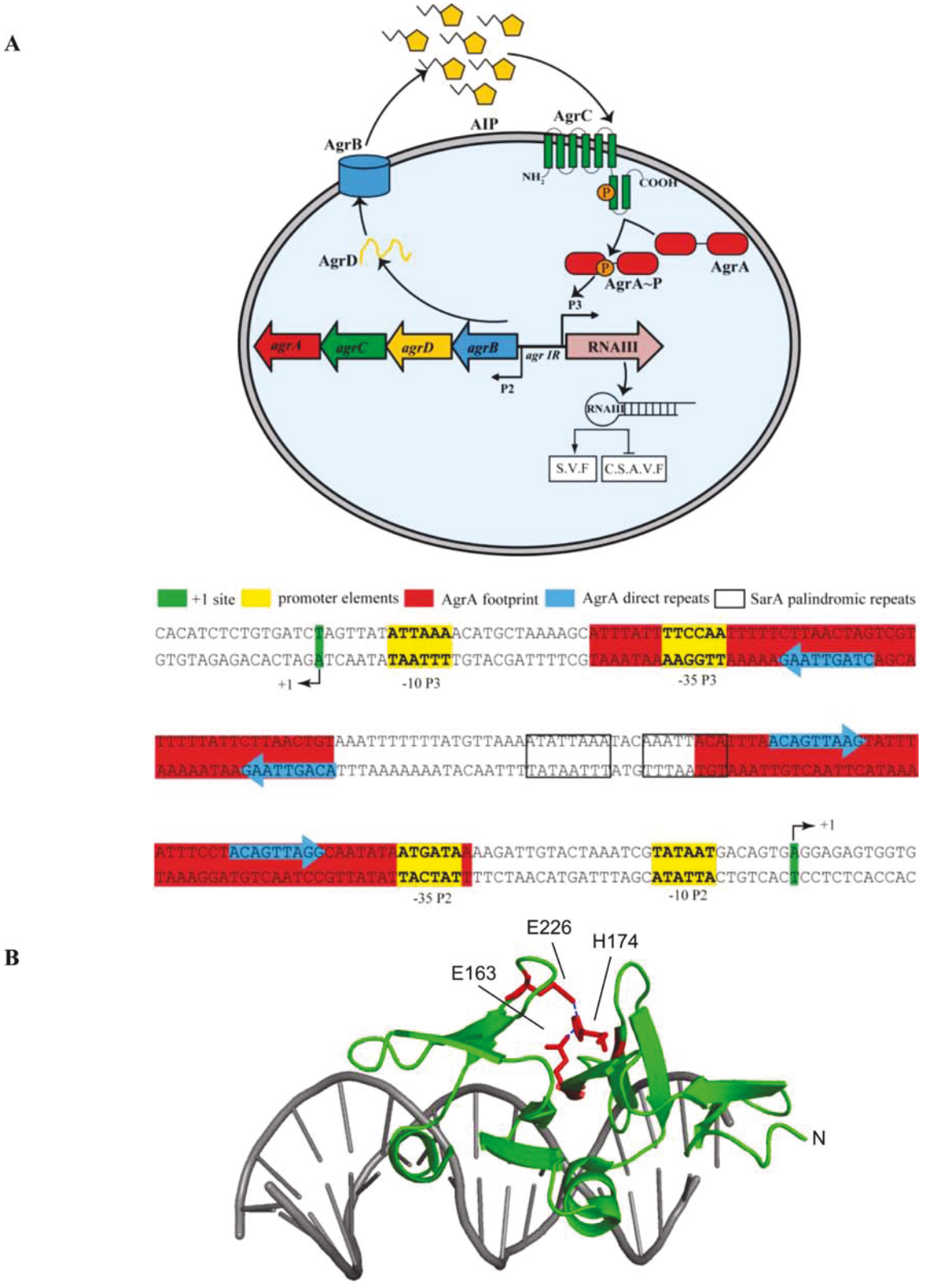
INTRODUCTION

Bacteria predominantly rely on two-component signal transduction systems (TCS) to sense and adapt gene expression patterns to constantly changing environments. The typical bacterial TCS comprises signal input and response output components, typically represented by a histidine kinase (HK) and a response regulator (RR), respectively. In response to a signal, the HK becomes phosphorylated and subsequently transphosphorylates its cognate RR, thereby activating the RR to elicit the output response. Most RRs contain two domains: a conserved amino-terminal regulatory domain and a variable carboxyl-terminal effector domain. The majority of RRs are transcription factors (hereafter referred to as RR-TF) with their carboxyl-terminal domain containing a DNA-binding motif, which allows recognition of tandem or inverted repeating DNA elements located upstream of promoters of genes and determinants for interaction with the RNA polymerase for modulating the transcriptional response. The carboxyl-terminal domain of the majority of RR-TFs contains DNA-binding motifs that belong to extensively characterized structural families and include the winged-helix motif of the OmpR/PhoB family, helix-turn-helix motif of the NtrC family and four-helix-helix-turn-helix of the NarL/FixJ family (1,2). In contrast, a small number of RR-TFs, mostly in γ-proteobacteria and firmicutes, have an unusual and relatively poorly characterized carboxyl-terminal domain, called the LytTR domain, which predominantly comprises β strands (3,4). Intriguingly, LytTR domain containing RR-TFs are disproportionately involved in the regulation of virulence gene expression. Hence, the role of LytTR domain containing RR-TFs has been widely studied in many bacterial pathogens (5–7). The molecular interactions between the LytTR domain and DNA have been recently elucidated from the crystal structure of the DNA-bound complex of the LytTR domain of *Staphylococcus aureus* accessory gene regulator A (AgrA), which is a cell-density (quorum) responsive global virulence-associated RR-TF (4,8).

The quorum-dependent regulation of AgrA activity in *S. aureus* is well understood (Figure 1 A): the binding of the quorum signal (the autoinducing peptide, a peptide thiolactone, product of the *agrD* gene; (8)) to the HK AgrC results in the autophosphorylation of AgrC. Phosphorylation of AgrA by AgrC 'activates' AgrA thus allowing AgrA to bind to specific sites in the intergenic region of the *agr* operon and activate transcription from divergent promoters, P2 and P3. The P2 transcript contains the *agrBDCA*

*To whom correspondence should be addressed. Tel: +44 207 594 1867; Fax: +44 207 594 3055; Email: s.r.wig@imperial.ac.uk

© The Author(s) 2014. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.



Downloaded from http://nar.oxfordjournals.org/ at Imperial College London on March 17, 2015

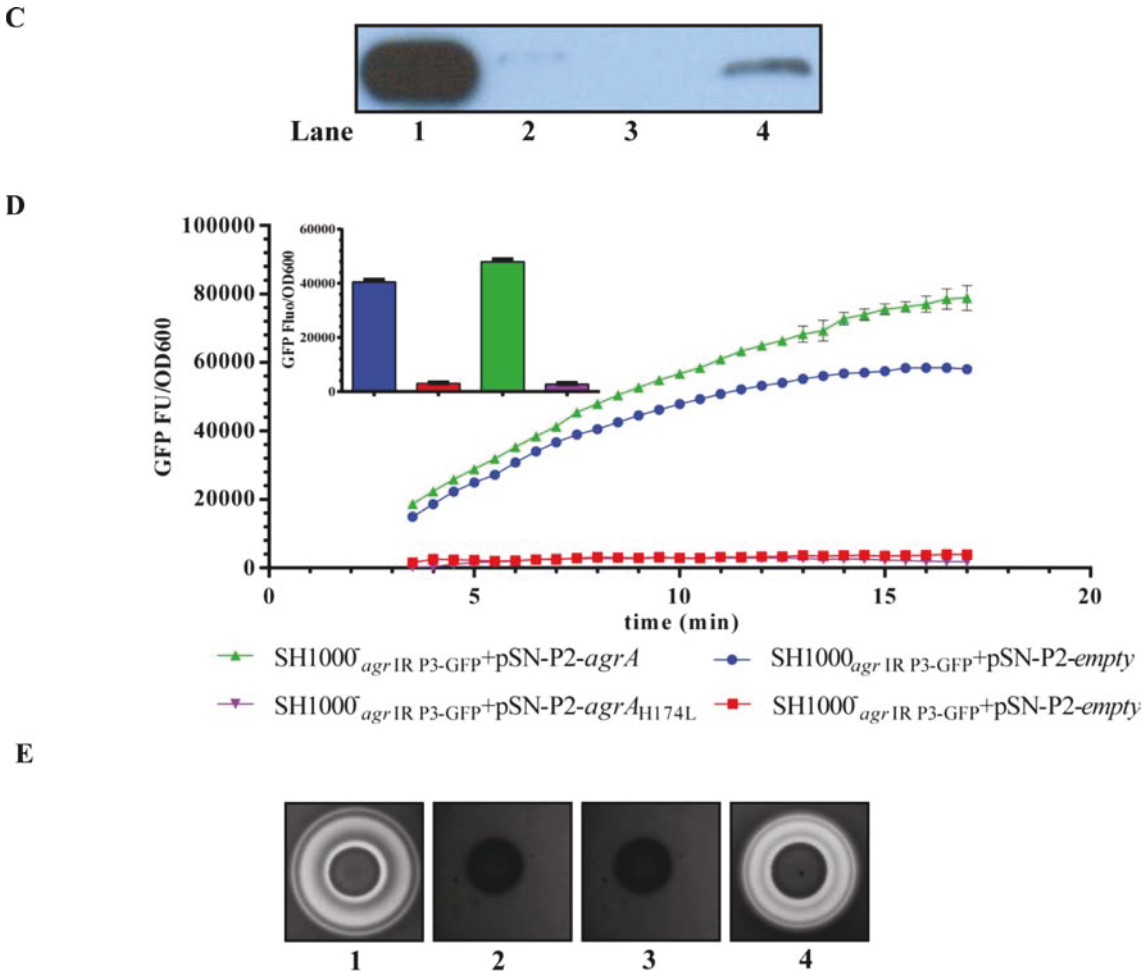


Figure 1. Establishing an experimental system to assess AgrA activity in *Staphylococcus aureus*. (**A**) Schematic representation of the *agr* operon organization and regulation in *S. aureus*. Secreted virulence factors (SVF) and cell-surface-associated virulence factors (CSAVF). The nucleotide sequence of the *S. aureus agr* operon intergenic region with the different regions of relevance to this study indicated. (**B**) Ribbon representation of the AgrA LytTR domain–DNA complex crystal structure (the LytTR domain and DNA are colored in green and gray, respectively). H174, E163 and E226 are highlighted in red and the salt bridge interactions indicated by dotted lines. (**C**) Western blot indicating AgrA protein level in whole cell lysate of *S. aureus* SH1000$^-$ + pSN-P2-*agrA* (lane 1), SH1000$^-$ + pSN-P2-*empty* (lane 2), SH1001 + pSN-P2-*empty* (lane 3) and SH1000 + pSN-P2-*empty* (lane 4) strains grown overnight in TSB. (**D**) Graph showing GFP expression [as GFP fluorescence units (GFP-FU)] as a function of growth (OD600) over time for *S. aureus* SH1000$^-_{\text{agr IR P3-GFP}}$ + pSN-P2-*agrA* (green), SH1000$^-_{\text{agr IR P3-GFP}}$ + pSN-P2-*empty* (red), SH1000$^-_{\text{agr IR P3-GFP}}$ + pSN-P2-*agrA*$_{\text{H174L}}$ (purple) and SH1000 $_{\text{agr IR P3-GFP}}$ + pSN-P2-*empty* (blue) strains grown in TSB. The bar chart in the insert represents the GFP expression of the same samples (color coded accordingly) at the 8 h time point. (**E**) Sheep blood agar hemolysis assay with SH1000$^-_{\text{agr IR P3-GFP}}$ + pSN-P2-*agrA* (panel 1), SH1000$^-_{\text{agr IR P3-GFP}}$ + pSN-P2-*agrA*$_{\text{H174L}}$ (panel 2), SH1000$^-_{\text{agr IR P3-GFP}}$ + pSN-P2-*empty* (panel 3) and SH1000$_{\text{agr IR P3-GFP}}$ + pSN-P2-*empty* (panel 4). Data for (C-E) were obtained from at least three biological replicates and for (C) and (E) representative results are shown.

genes; the P3 transcript (RNAIII) is the pleiotropic effector molecule of the *agr* response (8). RNAIII is directly responsible for post-transcriptional regulation of multiple virulence factors such as α- and δ-hemolysin (which is encoded within the RNAIII transcript) (8,9). AgrA also directly activates transcription of genes *psm*α, *psm*β and *psm-mec* that encode phenol-soluble modulins (PSMs) (which play a key role in immune evasion by *S. aureus* owing to their leukocidal activity) (10). Furthermore, the LytTR domain of AgrA also serves as a redox sensor and controls the expression of the *bsaA* gene that encodes the glutathione peroxidase, which allows bacterial cells to cope with oxidative stress (11). In this scenario, oxidative stress results in the formation of an intramolecular disulfide bond between aa C199 and C228 in the LytTR domain, which causes the dissociation of AgrA from its cognate DNA sequence located immediately downstream of the *bsaA* promoter and thereby leading to the derepression of *bsaA* transcription. AgrA binds as a dimer to two direct repeats located in the intergenic region of the *agr* operon immediately upstream of the P2 and P3 promoters, with the center of the promoter proximal direct repeats located approximately at position -60 with respect to the transcription start sites at +1 for P2 and P3 (Figure 1A) (12). The *agr* operon intergenic region also contains binding sites for several other global TFs, notably SarA, which binds as a dimer to sites located between the direct repeats bound by AgrA (Figure 1A) (13). It seems that SarA and AgrA co-activate transcription from the P2 promoter (14,15). Therefore, based on the proximity of both AgrA-binding sites to the P2 and P3 promoters, re-

Downloaded from http://nar.oxfordjournals.org/ at Imperial College London on March 17, 2015

spectively, it is possible that transcription activation at P2 and P3 involves protein–protein interaction between AgrA, SarA and the RNA polymerase (RNAp). Further, the -35 and -10 consensus promoter elements of all known AgrA-activated promoters in *S. aureus* have a suboptimal spacer length. For instance, the -35 and -10 consensus promoter elements of the P2 and P3 promoters of the *agr* operon are separated by 18 and 20 nucleotides, respectively, instead of the optimal 17 nucleotides long spacer region found at most promoters. Previous studies have suggested that transcription activation by AgrA involves a DNA-bending step (14). Such a mechanism could compensate for the suboptimal spacer length at the P2 and P3 promoters. Consistent with this view, shortening of the P2 and P3 promoter spacer region to the optimal 17 nucleotides length leads to increased AgrA-independent transcription from both promoters (16).

To date, only amino acid (aa) residues required for binding of LytTR domain containing TFs to DNA have been identified (4,5), but the aa residues in the LytTR domain involved in transcription activation are not known. In this study, we strategically targeted 74 of 101 aa residues in the LytTR domain (aa residues 137–238) of *S. aureus* AgrA for substitution with alanine and studied their contribution to transcription activation at the *agr* operon *in vivo* and *in vitro*. Our results, which represent the first comprehensive mutational analysis of an LytTR domain containing TF, are discussed in the context of structure–function relationships and biology of the *S. aureus* AgrA.

MATERIALS AND METHODS

Bacterial strains, plasmids and DNA manipulation

The bacterial strains and plasmids used in this study are listed in Supplementary Table S1 and sequences of primers used for DNA manipulation and cloning are listed in Supplementary Table S2. *Escherichia coli* and *S. aureus* were grown in Luria broth (LB) and tryptic soy broth (TSB), respectively. Where appropriate, antibiotics were added to the growth media at the following concentrations: ampicillin, 100 μg/ml (*E. coli*); kanamycin, 90 μg/ml (*S. aureus*), 50 μg/ml (*E. coli*) and chloramphenicol, 7.5 μg/ml (*S. aureus*). Anhydrotetracycline was added at a concentration of 200 ng/ml where appropriate. *Escherichia coli* XL1-blue, DC10B and BL21 (DE3) cells were transformed with plasmid DNA using the standard heat-shock transformation protocol. *Staphylococcus aureus* RN4220, SH1000, SH1000^{-} and SH1001 cells were transformed by electroporation with a Gene Pulser (Bio-Rad) with settings: 2.5 kV, 0.25 μFD, 100 Ω in a 1 mm cuvette. Following this, cells were recovered in brain-heart infusion (BHI) supplemented with 0.5 M sucrose for 1 h before being plated onto the appropriate selection plates. Transduction into target strains of *S. aureus* was carried out with φ11 following published protocols (17). All restriction enzymes and DNA modification enzymes were purchased from New England BioLabs. Plasmids and genomic DNA were extracted with the Qiagen plasmid miniprep kit and the Wizard bacterial genomic DNA kit (Promega), respectively, according to the manufacturers' instructions. The SH1000$^{-}{}_{\text{agr IR P3-GFP}}$ strain was constructed exactly as previously described by phage transduction of the reporter construct from SH1000$_{\text{agr IR P3-GFP}}$ (18). Plasmid pSN-P2-*agrA* was constructed by PCR amplifying the P2 promoter region from the pCL55$_{\textit{agr}\text{ IR P2-GFP}}$ (18) plasmid containing a region at the 3′ end homologous to the 5′ end of the *agrA* gene of *S. aureus* SH1000 strain and containing the RBS using primers P2-RBS F and P2-RBS R. The *agrA* gene was amplified from the chromosome of *S. aureus* SH1000 using primers P2-AgrA F and P2-AgrA R. The transcription terminator was amplified from the pCN44 (19) plasmid containing a region at the 5′ end homologous to the 3′ end of *agrA* using primers P2-TT F and P2-TT R. The three fragments were fused together by PCR using primers P2-RBS F, P2-RBS R and P2-TT R. The 1226 bp long *BamH*I and *Kpn*I digested PCR fragment was then ligated into the *BamH*I and *Kpn*I sites of pCN34 (19) plasmid. Plasmid pSN-*itet-agrA* was constructed by PCR amplification of the *agrA* gene and terminator region from the pSN-P2-*agrA* plasmid using primers AgrA pCN34itet F containing the RBS and AgrA pCN34itet R. The 1059 bp long fragment was digested with *Kpn*I and *Eco*RI and was then ligated into the *Kpn*I and *Eco*RI sites of the pCN34*itet* plasmid (20). Plasmid pSN-agrA was constructed by PCR amplifying *agrA* from the chromosome of *S. aureus* SH1000 using primers AgrA pTYB2 F and AgrA pTYB2 F and ligating the 731 bp long *Nde*I and *Sma*I digested PCR fragment into the *Nde*I and *Sma*I sites of pTYB2.

Purification of proteins

Staphylococcus aureus core RNAp and σ^{A} (from pJR28-[6His]*rpoD*) were prepared exactly as previously described by Reynolds & Wigneshweraraj (16). Recombinant AgrA was purified as follows: *Escherichia coli* strain ER2566 containing pSN-agrA was grown at 37°C. At $OD_{600\text{ nm}}$ ∼0.6, the cells were temperature shifted to 16°C for 30 min, and the expression of AgrA was induced with 0.25 mM IPTG. The cells were harvested after 17 h at 16°C. AgrA was purified using the IMPACT^{TM} kit (New England Biolabs) according to the manufacturer's instructions. Briefly, the cells were lysed in a buffer containing 20 mM Tris-HCl (pH9), 1 M NaCl and 1 mM EDTA (column buffer) and centrifuged to remove cellular debris. The supernatant was then loaded on a 10 ml gravity flow column (Bio-Rad) packed with 2 ml Chitin Resin (New England Biolabs). The column was washed with 20 bed volumes of column buffer and the protein was cleaved from the intein tag after incubation for 16 h at 4°C in three bed volumes of cleavage buffer [column buffer + 200 mM DTT]. The protein was concentrated using Amicon Ultracel-10K (Millipore) and dialysed in a storage buffer [10 mM Tris-HCl (pH8), 50 mM NaCl, 0.1 mM EDTA, 1 mM DTT and 20% (vol/vol) glycerol].

Bacterial growth and GFP expression assays

These were conducted exactly as described previously (18). Briefly, simultaneous growth ($OD_{600\text{ nm}}$) and GFP fluorescence measurements (with excitation and emission filters of 485 and 520 nm, respectively) were performed in 96-well black microtiter plates with clear bottoms (Corning) in a POLARstar Omega multiwell plate reader (BMG Labtech). At least three biological replicates (each defined as a single colony) were performed for each growth curve.

Downloaded from http://nar.oxfordjournals.org/ at Imperial College London on March 17, 2015

Blood agar hemolysis assay

Bacteria were grown for 16 h in TSB culture, then 20 μl aliquots were plated onto Columbia Agar containing 5% sheep's blood and left to grow for 16 h at 37°C and then 24 h at 4°C prior to image capture of the plates.

Western blotting for AgrA

Western blotting was performed using polyclonal antibodies against AgrA (raised in rabbits against recombinant AgrA by Eurogentec (used at 1:250 dilution) and anti-rabbit-horseradish peroxidase-conjugated antibodies (Dako; used at 1:3000 dilution) as primary and secondary antibodies, respectively, following standard laboratory protocols and as described by James *et al*. (18).

Electrophoretic gel mobility shift assays (EMSA)

Phosphorylation of AgrA was carried out by pre-incubating 3 μM AgrA with 50 mM acetyl phosphate and 5 mM $MgCl_2$ for 1 h at 37°C. Ten microliter binding reactions were set up in a reaction buffer [10 mM HEPES (pH 7.6), 50 mM KCl, 1 mM EDTA, 2 mM DTT, 0.5% (v/v) TritonX-100 and 5% (v/v) glycerol] using 1 μM final concentration of phosphorylated AgrA and 10 nM of a 214 bp long γ-32P-labeled DNA fragment representing the *agr* operon intergenic region including the P2 and P3 core promoter sequences up to position +11 of the P2 promoter and +15 of the P3 promoter. The DNA probe for the EMSA was prepared by PCR amplification of the *agr* operon intergenic region from the *S. aureus* SH1000 chromosome using primers P2-IR-P3 F and P2-IR-P3 R. The phosphorylated AgrA was incubated in reaction buffer for 10 min at 37°C before adding the DNA probe. The reactions were incubated at 37°C for another 10 min and stopped with native gel loading dye [reaction buffer + 50% (v/v) glycerol and 0.05% (w/v) bromophenol blue] and resolved on a 4–20% (w/v) gradient native polyacrylamide gel. The dried gel was visualized and quantified using a GE Typhoon FLA 2000 PhosphorImager and the Image Quant TL software, respectively.

In vitro transcription assays

The binding reactions for the *in vitro* transcription assays were conducted as described above; however, pJR_{P2+P3} was used as promoter template in transcription buffer [40 mM Tris–acetate (pH 7.9), 100 mM NaCl, 20 mM $MgCl_2$ and 0.2 mM DTT]. The *in vitro* transcription assay was conducted as described by Reynolds & Wigneshweraraj (16). Briefly, 50 nM *S. aureus* core RNAp and 200 nM *S. aureus* σ^A, 1 μM AgrA (phosphorylated as described above) and 10 nM pJR_{P2+P3} were incubated at 37°C for 5 min separately and then mixed and incubated for another 5 min. RNA synthesis was initiated by adding an elongation mix containing 0.5 mM ATP, CTP and GTP; 0.25 mM UTP; 0.75 μCi of α-^{32}P UTP and the reaction was incubated for a further 10 min at 37°C. The reactions were stopped with stop dye [3% (w/v) xylene cyanol, 3% (w/vl) bromophenol blue and 20 mM EDTA in deionized formamide)] and resolved on a 10% (w/v) urea-denaturing polyacrylamide gel. The dried gel was visualized and quantified using a GE Typhoon FLA 2000 PhosphorImager and the Image Quant TL software, respectively.

DNA-bending assays

The DNA-bending assays were conducted exactly as described by Reyes *et al*. (14). Plasmid pAM1847 was digested with *Eco*RI, *Hind*III, *Bst*NI, *Eco*RV, *Nhe*I and *Bam*HI. The fragments were purified using the Qiagen PCR purification kit and individually labeled with γ-^{32}P ATP. Seven nanograms of the labeled fragments were then incubated with 1 μM of phosphorylated AgrA in a reaction buffer with 0.3 μg of sonicated calf thymus at room temperature for 20 min. The reactions were stopped with native loading dye and resolved on a 4.5% (w/v) native polyacrylamide gel. The dried gel was visualized and quantified using a GE Typhoon FLA 2000 PhosphorImager and the Image Quant TL software, respectively.

Molecular dynamics (MD) simulations

The protonation states of residues present in the LytTR DNA-binding domain in PDB4G4K (4) were computed using H++ (21). The resulting atomic coordinates were embedded in a TIP3P (22) cubic water box extending for a minimum of 15 Å beyond the protein structure. The assembly was charge-neutralized, adjusted to 150 mM NaCl and parameterized with the AMBER99SB (23) force field. Simulations were run on Nvidia GPUs certified to carry out accurately reproducible calculations using the AMBER 12 package (24). The parameterized structures were energetically minimized by rapid descent and repeated annealing minimization before equilibration at 300°Kelvin and one atmosphere of pressure (NPT ensemble). For all MD simulations, the bond length between hydrogen and heavy atoms was fixed using the SHAKE algorithm (25). The time-step was 2 femtoseconds and the Particle Mesh Ewald (26) cut-off distance was set to 8 Å. The simulation engine was pmemd.cuda, a CUDA-accelerated MD production engine based on a hybrid single-/double precision (SPFP) algorithm. Conventional MD simulations were initially carried out for 50 ns to obtain representative values for the total potential and dihedral energy values for adjustment in the subsequent dual-boost accelerated MD (aMD) production runs. For boosting the dihedral potential, an energy contribution of 3.5 kcal/mol/residue per degree of freedom was assumed. A boost factor (α) of 0.2 was used for adjustment of the dihedral and potential energy in all aMD simulations (the complete data sets are available from one of the authors [ROJW] upon request). All aMD simulations lasted 100 ns, but are likely due to the acceleration to represent protein motions occurring in the tens to hundreds of microseconds range in real time (27). Trajectories were processed, visualized and analyzed using the Visual Molecular Dynamics (VMD) suite (28).

RESULTS

Establishing an experimental system to asses AgrA activity in *S. aureus*

Mutant *S. aureus* strain SH1000$^-$ displays an *agr*-defective phenotype and previous work by Tsompanidou *et al*.

Downloaded from http://nar.oxfordjournals.org/ at Imperial College London on March 17, 2015

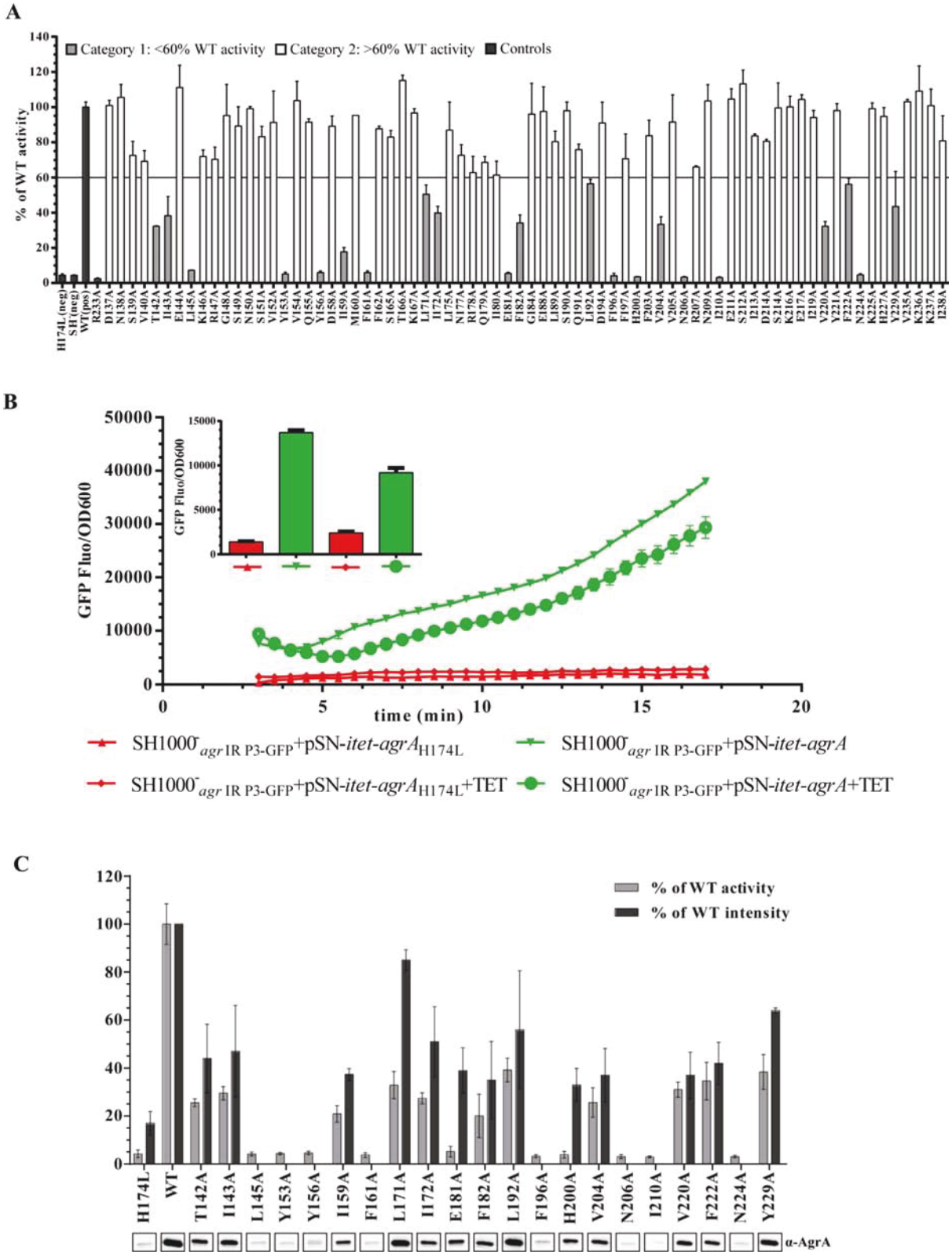
showed that wild-type level of *agr* activity can be restored in SH1000$^-$ in the presence of plasmid-borne AgrA (29). We obtained the nucleotide sequence of *agrA* from SH1000$^-$ and identified that it contains a leucine substitution at a conserved histidine at aa position 174 (H174) in the LytTR domain (Supplementary Figure S1). In the crystal structure of the AgrA LytTR domain, aa H174 forms a salt bridge with aa E163 and E226, which contributes to stabilizing the interface between three β sheets in the LytTR domain (Figure 1B) (4). Therefore, it is likely that a leucine substitution at aa H174L compromises the overall structural integrity of AgrA, rendering it unavailable for transcription activation. As expected, AgrA protein was detectable in whole-cell lysates of *agr*-positive SH1000 (the isogenic parent of SH1000$^-$), but was hardly detectable in whole-cell extracts of SH1000$^-$ on western blots using polyclonal anti-AgrA antibodies (Figure 1C). Since the nucleotide sequence upstream of *agrA* including intergenic region and the P2 and P3 promoter seqences were intact in SH1000$^-$, these results are consistent with the view that the H174L mutation destablizes the structural integrity of AgrA in SH1000$^-$ and thereby makes it unavailable to activate transcription from P2 and P3 promoters of the *agr* operon. To ascertain that this is indeed the case, we placed a transcriptional fusion of P3 to GFP at the *geh* locus on the SH1000$^-$ chromosome (creating SH1000$^-_{\text{agr IR P3-GFP}}$) and measured GFP fluorescence (indicating transcription from the P3 promoter) as a function of cell growth in the presence of plasmid-borne wild-type AgrA (pSN-P2-*agrA*) (Supplementary Table S1). In pSN-P2-*agrA*, the transcription of *agrA* is driven from its native promoter P2 and is dependent on functional AgrA. In other words, the transcription of *agrA* from pSN-P2-*agrA* will be autocatalytic as it is the case at the P2 promoter at the native *agr* locus. As shown in Figure 1D and as expected, P3 activity was not detected in SH1000$^-_{\text{agr IR P3-GFP}}$ containing pSN-P2-*empty* or pSN-P2-*agrA*$_{\text{H174L}}$. However, P3 activity was restored in SH1000$^-_{\text{agr IR P3-GFP}}$ containing the pSN-P2-*agrA*. Further, since *agr* dysfunction is associated with reduced β-hemolytic activity, we used a blood agar plate hemolysis assay to confirm that β-hemolytic activity is restored in SH1000$^-_{\text{agr IR P3-GFP}}$ containing pSN-P2-*agrA*, but not in the SH1000$^-_{\text{agr IR P3-GFP}}$ containing pSN-P2-*agrA*$_{\text{H174L}}$ or pSN-P2-*empty* (Figure 1E). We note a large difference seen in AgrA levels between in the *agr*-positive SH1000$_{\text{agr IR P3-GFP}}$ cells (Figure 1C, lane 4) and the complemented *agrA*-defective SH1000$^-_{\text{agr IR P3-GFP}}$ + pSN-P2-*agrA* cells (Figure 1C, lane 1). This is as expected since there are at least 25–30 additional copies (which equates to the copy number of pSN-P2-*agrA*) of *agrA* in SH1000$^-_{\text{agr IR P3-GFP}}$ + pSN-P2-*agrA* than in SH1000$_{\text{agr IR P3-GFP}}$, which will only contain one copy of *agrA*. However, rate-limiting factor for AgrA activation is AgrC, which will be present at wild-type levels in the SH1000$^-_{\text{agr IR P3-GFP}}$ + pSN-P2-*agrA* cells, it is unlikely that excess AgrA in SH1000$^-_{\text{agr IR P3-GFP}}$ + pSN-P2-*agrA* cells will have any detectable biological impact on AgrA function as it is evident in results shown in Figure 1D and E. In summary, the results unambigiously establish that SH1000$^-_{\text{agr IR P3-GFP}}$ in combination with plasmid-derived AgrA (from pSN-P2-*agrA*) can be used as a reporter strain to accurately measure AgrA activity in *S. aureus*.

Alanine-scanning mutagenesis analysis of AgrA identifies four amino acid residues potentially important for transcription activation in the LytTR domain

To indentify aa residues in the LytTR domain involved in transcription activation, we generated a mutant library of AgrA by alanine-scanning mutagenesis of the LytTR domain using pSN-P2-*agrA* as the template. Based on the crystal structure of the AgrA LytTR domain–DNA complex (4), we only targeted residues for alanine substitution that were neither implicated in DNA binding, nor known to be important for maintaining the structrual fold of the LytTR domain. We also constructed the previously described DNA-binding defective mutant AgrA$_{\text{R233A}}$ to use as a negative control. SH1000$^-_{\text{agr IR P3-GFP}}$ cells were transformed with the library comprising a total of 74 AgrA mutants in pSN-P2-*agrA* and mutant AgrA activity was determined by measuring GFP fluorescence as a function of cell density ($OD_{600\ nm}$) after 8 h of growth in rich media. As expected, GFP activity was bareley detected in SH1000$^-_{\text{agr IR P3-GFP}}$ containing pSN-P2-*agrA*$_{\text{R233A}}$ and pSN-P2-*agrA*$_{\text{H174L}}$ in comparison to SH1000$^-_{\text{agr IR P3-GFP}}$ containing pSN-P2-*agrA* (Figure 2 A). Based on the spread of the activities of the mutant AgrA library in SH1000$^-_{\text{agr IR P3-GFP}}$, the AgrA mutants were categorized into two activity groups: <60% (21 mutants) and >60% (51 mutants) activity relative to wild-type AgrA (Figure 2A). Mutants in the first group, which hereafter are referred to as putative transcription-activation-defective (TAD) mutants, were selected for further analysis.

To establish that the TAD property is not caused by the alanine substitution adversely affecting the stability of the mutant AgrA protein under the assay conditions (after 8 h of growth in rich media), we transferred the 21 putative TAD mutations into plasmid pSN-*itet-agrA*, so that expression of AgrA can be induced with anhydrous tetracycline and thus independent of the P2 promoter (see above). In other words, unlike with pSN-P2-*agrA*, transcription of *agrA* from pSN-*itet-agrA* will not be autocatalytic and therefore will be independent of AgrA. This approach will avoid complications in the interpretation of the results arising from a putative TAD mutant failing to or poorly activating its own transcription. Initially, we tested if P3 activity in SH1000$^-_{\text{agr IR P3-GFP}}$ can be restored with pSN-*itet-agrA* in the presence of anhydrotetracycline. As shown in Figure 2B, P3 activity was restored to comparable levels in SH1000$^-_{\text{agr IR P3-GFP}}$ containing pSN-*itet-agrA*, both in the presence and absence of anhydrotetracycline, but, as expected, not in SH1000$^-_{\text{agr IR P3-GFP}}$ containing pSN-*itet-agrA* $_{\text{H174L}}$ even in the presence of anhydrotetracycline. It seems that leaky expression of AgrA from pSN-*itet-agrA* is sufficient to restore full P3 activity in SH1000$^-_{\text{agr IR P3-GFP}}$ containing pSN-*itet-agrA*, therefore all downstream experiments involving pSN-*itet-agrA* were conducted in the absence of anhydrotetracycline. We next monitored, by Western blotting using anti-AgrA antibodies, the accumulation of AgrA in SH1000$^-_{\text{agr IR P3-GFP}}$ + pSN-*itet-agrA* (in the absence of anhydrotetracycline) over time to determine a time point at which saturation in AgrA levels has occurred. As shown in Supplementary Figure S2, AgrA levels begun to saturate approximately after 6 h of growth in rich media.

Downloaded from http://nar.oxfordjournals.org/ at Imperial College London on March 17, 2015



Downloaded from http://nar.oxfordjournals.org/ at Imperial College London on March 17, 2015

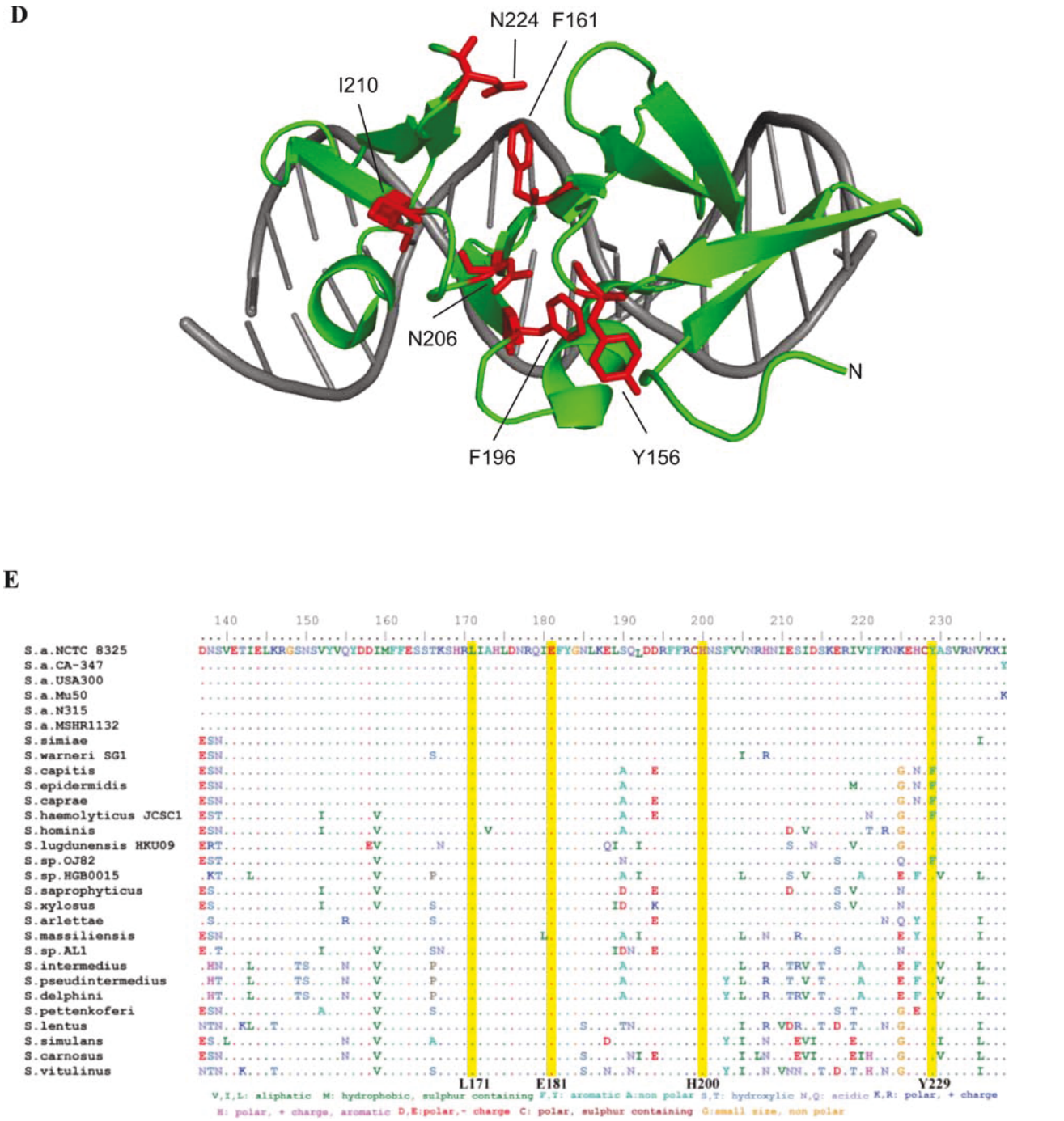


Figure 2. Systematic mutational analysis of *Staphylococcus aureus* AgrA LytTR domain. (**A**) Graph showing GFP expression [as GFP fluorescence units (GFP-FU)] as a function of growth (OD600) relative to the SH1000$^{-}{}_{\text{agr IR P3-GFP}}$ + pSN-P2-*agrA* (WT) for each single alanine mutant after 8 h of growth in TSB. AgrA mutants displaying more than 60% activity compared to the wild-type AgrA are shown in white, the AgrA mutants displaying less than 60% activity compared to the wild-type AgrA are shown in grey and the controls (SH1000$^{-}{}_{\text{agr IR P3-GFP}}$ + pSN-P2-*agrA*$_{\text{H174L}}$, SH1000$^{-}{}_{\text{agr IR P3-GFP}}$ + pSN-P2-*empty*, SH1000$^{-}{}_{\text{agr IR P3-GFP}}$ + pSN-P2-*agrA* and SH1000$^{-}{}_{\text{agr IR P3-GFP}}$ + pSN-P2-*agrA*$_{\text{R233A}}$) are shown in black. (**B**) Graphs showing GFP expression [as GFP fluorescence units (GFP-FU)] as a function of growth (OD600) over time for SH1000$^{-}{}_{\text{agr IR P3-GFP}}$ + pSN-*itet-agrA* (green lines) and SH1000$^{-}{}_{\text{agr IR P3-GFP}}$ + pSN-*itet-agrA*$_{\text{H174L}}$ (red lines) strains grown in TSB with and without anhydrotetracycline. The bar chart in the insert represents the GFP expression of the same samples (color coded accordingly) at the 8 h time point. (**C**) Graph showing GFP expression [as GFP fluorescence units (GFP-FU)] as a function of growth (OD600) relative to the SH1000$^{-}{}_{\text{agr IR P3-GFP}}$ + pSN-*itet-agrA* (WT) for each of the 21 single alanine mutant displaying less than 60% wild-type activity after 8 h of growth in TSB. A section of the western blot image indicating AgrA detected in whole cell lysates for each mutant is shown under the graph. For the 13 mutants that are detectably expressed, the quantification of the intensity of the band corresponding to AgrA mutants relative to the intensity of the band corresponding to the wild-type AgrA level is shown on the graph. Data for (A–C) were obtained from at least three biological replicates. **(D)** As in Figure 1B with the six of the eight aa residues (Y156, F161, F196, N206, I210 and N224), which when changed to alanine appear to impair the gross structural stability of AgrA indicated in red. **(E)** Multiple sequence alignment of the AgrA LytTR domain of representative staphylococci strains. Conserved residues are represented by a dot. The aa residues displaying similar colors have similar properties. The putative TAD mutants are highlighted in yellow.

Downloaded from http://nar.oxfordjournals.org/ at Imperial College London on March 17, 2015

Therefore, we used the 8 h post-inoculation time point to compare the amount of AgrA in the cell and GFP fluorescence (indicative of P3 activity) in SH1000$^{-}{}_{\text{agr IR P3-GFP}}$ cells producing the TAD mutants from pSN-*itet-agrA* with SH1000$^{-}{}_{\text{agr IR P3-GFP}}$ + pSN-*itet-agrA* expressing wild-type AgrA. As shown in Figure 2C, 8 of the 21 mutants are not detectably expressed, which suggests that alanine substitution at positions L145, Y153, Y156, F161, F196, N206, I210 and N224 in the LytTR domain of AgrA could impair the gross structural stability of AgrA under the assay conditions. Consistent with this view, side-chains of Y156, F161, F196, N206, I210 and N224 are either fully or partly buried within the hydrophobic core of AgrA (Figure 2D). The remaining 13 AgrA mutants are detectably expressed at varying levels compared to wild-type AgrA under the assay conditions. Notably, even though L171A, E181A, H200A and Y229A mutants are expressed to levels at least $\geq$ 35–50% of that of wild-type AgrA, a relatively reduced level of P3 activity (by $\geq$60% reduction in activity relative to wild-type AgrA) is seen in cells containing these mutants. Thus, the results suggest that aa residues L171, E181, H200 and Y229 in the LytTR domain could potentially play a significant role in transcription activation by AgrA. This conclusion is further substantiated by the 100% identity of L171, E181 and H200 among *agrA* alleles in representative staphylococcal species; Y229 is also highly conserved, being replaced by the similar phenylalanine in 5 of the 29 sequences (Figure 2E).

Conserved amino acid Y229 in the LytTR domain of AgrA is not required for DNA binding or bending but for transcription activation

In the crystal structure of the AgrA LytTR domain-DNA complex, aa residues L171, E181, H200 and Y229 are located proximal to the DNA (Figure 3A). We initially wanted to determine, using purified proteins, if alanine substitution at L171, E181, H200 and Y229 impairs the DNA binding activity of the mutant AgrA protein, and thereby contributes to the observed TAD phenotype. To test the DNA-binding activity of the L171A, E181A, H200A and Y229A AgrA mutants, we conducted electrophoretic mobility shift based protein-DNA-binding assays with a 214 bp long DNA probe containing the intergenic region of the *agr* operon and with the purified AgrA mutants. Since AgrA requires phosphorylation by AgrC for specific binding to DNA (see above), *in vitro* phosphorylation of AgrA was achieved using the small phosphodonor acetyl phosphate. The AgrA$_{\text{R233A}}$ mutant served as the negative control in the protein-DNA-binding assays. As shown in Figure 3B, and as expected, phosphorylation with acetyl phosphate markedly increased the binding of wild-type AgrA to the DNA probe (compare lanes 1 and 2) and AgrA$_{\text{R233A}}$ mutant did not detectably bind to the DNA probe in the presence of acetyl phosphate compared to wild-type AgrA (compare lanes 11 and 12), thus indicating the specific binding of AgrA to the DNA probe in the presence of acetyl phosphate. Of the four putative TAD mutants, AgrA$_{\text{Y229A}}$ bound to the DNA probe to a level comparable to the wild-type AgrA, whilst AgrA$_{\text{L171A}}$, AgrA$_{\text{E181A}}$ and AgrA$_{\text{H200A}}$ bound to the DNA probe with different degrees of reduced efficiency compared to wild-type AgrA (Figure 3B, lanes 3–10). Overall, the results indicate that an alanine substitution at aa residue Y229 in the LytTR domain of AgrA, whilst only moderately affecting DNA-binding, significantly impairs the ability of AgrA to activate transcription. Further, since the AgrA$_{\text{Y229A}}$ mutant binds DNA in a phosphorylation-dependent manner, we can exclude the possibility that, under the assay conditions, the TAD property of this mutant is due to defects associated with phosphorylation.

To ascertain that AgrA$_{\text{Y229A}}$ is a *bona fide* TAD mutant, we tested the ability of the AgrA$_{\text{Y229A}}$ to activate transcription from the *agr* operon P3 promoter *in vitro*. The AgrA$_{\text{L171A}}$ mutant was also included in the *in vitro* transcription assays as it displayed the best DNA binding activity ($\sim$50% wild-type activity; Figure 3B, compare lanes 2 and 4) compared to AgrA$_{\text{E181A}}$ and AgrA$_{\text{H200A}}$ mutants. As shown in Figure 3C, phosphorylated AgrA$_{\text{Y229A}}$ failed to activate transcription from the P3 promoter to a level comparable to that seen in reactions containing wild-type AgrA (compare lanes 2 and 3 with lanes 6 and 7). In contrast, phosphorylated AgrA$_{\text{L171A}}$ activated transcription from the P3 promoter to a level almost comparable to that of wild-type AgrA (compare lanes 2 and 3 with lanes 4 and 5), even though this mutant binds DNA approximately 50–60% less efficiently than the AgrA$_{\text{Y229A}}$ mutant or wild-type AgrA.

Tyrosine is a polar, hydrophobic and aromatic aa residue. To investigate the role of the tyrosine residue at position Y229 we made the AgrA$_{\text{Y229F}}$, AgrA$_{\text{Y229I}}$ and AgrA$_{\text{Y229G}}$ mutants and compared their ability to activate transcription from the P3 promoter *in vivo*. As shown in Figure 3D, substitution of Y229 by the non-polar, hydrophobic and aromatic aa phenylalanine did not significantly affect the ability of the AgrA$_{\text{Y229F}}$ mutant to activate transcription. Similarly, substitution of Y229 by the non-polar, hydrophobic, and non-aromatic aa isoleucine, which contains a bulky side chain as tyrosine did not significantly affect the ability of the AgrA$_{\text{Y229I}}$ to activate transcription. In contrast, substitution of Y229 by the small non-polar, hydrophobic, non-aromatic aa glycine significantly impaired the ability of AgrA$_{\text{Y229G}}$ mutant to activate transcription, as seen with AgrA$_{\text{Y229A}}$. Overall, the results suggest that the bulky hydrophobic side chain of tyrosine at position 229 in the LytTR domain in AgrA is required for maximal activation of transcription.

To further interrogate the interactions made by aa residue Y229 in AgrA, we carried out a series of fully atomistic accelerated molecular dynamics (aMD) stimulations. aMD is a powerful technique to explore the conformational space available to macromolecules (30) and is thus a particularly effective method for investigating structural changes induced by mutations. Simulation of the wild-type AgrA LytTR domain revealed that the core of the structure is very stable and deviates only within the expected range (root-mean square deviation $<$ 2.4 Å) from the original crystal structure during 100 ns of aMD simulation. In contrast, simulation of AgrA LytTR domain containing the *in silico* Y229-A substitution under identical conditions shows a substantial destabilization of the carboxyl-terminal domain (Figure 4; Supplementary movie 1). In the wild-type structure Y229 makes close contact with C199, H200 and I219 based on a dense network of van der Waals contacts, hydrogen-bonding and hydrophobic interactions.

Downloaded from http://nar.oxfordjournals.org/ at Imperial College London on March 17, 2015

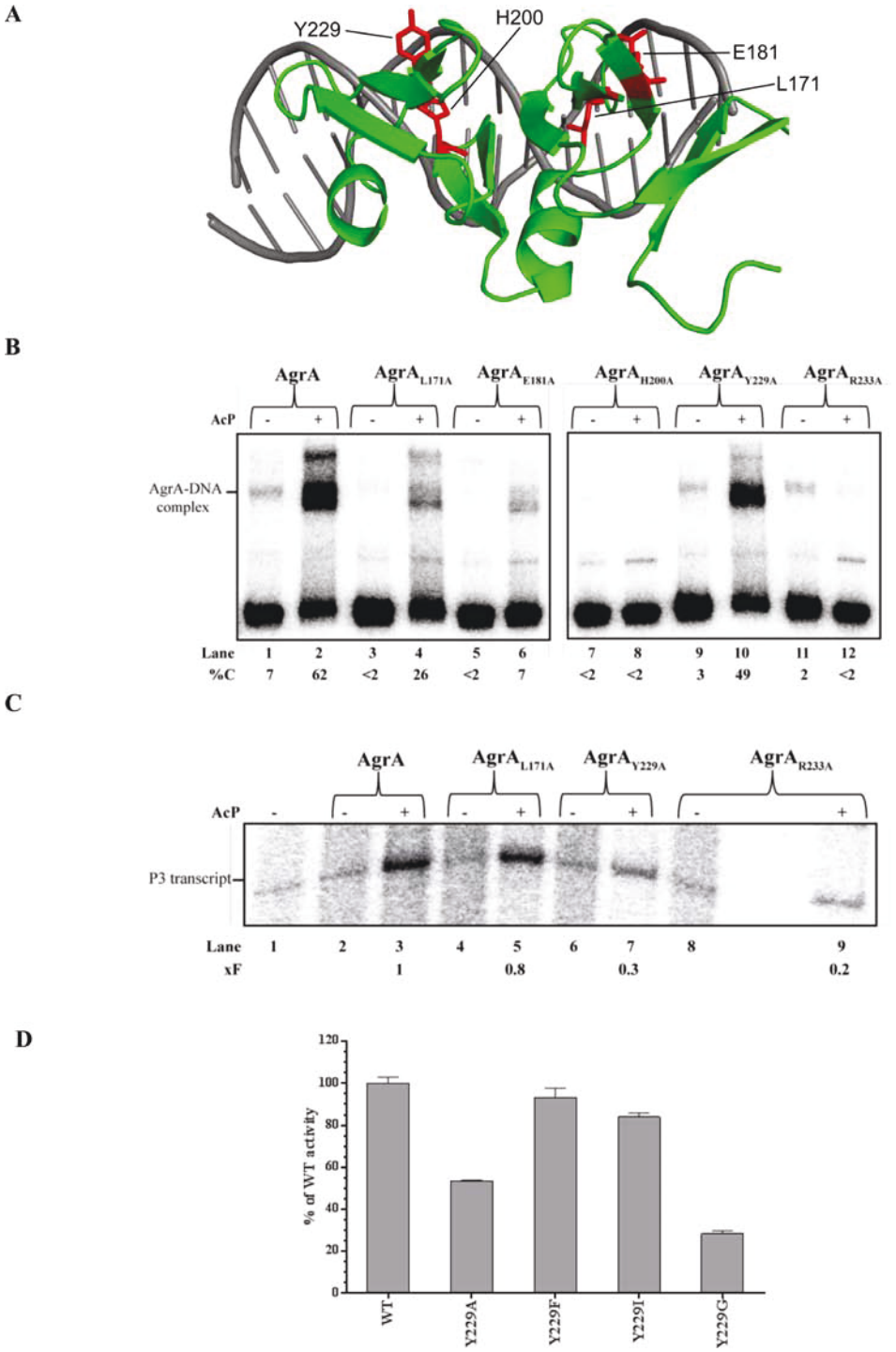


Figure 3. Conserved amino acid Y229 in the LytTR domain of AgrA is not required for DNA binding or bending but for transcription activation. (**A**) As in Figure 1B with the residues, which when substituted with alanine confer the putative TAD phenotype indicated in red. (**B**) Autoradiograph image of a 4–20% (w/v) native polyacrylamide gel comparing the ability of the four TAD mutants AgrA$_{L171A}$ (lanes 3, 4), AgrA$_{E181A}$ (lanes 5, 6), AgrA$_{H200A}$ (lanes 7, 8), AgrA$_{Y229A}$ (lanes 9, 10) with that of wild-type AgrA (lanes 1, 2) to bind to a DNA probe representing the intergenic region of the *agr* operon (as shown in Figure 1A) in absence (lanes 1, 3, 5, 7, 9, 11) and presence (lanes 2, 4, 6, 8, 10, 12) of acetyl phosphate. Lanes 11 and 12 contain the DNA-binding mutant AgrA$_{R233A}$, which is used here as a negative control. The percentage of DNA bound by AgrA (%C) in each reaction is given at the bottom of the gel. (**C**) Autoradiograph image of a 10% (v/v) denaturing urea gel showing the synthesis of the P3 transcript by wild-type AgrA (lanes 2, 3), AgrA$_{L171A}$ (lanes 4, 5), AgrA$_{Y229A}$ (lanes 6, 7) and AgrA$_{R233A}$ (lanes 8, 9) in the absence (lanes 2, 4, 6, 8) and presence (lanes 3, 5, 7, 9) of acetyl phosphate. The fold decrease in transcription (xF) compared to the P3 transcript in the presence of activated AgrA (lane 3) in each reaction is given at the bottom of the gel. (**D**) Graph showing GFP expression [as GFP fluorescence units (GFP-FU)] as a function of growth (OD600) relative to the SH1000$^{-}{}_{agr\ IR\ P3\text{-}GFP}$ + pSN-P2-*agrA* strain (WT) for SH1000$^{-}{}_{agr\ IR\ P3\text{-}GFP}$ containing pSN-P2-*agrA*$_{Y229A}$, pSN-P2-*agrA*$_{Y229F}$, pSN-P2-*agrA*$_{Y229I}$ and pSN-P2-*agrA*$_{Y229G}$ after 8 h of growth in TSB. For (B) and (C) at least two independent experiments were done to obtain the %C and xF values, respectively. The data obtained were all within 5–10% of the values shown. Data for (D) were obtained from at least three biological replicates.

Downloaded from http://nar.oxfordjournals.org/ at Imperial College London on March 17, 2015

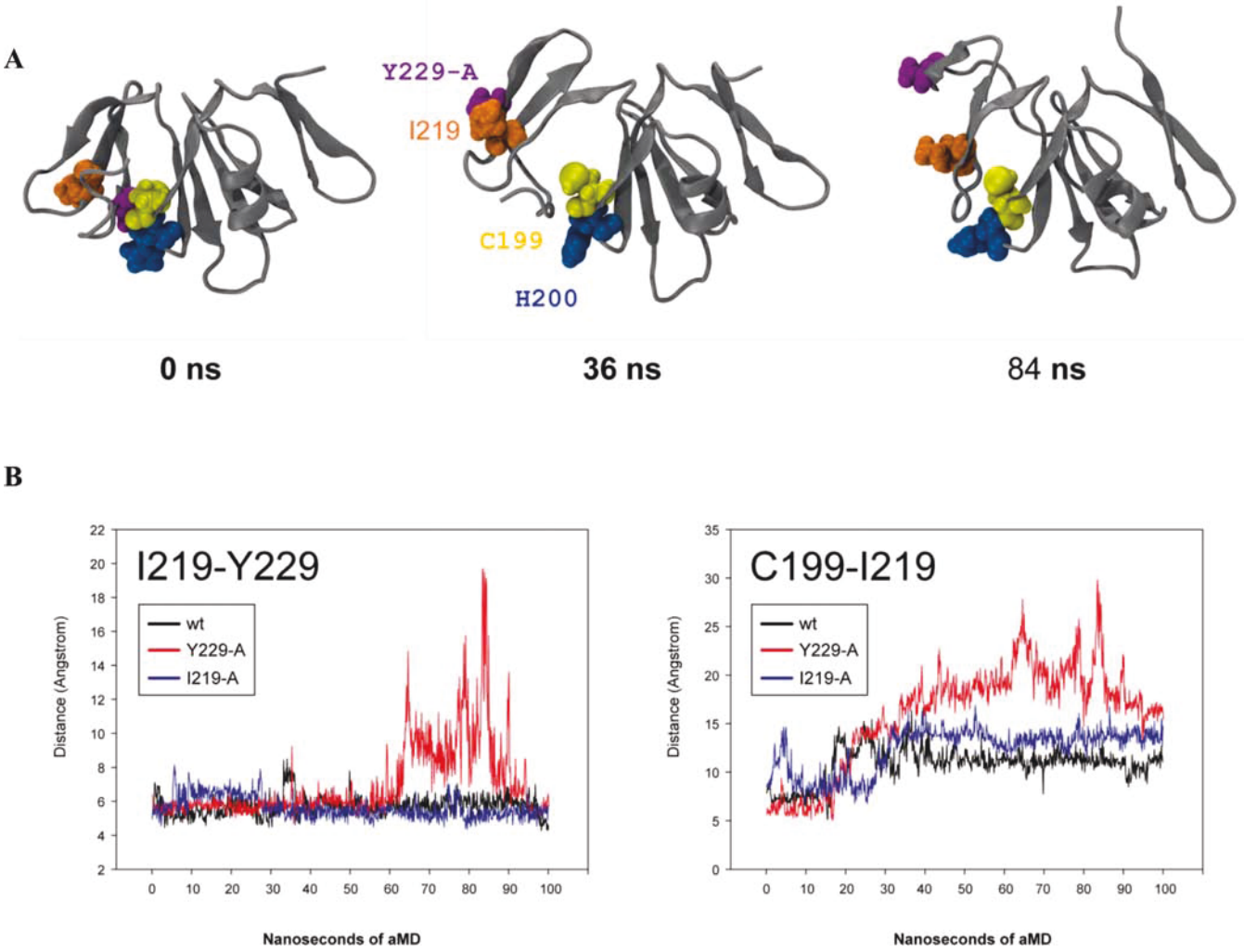


Figure 4. Molecular dynamics analysis of AgrA LytTR domain. (**A**) Snapshots of the aMD simulation of Y229-A at different time points. The polypeptide chain of the AgrA LytTR domain is shown as a grey ribbon. The side chains of the aas of interest are represented as van der Waals spheres (C199 yellow, H200 blue, I219 orange and the Y229-A *in silico* substitution in purple). At the beginning of the simulation (t = 0 ns), Y229-A is in close contact (as Y229 would be in the wild-type structure) with C199 and H200. This association dissociates during the course of the simulation, resulting in the carboxyl-terminal three β-sheets dissociating from the remainder of the LytTR domain (t = 36). At a later stage, Y229-A and I219 dissociate from each other (t = 84). (**B**) Quantitation of the aMD results shown in (A): The center-of-mass distances between I219 and Y229-A (left panel) and I219-A (right panel) are plotted against aMD simulation time for the wild-type (wt) structure (black), Y229-A (red) and I219-A (blue). The distance between I219 and Y229-A increases substantially at one stage of the simulation (from 55 ns onwards), but appears to reform in the final stages of the simulation. In contrast, the wild-type and I219-A simulations maintain a high degree of stability of this interaction. The distance between C199 and I219 serves as a measure of the dissociation of the three carboxyl-terminal β-sheets.

These contacts are severely disrupted in Y229-A, resulting in the rapid dissociation of the three C-terminal β-strands from the main β-sheet transversing the LytTR domain. At a later stage of the simulation, it also becomes evident that Y229-A further destabilizes the two carboxyl-terminal β strands and causes a reversible unfolding of this sub-structure due to the absence of interaction between I219 and Y229. Taking into account the structural relevance of the interaction between I219 and Y229, it appeared initially puzzling that the I219A substitution displayed no discernible defect in activating transcription from the P3 promoter (Figure 2A). Simulations of the *in silico* I219-A substitution revealed, however, that in this situation Y229 forms an alternative interaction network involving the adjacent residues E217 and R218 (Figure 4; Supplementary movie 2). It is therefore evident, from both theoretical as well as experimental observations, that I219 plays a structurally redundant role (which also explains the evolutionary variability in this position (Supplementary Figure S1)). Overall the aMD simulations show that an alanine substitution at Y229 is predicted to have a substantial impact on the carboxyl-terminal part of the LytTR domain of AgrA. The mutation causes a 'localized' structural destabilization that may prevent AgrA from adopting a conformation required for transcription activation, while still retaining DNA-binding activity.

A previous study by Reyes *et al*. reported that AgrA induced bending of DNA is the main driving force for activation of transcription from the P2 and P3 promoters (14). However, based on the results above (Figures 3D and

Downloaded from http://nar.oxfordjournals.org/ at Imperial College London on March 17, 2015

4), it is unlikely that Y229 is involved in the DNA bending activity of AgrA. Therefore, to directly rule out this possibility, we carried out the DNA bending assay previously used by Reyes *et al*. to compare the DNA bending activity of $AgrA_{Y229A}$ mutant with that of wild-type AgrA in the presence of acetyl phosphate. This assay uses the pAM1847 plasmid containing the AgrA tandem binding site upstream of the P2 promoter region cloned between the *Sac*I and *Bgl*I sites (schematic in Figure 5). Digestion of the plasmid with *Eco*RI, *Hind*III, *Bst*NI, *Eco*RV, *Nhe*I and *BamH*I results in DNA fragments of identical length and composition in which only the position of the AgrA binding sites is permuted with respect to the 5′ terminus of the fragments (schematic in Figure 5). Therefore, AgrA-mediated DNA bending can be monitored by nondenaturing polyacrylamide electrophoretic analysis of the AgrA–DNA complexes, as the mobility of these will be strongly dependent on the position of the bend in the DNA molecule as a result of AgrA binding. As shown in Figure 5, no detectable differences were seen in the pattern of the mobilities of the wild-type and mutant AgrA–DNA complexes, thus indicating that $AgrA_{Y229A}$ and wild-type AgrA bend the DNA equally well and that conserved aa Y229 in the LytTR domain of AgrA is not a major determinant of the DNA bending activity of AgrA.

DISCUSSION

Our analysis has provided several novel insights into AgrA function in *S. aureus* and other LytTR containing transcription factors:

The *agr* operon indirectly controls the expression of hemolysins in *S. aureus* via RNAIII and directly via AgrA (8). Lack of hemolytic activity is typical of strains with dysfunctional *agr* operon. A dysfunctional *agr* operon is considered to provide an adaptive advantage for survival in the infected host but counter-adaptive outside infected host tissues (31). Under laboratory growth conditions, spontaneous mutations in the *agr* operon genes can occur, which can confer a nonhemolytic phenotype and this has also been observed for the widely used laboratory strain SH1000, with the nonhemolytic variant designated $SH1000^{-}$ (29). Our results reveal that the molecular basis for the *agr* operon dysfunction in $SH1000^{-}$ is a point mutation at a conserved aa residue (H174) in AgrA, which disrupts salt–bridge interactions that stabilizes the interaction between three β sheets in the LytTR domain of AgrA. Further, a phenylalanine substitution at aa residues L171 in AgrA has been previously reported to confer the nonhemolytic phenotype (indicative of *agr* operon dysfunction) in a nosocomial methicillin-resistant *S. aureus* isolate (31). Intriguingly, our results reveal that even though $AgrA_{L171A}$ mutation displays a reduced (by ∼50%) DNA binding activity compared to wild-type AgrA, its ability to activate transcription is only moderately affected (∼20% reduction compared to wild-type AgrA). Thus, it is possible that an aromatic side chain at position 171 in the LytTR domain is more deleterious for AgrA activity than the presence of alanine.

The results reveal that alanine substitutions at aa residues L171, E181 or H200 significantly impair the DNA-binding activity of AgrA. It is possible that alanine substitutions at these aa residues indirectly affect the ability of mutant AgrA to become phosphorylated, however, this is unlikely given their proximity to DNA in the context of the structure of the AgrA LytTR domain–DNA complex (Figure 3A). The most conserved sequence motif within the LytTR domain is FhRhHRS (where 'h' indicates a hydrophobic aa) (3). In AgrA, the FhRhHRS corresponds to FFRCHNS (residues 196–202). In the crystal structure of the AgrA LytTR domain–DNA complex only aa residues R198, N201 and S202 were shown to be involved in interaction with DNA and alanine substitutions at N201 reduced the ability of the AgrA LytTR domain to bind DNA (4). These observations clearly imply a role for aa residue H200 within the highly conserved FFRCHNS motif in the binding of AgrA to DNA, and, based on the proximity of H200 to DNA in the AgrA LytTR domain–DNA complex, base-specific contacts between H200 and DNA cannot be excluded. In further support of this view, alanine substitution of H188 (equivalent residue of H200 in AgrA) in the LytTR domain of *Clostridium perfringens* RR-TF VirR, which regulates virulence and toxin gene expression, confers a loss of activity phenotype and residues R186 and S190 in the FhRhHRS motif of VirR has been shown to be involved in DNA-binding (5).

Importantly, the results identify the highly conserved aa residue Y229 in the LytTR domain of AgrA as a key determinant for maximal transcription activation by AgrA and the $AgrA_{Y229A}$ mutant still retains some level of activity (40–50% *in vivo* and 30% *in vitro* compared to wild-type AgrA) to activate transcription. In the structure of the AgrA LytTR domain–DNA complex, Y229 is adjacent to an aa residue critical for AgrA folding (C228) and DNA interaction (R218), thus indicating that Y229 is part of a multifunctional region in AgrA. Consistent with this notion, a recent study reported that a novel antibacterial compound, called savarin, binds to the LytTR domain of AgrA in a region proximal to aa residues C228, R218 and Y229 and thereby abrogates AgrA function (32). The proximity of both AgrA-binding sites to the P2 and P3 core promoter elements (Figure 1A) suggests that transcription activation at both promoters could involve direct protein–protein interaction between the region containing Y229 in AgrA and the RNAp and thus could occur via a simple 'recruitment' mechanism, whereby AgrA could facilitate the binding of RNAp to the promoter to yield a transcriptionally proficient promoter complex. Furthermore, a previous study by Reyes *et al*. reported that transcription from P2 and P3 agr operon promoters is differentially regulated, with the former (P2) dependent on AgrA and SarA and the later dependent only on AgrA for maximum promoter activity (14). Therefore, it is possible that the region containing Y229 in AgrA is involved in interaction with SarA for the activation and that the $SH1000^{-}{}_{\text{agr IR P3-GFP}}$ reporter strain used in this study indirectly indicates P2 activity (which drives its own transcriptio, see Materials and Methods). However, this is unlikely because $AgrA_{Y229A}$ mutant displays the same level of activity in the context of pSN-P2-*agrA* and pSN-*itet-agrA* in $SH1000^{-}{}_{\text{agr IR P3-GFP}}$ strain. The aMD analysis reveals that an alanine substitution at Y229 could cause a 'localized' structural destabilization that may prevent AgrA from adopting a conformation required for efficient tran-

Downloaded from http://nar.oxfordjournals.org/ at Imperial College London on March 17, 2015

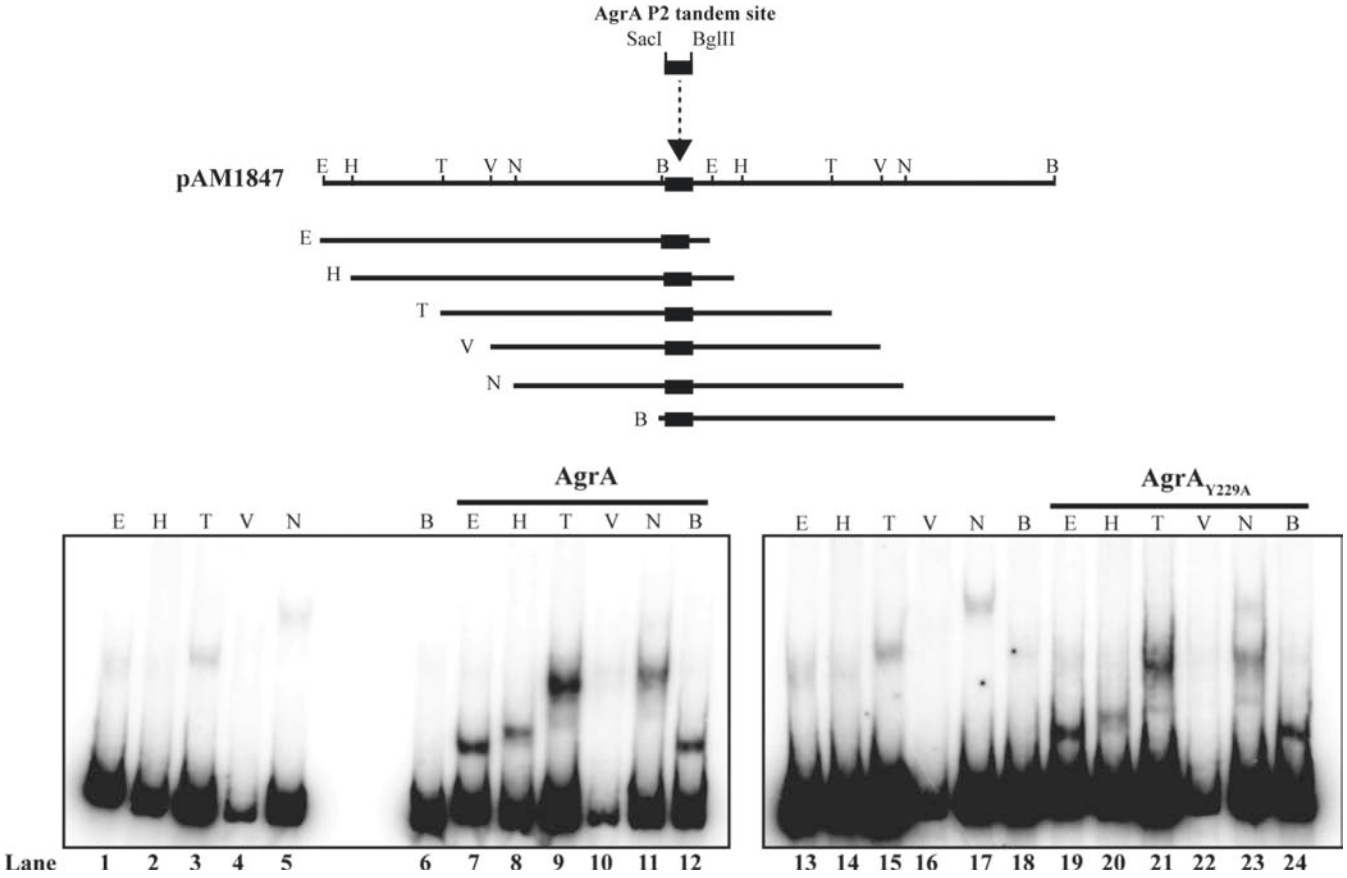


Figure 5. A schematic (based on Reyes *et al.* (14)) of the DNA bending vector pAM1847 containing the AgrA P2 tandem binding site (represented by a black rectangle) cloned into the *Sac*I-*Bgl*II site. The *Eco*RI-*Sac*I fragment on the 5′ side of the AgrA P2 binding site is identical to the *Bgl*II-*Bam*HI fragment on the 3′ side of the AgrA P2 binding site. Hence, digestion of the recombinant vector with any of the six restriction enzymes *Eco*RI (E), *Hind*III (H), *Bst*NI (N), *Eco*RV (V), *Nhe*I (N) or *Bam*HI (B) produces DNA fragments of identical length but with a different position of the AgrA P2 binding site with respect to the 5′ and 3′ end of the fragment. The autoradiograph image of a 4.5% (v/v) nondenaturing polyacrylamide gel shows the mobilities of phosphorylated AgrA$_{Y229A}$ and wild-type AgrA-DNA complexes bound to each of the DNA fragments generated upon digestion of pAM1847 with *Eco*RI (E), *Hind*III (H), *Bst*NI (N), *Eco*RV (V), *Nhe*I (N) or *Bam*HI (B). Lanes 1–6 and 13–18 contain no protein. Data from at least two independent experiments.

scription activation. This observation further substantiates the multifaceted role of Y229 in transcription activation and maintaining local structural integrity of AgrA. Future work will focus on molecular details of such potential interactions in order to further delineate the mechanism by which AgrA activates transcription.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank William Kelley and Ambrose Cheung for plasmid pAM1847, Jan Maarten van Dijl for the *Staphylococcus aureus* SH1000^{-} strain, Angelika Gründling for the pCN34 and pCN34*itet* plasmids. We are grateful to Dan Brown, Aisha Krishna, Andrew Edwards and Christoph Engl for useful discussions and comments on the manuscript.

FUNDING

Biotechnological and Biological Research Council [BB/I001492/1 to S.W.]; Medical Research Council [G1100057 to R.O.J.W.]; BBSRC PhD studentship to S.N. S.W. is a Wellcome Trust Investigator. Funding for open access charge: BBSRC.
Conflict of interest statement. None declared.

REFERENCES

1. Gao,R. and Stock,A.M. (2009) Biological insights from structures of two-component proteins. *Annu. Rev. Microbiol.*, **63**, 133–154.
2. Galperin,M.Y. (2006) Structural classification of bacterial response regulators: diversity of output domains and domain combinations. *J. Bacteriol.*, **188**, 4169–4182.
3. Nikolskaya,A.N. and Galperin,M.Y. (2002) A novel type of conserved DNA-binding domain in the transcriptional regulators of the AlgR/AgrA/LytR family. *Nucleic Acids Res.*, **30**, 2453–2459.
4. Sidote,D.J., Barbieri,C.M., Wu,T. and Stock,A.M. (2008) Structure of the Staphylococcus aureus AgrA LytTR domain bound to DNA reveals a beta fold with an unusual mode of binding. *Structure*, **16**, 727–735.
5. McGowan,S., Lucet,I.S., Cheung,J.K., Awad,M.M., Whisstock,J.C. and Rood,J.I. (2002) The FxRxHrS motif: a conserved region

Downloaded from http://nar.oxfordjournals.org/ at Imperial College London on March 17, 2015

essential for DNA binding of the VirR response regulator from Clostridium perfringens. *J. Mol. Biol.*, **322**, 997–1011.
6. Lizewski,S.E., Lundberg,D.S. and Schurr,M.J. (2002) The transcriptional regulator AlgR is essential for Pseudomonas aeruginosa pathogenesis. *Infect. Immun.*, **70**, 6083–6093.
7. Miller,D.P., Frederick,J.R., Sarkar,J. and Marconi,R.T. (2014) The Treponema denticola AtcR LytTR domain-containing response regulator interacts with three architecturally distinct promoter elements: implications for understanding the molecular signaling mechanisms that drive the progression of periodontal disease. *Mol. Oral Microbiol.*, **29,** 219–232.
8. Novick,R.P. (2003) Autoinduction and signal transduction in the regulation of staphylococcal virulence. *Mol. Microbiol.*, **48**, 1429–1449.
9. Fournier,B. (2008) Global regulators of Staphylococcus aureus virulence genes. In: Lindsay,JA (ed). Staphylococcus: Molecular Genetics. Caister Academic Press, Norfolk, pp. 131–183.
10. Queck,S.Y., Jameson-Lee,M., Villaruz,A.E., Bach,T.H., Khan,B.A., Sturdevant,D.E., Ricklefs,S.M., Li,M. and Otto,M. (2008) RNAIII-independent target gene control by the agr quorum-sensing system: insight into the evolution of virulence regulation in Staphylococcus aureus. *Mol. Cell*, **32**, 150–158.
11. Sun,F., Liang,H., Kong,X., Xie,S., Cho,H., Deng,X., Ji,Q., Zhang,H., Alvarez,S., Hicks,L.M. *et al.* (2012) Quorum-sensing agr mediates bacterial oxidation response via an intramolecular disulfide redox switch in the response regulator AgrA. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 9095–9100.
12. Koenig,R.L., Ray,J.L., Maleki,S.J., Smeltzer,M.S. and Hurlburt,B.K. (2004) Staphylococcus aureus AgrA binding to the RNAIII-agr regulatory region. *J. Bacteriol.*, **186**, 7549–7555.
13. Rechtin,T.M., Gillaspy,A.F., Schumacher,M.A., Brennan,R.G., Smeltzer,M.S. and Hurlburt,B.K. (1999) Characterization of the SarA virulence gene regulator of Staphylococcus aureus. *Mol. Microbiol.*, **33**, 307–316.
14. Reyes,D., Andrey,D.O., Monod,A., Kelley,W.L., Zhang,G. and Cheung,A.L. (2011) Coordinated regulation by AgrA, SarA, and SarR to control agr expression in Staphylococcus aureus. *J. Bacteriol.*, **193**, 6020–6031.
15. Cheung,A.L., Nishina,K.A., Trotonda,M.P. and Tamber,S. (2008) The SarA protein family of Staphylococcus aureus. *Int. J. Biochem. Cell Biol.*, **40**, 355–361.
16. Reynolds,J. and Wigneshweraraj,S. (2011) Molecular insights into the control of transcription initiation at the Staphylococcus aureus agr operon. *J. Mol. Biol.*, **412**, 862–881.
17. Novick,R.P. (1991) Genetic systems in staphylococci. *Methods Enzymol.*, **204**, 587–636.
18. James,E.H., Edwards,A.M. and Wigneshweraraj,S. (2013) Transcriptional downregulation of agr expression in Staphylococcus aureus during growth in human serum can be overcome by constitutively active mutant forms of the sensor kinase AgrC. *FEMS Microbiol. Lett.*, **349**, 153–162.
19. Charpentier,E., Anton,A.I., Barry,P., Alfonso,B., Fang,Y. and Novick,R.P. (2004) Novel cassette-based shuttle vector system for gram-positive bacteria. *Appl. Environ. Microbiol.*, **70**, 6076–6085.
20. Wormann,M.E., Corrigan,R.M., Simpson,P.J., Matthews,S.J. and Grundling,A. (2011) Enzymatic activities and functional interdependencies of Bacillus subtilis lipoteichoic acid synthesis enzymes. *Mol. Microbiol.*, **79**, 566–583.
21. Gordon,J.C., Myers,J.B., Folta,T., Shoja,V., Heath,L.S. and Onufriev,A. (2005) H++: a server for estimating pKas and adding missing hydrogens to macromolecules. *Nucleic Acids Res.*, **33**, W368–W371.
22. Jorgensen,W.L., Chandrasekhar,J., Madura,J.D., Impey,R.W. and Klein,M.L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, **79**, 926–935.
23. Ponder,J.W. and Case,D.A. (2003) Force fields for protein simulations. *Adv. Protein Chem.*, **66**, 27–85.
24. Salomon-Ferrer,R., Case,D.A. and Walker,R.C. (2013) An overview of the Amber biomolecular simulation package. *WIREs Comput. Mol. Sci.*, **3**, 195–210.
25. Ryckaert,J.P., Ciccotti,G. and Berendsen,H.J.C. (1977) Numerical-integration of Cartesian equations of motion of a system with constraints - molecular-dynamics of N-alkanes. *J. Comput. Phys.*, **23**, 327–341.
26. Essmann,U., Perera,L., Berkowitz,M.L., Darden,T., Lee,H. and Pedersen,L.G. (1995) A smooth particle mesh Ewald method. *J. Chem. Phys.*, **103**, 8577–8593.
27. Pierce,L.C., Salomon-Ferrer,R., Augusto,F., de Oliveira C., McCammon,J.A. and Walker,R.C. (2012) Routine access to millisecond time scale events with accelerated molecular dynamics. *J. Chem. Theory Comput.*, **8**, 2997–3002.
28. Humphrey,W., Dalke,A. and Schulten,K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.*, **14**, 33–38.
29. Tsompanidou,E., Sibbald,M.J., Chlebowicz,M.A., Dreisbach,A., Back,J.W., van Dijl,J.M., Buist,G. and Denham,E.L. (2011) Requirement of the agr locus for colony spreading of Staphylococcus aureus. *J. Bacteriol.*, **193**, 1267–1272.
30. Markwick,P.R. and McCammon,J.A. (2011) Studying functional dynamics in bio-molecules using accelerated molecular dynamics. *Phys. Chem. Chem. Phys.*, **13**, 20053–20065.
31. Shopsin,B., Drlica-Wagner,A., Mathema,B., Adhikari,R.P., Kreiswirth,B.N. and Novick,R.P. (2008) Prevalence of agr dysfunction among colonizing Staphylococcus aureus strains. *J. Infect. Dis.*, **198**, 1171–1174.
32. Sully,E.K., Malachowa,N., Elmore,B.O., Alexander,S.M., Femling,J.K., Gray,B.M., DeLeo,F.R., Otto,M., Cheung,A.L., Edwards,B.S. *et al.* (2014) Selective chemical inhibition of agr quorum sensing in Staphylococcus aureus promotes host defense with minimal impact on resistance. *PLoS Pathogens*, **10**, e1004174.

Downloaded from http://nar.oxfordjournals.org/ at Imperial College London on March 17, 2015