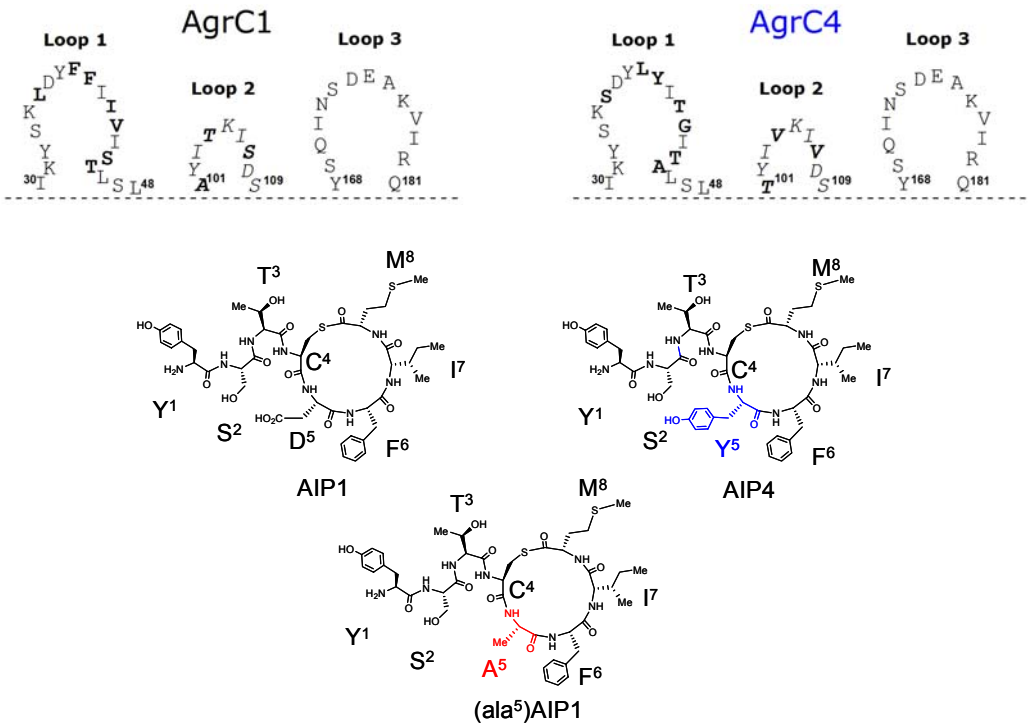
Functional analysis of the group specific interactions between AIP and AgrC in *Staphylococcus aureus*



Rasmus Overgaard Jensen, Cand. Polyt. (M.Sc.)

Submitted to The University of Nottingham for the degree of Doctor of Philosophy

September 2008

Declaration

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

Rasmus Overgaard Jensen

September 2008

Acknowledgements

First of all I would like to thank Institute of Infection, Immunity and Inflammation at University of Nottingham for funding. I would also like to thank Professor Paul Williams, Dr. Alan Cockayne, Dr. Weng Chan, Dr. Klaus Winzer and Dr. Simon R Clarke for supervision, scientific discussions and support through out my work. James Cottam, Rochelle Lamb and Saara Qazi also deserve mentioning for help in the laboratory and outside. Special thanks should go to Caroline Allen for putting up with me through hard months of writing and for catching the worst of the grammatical mistakes. And finally a big thanks to everybody else that has stood by me and given support through out the last four years, you know who you are, thank you

Dedicated to my mother, Christa Jensen

July 7th 1940 – January 12th 2008

Abbreviations

°C	degree centigrade
aa	amino acid
agr	accessory gene regulator
AHL	*N*-acylhomoserine lactone
AI-2	auto inducer 2
AIP	autoinducing peptide
Amp	ampicillin
BLAST	basic local alignment search tool
bp	base pair
BTH	bacterial two-hybrid system
CA-MRSA	community associated methicillin *staphylococcus aureus*
CHIPS	chemotaxis inhibitory protein
Cm	chloramphenicol
DCO	double cross over
dH_2O	demineralised water
DIG	digoxigenin
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
DTT	dithiothreitol
EC_{50}	median effective concentration (required to induce a 50% effect)
EDTA	ethylenediaminetetraacetic acid
Ery	erythromycin
FAME	fatty acid modifying enzyme
g	gram
Gfp	green fluorescent protein
gDNA	genomic DNA
h	hour

HA-MRSA	hospital associated methicillin *Staphylococcus aureus*
HPK	histidine protein kinase
Hpt	histidine phosphor transfer
IC_{50}	median inhibition concentration (required to reduce the effect by 50%)
Kan	kanamycin
kb	kilobases
kDa	kiloDalton
kV	kilovolt
L	liter
LB	Luria Bertani
M	molar
mA	miliampere
MALDI	matrix-assisted laser desorption/ionization
mg	miligram
min	minutes
ml	mililiters
mm	milimeters
mM	milimolar
mRNA	messenger ribonucleic acid
MRSA	methicillin *Staphylococcus aureus*
MSCRAMM	microbial surface components recognizing adhesive matrix molecules
ng	nanogram
o/n	over night
OD_{600}	optical density at a wavelength of 600 nanometers
ORF	open reading frame
PAGE	polyacrylamide gel electrophoreses
PBS	phosphate buffer saline
PBST	phosphate buffer saline tween20
PCR	polymerase chain reaction

PEG	polyethylene glycol
PVL	Panton-Valentine leukocidin
QS	quorum sensing
Qtof	quadropol time of flight
RLU	relative light units
RNA	ribonucleic acid
rpm	rounds per minute
RR	response regulator
s	seconds
SCC	staphylococcal cassette chromosome
SCO	single cross over
SCV	small colony variant
SDS	sodium dodecyl sulfate
TCS	two component system
TEMED	N,N,N',N'-tetramethylethylenediamine
Tet	tetracycline
TMH	transmembrane helix
TOF	time of flight
TSB	tryptic soy broth
TSST-1	toxic schock syndrome toxin 1
U	units
UV	ultraviolet
V	Volt
% vol/vol	percent volume per volume
% wt/vol	percent weight pr volume
X-gal	5-bromo-4-chloro-3-indolyl- beta-D-galactopyranoside
YMM	yeast minimal medium
YTH	yeast two-hybrid system
µF	microFarad
µg	microgram
µl	microliter

μm	micrometer
μM	micromolar

Table of Contents

Abstract

Staphylococcus aureus is an important human pathogen. The emergence of multiple antibiotic resistant bacteria and lack of new antibiotics has highlighted the need for better understanding of staphylococcal physiology, molecular biology and virulence. In *S. aureus* the *agr* quorum sensing (QS) system is a global regulator of virulence. In the *agr* system an autoinducing peptide (AIP) activates the histidine protein kinase (HPK), AgrC, leading to a switch from the production of colonization factors to exotoxins. The *S. aureus agr* system has diverged such that there are four different *agr* groups, each with a distinct AIP capable of activating its cognate AgrC but inhibiting the AgrC of the other groups. To investigate the molecular basis for the recognition of AIPs by AgrC, transmembrane topology modelling together with site-specific mutagenesis were used. The transmembrane topology of AgrC was predicted to consist of six transmembrane helices (TMHs) and three extracellular loops with both the N- and C-terminals on the cytoplasmic side. Since AIP-1 and AIP-4 differ by a single amino acid residue, the *S. aureus* AgrC1 and AgrC4 proteins were compared to identify extracellular amino acids likely to be involved in AIP recognition. Site-specific mutagenesis was used to exchange the key AgrC1 and AgrC4 amino acid residues in extracellular loops 1 and 2. The response of these AgrC mutants was evaluated using a novel bioluminescent AIP reporter. The data obtained showed that differential recognition of AIP-1 and AIP-4 depend primarily on three amino acid residues in loop 2, but that loop 1 plays an essential for activation but not for inhibition of AgrC. The data obtained also revealed that a single mutation in the AgrC1 loop2 results in conversion of (ala^5)AIP-1 from a potent antagonist to an activator, essentially forcing the evolution of a fifth *agr* group.

Attempts to identify AgrC in the cytoplasmic membrane using Western blotting failed, but data obtained using an N-terminal *gfp* tag showed that AgrC is evenly distributed through out the membrane.

Since the processing of AgrD by AgrB to generate an AIP requires at least 3 steps - two endopeptidase steps and the formation of a thiolactone bond to form the macrocycle, it is likely that other proteins are involved in the processing of AgrD and export of the AIP. To identify potential AgrB partners, yeast two hybrid assay was employed which revealed a potential role for the putative ABC transporter Rlp in the processing and/or secretion of AIP.

In summary, the data presented define the key amino acid residues involved in AIP/AgrC interactions and imply a role for proteins such as Rlp in AIP synthesis and export.

1 General Introduction

1.1 Staphylococci

Staphylococcus aureus is a coagulase positive member of the genus *Staphylococcus* which comprises of low G + C (30-40 mol%), non motile, non spore forming Gram-positive cocci with a diameter of approximately 1 μm belonging to the family *Staphylococcaceae*. The word staphylococcus is derived from the Greek word staphylos meaning a "bunch of grapes" and refers to the way that *S. aureus* divide in more than one plane, forming clusters resembling grape bunches under the light microscope (Kloos W.E., 1998).

The staphylococci are usually facultative anaerobes, growing better under aerobic than anaerobic conditions, capable of both respiration and fermentation. The cell wall of staphylococci contains both peptidoglycan and teichoic acid. The diamino acid in the peptidoglycan is L-lysine and the interpeptide bridge consists of oligoglycine (depending on the species some glycine may be substituted by L-serine or L-alanine). The genus is susceptible to lysis by lysostaphin, which cleaves the interpeptide glycine bridge, and is relatively resistant to lysozyme. Most species grow in the presence of 10% wt/vol sodium chloride over a temperature range of 10ºC to 40ºC (Kloos W.E., 1998).

Over 30 species of staphylococci have been identified (Kloos W.E., 1998) and they can be divided into coagulase-positive (e.g.,, *S. aureus*) and coagulase-negative (e.g.,, *S. epidermidis*) strains. The natural habitat of the staphylococci habitats includes the nasal cavity (anterior nares and pharynx), skin, mammary glands, genitourinary and respiratory tract of their host (Kloos W.E., 1998). About 30% of the human population are permanent or intermittent carriers of *S. aureus* in their nasal cavities (Wertheim *et al.*, 2005) and generally have a benign relationship with their hosts. However if the host is damaged by trauma, or the host's immune system is weakened the staphylococci can gain access to the host tissue and adopt the lifestyle of a pathogen (they are opportunistic pathogens).

1.2 *Pathogenicity of S. aureus*

S. aureus is a very successful opportunistic pathogen, its importance as a human pathogen was first described in 1884 by the Scottish surgeon Alexandra Ogston who discovered *S. aureus* in pus from surgical abscesses (Ogston, 1984). *S. aureus* is responsible for a wide range of diseases ranging from skin infections like boils, furuncles and cellulitis which are usually causes only a mild discomfort and the hosts immune system prevents the infections from becoming more severe (Lowy, 1998). More serious and potentially life threatening infections like bacteraemia, necrotising pneumonia and toxic shock syndrome are usually associated with people with multiple risk factors, e.g., trauma victims, drug

abusers, the new born and burn victims and patients with implanted medical devices (Lowy, 1998). Staphylococcal enterotoxins are also frequently associated with bacterial food poisoning.

Being colonised with *S. aureus* provides a reservoir for infection and subsequently increases the risk of infection (Wertheim *et al.*, 2005;Kluytmans *et al.*, 1997)and most *S. aureus* infections are due to already the present strain (Williams *et al.*, 1959;Gordon and Lowy, 2008). Colonisation also provides a source of *S. aureus* spreading to both health care and community settings.

The first step in infection is breach of the skin or mucous barriers allowing *S. aureus* access to the adjoining tissue and bloodstream. *S. aureus* establishes itself by binding to molecules such as collagen, fibronectin and fibrinogen. This is mediated by a group of cell surface proteins called "microbial surface components recognizing adhesive matrix molecules" (MSCRAMMs, Table 1-1 and Figure 1-1) (Patti *et al.*, 1994;Foster and Hook, 1998) and other adhesions, e.g., isdA. These adhesins are also responsible for the establishment of *S. aureus* infection on surgical implants such as catheters following conditioning with molecules such as fibrinogen. Once *S. aureus* has established itself in the tissue it is able to grow and evade the immune system. Biofilm formation on host and prosthetic surfaces renders it less susceptible to antimicrobials and facilitates evasion of host defences (Donlan and Costerton, 2002).

The formation of small-colony variants (SCVs, these are a slow growing subpopulation with a distinct phenotype and pathogenic traits able to revert to the more virulent wild-type phenotype) is another way of evading host defences and may contribute to persistent and recurring staphylococcal infections (Proctor and Peters, 1998;Kahl *et al.*, 1998;Proctor *et al.*, 1995).

IsdA is an iron-regulated surface protein which plays multiple roles in the virulence of *S. aureus.* IsdA binds fibrinogen and fibronectin , plays an important role in nasal colonisation (Clarke *et al.*, 2004;Taylor and Heinrichs, 2002) and protects the cell against apolactoferrin (Clarke and Foster, 2008), a human iron-binding glycoprotein protein with bactericidal properties. It also confers resistance against the free fatty acids and antimicrobial peptides making up a part of the innate defence of the skin by making the cells more hydrophilic (Clarke *et al.*, 2007).

The production of microcapsules (most clinical strains are sero type 5 or 8) produced via the *cap5* and *cap8* gene clusters (O'Riordan and Lee, 2004) is another way *S. aureus* can evade the host immune system. *S. aureus* also secrete a chemotaxis inhibitory protein and produces an extracellular adherence protein, which interfere with the recruitment of neutrophils to the site of infection. In addition to evading the immune system, *S. aureus* also produces leukocidins that destroy leukoctyes by forming pores in their membranes (Foster, 2005) (Table 1-1 and Figure 1-1).

After the establishment of an infection, *S. aureus* produces numerous exoproteins that enable it to invade and destroy host tissue and migrate to other sites within the host. By interacting with and activating the host immune system and coagulation pathways, *S. aureus* is capable of inducing septic shock (Timmerman *et al.*, 1993;Heumann *et al.*, 1994). The production of so called superantigens, e.g., TSST-1 (Toxic shock syndrome toxin) and Enterotoxins A-J, cause a sepsis-like syndrome by initiating a positive feedback loop with the immune system and thereby over activating it, creating a so called "cytokine storm" (McCormick *et al.*, 2001). *S. aureus* has also been reported escape the endosome from both professional and non-professional phagocytes and start replicating inside the cell, which could allow it to escape the immune system and serve as a reservoir for reoccurring infections (Qazi *et al.*, 2001;Ogawa *et al.*, 1985;Drake and Pang, 1988;Hamill *et al.*, 1986).

It should be mentioned that all the different virulence factors used in invading, evading and combating host defences are not necessarily present in all *S. aureus* strains. Hence, different strains vary in there ability to form biofilms, evade host defences, resist phagocytosis and induce toxic shock. Table 1-1 provides an overview of virulence factors, their frequency of occurrence and their regulation.

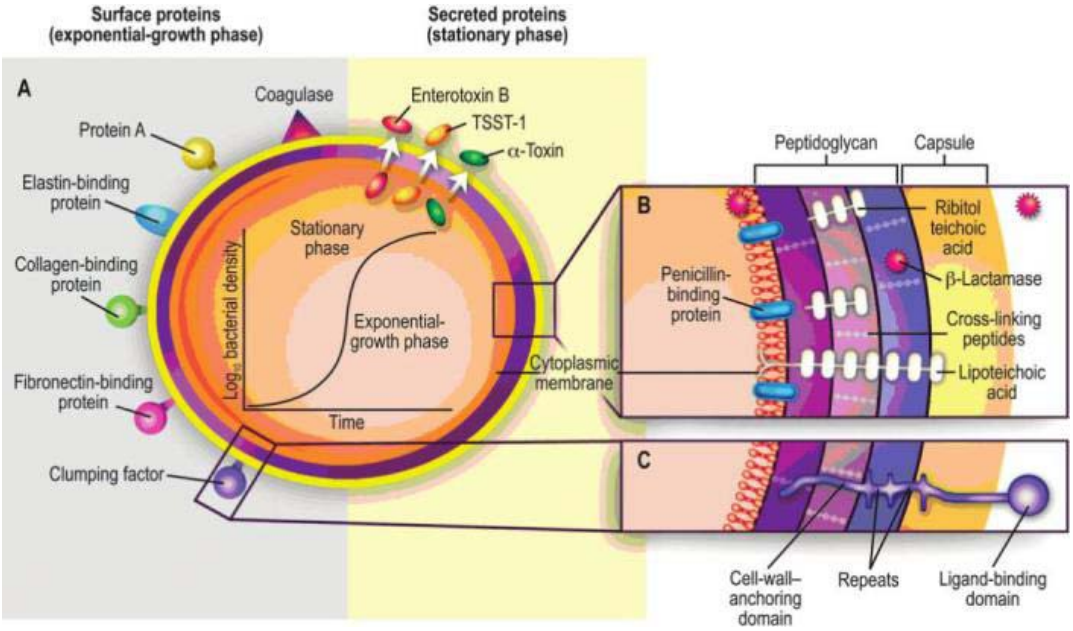


Figure 1-1: *S. aureus* surface proteins and exoproteins, their times of expression and location. In an infection the lag- and exponential-phase would be the colonisation phase and the stationary phase can be compared to an abscess or another site of infection where the bacteria will accumulate. Reproduced from (Gordon and Lowy, 2008)

Table 1-1: Virulence determinants regulated by SarA and *agr*, their function and their distribution in *S. aureus*. This is by no means a complete list but gives an overview of the arsenal of virulence factors at the disposal of *S. aureus*. Modified from (Peacock *et al.*, 2002) and (Cheung *et al.*, 2004). N/A = unknown; N/E = No effect; + = upregulated by ; - = down regulated by.

Name	Distribution (%)	Putative function	*agr*	SarA
Exoproteins				
Aureolysin	100	Metalloprotease	-	+
Enterotoxin B	7-9	Exotoxin with superantigen activity	+	+
Enterotoxin C	10-11	Exotoxin with superantigen activity	+	N/A
α-hemolysin	99-100	Cytolytic pore-forming toxin	+	+
β-hemolysin	79-89	Tissue invasion, Sphingmyelinase	+	+
γ-hemolysin	89-97	Potentiation of host ccell lysis, Bicomponent leukocidin	+	+
δ-hemolysin	69-80	Potentiation of β-hemolysin, Cytolytic toxin	+	+
Exfoliatins A and B	6-22, 0-3	Exotoxins with superantigen activity	+	N/A
FAME		Fatty acid modifying enzyme	+	+
HysA, (Hyaluronate lyase)		Tissue invasion	+	N/A
Lipase		Evasion of host defences	+	-
Phospholipase C		Lysis of host cells	+	N/A
Set8		Superantigen-like protein	+	N/A
Set9		Superantigen-like protein	+	N/A
SplA, B, D and F		Proteases	+	+
SspB		Cysteine protease	N/A	-
Staphylokinase		Evasion of host defences	+	N/A
TSST-1		Exotoxin with superantigen activity	+	+
V8 protease (serine protease)		Tissue invasion and modification of surface proteins	+	-
Panton valentine leukocidin		Bicomponent leukocidin	+	N/A
CHIPS		Chemotaxis inhibitory protein		
MSCRAMMs				
Clumping factor A	98-100	Adhesin to fibrinogen	N/E	N/A
Clumping factor B	100	Adhesin to fibrinogen	N/E	N/A
Coagulase		Binding to fibrinogen	-	+
FibA protein		Binding to fibrinogen		
Collagen-binding protein	32-52	Adhesin to collagen	±	-
Capsular polysaccharide type 5 and 8	16-26,55-65	Anti-phagocyte molecules	+	+
Fibronectin-binding protein A	87-98	Adhesin for fibronectin	-	+
Fibronectin-binding protein B		Adhesin for fibronectin	-	+
Protein A	90-94	Binds Fc domain of immunoglobulin and von Willibrand factor	-	-
SdrC	100	Putative adhesin	N/A	+
Vitronectin-binding protein			-	N/A
Map/Eap	93-96	MHC ClassII analogous / eltracellular adherence proteins		
Other agr /sar regulated factors				
PBP3		Penicillin binding protein	N/A	-
FemA		methicillin resistance expression factor	N/A	+
FemB		methicillin resistance expression factor	+	N/A

The growth phase dependent regulation of virulence factors in *S aureus* seen *in vitro*, where adhesins are expressed before haemolysins, proteases and other exo-proteins, have also been shown to occur *in vivo* (Novick, 2003). The importance of *agr* in animal models has been demonstrated in various studies. Thus, strain 8325-4 *Δagr* was demonstrated to be significantly less virulent than its wild type 8325-4 parent strain in a mouse arthritis model (Abdelnour *et al.*, 1993). Cheung *et al* (1994) later showed that *agr* also played an important role in a rabbit endocarditis model using RN6390 with 4 out of 11 animals developing edocarditis for *Δagr* compared to 9 out of 10 for the wild type strain at an inoculation dose of 10^4 CFU. A significant role for *agr* in a rabbit osteomyelitis model both at high ($\geq 2x10^6$) and at low dose ($\leq 2x10^5$ CFU) inocula, was also shown by the study of (Gillaspy *et al.*, 1995). However Schwan *et al.* (2003) compared different murine models and found that agr^- strains are attenuated in the abscess and systemic models, but not in the wound model.

1.2.1 MRSA

With the introduction of penicillin in the 1940s, treating *S. aureus* infections which previously were difficult to treat became routine. By the late 1940s and throughout the 1950s *S. aureus*, aided by selection pressure by the use, overuse and missuse of antibiotics, became resistant to penicillin. Methicillin was introduced in 1959 to treat penicillin resistant *S. aureus* infections but by 1961 the first case of methicillin resistant *Staphylococcus aureus* (MRSA) was reported in

the United Kingdom (Jevons, 1961), soon thereafter MRSA was reported in other European countries, Japan, Australia and USA. MRSA is now one of the leading causes of hospital acquired infections in the UK and most of the rest of the western world. In 2007 alone MRSA was mentioned on the death certificates of 1,593 patients in England and Wales (2008).

Resistance to methicillin is conferred by the *mecA* gene which is a part of a mobile genetic element called "staphylococcal cassette chromosome" (SCC). To date five different SCC*mec*'s of different size and composition have been identified (type I-V) (Witte *et al.*, 2008). The SCC*mec* is capable of horizontal gene transfer and is believed to have been transferred to *S. aureus* from a coagulase negative staphylococcus (Wisplinghoff *et al.*, 2003).

MRSAs have been divided into two groups; hospital associated (HA-MRSA) and community associated (CA-MRSA). HA-MRSA infections are generally resistant to multiple antibiotics.

In 1989 and 1991, the first cases of CA-MRSA were reported in the Australian indigenous community (Udo *et al.*, 1993) and in 1999 several cases of aggressive MRSA infections occurred in USA (Gosbell *et al.*, 2001), all without any established risk factors or prior exposure to a healthcare setting. Since then cases of CA-MRSA skin infections and necrotising pneumonia have been reported in

other countries (Vandenesch *et al.*, 2003). Most CA-MRSA strains carry the truncated type IV SCC*mec* and are susceptible to a range of antibiotics. The presence of the Panton-Valentine Leukocidin (PVL) has been associated with CA-MRSA, but its contribution to disease is still a highly debated subject.

1.3 Quorum sensing

The idea that bacteria can work together is not a new one, as far back as 1905 E.F. Smith stated: "I think that the multiple of bacteria are stronger than a few and thus by union are able to overcome obstacles too great for the few" (Smith EF, 1905). But it was not until 1962, that McVittie published on formation of fruiting bodies in *Myxococus Xanthus* (McVittie A et al., 1962), and presented evidence that a bacterial system utilising "chemotactic substance" was used for cell-cell communication and cooperation. These substances were first named pheromones by Karlson and Lusher (1959) (Karlson and Luscher, 1959). Nealson et al (1970) not only showed that *Vibrio fischeri* only produced light at high cell population density, but also the bacteria produced an extracellular substance that could induce bioluminescence at low cell densities. This is generally considered the discovery of quorum sensing (QS) (Nealson *et al.*, 1970), a term that was not used until 1994 (Fuqua *et al.*, 1994). The secreted chemical was purified and shown to be an *N*-acylhomoserine lactone (AHL) (Eberhard *et al.*, 1981). During the 1990s pheromones were shown to be produced by a wide range of Gram-negative

bacteria including *Erwinia carotovora, Pseudomonas aeruginosa* and *Serratia marcescens* (Bainton et al., 1992b;Bainton et al., 1992a). Since the discovery of QS it has been found to control many other properties, other than bioluminescence, such as swarming, competence, sporulation, secondary metabolism, expression of virulence factors and various other processes (Jones *et al.*, 1993;Passador *et al.*, 1993;Swift *et al.*, 1993;Zhang *et al.*, 1993).

For a molecule to be classified as a QS signal molecule four criteria have to be met: 1) the production of the molecule is dependent on growth phase, physiological conditions, or environmental changes; 2) The molecule accumulates extracellularly and is recognised by a specific receptor; 3) A clear cellular response is seen when the threshold concentration is reached and 4) The response is not just needed to metabolize or detoxify the QS signal molecule (Winzer *et al.*, 2002b).

Besides AHLs, 4-quinolones, fatty acids and fatty methyl esters have been shown to be QS molecules in Gram-negative bacteria. In Gram-negative bacteria the small organic signal molecules generally diffuse across the cytoplasmic membrane and bind to an intercellular receptor molecule, the response regulator. The activated response regulator, either directly or indirectly, regulates the expression of the target genes (Fuqua *et al.*, 2001;Swift *et al.*, 2001) (Figure 1-2).

Gram-positive bacteria appear to preferentially use post translationally modified peptides as QS molecules. Three classes of these molecules are currently known: 1) Oligopeptide lantibiotics (e.g., lactococcal nisin) (van, Jr. *et al.*, 1993;Quadri,

2002) 2) sixteen membered thiolactone peptides (ring structure of five amino acids with or with out a “tail”, e.g., The staphylococcal autoinducing peptides (AIPs)) (Ji *et al.*, 1997;McDowell *et al.*, 2001;Chan *et al.*, 2004) 3) and isoprenylated tryptophan peptides (e.g., *Bacillus subtilis* ComX) (Ansaldi *et al.*, 2002;Okada *et al.*, 2005).

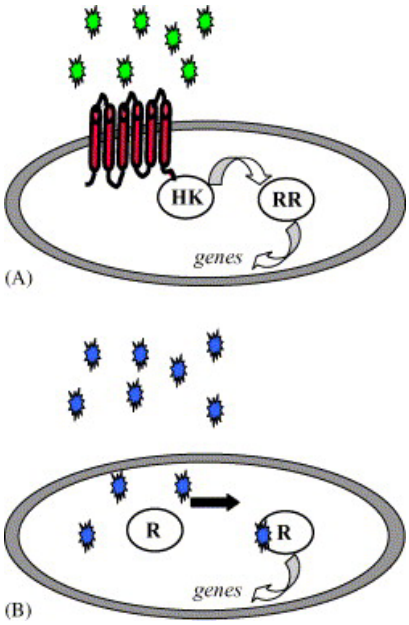


Figure 1-2: Quorum sensing in A) Gram-positive bacteria. The classic two component regulator system where the signal molecule binds to a sensor histidine kinase which activates a response regulator which then regulates the expression of the target genes (either directly or indirectly), and B) Gram-negative bacteria where the signal molecule binds directly to the receptor molecule which then regulates the expression of the target genes. Figure from (Lyon and Novick, 2004).

AIPs are secreted and activate a two component system by binding to a transmembrane receptor kinase leading to the phosphorylation of a response regulator. The activation of the two component system leads to the regulation of the target genes via an intracellular response pathway (Figure 1-2).
Streptomyces sp. uses γ-butyrolactones for signalling (Takano *et al.*, 2000) and is so far the only Gram-positive bacteria that has been found not to use peptides.

The QS molecule AI-2 (furanosyl borate diester) is a QS molecules signalling molecule used by some gram-positive and gram-negative bacteria (Winzer *et al.*, 2002a;Xavier and Bassler, 2003).

1.4 Two-component systems

The staphylococcal quorum sensing system (*agr*), in common with many other QS systems, employs two-component systems to relay the external stimulus to the cytoplasm. Two-component systems (TCS) consist of a membrane bound sensor histidine kinase (HK) and a response regulator (RR). The histidine kinase of the TCS is activated by an external stimulus which leads to the phosphorylation of the response regulator. The phosphorylated RR regulates its target gene by binding to promoter / operator DNA.

Histidine protein kinases are divided into two different classes depending on the organization of their functional domains. Class I type HPKs are the most common and contain an N-terminal membrane spanning sensor domain of variable length. The C-terminal domain makes up the so-called transmitter domain, which contains motifs (designated H, N, G1, F and G2) that are conserved in the histidine kinase super family. A variant of this canonical type displays both a phosphor-accepting and a histidine phosphor-transfer domain (HPt) fused to the C-terminal end (Foussard *et al.*, 2001). The class II HK's are solely involved in chemotaxis responses (Foussard *et al.*, 2001), and are not really relevant to the description of the *agr* QS locus.

1.5 The agr locus

A common QS system for many Gram-positives bacteria is the *agr* system and has been found not just in staphylococci but also in clostridia, lactocbacillus, bacillus, listeria, alkaliphilus, desulfitobacterium, moorella, ruminococcus, syntrophomonas, thermoanaerobacter (Winzer K, personal communication) and possibly more.

In *S. aureus* the production of surface adhesin proteins and secreted virulence factors are primarily under the control of the *agr* operon, which consists of two transcripts, RNAII and RNAIII. The RNAII mRNA contains four open reading

frames (ORFs), coding for AgrB, AgrD, AgrC and AgrA. RNAIII is a regulatory RNA molecule and is the effector molecule of the *agr* QS system. The regulatory RNA molecule RNAIII also contains a small ORF, *hld*, coding for δ-haemolysin (Janzon *et al.*, 1989).

AgrCA is a classic two component system, where the *agrC* gene codes for the receptor histidine kinase, and *agrA* codes for the response regulator. The *agrB* gene codes for a transmembrane cysteine protease that is involved in processing the AgrD propeptide into a mature autoinducing peptide (AIP), though the exact mechanism of this is still unclear. Type 1 staphylococcal signal peptidase (SpsB) is also thought to be involved by cleaving off the N-terminal leader of AgrD (Kavanaugh *et al.*, 2007). The AIP is predicted to bind to and activate AgrC which leads to phosphorylation of AgrA. AgrA in its phosphorylated state binds to the region between the two promoters and activates transcription from both P2 and P3 (Koenig *et al.*, 2004) in conjunction with SarA. Activation of transcription from the P2 promoter leads to a positive feedback loop. The P3 promoter drives the transcription of RNAIII, the effector molecule of the *agr* system (Novick *et al.*, 1993;Janzon and Arvidson, 1990). Transcription of RNAIII results in a decrease in the production of surface factors and an increase in the production of other extracellular virulence factors (Novick *et al.*, 1993) (Figure 1-3).

The *agr* locus contains a hypervariable region spanning the region from mid *agrB* over *agrD* and into *agrC* (Figure 1-3) (Ji *et al.*, 1997). The variability in this

region's primary amino acid sequence leads differences in the of the N-terminal region of AgrB, AgrD and the C-terminal domain of AgrC. This in turn leads to the production of four different AIPs and their cognate AgrCs. Based on the structure of these AIPs *S. aureus* can be divided into four *agr* groups (I-IV) (Ji *et al.*, 1997;McDowell *et al.*, 2001). Each AIP specifically activates their native AgrC and is an antagonist for the other groups AgrC; this is known as cross inhibition. AIP-1 is however a weak activator of AgrC4 and AIP-4 is a moderate activator of AgrC1 (Lyon *et al.*, 2002).

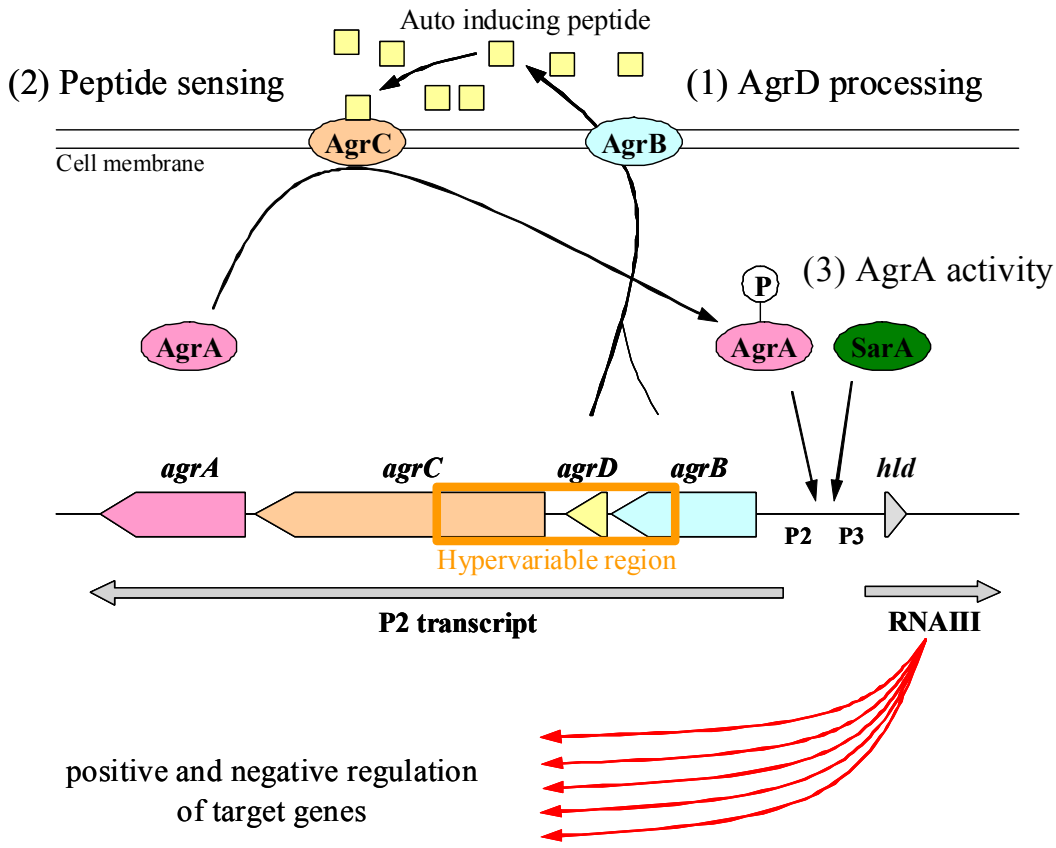


Figure 1-3: The *agr* locus and *agr* activation. Orange box shows the hypervariable region. (1) AgrB processes AgrD into active AIP (2) The AIP activates AgrC and AgrA is phosphorylated (3) P-AgrA binds to the *agr* promoter P2 leading to a positive feedback loop and P3 leading to transcription of RNAIII and *hld*.

1.5.1 AgrB

AgrB is a membrane cysteine peptidase required for processing of AgrD to generate active AIP (Ji *et al.*, 1995;Ji *et al.*, 1997). AgrB cleaves off the C terminal tail of AgrD (Qiu *et al.*, 2005). It is predicted, by a combination of hydrophobicity and PhoA fusion analysis, to contain six transmembrane domains. Four of these domains are hydrophobic and two are hydrophilic containing highly positively charged amino acid residues which are involved in the proteolytic cleavage of AgrD (Zhang *et al.*, 2002).

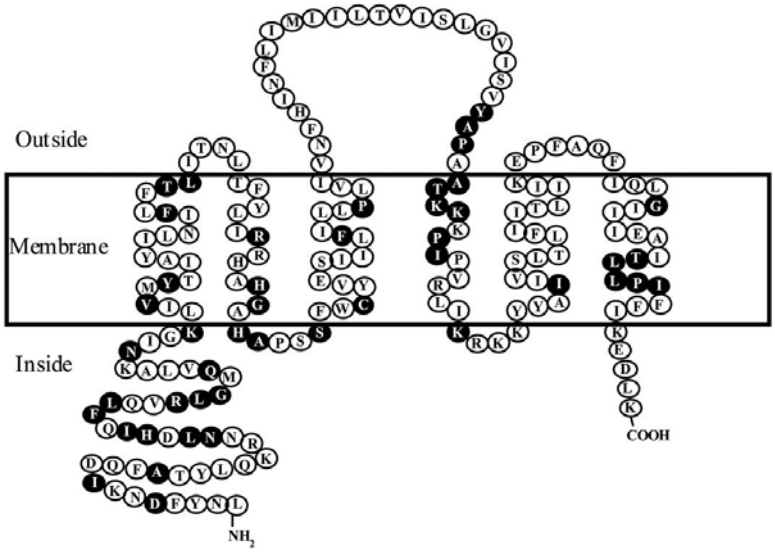


Figure 1-4: The two dimensional structure of AgrB1. Amino acids conserved among the different *agr* groups are marked in black (Qiu *et al.*, 2005).

The group specific processing of AgrD by AgrB is thought to be less stringent than that of activation of AgrC by AIP , with AgrB1 and AgrB3 being able to process both AgrD1 and AgrD3 but not AgrD2 (Ji *et al.*, 1997). The regions responsible for the group specific processing of AgrD for AgrB1 and AgrB have been identified. For AgrB1 the first transmembrane helix (TMH) and the first extracellular loop are important, whereas for AgrB2 it is two hydrophilic segments consisting of M67 to G75 (TMH2) and A126 to K141 (TMH4 together with cytoplasmic loop 2) (Zhang and Ji, 2004). It has been shown that alanine substitution of the histidine residue at position 77 (H77) and cysteine residue at position 84 (C84) totally eliminate the ability of AgrB1 to process AgrD1 to form the mature AIP (Qiu *et al.*, 2005).

1.5.2 AgrD

AgrD is processed by AgrB, and there is some indication that the signal peptidase 1 SpsB is involved (Kavanaugh *et al.*, 2007), to yield AIP (Ji *et al.*, 1995). This involved 3 steps, two endopeptidase steps and one cyclization (Figure 1-5).

For AgrD to be processed it has to be anchored in the membrane via the N-terminal amphipathic leader (Zhang *et al.*, 2004) where upon the C-terminal tail is cleaved off by AgrB (Zhang and Ji, 2004) this step is believed to be the catalyst for the formation of the thiolactone ring. The AIPs derived from the different AgrDs are all agonists for their cognate AgrCs and cross inhibit the other AgrC

groups (Ji *et al.*, 1997;McDowell *et al.*, 2001), with the exception of AIP-4 which is also a moderate activator of AgrC1 and AIP-1 which is a weak activator of AgrC4 (Lyon *et al.*, 2002), this may partially be due to the fact that AIP-1 and AIP-4 only differ by one amino acid residue at position 5, aspartic acid in AIP-1 and tyrosine in AIP-4. Structure function analysis of AIPs has shown that if the aspartic acid in position 5 of AIP-1 is changed to an alanine the resulting AIP derivitive, (ala^5)AIP-1, is an inhibitor of all four *S. aureus agr* groups (McDowell *et al.*, 2001).

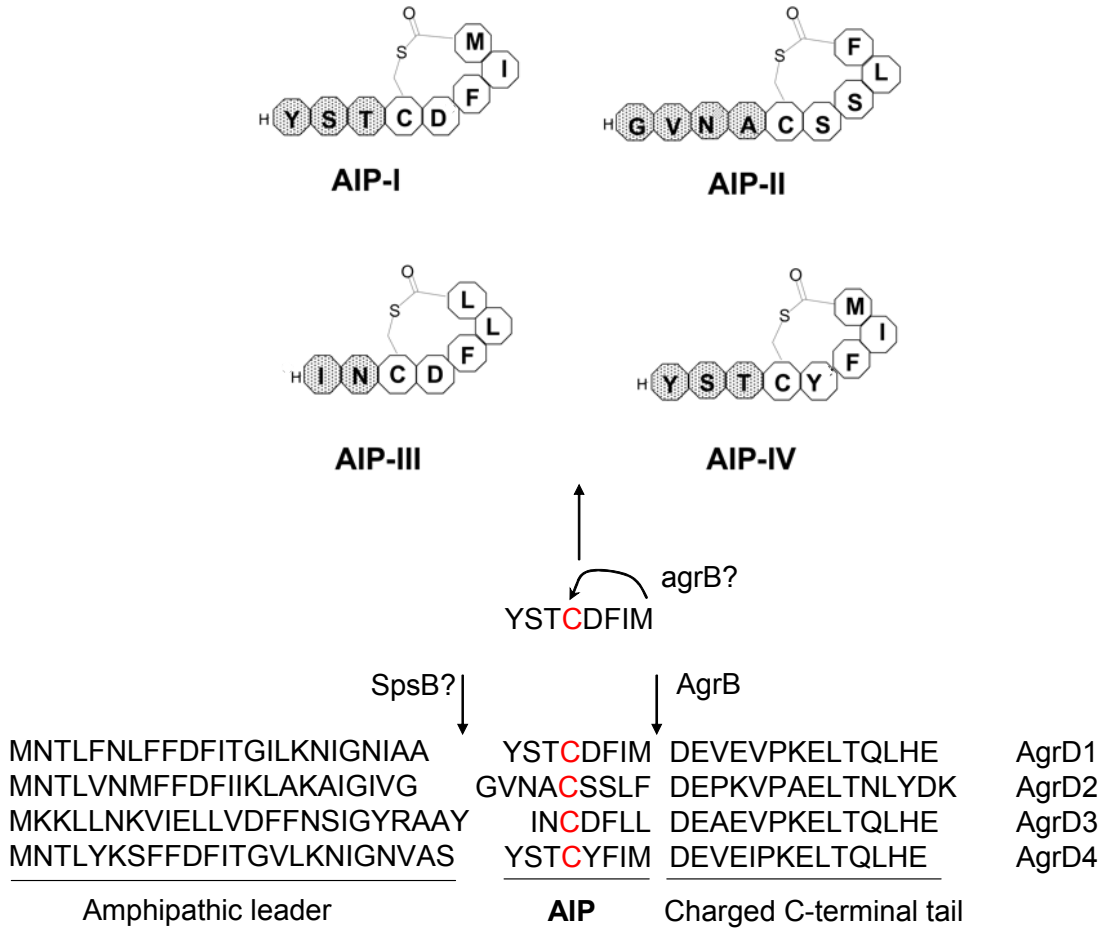


Figure 1-5: The domains and primary sequences of AgrD1-4. In order to yield active AIP two endopeptidase reactions and the formation of a thiolactone bond between the central cysteine (marked in red) and the C-terminal amino acid residue of the AIP, yielding the macrocycle. Note: AIP-1 and AIP-4 are closely related and only differ by one amino acid residue at position 5 (D to Y). Modified from (Lyon *et al.*, 2002).

1.5.3 AgrC

AgrC is a member of the class 10 receptor histidine protein kinases (HPKs) (Grebe and Stock, 1999) and consists of an N-terminal transmembrane sensor domain, that is a part of the *agr* hypervariable region (Figure 1-3), and a C-terminal HPK domain (Lina *et al.*, 1998) that is highly conserved among the staphylococci. The N-terminal sensor domain is responsible for the group specific recognition of AIPs (Wright, III *et al.*, 2004). AgrC is predicted to be found in the membrane as a dimer, but the structure of AgrC has yet to be solved. Lina, *et al* (1998) suggested, on the basis of hydrophobicity analysis and PhoA fusions, that it consisted of 5 transmembrane helices (Lina *et al.*, 1998). Later *in silico* prediction, however, suggests that AgrC consists of six TMHs and three outer loops (Geisinger *et al.*, 2008), a topology model that is also supported by findings using PhoA fusions (Lina *et al.*, 1998). These extracellular loops are found in the N-terminal sensor domain of AgrC. Chimeric studies of AgrC have indicated that the recognition specificity is located in the C-terminal part of the sensor domain (Wright, III *et al.*, 2004).

1.5.4 AgrA

AgrA is a member of the LytR response regulator family and when phosphorylated, binds to a consensus sequence consisting of two direct repeats (TA AC CA GTTN AG TG), that are separated by a 12-13 bp spacer region (Nikolskaya and Galperin, 2002). Two of these elements have been identified in

the P2-P3 intergenic region of the *agr* operon and RNAIII, one in the P2 promoter and one in the P3 promoter. AgrA binds to the P2 promoter with a higher affinity than to the P3 promoter (Koenig *et al.*, 2004) hence; auto-activation of the P2 promoter precedes P3 activation (Cheung *et al.*, 2004). The regulator SarA interacts with, or is displaced by AgrA and enhances AgrA induced transcription from both promoters in an as yet unknown manner (Cheung *et al.*, 1997;Rechtin *et al.*, 1999;Morfeldt *et al.*, 1996). The crystal structure of AgrA bound to DNA has recently been solved and has revealed a novel structure of a 10 strand β-fold that intercalates. Unlike other major groove DNA binding proteins, AgrA binds to a single face on the DNA, covering nearly 16 bp, inserting long loops into the major grooves making direct single base contacts, Figure 1-6 (Sidote *et al.*, 2008).

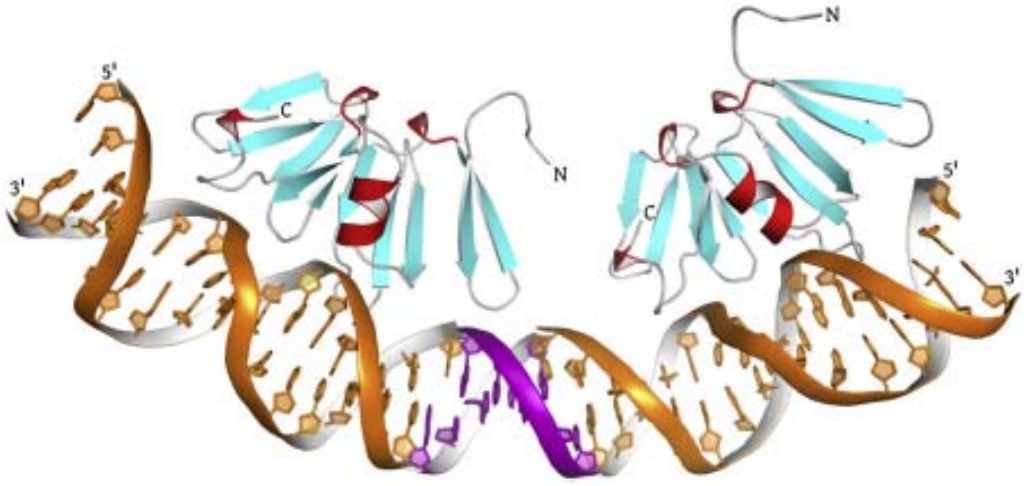


Figure 1-6: Dimer of the C-terminal DNA binding domain of AgrA bound to DNA. (Sidote *et al.*, 2008)

1.5.5 RNAIII

RNAIII is a 514 nucleotide regulatory RNA molecule and is the effector molecule of the *agr*-operon; it has a highly conserved secondary structure, although the primary sequence varies, among staphylococcal species (Novick *et al.*, 1993). RNAIII is a highly abundant, very stable RNA molecule with a half life of more than 45 min (Huntzinger *et al.*, 2005) that folds into a complex secondary structure (Benito *et al.*, 2000). Besides being a regulatory RNA, RNAIII also contains a short CDS containing the *hld* gene encoding δ-haemolysin (Janzon *et al.*, 1989) (Figure 1-7).

The exact mechanism of how RNAIII works is not yet fully understood but it is believed that functions as an antisense molecule, regulating translation of both individual exoproteins / toxins and pleiotropic regulators (e.g., SarT, SarS and Rot) (McNamara *et al.*, 2000;Tegmark *et al.*, 2000). RNAIII directly down regulates translation of protein A (Novick *et al.*, 1993;Benito *et al.*, 2000), and up regulates translation of α-haemolysin by binding to the *hla* mRNA by hindering the formation of a secondary translation blocking structure (Morfeldt *et al.*, 1995) (Figure 1-7).

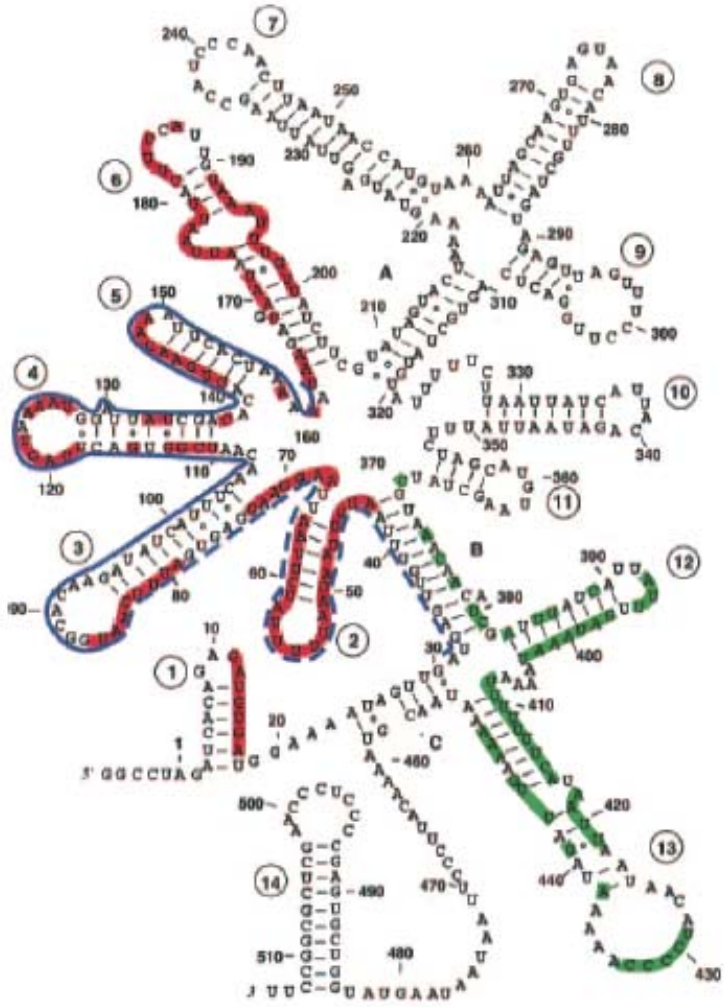


Figure 1-7: 2D structure of the functional molecule RNAIII (Benito *et al.*, 2000;Novick, 2003). The *hld* coding region is highlighted in solid blue and potential upstream translatable region in dashed blue. Red shading shows sequence that is complementary with the *hla* mRNA. Green shading shows sequence complementary with *spa* mRNA translation initiation region.

1.5.6 Role of *agr In Vivo*

The growth phase dependent regulation of virulence factors in *S aureus* seen *in vitro*, where adhesins are expressed before haemolysins, proteases and other exo-proteins, have also been shown to occur *in vivo* (Novick, 2003). The importance of *agr* in animal models has been demonstrated in various studies. Thus, strain 8325-4 *Δagr* was demonstrated to be significantly less virulent than its wild type 8325-4 parent strain in a mouse arthritis model (Abdelnour *et al.*, 1993). Cheung *et al* (1994) later showed that *agr* also played an important role in a rabbit endocarditis model using RN6390 with 4 out of 11 animals developing edocarditis for *Δagr* compared to 9 out of 10 for the wild type strain at an inoculation dose of 10^4 CFU. A significant role for *agr* in a rabbit osteomyelitis model both at high ($\geq 2x10^6$) and at low dose ($\leq 2x10^5$ CFU) inocula, was also shown by the study of (Gillaspy *et al.*, 1995). However Schwan *et al.* (2003) compared different murine models and found that agr^- strains are attenuated in the abscess and systemic models, but not in the wound model.

The results from various animal models seem to indicate that *agr* plays an important role *in vivo*; also most clinical isolates of *S. aureus* are agr^+ , but the occurrence of strains which either express *agr* poorly or not at all (Fowler, Jr. *et al.*, 2004;Sakoulas *et al.*, 2002) seems to raise questions about the role of *agr* in human disease. Traber *et al* (2008) carried out a study of 146 clinical isolates which showed that 15 % were non-haemolytic (often used as an *agr* indicator), a few of these could be complemented with WT *agr* but some were not

complementable; indicating that factors outside the *agr* locus are responsible for the phenotype. They also showed that the *agr*$^-$ strains isolates are most likely mutants that have appeared during the infection (Traber *et al.*, 2008). This is supported by a*gr* having been shown to be important during early stages of an infection (Wright, III *et al.*, 2005b).

Using a mouse abscess model Wright *et al.* (2005a) showed that *agr* was activated a few hours after injection of *S. aureus.* This burst lasted for a few hours and then subsided again, but a second burst of *agr* activity was observed approximately 72 hours after injection (Wright, III *et al.*, 2005a). This cessation of activity and the second burst can not be explained by the traditional model for *agr* activation but several environmental factors also influence the *agr* response. Low oxygen concentration inhibits the expression of RNAIII via the staphylococcal respiratory response (SrrAB) (Pragman *et al.*, 2004). The general stress response factor σ^B, which reacts to a variety of different environmental stress factors, influences the *agr* response via SarA (Cheung *et al.*, 1999) and it has been shown that a σ^B deletion mutant has decreased virulence in the murine model for septic arthritis (Jonsson *et al.*, 2004) but not in a murine subcutaneous skin abscess model (Horsburgh *et al.*, 2002). MgrA (also known as Rat or NorR) is a regulator repressor of *agr* (Truong-Bolduc *et al.*, 2003) via SarX and has been shown to sense reactive oxygen and nitrogen species (Chen *et al.*, 2006).

1.6 The SarA family

The staphylococcal SarA family is a family of DNA binding regulatory protein which are all homologous to the SarA regulatory protein. It has 11 members and can be divided into three subfamilies: 1) Single domain proteins; to this subfamily belong SarA, SarR, SarT, SarV, SarX and Rot which are 115 to 141 amino acid residues long and form active dimer. 2) Double domain proteins; to this subfamily belong SarS, SarU and SarY which are 247 to 250 long amino acid residues and are active as monomers with a structure that resembles a dimer from the "single domain family". They have two homologues but not identical domain, and 3) MarR homologues; this subfamily consists of MgrA and SarZ which are 147 and 148 amino acid residues long respectively and form active dimers (Table 1-2 and Figure 1-8).

The crystal structures have been solved for several of the SarA homologues and it is reasonable to assume that within the subfamilies the structure is similar. The structures show that the SarA family are winged helix proteins, where the single domain subfamily is a homo-dimeric winged helix structure possessing a central helical core and two winged helix motifs (Liu *et al.*, 2001;Liu *et al.*, 2006). The two domain subfamily is a monomeric structure forming a similar structure to the dimer of the single domain subfamily of two winged motifs, linked by a well ordered loop (Liu *et al.*, 2006;Liu *et al.*, 2001;Li *et al.*, 2003;Chen *et al.*, 2006). Members of the MarR homolog subfamily still form dimeric winged helix

structures but have more sequence similarity to the Gram-negative bacterial MarR proteins than to the other small SarA proteins (Ingavale *et al.*, 2005).

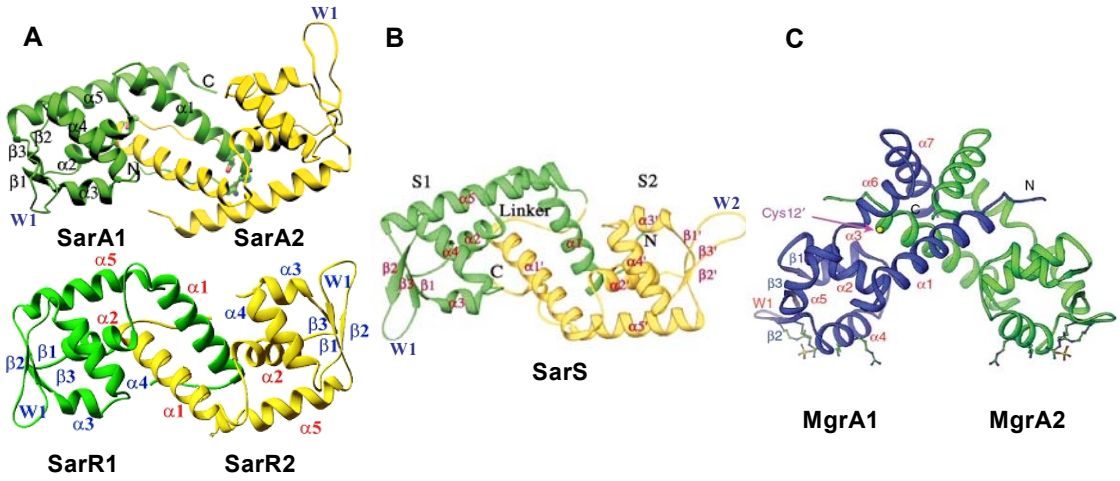


Figure 1-8: Comparison of (A) SarA (Liu *et al.*, 2006), SarR (Liu *et al.*, 2001) (single domain subfamily), (B) SarS (Li *et al.*, 2003) (Two domain family) and (C)MgrA (Chen *et al.*, 2006) (MarR homologue) winged helix structure. The different secondary structures (α-helices α1-7 β-sheets β1- and wings W1 and W2) and important features are marked; Linker and the cysteine involved in oxidation sensing in MgrA.

Table 1-2: The SarA protein family their subfamilies and regulatory functions. *MgrA is also known as Rat or NorR.

Protein	Subfamily	Putative function	Reference
SarA	Single Domain	Activates genes via *agr* and *agr* independent pathways	
SarR	Single Domain	Negative regulator of *sarA* and activator of *agr*	(Cheung *et al.*, 2004)
SarT	Single Domain	Activator of *sarS* and repressor of α-heamolysin	(Cheung *et al.*, 2004)
SarV	Single Domain	Regulator of autolysis repressed by SarA and MgrA	(Manna *et al.*, 2004)
SarX	Single Domain	Negative regulator of *agr* activated by MgrA	(Manna and Cheung, 2006a)
Rot	Single Domain	Repressor of toxin expression, works oppersit to *agr*	(Said-Salim *et al.*, 2003)
SarS	Double Domain	Activator of protein A	(Tegmark *et al.*, 2000)
SarU	Double Domain	Activates *agr*	(Cheung *et al.*, 2004)
SarY	Double Domain	Unknown	
MgrA*	MarR homologue	Regulator of autolysis and *agr*	(Truong-Bolduc *et al.*, 2003)
SarZ	MarR homologue	Activates *hla*	(Kaito *et al.*, 2006)

Structural data suggest that the SarA family all bind to DNA in a similar fashion, with the helix-turn-helix binding to the major grooves and the winged motif binding to the minor groves of the DNA.

In contrast to the similarity in binding domains, the activation domains differ among the family members. Based on the structures of SarA, SarS, SarR and

MgrA five different mechanisms of regulation have been proposed: 1) For MgrA an oxidation sensing mechanism for the cysteine (C12) residue has been proposed (Chen *et al.*, 2006); 2) Bending of target DNA to facilitate contact with the regulatory protein (Cheung *et al.*, 2004); 3) The formation of a complex of three SarA dimers encased in DNA blocking transcription from the promoter (Liu *et al.*, 2006); 4) Formation of heterodimer (for the single domain subfamily) between compatible family members (e.g., SarA and SarR) that interferes with the function of the homodimers (Cheung *et al.*, 2004); 5) Competitive displacement of one homodimer by another (e.g., SarR displacing SarA from the *agr* promoter) (Manna and Cheung, 2006b).

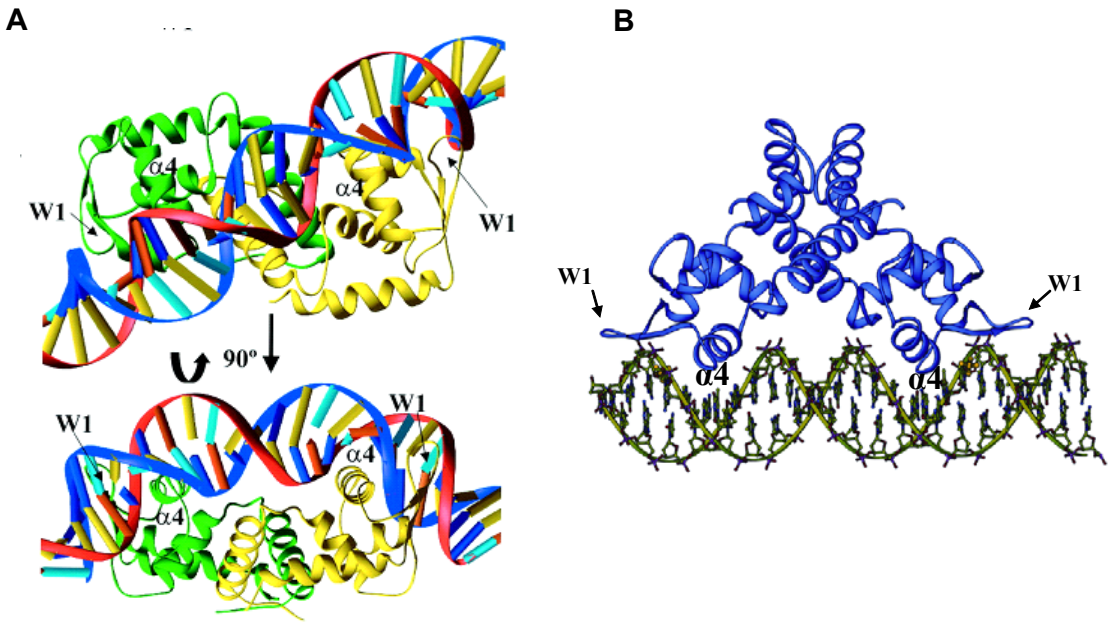


Figure 1-9: Modelling of (A) SarR (Liu *et al.*, 2001) and (B) MgrA (Chen *et al.*, 2006) binding to DNA. In both cases it is the fourth α-helix (α4) that fits in the major grooves of the DNA and the wings (W1) binds to the minor grooves.

1.6.1 SarA

SarA is the best studied protein in the SarA family. It is transcribed from the *sarA* locus which consist of *sarA* and three promoters (P1, P2 and P3) (Bayer *et al.*, 1996) and is required for optimal expression of *agr* (Cheung and Projan, 1994). Transcription peaks in the late exponential phase (Manna and Cheung, 2001) and controls over 100 genes (Dunman *et al.*, 2001). This regulatory effect is either through *agr* by binding to the *agr* promoter region, directly binding to the promoter of the target gene (e.g., *hla* and *spa*) or by stabilising mRNA during the logarithmic growth phase (Cheung *et al.*, 2004;Roberts *et al.*, 2006;Liu *et al.*, 2006) (Table 1-1). Various studies have identified the SarA binding site in the *agr* promoter region using footprinting (Chien and Cheung, 1998;Chien *et al.*, 1999;Rechtin *et al.*, 1999) and the 7 bp consensus binding sequence (ATTTTAT) has been proposed based on promoter selection assay (Sterba *et al.*, 2003).

Mutation studies of SarA revealed that substituting the Y18, E29, E36, K54, L74 or F80 for alanine reduces SarA ability to bind to DNA. Substituting R84 or R90 within the winged motif with alanine completely abolished SarA's ability to bind to DNA. In contrast substituting D88, E89 (part of the DER conserved motif), C9 or K23 with alanine did not alter SarAs ability to bind to DNA, but rendered SarA inactive and unable to complement a *ΔsarA* phenotype (Liu *et al.*, 2006).

1.6.2 Other SarA homologues

1.6.2.1 Single domain SarA homologues

SarR is a 115 amino acid residue protein that binds to the *sarA* promoter and represses the expression of SarA (Manna and Cheung, 2001). SarR also influences *agr* indirectly by repressing SarA but is also capable of binding directly to the *agr* promoter region (Manna and Cheung, 2006b).

SarT is a 118 amino acid residue protein and is repressed by *sarA* and *agr*. SarT represses *hla* and up-regulates *spa* by activating transcription of *sarS* (Schmidt *et al.*, 2001;Schmidt *et al.*, 2003).

SarV is a 116 amino acid residue protein. Its expression is very low under *in vitro* growth conditions and is repressed by MgrA and SarA (Manna *et al.*, 2004). S*arV* mutants are more resistant to detergent, or cell wall antibiotic-mediated lysis than wild type *S. aureus* and over expression lowers the parent strains resistance to lysis (Manna *et al.*, 2004). Therefore, *sarV* might be a part of the *mgrA* and *sarA* controlled autolysis pathway in *S. aureus*.

SarX is 141 amino acid residue long and is the largest of the single domain SarA homologues. SarX is an auto regulator that, in conjunction with MgrA, up-regulates its own expression. SarX is also a repressor of *agr* (Manna and Cheung, 2006a)

Rot was first reported to be 166 amino acid residues in length (Said-Salim *et al.*, 2003) but was later reported to be 133 amino acid residues (Manna and Ray, 2007). Rot is a repressor of toxin synthesis and up-regulates cell wall protein synthesis and has been shown to affect the transcription of 168 genes (Said-Salim *et al.*, 2003). Rot is a positive regulator of *sarS* (Said-Salim *et al.*, 2003) and SarA and SarR have been shown to bind to the *rot* promoter region (Hsieh *et al.*, 2008). Rot is transcribed throughout the different growth phases and is only modestly influenced by *agr* and SigB (Hsieh *et al.*, 2008). RNAIII has been shown to interact directly with *rot* mRNA and to inhibit *rot* translation (Geisinger *et al.*, 2006), indicating that a major part of *agr*-dependent regulation of *rot* is post transcriptional.

1.6.2.2 Two domain SarA homologues

SarS (also called SarH1) is the most studied of the two domain subfamily of SarA homologues. It is 250 amino acid residues long and activates *spa* and represses *hla* transcription (Tegmark *et al.*, 2000;Cheung *et al.*, 2001;Cheung *et al.*, 2001). Transcriptional analysis show that *sarS* is repressed by *agr* and MgrA (Tegmark *et al.*, 2000;Cheung *et al.*, 2001) and activated by ClpXP protease (Frees *et al.*, 2005) and SarT(Schmidt *et al.*, 2003).

SarU is 247 amino acid residues long and is transcribed divergently from *sarT*. SarU is an activator of *agr* and is repressed by SarT (Manna and Cheung, 2003). Since SarT is repressed by *agr* and SarU activates *agr* but is repressed by SarT, the *agr* locus could be activated by this positive feedback loop, in addition to the activation effected by AgrA and SarA.

SarY is a hypothetical SarA homologue and has only been predicted *in scilico* so far. It is predicted to be 247 amino acid residues (Manna and Cheung, 2003).

1.6.2.3 MarR homologues

MgrA, also known as Rat or NorR is the best studied of the MarR homologues. It is 147 amino acid residues in length and is a dimer in its functional form (Truong-Bolduc *et al.*, 2003;Chen *et al.*, 2006). MgrA was initially identified as a regulator of autolysins (Ingavale *et al.*, 2003;Manna *et al.*, 2004). MgrA also regulates *hla* (up regulated) and *spa* (down regulated) expression in both an *agr* dependent (by up regulating transcription of RNAIII) and an *agr* independent manner (Ingavale *et al.*, 2005). Besides having a regulatory effect on *agr,* MgrA also has an effect on the SarA homologue SarS and SarX (Ingavale *et al.*, 2005). Besides these effects MgrA has been observed to positively regulate capsular polysaccharide synthesis (Luong *et al.*, 2003). MgrA is also involved in regulation of antibiotic resistance and has been reported to positively regulate the efflux pumps NorA and AbcA (Truong-Bolduc *et al.*, 2003;Truong-Bolduc and

Hooper, 2007) and to negatively regulate three other efflux pumps NorB, Tet38 and NorC (Truong-Bolduc *et al.*, 2005;Truong-Bolduc *et al.*, 2006). Transcriptional analysis of *mgrA* has shown that MgrA effects 355 genes, 175 of which are positively regulated and 180 which are negatively regulated (Luong *et al.*, 2006). The phenotype of a *mgrA* mutant is very like that of an *agr* mutant.

SarZ is 148 amino acid residues and like MgrA is a member of the MarA subfamily of SarA homologues (Cheung *et al.*, 2004). It has been shown to be a positive regulator of *hla* and *sarZ* mutants have decreased virulence in both a silkworm and a mouse infection model (Kaito *et al.*, 2006).

All these different DNA binding proteins interact with each other, and many also directly or indirectly interact with *agr,* results in a complex virulence regulon in *S. aureus.* In Figure 1-10 the *sar-agr* regulatory network as it is currently understood.

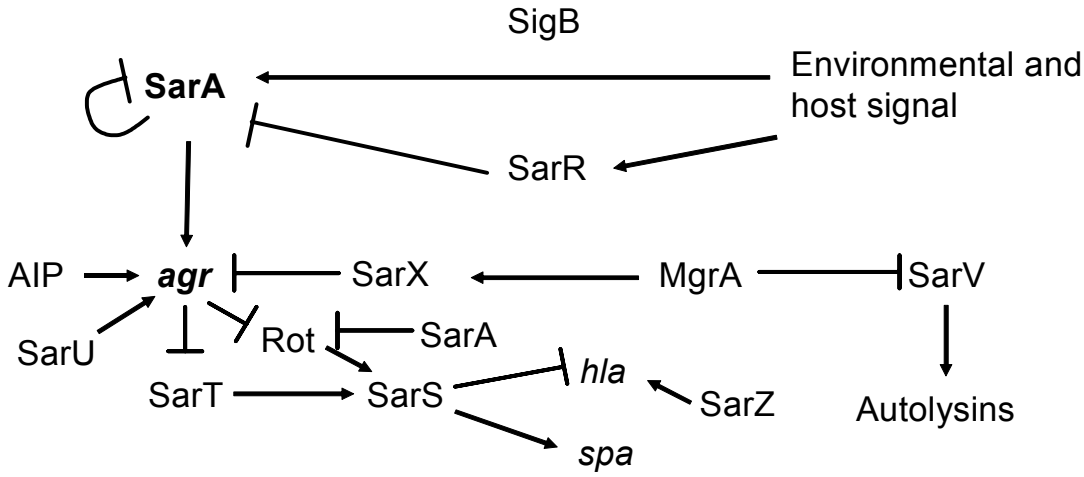


Figure 1-10:The *sar-agr* regulatory network, figure adopted from (Cheung *et al.*, 2008).

1.7 Other virulence regulating TCSs

The SaeRS (Giraudo *et al.*, 1994;Giraudo *et al.*, 1999) TCS is the second major global regulator of the *S. aureus* virulon. Its regulation of virulence factor production is *agr* independent but *agr* and *sae* mutually upregulates each other. Transcription of *sea* is influenced by various environmental stimuli such as salt concentration, pH, glucose and antibiotics (Novick and Jiang, 2003).

The *arl* TCS is a regulator of autolysins and the *norA* efflux pump (Fournier and Hooper, 2000). It is up-regulated by *agr* and repressed the transcription from the $P2_{agr}$ promoter, hence making up an autorepression circuit (Fournier *et al.*, 2001). Consequently, its effect on exoprotein production is probably through *agr*, but it acts independently on protein A production (Fournier *et al.*, 2001).

The SrrAB (also know as SrhRS) TCS is involved in the regulation of the *S. aureus* virulon and energy metabolism under microaerobic conditions (Yarwood *et al.*, 2001;Throup *et al.*, 2001). *agr* and *srr* cross-inhibit each other, and *srr* links the *agr* system to the energy metabolism of *S. aureus* (Novick, 2003).

1.8 Aims of the project

S. aureus is an incredible versatile pathogen which is proving increasingly difficult to control using the antimicrobials available. Antivirulence agents present a possible way of controlling infections without the selection pressure of conventional antibiotics. This study aims to elucidate some of the molecular mechanisms of the *agr* system which offers potential as an antimicrobial, antivirulence target, as follows:

- Use available prediction programs to determine the likely transmembrane topology of AgrC
    - To identify which amino acid residues are likely to play a part in the group specific activation of AgrC
- Develop and validate an easy to use AIP bioreporter
    - Requirements: no *agr* background, no need for additional substrates and high signal to background ratio
- Determine the contribution of each amino acid residue to the group specific activation of AgrC
- Identify the AgrC protein and elucidate the location and distribution of AgrC in the cytoplasmic membrane
- Investigate whether proteins other than AgrB are involved in AgrD processing

A detailed knowledge of the molecular mechanisms involved in the *agr* response should aid the rational design of inhibitors which control infection through attenuation of virulence

2 Materials and Methods

The bacterial strains, plasmids and primers used in this study are listed in Table 2-1 to Table 2-3.

Table 2-1: Strains

Name	Genotype	reference
Staphylococcus aureus strains		
RN4220	*rsbU*$^-$	Kreiswirth *et al.* 1983, *Nature* **305**, 709–712
RN6390B	Wild type,	Patel A. H. *et. al.* 1992, *Gene* **114**,25-34
8325-4	Wild type 8325 cured of prophages	Novick R.P. 1967, *Virology* **33**,155-166
ROJ37	RN6390B Δ*agr::ErmB* #1 *erm*^{R}	This study
ROJ38	RN6390B Δ*agr::ErmB* #2 *erm*^{R}	This study
ROJ39	RN4220 x Φ11(ROJ37) #1 *erm*^{R}	This study
ROJ40	RN4220 x Φ11(ROJ37) #1 *erm*^{R}	This study
ROJ48	ROJ40 pΔagr-lux (SCO) #1 *tet*^{R} *erm*^{R}	This study
Escherichia coli strains		
C41	BL21(DE3) *F*$^-$ *ompT gal hsdSB (rB-mB-) dcm lon λDE3* and an uncharacterised mutation described in the reference (Miroux, B. and Walker, J.E. 1996 *J. Mol. Biol*. **260**, 289–298)	Avidis
C43	as C41 but cured for pMW7(OGCP)	Avidis
TOP10	*F*$^-$ *mcrA* Δ*(mcr-hsdRMS-mcr-BC)* φ*80lacZ* Δ*M15* Δ*lacX74deoR recA1 araD139*Δ(*ara-leu*) *7697galU galK rpsL* (*Str*^{R})*endA1 nupG*	Invitrogen
TOP10F'	*F*(*lacI*^{q},Tn*10*(Tet^R))*mcrA* Δ*(mcr-hsdRMS-mcr-BC)* φ*80lacZ* Δ*M15* Δ*lacX74deoR recA1 araD139*Δ(*ara-leu*) *7697galU galK rpsL* (*StrR*)*endA1 nupG*	Invitrogen
JM109	F' *traD36 proA*$^+$*B*$^+$ *lacI*^{q} *Δ(lacZ)M15/ Δ(lac-proAB)*	Yanisch-Perron C. *et. al.*

	glnV44 e14- gyrA96 recA1 relA1 andA1 thi hsdR17	1985 *Gene* **22**,103-119
DHM1	F- *cya-854 recA1 endA1 gryA96* (*nal*^{R}) *thi1 hsdR17 spoT1 rfbD1glnV44*(*AS*)	*Proc. Natl. Acad. Sci. USA* Vol. 95, pp. 5752–5756, May 1998
BTH101	F- *cya-99 araD139 galE15 galK16 rspL1* (str^{R}) *hsdR2 mcrA1 mcrB*	*Proc. Natl. Acad. Sci. USA* Vol. 95, pp. 5752–5756, May 1998
Saccharomyces cerevisiae Strains		
PJ69-4A	*MAT**a** trp1-901 leu2-3,112 ura3-52 his3-200 gal4Δ gal80Δ LYS2::GAL1-HIS3 GAL2-ADE2 met2::GAL7-lacZ*	James P. *et al* 1996 Genetics 144 (4): 1425-1436

Table 2-2: Plasmids

Name	Genotype	reference
pCR2.1	TOPO cloning vector *kan*^{R} *bla*	Invitrogen
pSK5630	Shuttle vector *cat bla*	(Grkovic *et al.*, 2003)
pSK5632	Expression shuttle vector *cat bla*	(Grkovic *et al.*, 2003)
pMAD	Shuttle vector for gene replacement in low GC gram-positive bacteria *ermC bla*	(Arnaud *et al.*, 2004)
pΔagr-tetM	*agr* up and down stream regions *tetM*	Klaus Winzer
pΔagr-ErmB	*agr* up and down stream regions *ermB*	This study
pΔagr-lux	*agr* up and down stream regions *tetM luxABCDE*	Klaus Winzer
pΔagr-dual	*agr* up and down stream regions *tetM gfp luxABCDE*	Klaus Winzer
pGAD424	Col E1 and 2μ origin GAL4-AD *LEU2 bla*	Clonetech (Bartel *et al* 1993)
pGBT9	Col E1 and 2μ origin GAL4-BD *TRP1 bla*	Clonetech (Bartel *et al* 1993)
pGAD424-agrB	Col E1 and 2μ origin GAL4-AD::*agrB LEU2 bla*	Alan Cockayne
pGAD424-agrD	Col E1 and 2μ origin GAL4-AD::*agrD LEU2 bla*	Alan Cockayne
pGBT9-agrB	Col E1 and 2μ origin GAL4-BD::*agrB TRP1 bla*	Alan Cockayne
pGBT9-agrD	Col E1 and 2μ origin GAL4-BD::*agrD TRP1 bla*	Alan Cockayne
pGBT9-*S.a.*lib	*S. aureus* Newman strain genomic library fused to the GAL4-BD in pGBT9. *TRP1 bla*	Alan Cockayne
pBluescript SK +	High copy cloning vector	Fermentas
pBSK-AgrC1A	pBluescript SK + *agrC1 agrA*	
pBSK-AgrC4A	pBluescript SK + *agrC4 agrA*	
pSKermP2	pSK5630 *ermB* $P2_{agr}$-*agrC1 agrA*	This study
pAgrC1agrA	pSKermP2 *agrC1 agrA*	This study
pAgrC2agrA	pSKermP2 *agrC2 agrA*	This study

pAgrC3agrA	pSKermP2 *agrC3 agrA*	This study
pAgrC4agrA	pSKermP2 *agrC4 agrA*	This study
pAgrC1A101TagrA	pSKermP2 *agrC1A101T agrA*	This study
pAgrC1T104VagrA	pSKermP2 *agrC1T104V agrA*	This study
pAgrC1S107VagrA	pSKermP2 *agrC1S107V agrA*	This study
pAgrC1A101TT104VagrA	pSKermP2 *agrC1A101TT104 agrA*	This study
pAgrC1A101TS107VagrA	pSKermP2 *agrC1A101TS107 agrA*	This study
pAgrC1T104VS107VagrA	pSKermP2 *agrC1T104VS107V agrA*	This study
pAgrC1loop2gr4agrA	pSKermP2 *agrC1A101TT104V107V agrA*	This study
pAgrC4T101AagrA	pSKermP2 *agrC4T101A agrA*	This study
pAgrC4V104TagrA	pSKermP2 *agrC4V104T agrA*	This study
pAgrC4V107SagrA	pSKermP2 *agrC4V107S agrA*	This study
pAgrC4T101AV104TagrA	pSKermP2 *agrC4T101AV104T agrA*	This study
pAgrC4T101AV107SagrA	pSKermP2 *agrC4T101AV107S agrA*	This study
pAgrC4V104TV107SagrA	pSKermP2 *agrC4V104TV107S agrA*	This study
pAgrC4loop2gr1agrA	pSKermP2 *agrC4T101AV104TV107S agrA*	This study
pAgrC4S35LagrA	pSKermP2 *agrC4S35L agrA*	This study
pAgrC4Y39FagrA	pSKermP2 *agrC4Y39F agrA*	This study
pAgrC4L38FY39FagrA	pSKermP2 *agrC4L38FY39F agrA*	This study
pAgrC4G42VagrA	pSKermP2 *agrC4G42V agrA*	This study
pAgrC4loop1gr1agrA	pSKermP2 *agrC4R31KK33YS35LI38FY39FG42VV44SL45TS46LV48L agrA*	This study
pGfp-AgrC	pSKermP2 *gfp-agrC1 agrA* (N-terminal *gfp* fusion to *agrC1*)	This study

Table 2-3: Primers. * indicates primer is 5' phosphorylated.

Name	Sequence
DagrtstF	5'-GATTTACAATTGAATACGCCG
DagrtstR	5'-CTGTCAATTTGATAGCGGG
TetMseq1	5'-GGATATCAGGTTACCACTG
TetMseq2	5'-CCACGCTTCCTAATTCTG
ErmB-F	5'-AATTGGCATGAATTCGGTCTTGCGTATGGTTAACCC
ErmB-R	5'-TTTACTCGAGTAGAATTATTTCCTCCCGTT
pSK-FOR	5'-TCGTATGTTGTGTGGAATTG
pSK-REV	5'-GCTGCAAGGCGATTAAGTTG
pGBTf	5'-TGCCGTCACAGATAGATTGG
pGBTr	5'-CCTGAGAAAGCAACCTGACC
Drlp_up-F	5'-*ATTAATCAATAATGTTTTAAATAT
Drlp_up-R	5'-TACAGGATCCCAATATTGATTTTCATGCGT
Drlp_dwn-F	5'-TACAGAATTCCAGTAGTTATTGCAGTTATTTT
Drlp_dwm-R	5'-*TTTCTTCACATCTATCCCTA
DhutI_up-F	5'-TACAGGATCCCAGCCGTAGCCAAAATACTAGCC
DhutI_up-R	5'-TACAGAATTCAAAGCATAAAATAATTAGAAAGC
DhutI_dwn-F	5'-TACAGAATTCATCATTCATTATTTATCACCT
DhutI_dwn-R	5'-TACAGGATCCGTTGAGTGGTGATCCAAAAG
Drlp_up-F	5'-*ATTAATCAATAATGTTTTAAATAT
Drlp_up-R	5'-TACAGGATCCCAATATTGATTTTCATGCGT
Drlp_dwn-F	5'-TACAGAATTCCAGTAGTTATTGCAGTTATTTT
Drlp_dwm-R	5'-*TTTCTTCACATCTATCCCTA
agrCA1+4-R	5'-ACAAATAAGCTTCGCCGTTAACTGACTTTATTA
agrC1-F	5'-AGTAATGGATCCGAGTGTGATAGTAGGTGGAAT
agrC2-F	5'-AGTAATGGATCCGTTATTTTAGAAAGTGTGTAGC
agrC3-F	5'-AGTAATGGATCCGAGAAAGTGTGATAGTAAGTG
agrC4-F	5'-AGTAATGGATCCCAGAAAGTGTGATAGTAAGTG
agrC4loop1gr1-F	5'-*TATTTTTTCATCATAGTAATTTCGACATTATCGTTA
agrC4loop1gr1-R	5'-*GTCAAGTTTACTGTACTTGATACCACT
S35L-F	5'-*AAATTGGACTATTTGTATATCACG
L38F-F	5'-*AAATCTGACTATTTTTATATCACG

Y39F-F	5'-*TATTTGTTTATCACGGGTATTACG
G42V-F	5'-*TATTTGTATATCACGGTTATTACG
L38F_Y39F-F	5'-*TATTTTTTTATCACGGGTATTACG
Loop1-R1	5'-*ACTGTATTTGATACCACTAATAAT
loop1-R2	5'-*GTCAGATTTACTGTATTTGATACC
agrC1wt-R	5'-*ATATAACTATATACATGTAGTTAGC
agrC1 A101T-F	5'-*ATACATATATCACCAAAATTTCTGAT
agrC1 T104V-F	5'-*ATGCATATATCGTCAAAATTTCTGAT
agrC1 S107V-F	5'-*ATGCATATATCACCAAAATTGTTGAT
agrC1 A101T T104V-F	5'-*ATACATATATCGTCAAAATTTCTGAT
agrC1 A101T S107V-F	5'-*ATACATATATCACCAAAATTGTTGAT
agrC1 T104V S107V-F	5'-*ATGCATATATCGCCAAAATTGTTGAT
agrC1 A101T T104V S107V-F	5'-*ATACATATATCGCCAAAATTGTTGAT
agrC4wt-R	5'-*ATATAACTATATACATGTAATTTGC
agrC4 T101A-F	5'-*TCGCATATATTGTCAAAATCGTTGAT
agrC4 V104T-F	5'-*TCACATATATTACCAAAATCGTTGAT
agrC4 V107S-F	5'-*TCACATATATTGTCAAAATCTCTGAT
agrC4 T101A V104T-F	5'-*TCGCATATATTACCAAAATCGTTGAT
agrC4 T101A V107S-F	5'-*TCGCATATATTGTCAAAATCTCTGAT
agrC4 V104T V107S-F	5'-*TCACATATATTACCAAAATCTCTGAT
agrC4 T101A V104T V107S-F	5'-*TCGCATATATTACCAAAATCTCTGAT
Gfp-N-F	GACATCGGATCCGTTAACAAGGAGGAATAAAAAATG
Gfp-N-R	GACATCGAATTCTTTGTATAGTTCATCCATGCCATG
agrC1-N-F	GACATCGAATTCGTGGAATTATTAAATAGTTATAATTTTG
agrCA1-N-R	GACATCAAGCTTTTATATTTTTTTAACGTTTCT

2.1 Growth Media and antibiotics

All growth media was sterilised by autoclaving at 121˚C, 20 min at 15 p.s.i. unless otherwise stated.

2.1.1 Luria Bertani medium

Luria Bertani (LB) broth was prepared as described by Sambrock *et al* (1989) and consisted of 10 g/L tryptone (Oxoid) 5 g/L yeast extract (Berton, Dikinson and company) and 10 g/L NaCl (Fisher Scientific). LB agar was prepared by the addition of 15 g/L No. 1 Bacteriological Agar (Oxoid) to LB broth.

2.1.2 Tryptic Soya Broth

Tryptic soya broth was prepared as 15 g/L of Tryptic Soya Broth (Berton, Dikinson and company). TSB agar was prepared by the addition of 15 g/L No. 1 Bacteriological Agar (Oxoid) to TSB broth.

2.1.3 B2 medium

B2 broth consisted of 10 g/L Bacto Casamino Acids (Becton, Dickinson and company) 25 g/L Yeast Extract (Becton, Dickinson and company) 1 g/L K_2HPO_4 (Sigma) 25 g/L NaCl (Fisher Scientific). After autoclavation the sterilised medium was supplemented with 25 ml/L of filter sterilised 20% glucose (sigma).

2.1.4 LK medium

LK broth consists of 10 g/L Tryptone (Oxoid) 5 g/L Yeast extract (Becton, Dickinson and company) and 7 g/L KCl (sigma). LK agar was prepared by addition of 15 g/L No. 1 Bacteriological Agar (Oxoid) to LK broth

2.1.5 CYGP medium

CYGP broth was prepared as described by Ji *et al.* (1995) and consists of 10 g/L Casamino acids (Becton, Dickinson and company) 10 g/L Yeast Extract (Becton, Dickinson and company) 5 g/L Glucose (Sigma) and 5.9 g/L NaCl (Sigma) supplemented by 40 ml/L β-Glycerophosphate (1.5M).

2.1.6 YPD medium

YPD broth was prepared as described in Yeast Protocol Handbook (Clontech, 2001) and contains 20 g/L Peptone (Becton, Dickinson and company) and 10 g/L Yeast extract (Becton, Dickinson and company). After autoclavation the medium was supplemented with 50 ml/L 40% Glucose wt/vol. YPD agar was prepared by adding 20 g/L Technical agar No. 3 (Oxoid) to YPD broth.

2.1.7 YMM medium

YMM broth was prepared as described in Yeast Protocol Handbook (Clontech, 2001) and contains 6.7 g/L Yeast Nitrogen base w/o amino acids (Becton,

Dickinson and company) After autoclavation the medium was supplemented with 50 ml/L 40% Glucose wt/vol.

YMM agar was prepared by adding 20 g/L Technical agar No. 3 (Oxoid) to YMM broth.

2.1.8 Antibiotics

Growth media for *Escherichia coli* was supplemented as required with: ampicillin (Amp) at 100 μg/ml, chloramphenicol (Cm) at 30 μg/ml, tetracycline (Tet) at 10 μg/ml or kanamycin (Kan) at 50 μg/ml. Growth media for *Staphylococcus aureus* was supplemented as required with the following: Erythromycin (Ery) 10 μg/ml, chloramphenicol (Cm) at 10 μg/ml or tetracycline (Tet) at 10 μg/ml.

2.1.9 Yeast medium supplements

Growth media for *Saccharomyces cerevisiae* was supplemented as required with the following: L-Methionine (20 μg/ml) (Sigma), Uracil (20 μg/ml) (Sigma), L-Histidine-HCl (20 μg/ml) (Sigma), Adenine Sulphate (20 μg/ml) (Sigma), L-Tryptophan (20 μg/ml) (Sigma), L-Leucine (30 μg/ml) (Sigma)

For X-gal YMM agar: 80μg/ml X-gal, 20 μg/ml L-Methionine, 20 μg/ml Uracil, 20μg/ml Adenine Sulphate and 1:10 10X BU salts (70g/L $Na_2HPO_4 \cdot 7H_2O$ and 30g g/L NaH_2PO_4, pH 7).

2.2 Microbial growth conditions

E. coli and *S. aureus* were grown at 37˚C unless otherwise stated. *E. coli* and *S. aureus* liquid cultures were grown in broth with agitation at 200 rpm unless otherwise stated.

S. cerevisiae were grown at 30˚C. *S. cerevisiae* liquid cultures were grown in broth with agitation at 250 rpm unless otherwise stated.

Optical density of bacterial or yeast cultures was measured at 600 nm (OD_{600}), using a Pharmacia LKB Novaspec II spectrometer.

2.3 Manipulation of DNA

2.3.1 DNA restriction enzymes

Restriction enzymes were obtained from Promega, UK. DNA restriction digest contained 0.5-1 μg of DNA, 1U of restriction endonuclease and 1 x buffer (supplied by manufacture) made up to a final volume of 30-40 μl with sterile distilled water. Digests were incubated for between 2 to 18 h at the appropriate temperature. Restriction enzymes and cut off ends from PCR products were removed from the reaction using a QUIquick PCR purification kit as described in the manufacturers instructions (Qiagen Ltd, UK). For digested plasmid desired fragment was isolated on a 0.7-1% wt/vol agarose gel and gel purified as described in section 2.3.8

2.3.2 DNA Ligation

Ligation of DNA was carried out using either T4 DNA ligase supplied by Promega UK or Quick T4 DNA Ligase supplied by New England BioLabs (UK) Ltd. Purified DNA inserts were ligated to purified vector DNA in a ratio of 3:1. The ligation reaction mix also contained 1 x T4 DNA ligase buffer (supplied by manufacture) and 1 U T4 DNA ligase in a final volume of 10 µl made up with sterilised distilled water. Reactions were incubated overnight at 16˚C.

PCR products with phosphorylated 5' ends were ligated using Quick T4 ligase (NEB, UK). The ligation reaction mix contained approx. 25 ng DNA 1 x Quick Ligation Buffer and 0.5 µl Quick T4 DNA ligase in a final volume of 10 µl made up with sterilised distilled water. Reactions were incubated 5 min at room temperature (approx. 22˚C)

2.3.3 Synthesis of oligonucleotide primers

Oligonucleotide primers were synthesised at the Biopolymer Synthesis and Sequencing Unit, Queen's Medical Centre, University of Nottingham or at Sigma-Genosys (Suffolk, UK)

2.3.4 Polymerase Chain Reaction

Polymerase chain reaction (PCR) DNA amplifications were performed as described by (Sambrook *et al.*, 1989). PCR reactions for cloning and mutagenesis were carried out in 50 µl volumes using 1 U of Phusion™ High-Fidelity DNA polymerase (Finzymes supplied by NEB UK), 1 x HF Buffer, 200µM dNTPs

(Promega), 0.5-1 μM of each primer, 1 pg – 10 ng DNA template and sterilised distilled water to the final volume of 50 μl. Reaction mixtures were heated to 98˚C for 30 s to denature template. Primers were annealed at 48 - 60˚C for 20s and the product extended at 72˚C for 20 s per 1 kb DNA to be amplified. Denaturation, annealing and extension cycles were repeated 25-35 times. A final extension time of 5 min at 72˚C ensured completion of all strands.

Screening PCR and some cloning PCR reactions were carried out in 25-50 μl volumes using 1 U of *Taq* thermostable DNA polymerase (Promega UK.), 1 x PCR buffer (supplied with enzyme), 1 μM of each primer, 10-500 ng template DNA, 1,5 mM $MgCl_2$ and 0,2 nM dNTP (Promega UK.) Reaction mixtures were heated to 95˚C for 2 min to denature the template. Primers were annealed at 48-60˚C for 30 s and the product extended at 72˚C for 1 min per 1 kb DNA to be amplified. Denaturation, annealing and extension cycles were repeated 25 – 30 times. A final extention time of 5 min at 72˚C ensured completion of all strands.

The reactions were carried out in a Hybaid PCR Express Thermo (Hybaid, Ashford, UK). The elongation time and annealing temperatures were varied depending on the primers used, the stringency levels required and specified in the relevant results chapter.

2.3.5 Site directed mutagenesis by PCR

All site directed mutagenesis reactions was carried out using Phusion® Site-Directed Mutagenesis Kit (Finnzymes, Espoo, Finland) in accordance to the the

protocol supplied by the manufacture using phosphorylated primes designed according to the maufactures description (Finnzyme, Espoo, Finland). For template pBSK-AgrC1A and pBSK-AgrC4A (Table 2-2) was used.

2.3.6 Purification of PCR products

PCR products were purified by either agarose gel electrophoresis or using a QIAquick PCR purification kit as described by manufacturer's (Qiagen UK). 1 x volume of PCR product was mixed with 5 x volume of PBI buffer (supplied by manufacture), the mix were transferred to a spin column and spun at 14,000 rpm for 1 min and the flow though discarded. 750 µl of PE buffer (supplied by the manufacture) were added to the column and it was spun for 1 min at 14,000 rpm and the flow through discarded. The column was then spun for 1 min at 14,000 rpm to remove residues of the PE buffer and the dry column were inserted to a 1.5 ml microcentrifuge tube and 30 – 50 µl of 1/10 EB buffer (supplied by manufacture) was added. The column was incubated at room temperature for 1 min and spun at 14,000 rpm for 1 min. The eluate contain the purified PCR product.

2.3.7 DNA agarose gel electrophoresis

Agarose gels were prepared and run as described in (Sambrook *et al.*, 1989). Gels were prepared in 1 x TAE (40 mM Tris acetate, 2 mM EDTA, pH 8,5) to a final concentration of 0.7 – 1 % wt/vol agarose. To enable visualisation of the DNA,

ethidium bromide was added to the molten agarose at a final concentration of 0.4 µg / ml before the gel was poured. Samples were prepared in loading buffer (6 x stock: 0.4% orange G, 0.03% bromophenol blue, 0.03% xylene cyanol FF, 15% Ficoll® 400, 10mM Tris-HCl (pH 7.5) and 50mM EDTA (pH 8.0) supplied by Promega UK) and gels run in 1 x TAE at 80 – 100 V. A UV transluminator was used to visualise the DNA and bands were excised as required. DNA markers, 1 kb or 100 bp ladder (Promega UK), were used at 0.5 µg pr lane.

2.3.8 Purification of DNA from agarose gels

DNA was extracted from agarose gels using either QIAquick gel purification kit (Qiagen UK) or Wizard® SV Gel and PCR Clean-Up System (Promega UK) as described by the manufacturer. Briefly, the desired bands were cut out of the agarose gel and put in a 1.5 ml micro centrifuge tube. Both systems work by dissolving the gel in a buffer by heating it to 60˚C for 10 min in 30 µl of buffer QC pr 10 µg gel for QIAquick and 10 µl of Membrane Binding solution per 10 mg gel for SV gel. For the QIAquick system 10 µl of isopropanol per 10 mg of gel were added and mixed by vortexing. The dissolved gel was transferred to filter spin columns incubated at room temperature for 1 min and spun at 13,000 rpm for 1 min. The flow through was discarded and 700 µl of either PE buffer (QIAgen) or Membrane Wash Solution (Promega) was added and the columns were spun at 13.000 rpm for 1 min and the flow through discarded. The wash was repeated with 500 µl Membrane Wash Solution for the Wizard® SV system and spun at 13,000

rpm for 5 min and the flow through discarded. Columns were spun for 1 min to remove the last residues of the wash buffers and transferred to new labelled 1.5 ml microcentrifuge tubes. 30 – 50 µl of sterile destilled water or 1/10 Buffer EB (supplied by manufacture) was added directly to the membrane and the columns were incubated at room temperature for 1 min and spun at 13,000 rpm for 1 min. The eluted DNA was stores at 4˚C.

2.4 DNA sequencing

DNA sequencing was conducted at Geneservice (Cambridge, UK). Specific primers were used in the sequencing reactions, details of which are given in the appropriate chapter section.

2.5 DNA Sequence Analysis

DNA sequence data was analysed using Chromas Lite 2.0 (Technelysium Pty Ltd) and VectorNTI 10 Advance 10 (Invitrogen). When required, DNA sequences were obtained from the National Centre Biotechnology Information (NCBI) [http://www.ncbi.nlm.nih.gov]. Databases were also searched for sequence similarity using the basic alignment search tool (BLAST) on the NCBI webpage [http://www.ncbi.nlm.nih.gov/BLAST/].

2.6 Protein identification

Protein identification was conducted at Biopolymer Synthesis and Sequencing Unit, Queen's Medical Centre, University of Nottingham.

2.7 Manipulation of Escherichia coli

2.7.1 Electro-competent cells

Electro-competent *E. coli* JM109, TOP10 and C41 cells were prepared as described by (Ausubel *et al.*, 1988). Briefly, 1 l of LB broth was inoculated with 10 ml of an overnight *E. coli* culture. The culture was incubated at 37˚C, shaking at 200 rpm until an OD_{600} of 0.5 – 0.8 was reached. Following incubation on ice for 15 – 20 min, the cells were harvested by centrifugation at 5000 x *g*, 4˚C for 10 min. Cells were then washed in 1 l of ice-cold sterile distilled water, recentrifuged, washed with 500 ml ice-cold sterile distilled water, then resuspended in 20 ml ice-cold 10 % vol/vol glycerol. Cells were centrifuged once again prior to resuspension in a final volume of 1 ml of ice-cold 10 % vol/vol glycerol. Aliquots of 40 µl were transferred to cold microcentrifuge tubes and flash frozen in liquid nitrogen before storage at -80˚C.

2.7.2 Rubidium Chloride competent *E. coli*

Chemically competent *E. coli* JM109 cells were prepared as described by (Kushner, 1978). An *E. coli* overnight culture was used to inoculate 50 ml LB to an OD600 of approx 0.05 in a 500 ml Erlenmeyer flask. Cells were incubated at 37˚C, shaking at 200 rpm until an OD_{600} of 0.6 – 0.7 was reached. Following incubation on ice for 15 min, the cells were harvested by centrifugation 3000 x g for 15 min, the supernatant poured of and the cell pellet dried by inverting the falcon tube on a paper towel. Cells were resuspended by moderate vortexing in 17 ml *RF1* (see below for recipe) and incubated on ice for 15-30 min. Cells were centrifuged again and resuspened in 4 ml *RF2* (See below for recipe) and incubated on ice for 15 min. Aliquots of 400 µl were transferred to cold microcentrifuge tubes and flash frozen in liquid nitrogen before storage at -80˚C.

Table 2-4: RF1

Compound	Amount/ litre	Final concentration
RbCl	12 g	100 mM
$MnCl_2 \cdot 4H_2O$	9.9 g	50 mM
Potassium Acetate	30 ml of a 1M stock pH 7.5	30 mM
$CaCl_2 \cdot 2H_2O$	1.5 g	10 mM
Glycerol	150 g	15% wt/vol

Adjust the pH to 5,8 with 0.2 M acetic acid.

Table 2-5: RF2

Compound	Amount/ litre	Final concentration
MOPS	20 ml of a 0.5 M stock pH 6.8	10 mM
RbCl	1.2 g	10 mM
$CaCl_2 \cdot 2H2O$	11 g	75 mM
Glycerol	150 g	15% wt/vol

Adjust pH to 6,8 with NaOH.

2.7.3 Transformation of *Ecoli*

2.7.3.1 Electroporation

Ligation mix to be electroporated into competent cells was dialysed against sterile distilled water for 30 min using 0.025 μm Millipore filters (Millipore Corporation, USA) to remove any salts. Electrocompetent cells were thawed on ice and 2-5 μl dialysed ligation mix or 1 – 2 μl plasmid DNA added to the 40 μl aliquot. The mixture was incubated on ice, for 1 min before transferring a 2 mm electroporation cuvette (BioRad), on ice. A BioRad Gene Pulsar was used, according to manufacturer's instructions, to deliver an electric pulse of 2.5 kV (25μF, 200Ω). The cuvette was immediately removed and 1 ml of LB broth was added immediately to the electroporation cuvette. After incubation for approx 2 h at 37˚C (unless otherwise stated) 100 – 200 μl of the bacterial suspension was spread on to an appropriate selective plate.

2.7.3.2 Heat shock transformation

Cells were removed from the freezer and thawed at room temperature until just liquid, and placed back on ice. 200 µL aliquots were mixed with 1-10 µL of DNA and incubated on ice for 20 min. Cells were heat shocked by placing tubes in a 42˚C water bath for 90s, and then chilled by returning to ice immediately.

Cells were recovered by adding 500 µL of LB and incubated shaking at 37˚C for approx 1 h and plated on LB plates containing appropriate antibiotic

2.8 Manipulation of Staphylococcus aureus

2.8.1 Preparation of plasmids DNA

For purification of plasmid DNA from *S. aureus* a QIAprep Spin Miniprep Kit (Qiagen UK) was used with modification to the protocol. Briefly, 2-3 ml of a *S. aureus* overnight culture was harvested in a by centrifugation 14,000 rpm 1 min. The pellets was resuspended in 250 µl P1 buffer (supplied by the manufacturer) and 5 µl of a 5 mg/ml Lysostaphin (Sigma) was added, mixed and incubated at 37˚C for 30 min. 250 µl P2 buffer (supplied by the manufacturer) was added and mixed by gently inverting the tube 3-4 times, 350 µl N3 buffer was added and tube gently inverted 3-4 times to facilitate total mixing. Mixture was centrifuged for 10 min at 14,000 rpm to precipitate genomic DNA and cell debris. Supernatant was gently poured on to a spin column and spun at 14,000 rpm for 1 min and flow through was discarded. 750 µl of buffer PE (supplied by manufacturer, Ethanol

added) was added and column centrifuged again for 1 min at 14,0000 rpm. Flow through discarded and column spun again at 14,000 rpm for 1 min to get rid of wash buffer residues. Spin column was transferred to a new micro centrifuge tube and 30 – 50 µl of a 1:10 dilution of EB buffer was added. Columns were incubated at room temperature for 1 min and plasmid DNA recovered by centrifugation at 14,000 rpm for 1 min.

2.8.2 Preparation of genomic DNA

Genomic DNA was recovered from *S. aureus* using a QIAGEN DNeasy Tissue kit with a modified protocol. Briefly, *S. aureus* cells were grown at 37˚C for 4 – 6 h or statically overnight in LB or TSB and 2-3 ml culture was harvested by centrifugation at 14000 rpm for 1 min. The pellet was resuspend in 180 µl enzyme lysis buffer by vortexing and 5 µl of a 5 mg/ml lysostaphin in dH_2O (Sigma) plus 5 µl of a 1 mg/ml RNaseA in dH_2O (Sigma) was added and the cells were incubated for 30 min at 37˚C. 25 µl Proteinase K (supplied by manufacturer) and 200 µl Buffer AL was added and samples mixed by vortexing and incubated at 70˚C for 30 min. 200 µl of 100% ethanol was added and samples mixed thoroughly by vortexing. The mixture was transferred to DNeasy mini column sitting in a 2 ml collection tube and centrifuged at 8000 rpm for 1 min. The flow through and collection tube were both discarded. The column was transferred to a new 2 ml collection tube and 500 µl buffer AW1 was added and columns were

centrifuged at 8000 rpm for 1 min. The flow through and collection tube were again discarded. The column was transferred to a new 2 ml collection tube and 500 µl buffer AW2 was added and columns were centrifuged at 13,000 rpm for 3 min. The flow through and collection tube were agsin discarded. The column was transferred to a 1.5 ml microcentrifuge tube and 200 µl Buffer AE was added directly onto the column membrane, incubated for 1 min at room temperature and genomic DNA recovered by centrifuging the column at 8000 rpm for 1 min.

Enzyme lysis solution:

20mM Tris pH 8.0

2mM EDTA

1.2% vol/vol Triton X-100

Store at room temperature.

2.8.3 Preparation of electro-competent *S. aureus*

Electro-competent *S.aureus* RN4220, RN6390B and ROJ48 cells were prepared as described by (Schenk and Laddaga, 1992). Briefly, an overnight culture of the desired strain was grown in 30 ml of B2 broth for no more than 15 h. This overnight culture was used to inoculate 200 ml of pre-warmed B2 broth to an OD_{600} of approx. 0.25. The culture was incubated at 37˚C, shaking at 200 rpm until an OD_{600} of 0.35 – 0.4 was reached (approx 1 – 1.5 h). All of the washes and

incubations were done a room temperature unless otherwise stated. The cells were harvested by centrifugation at 10.000 x *g* for 10 min. Cells were then washed in 30 ml of sterile distilled water, recentrifuged, washed with 30 ml sterile distilled water. Centrifuged again and resuspended in 10 ml 10 % vol/vol glycerol. This step was repeated with resuspension in 5 ml of 10 % vol/vol glycerol and incubated at room temperature for 15 min. Cells were transferred to 1.5 ml microcentrifuge tubes and centrifuged once again at 10.000 x *g* prior to resuspension in a final volume of 1.5 ml of 10 % vol/vol glycerol. Aliquots of 80 µl were stored at -80˚C.

2.8.4 Electroporation of *S. aureus* cells

Electrocompetent cells were removed from the freezer and thawed at room temperature prior to use. The cells were mixed with approx 1 µg of plasmid DNA was added to 70 µl of electrocompetent *S. aureus* cells and transferred to a 0.2 mm electroporation cuvette (BioRad). A BioRad Gene Pulsar was used, according to manufacturer's instructions, to deliver an electric pulse of 2.3 kV (25µF, 100Ω). The cuvette was immediately removed and 1 ml of B2 broth was added. After incubation for approx 2 h at 37˚C (unless otherwise stated) 100 – 200 µl of the bacterial suspension was spread on to an appropriate selective plate, the rest of the cells were spun at 13.000 rpm for 1 min, supernatant discarded and resuspended in 200 µl B2 broth and spread in an appropriate selective plate.

2.8.5 Φ11 transduction

Phage Φ11 transduction of *Staphylococcus aureus* was carried out as described by P. Chan, Sheffield Uni (1996) as follows:

2.8.5.1 Phage Lysate preparation

Briefly, an *S. aureus* over night culture in TSB of 8325-4 or ROJ37 was grown at 37˚C. The overnight culture was used to inoculate 25 ml freshly prepared TSB containing 10 mM $MgSO_4$ and 10 mM $CaCl_2$ to an OD_{600} of 0.05 in a 100 ml conical flask and incubated shaking 200 rpm at 37˚C until an OD_{600} of 0.2 (approx 1 h). 10 ml of culture was subcultured in a fresh conical flask and 1 ml of Φ11 stock (kindly supplied by Dr. Alan Cockayne) was added. Both cultures, with and with out phage, was incubated for a minimum of 4 h, until lysis was apparent by comparing phage containing culture with uninfected control, at 37˚C with shaking at 200 rpm. Phage containing culture was centrifuged 10 min at 5000 rpm to remove unlysed cells and other cell debris. Supernatant, containing the lytic phage, was filter sterilised and labelled lysate A. The lysis step was repeated using 1 ml of Lysate A instead of 1 ml Φ11stock. The resulting lysate was labled Lysate B.

2.8.5.2 Determination of phage titres

Titres of Φ11 lysates was determined as follows. A 10 ml overnight culture of *S. aureus* 8325-4 was grown in TSB broth. 0.25 ml of the overnight culture was used

to inoculate 25 ml of fresh TSB and incubated shaking at 37˚C to an OD_{600} of 0.2 after 2 h. 100 ml TSB bottom agar was (10 g/L agar No 1) melted, cooled down to approx 50˚C and and 5 ml of a sterile 1 M $CaCL_2$ solution was added, plates poured and left to solidify. At the same time TSB top agar (7 g/L agar No 1) was melted and cooled to 50˚C before 1 ml of sterile 1 M $CaCL_2$ and 1 ml of sterile 1 M $MgSO_4$ was added. The top agar was kept melted while TSB bottom agar solidified. 100 ml of the TSB top agar was seeded with 20 ml early log phase *S aureus* 8325-4 (OD_{600} approx 0.2) and 5 ml of the mixture was evenly distributed on top of a TSB bottom agar plate and left to solidify. A series dilution of the Φ11 lysate to be tested was made and 10 µl of each dilution spotted on the the freshly made plates. When lysate was absorbed plates was incubated at 37˚C over night. The number of plaques on each plate was counted and the number of plaque forming units per ml was calculated.

2.8.5.3 Phage transduction

S. aureus RN6390B cells were phage transduced as described by P Chan. An overnight culture of RN6390B was grown in 20 ml LK broth. The cells were harvested by centrifugation for 10 min at 10000 rpm and resuspended in 1 ml LK broth. Both a sample and a control were mixed. Sample consisted of 500 µl cells, 1 ml LK with 10 mM $CaCl_2$ and 500 µl Φ11 lysate B. Control consisted of 500 µl cells and 1.5 ml LK with 10 mM $CaCl_2$. Both sample and control was incubated first 25 min stationary and the 15 min shaking at 37˚C. After incubation sample

and control was transferred to an ice bucket and 1 ml of ice cold 0.02 M sodium citrate was added. Both Sample and control was centrifuged for 10 min at 10000 rpm at 4˚C and resuspended in 1 ml ice cold 0.02 M sodium citrate and incubated on ice for 2-3 h. Aliquots of 100 µl was spread on selective LK plate's containing 0.05% wt/vol sodium citrate. Plates were incubated at 37°C for 12-72 h and transduction of the desired DNA was verified by Southern blot.

2.9 Manipulation of Saccharomyces cerevisiae

2.9.1 Lithium acetate transformation

Co-transformation of two-hybrid plasmids into *Saccharomyces cerevisiae* strain PJ694A was performed as described by (Gietz *et al.*, 1992). Briefly, 1 ml of YPD broth was inoculated with several 2-3 mm colonies of PJ69-4A and vortexed vigorously for approx 5 min. The cell suspension was transferred into a conical flask containing 50 ml of YPD broth and incubated over night (approx 17 h) at 30˚C with shaking at 250 rpm. 30 ml of the overnight culture was transferred to a conical flask containing 300 ml of YPD, OD_{600} 0.2 – 0.3. The culture was incubated for 3 h at 30˚C shaking at 230 rpm until OD_{600} of 0.4 – 0.6 was reached. The culture was transferred to 50 ml falcon tubes and centrifuged at 1000 x *g* for 5 min at room temperature. The supernatant was discarded and the cells resuspended and pooled in sterile TE, final volume of 40 ml and centrifuged again at 1000 x *g* for 5 min, room temperature. The supernatant was discarded and the cells were resuspended in 1.5 ml of freshly prepared sterile 1 x TE / 1 x lithium acetate.

Approx 0.1 µg of each two-hybrid plasmid and 0.1 mg of herring testes carrier DNA were mixed in a fresh 1.5 ml tube and 0.1 ml of competent *S. cerevisiae* PJ69-4A cells was added and the suspension mixed by vortexing. 0.6 ml of PEG / lithium acetate solution was added and the tubes were vortexed for approx 10 s and incubated at 30˚C shaking at 200 rpm for 30 min. 70 µl of DMSO was added and the suspension mixed well by gently inverting the tubes. The transformation mixture was heated to 42˚C for 15 min and chilled on ice for 1-2 min. Cells were centrifuged at 14000 rpm for 5 s and the supernatant removed. Cells were then resuspended in 0.5 ml of sterile 1 x TE buffer and aliquots of 100 µl of 10^0, 10^1, 10^2, and 10^3 times dilutions were plated onto selective YMM plates. The plates were incubated at 30˚C for 2-4 days.

2.9.2 Preparation of plasmid DNA from *S. cerevisiae*

Preperation of plasmids from was done as described in (Holm *et al.*, 1986)*S. cerevisiae* PJ69-4A containing two-hybrid plasmids was patched on YMM MUHA (Yeast minimal medium with methionine, uracil, histidine and alanine) plates and incubated at 30˚C for 3-4 days. Approx 10 mm^2 of the patches was scraped off and resuspended in 50 µl TE buffer in a 1.5 ml microcentrifuge tube. 10 µl of lyticase solution (5 U/µl lyticase (sigma) in TE buffer), to lyse the yeast cells, was added to each tube and vortexed. Tubes was incubated for 30 – 60 min at 37˚C shaking, 200 rpm. 10 µl 20% SDS was added and tubes was vortexed for 1 min and the cells subjected to a freeze thaw cycle (-20˚C) before vortexing. 130 µl

of TE, pH 7, was added to each tube to bring the volume up to 200 µl. Next 200 µl of phenol:chloroform:isoamyl alcohol (25:24:1) was added and tubes was vortexed for approx 5 min and centrifuged at 14000 rpm for 10 min. The upper aqueous phase was transferred to a fresh micro centrifuge tube and 8 µl of 10 M ammonium acetate and 500 µl of 96% vol/vol ethanol were added to precipitate the DNA. The tubes were incubated at -70˚C for 1 h and centrifuged at 14000 rpm for 10 min. The supernatant was discarded and the pellet dried completely before being resuspended in 20 µl dH_2O.

2.9.3 Yeast two-hybrid screening

Interactions between the fusion proteins expressed by the two-hybrid vectors transformed into the yeast *S. cerevisiae* PJ69-4A were evaluated as follows. For library screening for intactions with AgrB the co-transformants was plated onto selective plates YMM MUA (Yeast Minimal media + Methionine +Uracil +Alanine) which not just selects for the two plasmids but also for interactions between AgrB and the protein expressed from the library plasmid. 150 colonies which showed potential interaction were picked and streaked onto more stringent plates YMM MUA X-gal, YMM MUH and YMM MU, with YMM MU being the most stringing and incubated at 30°C for 72 h. Colonies showing positive for interactions, by turning blue on X-gal and growing on the most stringent selection medium, were restreaked on YMM MUA plates and incubated for 72 h at 30°C, this was repeated twice to make sure that each cell only contained one copy of the

library plasmid. Plasmid DNA was isolated from overnight cultures of streaks showing positive for interactions. Insert of the library plasmid was PCR amplified using primers pGBTf and pGBTr (Table 2-3). 5 µl the resulting PCR fragments was digested with the restriction enzyme *HaeIII* and cut and uncut PCR products were run on a 1% analytical agarose gel. PCR products revealing the same restriction pattern was assumed to have come from identical library inserts. Two of each PCR product showing an identical restriction pattern was sent for sequencing to identify the insert.

2.9.4 Analysis of Yeast Two-Hybrid data

Sequence data was analysed using both blastn and blastx (2006) and the VectorNTI 10.0 software packet (Invitrogen)

2.10 Southern blot

Southern hybridisation was carried out as described in Sambrook *et al.* (1989).

2.10.1 DIG labelling of DNA probe

Probes for southern blots were labled with digoxygenin (DIG) using PCR DIG Labelling Mix (Roche Applied Science, UK) as recommended by the manufacture. Briefly, a PCR reaction was carried out using using 1 U of *Taq* thermostable DNA polymerase (Promega UK.), 1 x PCR buffer (supplied with enzyme), 1 µM of each primer, 10-500 ng template DNA, 1.5 mM $MgCl_2$ and 200

µM dNTP (10 µl) PCR DIG Labelling Mix and sterile dH2O to 100 µl. The reaction mixture was heated to 95˚C for 2 min to denature the template. Primers were annealed at 48-60˚C for 30s and the product extended at 72˚C for 1 min. Denaturation, annealing and extention cycles were repeated 25 – 30 times. A final extention time of 5 min at 72˚C ensured completion of all strands.

2.10.2 Digestion of DNA and gel electrophoresis

Genomic DNA isolated from *S. aureus* strains RN6390B and RN4220, were digested with *PstI* and loaded on a 0.7% agarose gel. The gel was run for 3h at 100 V (approx 3.5 V/cm). After electrophoresis the DNA agarose gel was pre-treated with 250 mM HCl for 10 min, shaking gently at room temperature to depurinate the DNA. The DNA was denatured by submerging the gel in Denaturation solution (0.5 M NaOH, 1.5 M NaCl) for 2 x 15 min, gently shaking at room temperature and the gel was rinsed with sterile dH_20. Gel was submerged in Neutralisation solution (0.5 M Tris-HCl, 1.5 M NaCl) for 2 x 15 min gently shaking at room temperature and equilibrated in 20 x SSC (3M NaCl, 0.3 M sodium citrate, pH 7.0) for 10 min. DNA was transferred onto a positively charged Nylon membrane (Roche, USA) using the setup illustrated in Figure 2-1. The gel was placed on top of a piece of Whatman 3MM paper soaked in 20 x SSC and placed on a "bridge" resting in a shallow reservoir of 20 x SSC. A piece of positively charged Nylon membrane the size of the gel was placed on top of the gel and a gel size piece of Whatman paper was placed on top. A total of 4 gel size

pieces of blotting paper (Sigma) was placed on top and the blotting sandwich was finished of by placing a glass plate and a weight (approx 500g) on top of the sandwich. The DNA was left to transfer overnight.

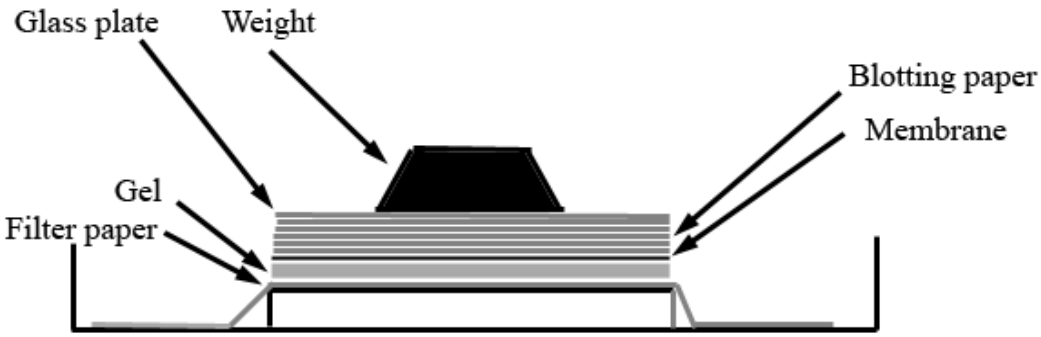


Figure 2-1: Schematic of a southern blot transfer set up

2.10.3 Hybridization

The blotting sandwich was disassembled and the DNA was cross linked to the still damp membrane by exposing the membrane to UV light for 2 min, DNA side up. The membrane was rinsed in distilled water and left to air dry. The dry blot was placed in a glass roller bottle, 20 ml of prewarmed DIG Easy hyb (Roche, USA) was added and the blot was incubated in a hybridisation oven at 42˚C for 30 min. In the mean time the prepared DIG labelled DNA probe was denatured by heating in a boiling water bath for 5 min and quenching on ice for 2 min and the probe was added to 7 ml DIG Easy hyb (Roche). The DIG Easy hyb from the prehybridisation step was discarded and the 7 ml DIG Easy hyb containing the

labelled probe was added to the roller bottle contaning the membrane. The membrane with probe was left to hybridise overnight in at 42˚C in a hybridisation oven.

The hybridisation mix was poured into a 50 ml falcon tube and stored at -20C. The blot was rinsed in low stringency buffer (2 x SSC, 0.1% SDS) and 50 ml of low stringency buffer was added and the blot was washed at 65°C for 30 min. The buffer was poured of and replaced with 50 ml of preheated of high stringency buffer (0.5 x SSC, 0.1% wt/vol SDS) and the blot incubated at 65°C for 2 x 15 min.

2.10.3.1 Detection of hybridised bands

The membrane was equilibrated in 50 ml washing buffer (0.1 M maleic acid, 0.15 M NaCl; pH 7.5; 0.3% Tween) for 2 min, the washing solution was poured of and 50 ml of blocking solution (1 x Blocking Solution (Roche, USA) in 0.1 M maleic acid, 0.15 M NaCl; pH 7.5) and incubated at room temperature for 1 h. The blocking solution was removed of and replaced 20 ml antibody solution (1:10000 alkaline phosphorase conjugated anti-digoxigenin ab in blocking solution) and incubated for 30 min at room temperature. The antibody solution was discarded and the membrane was washed 2 x 15 min in 50 ml washing solution. The membrane was equilibrated in 20 ml detection buffer (0.1 M Tris-HCl, 0.1 M NaCl, pH 9.5) for 3 min. The membrane was placed on an acetate sheet and 1.5 ml of a 0.25 M CDP-Star™ solution was evenly distributed over the membrane before another sheet of acetate was placed on top and the membrane was

incubated at room temperature for 5 min. The blot was visualised by taking a picture of the membrane using a light camera (EG&G BERTHOLD Luminograph LB980).

2.11 Membrane protein preparation

Crude membrane protein preparations were prepared as follows. A volume equevilent to an OD_{600} of 5 was pelleted from an overnight culture of *S. aureus* RN6390B, ROJ38 or ROJ38 pAgrC1agrA and the supernatant removed. The pellet was resuspended in 250µl of in PBS 80 µg/ml Lysostaphin by vortexing and incubated for 30 min at 37°C. Cooled on ice and sonicated for 2 x 15 s at 6-8 MSE. Centrifuged for 10 min at 13.000 g and the supernatant removed (soluble proteins). The Pellet was washed with 1 ml PBS, centrifuged at 13.000 g for 10 min resuspended and in 200 µl 1X SDS-PAGE sample buffer. 15 µl was loaded in each gel lane.

2.11.1 Elugent extraction of membrane proteins

Elugent extracts were prepared as described for membrane protein preperations with the pellet in the last step being resuspended in 100 µl 2 % vol/vol Elugent (Calbiochem, MERCK, UK,), instead of 1X SDS-PAGE buffer, and incubated at 37°C for 30 min and centrifuged at 13.000 g for 10 min. The pellet was resuspended and in 200 µl 2X SDS-PAGE sample buffer. 15 µl was loaded in each gel lane.

2.12 SDS-polyacrylamide gel electrophoresis

SDS-PAGE was conducted using the method of (Laemmli, 1970). 10 % acrylamide gels were prepared using 3.3 ml BioRad Protogel (30:08), 1.25 ml 3 M Tris-HCL pH 8.8, 4.8 ml sterile distilled water, 100 µl 10 % wt/vol sodium dodecyl sulphate (SDS), 75 µl 10 % wt/vol ammonium persulphate and 5 µl NNN'N'-tetramethylethylene diamine (TEMED). The mixture was then poured into a BioRad Mini-protean II casting tray, and overlaid with water saturated with butanol to minimise shrinkage of the gel due to evaporation. Upon polymerisation of the separating gel the butanol saturated water was discarded and the stacking gel was poured onto the separating gel. The 5 % acrylamide stacking gel was prepared using 750 µl Protogel (National Diagnostics, USA), 1.5 ml 0.5 M Tris-HCl pH 6.8, 60 µl 10 % wt/vol SDS, 10.3 ml water, 150 µl 10 % wt/vol ammonium persulphate and 6 µl TEMED.

Samples for analysis was mixed with an equal volume of 2 x sample buffer (0.25 M Tris-HCl, pH 6.8, 20 % vol/vol Glycerol, 4 % wt/vol SDS. 0.001 % wt/vol bromphenol blue, 100 mM DTT) and then boiled for 10 min. Samples were then loaded on to the gel, and the gel run at 150 V in Tris-glycine running buffer (2.88 % wt/vol glycin, 0.6 % wt/vol Tris, 0.2 % wt/vol SDS, pH 8.3). Gels were run until the dye present in the sample buffer tracked to the bottom of the gel.

2.13 Coomassie staining of polyacrylamide gels

SDS-PAGE gels were first fixed in destain solution (20 % vol/vol methanol; 7.5 % vol/vol acetic acid) for 10-15 min, then stained with Coomassie Brilliant Blue (20 % vol/vol methanol; 7.5 % vol/vol acetic acid, 0.025% Coomassie Brilliant Blue R250) and the subsequently destained in destain solution to reveal the protein profile and reduce background staining.

2.14 Protein identification

Protein identification of bands from Coomassie stained SDS-PAGE gels was conducted by MALDi TOF or QTof mass spectroscopy at Biopolymer Synthesis and Sequencing Unit, Queen's Medical Centre, University of Nottingham. Whole SDS-PAGE gels were delivered together with an image of the gel with the bands to be sequenced marked.

2.15 Western blotting

Western blotting was carried out as described in Sambrook *et al.*, (1989).

Proteins to be analysed by Western blotting were first separated by SDS-PAGE (2.12). Proteins were the transferred onto a BioTrace®NT nitrocellulose membrane (PALL life sciences, USA) using a BioRad II Mini-Cell and Electroblotting Module according to the manufacturers instructions. Transfer was carried out at 140 mA for 1.5 h at room temperature in Western blotting buffer (2.88 % wt/vol glycin, 0.6 % wt/vol Tris-HCL).

Blots were removed from the apparatus and blocked in PBS (pH 7.4) containing 5 % wt/vol skimmed milk powder overnight at 4˚C then incubated with primary antibody (1:500 dilution of rat antiserum immunised with the C-terminal fragment of AgrC) in PBS 5 % wt/vol skimmed milk powder for 1 h. Blots were subjected to 2 x 15 min washes in PBST (PBS pH 7.4 with 0.05% vol/vol Tween20) and then probed with the same buffer containing the secondary antibody (1:2000 rabit-anti-rat alkaline phosphatase conjugated (DAKO A/S, Denmark) in PBST) for 1 h. Blots were once again washed 2 x 15 min in PBST prior to being developed using ECL™ substrate (GE healthcare). Positive reactions were visualised by exposure to autoradiograph film (Kodak).

2.16 Bioluminescence agr activity assay

Reporter strain containing plasmids pSKermP2 with either wild type *agrC* or mutated *agrC* and *agrA* were grown o/n @ 37°C in CYGP medium supplemented with appropriate antibiotics. In the morning cultures were diluted 1/25 and grown @ 37°C for 2h after which 10 µl of the cultures were diluted 1/50 into wells in a 96 well microtiter plates containing a series dilution of AIP obtained from Dr. Weng Chan (CBS, University of Nottingham). Each plate contained two sets of a series dilution of AIP in triplicates. CYGP containing the appropriate antibiotic was then added to a total volume of 200 µl. The microtiter plates were then incubated in an Anthos Lucy 1 luminometer (Labtech) where light measurements and OD_{492} were taken every 30 min for 24h.

Data were plotted as Relative Light Units pr cell density (RLU/OD) over time.

EC_{50} and IC_{50} values were extracted from the sigmoidal dose-response curves using the PRISM2 PROGRAM (Graphpad). All assays were carried out at least in triplicate.

2.17 Light camera pictures

Pictures of plates with luminescent bacteria and southern blots were taken with an EG&G BERTHOLD Luminograph LB980 light camera connected to a computer using Luminograph LB980 v. 2.6 software.

2.18 Fluorescence microscopy

Pictures of *S. aureus* and *E. coli* carrying the plasmid pGfp-AgrCA were obtained using a Nikon eclipse TE2000-S (Nikon) fluorescence microscope linked to a computer running IPlab v3.7 (Scanalytics Inc, BD Biosciences, USA). Manipulation of the obtained pictures was done using IPlab v3.7.

3 Structure function study of the interactions of the histidine kinase sensor AgrC and AIP ligands

3.1 Introduction

The production of many surface proteins and secreted toxins in *Staphylococcus aureus* is controlled in a quorum sensing-dependent manner via the *agr* (accessory gene regulator) system. This system is activated by a diffusible signal peptide termed an Auto-Inducing Peptide (AIP) which is predicted to bind to and activate the histidine sensor kinase, AgrC. Four different AIPS have so far been identified (Figure 3-1), dividing *S. aureus* into four groups. Each group's native AIP activates its own AgrC and inhibits activation of those of other groups, a process known as cross-inhibition (Ji *et al.*, 1997;McDowell *et al.*, 2001). Activation of AgrC by an AIP leads to the phosphorylation of the response regulator AgrA, which binds to the intergenic promoter region and activates both *agr* promoters P2 and P3, leading to increased transcription from both promoters. The P3 transcript is the functional RNAIII which is responsible for the regulation of the genes in the *agr* regulon (Balaban and Novick, 1995;Ji *et al.*, 1995;McDowell *et al.*, 2001).

AgrC was first predicted to have five transmembrane domains and two extra-cellular loops by Lina *et al* (1998) based on data obtained using PhoA fusions and

Kyle-Doolittle hydrophobicity plots. Recently this suggested model has been questioned.

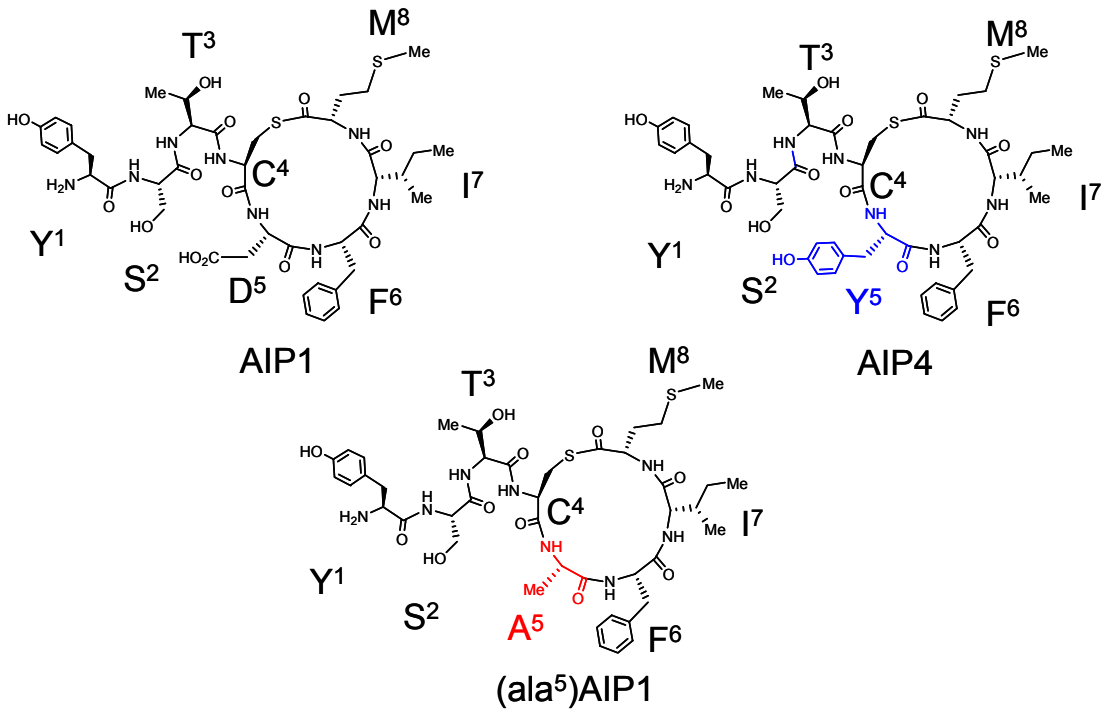


Figure 3-1: Structures of AIPs, (top) AIP-1 and AIP-4 and the structure of the modified AIP, (ala^5)AIP-1. In blue is shown the difference between AIP-1 which has an aspartic acid at position 5 and AIP-4 which has a tyrosine at positions 5. The alanine substitution of position 5 (ala^5)AIP and is known to be a potent inhibitor of all known *agr* groups (McDowell *et al.*, 2001;Lyon *et al.*, 2002) shown in red.

The initial aim of this work was to construct and validate an easy to use and reliable reporter to the study of AIP mediated activation and inhibition of AgrC, and to identify the key amino acids involved in these processes. For this purpose AgrC1/AIP-1 and AgrC4/AIP-4 were chosen due to their high degree of similarity of both AIPs and AgrCs. Only the amino acid in position five differs between the AIPs, aspartic acid (a polar acidic amino acid) in AIP-1 and tyrosine (a polar neutral amino acid) in AIP-4.

3.2 Results

3.2.1 Prediction of the transmembrane topology of AgrC

To gain insight into the key amino acid residues in AgrC likely to be involved in AIP recognition, the transmembrane topology of AgrC was first modelled using online prediction tools.

AgrC is encoded by the third gene in the *agr* P2 transcript. The size of the *agrC* gene depends on the *agr* group; *agrC1* is either 1245 bp or 1293 bp, *agrC2* is 1116 bp, *agrC3* and *agrC4* are 1293 bp. The *agrC* gene products are either a 371, 414 or 430 amino acid residue transmembrane proteins predicted to have molecular weight of 42.45 kDa, 48.0 kDa or 49.89 kDa. AgrC consists of two major domains, an N-terminal membrane spanning sensor domain and a C-terminal cytoplasmic histidine kinase domain, each approximately making up half the protein.

3.2.1.1 Selection of prediction server

To derive the most likely AgrC topology, the online prediction tools listed below were used (See http://www.expasy.org/tools/#topology and http://saier-144-37.ucsd.edu/memsat.html).

Table 3-1: Transmembrane topology prediction servers available online and used for the transmembrane topology prediction for staphylococcal AgrC in this study.

Name	Web address (as of 08/2008)	Reference
DAS (Stockholm University, Sweden)	http://www.sbc.su.se/~miklos/DAS/	(Cserzo *et al.*, 1997)
PredictProtein (Columbia University, USA)	http://www.predictprotein.org/	(Rost *et al.*, 1996)
SOSUI (Nagoya University, Japan)	http://bp.nuap.nagoya-u.ac.jp/sosui/	(Hirokawa *et al.*, 1998)
TMAP (Karolinska Institut; Sweden)	http://bioinfo4.limbo.ifm.liu.se/tmap/index.html	(Persson and Argos, 1994;Persson and Argos, 1996)
TMHMM (CBS; Denmark)	http://www.cbs.dtu.dk/services/TMHMM-2.0/	(Sonnhammer *et al.*, 1998;Krogh *et al.*, 2001)
TMpred (EMBnet-CH)	http://www.ch.embnet.org/software/TMPRED_form.html	(Hofmann and Stoffel W., 1994)
TopPred (France)	http://mobyle.pasteur.fr/cgi-bin/MobylePortal/portal.py?form=toppred	(von Heijne G., 1992;Claros and von, 1994)
MEMSAT (USA)	http://saier-144-37.ucsd.edu/memsat.html	(Jones *et al.*, 1994)
HMMTOP (Hungarian Academy of Sciences)	http://www.enzim.hu/hmmtop/html/adv_submit.html	(Tusnady and Simon, 2001)

The predicted amino acid sequence for *S. aureus* Newman AgrC1 was evaluated via each of eight prediction programs listed in Table 3-1.

A summary of the results is shown in Table 3-2 (for full results see Appendix A). For DAS two different cut off values are given by default giving two different possible predictions. For MEMSAT, a 7 TMH was given as the favoured structure but the score shown in the final result corresponded to the 6 TMH shown. HMMTOP was the only topology prediction server that predicted both the N- and the C- terminal to be extracellular.

Table 3-2: Summary of transmembrane topology predictions from selected online programs. Name of server, number of predicted transmembrane helices (TMH) and which amino acids are likely to be located in the TMH are shown for each program.

Prediction server name	No of TMH	Amino acid predicted to be in the TMH (N to C)							
		TMH1	THM2	THM3	THM4	THM5	THM6	THM7	THM8
TMHMM v2	6	7-29	49-71	78-100	110-132	145-167	188-204		
toppred	6	8-28	54-74	84-104	109-129	149-169	182-202		
TMpred	6	9-27	53-72	79-103	113-132	149-168	184-205		
Tmap	6	11-39	50-78	83-106	112-140	150-178	188-216		
SOSUI	7	6-28	34-56	60-82	85-107	110-132	148-170	182-204	
PHDhtm	6	12-30	41-65	70-94	110-133	149-167	183-201		
DAS (1.7)	6	18-37	47-62	67-113	119-142	158-177	193-212		
DAS (2.2)	8	19-36	48-60	69-82	89-98	105-111	120-141	159-176	194-211
MEMSAT	7	9-27	35-51	58-74	87-103	110-132	149-168	182-205	
MEMSAT	6	4-30	56-72	79-103	110-132	149-168	182-205		
HMMTOP	6	7-30	39-62	79-103	110-133	150-169	182-205		

From the eight prediction programs, six gave one possible AgrC transmembrane topology model and two gave two alternatives.

Information from multiple programs should help with an overall prediction and of the eleven predicted topologies, eight gave a six TMH topology, two gave a seven TMH topology and one predicted an eight TMH model. Based on this comparison, the age of the prediction programs and Möller *et al*'s (2001) large study of various prediction servers and evaluation of their accuracy using sequences from proteins with known structure, where the TMHMM 2.0 was shown to be the most accurate service available at the time of the article, (Moller *et al.*, 2001) TMHMM 2.0 was used for the further studies of the AgrC membrane topology and the AgrC-AIP structure function study.

3.2.2 Predicted transmembrane topology of several staphylococcal AgrCs

The next step in this study was to predict the transmembrane topology of different AgrC using sequences from available staphylococcal genomes (www.ncbi.nlm.nih.gov/ sites/entrez?db=genome). Sequences used in this study were five *S. aureus* AgrC1s, three *S. aureus* AgrC2s, three *S. aureus* AgrC3s, three *S. aureus* AgrC4s (None of the *S. aureus* genomes published is *agr* group four so the three sequences were selected by a gene search for AgrC and identified by the AgrD sequence as *agr* group four), two *S. epidermidis* AgrCs, one *S.*

haemolyticus AgrC, and one *S.saprophyticus* AgrC. The results are shown in Figure 3-2 to Figure 3-4

S. aureus AgrC1

Modelling the two different attenuated AgrC1 (414 aa: NCTC 8325 and USA300 and 430: Newman, COL and USA300_TCH1516, Source; NCBI, accession no. YP_500745, YP_494642, YP_001332979 and YP_001575910) from completed *S. aureus* genomes resulted in two different predicted transmembrane topologies, both with six transmembrane spanning helices and three extracellular loops with the N- and C-terminal amino acid residues located on the cytoplasmic side of the membrane. This six TMH structure for the 430 amino acid residue AgrC1 is consistent with data obtained using PhoA fusions which indicated that ^{33}Y, ^{105}K and ^{176}A is located on the outside of the membrane (Lina *et al.*, 1998) (Figure 3-3 A).

S. aureus AgrC2

Since AgrC2 (Accession no. Mu50: NP_372562, Mu3: YP_001442613 and N315: NP_375146) DNA sequence is shorter than the other three groups it is predicted to have four TMH resulting in two extra-cellular loops with N- and C-terminal ends of the protein on the cytoplasmic side of the membrane (Figure 3-3 B).

S. aureus AgrC3

The AgrC3 sequence (Accession no. Mw2: NP_646779, MRSA252: YP_041488 and MSSA476: YP_044054) has 430 amino acids and the prediction suggested six TMH and three extracellular loops with N- and C-terminal ends of the protein on the cytoplasmic side of the membrane (Figure 3-3 C).

S. aureus AgrC4

The AgrC4 sequence (Accession no. RN4850: ABB29292 ,H460: ABB17521 and H417: ABB17515) has 430 amino acids and the prediction suggested six TMH and three extracellular loops with N- and C-terminal ends of the protein on the cytoplasmic side of the membrane (Figure 3-3 D).

Figure 3-2: Predicted six transmembrane helices and three extracellular loops topology of AgrC1 and AgrC4. Circled amino acids are those that differ between AgrC group 1 and 4. * represents published PhoA fusions (Lina *et al.*, 1998) and all collaborated the structure predicted by TMHMM 2.0 shown in the figure.

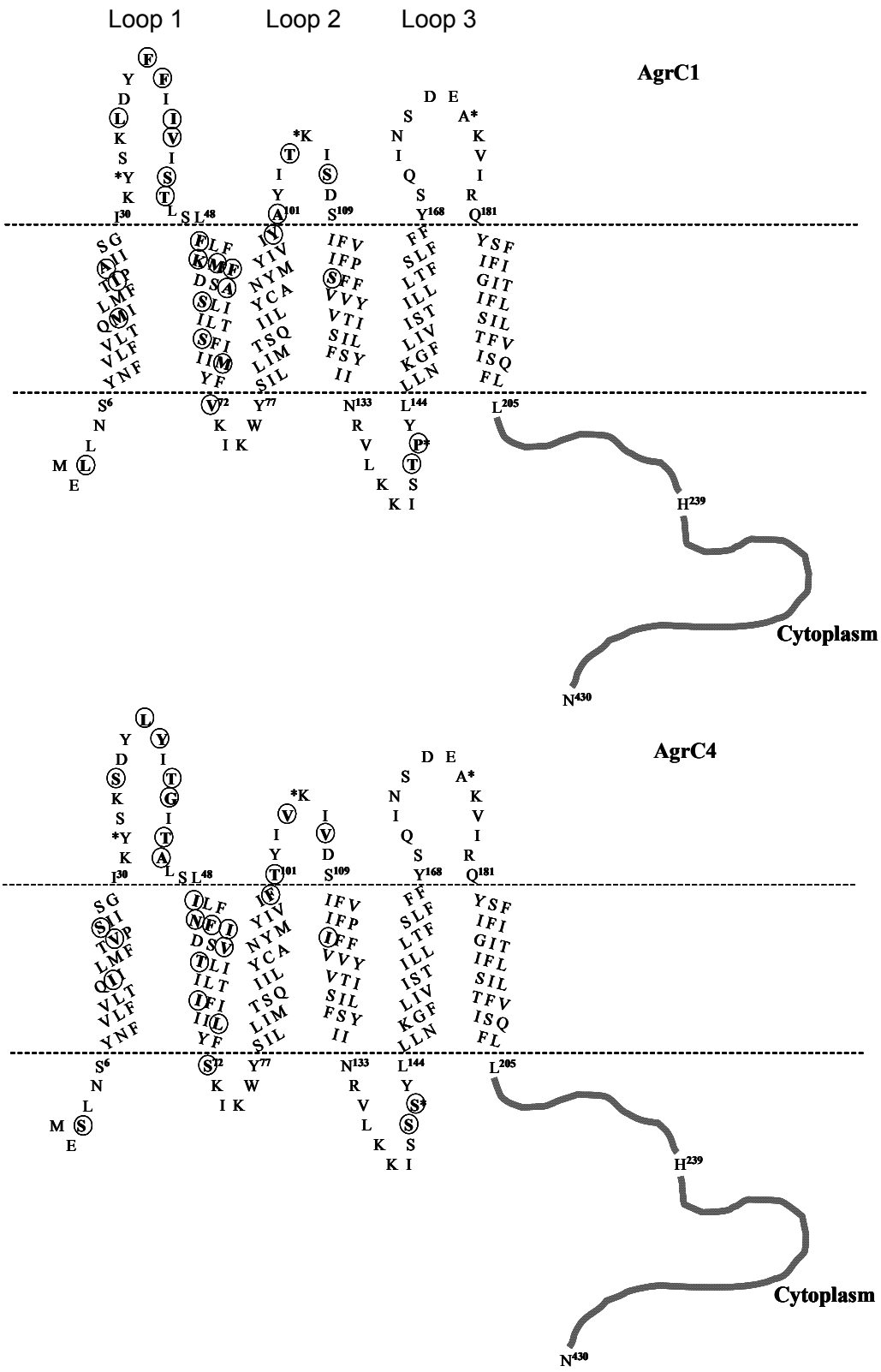
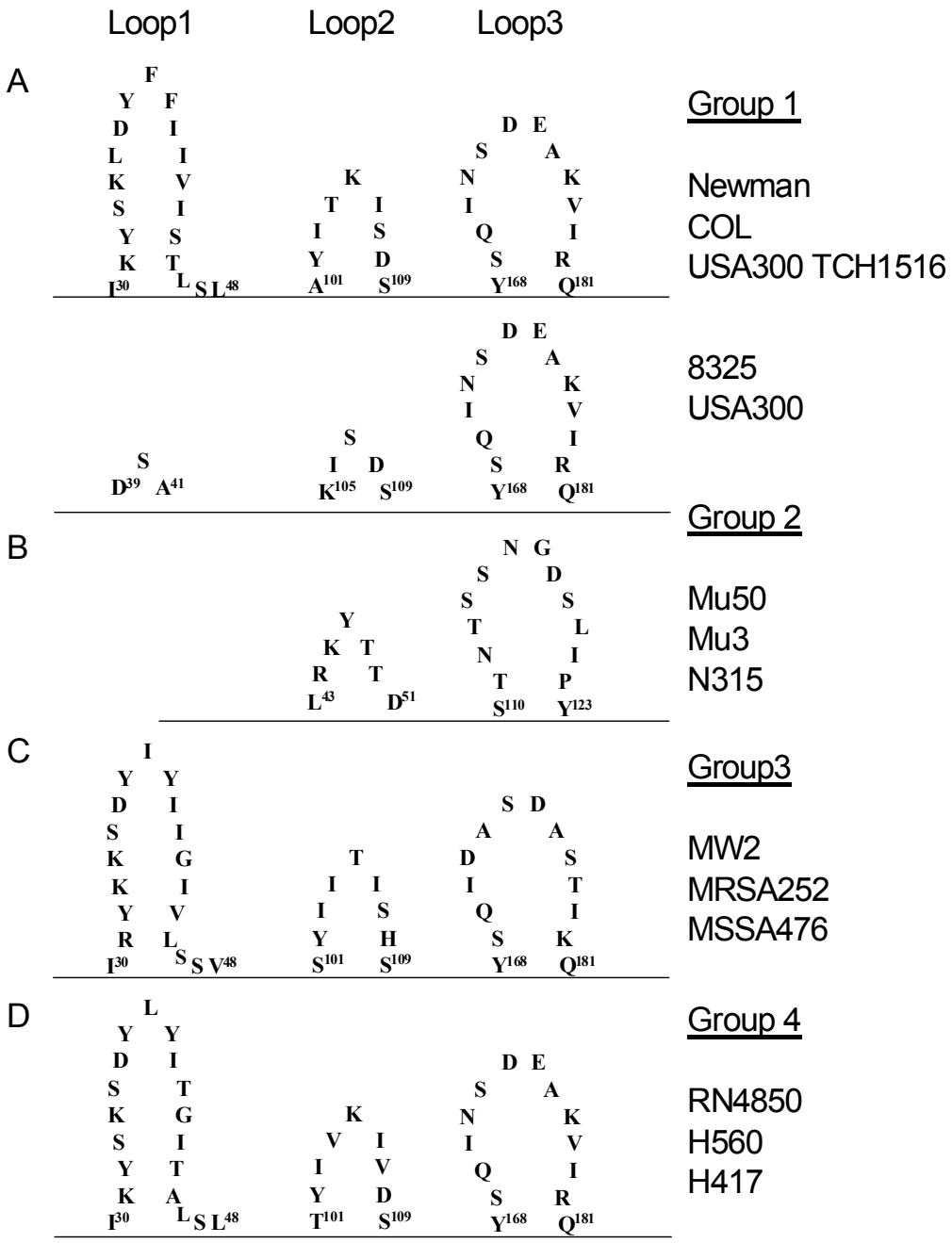

Loop 1
Loop 2
Loop 3
AgrC1
Cytoplasm
AgrC4
Cytoplasm

Figure 3-3: Predicted extracellular loops for *S. aureus*. A) AgrC1 from *S. aureus* Newman, COL and USA300 TCH1516 all have annotated genes coding for a 430 aa AgrC1. AgrC1 from *S. aureus* 8325 and USA300 both have annotated genes coding for a 414 aa AgrC1. B) Extracellular loops 1 and 2 for AgrC2. Due to the shorter annotation of the *agrC2* gene AgrC2 is only predicted to be 371 amino acids and contain four TMH insteas of six as predicted for AgrC1, AgrC3 and AgrC4. C) Predicted extracellular loops 1 to 3 for AgrC3. D) Predicted extracellular loops 1 to 3 for AgrC4.


Loop1
Loop2
Loop3
A
B
C
D
Group 1
Newman
COL
USA300 TCH1516
8325
USA300
Group 2
Mu50
Mu3
N315
Group3
MW2
MRSA252
MSSA476
Group 4
RN4850
H560
H417
I30 S L48
A101 S109
Y168 Q181
D39 A41
K105 S109
Y168 Q181
L43 D51
S110 Y123
I30 S S V48
S101 S109
Y168 Q181
I30 L S L48
T101 S109
Y168 Q181

Staphylococcus epidermidis AgrC

The *S. epidermidis* AgrC (Accession no. ATCC12228: NP_765192 and RP62A: YP_189058) is 429 amino acid long and has only ~51-54% sequence homology with *S. aureus* AgrCs with most variation in the N-terminal sensor domain. It is predicted to have seven TMH with the N-terminus and three loops on the extracellular side and the C-terminal histidine kinase domain on the cytoplasmic side of the membrane (Figure 3-4 A).

Staphylococcus haemolyticus and *Staphylococcus saprophyticus* AgrC

Both *S. haemolyticus* and *S. saprophyticus's* AgrC (Accession no. YP_252909 and YP_300930) are 431 amino acids and share ~52% and 47% sequence homology with *S. aureus* AgrC respectively. Both are predicted to have six TMH and three extracellular loops with both N-terminal and C-terminal ends on the cytoplasmic side of the membrane (Figure 3-4 B and C).

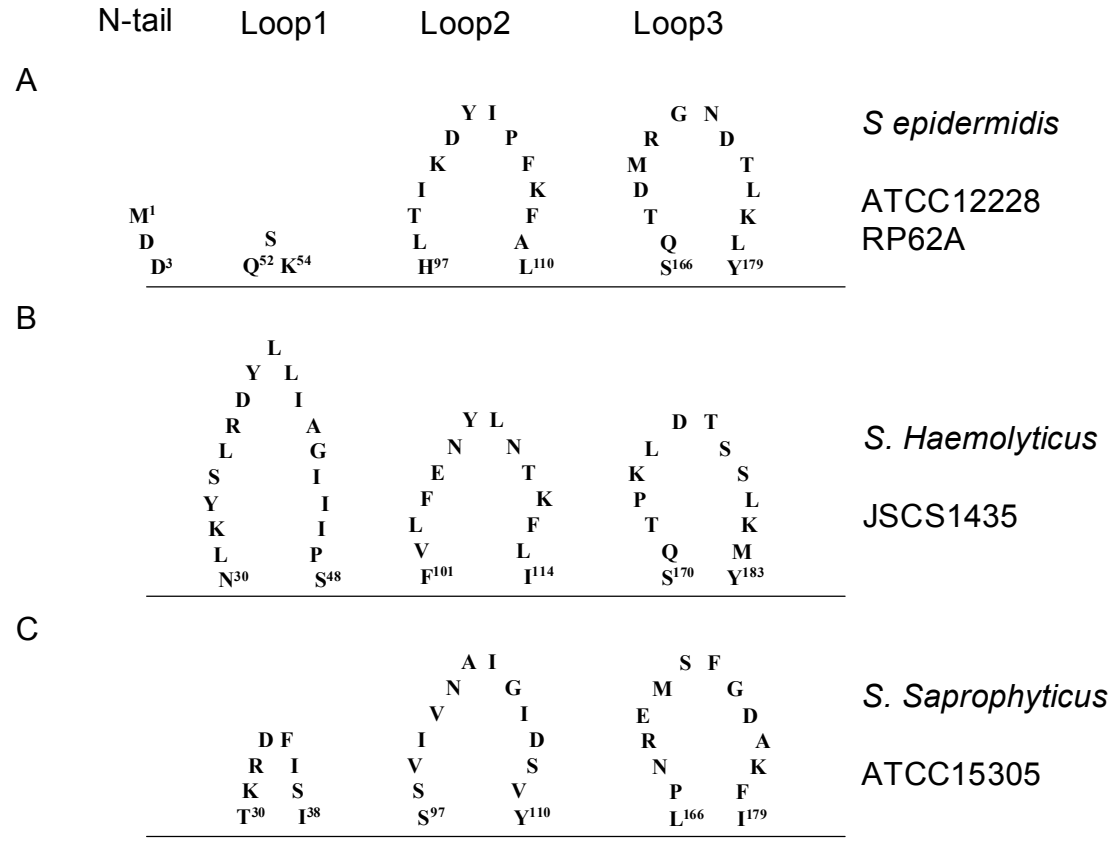


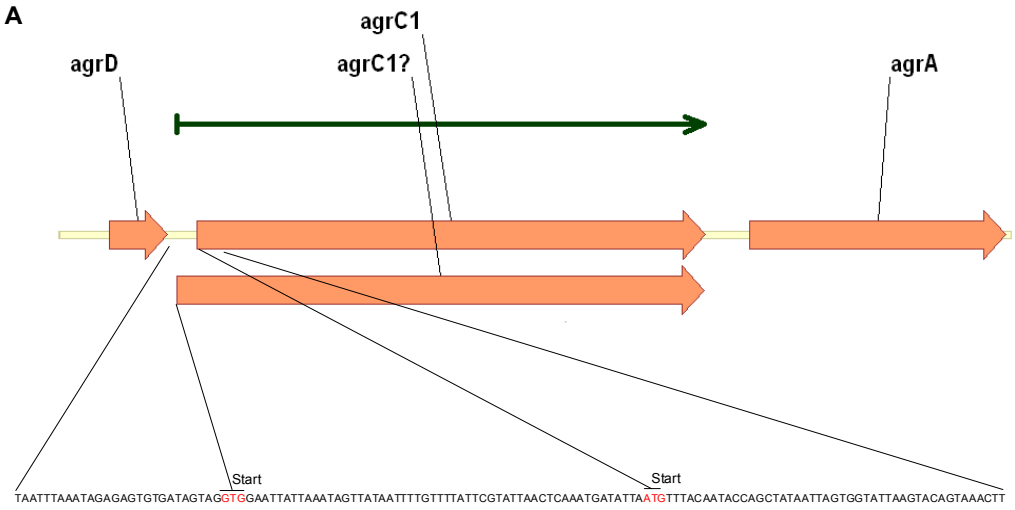
Figure 3-4: Predicted extracellular loops of A) *S. epidermidis* B) *S. hemolyticus* and C) *S. saprohyticus*.

For all the predicted staphylococcal AgrC structures loop3 contains 14 amino acids residues and loop 2 is 9-14 amino acid residues, where as loop 1 varies a great deal more, from 3 to 19 residues.

3.2.2.1 Alternative open reading frames for AgrC1 confer identical transmembrane topology

That there are two different genes coding for two different size AgrC1s both activated by the same AIP seems highly unlikely. Sequence analysis of the available *agrC1* genes from fully sequenced genomes revealed that there are alternative open reading frames for AgrC1 in both *S. aureus* 8325 and USA300, both having annotated AgrC1s of 414 amino acid residues giving 8325 and USA300 AgrC1s of 430 amino acid recidues, which is comparable with Newman, COL and USA300 TCH1516 and all AgrC3 and AgrC4s. Besides that it seems unlikely that there are two AgrC1s, more importantly this "new" AgrC1 is predicted to posses the same transmembrane topology as the AgrC1 from Newman, COL, USA300 TCH1516, AgrC3 and AgrC4.

Figure 3-5: A) graphic representation of *agrD, agrC and agrA* from *S. aureus* 8325 and USA300 with the two different size open reading frames for AgrC1; "agrC1" being the annotated ORF from NCBI (www.ncbi.nlm.nih.gov) and "agrC1?" being the larger ORF predicted using VectorNTI (Invitrogen, CA, USA) which has the same transmembrane topology as AgrC1 from strains Newman, COL and USA300 TCH1516 and AgrC3 and AgrC4. B) The general transmembrane topology of *S. aureus* AgrC1, AgrC3 and AgrC4.


A
agrC1
agrD
agrC1?
agrA
Start
Start
TAATTTAAATAGAGAGTGTGATAGTAGGTGGAATTATTAAATAGTTATAATTTTGTTTTATTCGTATTAACTCAAATGATATTAATGTTTACAATACCAGCTATAATTAGTGGTATTAAGTACAGTAAACTT
B
Loop1
Loop2
Loop3
TMH 1
TMH 2
TMH 3
TMH 4
TMH 5
TMH 6
Membrane
Cytoplasm
2HN
COOH

3.2.3 Identification of amino acids residues involved in the group specific activation and inhibition of AgrC

3.2.3.1 Construction of an AIP bioreporter

To facilitate identification of the amino acid residues involved in the AgrC/AIP interactions, a reliable low background AIP-bioreporter was needed, into which plasmids expressing *agrC* variants could easily be introduced. The initial strategy was to use suicide plasmids containing the *gfp* gene and the *luxABCDE* operon (pΔagr-dual, Figure 3-6) or just the *luxABCDE* operon (pΔagr-lux, Figure 3-6 and Figure 3-7) under the control of the $P3_{agr}$ promotor, *tetM* and the 1kb up- and down-stream regions of the *agr* locus to replace the entire *agr* locus. This was achieved by transforming the plasmids directly into *S. aureus* RN4220, resulting in single cross overs (SCO). The intention was to then transduce the engineered *agr* locus into *S. aureus* RN6390B using Φ11 and to force a second recombination event resulting in the complete replacement of *agr* with the *gfp-luxABCDE* / *luxABCDE* reporter genes.

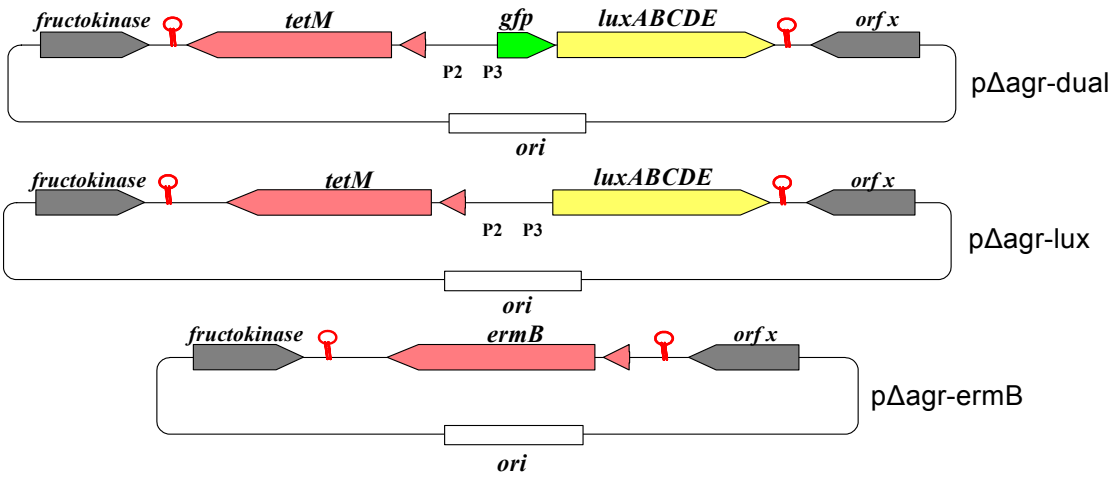


Figure 3-6: Plasmids used in the construction of the *agr* P3-*luxABCDE* reporter strain ROJ48. Plasmid pΔagr-dual was constructed by Klaus Winzer. Plasmid pΔagr-lux was made by excising the *gfp* gene using *SmaI* and *PvuII*. pΔagr-ermB is a derivative of pΔagr-tetM (constructed by Klaus Winzer, not shown) where the *tetM* gene has been excised using restriction enzymes *XhoI* and *EcoRI* and replaced by the *ermB* gene using the same restriction sites. The *ermB* gene was PCR amplified from pEC4 (Bruckner, 1997).

RN4220 transformants were screened by PCR for the presence of the *tetM* cassette on the chromosome. Four of the transformants which contained the *tetM* cassette were used to make Φ11 lysates. The lysates were used to transduce *S. aureus* RN6390B and the resulting transductants were screened by PCR for the *tetM* cassette. However this strategy failed to produce the desired double cross overs (DCO), at least at a frequency that was suitable for our screening methods (data not shown).

To make the screening process easier a *Δagr::ermB* stain was made by transforming pΔagr-ermB (Figure 3-6 and Figure 3-7) into RN4220, and PCR screening colonies growing on LB containing 5μg/ml erythromycin for the presence of the *ermB* gene on the chromosome. Φ11 lysates were prepared from four transformants positive for the *ermB* gene. The lysates were used to transduce *S. aureus* RN6390B. Transductants were screened by Southern blot for double cross over events, i.e., strains where the *agr* locus had been replaced by the *ermB* gene. A Southern blot of 6 transductants and their RN4220 SCO donor strains showed two of the transductants were DCOs (Figure 3-8).

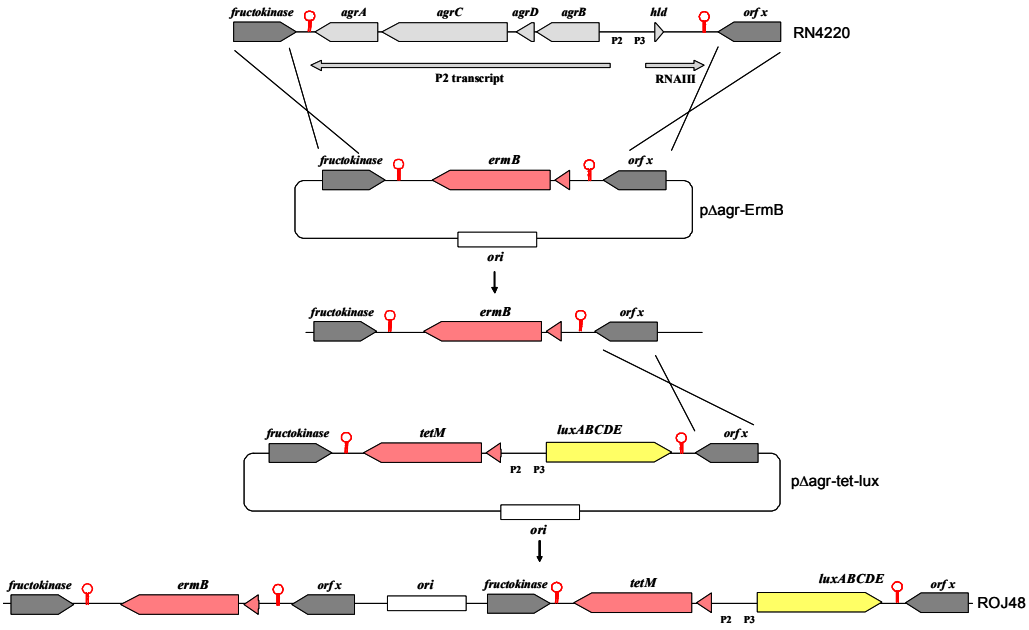


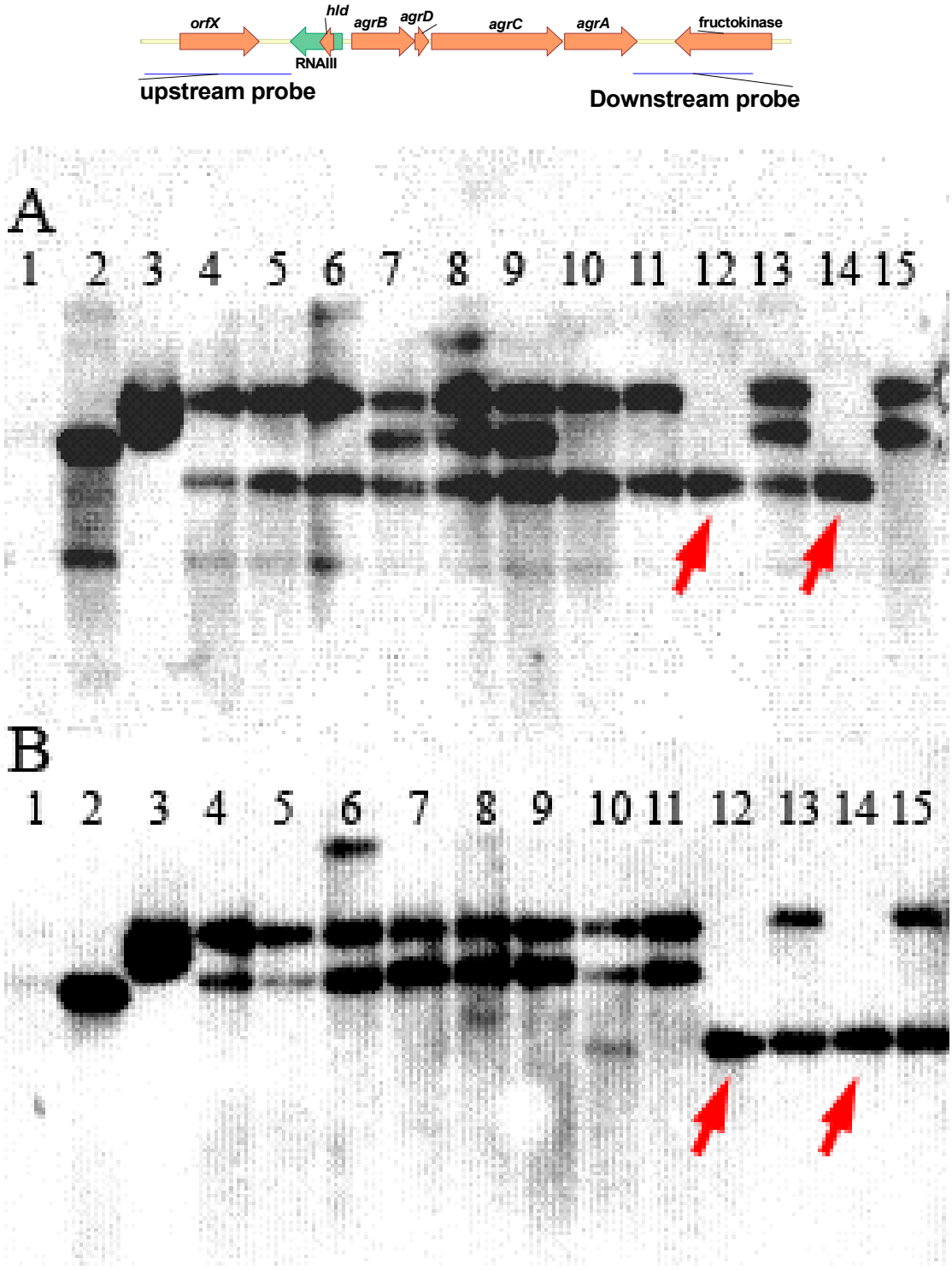
Figure 3-7: Diagram of the steps involved in the construction of the bio-reporter ROJ48. First a complete *agr* deletion was made in *S. aureus* RN4220 using plasmid pΔagr-ermB yielding the strain ROJ39. The plasmid pΔagr-tet-lux containing the *luxABCDE* operon under the control of the $P3_{agr}$ promoter was then introduced into ROJ39 yielding the final bio-reporter ROJ48.

Homologous recombination can occur when DNA duplexes align at regions of homologous sequence. It involves 5 steps 1) alignment of the homologous DNA duplex, 2) cleavage of a single strand from each duplex, 3) strand invasion, 4) sealing, resulting in a so called Holliday intermediate and 5) resolving of the Holliday structure. This can happen in two different ways one leading either to two molecules identical to the two starting molecules (e.g.,, a plasmid and a circular bacterial chromosome) or leading to the integration of the plasmid into the chromosome (Figure 3.7).

For gene replacement using suicide plasmids two of these events are necessary. The first recombination event will integrate the whole plasmid into either of the homologous regions (in figure 3.7 either in "*fructokinase*" of the "*orfX*"). The second (between the two other homologous regions) will delete the gene / region flanked by the two regions (figure 3.7 top) (Stryer L, 1997).

Figure 3-8: Southern blot of *PstI* digested genomic *S. aureus* DNAs and plasmid controls. A) Probed with ~1kb DIG labelled DNA upstream of *agrA* probe B) Probed with ~1kb DIG labelled downstream of agr probe. Red arrows indicate lanes with single band indicating a double cross over.

Lane 1: Marker (not DIG labelled). 2: RN6390B. 3: pΔagr-ErmB 4:RN4220 pΔagr-ErmB #1. 5: RN6390B X Φ11(RN4220 pΔagr-ErmB #1) #1. 6: RN6390B X Φ11(RN4220 pΔagr-ErmB #1) #2. 7: RN4220 pΔagr-ErmB #2. 8: RN6390B X Φ11(RN4220 pΔagr-ErmB #2) #1. 9: RN6390B X Φ11(RN4220 pΔagr-ErmB #2) #1. 10: RN4220 pΔagr-ErmB #3. 11: RN6390B X Φ11(RN4220 pΔagr-ErmB #3) #1. 12: RN6390B X Φ11(RN4220 pΔagr-ErmB #3) #2. 13: RN4220 pΔagr-ErmB #4. 14: RN6390B X Φ11(RN4220 pΔagr-ErmB #4) #1. 15: RN6390B X Φ11(RN4220 pΔagr-ErmB #4) #2


orfX
hld
agrB
agrD
agrC
agrA
fructokinase
RNAIII
upstream probe
Downstream probe
A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
B
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

This Southern blot shows two RN6390B transductants where only one of each flanking region is a different size to that of the parent RN6390B strain, indicating that they are DCOs. It also shows that the RN4220 strains have two or three copies of each flanking region; indicating they are SCOs.

Multiple copies of one of the flanking regions resulted from the suicide plasmid only inserting itself into one of the flanking regions, i.e., if a suicide plasmid inserts itself in the upstream region as a single event, the entire plasmid inserts itself on the chromosome (a single cross over, SCO). Two copies of the downstream region will be present on the chromosome until a second homologous recombination event has taken place between the two downstream regions. This second recombination event leads to the excision of the gene (or genes) flanked by the two regions on the chromosome and insertion of the genes and flanked by the homologous regions on the suicide plasmid, see Figure 3-7.

The *agr* flanking regions of the RN6390B *Δagr::ermB* transformants were sequenced. This showed that both flanking regions were defined, i.e., they were neither absent nor were there any additional nucleotides in the flanking regions (see appendix C for sequence alignment.). One strain, ROJ37, was selected for further studies. The *Δagr* deletion from ROJ37 was introduced into RN4220 by Φ11 transduction yielding the RN4220 *Δagr* strain ROJ39 (Figure 3-7).

To assay for the deletion of the *agr* locus through loss of AIP production all *Δagr* strain supernatants were tested for AIPs using an established β-lactamase assay (Novick *et al.*, 1995;McDowell *et al.*, 2001) and were found negative for AIP production (supernatant tested by Siti Hanna Muharram, data not shown). To introduce the reporter genes, plasmids pΔagr-dual and pΔagr-lux (Figure 3-6) were transformed into ROJ39 (RN4220 *Δagr::ermB*) and plated on LB 10 µg/ml tetracycline. This resulted in the *Δagr::ermB* $P3_{agr}$*-luxABCDE* (ROJ48) or *Δagr::ermB* $P3_{agr}$*-gfp-luxABCDE* (ROJ51) single cross over reporter strain. Φ11 lysates were made from both and transduced into RN6390B yielding RN6390B *Δagr* $P3_{agr}$*-luxABCDE* (ROJ134) and RN6390B *Δagr* $P3_{agr}$*-gfp-luxABCDE* (ROJ137). Each of these strains contains the entire AIP reporter plasmid which has been integrated on the chromosome via a single homologous recombination event. Attempts to force a second recombination event were not successful consequently the single cross over strain was used for further studies.

The two RN4220 derived reporter stains (ROJ48, *lux* only and ROJ51, *gfp-lux*) and one of the RN6390B derived strains (ROJ134, *lux* only) were tested for reporter activity by transforming plasmid pAgrC1agrA (Table 2-2) and running bioluminescence assays (section 2.16) with and with out the addition of 270 nM AIP-1. From these tests the bio-reporter ROJ48 was chosen for the study as it had the highest signal to background ratio (See Figure 3-9).

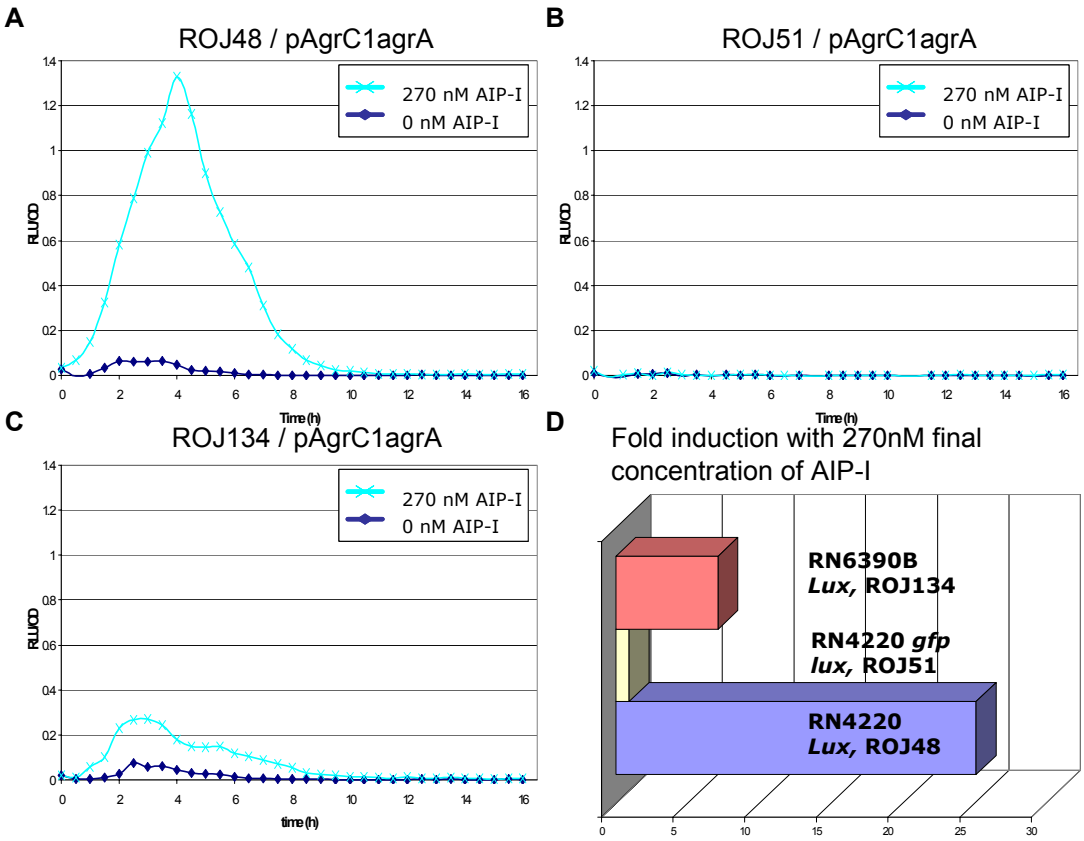


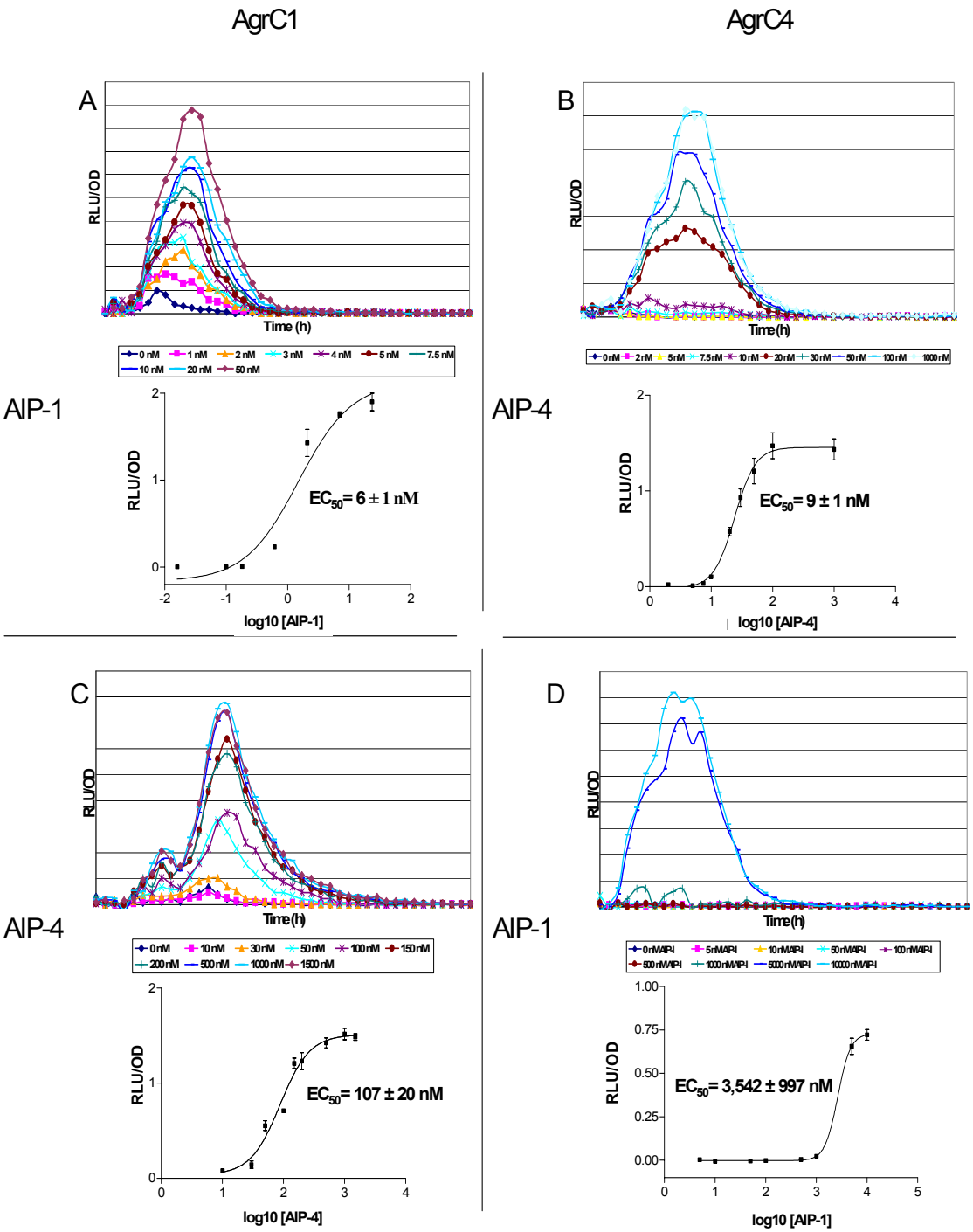
Figure 3-9: Activation of the bioreporters A) ROJ48 (RN4220 *Δagr::ErmB* P3$_{agr}$-*luxABCDE*), B) ROJ51 (RN4220 *Δagr::ErmB* P3$_{agr}$-*gfp-luxABCD* and C) ROJ134 (RN6390B *Δagr::ErmB* P3$_{agr}$-*luxABCD*) by AIP-1. Light output and OD were recorded via a Lucy1 (an automated bioluminometer–photometer, Anthos, Salzburg, Austria) every 30 min. The ratio between the maximum light output of the reporter activated with AIP-1 and the non activated reporter is shown in D.

3.2.3.2 Validation of the AIP-bioreporter

Before the newly constructed AIP bioreporter *S. aureus* ROJ48 could be used for the AgrC structure-activity studies, we needed to make sure that the data obtained from this reporter were comparable with that obtained using the established $P3_{agr}$ β-lactamase reporter (Novick *et al.*, 1995;McDowell *et al.*, 2001). For this purpose two wild type *agrC* genes (*agrC1,* and *agrC4)* were introduced on low-copy vectors together with *agrA* (pAgrC1agrA and pAgrC4agrA respectively, Table 2-2) into *S. aureus* ROJ48.

The resulting strains were then tested for $P3_{agr}$ activation by a range of concentrations of both AIP-1 and AIP-4 and the antagonistic effect (ala^5)AIP-1, figures 3-10 and 3-11 and Table 3-3. EC_{50} and IC_{50} obtained and compared with the values reported in the literature.

Figure 3-10: AIP dependent activation of ROJ48 containing A) pAgrC1agrA activated by AIP-1 curve B) pAgrC4agrA activated by AIP-4 C) pAgrC1agrA activated by AIP-4 and D) pAgrC4agrA activated by AIP-1. Top part is activation plotted over time in the presence or absence (ala^{5})AIP-1 from 0nM to 10µM underneath is the corresponding does response curves .


AgrC1
AgrC4
A
B
RLU/OD
Time (h)
0 nM 1 nM 2 nM 3 nM 4 nM 5 nM 7.5 nM 10 nM 20 nM 50 nM
AIP-1
AIP-4
EC50= 6 ± 1 nM
EC50= 9 ± 1 nM
log10 [AIP-1]
log10 [AIP-4]
C
D
0nM 10nM 30nM 50nM 100nM 150nM 200nM 500nM 1000nM 1500nM
EC50= 107 ± 20 nM
EC50= 3,542 ± 997 nM

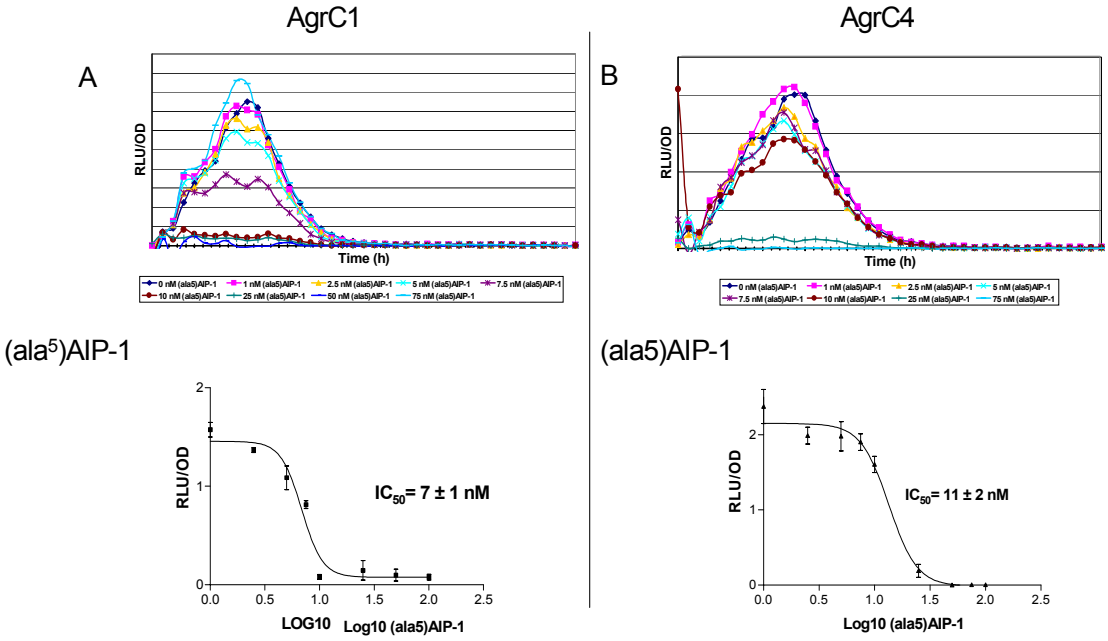


Figure 3-11: AIP dependent competitive inhibition of ROJ48 containing either A) pAgrC1agrA or B) pAgrC4agrA. Top part is activation plotted over time in the presence or absence (ala^5)AIP-1 from 0nM to 100nM underneath is the corresponding does response curves. The reporter was activated with ten times EC_{50} of either AIP-1 or AIP-4, 60 nM and 90 nM respectively.

For a comparison of previously published EC_{50} and IC_{50}s with the data obtained from ROJ48 see Table 3-3.

Table 3-3: Summary of EC_{50} and IC_{50} values for AIP-1 and AIP-4 used as an agonist for either AgrC1 or AgrC4 respectively in activation assays and (ala^5)AIP-1 an antagonist in competitive inhibition assays.

	Published			*agrP3::luxABCDE*		
	EC_{50}		IC_{50}	EC_{50}		IC_{50}
	AIP-1	AIP-4	(ala^5)AIP-1	AIP-1	AIP-4	(ala^5)AIP-1
AgrC1	10.2 ± nM[3)] 19 ± nM[2)] 28 ± nM[1)]	62 ± nM[1)]	5n ± 2 nM[1)] 33 nM[2)]	6 ± nM	107 ± nM	7 ±1 nM
AgrC4	26 ± μM[1)]	13 ± nM[1)]	3 ± 2 nM[1)]	3.5 ± μM	9 ± nM	11 ± 2 nM

1) Lyon *et al* (2002), 2)McDowell *et al.* (2001), 3) Mayville *et al.*, (1999)

The data in Table 3-3 show that AIPs activated their native AgrCs in the low nanomolar range and that (ala^5)AIP-1 inhibits both AgrC1 and AgrC4. This shows that ROJ48 AIP-bioreporter has a high degree of selectivity and sencitivity and is activated and inhibited at similar AIP concentrations to those obtained using a β-lactamase reporter.

However, the advantage of using ROJ48 is that there is no background *agr* activity and no additional reagents are required

3.2.4 Role of the predicted extra-cellular loops in group specific activation and inhibition of AgrC

3.2.4.1 Introduction

To identify the amino acids important for group specific activation and inhibition of AgrC, it was decided to base the analysis on a comparison of AgrC1 and AgrC4 as they are the most closely related AgrC receptors and their AIPs differ with only one amino acid. Figure 3-2, show that there are six amino acids that differ in the first extra-cellular loop, three in the second and the third extra-cellular loop is identical in AgrC1 and AgrC4. Since the second loop has the least variable it was decided to use this as the starting point for the investigation.

Plasmid constructs carrying *agrC* and *agrA* similar to the those used for the validation of ROJ48 (see section 3.2.3.2) were used. A total of 14 constructs were made using AgrC1 and AgrC4 as backbones and changing one, two or all three amino acids that differ between the two groups using site-directed mutagenesis (primers listed in Table 2-3). These modified AgrC AgrA constructs under the control of the $P2_{agr}$ promoter were introduced into ROJ48 as described in section

2.8.4 and 2.16. EC_{50}'s and IC_{50}'s for these constructs were determined for AIP-1, AIP-4 and (ala^5)AIP-1.

3.2.4.2 Contribution of the second extra-cellular loop of AgrC4 to AIP-dependent activation

Using AgrC4 as the AgrC backbone the three different amino acids in the second extra-cellular loop were changed to those of group 1; ^{101}T to ^{101}A, ^{104}T to ^{104}V and ^{107}S to ^{107}V.

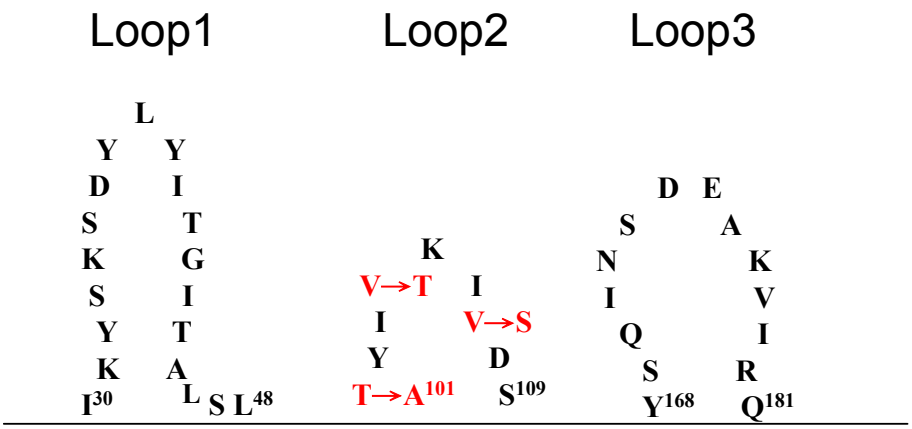


Figure 3-12: Amino acids changed in the second extracellular loop of AgrC4.

Subsequently single and double mutations were made and assayed to determine the contribution of each amino acid to AgrC activation. The results are shown below in Figure 3-13 to Figure 3-15 and summarised in Table 3-4.

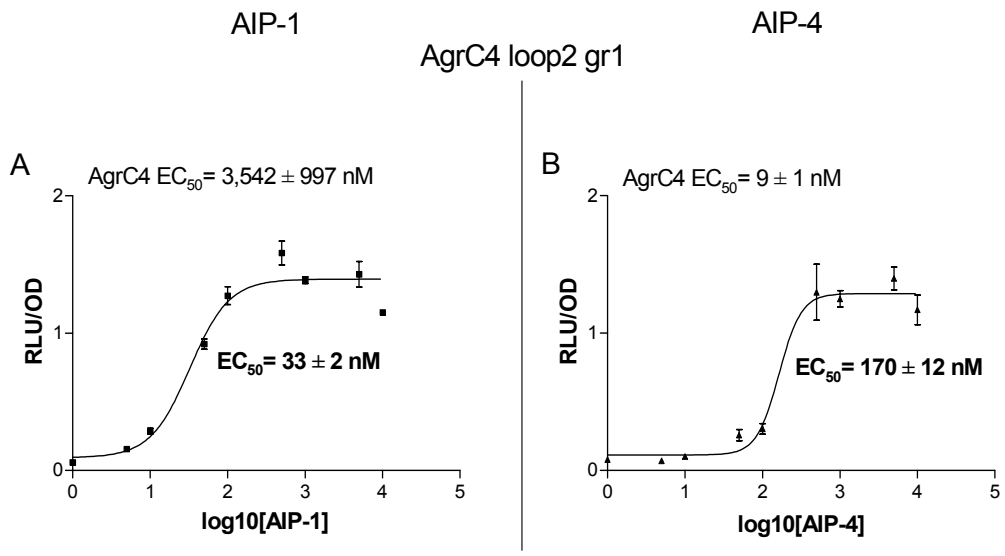


Figure 3-13: Dose response curves for activation of ROJ48 bio-reporter via AgrC4 loop2 gr1 (AgrC4T101AT104VS107V) by A) AIP-1 or B) AIP-4. AIP concentrations ranged from 0nM to 10µM. Error bars represent standard deviations where n=3.

From the data used to plot Figure 3-13, EC_{50}s were calculated for AIP-1 and AIP-4 acting on AgrC4loop2gr1 (Table 3-4) using Prism2 (GraphPad, San Diego, USA). The EC_{50} for AIP-1 was calculated to be 33 nM and the EC_{50} for AIP-4 was calculated to be 170 nM. This shows a clear shift towards group 1 behaviour with a 107-fold increase in EC_{50} for AIP-1 and a 19 fold decrease for AIP-4. This shows that the second extra-cellular loop of AgrC is important for group specific activation.

To further investigate the respective roles of the three amino acids that differ between AgrC1 and AgrC4 in loop2, a series of single and double mutations were

made in AgrC4. The effect of AIP-1 and AIP-4 on the activation of these mutated AgrC4s was tested using the bioreporter ROJ48 (Figure 3-14 and Figure 3-15).

Table 3-4 Activation of native and mutated AgrC group 1 and 4 by AIP-1 and AIP-4 respectively (EC_{50} nM; n=3), and the fold increase or decrease observed for the activation of mutated AgrC. ↑ indicates an increased activation of the mutated AgrC (decrease in the EC_{50}). ↓ indicates a decreased ability to activate the mutated AgrC (increase in the EC_{50}). → indicates no difference in the ability to activate the mutated AgrC.

Native and mutated AgrC4	EC_{50} (nM)		Fold increase / decrease in activity	
	AIP-1	AIP-4	AIP-1	AIP-4
AgrC4 (wild type)	3542 ± 997	9 ± 1	-	-
T101A V104T V107S (loop2 gr1)	33 ± 2	170 ± 12	107 ↑	19 ↓
T101A	754 ± 47	29 ± 1	4.7 ↑	3.2 ↓
V104T	6357 ± 349	512 ± 57	1.8 ↓	59 ↓
V107S	451 ± 30	9 ± 3	7.8 ↑	→
T101A V104T	12064 ± 1654	587 ± 10	3.4 ↓	62 ↓
T101A V107S	407 ± 45	22 ± 3	8.7 ↑	2.4 ↓
V104T V107S	4335 ± 457	551 ± 48	1.2 ↓	61 ↓
Native and mutated AgrC1				
AgrC1 (wild type)	6 ± 1	107 ± 20	-	-
A101T T104V S107V (loop2 gr4)	110 ± 7	94 ± 7	18 ↓	1.1 ↑
A101T	30 ± 2	22 ± 5	5 ↓	4.8 ↑
T104V	235 ± 38	269 ± 49	39 ↓	2.5 ↓
S107V	150 ± 15	408 ± 4	25 ↓	3.8 ↓
A101T T104V	47 ± 2	15 ± 2	7.8 ↓	7 ↑
A101T S107V	99 ± 10	19 ± 3	16.5 ↓	5.6 ↑
T104V S107V	120 ± 16	466 ± 9	20 ↓	4.3 ↓

Figure 3-14: Dose response curves for activation of ROJ48 via AgrC4 carrying single amino acid mutations in loop2. A) AgrC4T101A activated by AIP-1 B) AgrC4T101A activated by AIP-4 C) AgrCV104T activated by AIP-1 D) AgrC4V104T activated by AIP-4 E) AgrC4V107S activated by AIP-1 F) AgrC4V107S activated by AIP-4. Added AIP concentration ranged from 0nM to 10μM in a total. Error bars represent standard deviations where n=3.

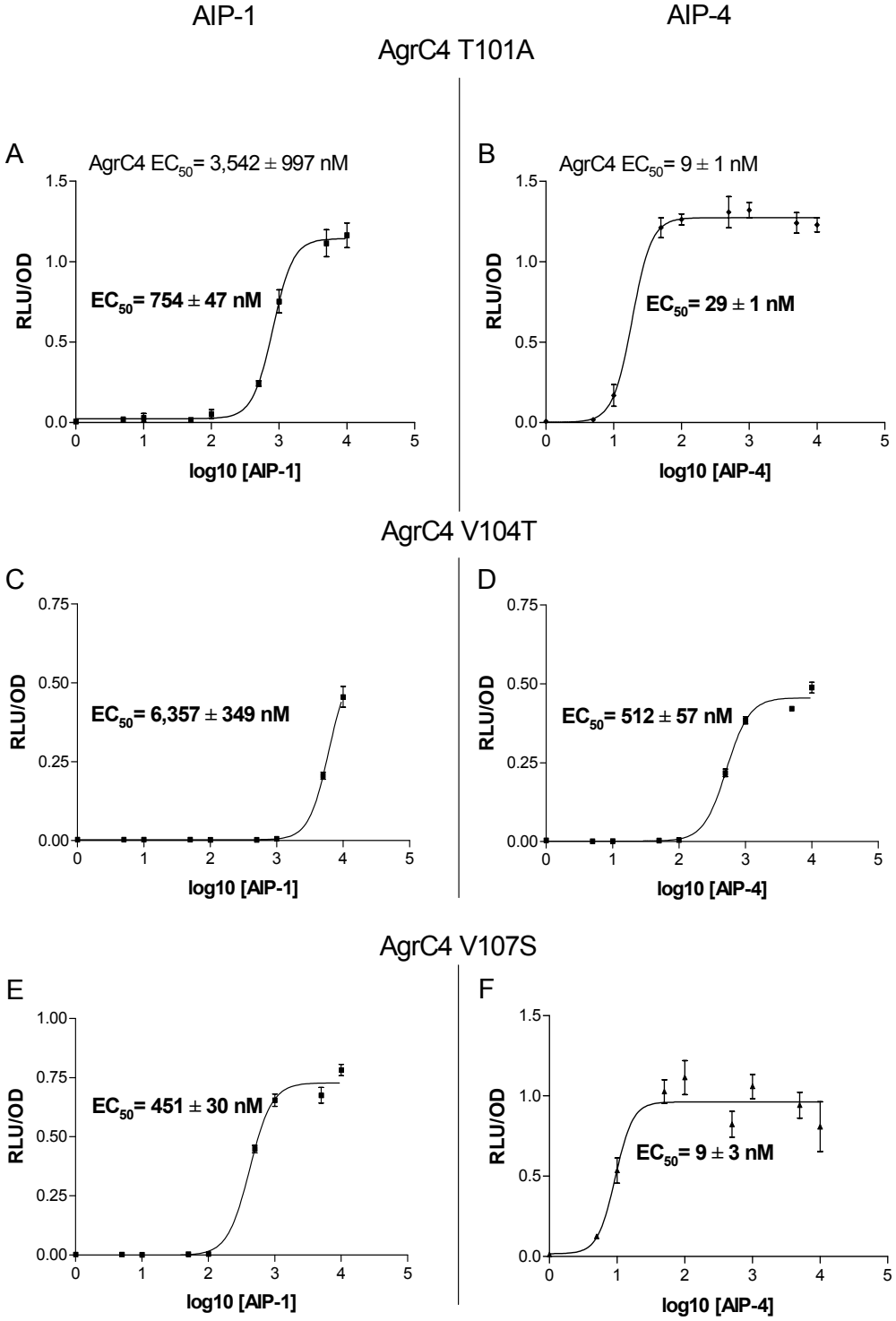
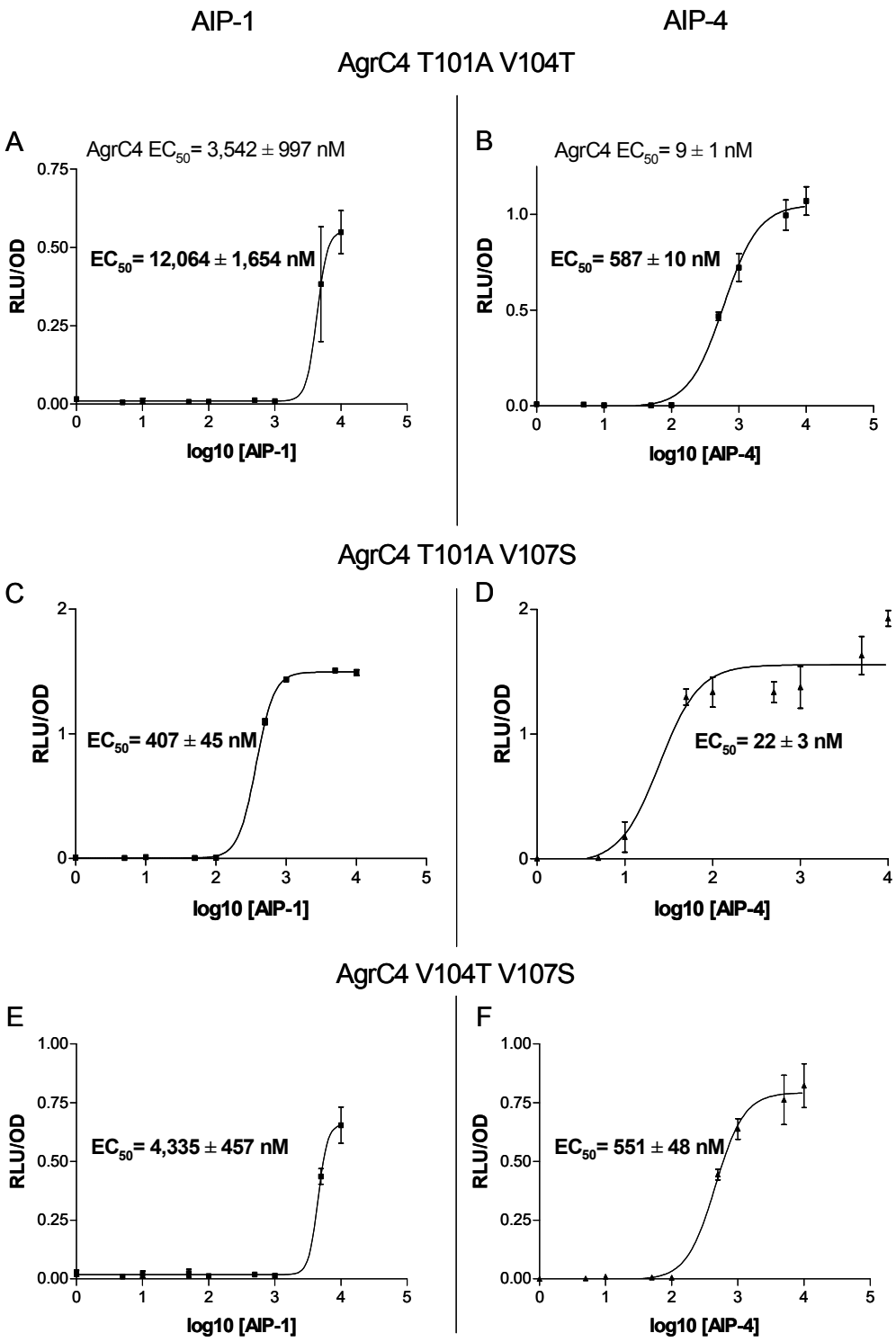

AIP-1
AIP-4
AgrC4 T101A
A
AgrC4 EC50= 3,542 ± 997 nM
EC50= 754 ± 47 nM
RLU/OD
log10 [AIP-1]
B
AgrC4 EC50= 9 ± 1 nM
EC50= 29 ± 1 nM
RLU/OD
log10 [AIP-4]
AgrC4 V104T
C
EC50= 6,357 ± 349 nM
RLU/OD
log10 [AIP-1]
D
EC50= 512 ± 57 nM
RLU/OD
log10 [AIP-4]
AgrC4 V107S
E
EC50= 451 ± 30 nM
RLU/OD
log10 [AIP-1]
F
EC50= 9 ± 3 nM
RLU/OD
log10 [AIP-4]

Figure 3-15: Dose response curves for activation of the ROJ48 via AgrC4 carrying double amino acid mutations in loop2. A) AgrC4T101AV104T activated by AIP-1 B) AgrC4T101AV104T activated by AIP-4 C) AgrCT101AV107S activated by AIP-1 D) AgrC4T101AV107S activated by AIP-4 E) AgrC4V104TV107S activated by AIP-1 F) AgrC4V107SV104T activated by AIP-4. Added AIP concentration ranged from 0nM to 10μM. Error bars represent standard deviation where n=3.


AIP-1
AIP-4
AgrC4 T101A V104T
A
AgrC4 EC50= 3,542 ± 997 nM
EC50= 12,064 ± 1,654 nM
RLU/OD
log10 [AIP-1]
B
AgrC4 EC50= 9 ± 1 nM
EC50= 587 ± 10 nM
RLU/OD
log10 [AIP-4]
AgrC4 T101A V107S
C
EC50= 407 ± 45 nM
RLU/OD
log10 [AIP-1]
D
EC50= 22 ± 3 nM
RLU/OD
log10 [AIP-4]
AgrC4 V104T V107S
E
EC50= 4,335 ± 457 nM
RLU/OD
log10 [AIP-1]
F
EC50= 551 ± 48 nM
RLU/OD
log10 [AIP-4]

From the data used to plot Figure 3-14 and Figure 3-15, EC_{50}s were calculated for AIP-1 and AIP-4 acting on AgrC4 with single or double mutations in the second extra-cellular loop (Table 3-4). From the EC_{50}s it can be seen that changing the valine at position 104 to a threonine (V104T) resulted in a 1.8 fold decrease in EC_{50} when using AIP-1 to activate AgrC and a 57 fold decrease for AIP-4.

A similar result was obtained if the valines at positions 104 and at position 107 were changed to threonine and serine, respectively. The EC_{50} for AIP-1 decreased 1.2 fold from 3542 nM to 4335 nM and for AIP-4 decreased by 61 fold from 9 nM to 551 nM (Table 3-4). If the valine for threonine substitution at position 104 was combined with changing the threonine at position 101 to an alanine, the decrease in EC_{50} was even more pronounced, resulting in a 3.4 fold decrease in EC_{50} for AIP-1 from 3342 nM to 12064 nM and a 62 fold decrease for AIP-4 from 9 nM to 587 nM. In contrast, when the threonine at position 101 was substituted with an alanine (T101A) or the valine at position 107 was substituted with a serine (V107S), a 4.7 fold increase in the EC_{50} were obtained (3542 nM to 754 nM) and 7.8 fold (3542 nM to 451 nM) respectively. This increase in EC_{50} for AIP-1 was accompanied by a relatively small reduction in the EC_{50} for the T101A substitution and no change for V107S substitution when AIP-4 was the agonist.

These observations suggest that for AgrC4 ^{104}V plays an important role in activation where as it is ^{101}T and ^{107}V that are key to the group specificity.

3.2.4.3 Contribution of the second extra-cellular loop of AgrC1 to AIP-dependent activation

To complete the analysis of the second extracellular loop, a set of mutants using AgrC1 as the AgrC backbone were made, complementary to those previously described for AgrC4. AIP-1 and AIP-4 dependent activation of these construct was tested using ROJ48. The results of these experiments are shown in Figure 3-16 to Figure 3-18 and data summarised in Table 3-4.

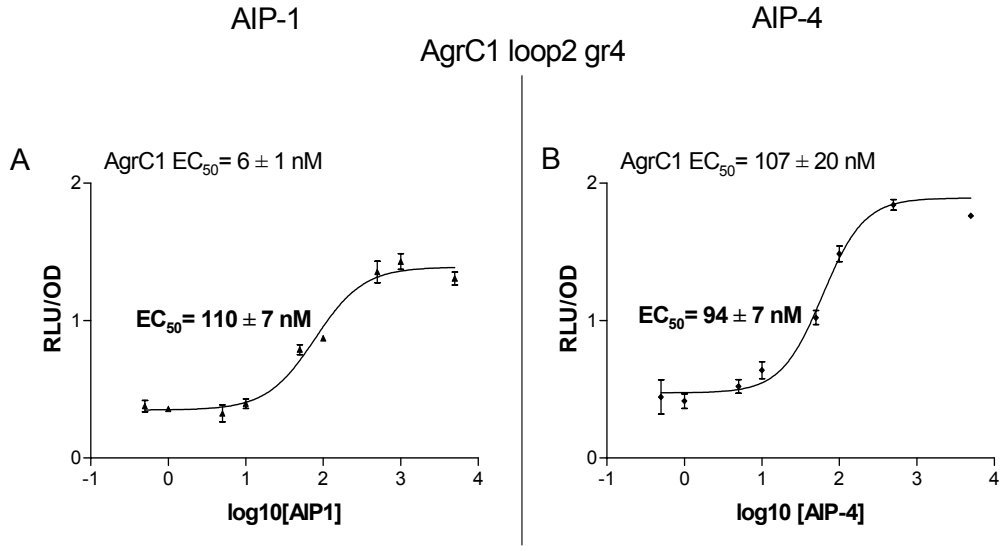


Figure 3-16: Dose response curves for activation of the ROJ48 via AgrC1 loop2 gr4 (AgrC1A101TV104TV107S) by A) AIP-1 or B) AIP-4. Exogenous AIP concentrations ranged from 0nM to 10µM. Error bars represent standard deviation where n=3.

Figure 3-17: Dose response curves for activation of ROJ48 via AgrC1 carrying a single amino acid substitution in loop2 A) AgrC1 A101T activated by AIP-1 B) AgrC1 A101T activated by AIP-4 C) AgrC1 T104V activated by AIP-1 D) AgrC1 T104V activated by AIP-4 E) AgrC1 S107V activated by AIP-1 F) AgrC1 S107V activated by AIP-4. Added AIP concentration ranged from 0nM to 10μM. Error bars represent standard deviation where n=3.

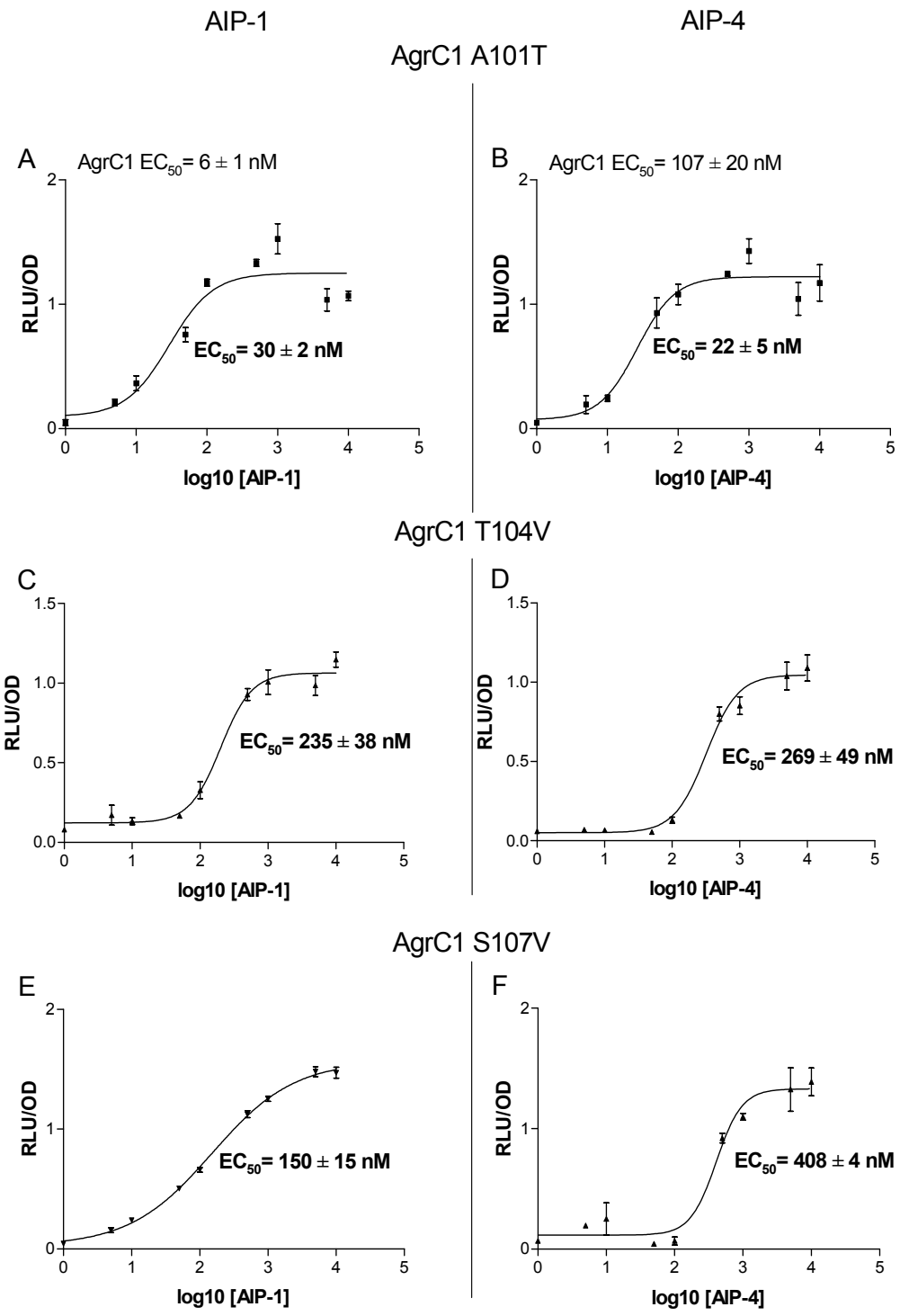
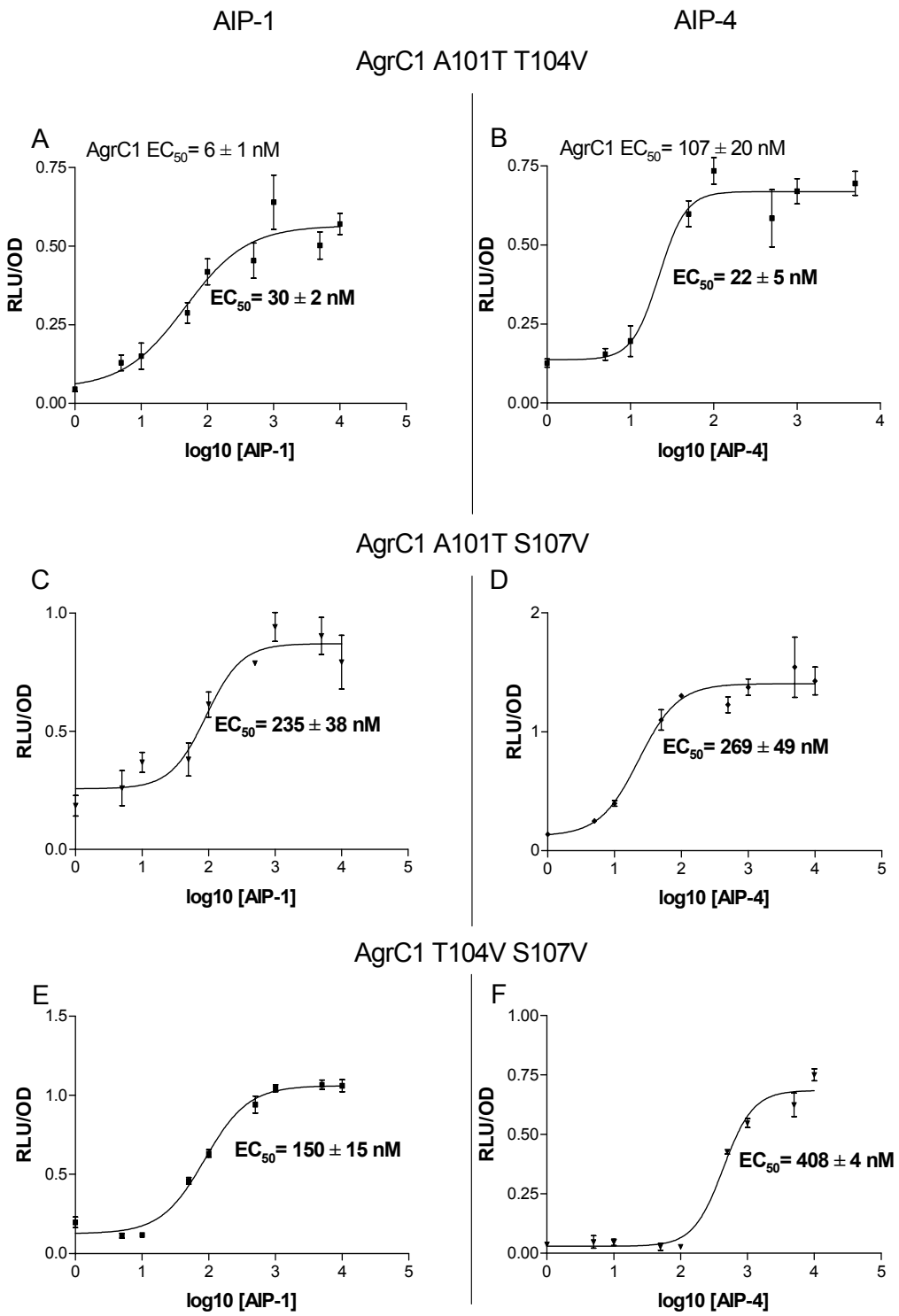

AIP-1
AIP-4
AgrC1 A101T
A
AgrC1 EC50= 6 ± 1 nM
RLU/OD
EC50= 30 ± 2 nM
log10 [AIP-1]
B
AgrC1 EC50= 107 ± 20 nM
RLU/OD
EC50= 22 ± 5 nM
log10 [AIP-4]
AgrC1 T104V
C
RLU/OD
EC50= 235 ± 38 nM
log10 [AIP-1]
D
RLU/OD
EC50= 269 ± 49 nM
log10 [AIP-4]
AgrC1 S107V
E
RLU/OD
EC50= 150 ± 15 nM
log10 [AIP-1]
F
RLU/OD
EC50= 408 ± 4 nM
log10 [AIP-4]

Figure 3-18: Dose response curves for activation of ROJ48 via AgrC1 carrying two amino acid substitutions in loop2 A) AgrC1 A101TT104V activated by AIP-1 B) AgrC1 A101TT104V activated by AIP-4 C) AgrC1 A101TS107V activated by AIP-1 D) AgrC1 A101TS107V activated by AIP-4 E) AgrC1 T104VS107V activated by AIP-1 F) AgrC1 T104VS107V activated by AIP-4. Added AIP concentration ranged from 0nM to 10µM. Error bars represent standard deviation where n=3


AIP-1
AIP-4
AgrC1 A101T T104V
A
AgrC1 EC50= 6 ± 1 nM
EC50= 30 ± 2 nM
RLU/OD
log10 [AIP-1]
B
AgrC1 EC50= 107 ± 20 nM
EC50= 22 ± 5 nM
log10 [AIP-4]
AgrC1 A101T S107V
C
EC50= 235 ± 38 nM
D
EC50= 269 ± 49 nM
AgrC1 T104V S107V
E
EC50= 150 ± 15 nM
F
EC50= 408 ± 4 nM

The data shown in Figure 3-16 to Figure 3-18 were used to calculate EC_{50}s for the different AgrC1 mutations and these are summarised in Table 3-4.

These data shows that changing the second extra-cellular loop of AgrC1 to that of AgrC4 does not result in the same significant shift in group specific activity observed with AgrC4. Here we observed little difference in EC_{50} for AIP-4 acting on the hybrid receptor (from EC_{50} 107 nM to 94 nM). However, as for AgrC1 there was an 18 fold decrease in the activity of the native AIP (AIP-1) (from EC_{50} 6 nM to 110 nM).

If the alanine at position 101 is changed to a threonine (A101T) a 5 fold decrease in the EC_{50} from 6 nM to 30 nM for AIP-1, and a 4.8 fold increase for AIP-4 from 107 to 22 nM for AIP-4 was noted. This was the only increase in EC_{50} for AIP-4 noted for any of the single amino acid substitutions. A 2.5 fold decrease in EC_{50} was observed if the threonine at position 104 was changed to a valine (T104V; EC_{50} 107 nM to 269 nM) and a 3.8 fold decrease from 107 nM to 408 nM if the serine at position 107 was changed to a valine (S107V). A similar but much more pronounced effect was noted for AIP-1 as an agonist with a 39 fold decrease in EC_{50} from 6 nM to 235 nM for T104Vand a 25 fold decrease in EC_{50} from 6 nM to 150 nM for S107V.

For double mutations a similar picture was observed where the EC_{50} for the native agonist (AIP-1) was reduced for all constructs: 7.8 fold increase for A101TT104V

(EC_{50} of 6 nM to 47 nM), 16.5 fold for A101T S107V (EC_{50} of 6 nM to 99 nM), and 20 fold for T104V S107V (EC_{50} of 6 nM to 120 nM). For AIP-4 a 7 fold increase in EC_{50} from 107 nM to 15 nM if both the alanine at position 101 were changed to a threonine and the threonine at position 104 was changed to a valine. A 5.6 fold increase in EC_{50} from 107 nM to 19 nM was observed if the alanine at position 101 was exchanged for threonine and the serine at position 107 to valine.

This suggests that for AgrC1 ^{104}T and ^{107}S and for AgrC4 ^{104}V play an important role in the activation and that for AgrC1 ^{101}A and for AgrC4 ^{101}V and ^{107}V is important for group specificity.

3.2.4.4 Contribution of the predicted first extra-cellular loop

Since exchanging the second extra-cellular loop did not result in complete change of group specificity, an AgrC4 loop1 group1 receptor was constructed to investigate the role of the first extra-cellular loop in the group specific *agr* response. Figure 3-19 shows the dose response curve for activation of AgrC4 and AgrC4 loop1 group 1 with AIP-4.

.

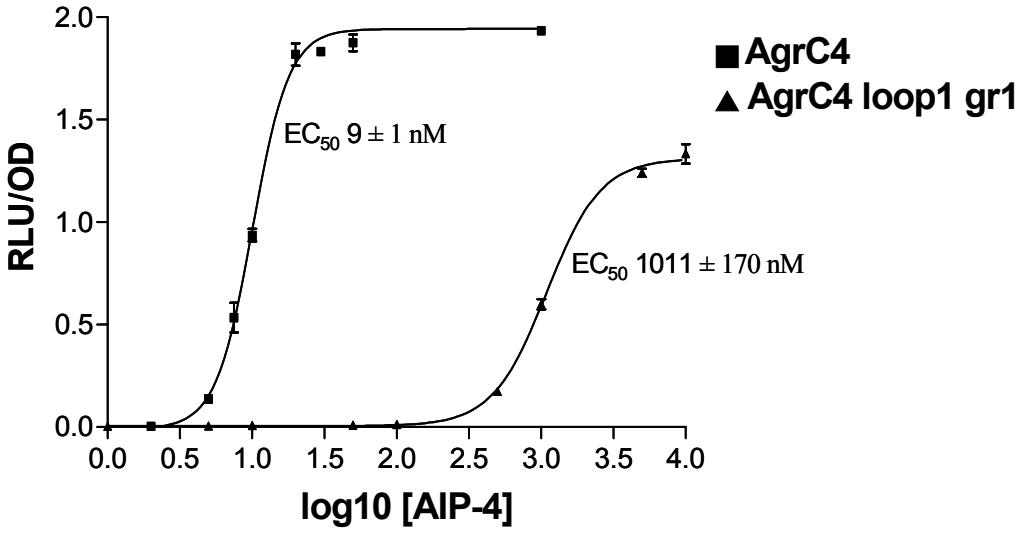


Figure 3-19: Dose response curve for the activation of ROJ48 via AgrC4 or AgrC4 loop1 gr1 activated with AIP-4 at various concentrations ranging from 0 nM to 10 µM. Error bars represent standard deviation where n=3

This clearly shows a reduction in EC_{50} from 9 nM for AgrC4 to 1011 nM for AgrC4 loop1 gr4 indicating that the first extra-cellular loop also plays an

important role in activation of AgrC by its native ligand. The AgrC4 loop1 gr1 construct was also tested with AIP-1 as an activator but showed no activation at a concentration of 10 μM (data not shown), unlike wild type AgrC4 for which AIP-1 is a weak agonist with an EC_{50} of 3.5μM. This indicates that the first extra-cellular loop is required for activation of AgrC.

The cross group inhibitor (ala^5)AIP-1, was used to investigate whether changes in the first extra-cellular loop only effected the activation of the receptor or, as with changes in the second extra-cellular loop the changes also had impacted on inhibition. For this, a competitive inhibition assay (described in section 3.2.3.2) where the AgrC4 loop1 gr1 was activated with ten fold EC_{50} of AIP-4 and (ala^5)AIP-1 was added in concentrations ranging form 0nM to 10 μM. Results are shown in Figure 3-20.

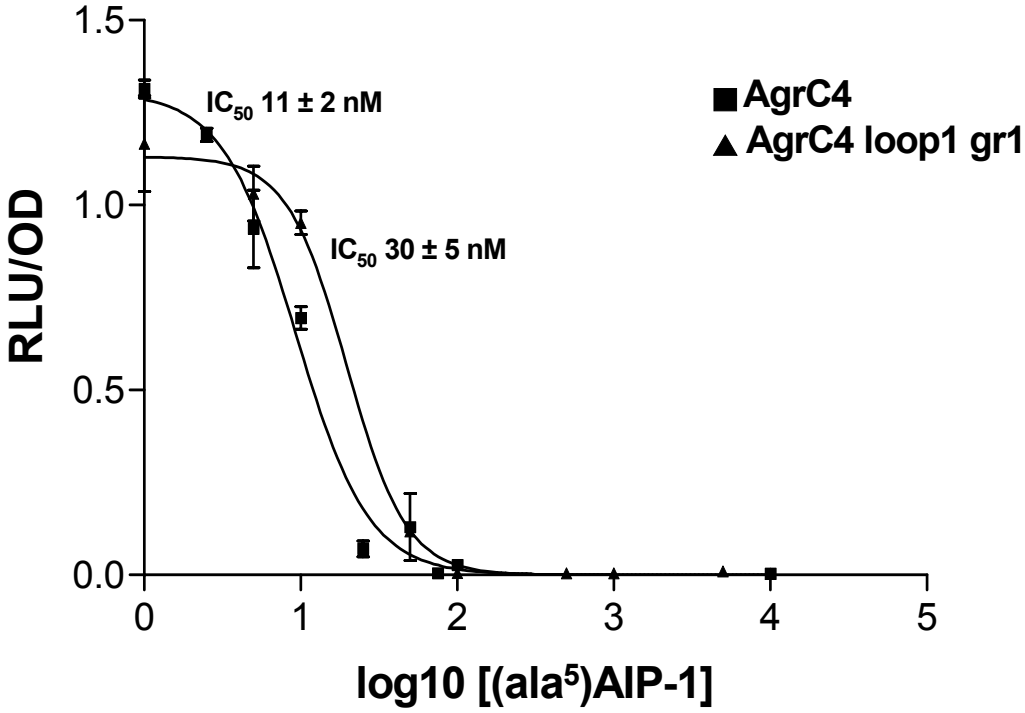


Figure 3-20: Dose response curve for the inhibition of ROJ48 via AgrC4 or AgrC4 loop1 group1 activated with fold EC_{50} AIP-4. Exogenous concentration of (ala5)AIP-1 ranged from 0 nM to 10 µM. Error bars represent standard deviation where n=3.

Figure 3-20 shows that the loop change does alter the ability of (ala^5)AIP-1 to inhibit activation, but is still a very potent antagonist indicating that the first extra-cellular loop is not important for the inhibition of AgrC receptor activation.

3.2.5 Forced evolution of new *agr* group

To investigate further the effect of the changes made in AgrC1 and AgrC4, receptor inhibition assays using the cross group *agr* inhibitor (ala^5)AIP-1 were carried out. The (ala^5)AIP-1 compound is a potent inhibitor of all four *agr* groups (McDowell *et al.*, 2001;Lyon *et al.*, 2002). Firstly, a complete loop swap of the second AgrC extra-cellular loop was tested (see Figure 3-21).

This showed that (ala^5)AIP-1 had no antagonistic effect on either AgrC4loop2gr1 or AgrC1loop2gr4. When tested as an agonist with these constructs (ala^5)AIP-1 was now found to be a potent agonist instead of a potent antagonist with an EC_{50} of 389 nM for AgrC4loop2gr1 and an EC_{50} of 46 nM for AgrC1loop2gr4 (Figure 3-21).

To try to pinpoint which amino acids were responsible for this reversal of activity, assays with (ala^5)AIP-1 used both as an agonist and an antagonist was carried out for all available AgrC mutants (Figure 3-21 to Figure 3-23).

Table 3-5: Summary of EC_{50} and IC_{50} (nM) for the activation / inhibition of the AgrC4 and AgrC1 wild type and mutated receptors.

Native and mutant AgrC4 and AgrC1 proteins	Activity in nM	
	Activation EC_{50}	Inhibition IC_{50}
AgrC4 (wild type)		11 ± 2
T101A V104T V107S (loop2 gr1)	389 ± 12	
T101A		73 ± 13
V104T		16 ± 2
V107S		4718 ± 318
T101A V104T		28 ± 5
T101A V107S		255 ± 15
V104T V107S		10 ± 3
AgrC1 (wild type)		7 ± 1
A101T T104V S107V (loop2 gr4)	46 ± 4	
A101T	120 ± 24	
T104V	556 ± 38	
S107V	4054 ± 785	
A101T T104V	26 ± 1	
A101T S107V	16 ± 1	
T104V S107V	1235 ± 303	

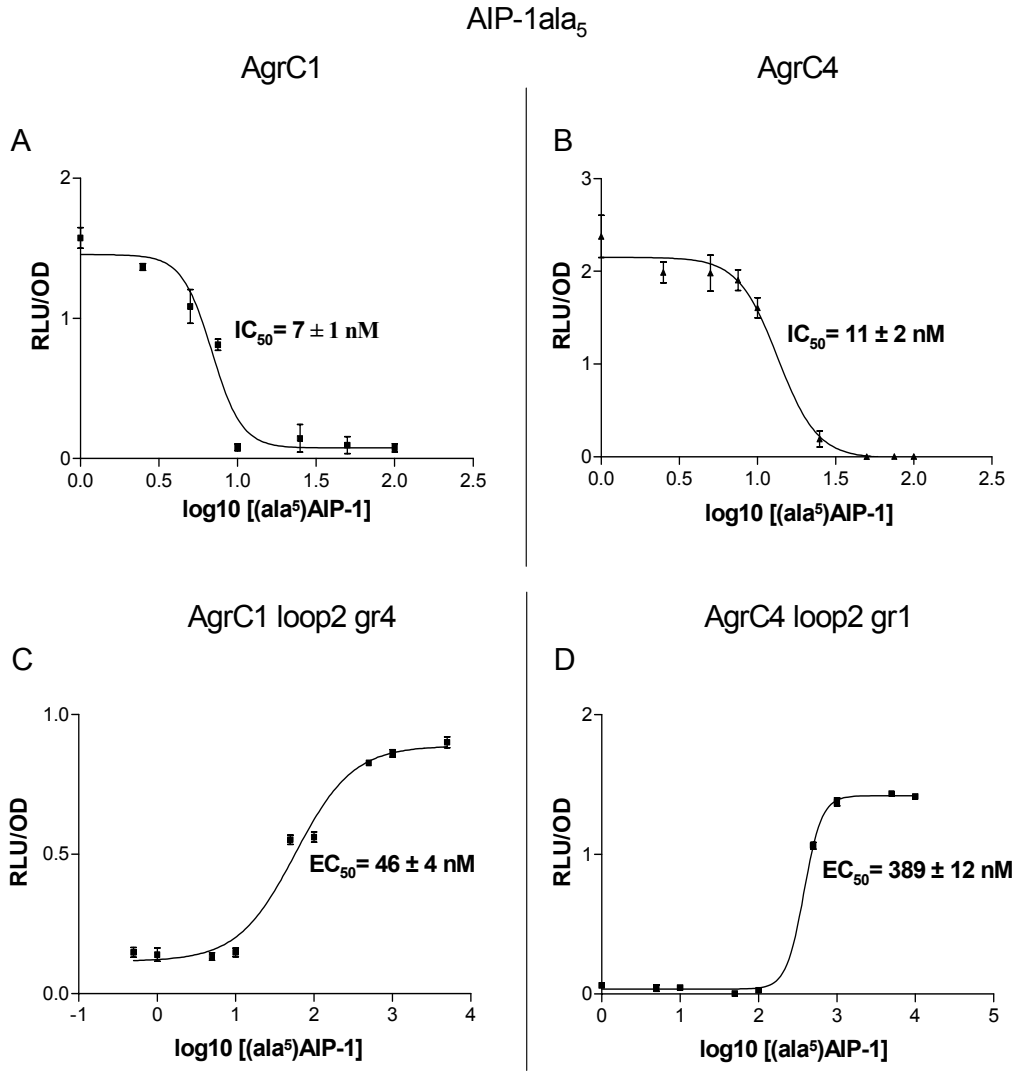


Figure 3-21: Dose response curves for the inhibition of ROJ48 via A) AgrC1 activated with 60 nM AIP-1 B) AgrC4 activated with 90 nM AIP-4. (ala^5)AIP-1 was added at concentrations ranging from 0 nM to 10 µM. C) and D) show the dose response curves for activation of ROJ48 via AgrC1loop2gr4 (C) and AgrC4loop2gr1 (D) using (ala^5)AIP-1 as an agonist. (ala^5)AIP-1 was added at concentrations ranging from 0 nM to 10 µM. Error bars represent standard deviation where n=3

Figure 3-22: Dose response curves of the inhibition of ROJ48 via AgrC4 carrying single or double amino acid substitutions in the second extra-cellular loop using (ala^{5})AIP-1 as the antagonist. A) AgrC4 T101A, B) AgrC4 V104T, C) AgrC4 V107S D) AgrC4 T101A V104T E) AgrC4 T101A V107S F) AgrC4 V104T V107S. All AgrC4 constructs were activated by ten 10X EC_{50}. Concentration of added (ala5)AIP-1 ranged from 0 nM to 10 µM. Error bars represent standard deviation where n=3

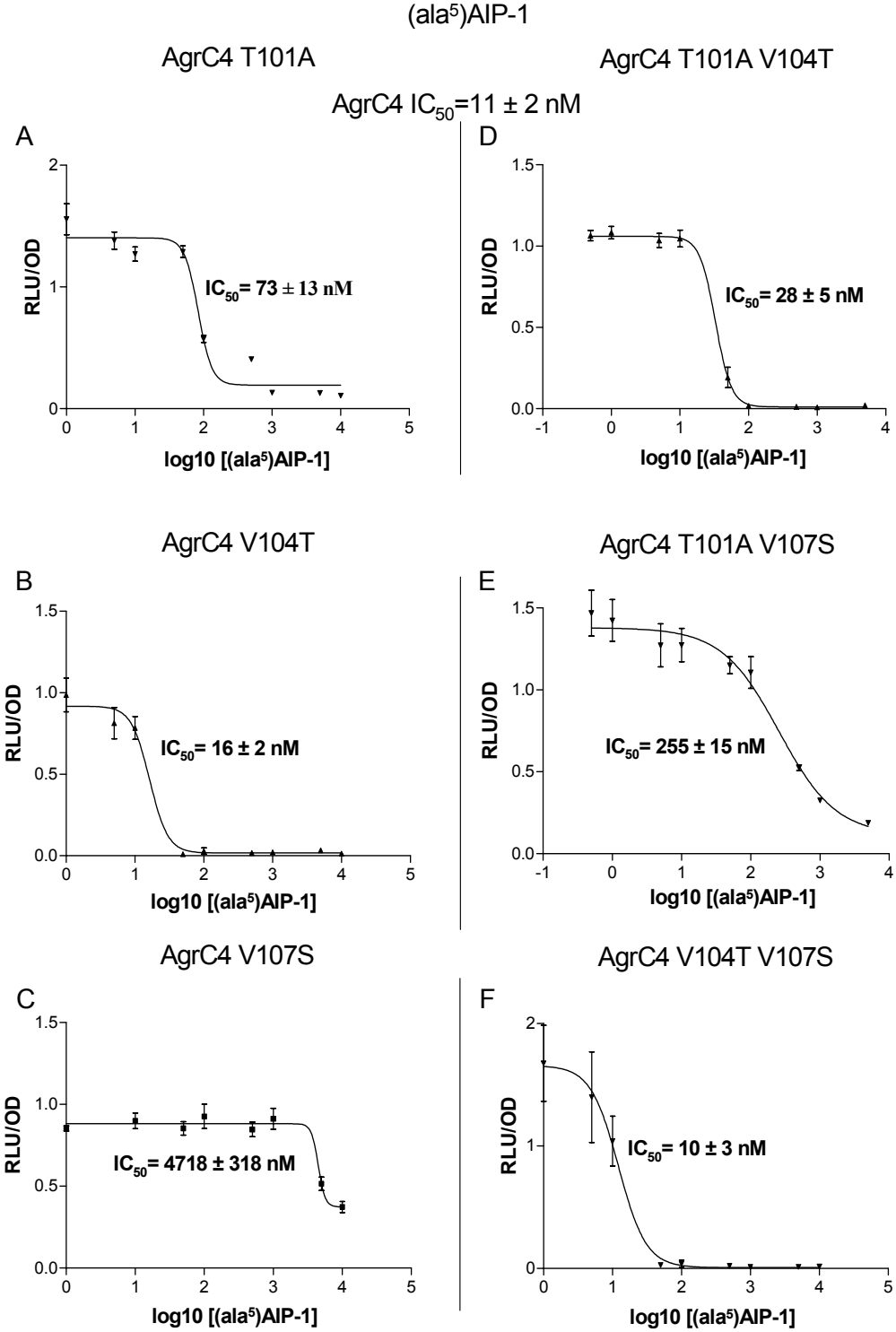
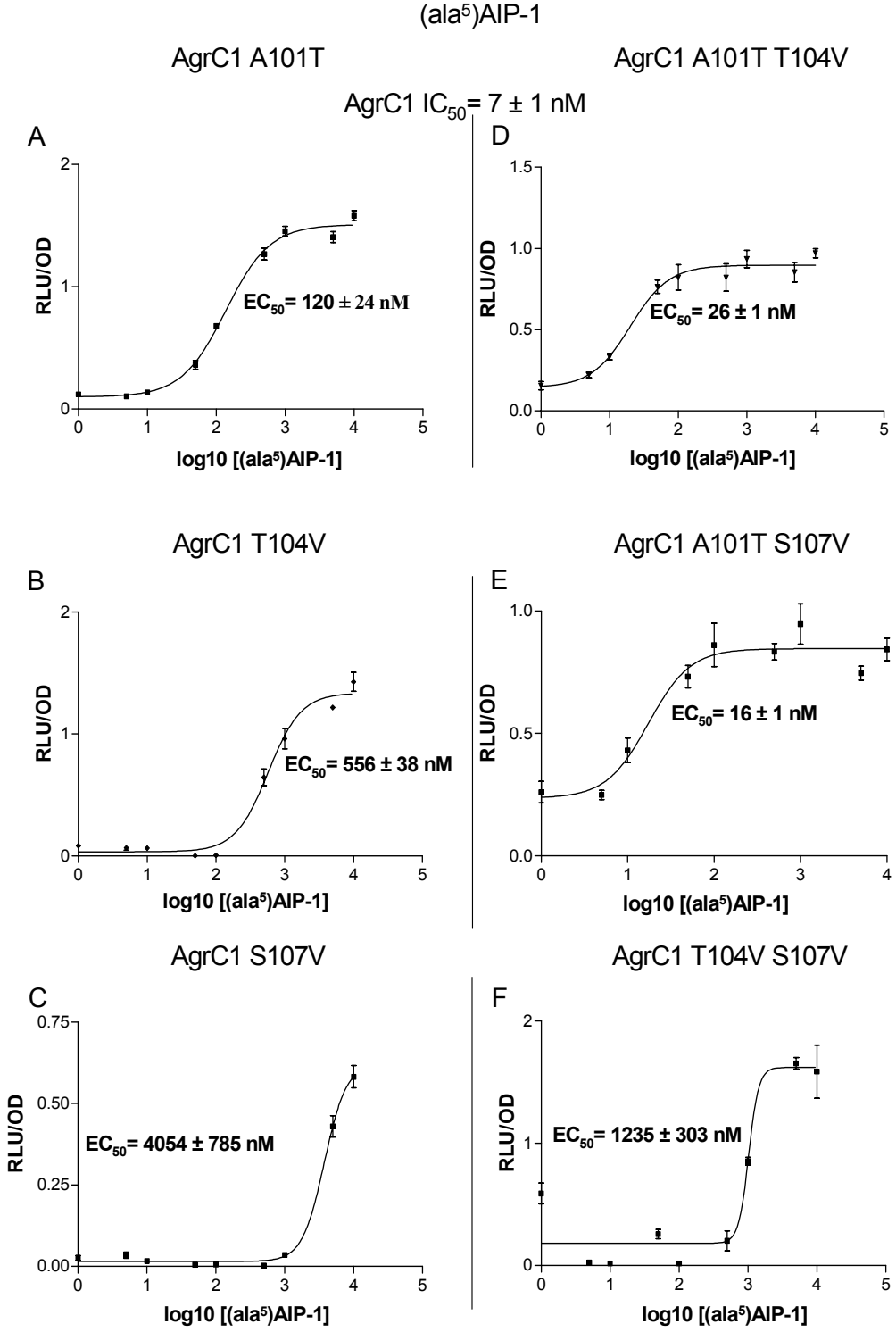

(ala5)AIP-1
AgrC4 T101A
AgrC4 T101A V104T
AgrC4 IC50=11 ± 2 nM
A
D
RLU/OD
IC50= 73 ± 13 nM
IC50= 28 ± 5 nM
log10 [(ala5)AIP-1]
AgrC4 V104T
AgrC4 T101A V107S
B
E
IC50= 16 ± 2 nM
IC50= 255 ± 15 nM
AgrC4 V107S
AgrC4 V104T V107S
C
F
IC50= 4718 ± 318 nM
IC50= 10 ± 3 nM

Figure 3-23: Dose response curves of the activation of ROJ48 via AgrC1 carrying single or double amino acid substitutions in the second extra-cellular loop using (ala^{5})AIP-1 and an agonist. A) AgrC1 A101T, C) AgrC1 T104V, E) AgrC1 S107V B) AgrC1 A101T D) AgrC1 T104V F) AgrC1 S107V. Concentration of added (ala^{5})AIP-1 is ranged from 0 nM to 10 µM. Error bars represent standard deviation where n=3


(ala5)AIP-1
AgrC1 A101T
AgrC1 A101T T104V
AgrC1 IC50= 7 ± 1 nM
A
D
RLU/OD
EC50= 120 ± 24 nM
EC50= 26 ± 1 nM
log10 [(ala5)AIP-1]
AgrC1 T104V
AgrC1 A101T S107V
B
E
EC50= 556 ± 38 nM
EC50= 16 ± 1 nM
AgrC1 S107V
AgrC1 T104V S107V
C
F
EC50= 4054 ± 785 nM
EC50= 1235 ± 303 nM

From these data it is clear that for AgrC1 it is only necessary to change one single amino acid to make a receptor that responds to (ala^5)AIP-1 as an agonist, whereas for AgrC4 a change of three amino acids is required.

Altering the serine in position 107 to a valine on its own, or with the threonine at position 104 to a valine for AgrC1, makes (ala^5)AIP-1 a very weak AgrC1 agonist with EC_{50}s of 4 μM and 1.2 μM respectively. Whereas changing either the alanine at position 101 to a threonine or the threonine at position 104 to a valine, makes a receptor where (ala^5)AIP-1 is a moderate agonist with an EC_{50} of 120 nM or 556 nM respectively. The most potent agonistic activity of (ala^5)AIP-1 was seen when alanine at position 101 was changed to a threonine together with the serine at position 107 to a valine, this makes a receptor where (ala^5)AIP-1 is a potent agonist with an EC_{50} of 16 nM. Changing both the alanine at position 101 to a threonine and the threonine at position 104 to a valine or all three amino acids (A101T T104V S107V) (Figure 3-21 and Figure 3-23) also gives a receptor where (ala^5)AIP-1 is a potent agonist with an EC_{50} of 26 nM or 46 nM respectively.

For AgrC4 only the full loop change provides a receptor where (ala^5)AIP-1 functions as an agonist. Changing the valine at position 104 to a threonine on its own and in conjunction with changing the serine at position 107 to a valine, gave little or no change in potency of antagonistic activity (ala^5)AIP-1 with IC_{50}s of 16 nM and 10 nM respectively using 90 nM AIP-4 to activate the receptor and only a slight difference in IC_{50}, from 11 nM to 28 nM, is obtained if both the threonine at

position 101 was changed to an alanine and the valine at position 104 to a threonine. If the threonine at position 101 was changed to an alanine a larger change was noted with an IC_{50} of 73 nM compared to 11 nM for AgrC4. However changing the valine at position 107 to a serine yielded a receptor where (ala^5)AIP-1 was a very weak antagonist with an IC_{50} of 4.7 μM, whereas changing both the threonine at position 101 to an alanine and the valine at position 107 to a serine gives a receptor where (ala^5)AIP-1 is an intermediate antagonist with an IC_{50} of 255 nM.

These results show that a single point mutation can make an AgrC receptor recognise a potent antagonist as an agonist, essentially creating a fifth *agr* group. For AgrC1 the changing of the alanine at position 101 to a threonine is the key to this switch in recognition of agonist. This effect is boosted by changing either the threonine at position 104 or the serine at position 107 to valine at the same time. For AgrC4 it is the valine at position 107 which is important for the antagonistic effect of (ala^5)AIP-1 and a change of three amino acids is required to make it recognise (ala^5)AIP-1 as an agonist.

3.3 Discussion

AgrC belongs to a family of QS peptide histidine protein kinases (HPK). The 3D structure of AgrC has not yet been solved but using different prediction methods several different topology models have been proposed where AgrC has either five (Lina *et al.*, 1998) or six TMH (Geisinger *et al.*, 2008).

Here eight different software tools have been used to predict the transmembrane topology of AgrC and each six transmembrane helices (24 to 18 amino acid sequence domains), three extracellular loops and a cytoplasmic histidine-kinase domain. The general consensus of these models, and the predictions and large scale comparison of different predictions reported by Möller *et al* (2008), indicates that the TMHMM 2.0 program is the most accurate prediction tool available at present. The five TMH model proposed by Lina *et al* (1998) is most likely to be incorrect as it is only based on a Kyte-Doolittle hydropathy map and incorporates a 12 amino acid residue helix (helix II) which is incapable of spanning the membrane. However, each prediction program predicts six TMH and three extracellular loops, the size of the loops differs in number of predicted amino acid residues present within each loop. Loop1 is predicted to contain from 5 to 25 amino acid residues, loop2 from 2 to 15 amino acid residues and loop3 from 9 to 20 amino acid residues.

PhoA fusion data of AgrC expressed in *E. coli* have indicated that amino acid residues 33, 105 and 176 are located externally and residue 142 is located on the cytoplasmic side of the membrane, a finding that is in agreement with the models predicted in section 3.2.2 and the HMMTOP 2.0 model proposed by Geisinger *et al* (2008). Since AIPs interact with AgrC at the cell surface and do not enter the cell, and AIP-1 and AIP-4 differ by only one amino acid, it is likely that differences in the AgrC1 and AgrC4 sequences will play an important role in the group specific AIP recognition. Using both the TMHMM 2.0 (used here) and the Geisinger *et al* (2008) prediction models for AgrC, the major changes are located in the predicted extracellular loops 1 and 2 and whereas extracellular loop3 is identical; these differences are likely to account for AIP recognition.

In the prediction by TMHMM 2.0, loop1 consists of 19 amino acid residues for both AgrC1 and AgrC4 but differs in 7 amino acids where 3 are conservative changes (^{38}F → ^{38}L, ^{42}V → ^{42}G and ^{44}S → ^{44}T) and 3 results in the exchange of hydrophobic for polar amino acids (^{35}L → ^{35}S, ^{39}F → ^{39}Y and ^{41}I → ^{41}T) and 1 results in the exchange of a polar for a hydrophobic amino acid (^{45}T → ^{45}A). Alternatively Geisinger proposes a 10 amino acid loop1 (from ^{29}G to ^{38}F) with only two amino acid differences (^{35}L → ^{35}S and ^{38}F → ^{38}L). For extracellular loop2 TMHMM 2.0 predicts 9 amino acid residues on the out side (^{101}A to ^{109}S) with 3 amino acids differing between AgrC1 and AgrC4 (^{101}A →^{101}T, ^{104}T → ^{104}V and ^{107}S → ^{107}V), whereas Geisinger predicts 10 amino acid residues on the outside (^{101}A to ^{110}I) with the same amino acids differing. One of these changes

results in the exchange of a polar amino acid to a hydrophobic amino acid (^{101}A $\rightarrow$^{101}T) and the other two results in the exchange of hydrophobic amino acids to polar amino acids (^{104}T $\rightarrow$ ^{104}V and ^{107}S $\rightarrow$ ^{107}V). This observation is likely to be significant as the Asp5 (polar) in AIP-1 is exchanged with a Tyr (hydrophobic) in AIP-4. These differences in the AgrC receptor are likely to play a pivotal role in the group specific AIP recognition of AIP-1 and AIP-4. There is, however, a discrepancy between the topology model proposed in Geisinger *et al* (2008) and the topology model from the HMMTOP 2.0 program cited to have been used for the modelling (Tusnady and Simon, 2001). The HMMTOP 2.0 predicts the N- and C-terminals to be extracellular, where as most other prediction programs predict them to be cytoplasmic. If the model is inverted such that N- and C-terminal ends are on the cytoplasmic side of the membrane the models still correlate, Geisinger *et al*'s (2008) model presents loop1 as residues 29-38, loop 2 as residues 101-110 and loop 3 as residues 169-181, whereas the HMMTOP 2.0 predict loop1 to be residue 30-39, loop2 to be residue 103-110 and loop3 to be residue 169-182. Figure 3-24 show a comparison of the discussed models.

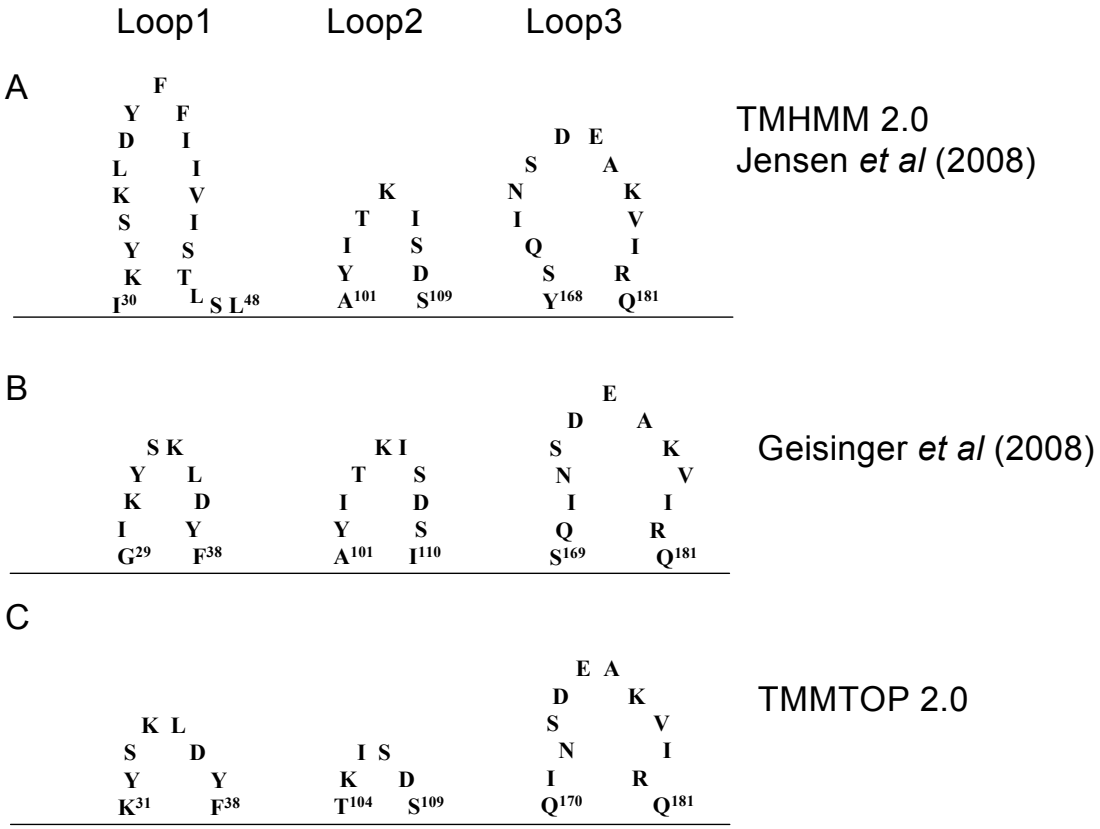


Figure 3-24: Comparison of the predicted extracellular loops. A) Prediction using TMHMM 2.0 used by Jensen *et al* 2008. B) Geisinger *et al*'s (2008) prediction using TMMTOP 2.0. C) Prediction obtained in this study using TMMTOP 2.0.

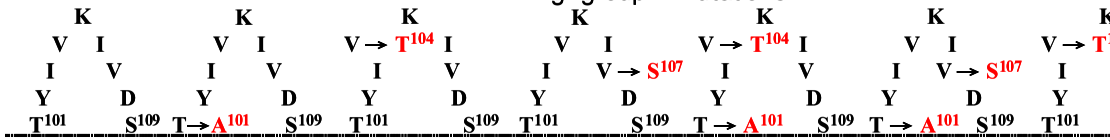
Using the TMHMM 2.0 prediction server the transmembrane topology of AgrC2 and AgrC3 from *S. aureus*, AgrC from *S. epidermidis*, AgrC from *S. heamolyticus* and the AgrC from *S. saprophyticus* were predicted to have six transmembrane helices and three extracellular loops, with the exception of *S. aureus* AgrC2 and *S. epidermidis* AgrC which were predictied to have four TMH and two extracellular loops and seven TMH and three extracellular loops with a N-terminal extracellular "tail" respectively. This difference in AgrC topology could explain why the *S.*

aureus AgrC inhibitor (ala^5)AIP-1 does not have any antagonistic effect on *S. epidermidis* AgrC activation. All structures of staphylococcal AgrCs conforms to the HPK_{10} group and are all very similar in predicted topology. This is hinting at some common way for AIP to bind to the AgrC for all staphylococci and as the macrocyclic ring structure is a common feature, it would make sense if loops 2 and 3 make up the binding pocket and loop 1 interacts with the tail.

To elucidate the contribution of the extracellular loops 1 and 2 to the group specific recognition of AIPs by AgrC, a new AIP-bioreporter wherein AgrC could readily be mutated was required. To achieve this, the *agr* locus was first deleted from *S. aureus* strain RN4220. The *luxABCDE* operon was placed on the chromosome under the control of the $P3_{agr}$ promoter. *agrC* and *agrA* genes were introduced on a plasmid under the control of the $P2_{agr}$ promoter, and the reporter activated by the addition of an exogenous AIP. Figure 3-10 and Figure 3-11 show the response of the reporter as a function of exogenous AIP concentration. Using AIP-1 to activate the reporter via AgrC1 and AgrC4 EC_{50}s of 6 ± 1 nM and 3542 ± 997 nM were obtained, respectively. Similar data were obtained using AIP-4 to activate the reporter via AgrC1 and AgrC4 giving EC_{50}s of 107 ± 10 nM and 9 ± 1 nM respectively (Table 3-3). In the presence of the potent cross group inhibitor (ala^5)AIP-1, the respective reporter fusions were strongly inhibited (Figure 3-11) with IC_{50} values of 7 ± 1nM and 11 ± 2 nM respectively (Table 3-3). These data are consistent with that previously published using the alternative $P3_{agr}$ *blaZ* reporter gene fusion (Mayville *et al.*, 1999;McDowell *et al.*, 2001).

Figure 3-25: Mutations and combination of mutations made in the extracellular loop 2 of AgrC4 and AgrC1 and the resulting receptor's EC_{50} and IC_{50} values for activation / inhibition with AIP-1 AIP-4 and (ala^5)AIP-1. All values are in nM and taken from Table 3-4 and 3-5.

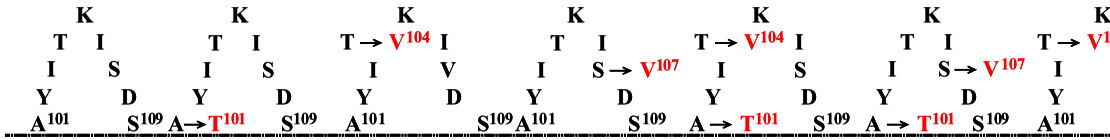
Agr group 4 mutations



AIP-1	EC_{50} = 3542	EC_{50} = 754	EC_{50} = 6357	EC_{50} = 451	EC_{50} =12064	EC_{50} = 407	EC_{50} = 4335	EC_{50} = 33
AIP-4	EC_{50} = 9	EC_{50} = 29	EC_{50} = 512	EC_{50} = 9	EC_{50} = 587	EC_{50} = 22	EC_{50} = 551	EC_{50} = 107
(ala^5)AIP-1	IC_{50} = 11	IC_{50} = 73	IC_{50} = 16	IC_{50} = 4718	IC_{50} = 28	IC_{50} = 255	IC_{50} = 10	EC_{50} = 389

nM

Agr group 1 mutations



AIP-1	EC_{50} = 6	EC_{50} = 30	EC_{50} = 235	EC_{50} = 150	EC_{50} = 47	EC_{50} = 99	EC_{50} = 120	EC_{50} = 110
AIP-4	EC_{50} = 107	EC_{50} = 22	EC_{50} = 269	EC_{50} = 408	EC_{50} = 15	EC_{50} = 19	EC_{50} = 466	EC_{50} = 94
(ala^5)AIP-1	IC_{50} = 7	EC_{50} = 46	EC_{50} = 120	EC_{50} = 556	EC_{50} = 4054	EC_{50} = 26	EC_{50} = 16	EC_{50} = 1235

nM

To elucidate the contribution of the AgrC extracellular loop2 to group specific AIP recognition of AIP-1 and AIP-4, first the three differing amino acid residues in AgrC4 loop2 were exchanged to those of AgrC1 loop2 (i.e., at position 101, 104 and 107) and introduced into the *S. aureus* bioreporter ROJ48. This resulted in a 107 fold increase in EC_{50} for AIP-1 and a ~19 fold decrease for AIP-4, see Table 3-4 A similar but less pronounced change in EC_{50}s was seen if the three differing amino acid residues inAgrC1 loop2 were changed to those of AgrC4 loop2 with a 1.1 fold increase in EC_{50} for AIP-4 and an 18 fold decrease in EC_{50} for AIP-1, see Table 3-4. These results indicate that the differential recognition of AIP-1 and AIP-4 by their cognate receptors, depends to a large extent on extracellular loop2.

To explore the relative importance of the individual amino acid residues at position 101, 104 and 107, a series of single and double mutants in both AgrC1 and AgrC4 were made. Table 3-4 shows that changing the threonine in position 101 to an alanine, the valine in position 107 to a threonine, or both of them together accounts for the increase in EC_{50} for AIP-1 where as changing the valine in position 104 to a threonine results in a decrease in EC50 for both AIP-1 and AIP-4, Table 3-4. Similar changes in the EC50s are observed when changing either one or two of the differing amino acids at a time in AgrC1 to those of AgrC4. Where changing the alanine in position 101 to a threonine gives an increase in the EC_{50} for AIP-4 but a decrease for AIP-1 both on its own and in combination with changing the threonine in position 104 to a valine or the serine

in position 107 to a valine. A decrease in EC_{50} is also seen for both AIP-1 and AIP-4 if the threonine in position 104 is changed to a valine.

The difference from using AgrC4 as a backbone is that when the serine at position 107 is changed to a valine an increase in EC_{50} for AIP-4 would be expected since the reverse change for AgrC4 gave an increase in EC50 for AIP-1, however here a decrease in the EC_{50} for both AIP-1 and AIP-4 is observed (Table 3-4). These observations suggest that the amino acid residue at position 101 is key to the differential recognition of AIP for both AgrC1 and AgrC4 and that the amino acid residue at 107 also is key for the differential recognition for AgrC4 whereas position 104 is essential for the efficient activation of both AgrCs by both AIP.

These data are consistent with the results obtained for chimeric AgrC1-4 and AgrC4-1 proteins by Wright *et al* (Wright, III *et al.*, 2004). Here the authors fused N-terminal and C-terminal domains from AgrC1 and AgrC4 at position 86-93, which are identical in AgrC1 and AgrC4, so that the AgrC4-1 chimera incorporates the extracellular loop2 from AgrC1 and AgrC1-4 incorporates the extracellular loop2 from AgrC4. Wright *et al* (2004) found that AgrC4-1 exhibits an EC_{50} of 26 nM for AIP-1 and EC_{50} of 120 nM for AIP-4 which is in good agreement with the EC_{50}s of 33 nM and 170 nM, respectively, for the AgrC4 mutant protein in which all three amino acids were changed to those of AgrC1. However, the EC_{50}s for the AgrC1-4 reported by Wright *et al* (2004) does not correlate with the EC_{50}s presented here for the mutated AgrC1 protein where all

three differing amino acid residues were changed to those of AgrC4, Wright *et al* (2004) reported an EC_{50} of 1220 nM for AIP-1 acting on the AgrC1-4 and an EC_{50} of 6 nM for AIP-4, here the comparative EC_{50}s are 110 nM for AIP-1 and 94 for AIP-4.

Geisinger *et al* (2008) have recently published a detailed study of AgrC1 and AgrC4 group specific recognition of AIP-1 and AIP-4. Some of their data corroborates the date presented here in, that the second extracellular loop of AgrC plays an important role in the group specific recognition of AIPs and, apart from residue 101, 104 and 107, they also identified residue 116 as contributing to the differential AIP recognition. However, the data differ with respect to the impact of V107S and V104T V107S single and double mutations on the activation of AgrC4 by AIP-1. There are no obvious explanations for these differences.

The contribution of the extracellular loop1 predicted by TMHMM 2.0 was investigated by exchanging the seven differing amino acid residues (position 35, 38, 39, 41, 42, 44 and 45) in AgrC4 to those of AgrC1 and introducing the mutated AgrC4 into the bioreporter ROJ48. Figure 3-19 shows that the EC50 for AIP-4 is reduced >100 fold, where as AIP-1 does not activate the AgrC4 loop1 mutant. This is probably due to the absence of the molecular features of the AgrC1 loop2. In contrast the change in IC_{50} for the (ala^5)AIP-1 when AgrC4 loop1 is exchanged for the AgrC1 loop1 is very small (AgrC4 IC50 11 nM mutated AgrC4 IC50 30 nM, Figure 3-20). These data indicate that the putative

extracellular loop1 plays a critical role in AgrC activation. Although the tripeptide exocyclic tail (the "message" region of the AIP) of AIP-1 and AIP-4 are identical in their orientation with respect to the macrocyclic part of the AIP is likely to be different, and loop1 may very well be involved in the recognition of this "address" region (Chan *et al.*, 2004).

Alternatively, the change of amino acid residues in the extracellular loop1 may result in structural differences leading to the loss of activity, or the changes could have led to conformational or topology changes in loop1 which could influence the molecular composition of the flanking transmembrane domains. The data for the changes in the extracellular loop1 together with that from the changes in the extracelluar loop2 suggest that it is the interaction between the AgrC loop2 and the cognate AIP thiolactone domain that may be the rate-limiting step.

It has been shown that replacement of aspartic acid in the fifth position of AIP-1 with alanine creates an inhibitor of all four *S. aureus agr* groups (McDowell *et al.*, 2001;Scott *et al.*, 2003;Wright, III *et al.*, 2004). In addition Wright *et al* (2004) observed that the (ala^5)AIP-1 could not activate either the AgrC1-4 or the AgrC4-1 chimera (Wright, III *et al.*, 2004). Here, by simply changing the alanine in position 101 of AgrC1 to the corresponding threonine of AgrC4 the specificity of (ala^5)AIP-1 has been changed from a potent inhibitor to an activator. In contrast changing the threonine in position 101 of the AgrC4 to the corresponding alanine of AgrC1 do not change the specificity of (ala^5)AIP-1 (see Table 3-5). Changing

the alanine in position 101 to a threonine in conjunction with changing either the threonine in position 104 or the serine in position 107 to a valine results in an increase in the potency of (ala^5)AIP-1 as an agonist. However, if all three changes are made at the same time (A101T, T104V and S107V) the activity of (ala^5)AIP-1 as an agonist is almost completely lost. Hence, to function as a potent agonist for AgrC1, (ala^5)AIP-1 requires the exchange of a non-polar amino acid in position 101 to a polar amino acid and the retention of a polar amino acid residue in either position 104 or 107, but benefits from exchanging one of them for a non-polar (hydrophobic) amino acid residue.

Structure activity studies of AIP-1 have established that the intact macrocycle (the "address" region) is essential for binding to AgrC while the N-terminal "tail" (the "message" region) is required for activation of AgrC (Chan *et al.*, 2004). NMR studies of AIP-1 and (ala^5)AIP-1 have shown that the structural difference that converts AIP-1 from an activator to an inhibitor of AgrC1 is the orientation of the exocyclic tail with respect to the marcrocycle (Chan *et al.*, 2004). Since the macrocycle alone can function as a weak inhibitor these data suggest that the activation of AgrC1 mutants could be a consequence of the receptor adopting a partially activated conformation. However since that the macrocycle on its own is only a very weak activator of AgrC1 A101T this indicated that the tripeptide tail plays an important role in receptor activation. At a genetic level it is interesting that the change in specificity of the AgrC1 receptor requires only one single base change (G to A) to modify the codon for alanine (GCA) to that of threonine

(ACA). The *agrD1* gene also only requires one (A to C) mutation to modify the codon for aspartic acid to alanine to create (ala^5)AIP-1 (GAC to GCC). Hence, for such a new *agr* system to evolve naturally only two complementary point mutations would be required to make a fully functional *agr* system, and a new *agr* group. The natural cause of *agr* divergence is, at present, not known (Dufour *et al.*, 2002), although competitive interference between different *S. aureus agr* groups has been reported (Fleming *et al.*, 2006).

4 Visualisation of the AgrC protein

4.1 Introduction

Full length AgrC has not previously been purified either in its native form from *S. aureus* or as a recombinant protein from any other organism. However, Lina *et al* (1998) raised rabbit anti-sera against a fusion of maltose binding protein (MBP) and the cytoplasmic C-terminal domain of AgrC (from ^{173}S to the stop codon). This antisera recognized a 46 kDa band on Western blots which was absent from an isogenic *agr* null mutant. Apart from these data, the only structural information for AgrC is that predicted using transmembrane topology programmes (Lina *et al.*, 1998;Geisinger *et al.*, 2008)and from PhoA fusions expressed in *E. coli.* (Lina *et al.*, 1998). Consequently, a method for overexpressing and purifying functional AgrC would greatly assist attempts to solve the structure of AgrC and aid characterisation of the specific AIP binding site(s).

In this chapter, the expression and identification of AgrC in both *E. coli* and *S.aureus* was attempted using Western blot, photo-activatable probes and green fluorescent protein (Gfp) fusions.

4.2 Results

4.2.1 Tagging AgrC for purification and identification of the AIP binding site(s)

Initially the plan was to tag AgrC with a His-tag or using the Impact system (This systems utilize engineered protein splicing elements (inteins) fused to a chitin binding domain as affinity tags. Which allows purification in a single chromatographic step) such that it could be affinity purified before and after reaction with photo-activatable AIP derivatives (Figure 4-2). After purification, the native AgrC or AgrC-AIP complex would be proteolytically cleaved and subjected to mass spectrometry such that MS fragments bearing the covalently attached AIP could be identified and their structure determined to identify the key AgrC amino acid residues involved in AIP binding.

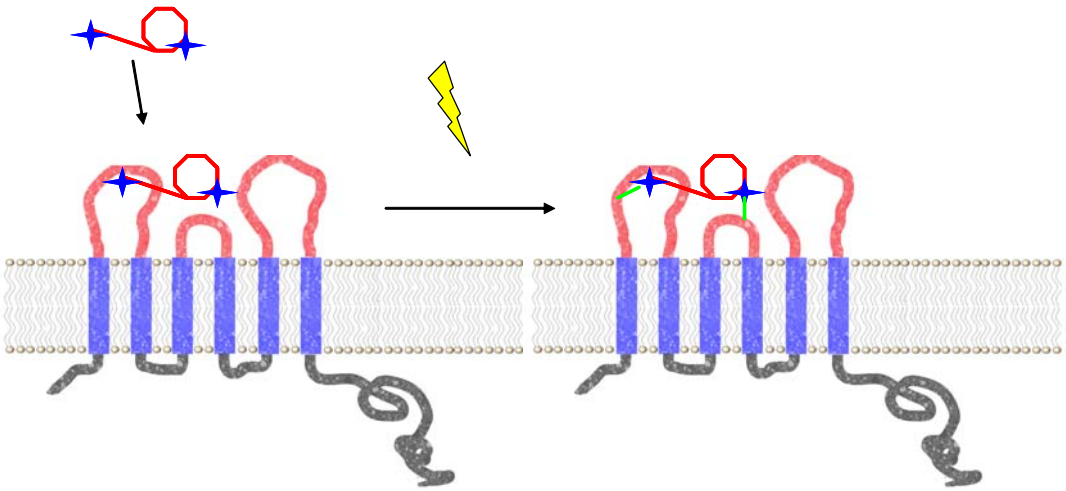


Figure 4-1: Use of photoactivatable AIP for detection of binding site in AgrC. Blue star denominates photoactivatable group either in the macrocyclic ring or in the tail. Green line is covalent bonds formed by the photoactivation.

The photoactivatable AIPs to be used were *N*-4-(4-benzoylphenoxy)butyryl(*des*-Tyr[1]) and 4-benzoyl-L-phenylalanine (synthesised by Dr Weng Chan;).

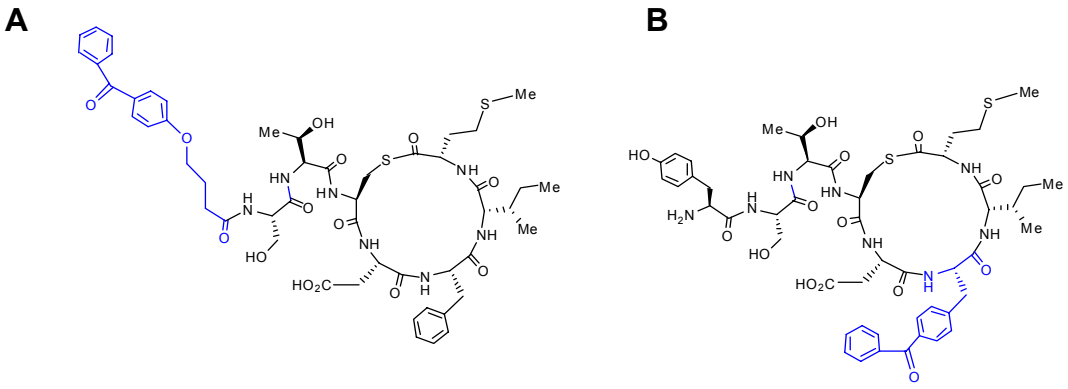


Figure 4-2: Photoactivatable AIPs A) *N*-4-(4-benzoylphenoxy)butyryl(*des*-Tyr[1])AIP-1 B) 4-benzoylPhe[6])AIP-1.

To construct tagged AgrC proteins, a 6 x His-tagged *agrC* gene (primers Table 2-3) was introduced into the *E. coli-S.aureus* shuttle expression vector, pSK5632 (Grkovic *et al.*, 2003). This proved more difficult than anticipated, and cloning the tagged *agrC* directly into pSK5632 proved impossible. The 6 x His-tagged *agrC* was, therefore, cloned into a Topo-vector for sub-cloning into an alternative expression vector. This strategy proved successful until the sub-cloning step where again no clones were obtained. An alternative method of tagging was then tried, the Impact™ (NEB, UK) (primers Table 2-3), but also proved unsuccessful. Both cloning strategies were attempted using several different *E. coli* host strains including JM109 and TOP10F' as well as C41 and C43 which are *E. coli* strains specifically developed for expressing toxic recombinant membrane proteins (Miroux and Walker, 1996). None of these host/vector combinations successfully yielded tagged AgrC protein.

4.2.2 Visualisation of AgrC using Western blot and SDS PAGE gel electrophoresis

After unsuccessfully attempting to tag and express AgrC in both *E. coli* and *S. aureus*, Western blot analysis using an antiserum raised in rats against an MBP-C-terminal AgrC fusion (MBP-AgrCser$_{173}$, constructed by Dr Alan Cockayne) was used to visualise AgrC (section 2.15) and identify the corresponding band on a Coomassie blue stained SDS-PAGE gel for subsequent use with the the photoactivatable AIPs.

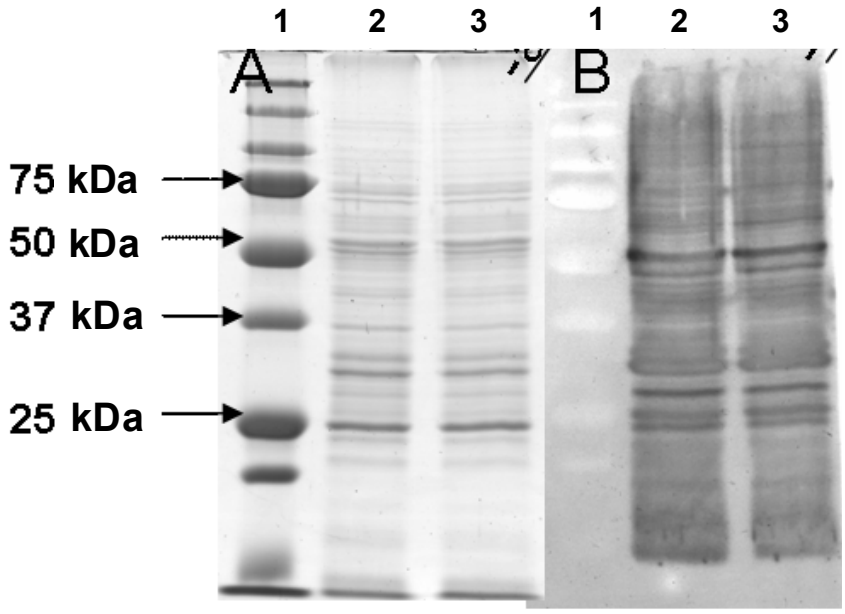


Figure 4-3: Lane 1) Biorad Precision Plus Protein Standard. Lane 2 and 3) Crude membrane extract from; 2) ROJ40 (RN4220 *Δagr*) and 3) ROJ134 (RN4220 *Δagr*) pAgrC1agrA. (A) Coomasie blue stained gel B) Western blot, film exposed for 30 s

Figure 4-3 shows that either the primary, secondary or both antibodies conjugates reacted with multiple staphylococcal proteins. To determine whether the problem was only due to one or both of the antibodies, a western blot probed only with the secondary antibody was performed (Figure 4-4) .

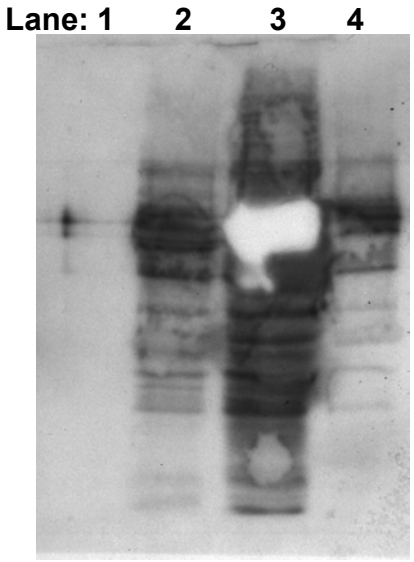


Figure 4-4: Crude membrane prep Western blot probed with just rabbit-anti-rat-HRP secondary antibody. Lane 1) Biorad low range SDS-PAGE Standard 2) ROJ40 3) RN6090B 4) ROJ40 (ala^5)AIP-1.

As the secondary antibody alone recognised many different staphylococcal proteins (Figure 4-4) the antiserum was absorbed with ROJ38 (RN6390B *Δagr*) or ROJ40 (RN4220 *Δagr*) whole cells and incubating either for 2 h at room temperature or overnight at 4°C. To clean up the crude membrane protein preparation, Elugent™ extraction was undertaken in the hope that AgrC would be selectively extracted (section 2.11).

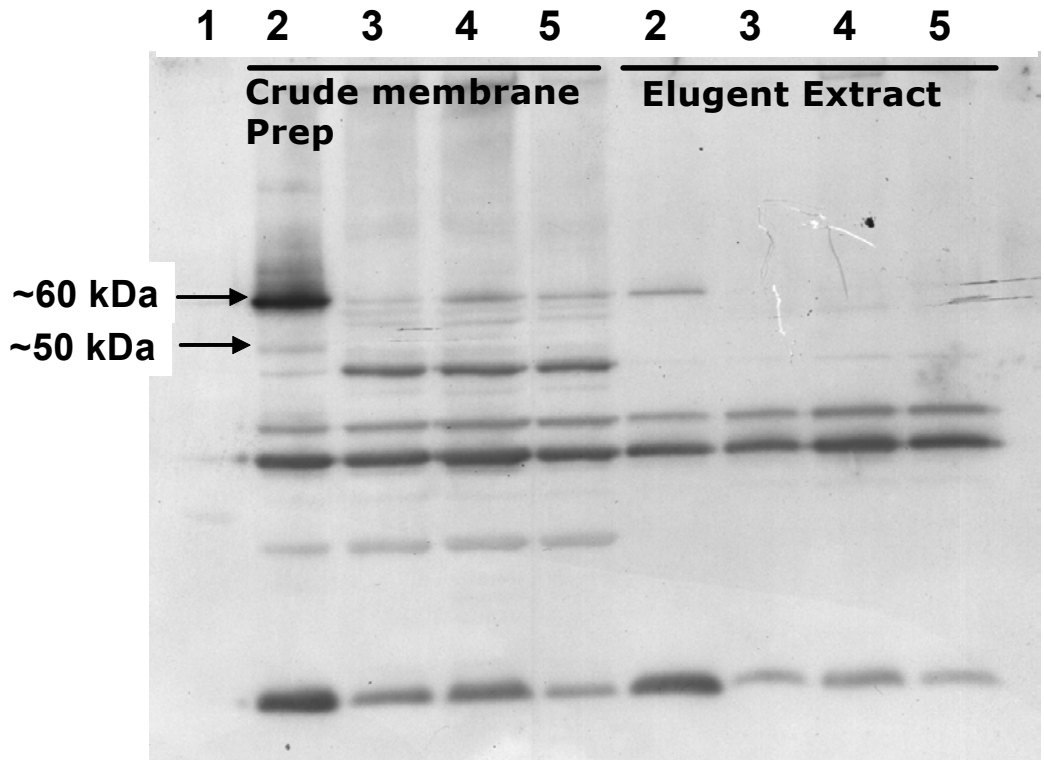


Figure 4-5: Western blot of both crude membrane prep and Elugent™ extract using absorbed antibodies. Lane 1) Biorad low range SDS-PAGE Standard 2) RN6390B 3) ROJ38 4) ROJ38 pAgrC1agrA and 5) ROJ38 pAgrC1agrA with 100nM AIP. Black arrows indicates putative AgrC protein

On the western blot shown in Figure 4-5 AIP-1 was added to strains carrying the plasmids pAgrC1AgrA to induce the transcription of *agrC* from the $P2_{agr}$ promoter. There are two bands (approximately ~60 kDa and ~50kDa) present in the lane 2; wild type *S. aureus* RN6390B that are absent in lane 3; *S. aureus* ROJ38 (RN6390B *Δagr*) which is also visible in lanes 4 and 5; ROJ38 pAgrC1AgrA with out and with AIP-1 that may be AgrC, however one protein band appears to be too large (AgrC is predicted to be 49 kDa). Both bands are also visible in lanes 2,4 and 5 but absent from lane 3 with the Elugent™ extracts.

Unfortunately these can not be identified on a Coomassie blue stained SDS-PAGE gel (Figure 4-6).

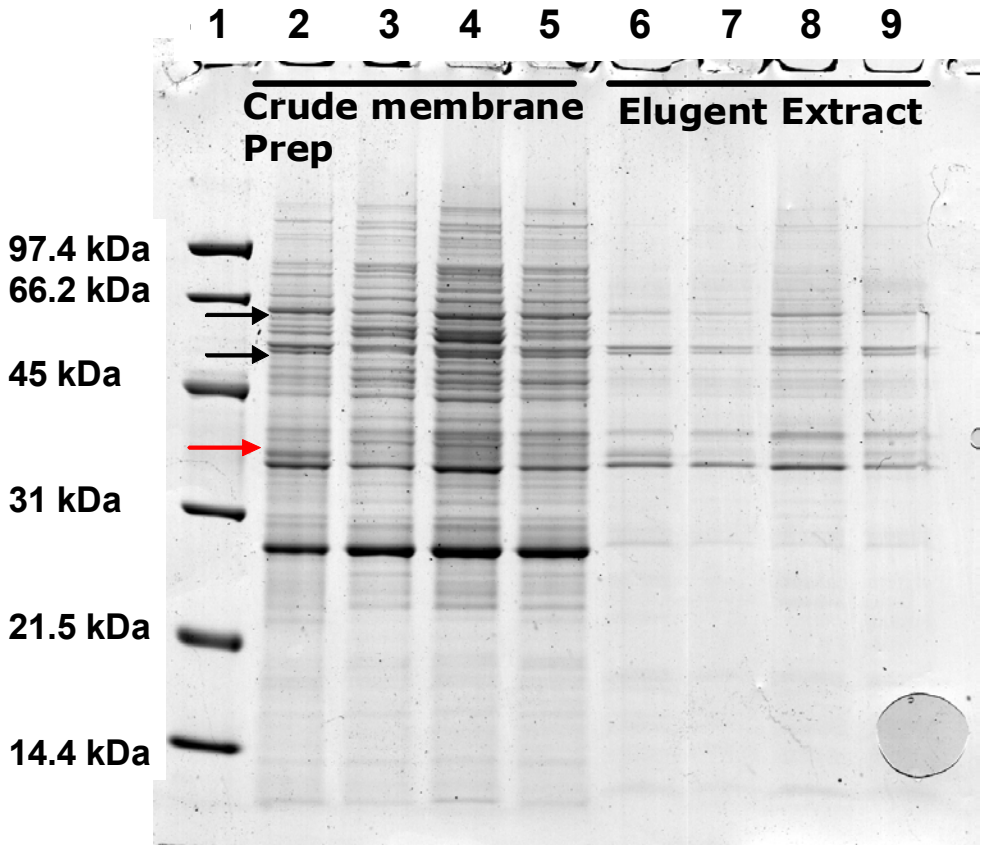


Figure 4-6: Coomasie blue stained 12% acrylamide gel of crude membrane prep (lane 2-5) and Elugent™ extract (lane 6-9). Lane 1) Biorad low range SDS-PAGE Standard, 2 & 6) RN6390B 3 & 7) ROJ38 4 & 8) ROJ38 pAgrC1agrA and 5 & 9) ROJ38 pAgrC1agrA with 100nM AIP. Black arrows indicates AgrC candidates observed in blot Figure 4-5. Red arrow indicates band submitted for MALDI-TOF protein identification.

How ever, there is a band that is visible in *S. aureus* RN6390B (lane 2; red arrow), which is absent in the lane 3 and 7 but present in lanes 2,4,5,6,8 and 9. Even

though this band appears to be too small to be AgrC (~38 kDa) it was submitted for MALDI-TOF. To ensure there was enough protein a new gel was run with three lanes of membrane preparations from RN6390B and ROJ38 pAgrC1agrA with 100nM AIP-1 and two lanes of Elugent™ extract from ROJ38 pAgrC1agrA with 100nM AIP-1. The protein identification results were: RN6390B and ROJ38 pAgrC1agrA samples gave were mixed bands making identification uncertain. The only staphylococcal protein positively identified was Pseudouridine synthase (Rsu). The ROJ38 pAgrC1agrA Elugent™ extract gave a confident match as a hypothetical protein gene code MW1525 from *S .aureus* strain MW2.

Since none of the bands sent for mass spectrometry proved to be AgrC, the photoactivatable AIP strategy was abandoned.

4.2.3 AgrC-Gfp fusions

To investigate whether AgrC is localised to a specific region in the cell, both C-terminal and N-terminal AgrC-Gfp (green fluorescent protein) fusions were constructed in the same plasmid backbone as that used for the AIP bioreporter assays (pSKermp2), where the fusion is under the control of the $P2_{agr}$ promoter (for primers Table 2-3).

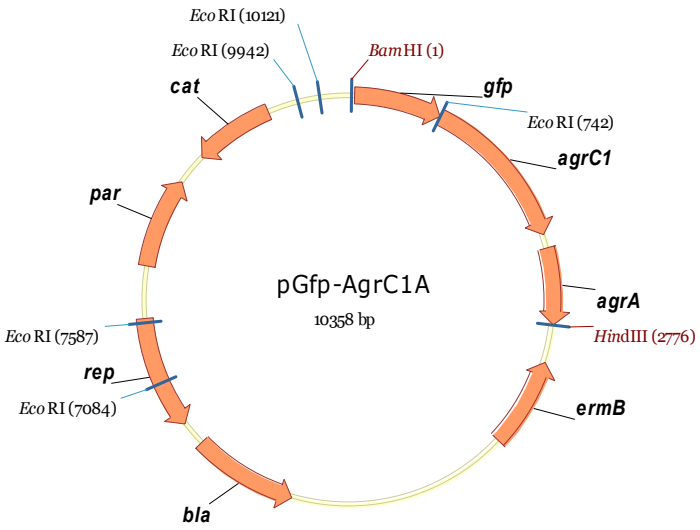


Figure 4-7: Plasmid map of pGfp-AgrC1A

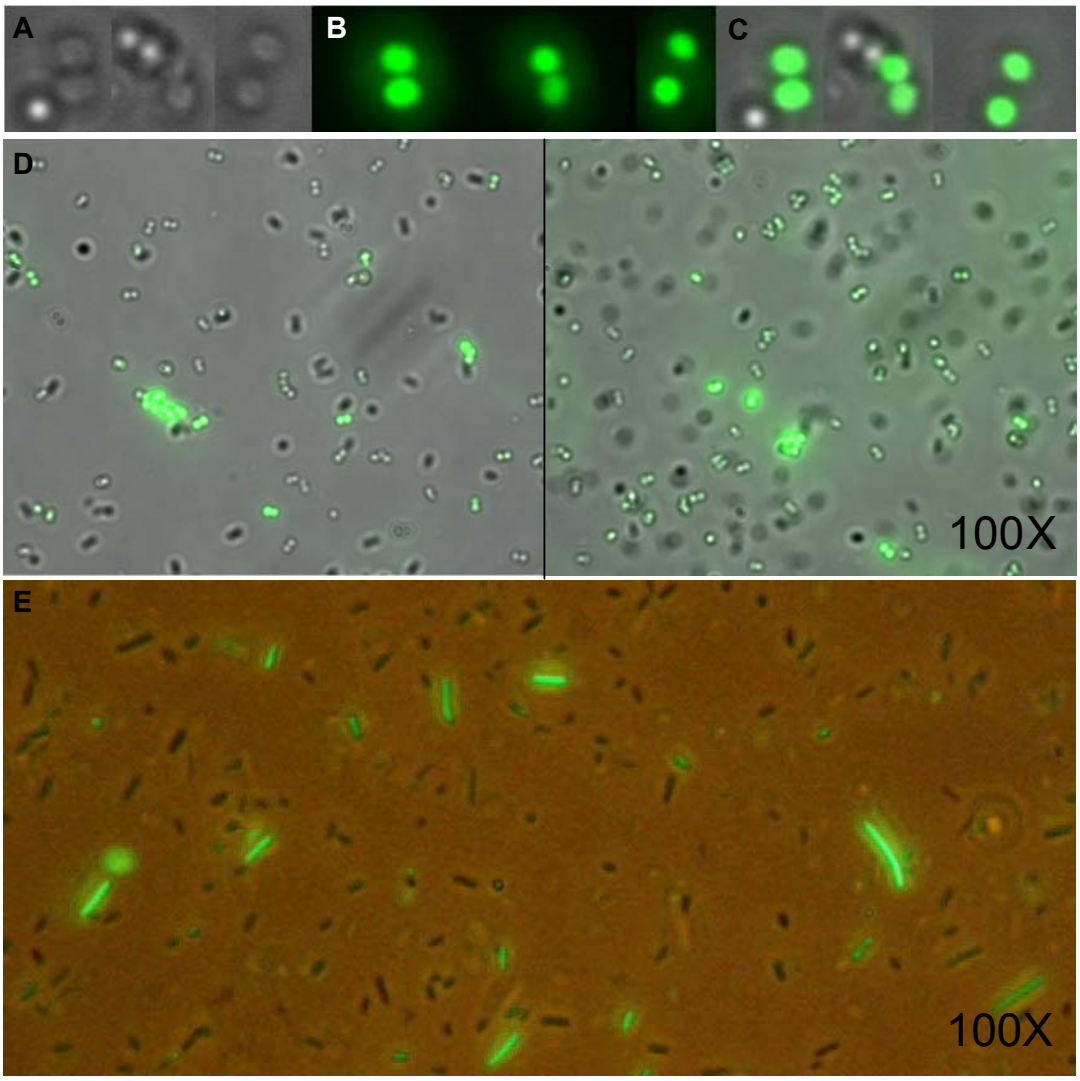
Unfortunately attempts to generate the C-terminal fusion proved unsuccessful, however, transformants for the N-terminal fusion were obtained (plasmid pGfp-AgrC1A, Figure 4-7). Restriction mapping of the plasmid proved difficult but the insert size was correct (plasmid sent for sequencing, but no results have been received yet) and since the transformants grew on all the antibiotics (the plasmids

carried resistance genes to ampicillin, chloramphenicol and erythromycin) it was decided to transform the fusions into *S. aureus* RN4220 and ROJ39 (RN4220 *Δagr*) for further study.

Microscopy of *E. coli* JM109/ pGfp-AgrC1A show that the fusion is uniformly distributed in *E. coli*. However, the *E. coli* cells showed signs of being slightly division deficient, with many cells being abnormally elongated; these were also the cells that fluoresced the strongest (Figure 4-8E).

Figure 4-8A-C shows that in *S. aureus* RN4220 AgrC is uniformly distributed throughout the cell. To test whether the AgrC-Gfp fusion was functional it was transformed into ROJ39 and grown with and with out AIP-1 (Figure 4-8D). A few differences in fluorescing cells were observed, however, the difference was small. The pGfp-Agr plasmid was also transformed into *S. aureus* ROJ48 and used in a bioluminescence assay to determine whether the fusion was functional. This, however, did not show any *agr* activity even at 10 μM AIP-1.

Figure 4-8: Light and fluorescent micrographs of N-terminal *gfp-agrC* fusions. A) cut outs of Bright field images B) is fluorescence images and C) merge of bright field and fluorescence images of RN4220 with pGfp-AgrC. D) Merge of bright field and fluorescence microscopy pictures of ROJ39 with pGfp-AgrC grown with out (left) and with AIP-1 (right) E) *E. coli* JM109 with pGfp-AgrC


A
B
C
D
100X
E
100X

4.3 Discussion

Attempts at tagging and purifying AgrC unfortunately failed, most likely due to the toxicity of AgrC for *E. coli*. Alternatives to solve this problem could be to use a different vector / host expression system such as the lactobacillus pNZ8020 expression vector (de Ruyter *et al.*, 1996) or a yeast expression system.

Lina *et al* (1998) showed western blots using antibodies raised in New Zealand with rabbits against the C-terminal domain of AgrC (Lina *et al.*, 1998), however, this did not prove reproducible with the antibodies raised in rats in our laboratory. Preliminary data generated using Gfp tagging of AgrC revealed that AgrC is uniformly distributed through out *S. aureus* and, since it is fluorescent, the Gfp protein must have folded correctly when fused to AgrC. A C-terminal fusion also needs to be made where AgrC will hopefully still be active verifying that AgrC-Gfp has inserted correctly into the membrane. Use of this fusion for confocal microscopy would verify AgrCs localisation in the membrane. The Gfp tagged AgrC could also be used for co-immunoprecipitation to further study AgrC AIP interactions using the available photoactivatable AIPs.

5 AgrB-D interactions and identification of a novel protein potientially interacting with AgrB

5.1 Introduction

AgrD is the precursor molecule for functional AIP. To yield the functional AIP three events are necessary; two endopeptidase events removing the 24 amino acid residue N-terminal amphipathic leader and the C-terminal charged tail and the formation of the thioester bond between α-carboxyl group of the C-terminal amino residue the thiol group on the central cysteine to form a macrocyclic and the functional AIP will need to be secreted.

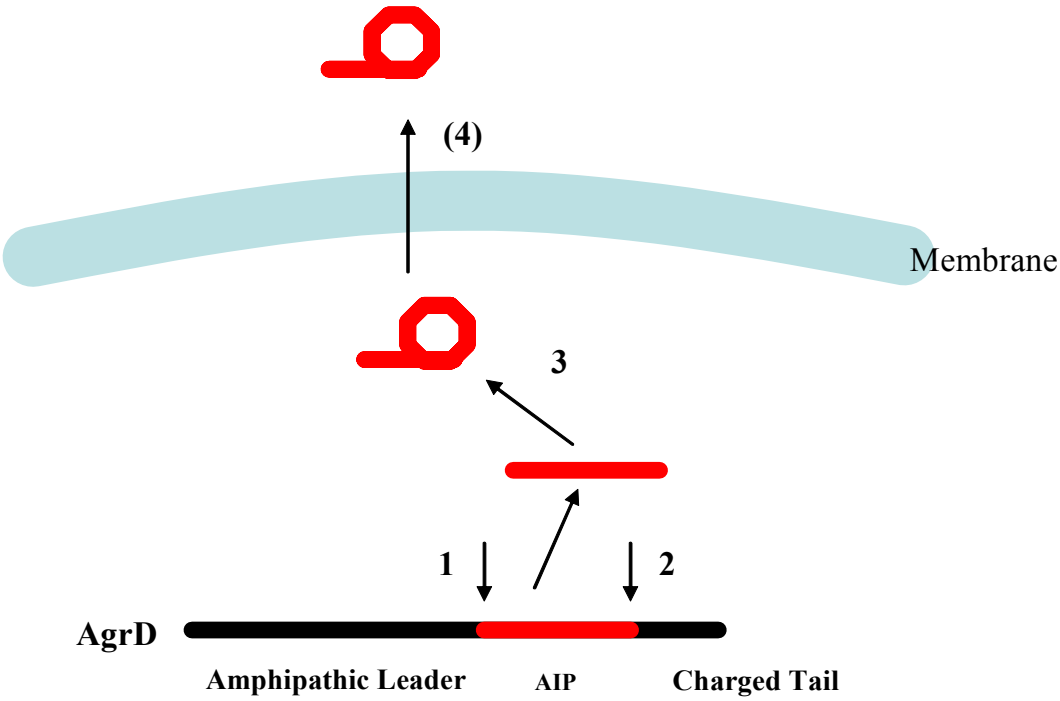


Figure 5-1: Steps involved in the production and secretion of functional AIP. 1) Removal of the N-terminal leader 2) removal of the C-terminal tail 3) formation of the macrocycle.

AgrB has been hypothesised to process the AgrD polypeptide and secrete functional auto inducing peptide (AIP) (Ji *et al.*, 1995;Ji *et al.*, 1997). It is, however, very unlikely that AgrB can perform all of these steps on its own, thus it has been hypothesised that AgrB requires the help of other proteins to fulfil this role (Garvis *et al.*, 2002;Novick, 2003), either for the cleaving of the AgrD polypeptide, forming the macrocycle to yield the functional AIP or with the secretion of the active AIP (Ji *et al.*, 1995;Mayville *et al.*, 1999). AgrB has been shown to have peptidase activity and cleave AgrD to remove the C-terminal tail

(Zhang *et al.*, 2002;Qiu *et al.*, 2005), and recently Kavanaugh *et al.* (2007) demonstrated that type 1 signal peptidase likely cleaves off the N-terminal leader.

5.1.1 Two hybrid systems

Two hybrid systems are yeast or bacterial systems used to look for protein-protein interactions (some systems can also be used for protein-DNA interactions). They work by fusing the proteins of interest inframe to each domain of a two domain protein and if the fusions interact the function of the protein is restored leading to a screenable event.

5.1.1.1 Yeast two hybrid systems

The Matchmaker™ yeast two hybrid system, chosen to show direct interaction between AgrB and AgrD and for the library screening for interaction between AgrB and other protein(s), if any, that might be involved in the processing or secretion of the AIP, is based on the GAL4 regulatory system, where the GAL4 protein binds to the UAS (upstream activating sequence) and activates the transcription of the target gene. The separation of GAL4 into two distinct domains, a DNA binding domain (BD) and an activating domain (AD), is exploited by fusing the library DNA fragments or your gene of interest (here *agrB*) in frame to either the binding domain or the activating domain. If the two proteins of interest interact, the functionality of the GAL4 protein is restored and the reporter genes are transcribed (Figure 5-2).

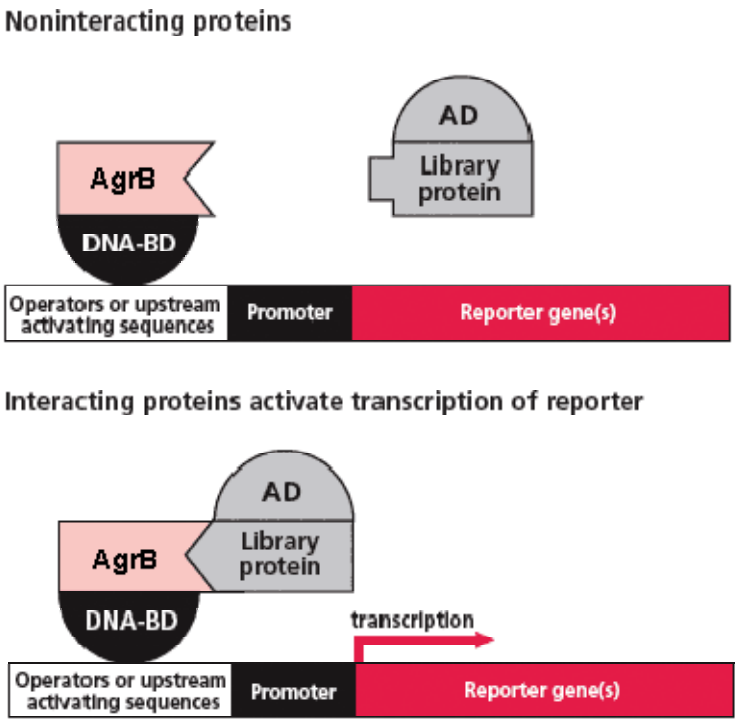


Figure 5-2: Principle of a yeast two hybrid system; the DNA binding domain (BD) fused to the "bait protein" (AgrB) binds to the UAS. And if there are interactions between AgrB and the library protein fused to the activating domain (AD) the activating domain will be positioned so that transcription of the reporter genes can be initiated from the promoter leading. This enables the yeast *Saccharomyces cerevisiae* PJ69-4A to grow on minimal plates missing nutrients otherwise required to grow on minimal medium and to the transcription of the *lacZ* gene.

5.1.1.2 Bacterial two hybrid systems

The principle of restoring the function of a protein by fusing proteins of interest to two distinct domains of the same protein and the activity of this protein is restored of the proteins of interest is interacting is also utilised in a bacterial two-hybrid system. Here, instead of restoring the activator GAL4, it is the CyaA's (adenylate cyclase) activity that is restored. CyaA is involved in a signal transduction cascade converting AMP (adenine mono-phosphate) into cAMP (cyclic AMP). An adenylate cyclase-deficient *E. coli* strain (*E. coli* lacks CaM (calmodulin) or CaM-related proteins which are required for CyaA function) is used as a reporter strain. The proteins investigated are fused to either the T25 or the T18 fragment of CyaA and if interactions occur, the two domains are brought together and CyaA function is restored leading to the production of cAMP. As cAMP is required for *E. coli* to ferment maltose or lactose, this interaction can be visualised using MacConkey plates supplemented with maltose, LB X-gal or by growth selection on minimal media plates supplemented with maltose or lactose as sole carbon source.

5.2 Results

5.2.1 Yeast two hybrid system library screening

5.2.1.1 Interactions between AgrB and AgrD

As direct interactions between AgrB and AgrD have not yet been definitively demonstrated *agrB* and *agrD* were cloned in-frame into pGBT9 and pGAD424 respectively and transformed into *S. cerevisiae* PJ69-4A. No colonies were recovered that show any potential interactions. To rule out the possibility that AgrB and / or AgrD is not functional in yeast putative AgrB /AgrD interactions were investigated using a bacterial two hybrid system (Karimova *et al.*, 1998). AgrB and AgrD was cloned into the plasmids pKT25 and pUT18 and transformed into the *E. coli* strain BTH101 (Karimova *et al.*, 1998). However, no interactions between AgrB and AgrD were found.

5.2.1.2 Yeast two hybrid system library screening

To facilitate "fishing" for novel proteins interacting with AgrB, a genomic DNA library was made by digesting gDNA from *S.aureus* strain Newman with *EcoRI* and ligating it into the plasmid pGBT9 creating pGBT9-*SA*-gDNA-Lib, a BD-library hybrid, a library of all *S. aureus* Newman protein fused to the binding domain of GAL4 (Constructed by Alan Cockayne). The bait fusion of the two

hybrid system was made by fusing *agrB* in frame into pGAD424 using *Eco*RI and *Bam*HI creating pGAD424-AgrB, an AD-AgrB hybrid.

The two plasmids were co-transformed into the *S. cerevisiae* PJ69-4A carrying the reporter genes *HIS3* under the control of the *GAL1* promoter, the reporter gene *ADE2* under the control of the *GAL2* promoter, and a *lacZ* reporter under the control of the *GAL7* promoter. In minimal media this strain requires tryptophan, leucine, methionine, uracil, adenine and histidine for growth. The leucine and tryptophan dependency is used for selection as pGBT9 carries the complementary tryptophan gene (*TRP1*) and pGAD424 carries the complementary leucine gene (*LEU2*). The presence of three reporters all responding to the same inducer gives the opportunity to select for interactions more or less stringently, see Table 5-1.

Table 5-1: Yeast medium used for selection of interactions between proteins in the yeast two hybrid system.

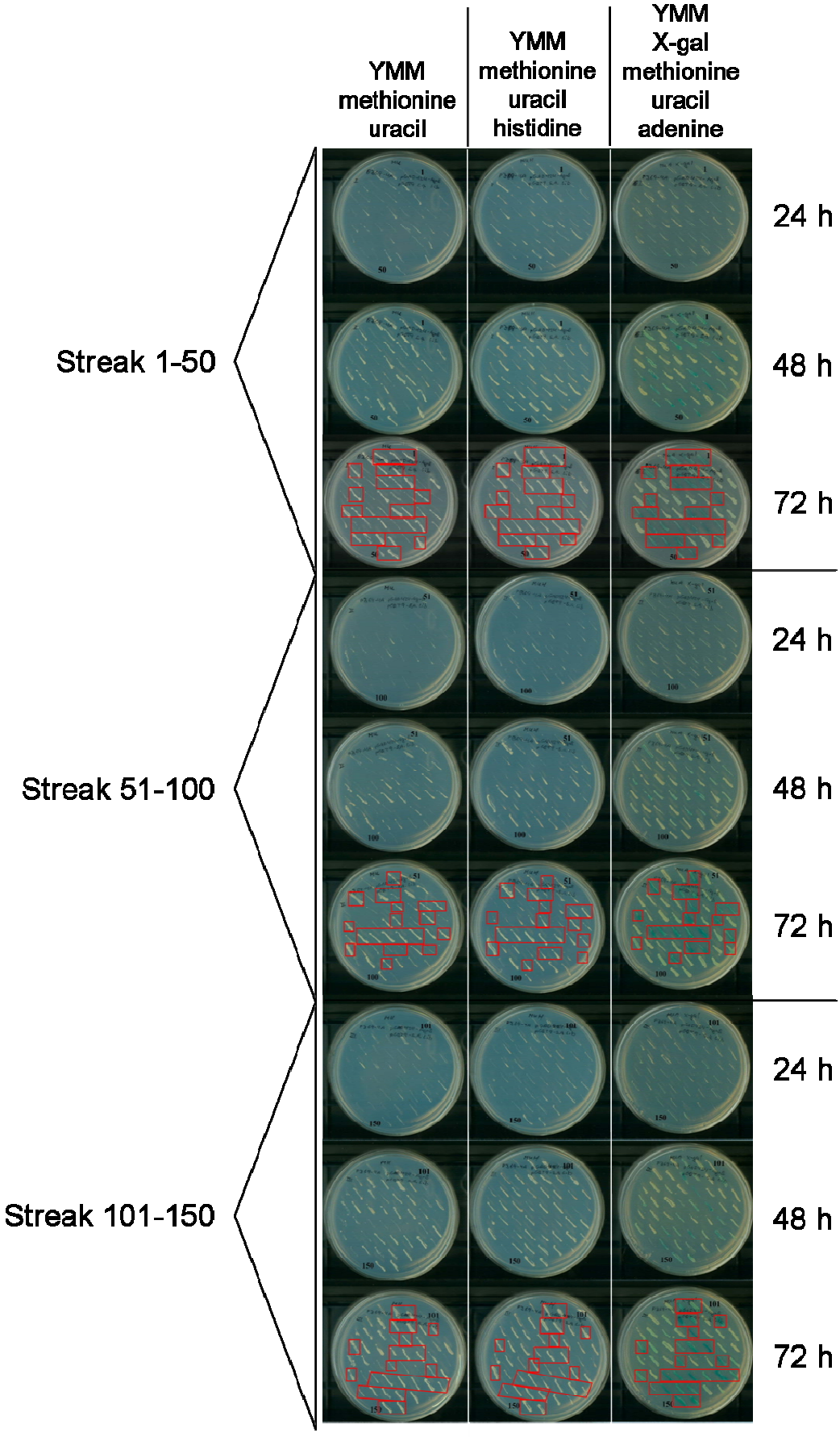
Yeast Minimal Medium with:	Stringency
Methionine, uracil and adenine	Least stringent
Methionine, uracil and histidine	
Methionine and uracil	Most stringent

To select for co-transformants the yeast was plated at 30°C for 72 h on Yeast Minimal Medium (YMM) with methionine uracil, and adenine, selecting not only for the presence of the two plasmids but also for interactions with AgrB.

150 co-transformants were streaked on to three sets of plates; YMM X-gal with methionine, uracil and adenine, YMM with methionine uracil and histidine, and YMM with methionine and uracil. The plates were incubated at 30°C and growth and colour were followed for 72 h, see Figure 5-3.

After 72 h 70 co-transformants showed potential interactions between AgrB-AD and the gDNA-BD fusions by growth on one of the two growth selective mediums and by turning blue on the X-gal plates (Figure 5-3). These were re-streaked on to YMM methionine, uracil, histidine, adenine and leucine and incubated at 30ºC for 48 h. The purpose of this was to ensure each strain only contains one copy of pGBT9-*SA*-gDNA-Lib plasmid.

Figure 5-3: Yeast PJ69-4A containing plasmids pGBT9-*SA*-gDNA-Lib and pGAD424-AgrB plated on three different selective plates selecting for protein-protein interactions with AgrB. YMM methionine and uracil (most stringent) YMM methionine uracil and histidine and YMM X-gal methionine uracil and adenine (Least stringent for growth, blue if interactions are present). Photos were taken every 24 h for 72 h to document growth and colour change. After 72 h streaks that were blue on the X-gal medium and grew on one of the two growth selective medias were considered potential candidates for interactions between AgrB-AD and the protein expressed by the library plasmid. Potential candidates are shown in red boxes.


YMM
methionine
uracil
YMM
methionine
uracil
histidine
YMM
X-gal
methionine
uracil
adenine
Streak 1-50
Streak 51-100
Streak 101-150
24 h
48 h
72 h
24 h
48 h
72 h
24 h
48 h
72 h

Plasmids from each of the 70 strains were isolated and the library insert was amplified by PCR, using primers pGBT9f and pGBT9r (Table 5-2). The resulting PCR products were purified and digested with the frequent cutting restriction enzyme *HaeIII*. The digests and their parent PCR products were then run on a 1% analytical agarose gels. PCR products of the same size and with the same digestion patters were presumed to be identical. Pairs of identical PCR products were sent off for sequencing (Figure 5-4).

Sequences (appendix B) were blasted against the *S. aureus* genome to identify which gene the library insert originated from (see Table 5-2.)

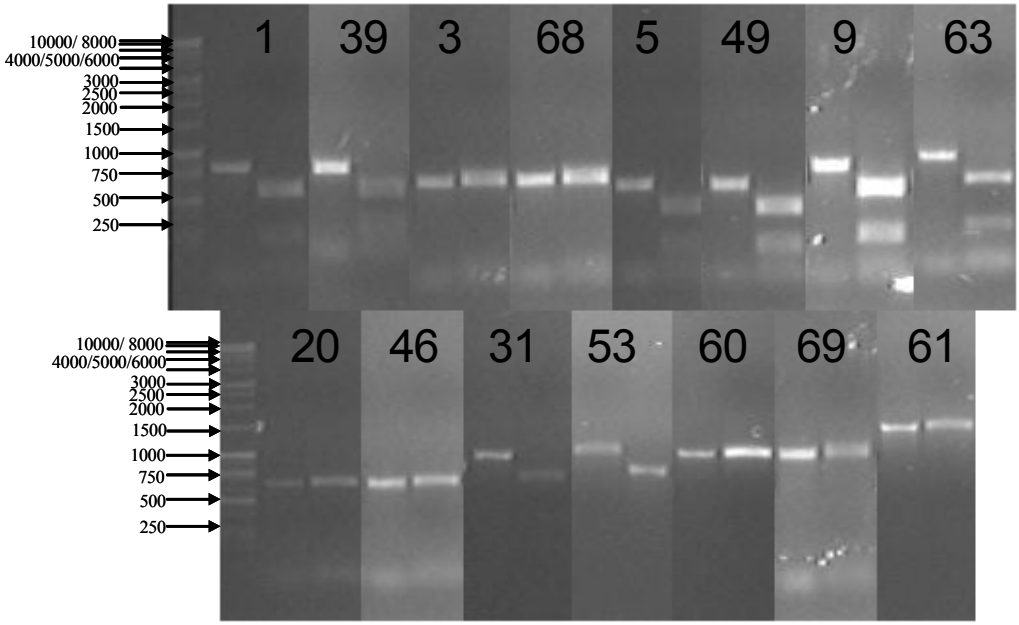


Figure 5-4: PCR products sent for sequencing on the left hand side of each double lane and the *HaeIII* digest on the right hand side of each double lane of PCR products. Numbers above the lanes indicate the number given to the potential interaction during the YTH screening.

Table 5-2: List of genes of interest that showed up when sequences were blasted against *Staphylococcus aureus'* genome using blastn (http://www.ncbi.nlm.nih.gov/BLAST/Blast.cgi). Only results with the highest scores and e-values are shown. Score: The higher the score the better a match the blasted sequence matches the sequence in the database. E-value: Is the expectation value and is number of different alignents with scores equivalent to or better than the score that are expected to occur in a database search by chance. Hence, the lower the E value, the more significant the score

Sequence no	Blast result (Features in subject sequence)	Score, e-value
1 & 39	RGD-containing lipoprotein	954 , 0.0
3 & 68	imidazolonepropionase	567 , 6e-159
5 & 49	glucose-6-phosphate 1-dehydrogenase	650 , 0.0
9 & 63	SA0619 (hypothetical protein, similar to low-affinity inorganic phosphate transporter)	710 , 0.0
20 & 46	SA0958 (myo-inositol-1(or 4)-monophosphatase homolog)	464 , 5e-128
31 & 53	glucose-6-phosphate 1-dehydrogenase	975 , 0.0
60 & 69	glucose-6-phosphate 1-dehydrogenase	975 , 0.0
61	SA0211 (conserved hypothetical protein)	2044 , 0.0

Due to the way the genomic library was constructed, the above candidates potentially interacting with AgrB were not necessarily transcribed in frame, i.e., the fusion protein might not be transcribed in the same reading frame as the original gene giving rise to a nonsense protein. Each sequence was compared at the fusion site to its original gene, using the sequence analysis software VectorNTI (Invitrogen, USA) to determine which candidates was transcribed in the correct reading frame. The result of this analysis is shown in Table 5-3.

Table 5-3: Potential interactions transcribed in frame with the fusion

Sequence no.	S.A. gene no	S.A. geneproduct
1 & 39	SA0201	RGD-containing lipoprotein
20 & 46	SA0958	SA0958 (myo-inositol-1(or 4)-monophosphatase homolog)
3 & 68	SA2121	Imidazolonepropionase

5.2.2 Verification of candidate genes involvement in *agr* function

The proteins listed in Table 5-3 were only potential candidates for interactions with AgrB and these interactions require verification using alternative assays. Ideally this would be done by purifying the proteins and using a method such as surface plasmon resonance (SPR). However, since AgrB is a large membrane

protein which is difficult to express and purify, it was decided to make in frame knock out mutants of the three candidate genes and test these mutants for production of AIP, α-haemolysin and protein A production. This would tell us whether or not the protein in question contributed AIP production / secretion and *agr* function. This is work in progress.

5.3 Discussion

Identifying direct interactions between AgrB and AgrD, for this there could be several reasons for the lack of interactions; (1.) they do not interact, (2.) the binding domain and the activating domain from GAL4 are too bulky to allow for the small AgrD to gain access to AgrB, (3.) AgrB processes AgrD too quickly for there to be any detectable interactions or (4.) the proteins are not functional in yeast / *E. coli*.

Three different proteins that potentially interact with AgrB were identified using a yeast two-hybrid system. Two of the three SA0958 (gene no from *S. aureus* strain N315) and *hutI* are unlikely to have a direct function in quorum sensing let alone AIP production. SA0958 is a myo-inositol-1(or 4)-monophosphatase homolog which is involved in the inositol phosphate metabolism and belongs to a family of bivalent metal-ion dependant phosphatases. HutI is an imidazolonepropionase and is a part of the histidine utilisation operon. *Rlp*, the third protein found, is a RGD-containing lipoprotein predicted to belong to the SBP_bac_5 super family.

This is a family of ABC-type di-peptide periplasmic transporter components. As Rlp is a hypothetical protein of a predicted function, it is plausible that it could transport polypeptides instead of just dipeptides. Hence, a hypothesis for AIP synthesis could be AgrB cleaves the charged C-terminal and makes the cyclic thiolactone then transports the intermediate product size out, in conjunction with

Rlp and once outside the cell SpsB (*S. aureus* signal peptidase 1) cleaves the N-terminal yielding a functional AIP (figure 5-5).

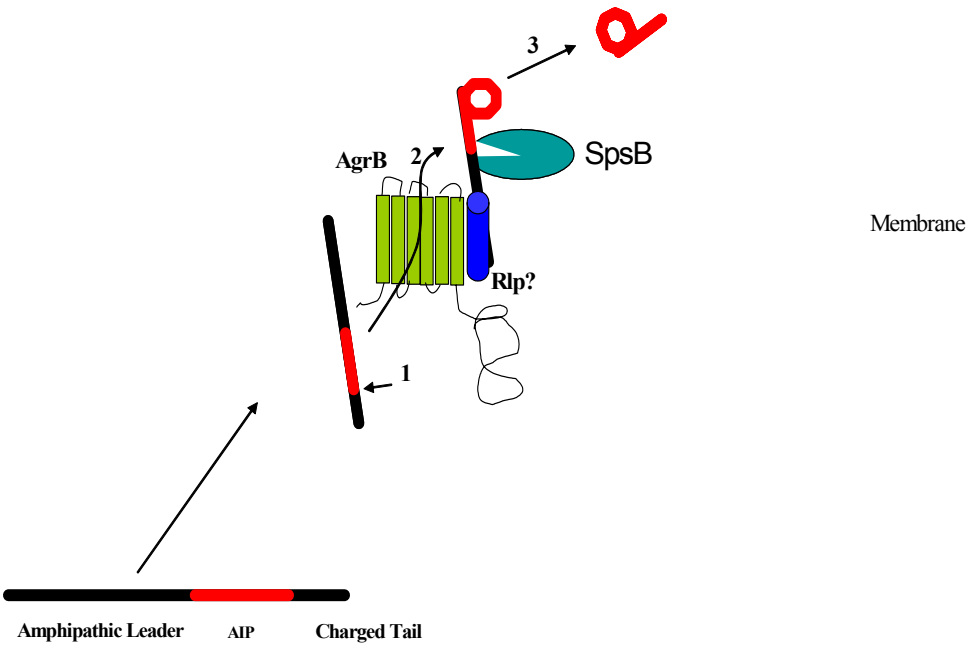


Figure 5-5: Hypothesis on the production and secretion of AIP. 1) C-terminal tail is cleaved of by AgrB 2) AgrB forms the macrocycle and exports the intermediate to the cell surface in conjunction with Rlp 3) SpsB cleaves off the active AIP.

There is one flaw in this hypothesis and that is that SpsB have only been shown to cleave AgrD1 and the cleavage site does not appear to be present on other AgrD

sequences. This, however, could a group specific event and SpsB (or other peptidases) in other *agr* groups have a different recognition sequence.

6 General discussion

S. aureus is an opportunistic pathogen that usually does not present a danger to healthy individuals. Over the past four decades, antibiotic resistant strains have steadily become more and more of a problem in health care settings and within the last two decades aggressive, highly virulent strains have also emerged in community settings, infecting young healthy individuals. As it is highly adaptable and capable of horizontal transfer of antibiotic resistance genes, new ways of combating the organism are urgently needed. One proposed way is through virulence inhibition. *S. aureus* infection / virulence can be divided into two steps: 1) Establishing the infection through invasion and adherence to the host tissue utilizing numerous cell wall associated adhesion factors (MSCRAMMs), and 2) Spreading and exo-toxin production. The switch between these two phases is mainly controlled by the *agr* locus and it has been shown that the *agr* deficient mutants have a less virulent phenotype than *agr* wild type (Cheung *et al.*, 1994;Bunce *et al.*, 1992). It has also been shown that blocking the *agr* response by co-administrating AIP-2 with *S. aureus* group I attenuates the infection (Mayville *et al.*, 1999). These properties of *S. aureus* infections make the *agr* response a good potential target for anti-virulence drugs.

S. aureus is divided into four *agr* groups with each group producing a unique AIP which is recognised by the sensor domain of the AgrC histidine kinase and can cross inhibit the other groups AgrCs. Therefore, in order to make an *agr* inhibitor

that would be effective in a possible clinical setting, it would need to inhibit all four groups. One such inhibitor, (ala^5)AIP-1 was made by McDowell *et al* (McDowell *et al.*, 2001) and *in vitro,* even at low concentrations, strains from all four *agr* groups show an *agr* negative phenotype. However, very little is known about the detailed mechanism involved in the group specific *agr* response, and a greater understanding of this could help to make better virulence inhibitors.

In this study I have tried to elucidate first, the potential transmembrane topology of AgrC in order to get a better idea of which amino acid residues are available to interactions with the AIP, and secondly to elucidate the role of individual amino acids in the group specific *agr* recognition of AIP by the AgrC receptor. For this, *agr* groups I and IV were chosen as they are the most closesly related of the four groups, differing by only one residue in the AIP. There are a multitude of topology prediction tools available online for using different methods and algorithms with varying degrees of success, but the general trend was a topology for AgrC consisting of six TMH and three extracellular loops with both cytoplamic N- and C-terminals. This is in contrast to previous published prediction for AgrC by Lina *et al* where AgrC1 was predicted to have 5 TMH and a extracellular N-terminal and cytoplamic C-terminal (Lina *et al.*, 1998), but is in agreement with a more recently published topology by Geisinger *et al* (Geisinger *et al.*, 2008). It was also shown that most staphylococcal AgrCs seem to have the same general topology, with only *S. epidermidis* having seven TMH and an extracellular N-terminal, but still three extracellular loops.

Having a good idea of the transmembrane topology of AgrC and knowledge of which amino acid residues are available for interactions with the AIP on the extracellular surface of *S. aureus,* the next step was to try to elucidate the individual residues role in the group specific response. For this an AIP bioreporter based on the *luxABCDE* bioluminescence cassette was built in an *agr* negative version of the transformable *S. aureus* strain RN4220. Into this reporter strain various versions of *agrC1* and *agrC4,* both wild type and mutated, were introduced on a plasmid and expressed from the P2*agr* promoter in conjunction with *agrA*. The effectiveness of AIP-1 and AIP-4 in activating these different AgrC receptors was then measured in a bioluminescence assay. The data collected from these assays indicate that the second extracellular loop is responsible for the group specific recognition of AIP-1 and AIP-4 by AgrC1 and AgrC4, more specifically the residue at position 101 and 107 is important for the differential recognition of AIP-1 and AIP-4 for AgrC4 and AgrC1, however the role of residue 107 for AgrC1 is not as significant as for AgrC4, hence residue 101 may play a larger role here. I have also shown that the first extracellular loop plays an important role in activation of AgrC but less so in the inhibition of AgrC. The finding that the second extracellular loop is responsible for the group specific AIP recognition are in agreement with Wright *et al*'s findings using chimeric AgrC receptors (Wright, III *et al.*, 2004) and most of the findings are in agreement with the recently published data from Geisinger *et al* that also found that residues 100 and 116 play a role in the group specific recognition of AIP-1 and AIP-4.

To study the effects of mutations in the AgrC receptor on the inhibition by possible virulence-inhibiting drugs, the *S. aureus agr* inhibiting (ala^5)AIP-1 was tested against the available AgrC mutants. The data from this part of the study surprisingly showed that (ala^5)AIP-1 is a moderate activator of the AgrC4 loop2 gr1 mutant and that it is a weak to potent activator of all the AgrC1 mutants, showing most potent activation of the AgrC1A101TS107S (the two residues suggested to be responsible of the group specific AIP recognition of AgrC1 and 4). By changing just the residue at position 101, (the residue that seems to play the largest role in the group specific recognition of AIP-1), made the AgrC1 receptor recognise (ala^5)AIP-1 as a fairly potent agonist. This has never before been observed. Then, Wright *et al* observed that (ala^5)AIP-1 activated their AgrC1-3 and AgrC4-3 chimera but the AgrC1-4 or AgrC4-1 chimera was not activated by (ala^5)AIP-1 (Wright, III *et al.*, 2004).

This part of the study has given a more detailed insight to how the activation of AgrC occurs and can hopefully help in rational design of anti-virulence drugs down the line. Also, it has demonstrated that the theory that all *S. aureus agr* groups originate from the same ancestor is plausible as only two point mutations would be necessary to make a new group who's AIP activates its own AgrC and cross inhibits other groups AgrC. This discovery also raises the question that resistance to potential AIP analogue anti-virulence drugs would also render them ineffective. This is, however, unlikely: 1) if AgrC is mutated so that it recognises

the drug as an agonist instead of an antagonist the *agr* system would be continuously switched on and all the exo-toxins would be continuously expressed, which is very energy expensive for *S. aureus*. 2) The adhesins and immune evasions normally associated with the early part of an *S. aureus* infection would not be expressed which would make *S. aureus* easily detected by the immune system and *S. aureus* would be less able to establish itself in the host. It would not, therefore, be advantageous for the bacteria to develop resistance in this way. Other resistance mechanisms are of course possible but since anti-virulence drugs do not directly select for resistance this must be presumed to be fairly unlikely. There is, however, a potential down side to inhibiting the *agr* system in *Staphylococcus*. The inhibition of *agr* can lead to biofilm formation or thicker biofilms which is more persistent to both the immune response and antibiotics. However, as *agr* have only been shown to be expressed for a matter of hours in the beginning of an infection and that several groups have shown that *agr* is required for the establishment of a fully robust infection (Booth *et al.*, 1995;Abdelnour *et al.*, 1993;Gillaspy *et al.*, 1995), the role of anti-*agr* drugs seems limited to the first hours of the infection. At first this does not seem to have much application, but with the emergence of highly virulent community associated MRSA and MSSA strains that can lead to the death of the infected person within 48 h, preventing the expression of toxins and thereby buying the patient and doctors a bit more time to treat the infection with conventional antibiotics could prove a very benficial. On the other hand blocking *agr* would lead to the continuation of the production of adhesins and immune adversins, and could lead to more biofilm formation which

could result in chronic infection. An argument for using *agr* activators in the case of biofilm infection could also be made. Activating *agr* under these circumstances could lead to the dispersion of the biofilm and allow conventional antibiotics to work. Here again there is the problem that this treatment could have undesirable effects, as it would most likely lead to high toxin production and subsequent toxic shock.

The role of *agr* in biofilm formation has, however, been questioned by (Johnson *et al.*, 2008) who showed that *agr* mutants produce less biofilm under iron restricted conditions, which at least some areas of the human body can be assumed to be, giving a potential case for the use of anti-*agr* drugs against biofilms.

As there seems to be a case both for and against this type of treatment in *S. aureus* infections, only time and further research will show whether this will be a fruitful road to take

A research problem that has been identified and attempted to be resolved in this study is the visualisation of AgrC. Lina *et al* are the only group ever to show a western blot of AgrC (Lina *et al.*, 1998) and AgrC has never been purified. Here we tried to recreate Lina *et al*'s western blot of AgrC using recombinant C-terminal part of AgrC to raise AgrC-anti-sera in rats. This, however, proved unsuccessful probably because the rats used had been in contact with *S. aureus* from the environment, and cleaning up the antibodies with whole cell protein preps from *agr* negative *S. aureus* did not prove more successful. From personal communication with Professor Paul Williams and conversations with people at

staphylococcal conferences, it is apparent that many people have tried to tag and purify AgrC over recent years unsuccessfully. Our attempt, with 6X-HIS tag and the NEB Impact™ system also proved unsuccessful. Work on several other methods of visualising AgrC is currently in progress in the group, including expressing HIS-tagged AgrC from its native promoter and gfp tagging both the N-terminal and C-terminal end of AgrC for microscopy purposes.

Another possible target for anti *agr* drugs could be AgrB which is required for AIP production. Qiu *et al* have showed that AgrB specifically cleaves off the C-terminal charged tail of AgrB (Qiu *et al.*, 2005), and Kavanaugh *et al* (2007) have showed a possible role for SspB in the processing of AgrB (Kavanaugh *et al.*, 2007). However, all steps have not yet been clarified in the production of AIP. To investigate if any other proteins are directly involved in the processing of AgrD into AIP a yeast two hybrid system was used to screen for proteins that potentially interact with AgrB. Three proteins were found, but only Rlp, a predicted ABC dipeptide transporter, appeared to be a real possibility. If this potential transporter is involved in AIP production, it is possibly in the secretion of AIP, it could prove a potential target for anti *agr* drugs inhibiting the secretion of AIP and its accumulation in the extracellular space and preventing it from activating AgrC. Currently we are working on knocking *rlp* out in *S. aureus* in order to test if it has any influence on AIP production.

Since the golden heydays of new antibiotics in the 1950's only four new classes have emerged, two in the sixties and then none until oxazolidione in 2000 and

lipopetides in 2003 (Moreillon, 2008). Most new antibiotics and those in clinical trails are members of already known classes and it can therefore be assumed that resistance to these drugs will be a problem sooner or later. Williams P (Williams, 2002) suggested targeting the quorum sensing systems, and thereby the virulence, of *S. aureus* as a way of circumventing existing resistance mechanisms. Various other anti-virulence strategies are also being pursued, these include; anti-toxin-antibodies or toxin target homologues neutralising toxins by blocking their function, targeting toxin transcription factors, anti-adherence drugs mimicking the host target or blocking the pilus assembly(in pili producing bacteria), targeting secretion and anti-QS anti TCS drugs are among the strategies being perused (Reviewed in (Cegelski *et al.*, 2008). Besides the advantage of reduced chance of resistance, anti-virulence drugs also have the advantage of not killing off the natural micro-flora and opening the door for secondary infections. There are, however, a number of potential problems that need to be overcome when developing these drugs. Largest of these seem to be pharmacodynamics and the potential for an inmmune systems response to the drug. As many of the current drugs are peptides or relatively large molecules they tend to have relatively poor pharmacodynamics, as they often will have problems being taken up or defused through barriers and are susceptible to digestive and metabolic breakdown. Also large molecules may be detected by the immune system and could be neutralised by antibodies as a consequence hereof or initiate an adverse immune response (Patrick, 2001). It would also be over optimistic not to consider the possibility that

resistance to these drugs could arise, even though they would not actively select for this resistance by killing non resistant cells.

The research presented in this thesis help give a better understanding of the mechanisms involved in *agr* QS on a molecular level and have given a small clue more on how AgrD is processed into AIP and secreted. These findings will aid hopefully in the design of anti-virulence drugs in the future.

References

NCBI (www.ncbi.nih.gov). 20-4-2006. internet.
Ref Type: Internet Communication

MRSA deaths. National Statistics webpage . 28-8-2008.
Ref Type: Internet Communication

Abdelnour,A., Arvidson,S., Bremell,T., Ryden,C., and Tarkowski,A. (1993) The accessory gene regulator (agr) controls Staphylococcus aureus virulence in a murine arthritis model. *Infect Immun* **61**: 3879-3885.

Ansaldi,M., Marolt,D., Stebe,T., Mandic-Mulec,I., and Dubnau,D. (2002) Specific activation of the Bacillus quorum-sensing systems by isoprenylated pheromone variants. *Mol Microbiol* **44**: 1561-1573.

Arnaud,M., Chastanet,A., and Debarbouille,M. (2004) New vector for efficient allelic replacement in naturally nontransformable, low-GC-content, gram-positive bacteria. *Appl Environ Microbiol* **70**: 6887-6891.

Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A., and Struhl,K. Current protocols in molecular biology. 1988. Chichester UK, John Wiley and Sons.
Ref Type: Generic

Bainton,N.J., Bycroft,B.W., Chhabra,S.R., Stead,P., Gledhill,L., Hill,P.J. *et al.* (1992a) A general role for the lux autoinducer in bacterial cell signalling: control of antibiotic biosynthesis in Erwinia. *Gene* **116**: 87-91.

Bainton,N.J., Stead,P., Chhabra,S.R., Bycroft,B.W., Salmond,G.P., Stewart,G.S., and Williams,P. (1992b) N-(3-oxohexanoyl)-L-homoserine lactone regulates carbapenem antibiotic production in Erwinia carotovora. *Biochem J* **288 (Pt 3)**: 997-1004.

Balaban,N., and Novick,R.P. (1995) Autocrine regulation of toxin synthesis by Staphylococcus aureus. *Proc Natl Acad Sci U S A* **92**: 1619-1623.

Bayer,M.G., Heinrichs,J.H., and Cheung,A.L. (1996) The molecular architecture of the sar locus in Staphylococcus aureus. *J Bacteriol* **178**: 4563-4570.

Benito,Y., Kolb,F.A., Romby,P., Lina,G., Etienne,J., and 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.

Booth,M.C., Atkuri,R.V., Nanda,S.K., Iandolo,J.J., and Gilmore,M.S. (1995) Accessory gene regulator controls Staphylococcus aureus virulence in endophthalmitis. *Invest Ophthalmol Vis Sci* **36**: 1828-1836.

Bruckner,R. (1997) Gene replacement in Staphylococcus carnosus and Staphylococcus xylosus. *FEMS Microbiol Lett* **151**: 1-8.

Bunce,C., Wheeler,L., Reed,G., Musser,J., and Barg,N. (1992) Murine model of cutaneous infection with gram-positive cocci. *Infect Immun* **60**: 2636-2640.

Cegelski,L., Marshall,G.R., Eldridge,G.R., and Hultgren,S.J. (2008) The biology and future prospects of antivirulence therapies. *Nat Rev Microbiol* **6**: 17-27.

Chan,W.C., Coyle,B.J., and Williams,P. (2004) Virulence regulation and quorum sensing in staphylococcal infections: competitive AgrC antagonists as quorum sensing inhibitors. *J Med Chem* **47**: 4633-4641.

Chen,P.R., Bae,T., Williams,W.A., Duguid,E.M., Rice,P.A., Schneewind,O., and He,C. (2006) An oxidation-sensing mechanism is used by the global regulator MgrA in Staphylococcus aureus. *Nat Chem Biol* **2**: 591-595.

Cheung,A.L., Bayer,A.S., Zhang,G., Gresham,H., and Xiong,Y.Q. (2004) Regulation of virulence determinants in vitro and in vivo in Staphylococcus aureus. *FEMS Immunol Med Microbiol* **40**: 1-9.

Cheung,A.L., Bayer,M.G., and Heinrichs,J.H. (1997) sar Genetic determinants necessary for transcription of RNAII and RNAIII in the agr locus of Staphylococcus aureus. *J Bacteriol* **179**: 3963-3971.

Cheung,A.L., Chien,Y.T., and Bayer,A.S. (1999) Hyperproduction of alpha-hemolysin in a sigB mutant is associated with elevated SarA expression in Staphylococcus aureus. *Infect Immun* **67**: 1331-1337.

Cheung,A.L., Eberhardt,K.J., Chung,E., Yeaman,M.R., Sullam,P.M., Ramos,M., and Bayer,A.S. (1994) Diminished virulence of a sar-/agr- mutant of Staphylococcus aureus in the rabbit model of endocarditis. *J Clin Invest* **94**: 1815-1822.

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.

Cheung,A.L., and Projan,S.J. (1994) Cloning and sequencing of sarA of Staphylococcus aureus, a gene required for the expression of agr. *J Bacteriol* **176**: 4168-4172.

Cheung,A.L., Schmidt,K., Bateman,B., and Manna,A.C. (2001) SarS, a SarA homolog repressible by agr, is an activator of protein A synthesis in Staphylococcus aureus. *Infect Immun* **69**: 2448-2455.

Chien,Y., and Cheung,A.L. (1998) Molecular interactions between two global regulators, sar and agr, in Staphylococcus aureus. *J Biol Chem* **273**: 2645-2652.

Chien,Y., Manna,A.C., Projan,S.J., and Cheung,A.L. (1999) SarA, a global regulator of virulence determinants in Staphylococcus aureus, binds to a conserved motif essential for sar-dependent gene regulation. *J Biol Chem* **274**: 37169-37176.

Clarke,S.R., and Foster,S.J. (2008) IsdA protects Staphylococcus aureus against the bactericidal protease activity of apolactoferrin. *Infect Immun* **76**: 1518-1526.

Clarke,S.R., Mohamed,R., Bian,L., Routh,A.F., Kokai-Kun,J.F., Mond,J.J. *et al.* (2007) The Staphylococcus aureus surface protein IsdA mediates resistance to innate defenses of human skin. *Cell Host Microbe* **1**: 199-212.

Clarke,S.R., Wiltshire,M.D., and Foster,S.J. (2004) IsdA of Staphylococcus aureus is a broad spectrum, iron-regulated adhesin. *Mol Microbiol* **51**: 1509-1519.

Claros,M.G., and von,H.G. (1994) TopPred II: an improved software for membrane protein structure predictions. *Comput Appl Biosci* **10**: 685-686.

Cserzo,M., Wallin,E., Simon,I., von,H.G., and Elofsson,A. (1997) Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method. *Protein Eng* **10**: 673-676.

de Ruyter,P.G., Kuipers,O.P., and De Vos,W.M. (1996) Controlled gene expression systems for Lactococcus lactis with the food-grade inducer nisin. *Appl Environ Microbiol* **62**: 3662-3667.

Donlan,R.M., and Costerton,J.W. (2002) Biofilms: survival mechanisms of clinically relevant microorganisms. *Clin Microbiol Rev* **15**: 167-193.

Drake,T.A., and Pang,M. (1988) Staphylococcus aureus induces tissue factor expression in cultured human cardiac valve endothelium. *J Infect Dis* **157**: 749-756.

Dufour,P., Jarraud,S., Vandenesch,F., Greenland,T., Novick,R.P., Bes,M. *et al.* (2002) High genetic variability of the agr locus in Staphylococcus species. *J Bacteriol* **184**: 1180-1186.

Dunman,P.M., Murphy,E., Haney,S., Palacios,D., Tucker-Kellogg,G., Wu,S. *et al.* (2001) Transcription profiling-based identification of Staphylococcus aureus genes regulated by the agr and/or sarA loci. *J Bacteriol* **183**: 7341-7353.

Eberhard,A., Burlingame,A.L., Eberhard,C., Kenyon,G.L., Nealson,K.H., and Oppenheimer,N.J. (1981) Structural identification of autoinducer of Photobacterium fischeri luciferase. *Biochemistry* **20**: 2444-2449.

Fleming,V., Feil,E., Sewell,A.K., Day,N., Buckling,A., and Massey,R.C. (2006) Agr interference between clinical Staphylococcus aureus strains in an insect model of virulence. *J Bacteriol* **188**: 7686-7688.

Foster,T.J. (2005) Immune evasion by staphylococci. *Nat Rev Microbiol* **3**: 948-958.

Foster,T.J., and Hook,M. (1998) Surface protein adhesins of Staphylococcus aureus. *Trends Microbiol* **6**: 484-488.

Fournier,B., and Hooper,D.C. (2000) A new two-component regulatory system involved in adhesion, autolysis, and extracellular proteolytic activity of Staphylococcus aureus. *J Bacteriol* **182**: 3955-3964.

Fournier,B., Klier,A., and Rapoport,G. (2001) The two-component system ArlS-ArlR is a regulator of virulence gene expression in Staphylococcus aureus. *Mol Microbiol* **41**: 247-261.

Foussard,M., Cabantous,S., Pedelacq,J., Guillet,V., Tranier,S., Mourey,L. *et al.* (2001) The molecular puzzle of two-component signaling cascades. *Microbes Infect* **3**: 417-424.

Fowler,V.G., Jr., Sakoulas,G., McIntyre,L.M., Meka,V.G., Arbeit,R.D., Cabell,C.H. *et al.* (2004) Persistent bacteremia due to methicillin-resistant Staphylococcus aureus infection is associated with agr dysfunction and low-level in vitro resistance to thrombin-induced platelet microbicidal protein. *J Infect Dis* **190**: 1140-1149.

Frees,D., Sorensen,K., and Ingmer,H. (2005) Global virulence regulation in Staphylococcus aureus: pinpointing the roles of ClpP and ClpX in the sar/agr regulatory network. *Infect Immun* **73**: 8100-8108.

Fuqua,C., Parsek,M.R., and Greenberg,E.P. (2001) Regulation of gene expression by cell-to-cell communication: acyl-homoserine lactone quorum sensing. *Annu Rev Genet* **35**: 439-468.

Fuqua,W.C., Winans,S.C., and Greenberg,E.P. (1994) Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. *J Bacteriol* **176**: 269-275.

Garvis,S., Mei,J.M., Ruiz-Albert,J., and Holden,D.W. (2002) Staphylococcus aureus svrA: a gene required for virulence and expression of the agr locus. *Microbiology* **148**: 3235-3243.

Geisinger,E., Adhikari,R.P., Jin,R., Ross,H.F., and Novick,R.P. (2006) Inhibition of rot translation by RNAIII, a key feature of agr function. *Mol Microbiol* **61**: 1038-1048.

Geisinger,E., George,E.A., Muir,T.W., and Novick,R.P. (2008) Identification of ligand specificity determinants in AgrC, the Staphylococcus aureus quorum-sensing receptor. *J Biol Chem* **283**: 8930-8938.

Gietz,D., St,J.A., Woods,R.A., and Schiestl,R.H. (1992) Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res* **20**: 1425.

Gillaspy,A.F., Hickmon,S.G., Skinner,R.A., Thomas,J.R., Nelson,C.L., and Smeltzer,M.S. (1995) Role of the accessory gene regulator (agr) in pathogenesis of staphylococcal osteomyelitis. *Infect Immun* **63**: 3373-3380.

Giraudo,A.T., Calzolari,A., Cataldi,A.A., Bogni,C., and Nagel,R. (1999) The sae locus of Staphylococcus aureus encodes a two-component regulatory system. *FEMS Microbiol Lett* **177**: 15-22.

Giraudo,A.T., Raspanti,C.G., Calzolari,A., and Nagel,R. (1994) Characterization of a Tn551-mutant of Staphylococcus aureus defective in the production of several exoproteins. *Can J Microbiol* **40**: 677-681.

Gordon,R.J., and Lowy,F.D. (2008) Pathogenesis of methicillin-resistant Staphylococcus aureus infection. *Clin Infect Dis* **46 Suppl 5**: S350-S359.

Gosbell,I.B., Mercer,J.L., Neville,S.A., Crone,S.A., Chant,K.G., Jalaludin,B.B., and Munro,R. (2001) Non-multiresistant and multiresistant methicillin-resistant Staphylococcus aureus in community-acquired infections. *Med J Aust* **174**: 627-630.

Grebe,T.W., and Stock,J.B. (1999) The histidine protein kinase superfamily. *Adv Microb Physiol* **41**: 139-227.

Grkovic,S., Brown,M.H., Hardie,K.M., Firth,N., and Skurray,R.A. (2003) Stable low-copy-number Staphylococcus aureus shuttle vectors. *Microbiology* **149**: 785-794.

Hamill,R.J., Vann,J.M., and Proctor,R.A. (1986) Phagocytosis of Staphylococcus aureus by cultured bovine aortic endothelial cells: model for postadherence events in endovascular infections. *Infect Immun* **54**: 833-836.

Heumann,D., Barras,C., Severin,A., Glauser,M.P., and Tomasz,A. (1994) Gram-positive cell walls stimulate synthesis of tumor necrosis factor alpha and interleukin-6 by human monocytes. *Infect Immun* **62**: 2715-2721.

Hirokawa,T., Boon-Chieng,S., and Mitaku,S. (1998) SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics* **14**: 378-379.

Hofmann,K., and Stoffel W. (1994) TMbase - A database of membrane spanning proteins segments. *Biol Chem Hoppe-Seyler* **374**.

Holm,C., Meeks-Wagner,D.W., Fangman,W.L., and Botstein,D. (1986) A rapid, efficient method for isolating DNA from yeast. *Gene* **42**: 169-173.

Horsburgh,M.J., Aish,J.L., White,I.J., Shaw,L., Lithgow,J.K., and Foster,S.J. (2002) sigmaB modulates virulence determinant expression and stress resistance: characterization of a functional rsbU strain derived from Staphylococcus aureus 8325-4. *J Bacteriol* **184**: 5457-5467.

Hsieh,H.Y., Tseng,C.W., and Stewart,G.C. (2008) Regulation of Rot expression in Staphylococcus aureus. *J Bacteriol* **190**: 546-554.

Huntzinger,E., Boisset,S., Saveanu,C., Benito,Y., Geissmann,T., Namane,A. *et al.* (2005) Staphylococcus aureus RNAIII and the endoribonuclease III coordinately regulate spa gene expression. *EMBO J* **24**: 824-835.

Ingavale,S., van,W.W., Luong,T.T., Lee,C.Y., and Cheung,A.L. (2005) Rat/MgrA, a regulator of autolysis, is a regulator of virulence genes in Staphylococcus aureus. *Infect Immun* **73**: 1423-1431.

Ingavale,S.S., van,W.W., and Cheung,A.L. (2003) Characterization of RAT, an autolysis regulator in Staphylococcus aureus. *Mol Microbiol* **48**: 1451-1466.

Janzon,L., and 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. *EMBO J* **9**: 1391-1399.

Janzon,L., Lofdahl,S., and Arvidson,S. (1989) Identification and nucleotide sequence of the delta-lysin gene, hld, adjacent to the accessory gene regulator (agr) of Staphylococcus aureus. *Mol Gen Genet* **219**: 480-485.

Jevons,M.P. (1961) " Celbenin "-resistant Staphylococci. *British Medical Journal*: 124-125.

Ji,G., Beavis,R., and Novick,R.P. (1997) Bacterial interference caused by autoinducing peptide variants. *Science* **276**: 2027-2030.

Ji,G., Beavis,R.C., and Novick,R.P. (1995) Cell density control of staphylococcal virulence mediated by an octapeptide pheromone. *Proc Natl Acad Sci U S A* **92**: 12055-12059.

Johnson,M., Cockayne,A., and Morrissey,J.A. (2008) Iron-regulated biofilm formation in Staphylococcus aureus Newman requires ica and the secreted protein Emp. *Infect Immun* **76**: 1756-1765.

Jones,D.T., Taylor,W.R., and Thornton,J.M. (1994) A model recognition approach to the prediction of all-helical membrane protein structure and topology. *Biochemistry* **33**: 3038-3049.

Jones,S., Yu,B., Bainton,N.J., Birdsall,M., Bycroft,B.W., Chhabra,S.R. *et al.* (1993) The lux autoinducer regulates the production of exoenzyme virulence determinants in Erwinia carotovora and Pseudomonas aeruginosa. *EMBO J* **12**: 2477-2482.

Jonsson,I.M., Arvidson,S., Foster,S., and Tarkowski,A. (2004) Sigma factor B and RsbU are required for virulence in Staphylococcus aureus-induced arthritis and sepsis. *Infect Immun* **72**: 6106-6111.

Kahl,B., Herrmann,M., Everding,A.S., Koch,H.G., Becker,K., Harms,E. *et al.* (1998) Persistent infection with small colony variant strains of Staphylococcus aureus in patients with cystic fibrosis. *J Infect Dis* **177**: 1023-1029.

Kaito,C., Morishita,D., Matsumoto,Y., Kurokawa,K., and Sekimizu,K. (2006) Novel DNA binding protein SarZ contributes to virulence in Staphylococcus aureus. *Mol Microbiol* **62**: 1601-1617.

Karimova,G., Pidoux,J., Ullmann,A., and Ladant,D. (1998) A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc Natl Acad Sci U S A* **95**: 5752-5756.

Karlson,P., and Luscher,M. (1959) Pheromones': a new term for a class of biologically active substances. *Nature* **183**: 55-56.

Kavanaugh,J.S., Thoendel,M., and Horswill,A.R. (2007) A role for type I signal peptidase in Staphylococcus aureus quorum sensing. *Mol Microbiol* **65**: 780-798.

Kloos W.E. (1998) Staphylococcus. In Topley and Wilson's Microbiology and Microbial Infections. Collier,L., Balows,A., and Sussman,M. (eds). London: Edward Arnold, pp. 577-630.

Kluytmans,J., van,B.A., and Verbrugh,H. (1997) Nasal carriage of Staphylococcus aureus: epidemiology, underlying mechanisms, and associated risks. *Clin Microbiol Rev* **10**: 505-520.

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.

Krogh,A., Larsson,B., von,H.G., and Sonnhammer,E.L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* **305**: 567-580.

Kushner,S.R. An improved method for transformation of *Escherichia coli* with ColE1-derived plasmids. Boyer, H. B. and Nicosta, S. 17-23. 1978. Amsterdam, Elsevier. Genetic Engineering.
Ref Type: Report

Laemmli,U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**: 680-685.

Li,R., Manna,A.C., Dai,S., Cheung,A.L., and Zhang,G. (2003) Crystal structure of the SarS protein from Staphylococcus aureus. *J Bacteriol* **185**: 4219-4225.

Lina,G., Jarraud,S., Ji,G., Greenland,T., Pedraza,A., Etienne,J. *et al.* (1998) Transmembrane topology and histidine protein kinase activity of AgrC, the agr signal receptor in Staphylococcus aureus. *Mol Microbiol* **28**: 655-662.

Liu,Y., Manna,A., Li,R., Martin,W.E., Murphy,R.C., Cheung,A.L., and Zhang,G. (2001) Crystal structure of the SarR protein from Staphylococcus aureus. *Proc Natl Acad Sci U S A* **98**: 6877-6882.

Liu,Y., Manna,A.C., Pan,C.H., Kriksunov,I.A., Thiel,D.J., Cheung,A.L., and Zhang,G. (2006) Structural and function analyses of the global regulatory protein SarA from Staphylococcus aureus. *Proc Natl Acad Sci U S A* **103**: 2392-2397.

Lowy,F.D. (1998) Staphylococcus aureus infections. *N Engl J Med* **339**: 520-532.

Luong,T.T., Dunman,P.M., Murphy,E., Projan,S.J., and Lee,C.Y. (2006) Transcription Profiling of the mgrA Regulon in Staphylococcus aureus. *J Bacteriol* **188**: 1899-1910.

Luong,T.T., Newell,S.W., and Lee,C.Y. (2003) Mgr, a novel global regulator in Staphylococcus aureus. *J Bacteriol* **185**: 3703-3710.

Lyon,G.J., and Novick,R.P. (2004) Peptide signaling in Staphylococcus aureus and other Gram-positive bacteria. *Peptides* **25**: 1389-1403.

Lyon,G.J., Wright,J.S., Muir,T.W., and Novick,R.P. (2002) Key determinants of receptor activation in the agr autoinducing peptides of Staphylococcus aureus. *Biochemistry* **41**: 10095-10104.

Manna,A., and Cheung,A.L. (2001) Characterization of sarR, a modulator of sar expression in Staphylococcus aureus. *Infect Immun* **69**: 885-896.

Manna,A.C., and Cheung,A.L. (2003) sarU, a sarA homolog, is repressed by SarT and regulates virulence genes in Staphylococcus aureus. *Infect Immun* **71**: 343-353.

Manna,A.C., and Cheung,A.L. (2006a) Expression of SarX, a negative regulator of agr and exoprotein synthesis, is activated by MgrA in Staphylococcus aureus. *J Bacteriol* **188**: 4288-4299.

Manna,A.C., and Cheung,A.L. (2006b) Transcriptional regulation of the agr locus and the identification of DNA binding residues of the global regulatory protein SarR in Staphylococcus aureus. *Mol Microbiol* **60**: 1289-1301.

Manna,A.C., Ingavale,S.S., Maloney,M., van,W.W., and Cheung,A.L. (2004) Identification of sarV (SA2062), a new transcriptional regulator, is repressed by SarA and MgrA (SA0641) and involved in the regulation of autolysis in Staphylococcus aureus. *J Bacteriol* **186**: 5267-5280.

Manna,A.C., and Ray,B. (2007) Regulation and characterization of rot transcription in Staphylococcus aureus. *Microbiology* **153**: 1538-1545.

Mayville,P., Ji,G., Beavis,R., Yang,H., Goger,M., Novick,R.P., and Muir,T.W. (1999) Structure-activity analysis of synthetic autoinducing thiolactone peptides from Staphylococcus aureus responsible for virulence. *Proc Natl Acad Sci U S A* **96**: 1218-1223.

McCormick,J.K., Yarwood,J.M., and Schlievert,P.M. (2001) Toxic shock syndrome and bacterial superantigens: an update. *Annu Rev Microbiol* **55**: 77-104.

McDowell,P., Affas,Z., Reynolds,C., Holden,M.T., Wood,S.J., Saint,S. *et al.* (2001) Structure, activity and evolution of the group I thiolactone peptide quorum-sensing system of Staphylococcus aureus. *Mol Microbiol* **41**: 503-512.

McNamara,P.J., Milligan-Monroe,K.C., Khalili,S., and Proctor,R.A. (2000) Identification, cloning, and initial characterization of rot, a locus encoding a regulator of virulence factor expression in Staphylococcus aureus. *J Bacteriol* **182**: 3197-3203.

McVittie A, MESSIK,F., and ZAHLER,S.A. (1962) Developmental biology of Myxococcus. *J Bacteriol* **84**: 546-551.

Miroux,B., and Walker,J.E. (1996) Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol* **260**: 289-298.

Moller,S., Croning,M.D., and Apweiler,R. (2001) Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics* **17**: 646-653.

Moreillon,P. (2008) New and emerging treatment of Staphylococcus aureus infections in the hospital setting. *Clin Microbiol Infect* **14 Suppl 3**: 32-41.

Morfeldt,E., Taylor,D., von,G.A., and Arvidson,S. (1995) Activation of alpha-toxin translation in Staphylococcus aureus by the trans-encoded antisense RNA, RNAIII. *EMBO J* **14**: 4569-4577.

Morfeldt,E., Tegmark,K., and Arvidson,S. (1996) Transcriptional control of the agr-dependent virulence gene regulator, RNAIII, in Staphylococcus aureus. *Mol Microbiol* **21**: 1227-1237.

Nealson,K.H., Platt,T., and Hastings,J.W. (1970) Cellular control of the synthesis and activity of the bacterial luminescent system. *J Bacteriol* **104**: 313-322.

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.

Novick,R.P. (2003) Autoinduction and signal transduction in the regulation of staphylococcal virulence. *Mol Microbiol* **48**: 1429-1449.

Novick,R.P., and Jiang,D. (2003) The staphylococcal saeRS system coordinates environmental signals with agr quorum sensing. *Microbiology* **149**: 2709-2717.

Novick,R.P., Projan,S.J., Kornblum,J., Ross,H.F., Ji,G., Kreiswirth,B. *et al.* (1995) The agr P2 operon: an autocatalytic sensory transduction system in Staphylococcus aureus. *Mol Gen Genet* **248**: 446-458.

Novick,R.P., Ross,H.F., Projan,S.J., Kornblum,J., Kreiswirth,B., and Moghazeh,S. (1993) Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. *EMBO J* **12**: 3967-3975.

O'Riordan,K., and Lee,J.C. (2004) Staphylococcus aureus capsular polysaccharides. *Clin Microbiol Rev* **17**: 218-234.

Ogawa,S.K., Yurberg,E.R., Hatcher,V.B., Levitt,M.A., and Lowy,F.D. (1985) Bacterial adherence to human endothelial cells in vitro. *Infect Immun* **50**: 218-224.

Ogston,A. (1984) Classics in infectious diseases. "On abscesses". Alexander Ogston (1844-1929). *Rev Infect Dis* **6**: 122-128.

Okada,M., Sato,I., Cho,S.J., Iwata,H., Nishio,T., Dubnau,D., and Sakagami,Y. (2005) Structure of the Bacillus subtilis quorum-sensing peptide pheromone ComX. *Nat Chem Biol* **1**: 23-24.

Passador,L., Cook,J.M., Gambello,M.J., Rust,L., and Iglewski,B.H. (1993) Expression of Pseudomonas aeruginosa virulence genes requires cell-to-cell communication. *Science* **260**: 1127-1130.

Patrick,G.L. (2001) An Introduction to Medicinal Chemistry New York: Oxford University Press.

Patti,J.M., Allen,B.L., McGavin,M.J., and Hook,M. (1994) MSCRAMM-mediated adherence of microorganisms to host tissues. *Annu Rev Microbiol* **48**: 585-617.

Peacock,S.J., Moore,C.E., Justice,A., Kantzanou,M., Story,L., Mackie,K. *et al.* (2002) Virulent combinations of adhesin and toxin genes in natural populations of Staphylococcus aureus. *Infect Immun* **70**: 4987-4996.

Persson,B., and Argos,P. (1994) Prediction of transmembrane segments in proteins utilising multiple sequence alignments. *J Mol Biol* **237**: 182-192.

Persson,B., and Argos,P. (1996) Topology prediction of membrane proteins. *Protein Sci* **5**: 363-371.

Pragman,A.A., Yarwood,J.M., Tripp,T.J., and Schlievert,P.M. (2004) Characterization of virulence factor regulation by SrrAB, a two-component system in Staphylococcus aureus. *J Bacteriol* **186**: 2430-2438.

Proctor,R.A., and Peters,G. (1998) Small colony variants in staphylococcal infections: diagnostic and therapeutic implications. *Clin Infect Dis* **27**: 419-422.

Proctor,R.A., van,L.P., Kristjansson,M., Maslow,J.N., and Arbeit,R.D. (1995) Persistent and relapsing infections associated with small-colony variants of Staphylococcus aureus. *Clin Infect Dis* **20**: 95-102.

Qazi,S.N., Counil,E., Morrissey,J., Rees,C.E., Cockayne,A., Winzer,K. *et al.* (2001) agr expression precedes escape of internalized Staphylococcus aureus from the host endosome. *Infect Immun* **69**: 7074-7082.

Qiu,R., Pei,W., Zhang,L., Lin,J., and Ji,G. (2005) Identification of the putative staphylococcal AgrB catalytic residues involving the proteolytic cleavage of AgrD to generate autoinducing peptide. *J Biol Chem* **280**: 16695-16704.

Quadri,L.E. (2002) Regulation of antimicrobial peptide production by autoinducer-mediated quorum sensing in lactic acid bacteria. *Antonie Van Leeuwenhoek* **82**: 133-145.

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.

Roberts,C., Anderson,K.L., Murphy,E., Projan,S.J., Mounts,W., Hurlburt,B. *et al.* (2006) Characterizing the effect of the Staphylococcus aureus virulence factor regulator, SarA, on log-phase mRNA half-lives. *J Bacteriol* **188**: 2593-2603.

Rost,B., Fariselli,P., and Casadio,R. (1996) Topology prediction for helical transmembrane proteins at 86% accuracy. *Protein Sci* **5**: 1704-1718.

Said-Salim,B., Dunman,P.M., McAleese,F.M., Macapagal,D., Murphy,E., McNamara,P.J. *et al.* (2003) Global regulation of Staphylococcus aureus genes by Rot. *J Bacteriol* **185**: 610-619.

Sakoulas,G., Eliopoulos,G.M., Moellering,R.C., Jr., Wennersten,C., Venkataraman,L., Novick,R.P., and Gold,H.S. (2002) Accessory gene regulator (agr) locus in geographically diverse Staphylococcus aureus isolates with reduced susceptibility to vancomycin. *Antimicrob Agents Chemother* **46**: 1492-1502.

Sambrook,J., Fritsch,E.F., and Maniatis,T. Molecular Cloning, a Laboratory Manual. 2nd. 1989. Cold Spring Harbor, Cold Spring Harbor Laboratory.
Ref Type: Generic

Schenk,S., and Laddaga,R.A. (1992) Improved method for electroporation of Staphylococcus aureus. *FEMS Microbiol Lett* **73**: 133-138.

Schmidt,K.A., Manna,A.C., and Cheung,A.L. (2003) SarT influences sarS expression in Staphylococcus aureus. *Infect Immun* **71**: 5139-5148.

Schmidt,K.A., Manna,A.C., Gill,S., and Cheung,A.L. (2001) SarT, a repressor of alpha-hemolysin in Staphylococcus aureus. *Infect Immun* **69**: 4749-4758.

Scott,R.J., Lian,L.Y., Muharram,S.H., Cockayne,A., Wood,S.J., Bycroft,B.W. *et al.* (2003) Side-chain-to-tail thiolactone peptide inhibitors of the staphylococcal quorum-sensing system. *Bioorg Med Chem Lett* **13**: 2449-2453.

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.

Smith EF. Bacteria in relation to plant disease vol 1. 1905. Washington DC, Carnegie Institution. Carnegie Institution Report.
Ref Type: Report

Sonnhammer,E.L., von,H.G., and Krogh,A. (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol* **6**: 175-182.

Sterba,K.M., Mackintosh,S.G., Blevins,J.S., Hurlburt,B.K., and Smeltzer,M.S. (2003) Characterization of Staphylococcus aureus SarA binding sites. *J Bacteriol* **185**: 4410-4417.

Stryer L (1997) Gene Rearrangements. In Biochemistry. New York: W. H. Freeman and Company, pp. 819-840.

Swift,S., Downie,J.A., Whitehead,N.A., Barnard,A.M., Salmond,G.P., and Williams,P. (2001) Quorum sensing as a population-density-dependent determinant of bacterial physiology. *Adv Microb Physiol* **45**: 199-270.

Swift,S., Winson,M.K., Chan,P.F., Bainton,N.J., Birdsall,M., Reeves,P.J. *et al.* (1993) A novel strategy for the isolation of luxI homologues: evidence for the widespread distribution of a LuxR:LuxI superfamily in enteric bacteria. *Mol Microbiol* **10**: 511-520.

Takano,E., Nihira,T., Hara,Y., Jones,J.J., Gershater,C.J., Yamada,Y., and Bibb,M. (2000) Purification and structural determination of SCB1, a gamma-butyrolactone that elicits antibiotic production in Streptomyces coelicolor A3(2). *J Biol Chem* **275**: 11010-11016.

Taylor,J.M., and Heinrichs,D.E. (2002) Transferrin binding in Staphylococcus aureus: involvement of a cell wall-anchored protein. *Mol Microbiol* **43**: 1603-1614.

Tegmark,K., Karlsson,A., and Arvidson,S. (2000) Identification and characterization of SarH1, a new global regulator of virulence gene expression in Staphylococcus aureus. *Mol Microbiol* **37**: 398-409.

Throup,J.P., Zappacosta,F., Lunsford,R.D., Annan,R.S., Carr,S.A., Lonsdale,J.T. *et al.* (2001) The srhSR gene pair from Staphylococcus aureus: genomic and proteomic approaches to the identification and characterization of gene function. *Biochemistry* **40**: 10392-10401.

Timmerman,C.P., Mattsson,E., Martinez-Martinez,L., De,G.L., Van Strijp,J.A., Verbrugh,H.A. *et al.* (1993) Induction of release of tumor necrosis factor from human monocytes by staphylococci and staphylococcal peptidoglycans. *Infect Immun* **61**: 4167-4172.

Traber,K.E., Lee,E., Benson,S., Corrigan,R., Cantera,M., Shopsin,B., and Novick,R.P. (2008) agr function in clinical Staphylococcus aureus isolates. *Microbiology* **154**: 2265-2274.

Truong-Bolduc,Q.C., Dunman,P.M., Strahilevitz,J., Projan,S.J., and Hooper,D.C. (2005) MgrA is a multiple regulator of two new efflux pumps in Staphylococcus aureus. *J Bacteriol* **187**: 2395-2405.

Truong-Bolduc,Q.C., and Hooper,D.C. (2007) The transcriptional regulators NorG and MgrA modulate resistance to both quinolones and beta-lactams in Staphylococcus aureus. *J Bacteriol* **189**: 2996-3005.

Truong-Bolduc,Q.C., Strahilevitz,J., and Hooper,D.C. (2006) NorC, a new efflux pump regulated by MgrA of Staphylococcus aureus. *Antimicrob Agents Chemother* **50**: 1104-1107.

Truong-Bolduc,Q.C., Zhang,X., and Hooper,D.C. (2003) Characterization of NorR protein, a multifunctional regulator of norA expression in Staphylococcus aureus. *J Bacteriol* **185**: 3127-3138.

Tusnady,G.E., and Simon,I. (2001) The HMMTOP transmembrane topology prediction server. *Bioinformatics* **17**: 849-850.

Udo,E.E., Pearman,J.W., and Grubb,W.B. (1993) Genetic analysis of community isolates of methicillin-resistant Staphylococcus aureus in Western Australia. *J Hosp Infect* **25**: 97-108.

van,d.M., Jr., Polman,J., Beerthuyzen,M.M., Siezen,R.J., Kuipers,O.P., and De Vos,W.M. (1993) Characterization of the Lactococcus lactis nisin A operon genes nisP, encoding a subtilisin-like serine protease involved in precursor processing, and nisR, encoding a regulatory protein involved in nisin biosynthesis. *J Bacteriol* **175**: 2578-2588.

Vandenesch,F., Naimi,T., Enright,M.C., Lina,G., Nimmo,G.R., Heffernan,H. *et al.* (2003) Community-acquired methicillin-resistant Staphylococcus aureus carrying Panton-Valentine leukocidin genes: worldwide emergence. *Emerg Infect Dis* **9**: 978-984.

von Heijne G. (1992) Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. *J Mol Biol* **225**: 487-494.

Wertheim,H.F., Melles,D.C., Vos,M.C., van,L.W., van,B.A., Verbrugh,H.A., and Nouwen,J.L. (2005) The role of nasal carriage in Staphylococcus aureus infections. *Lancet Infect Dis* **5**: 751-762.

Williams,P. (2002) Quorum sensing: an emerging target for antibacterial chemotherapy? *Expert Opin Ther Targets* **6**: 257-274.

Williams,R.E., JEVONS,M.P., SHOOTER,R.A., HUNTER,C.J., GIRLING,J.A., GRIFFITHS,J.D., and TAYLOR,G.W. (1959) Nasal staphylococci and sepsis in hospital patients. *Br Med J* **2**: 658-662.

Winzer,K., Hardie,K.R., Burgess,N., Doherty,N., Kirke,D., Holden,M.T. *et al.* (2002a) LuxS: its role in central metabolism and the in vitro synthesis of 4-hydroxy-5-methyl-3(2H)-furanone. *Microbiology* **148**: 909-922.

Winzer,K., Hardie,K.R., and Williams,P. (2002b) Bacterial cell-to-cell communication: sorry, can't talk now - gone to lunch! *Curr Opin Microbiol* **5**: 216-222.

Wisplinghoff,H., Rosato,A.E., Enright,M.C., Noto,M., Craig,W., and Archer,G.L. (2003) Related clones containing SCCmec type IV predominate among clinically significant Staphylococcus epidermidis isolates. *Antimicrob Agents Chemother* **47**: 3574-3579.

Witte,W., Cuny,C., Klare,I., Nubel,U., Strommenger,B., and Werner,G. (2008) Emergence and spread of antibiotic-resistant Gram-positive bacterial pathogens. *Int J Med Microbiol* **298**: 365-377.

Wright,J.S., III, Jin,R., and Novick,R.P. (2005a) Transient interference with staphylococcal quorum sensing blocks abscess formation. *Proc Natl Acad Sci U S A* **102**: 1691-1696.

Wright,J.S., III, Lyon,G.J., George,E.A., Muir,T.W., and Novick,R.P. (2004) Hydrophobic interactions drive ligand-receptor recognition for activation and inhibition of staphylococcal quorum sensing. *Proc Natl Acad Sci U S A* **101**: 16168-16173.

Wright,J.S., III, Traber,K.E., Corrigan,R., Benson,S.A., Musser,J.M., and Novick,R.P. (2005b) The agr radiation: an early event in the evolution of staphylococci. *J Bacteriol* **187**: 5585-5594.

Xavier,K.B., and Bassler,B.L. (2003) LuxS quorum sensing: more than just a numbers game. *Curr Opin Microbiol* **6**: 191-197.

Yarwood,J.M., McCormick,J.K., and Schlievert,P.M. (2001) Identification of a novel two-component regulatory system that acts in global regulation of virulence factors of Staphylococcus aureus. *J Bacteriol* **183**: 1113-1123.

Zhang,L., Gray,L., Novick,R.P., and Ji,G. (2002) Transmembrane topology of AgrB, the protein involved in the post-translational modification of AgrD in Staphylococcus aureus. *J Biol Chem* **277**: 34736-34742.

Zhang,L., and Ji,G. (2004) Identification of a staphylococcal AgrB segment(s) responsible for group-specific processing of AgrD by gene swapping. *J Bacteriol* **186**: 6706-6713.

Zhang,L., Lin,J., and 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. *J Biol Chem* **279**: 19448-19456.

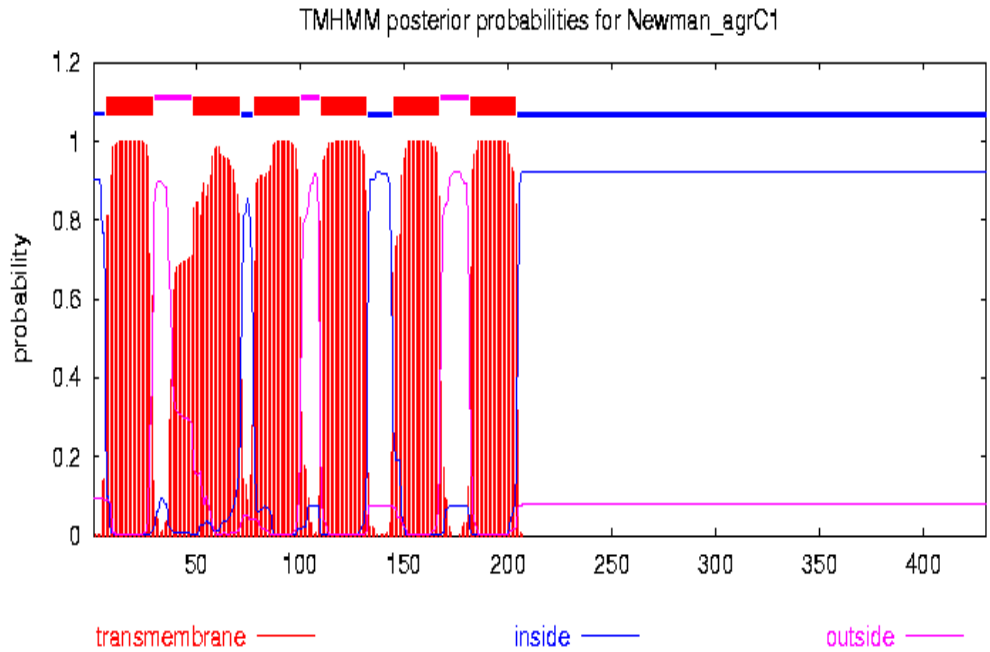
Zhang,L., Murphy,P.J., Kerr,A., and Tate,M.E. (1993) Agrobacterium conjugation and gene regulation by N-acyl-L-homoserine lactones. *Nature* **362**: 446-448.

Appendix A

TMHMM 2.0

HELP with output formats

```
# Newman_agrC1 Length: 430
# Newman_agrC1 Number of predicted TMHs:  6
# Newman_agrC1 Exp number of AAs in TMHs: 142.64221
# Newman_agrC1 Exp number, first 60 AAs:  40.58505
# Newman_agrC1 Total prob of N-in:        0.90392
# Newman_agrC1 POSSIBLE N-term signal sequence
Newman_agrC1        TMHMM2.0   inside      1    6
Newman_agrC1        TMHMM2.0   TMhelix           7    29
Newman_agrC1        TMHMM2.0   outside     30    48
Newman_agrC1        TMHMM2.0   TMhelix           49    71
Newman_agrC1        TMHMM2.0   inside      72    77
Newman_agrC1        TMHMM2.0   TMhelix           78    100
Newman_agrC1        TMHMM2.0   outside     101    109
Newman_agrC1        TMHMM2.0   TMhelix           110    132
Newman_agrC1        TMHMM2.0   inside      133    144
Newman_agrC1        TMHMM2.0   TMhelix           145    167
Newman_agrC1        TMHMM2.0   outside     168    181
Newman_agrC1        TMHMM2.0   TMhelix           182    204
Newman_agrC1        TMHMM2.0   inside      205    430
```



plot in postscript, script for making the plot in gnuplot, data for plot

---

Toppred
Algorithm specific parameters:

Full window size : 21
Core window size : 11
Wedge window size: 5
Using hydrophobicity file: GES-scale

Cutoff for certain transmembrane segments: 1.00
Cutoff for putative transmembrane segments: 0.60
Critical distance between 2 transmembrane segments: 2

Critical loop length: 60

Kingdom: procaryote

Using cyt/ext file: CYTEXT-scale

Sequence : unknown (430 res)
MELLNSYNFVLFVLTQMILMFTIPAIISGIKYSKLDYFFIIVISTLSLFLFKMF
DSASLI
ILTSFIIIMYFVKIKWYSILLIMTSQIILYCANYMYIVIYAYITKISDSIFVIFPS
FFVV
YVTISILFSYIINRVLKKISTPYLILNKGFLIVISTILLLTFSLFFFYSQINSDEA
KVIR
QYSFIFIGITIFLSILTFVISQFLLKEMKYKRNQEEIETYYEYTLKIEAINNEM
RKFRHD
YVNILTTLSEYIREDDMPGLRDYFNKNIVPMKDNLQMNAIKLNGIENLKV
REIKGLITAK
ILRAQEMNIPISIEIPDEVSSINLNMIDLSRSIGIILDNAIEASTEIDDPIIRVAFI
ESE
NSVTFIVMNKCADDIPRIHELFQESFSTKGEGRGLGLSTLKEIADNADNVL
LDTIIENGF
FIQKVEIINN

Found: 6 segments

Candidate membrane-spanning segments:

 Helix Begin - End   Score Certainity
    1   8 - 28   2.092 Certain

2 54 - 74 1.886 Certain
3 84 - 104 1.004 Certain
4 109 - 129 2.219 Certain
5 149 - 169 2.438 Certain
6 182 - 202 2.417 Certain

Total of 1 structures are to be tested

HEADER START STOP LEN PROB HP DARGLYS DCYTEXT DNCHARGE DNNEGPOS
TOPOLOGY 1 1.00 1.00 -0.35 -2.00 1.00
TOPOLOGY N-in N-in N-out
CYT_LOOP 1 7 7 0.00 (0.25)
TRANSMEM 8 28 21 1.00 2.09
EXT_LOOP 29 53 25 3.00 (0.52)
TRANSMEM 54 74 21 1.00 1.89
CYT_LOOP 75 83 9 2.00 (0.24)
TRANSMEM 84 104 21 1.00 1.00
EXT_LOOP 105 108 4 1.00 (-0.19)
TRANSMEM 109 129 21 1.00 2.22
CYT_LOOP 130 148 19 4.00 (0.43)
TRANSMEM 149 169 21 1.00 2.44
EXT_LOOP 170 181 12 2.00 (-0.15)
TRANSMEM 182 202 21 1.00 2.42
CYT_LOOP 203 430 228 (26.00) -0.35
TMpred
TMpred output for AgrC1
[ISREC-Server] Date: Thu Sep 18 15:45:06 2008

---

tmpred -par=matrix.tab -html -min=17 -max=33 -def -in=wwwtmp/.TMPRED.6553.6779.seq -out=wwwtmp/.TMPRED.6553.6779.out -out2=wwwtmp/.TMPRED.6553.6779.out2 -out3=wwwtmp/.TMPRED.6553.6779.txt >wwwtmp/.TMPRED.6553.6779.err
Sequence: MEL...INN, length: 430
Prediction parameters: TM-helix length between 17 and 33

1.)Possible transmembrane helices
The sequence positions in brackets denominate the core region.
Only scores above 500 are considered significant.
Inside to outside helices : 6 found
from to score center
9 (9) 27 (27) 2110 19
53 (53) 72 (72) 2198 62
79 (79) 103 (99) 1952 90
110 (112) 128 (128) 2285 120
149 (151) 168 (168) 2620 159

185 (185) 205 (202)   2707   194

Outside to inside helices :   6 found
     from   to   score center
  9 (9) 30 (27)   1687   19
 53 (53) 72 (72)   2321   61
 74 (80) 102 (97)   1365   89
113 (113) 132 (129)   2464   121
149 (149) 168 (168)   2518   159
184 (184) 205 (205)   2548   194

2.) Table of correspondences
Here is shown, which of the inside->outside helices correspond to which of the outside->inside helices.
Helices shown in brackets are considered insignificant.
A "+"-symbol indicates a preference of this orientation.
A "++"-symbol indicates a strong preference of this orientation.

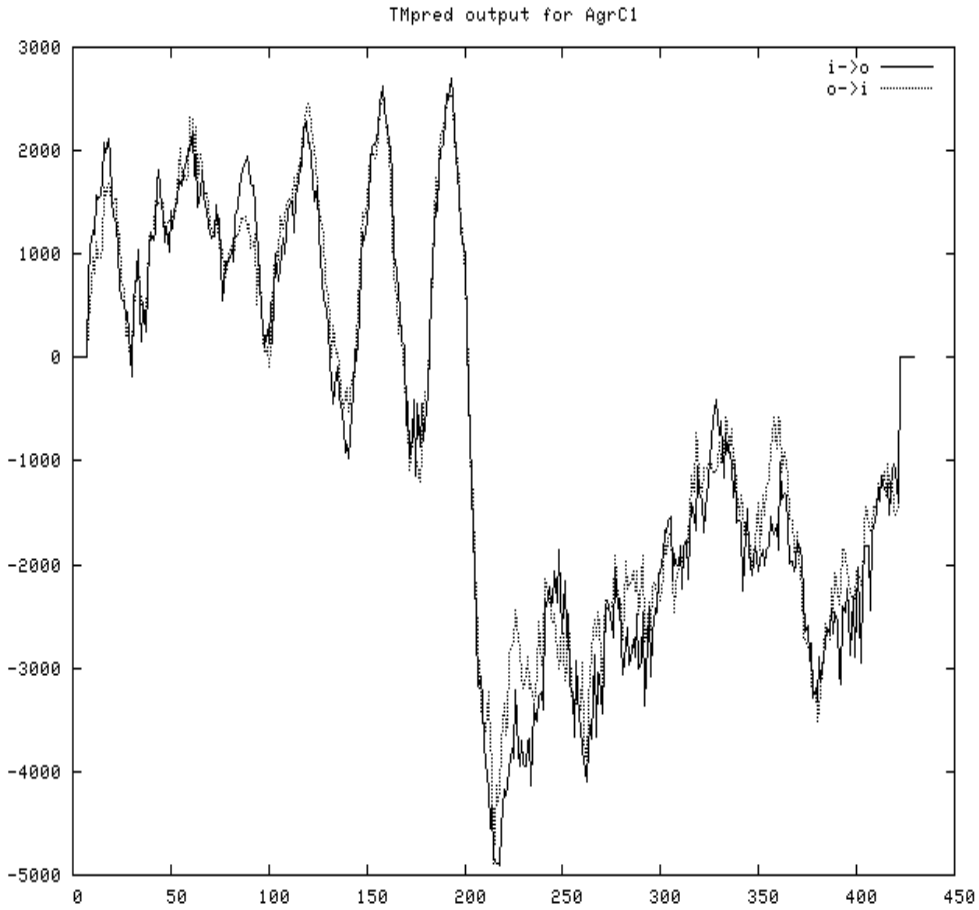
        inside->outside | outside->inside
   9- 27 (19) 2110 ++ | 9- 30 (22) 1687
  53- 72 (20) 2198 | 53- 72 (20) 2321 +
  79- 103 (25) 1952 ++ | 74- 102 (29) 1365
 110- 128 (19) 2285 | 113- 132 (20) 2464 +
 149- 168 (20) 2620 + | 149- 168 (20) 2518
 185- 205 (21) 2707 + | 184- 205 (22) 2548

3.) Suggested models for transmembrane topology
These suggestions are purely speculative and should be used with extreme caution since they are based on the assumption that all transmembrane helices have been found.
In most cases, the Correspondence Table shown above or the prediction plot that is also created should be used for the topology assignment of unknown proteins.

2 possible models considered, only significant TM-segments used

-----> STRONGLY prefered model: N-terminus inside
 6 strong transmembrane helices, total score : 14015
 # from   to length score orientation
 1   9   27 (19)   2110 i-o
 2   53   72 (20)   2321 o-i
 3   79 103 (25)   1952 i-o
 4 113 132 (20)   2464 o-i
 5 149 168 (20)   2620 i-o
 6 184 205 (22)   2548 o-i

------> alternative model
6 strong transmembrane helices, total score : 12760
from to length score orientation
1 9 30 (22) 1687 o-i
2 53 72 (20) 2198 i-o
3 74 102 (29) 1365 o-i
4 110 128 (19) 2285 i-o
5 149 168 (20) 2518 o-i
6 185 205 (21) 2707 i-o



You can get the prediction graphics shown above in one of the following formats:
GIF-format
Postscript-format
numerical format

Back to ISREC home page

Tmap
Results from TMAP prediction:

Predictions for the sequence is given below.
A PostScript file showing the P values plotted against sequence positions can be downloaded by clicking here.

Please cite Persson, B. & Argos, P. (1994) J. Mol. Biol. 237, 182-192. and
Persson, B. & Argos, P. (1996) Prot. Sci. 5, 363-371.
Please notify me via E-mail if the server does not behave properly.

Thank you for using TMAP!
Bengt Persson

RESULTS from program TMAP, edition 51

Numbers give: a) number of transmembrane segment
b) start of TM segment (alignment position / residue number)
c) end of TM segment (alignment position / residue number)
d) length of TM segment within parentheses

NO CERTAIN TOPOLOGY PREDICTION (contributions: 6 in 6 out K out R in)

PREDICTED TRANSMEMBRANE SEGMENTS FOR SEQUENCE
/var/www/html/temp/1221747504.99968.msf

TM 1: 11 - 39 (29.0)
TM 2: 50 - 78 (29.0)
TM 3: 83 - 106 (24.0)
TM 4: 112 - 140 (29.0)
TM 5: 150 - 178 (29.0)
TM 6: 188 - 216 (29.0)

PREDICTED TRANSMEMBRANE SEGMENTS FOR PROTEIN

TM 1: 11 - 39 (29)
TM 2: 50 - 78 (29)
TM 3: 83 - 106 (24)
TM 4: 112 - 140 (29)
TM 5: 150 - 178 (29)
TM 6: 188 - 216 (29)

SOSUI
SOSUI Result

Query title : >Newman_agrC1
Total length : 430 A. A.
Average of hydrophobicity : 0.493721

This amino acid sequence is of a MEMBRANE PROTEIN
which have 7 transmembrane helices.

No.	N terminal	transmembrane region	C terminal	type	length
1	6	SYNFVLFVLTQMILMFTIPAIIS	28	PRIMARY	23
2	34	KLDYFFIIVISTLSLFLFKMFDS	56	PRIMARY	23
3	60	IILTSFIIIMYFVKIKWYSILLI	82	PRIMARY	23
4	85	SQIILYCANYMYIVIYAYITKIS	107	PRIMARY	23
5	110	IFVIFPSFFVVYVTISILFSYII	132	PRIMARY	23
6	148	KGFLIVISTILLLTFSLFFFYSQ	170	PRIMARY	23
7	182	YSFIFIGITIFLSILTFVISQFL	204	PRIMARY	23

PHDhtm

```
********************************************************************
**********
*                                                                  *
*   PHD: Profile fed neural network systems from HeiDelberg        *
*   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
*
*                                                                  *
*   Prediction of:                                                 *
*       secondary structure,          by PHDsec                    *
*       solvent accessibility,        by PHDacc                    *
*       and helical transmembrane regions,  by PHDhtm
*                                                                  *
*   Author:
*       Burkhard Rost
*     EMBL, 69012 Heidelberg, Germany                              *
*     Internet: Rost@EMBL-Heidelberg.DE
*                                                                  *
*   All rights reserved.                                           *
*                                                                  *
********************************************************************
**********
*                                                                  *
*   The network systems are described in:                          *
*                                                                  *
*   PHDsec:   B Rost & C Sander: JMB, 1993, 232, 584-599.
*             B Rost & C Sander: Proteins, 1994, 19, 55-72.
*   PHDacc:   B Rost & C Sander: Proteins, 1994, 20, 216-226.      *
*   PHDhtm:   B Rost et al.:     Prot. Science, 1995, 4, 521-533.
*                                                                  *
********************************************************************
**********
*                                                                  *
*   Some statistics                                                *
*   ~~~~~~~~~~~~~~~~                                               *
*                                                                  *
*   Percentage of amino acids:                                     *
*   +-------------+--------+--------+--------+--------+--------+  *
*   | AA:         |   I    |   L    |   S    |   F    |   N    |  *
*   | % of AA:    |  17.4  |  10.5  |   8.1  |   7.7  |   6.3  |  *
*   +-------------+--------+--------+--------+--------+--------+  *
*   | AA:         |   E    |   K    |   Y    |   V    |   T    |  *
*   | % of AA:    |   6.0  |   5.8  |   5.1  |   5.1  |   5.1  |  *
*   +-------------+--------+--------+--------+--------+--------+  *
*   | AA:         |   D    |   M    |   A    |   R    |   G    |  *
*   | % of AA:    |   4.4  |   3.5  |   3.5  |   3.0  |   2.8  |  *
```

```
*   +--------------+--------+--------+--------+--------+--------+       *
*   | AA:          |   Q    |   P    |   H    |   C    |   W    |       *
*   | % of AA:     |   2.3  |   2.1  |   0.5  |   0.5  |   0.2  |       *
*   +--------------+--------+--------+--------+--------+--------+       *
*                                                                        *
*   Percentage of helical trans-membrane predicted:                      *
*   +--------------+--------+--------+                                   *
*   | SecStr:      |   H    |   L    |                                   *
*   | % Predicted: |  46.3  |  53.7  |                                   *
*   +--------------+--------+--------+                                   *
*                                                                        *
**************************************************************************
**********
*                                                                        *
*   PHD output for your protein                                          *
*   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~                                      *
*                                                                        *
*   Thu Sep 18 16:19:06 2008                                             *
*   Jury on:       4    different architectures (version   8.94_69  ).   *
*   Note: differently trained architectures, i.e., different versions can *
*   result in different predictions.                                     *
*                                                                        *
**************************************************************************
**********
*                                                                        *
*   About the protein                                                    *
*   ~~~~~~~~~~~~~~~~~~                                                   *
*                                                                        *
*   HEADER                                                               *
*   COMPND                                                               *
*   SOURCE                                                               *
*   AUTHOR                                                               *
*   SEQLENGTH   430                                                      *
*   NCHAIN       1 chain(s) in Newman_agrC1 data set                     *
*   NALIGN       8                                                       *
*   (=number of aligned sequences in HSSP file)                          *
*                                                                        *
*                                                                        *
*   protein:      Newman_       length     430                         *
*                                                                        *
---
--- ----------------------------------------------------------
--- PhdTopology prediction of transmembrane helices and topology
--- ----------------------------------------------------------
---
```

```
--- PhdTopology REFINEMENT AND TOPOLOGY HEADER:
ABBREVIATIONS
---
--- NHTM_BEST    : number of transmembrane helices best model
--- NHTM_2ND_BEST: number of transmembrane helices 2nd best model
--- REL_BEST     : reliability of best model (0 is low, 9 high)
--- HTMTOP_PRD   : topology predicted ('in': intra-cytoplasmic)
--- HTMTOP_RID   : difference between positive charges
--- HTMTOP_RIP   : reliability of topology prediction (0-9)
--- MOD_NHTM     : number of transmembrane helices of model
--- MOD_STOT     : score for all residues
--- MOD_SHTM     : score for HTM added at current iteration step
--- MOD_N-C      : N  -  C  term of HTM added at current step
---
--- ALGORITHM REF: The refinement is performed by a dynamic pro-
--- ALGORITHM    : gramming-like procedure: iteratively the best
--- ALGORITHM    : transmembrane helix (HTM) compatible with the
--- ALGORITHM    : network output is added (starting from the  0
--- ALGORITHM    : assumption, i.e.,  no HTM's  in the protein).
--- ALGORITHM TOP: Topology is predicted by the  positive-inside
--- ALGORITHM    : rule, i.e., the positive charges are compiled
--- ALGORITHM    : separately  for all even and all odd  non-HTM
--- ALGORITHM    : regions.  If the difference (charge even-odd)
--- ALGORITHM    : is < 0, topology is predicted as 'in'.   That
--- ALGORITHM    : means, the protein N-term starts on the intra
--- ALGORITHM    : cytoplasmic side.
---
--- PhdTopology REFINEMENT HEADER: SUMMARY
 MOD_NHTM MOD_STOT MOD_SHTM MOD_N-C
      1   0.618    0.902   183 -   201
      2   0.663    0.891   108 -   132
      3   0.700    0.891   148 -   167
      4   0.732    0.890    42 -    59
      5   0.777    0.887    64 -    88
      6   0.811    0.881    14 -    32
---
--- PhdTopology REFINEMENT AND TOPOLOGY HEADER: SUMMARY
--- NHTM_BEST    : 6
--- NHTM_2ND_BEST: 5
--- REL_BEST     : 3
--- HTMTOP_PRD   : in
--- HTMTOP_RID   : -3.438
--- HTMTOP_RIP   : 3
---
--- PhdTopology REFINEMENT AND TOPOLOGY PREDICTION: SYMBOLS
--- AA           : amino acid in one-letter code
```

```
--- PHD htm      : HTM's predicted by the PHD neural network
---              system (H=HTM, ' '=not HTM)
--- Rel htm      : Reliability index of prediction (0-9, 0 is low)
--- detail       : Neural network output in detail
--- prH htm      : 'Probability' for assigning a helical trans-
---              membrane region (HTM)
--- prL htm      : 'Probability' for assigning a non-HTM region
---         note: 'Probabilites' are scaled to the interval
---              0-9, e.g.,, prH=5 means, that the first
---              output node is 0.5-0.6
--- subset       : Subset of more reliable predictions
--- SUB htm      : All residues for which the expected average
---              accuracy is > 82% (tables in header).
---         note: for this subset the following symbols are used:
---           L: is loop (for which above ' ' is used)
---          '.': means that no prediction is made for this,
---              residue as the reliability is:  Rel < 5
--- other        : predictions derived based on PHDhtm
--- PHDFhtm      : filtered prediction, i.e., too long HTM's are
---              split, too short ones are deleted
--- PHDRhtm      : refinement of neural network output
--- PHDThtm      : topology prediction based on refined model
---              symbols used:
---            i: intra-cytoplasmic
---            T: transmembrane region
---            o: extra-cytoplasmic
---
--- PhdTopology REFINEMENT AND TOPOLOGY PREDICTION
                ....,....1....,....2....,....3....,....4....,....5....,....6
       AA
|MELLNSYNFVLFVLTQMILMFTIPAIISGIKYSKLDYFFIIVISTLSLFLFKMF
DSASLI|
       PHD htm |
HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
HHHH|
       Rel htm
|99998887764035678888888887777655555667777777777777777777777|
 detail:
       prH htm
|000000011125678899999999988888777778888888888888888888888888|
       prL htm
|99999998887432110000000001111122222111111111111111111111111|
 subset: SUB htm
|LLLLLLLLL......HHHHHHHHHHHHHH.....HHHHHHHHHHHHHHHHHHHH
HHHH|
```

PHDRhtm | HHHHHHHHHHHHHHHHHHH
HHHHHHHHHHHHHHHHHHH |
PHDThtm
|iiiiiiiiiiiiTTTTTTTTTTTTTTTTTTTTooooooooTTTTTTTTTTTTTTTTTTi|
....,....7....,....8....,....9....,....10...,....11...,....12
AA
|ILTSFIIIMYFVKIKWYSILLIMTSQIILYCANYMYIVIYAYITKISDSIFVIFPS
FFVV|
PHD htm
|HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
HHHHHHHHHHHHHHHH|
Rel htm
|77777777777777777777777777777777777777777666677777777777777|
detail:
prH htm
|88888888888888888888888888888888888888888888888888888888|
prL htm
|11111111111111111111111111111111111111111111111111111111|
subset: SUB htm
|HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH....H
HHHHHHHHHHHHHH|
PHDRhtm | HHHHHHHHHHHHHHHHHHHHHHHHHH
HHHHHHHHHHHHHH|
PHDThtm
|iiiTTTTTTTTTTTTTTTTTTTTTTTTooooooooooooooooooTTTTTTTTTTTT
T|
....,....13...,....14...,....15...,....16...,....17...,....18
AA
|YVTISILFSYIINRVLKKISTPYLILNKGFLIVISTILLLTFSLFFFYSQINSDEA
KVIR|
PHD htm |HHHHHHHHHHHHHHHHHHH
HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH |
Rel htm
|77777777787777653001012455567777777777778888877630235676653 2|
detail:
prH htm
|88888888988888765545567777888888888888899999888653321111233|
prL htm
|11111111011111234454432222111111111111000001113466788887 66|
subset: SUB htm
|HHHHHHHHHHHHHHH.............HHHHHHHHHHHHHHHHHHH.....L.....|
PHDRhtm |HHHHHHHHHHHH
HHHHHHHHHHHHHHHHHHHHH |
PHDThtm
|TTTTTTTTTTTiiiiiiiiiiiiiiTTTTTTTTTTTTTTTTTTooooooooooooo|
....,....19...,....20...,....21...,....22...,....23...,....24

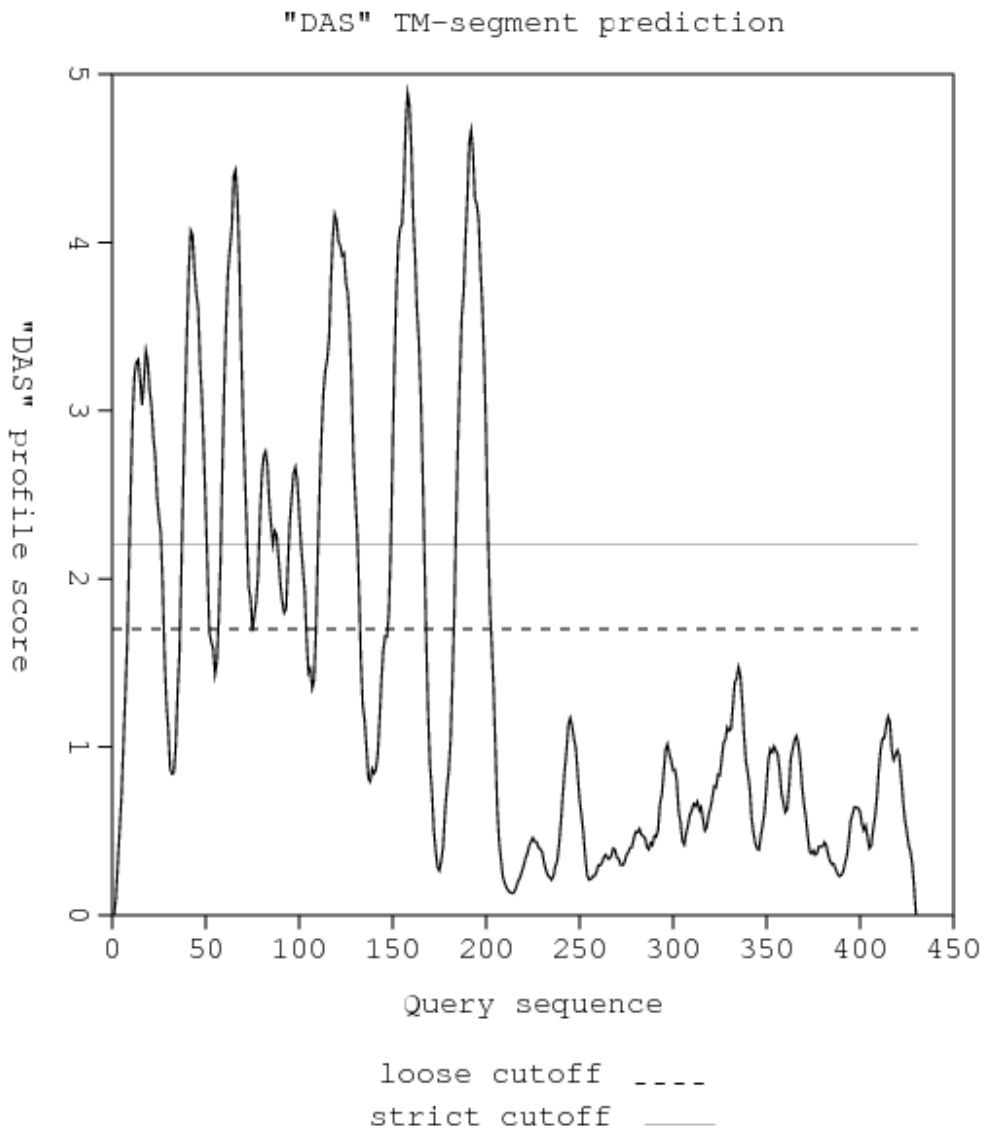
```
        AA
|QYSFIFIGITIFLSILTFVISQFLLKEMKYKRNQEEIETYYEYTLKIEAINNEM
RKFRHD|
        PHD htm |HHHHHHHHHHHHHHHHHHHHHHHH                          |
        Rel htm
|25778888887777788887765305678899999999999999999999999999999|
 detail:
        prH htm
|67889999998888899998887642110000000000000000000000000000000|
        prL htm
|32110000001111100001112357889999999999999999999999999999999|
 subset: SUB htm
|..HHHHHHHHHHHHHHHHHHHH.....LLLLLLLLLLLLLLLLLLLLLLLLLLLLL
LLLLL|
        PHDRhtm |  HHHHHHHHHHHHHHHHHHHH                                |
        PHDThtm |ooTTTTTTTTTTTTTTTTTTiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiii|
                ....,....25...,....26...,....27...,....28...,....29...,....30
        AA
|YVNILTTLSEYIREDDMPGLRDYFNKNIVPMKDNLQMNAIKLNGIENLKV
REIKGLITAK|
        PHD htm |                                                  |
        Rel htm
|99999999999999999999999999999999999999999999999999998899|
 detail:
        prH htm
|00000000000000000000000000000000000000000000000000000000|
        prL htm
|99999999999999999999999999999999999999999999999999999999|
 subset: SUB htm
|LLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLL
LLLLLL|
        PHDRhtm |                                                    |
        PHDThtm |iiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiii|
                ....,....31...,....32...,....33...,....34...,....35...,....36
        AA
|ILRAQEMNIPISIEIPDEVSSINLNMIDLSRSIGIILDNAIEASTEIDDPIIRVAFI
ESE|
        PHD htm |                          HHHHHH                HHHHH|
        Rel htm
|99999999999999999999998875321210112213667778888888875123210|
 detail:
        prH htm
|00000000000000000000000012334345556653111110000000012466655|
        prL htm
|99999999999999999999999987665654443346888899999999987533344|
```

```
 subset: SUB htm
|LLLLLLLLLLLLLLLLLLLLLLLL.................LLLLLLLLLLLL.......|
        PHDRhtm |                                              |
        PHDThtm |iiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiii|
                ....,....37...,....38...,....39...,....40...,....41...,....42
        AA
|NSVTFIVMNKCADDIPRIHELFQESFSTKGEGRGLGLSTLKEIADNADNVL
LDTIIENGF|
        PHD htm |HHHHHH                                        |
        Rel htm
|122331158889999999999999999999999999999999999999999999999999|
 detail:
        prH htm
|566665420000000000000000000000000000000000000000000000000000|
        prL htm
|433334579999999999999999999999999999999999999999999999999999|
 subset: SUB htm
|........LLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLL
LL|
        PHDRhtm |                                              |
        PHDThtm |iiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiii|
                ....,....43...,....44...,....45...,....46...,....47...,....48
        AA      |FIQKVEIINN|
        PHD htm |          |
        Rel htm |9999999999|
 detail:
        prH htm |0000000000|
        prL htm |9999999999|
 subset: SUB htm |LLLLLLLLLL|
        PHDRhtm |          |
        PHDThtm |iiiiiiiiii|
---
--- PhdTopology REFINEMENT AND TOPOLOGY PREDICTION END
---
```

DAS
Potential transmembrane segments

Start	Stop	Length	~	Cutoff
9	27	19	~	1.7
10	26	17	~	2.2
37	51	15	~	1.7
38	50	13	~	2.2
57	74	18	~	1.7
59	72	14	~	2.2
76	103	28	~	1.7
79	88	10	~	2.2
95	101	7	~	2.2
110	132	23	~	1.7
111	131	21	~	2.2
148	167	20	~	1.7
150	166	17	~	2.2
183	202	20	~	1.7
184	201	18	~	2.2

The DAS curve for your query:



The DAS curve of the query is also available in PostScript format.

You may wish to compare your curve with those obtained for the test set. It gives you an idea about the confidence of the prediction.

MEMSAT
THE TOPOLOGY RESULT OF YOUR SEQUENCE

The results that may be important to you are printed at the bottom of your browser,
The first section after the FINAL PREDICTION is the topology of the protein The second is the visual location of the predicted transmembrane helix The third is the fasta sequence of your protein

Any Suggestion On the Web Please Mail to
yzhai@biomail.ucsd.edu

MEMSAT Version 1.5 (ACADEMIC VERSION) - Copyright 1993 D. Jones

COMMERCIAL USE OF THIS PROGRAM IS FORBIDDEN

Newman_agrC1
430 residues read from file.

Helix 1 from 184 (in) to 205 (out) : 10028
+++-

Score = -35.667000
Helix 1 from 184 (out) to 205 (in) : 10136

OOOOXXXXXXXXXXXXXXXXXIIII+++
+++++++++++++++++++++
Score = -20.294000
Helix 1 from 9 (in) to 27 (out) : 3261
Helix 2 from 35 (out) to 51 (in) : 3437
Helix 3 from 58 (in) to 74 (out) : 2511
Helix 4 from 81 (out) to 103 (in) : 2601
Helix 5 from 110 (in) to 132 (out) : 5463
Helix 6 from 149 (out) to 168 (in) : 5833
Helix 7 from 182 (in) to 205 (out) : 5498
Helix 8 from 216 (out) to 232 (in) : -8175

Helix 9 from 241 (in) to 257 (out) : -5566
Helix 10 from 268 (out) to 285 (in) : -6740
Helix 11 from 295 (in) to 313 (out) : -5191
Helix 12 from 321 (out) to 337 (in) : -1626
Helix 13 from 351 (in) to 368 (out) : -2052
Helix 14 from 384 (out) to 400 (in) : -5963
Helix 15 from 407 (in) to 423 (out) : -3174
++++++++IIIIXXXXXXXXXXXXOOOO-------OOOOXXXXXXXXXXIIII++++++IIIIXXXXXXX
OOOOXXXXXXXXXXIIII++++++++IIIIXXXXXXXXXXOOOO----------OOOOXXXXXXXXXX
Score = -100000.000000
Helix 1 from   9 (out) to 27 (in) : 3774
Helix 2 from 35 (in) to 51 (out) : 3682
Helix 3 from 58 (out) to 74 (in) : 3229
Helix 4 from 87 (in) to 103 (out) : 2848
Helix 5 from 110 (out) to 132 (in) : 5667
Helix 6 from 149 (in) to 168 (out) : 5860
Helix 7 from 182 (out) to 205 (in) : 5466
Helix 8 from 236 (in) to 252 (out) : -4560
Helix 9 from 268 (out) to 285 (in) : -6740
Helix 10 from 295 (in) to 313 (out) : -5191
Helix 11 from 321 (out) to 337 (in) : -1626
Helix 12 from 351 (in) to 368 (out) : -2052
Helix 13 from 384 (out) to 400 (in) : -5963
Helix 14 from 407 (in) to 423 (out) : -3174
--------OOOOXXXXXXXXXXXXIIII+++++++IIIIXXXXXXXXXXOOOO------OOOOXXXXXXX
OOOOXXXXXXXXXXXXXXXXXXIIII++++++++++++++++++++++++++++++IIIIXXXXXXX
OOOOXXXXXXXXXXIIII++++++IIIIXXXXXXXXXXOOOO-------
Score = -100000.000000
Helix 1 from   9 (in) to 27 (out) : 3261
Helix 2 from 35 (out) to 51 (in) : 3437
Helix 3 from 58 (in) to 74 (out) : 2511
Helix 4 from 81 (out) to 103 (in) : 2601
Helix 5 from 110 (in) to 132 (out) : 5463
Helix 6 from 149 (out) to 168 (in) : 5833
Helix 7 from 182 (in) to 205 (out) : 5498
Helix 8 from 236 (out) to 252 (in) : -5005
Helix 9 from 295 (in) to 313 (out) : -5191
Helix 10 from 321 (out) to 337 (in) : -1626
Helix 11 from 351 (in) to 368 (out) : -2052
Helix 12 from 384 (out) to 400 (in) : -5963
Helix 13 from 407 (in) to 423 (out) : -3174
++++++++IIIIXXXXXXXXXXXXOOOO-------OOOOXXXXXXXXXXIIII++++++IIIIXXXXXXX
OOOOXXXXXXXXXXIIII++IIIIXXXXX
Score = -100000.000000
Helix 1 from   9 (out) to 27 (in) : 3774

Helix 2 from 35 (in) to 51 (out) : 3682
Helix 3 from 58 (out) to 74 (in) : 3229
Helix 4 from 87 (in) to 103 (out) : 2848
Helix 5 from 110 (out) to 132 (in) : 5667
Helix 6 from 149 (in) to 168 (out) : 5860
Helix 7 from 182 (out) to 205 (in) : 5466
Helix 8 from 295 (in) to 313 (out) : -5191
Helix 9 from 321 (out) to 337 (in) : -1626
Helix 10 from 351 (in) to 368 (out) : -2052
Helix 11 from 384 (out) to 400 (in) : -5963
Helix 12 from 407 (in) to 423 (out) : -3174
--------OOOOXXXXXXXXXXXXIIII+++++++IIIIXXXXXXXXXXOOOO------OOOOXXXXXXX
OOOOXXXXXXXXXXXXXXXXIIII++
OOOOXXXXXXXXXXIIII++++++IIIIXXXXXXXXXXOOOO-------
Score = -100000.000000
Helix 1 from   9 (in) to 27 (out) : 3261
Helix 2 from 35 (out) to 51 (in) : 3437
Helix 3 from 58 (in) to 74 (out) : 2511
Helix 4 from 81 (out) to 103 (in) : 2601
Helix 5 from 110 (in) to 132 (out) : 5463
Helix 6 from 149 (out) to 168 (in) : 5833
Helix 7 from 182 (in) to 205 (out) : 5498
Helix 8 from 236 (out) to 252 (in) : -5005
Helix 9 from 319 (in) to 337 (out) : -1705
Helix 10 from 351 (out) to 368 (in) : -2798
Helix 11 from 406 (in) to 422 (out) : -2831
++++++++IIIIXXXXXXXXXXXXOOOO-------OOOOXXXXXXXXXXIIII++++++IIIIXXXXXXX
OOOOXXXXXXXXXXIIII++
Score = -100000.000000
Helix 1 from   9 (out) to 27 (in) : 3774
Helix 2 from 35 (in) to 51 (out) : 3682
Helix 3 from 58 (out) to 74 (in) : 3229
Helix 4 from 87 (in) to 103 (out) : 2848
Helix 5 from 110 (out) to 132 (in) : 5667
Helix 6 from 149 (in) to 168 (out) : 5860
Helix 7 from 182 (out) to 205 (in) : 5466
Helix 8 from 319 (in) to 337 (out) : -1705
Helix 9 from 351 (out) to 368 (in) : -2798
Helix 10 from 406 (in) to 422 (out) : -2831
--------OOOOXXXXXXXXXXXXIIII+++++++IIIIXXXXXXXXXXOOOO------OOOOXXXXXXX
OOOOXXXXXXXXXXXXXXXXIIII++
OOOOXXXXXXXXXXIIII++++++++++++++++++++++++++++++++++IIIIXXXXXXXX
Score = -100000.000000
Helix 1 from   9 (in) to 30 (out) : 3980
Helix 2 from 56 (out) to 72 (in) : 4350

Helix 3 from 79 (in) to 103 (out) : 3285
Helix 4 from 110 (out) to 132 (in) : 5667
Helix 5 from 149 (in) to 168 (out) : 5860
Helix 6 from 182 (out) to 205 (in) : 5466
Helix 7 from 319 (in) to 337 (out) : -1705
Helix 8 from 351 (out) to 368 (in) : -2798
Helix 9 from 406 (in) to 422 (out) : -2831
++++++++IIIIXXXXXXXXXXXXXXXOOOO----------OOOOXXXXXXXXXIIII++
OOOOXXXXXXXXXXXXXXXXIIII++
OOOOXXXXXXXXXXIIII++++++++++++++++++++++++++++++++++++++IIIIXXXXXXXX
Score = -100000.000000
Helix 1 from   9 (out) to 27 (in) : 3774
Helix 2 from 35 (in) to 51 (out) : 3682
Helix 3 from 58 (out) to 74 (in) : 3229
Helix 4 from 87 (in) to 103 (out) : 2848
Helix 5 from 110 (out) to 132 (in) : 5667
Helix 6 from 149 (in) to 168 (out) : 5860
Helix 7 from 182 (out) to 205 (in) : 5466
Helix 8 from 319 (in) to 337 (out) : -1705
--------OOOOXXXXXXXXXXXXIIII+++++++IIIIXXXXXXXXXXOOOO------OOOOXXXXXXX
OOOOXXXXXXXXXXXXXXXXIIII++
Score = -100000.000000
Helix 1 from   9 (in) to 30 (out) : 3980
Helix 2 from 56 (out) to 72 (in) : 4350
Helix 3 from 79 (in) to 103 (out) : 3285
Helix 4 from 110 (out) to 132 (in) : 5667
Helix 5 from 149 (in) to 168 (out) : 5860
Helix 6 from 182 (out) to 205 (in) : 5466
Helix 7 from 319 (in) to 337 (out) : -1705
++++++++IIIIXXXXXXXXXXXXXXXOOOO----------OOOOXXXXXXXXXIIII++
OOOOXXXXXXXXXXXXXXXXIIII++
Score = -100000.000000
Helix 1 from   9 (out) to 30 (in) : 4166
Helix 2 from 56 (in) to 72 (out) : 4318
Helix 3 from 110 (out) to 132 (in) : 5667
Helix 4 from 149 (in) to 168 (out) : 5860
Helix 5 from 182 (out) to 205 (in) : 5466
Helix 6 from 319 (in) to 337 (out) : -1705
--------OOOOXXXXXXXXXXXXXXXIIII++++++++++++++++++++++++IIIIXXXXXXXXXX
OOOOXXXXXXXXXXXXXXXXIIII++
Score = -100000.000000
Helix 1 from   9 (in) to 30 (out) : 3980
Helix 2 from 56 (out) to 72 (in) : 4350
Helix 3 from 110 (in) to 132 (out) : 5463
Helix 4 from 149 (out) to 168 (in) : 5833

Helix 5 from 182 (in) to 205 (out) : 5498
++++++++IIIIXXXXXXXXXXXXXXXOOOO----------OOOOXXXXXXXXXXIIII++

Score = 16.868000
Helix 1 from   9 (out) to 30 (in) : 4166
Helix 2 from 110 (in) to 132 (out) : 5463
Helix 3 from 149 (out) to 168 (in) : 5833
Helix 4 from 182 (in) to 205 (out) : 5498
--------OOOOXXXXXXXXXXXXXXIIII++

Score = 10.220000
Helix 1 from 110 (in) to 132 (out) : 5463
Helix 2 from 149 (out) to 168 (in) : 5833
Helix 3 from 182 (in) to 205 (out) : 5498
++

Score = 0.360000
Helix 1 from 110 (out) to 132 (in) : 5667
Helix 2 from 182 (in) to 205 (out) : 5498
----------OOOO

Score = -8.163000
Helix 1 from   9 (out) to 27 (in) : 3774
Helix 2 from 35 (in) to 51 (out) : 3682
Helix 3 from 58 (out) to 74 (in) : 3229
Helix 4 from 87 (in) to 103 (out) : 2848
Helix 5 from 110 (out) to 132 (in) : 5667
Helix 6 from 149 (in) to 168 (out) : 5860
Helix 7 from 182 (out) to 205 (in) : 5466
Helix 8 from 217 (in) to 234 (out) : -7586
Helix 9 from 241 (out) to 257 (in) : -6320
Helix 10 from 268 (in) to 285 (out) : -6744
Helix 11 from 295 (out) to 313 (in) : -5146
Helix 12 from 321 (in) to 337 (out) : -1881
Helix 13 from 351 (out) to 368 (in) : -2798
Helix 14 from 384 (in) to 400 (out) : -6639
Helix 15 from 407 (out) to 423 (in) : -4021
--------OOOOXXXXXXXXXXXXIIII+++++++IIIIXXXXXXXXXXOOOO------OOOOXXXXXXX
OOOOXXXXXXXXXXXXXXXXXIIII++++++++++IIIIXXXXXXXXXXOOOO------OOOOX
OOOOXXXXXXXXXIIII+++++++
Score = -100000.000000
Helix 1 from   9 (in) to 27 (out) : 3261
Helix 2 from 35 (out) to 51 (in) : 3437
Helix 3 from 58 (in) to 74 (out) : 2511
Helix 4 from 81 (out) to 103 (in) : 2601

Helix 5 from 110 (in) to 132 (out) : 5463
Helix 6 from 149 (out) to 168 (in) : 5833
Helix 7 from 182 (in) to 205 (out) : 5498
Helix 8 from 216 (out) to 232 (in) : -8175
Helix 9 from 241 (in) to 257 (out) : -5566
Helix 10 from 268 (out) to 285 (in) : -6740
Helix 11 from 295 (in) to 313 (out) : -5191
Helix 12 from 321 (out) to 337 (in) : -1626
Helix 13 from 351 (in) to 368 (out) : -2052
Helix 14 from 406 (out) to 422 (in) : -3101
++++++++IIIIXXXXXXXXXXXXOOOO-------OOOOXXXXXXXXXXXIIII++++++IIIIXXXXXXX
OOOOXXXXXXXXXXIIII++++++++IIIIXXXXXXXXXXOOOO----------OOOOXXXXXXXXXX
Score = -100000.000000
Helix 1 from 9 (out) to 27 (in) : 3774
Helix 2 from 35 (in) to 51 (out) : 3682
Helix 3 from 58 (out) to 74 (in) : 3229
Helix 4 from 87 (in) to 103 (out) : 2848
Helix 5 from 110 (out) to 132 (in) : 5667
Helix 6 from 149 (in) to 168 (out) : 5860
Helix 7 from 182 (out) to 205 (in) : 5466
Helix 8 from 236 (in) to 252 (out) : -4560
Helix 9 from 268 (out) to 285 (in) : -6740
Helix 10 from 295 (in) to 313 (out) : -5191
Helix 11 from 321 (out) to 337 (in) : -1626
Helix 12 from 351 (in) to 368 (out) : -2052
Helix 13 from 406 (out) to 422 (in) : -3101
--------OOOOXXXXXXXXXXXIIII+++++++IIIIXXXXXXXXXXOOOO------OOOOXXXXXXX
OOOOXXXXXXXXXXXXXXXXIIII++++++++++++++++++++++++++++++IIIIXXXXXXX
OOOOXXXXXXXXIIII++++++++
Score = -100000.000000
Helix 1 from 9 (in) to 27 (out) : 3261
Helix 2 from 35 (out) to 51 (in) : 3437
Helix 3 from 58 (in) to 74 (out) : 2511
Helix 4 from 81 (out) to 103 (in) : 2601
Helix 5 from 110 (in) to 132 (out) : 5463
Helix 6 from 149 (out) to 168 (in) : 5833
Helix 7 from 182 (in) to 205 (out) : 5498
Helix 8 from 236 (out) to 252 (in) : -5005
Helix 9 from 295 (in) to 313 (out) : -5191
Helix 10 from 321 (out) to 337 (in) : -1626
Helix 11 from 351 (in) to 368 (out) : -2052
Helix 12 from 406 (out) to 422 (in) : -3101
++++++++IIIIXXXXXXXXXXXXOOOO-------OOOOXXXXXXXXXXXIIII++++++IIIIXXXXXXX
OOOOXXXXXXXXXXIIII++IIIIXXXXX
Score = -100000.000000

Helix 1 from   9 (out) to 27 (in) : 3774
Helix 2 from 35 (in) to 51 (out) : 3682
Helix 3 from 58 (out) to 74 (in) : 3229
Helix 4 from 87 (in) to 103 (out) : 2848
Helix 5 from 110 (out) to 132 (in) : 5667
Helix 6 from 149 (in) to 168 (out) : 5860
Helix 7 from 182 (out) to 205 (in) : 5466
Helix 8 from 295 (in) to 313 (out) : -5191
Helix 9 from 321 (out) to 337 (in) : -1626
Helix 10 from 351 (in) to 368 (out) : -2052
Helix 11 from 406 (out) to 422 (in) : -3101
--------OOOOXXXXXXXXXXXXIIII+++++++IIIIXXXXXXXXXXOOOO------OOOOXXXXXXX
OOOOXXXXXXXXXXXXXXXXIIII++
OOOOXXXXXXXXXIIII++++++++
Score = -100000.000000
Helix 1 from   9 (in) to 27 (out) : 3261
Helix 2 from 35 (out) to 51 (in) : 3437
Helix 3 from 58 (in) to 74 (out) : 2511
Helix 4 from 81 (out) to 103 (in) : 2601
Helix 5 from 110 (in) to 132 (out) : 5463
Helix 6 from 149 (out) to 168 (in) : 5833
Helix 7 from 182 (in) to 205 (out) : 5498
Helix 8 from 236 (out) to 252 (in) : -5005
Helix 9 from 319 (in) to 337 (out) : -1705
Helix 10 from 351 (out) to 368 (in) : -2798
++++++++IIIIXXXXXXXXXXXXOOOO-------OOOOXXXXXXXXXIIII++++++IIIIXXXXXXX
OOOOXXXXXXXXXIIII++
Score = -100000.000000
Helix 1 from   9 (out) to 27 (in) : 3774
Helix 2 from 35 (in) to 51 (out) : 3682
Helix 3 from 58 (out) to 74 (in) : 3229
Helix 4 from 87 (in) to 103 (out) : 2848
Helix 5 from 110 (out) to 132 (in) : 5667
Helix 6 from 149 (in) to 168 (out) : 5860
Helix 7 from 182 (out) to 205 (in) : 5466
Helix 8 from 319 (in) to 337 (out) : -1705
Helix 9 from 351 (out) to 368 (in) : -2798
--------OOOOXXXXXXXXXXXXIIII+++++++IIIIXXXXXXXXXXOOOO------OOOOXXXXXXX
OOOOXXXXXXXXXXXXXXXXIIII++
OOOOXXXXXXXXXXXIIII++
Score = -100000.000000
Helix 1 from   9 (in) to 30 (out) : 3980
Helix 2 from 56 (out) to 72 (in) : 4350
Helix 3 from 79 (in) to 103 (out) : 3285
Helix 4 from 110 (out) to 132 (in) : 5667

Helix 5 from 149 (in) to 168 (out) : 5860
Helix 6 from 182 (out) to 205 (in) : 5466
Helix 7 from 319 (in) to 337 (out) : -1705
Helix 8 from 351 (out) to 368 (in) : -2798
++++++++IIIIXXXXXXXXXXXXXXXOOOO----------OOOOXXXXXXXXXXIIII++
OOOOXXXXXXXXXXXXXXXXIIII++
OOOOXXXXXXXXXXIIII++
Score = -100000.000000
Helix 1 from 9 (out) to 27 (in) : 3774
Helix 2 from 35 (in) to 51 (out) : 3682
Helix 3 from 58 (out) to 74 (in) : 3229
Helix 4 from 87 (in) to 103 (out) : 2848
Helix 5 from 110 (out) to 132 (in) : 5667
Helix 6 from 149 (in) to 168 (out) : 5860
Helix 7 from 182 (out) to 205 (in) : 5466
--------OOOOXXXXXXXXXXXXIIII+++++++IIIIXXXXXXXXXXOOOO------OOOOXXXXXXX
OOOOXXXXXXXXXXXXXXXXIIII++
++++++++++++++++++++
Score = 36.574000
Helix 1 from 9 (in) to 30 (out) : 3980
Helix 2 from 56 (out) to 72 (in) : 4350
Helix 3 from 79 (in) to 103 (out) : 3285
Helix 4 from 110 (out) to 132 (in) : 5667
Helix 5 from 149 (in) to 168 (out) : 5860
Helix 6 from 182 (out) to 205 (in) : 5466
++++++++IIIIXXXXXXXXXXXXXXXOOOO----------OOOOXXXXXXXXXXIIII++
OOOOXXXXXXXXXXXXXXXXIIII++
++++++++++++++++++++
Score = 31.547000
Helix 1 from 9 (out) to 30 (in) : 4166
Helix 2 from 56 (in) to 72 (out) : 4318
Helix 3 from 110 (out) to 132 (in) : 5667
Helix 4 from 149 (in) to 168 (out) : 5860
Helix 5 from 182 (out) to 205 (in) : 5466
--------OOOOXXXXXXXXXXXXXXXIIII++++++++++++++++++++++++IIIIXXXXXXXXXX
OOOOXXXXXXXXXXXXXXXXIIII++
++++++++++++++++++++
Score = 25.369000
Helix 1 from 9 (in) to 30 (out) : 3980
Helix 2 from 110 (out) to 132 (in) : 5667
Helix 3 from 149 (in) to 168 (out) : 5860
Helix 4 from 182 (out) to 205 (in) : 5466
++++++++IIIIXXXXXXXXXXXXXXXOOOO----------
OOOOXXXXXXXXXXXXXXXXIIII++
++++++++++++++++++++

Score = 15.639000
Helix 1 from 110 (out) to 132 (in) : 5667
Helix 2 from 149 (in) to 168 (out) : 5860
Helix 3 from 182 (out) to 205 (in) : 5466
----------OOOO
OOOOXXXXXXXXXXXXXXXXIIII++
++++++++++++++++++++
Score = 8.526000
Helix 1 from 110 (in) to 132 (out) : 5463
Helix 2 from 149 (out) to 168 (in) : 5833
++
OOOOXXXXXXXXXXXXIIII++
++
Score = -3.605000
 1 helices (+) : Score = -35.667
 1 helices (-) : Score = -20.294
 2 helices (+) : Score = -3.605
 2 helices (-) : Score = -8.163
 3 helices (+) : Score = 0.36
 3 helices (-) : Score = 8.526
 4 helices (+) : Score = 15.639
 4 helices (-) : Score = 10.22
 5 helices (+) : Score = 16.868
 5 helices (-) : Score = 25.369
 6 helices (+) : Score = 31.547
 6 helices (-) : Score = -100000
 7 helices (+) : Score = -100000
 7 helices (-) : Score = 36.574
 8 helices (+) : Score = -100000
 8 helices (-) : Score = -100000
 9 helices (+) : Score = -100000
 9 helices (-) : Score = -100000
10 helices (+) : Score = -100000
10 helices (-) : Score = -100000
11 helices (+) : Score = -100000
11 helices (-) : Score = -100000
12 helices (+) : Score = -100000
12 helices (-) : Score = -100000
13 helices (+) : Score = -100000
13 helices (-) : Score = -100000
14 helices (+) : Score = -100000
14 helices (-) : Score = -100000
15 helices (+) : Score = -100000
15 helices (-) : Score = -100000

FINAL PREDICTION
==========
1: (out) 9-27 (3.77)
2: 35-51 (3.68)
3: 58-74 (3.23)
4: 87-103 (2.85)
5: 110-132 (5.67)
6: 149-168 (5.86)
7: 182-205 (5.47)

--------OOOOXXXXXXXXXXXXIIII+++++++IIIIXXXXXXXXXXOOOO------OOOOXXXXXXX
OOOOXXXXXXXXXXXXXXXXIIII+++
++++++++++++++++++++
MELLNSYNFVLFVLTQMILMFTIPAIISGIKYSKLDYFFIIVISTLSLFLFKMFDSASLIILTSFI
YTLKIEAINNEMRKFRHDYVNILTTLSEYIREDDMPGLRDYFNKNIVPMKDNLQMNAIKL

HMMTOP 2.0

Protein: Newman_agrC1
Length: 430
N-terminus: OUT
Number of transmembrane helices: 6
Transmembrane helices: 7-30 39-62 79-103 110-133 150-169 182-205

Total entropy of the model: 17.0116
Entropy of the best path: 17.0140

The best path:

seq MELLNSYNFV LFVLTQMILM FTIPAIISGI KYSKLDYFFI IVISTLSLFL 50
pred OOooooHHHH HHHHHHHHHH HHHHHHHHHH iiiiiiiiHH HHHHHHHHHH

seq FKMFDSASLI ILTSFIIIMY FVKIKWYSIL LIMTSQIILY CANYMYIVIY 100
pred HHHHHHHHHH HHoooooooo oooooooHH HHHHHHHHHH HHHHHHHHHH

seq AYITKISDSI FVIFPSFFVV YVTISILFSY IINRVLKKIS TPYLILNKGF 150
pred HHHiiiiiiH HHHHHHHHHH HHHHHHHHHH HHHooooooo ooooooooH

seq LIVISTILLL TFSLFFFYSQ INSDEAKVIR QYSFIFIGIT IFLSILTFVI 200
pred HHHHHHHHHH HHHHHHHHHi iiiiiiiiii iHHHHHHHHH HHHHHHHHHH

seq SQFLLKEMKY KRNQEEIETY YEYTLKIEAI NNEMRKFRHD YVNILTTLSE 250
pred HHHHHooooo ooooooooo OOOOOOOOOO OOOOOOOOOO OOOOOOOOOO

seq YIREDDMPGL RDYFNKNIVP MKDNLQMNAI KLNGIENLKV REIKGLITAK 300
pred OOOOOOOOOO OOOOOOOOOO OOOOOOOOOO OOOOOOOOOO OOOOOOOOOO

seq ILRAQEMNIP ISIEIPDEVS SINLNMIDLS RSIGIILDNA IEASTEIDDP 350

pred OOOOOOOOOO OOOOOOOOOO OOOOOOOOOO OOOOOOOOOO
OOOOOOOOOO

seq IIRVAFIESE NSVTFIVMNK CADDIPRIHE LFQESFSTKG
EGRGLGLSTL 400
pred OOOOOOOOOO OOOOOOOOOO OOOOOOOOOO OOOOOOOOOO
OOOOOOOOOO

seq KEIADNADNV LLDTIIENGF FIQKVEIINN 430
pred OOOOOOOOOO OOOOOOOOOO OOOOOOOOOO

Appendix B

>1
GAAAATGCAGGATTTAACGTGAAGATGAATATCCAACCAGATTATGGT
AGCTATCGTACACAACGTCAAGCCGGCAATTATGATATCCAAATTGAT
GACTGGATGACAGTGTTTGGTGACCCGAACTATGCTATGACGGCATTA
TTTAGTTCTACAGGATCAAATAGTTTATTGAAAGATAAACATGTAGAC
CAGTTGTTAAATAAAGCTTCTACTCAAAATGAAGCAGATGTTAAACAA
ACATATAAGCAAATTGAAGATGAAGTTGTATTTGATAAAGGGTATATG
GCGCCTTTATATGGATCAAAAAAGAATTTAGTATATGACAATAAAGTG
TTAGATAAAAATAGTGTTGGATTGCCAAATTCACGTGCATTAATATGG
CAACAATTTGATTACAACAATAGTAGAGAACGAGATACGCGGCCACTT
GTGATGACACAACAAGATGGTGAAATTCCTACATTGGATCCAATACGT
T

>39
GAAAATGCAGGATTTAACGTGAAGATGAATATCCAACCAGATTATGGT
AGCTATCGTACACAACGTCAAGCCGGCAATTATGATATCCAAATTGAT
GACTGGATGACAGTGTTTGGTGACCCGAACTATGCTATGACGGCATTA
TTTAGTTCTACAGGATCAAATAGTTTATTGAAAGATAAACATGTAGAC
CAGTTGTTAAATAAAGCTTCTACTCAAAATGAAGCAGATGTTAAACAA
ACATATAAGCAAATTGAAGATGAAGTTGTATTTGATAAAGGGTATATG
GCGCCTTTATATGGATCAAAAAAGAATTTAGTATATGACAATAAAGTG
TTAGATAAAAATAGTGTTGGATTGCCAAATTCACGTGCATTAATATGG
CAACAATTTGATTACAACAATAGTAGAGAACGAGATACGCGGCCACTT
GTGATGACACAACAAGATGGTGAAATTCCTACATTGGATCCAATACGT
T

>3
AAAGTGTCTAATCGTTTAGCTGAGAAATATGATTTAGACATGAAACAT
ACTTTCCTAGGGCCTCATGCTGTACCTAAAGAGGCAAGTTCAAATGAG
GCATTTTTAGAAGAAATGATTGCGTTACTTCCGGAAGTAAAACAATAT
GCAGACTTTGCGGATATTTTCTGTGAAACAGGTGTATTTACAATAGAA
CAATCGCAACATTATATGCAAAAAGCCAAAGAAGCAGGTTTTAAAGTG
AAAATACATGCGGATGAAATTGATCCGTTAGGCGGACTGGAATTAG

>68
AAAGTGTCTAATCGTTTAGCTGAGAAATATGATTTAGACATGAAACAT
ACTTTCCTAGGGCCTCATGCTGTACCTAAAGAGGCAAGTTCAAATGAG
GCATTTTTAGAAGAAATGATTGCGTTACTTCCGGAAGTAAAACAATAT
GCAGACTTTGCGGATATTTTCTGTGAAACAGGTGTATTTACAATAGAA
CAATCGCAACATTATATGCAAAAAGCCAAAGAAGCAGGTTTTAAAGTG
AAAATACATGCGGATGAAATTGATCCGTTAGGCGGACTGGAATTAG

>5
TCTTACTCAATGAGCGCTCAAGATAAAATGAATACTGTAGATGCATAT
GAAAATCTATTATTTGATTGTCTTAAAGGTGATGCCACTAACTTCACGC
ACTGGGAAGAATTAAAATCAACATGGAAATTTGTTGATGCAATTCAAG
ATGAATGGAATATGGTTGATCCAGAATTCCCTAACTATGAATCAGGTA
CTAATGGTCCATTAGAAAGTGATTTACTACTTGCTCGTGATGGTAACCA
TTGGTGGGACGATATTCAATAATTGAATTAAAACGCACATGTTAAACA
AAAATAAATGAGCGAATGACAAGGATTGATTACGAATCAAGTAAATC
AATTGATTTAATTGATGAAGCAAGTTCTAAAGTAAGACTTAAGAGTCA
TACGACACCTAATAATTTAAAAGAAATTGAACAAGAAATTGAAAAAGT
TAAAAATGAAAAAGATGCCGCAGTACATGCTCAAGAGTTTGAAAATGC
TGCTAACCTGCGTGATAAACAAACAAAACTTGAAAAGCAATATGAAG
AAGCTAAAAATGAATGGAAGAATGCACAAAATGGCATGTCAACTTCAT
TGTCAGAAGAAGATATTGCTGAAGTTATTGCAGGATGGACAGGTATCC
CATTAACTAAAATCAATGAAACAGAATCTGAAAAACTTCTTAGTCTAG
AAGATACATTACATGAGAGAGTTATTGGGGCAAAAAGATGCTGTTAAT
TCAATCAGTAAAGCGGTTAGACGTGCCCGTGCAGGGGTTAAAAGATCC
TAAACGACCAATTCCCGGGGA

>49
TCTTACTCAATGAGCGCTCAAGATAAAATGAATACTGTAGATGCATAT
GAAAATCTATTATTTGATTGTCTTAAAGGTGATGCCACTAACTTCACGC
ACTGGGAAGAATTAAAATCAACATGGAAATTTGTTGATGCAATTCAAG
ATGAATGGAATATGGTTGATCCAGAATTCCCTAACTATGAATCAGGTA
CTAATGGTCCATTAGAAAGTGATTTACTACTTGCTCGTGATGGTAACCA
TTGGTGGGACGATATTCAATAATTGAATTAAAACGCACATGTTAAACA
AAAATAAATGAGCGAATGACAAGGATTGATTACGAATCAAGTAAATC
AATTGATTTAATTGATGAAGCAAGTTCTAAAGTAAGACTTAAGAGTCA
TACGACACCTAATAATTTAAAAGAAATTGAACAAGAAATTGAAAAAGT
TAAAAATGAAAAAGATGCCGCAGTACATGCTCAAGAGTTTGAAAATGC
TGCTAACCTGCGTGATAAACAAACAAAACTTGAAAAGCAATATGAAG
AAGCTAAAAATGAATGGAAGAATGCACAAAATGGCATGTCAACTTCAT
TGTCAGAAGAAGATATTGCTGAAGTTATTGCAGGATGGACAGGTATCC
CATTAACTAAAATCAATGAAACAGAATCTGAAAAACTTCTTAGTCTAG
AAGATACATTACATGAGAGAGTTATTGGGGCAAAAAGATGCTGTTAAT
TCAATCAGTAAAGCGGTTAGACGTGCCCGTGCAGGGGTTAAAAGATCC
TAAACGACCAATTCCCGGGGA

>9
TCTTACTCAATGAGCGCTCAAGATAAAATGAATACTGTAGATGCATAT
GAAAATCTATTATTTGATTGTCTTAAAGGTGATGCCACTAACTTCACGC
ACTGGGAAGAATTAAAATCAACATGGAAATTTGTTGATGCAATTCAAG
ATGAATGGAATATGGTTGATCCAGAATTCCCTAACTATGAATCAGGTA
CTAATGGTCCATTAGAAAGTGATTTACTACTTGCTCGTGATGGTAACCA
TTGGTGGGACGATATTCAATAATTGAATTAAAACGCACATGTTAAACA
AAAATAAATGAGCGAATGACAAGGATTGATTACGAATCAAGTAAATC
AATTGCAGTACCAAGCCCCATTGCTGTCGCACAGGCAAATTTTACCCA
TAACTGTGGTTCAACACTGCCATCATTCTGTACATTAGCGACAATCAAT
GCCAACGTAATAATACCCATTGATTTTTGCGCATCATTCGTACCGTGAG
AGAATGATTGTAACGCTGCTGTGAAAATTTGGAAAAATCTAAAGTTAC
GATTCGCTCTTGTTAAATTTGCATTTTTAAAGATAACTTTAAAAATTGA
ATACATCAAGAAACCAACACAAAATGCGATAATCGGTGAAACGATTA
ATACAATAATAATTTTTGTGAAACCTTGGTAATGTAACACTCCAAATG
AGCCTTCAGATGCGATTGCTGCACCCG

>63
TCTTACTCAATGAGCGCTCAAGATAAAATGAATACTGTAGATGCATAT
GAAAATCTATTATTTGATTGTCTTAAAGGTGATGCCACTAACTTCACGC
ACTGGGAAGAATTAAAATCAACATGGAAATTTGTTGATGCAATTCAAG
ATGAATGGAATATGGTTGATCCAGAATTCCCTAACTATGAATCAGGTA
CTAATGGTCCATTAGAAAGTGATTTACTACTTGCTCGTGATGGTAACCA
TTGGTGGGACGATATTCAATAATTGAATTAAAACGCACATGTTAAACA
AAAATAAATGAGCGAATGACAAGGATTGATTACGAATCAAGTAAATC
AATTGCAGTACCAAGCCCCATTGCTGTCGCACAGGCAAATTTTACCCA
TAACTGTGGTTCAACACTGCCATCATTCTGTACATTAGCGACAATCAAT
GCCAACGTAATAATACCCATTGATTTTTGCGCATCATTCGTACCGTGAG
AGAATGATTGTAACGCTGCTGTGAAAATTTGGAAAAATCTAAAGTTAC
GATTCGCTCTTGTTAAATTTGCATTTTTAAAGATAACTTTAAAAATTGA
ATACATCAAGAAACCAACACAAAATGCGATAATCGGTGAAACGATTA
ATACAATAATAATTTTTGTGAAACCTTGGTAATGTAACACTCCAAATG
AGCCTTCAGATGCGATTGCTGCACCCG

>20
ATGGAGCAAAATTTAACAATTGAAACAAAGTCAAATCCGAATGACCTT
GTTACAAATGTAGATAAAGCAACAGAAGATTTCATTTTTGATACAATT
TTAGAAACATATCCCAATCATCAAGTATTAGGTGAAGAAGGGCATGGT
CATGACATCGATACTTCCAAAGGTACGGTATGGATTGTTGACCCAATA
GACGGTACATTGAATTTTGTTCATCAACAAGAAAATTTCGCAATTT

>46
ATGGAGCAAAATTTAACAATTGAAACAAAGTCAAATCCGAATGACCTT
GTTACAAATGTAGATAAAGCAACAGAAGATTTCATTTTTGATACAATT
TTAGAAACATATCCCAATCATCAAGTATTAGGTGAAGAAGGGCATGGT
CATGACATCGATACTTCCAAAGGTACGGTATGGATTGTTGACCCAATA
GACGGTACATTGAATTTTGTTCATCAACAAGAAAATTTCGCAATTT

>31
TCTTACTCAATGAGCGCTCAAGATAAAATGAATACTGTAGATGCATAT
GAAAATCTATTATTTGATTGTCTTAAAGGTGATGCCACTAACTTCACGC
ACTGGGAAGAATTAAAATCAACATGGAAATTTGTTGATGCAATTCAAG
ATGAATGGAATATGGTTGATCCAGAATTCCCTAACTATGAATCAGGTA
CTAATGGTCCATTAGAAAGTGATTTACTACTTGCTCGTGATGGTAACCA
TTGGTGGGACGATATTCAATAATTGAATTAAAACGCACATGTTAAACA
AAAATAAATGAGCGAATGACAAGGATTGATTACGAATCAAGTAAATC
AATTGTTCTATCATTCGCTTTTTTTCTAGACTATATTATGAAATTATATT
TTACAATGCCCAAAACTATTTTAATAATCATTGAACAAATGGGTGTAT
AATTTATAGAAATAATGTAGAATAAAAATAAATGATTGAATTAATTGG
AGTGAAAGTTTTGGACGTTATCAAGCAAATACAACAGG

>53
TCTTACTCAATGAGCGCTCAAGATAAAATGAATACTGTAGATGCATAT
GAAAATCTATTATTTGATTGTCTTAAAGGTGATGCCACTAACTTCACGC
ACTGGGAAGAATTAAAATCAACATGGAAATTTGTTGATGCAATTCAAG
ATGAATGGAATATGGTTGATCCAGAATTCCCTAACTATGAATCAGGTA
CTAATGGTCCATTAGAAAGTGATTTACTACTTGCTCGTGATGGTAACCA
TTGGTGGGACGATATTCAATAATTGAATTAAAACGCACATGTTAAACA
AAAATAAATGAGCGAATGACAAGGATTGATTACGAATCAAGTAAATC
AATTGTACACGTCCAAACTTTTTTATAGCAATTATTTGCTTATAAGTTA
ATAATTGAACTTTTTAAAATATTGTCTCATACTTTAAAAATCCAGTTTC
TCACTCATTACCTTCATCCATCTCGAACAAATATTCTCACACCATTATC

AATAAGGTTGACTTAGTTGAGAATACAAATCTTCTTTTTTTACATTTTTT
ACATTAAAATCAAGATATTAAATATAGACACAACCTTTTAAATCATCA
ACTACAAAATCCATTATCAAAGCACTTTATATTATCTTTCTTTATTTTAA
AAGTAATCATTCACTTTACAAAGTTATCCAGTTGACTATTCTGGATCAA
AATGTAATAATGCTAACTAAGAGATTAGAATTATTATAATTAAGCGCA
AT

>60
TCTTACTCAATGAGCGCTCAAGATAAAATGAATACTGTAGATGCATAT
GAAAATCTATTATTTGATTGTCTTAAAGGTGATGCCACTAACTTCACGC
ACTGGGAAGAATTAAAATCAACATGGAAATTTGTTGATGCAATTCAAG
ATGAATGGAATATGGTTGATCCAGAATTCCCTAACTATGAATCAGGTA
CTAATGGTCCATTAGAAAGTGATTTACTACTTGCTCGTGATGGTAACCA
TTGGTGGGACGATATTCAATAATTGAATTAAAACGCACATGTTAAACA
AAAATAAATGAGCGAATGACAAGGATTGATTACGAATCAAGTAAATC
AATTGTTCTATCATTCGCTTTTTTTCTAGACTATATTATGAAATTATATT
TTACAATGCCCAAAACTATTTTAATAATCATTGAACAAATGGGTGTAT
AATTTATAGAAATAATGTAGAATAAAAATAAATGATTGAATTAATTGG
AGTGAAAGTTTTGGACGTTATCAAGCAAATACAACAGG

>69
TCTTACTCAATGAGCGCTCAAGATAAAATGAATACTGTAGATGCATAT
GAAAATCTATTATTTGATTGTCTTAAAGGTGATGCCACTAACTTCACGC
ACTGGGAAGAATTAAAATCAACATGGAAATTTGTTGATGCAATTCAAG
ATGAATGGAATATGGTTGATCCAGAATTCCCTAACTATGAATCAGGTA
CTAATGGTCCATTAGAAAGTGATTTACTACTTGCTCGTGATGGTAACCA
TTGGTGGGACGATATTCAATAATTGAATTAAAACGCACATGTTAAACA
AAAATAAATGAGCGAATGACAAGGATTGATTACGAATCAAGTAAATC
AATTGTTCTATCATTCGCTTTTTTTCTAGACTATATTATGAAATTATATT
TTACAATGCCCAAAACTATTTTAATAATCATTGAACAAATGGGTGTAT
AATTTATAGAAATAATGTAGAATAAAAATAAATGATTGAATTAATTGG
AGTGAAAGTTTTGGACGTTATCAAGCAAATACAACAGG

>61
AGCAAACAACCGAATCAAATTAATGATTGGGGAACATTTGATCATACT
AAATTTGATGTCGATGATCATGTTACTAGTTATATGACATTTGCCAATC
GAGCAAGCATGCAGTTTGAATGTTCGTGGTCTGCAAATATCAAAGAAG
ATAAGGTTCACGTTAGTTTATCAGGAGAAGATGGCGGTATCAATTTAT
TTCCATTTGAAATATATGAGCCCCGCTTTGGAACTATTTTTGAAAGCAA
AGCTAATGTTGAGCATAACGAAGACATTGCTGGTGAGAGACAGGCGC
GTAACTTTGTCAATGCGTGTTTAGGGATAGAAGAGATTGTGGTGAAAC
CGGAAGAAGCACGCAATGTAAATGCCCTTATAGAAGCGATTTATCGTA
GCGATCTTGATAACAAGAGCATACAACTTTAATGATTATCATATATGA
TACAAAATTCTCAATATAAAAAGAAGGAGTGCTTTTCAATGAAAATAG
GTGTATTTTCAGTATTATTTTACGATAAAAATTTTGAAGATATGTTAGA
TTATGTCTCAGAATCTGGATTGGATATGATTGAAGTTGGAACAGGTGG

TAACCCAGGAGATAAATTTTGTAAGTTAGATGAGTTGTTAGAAAATGA
AGACAAGCGCCAAGCATTTATGAAGTCAATCACAGACAGAGGCTTACA
AATAAGTGGTTTCAGTTGTCATAACAATCCAATTTCTCCAGATCCGATA
GAAGCGAAAGAAGCCGATGAAACGTTACGTAAAACAATCCGTTTAGC
AAATCTATTAGACGTGCCAGTTGTTAATACATTTTCTGGCATTGCAGGA
TCAGATGATACCGCTAAAAAGCCTAATTGGCCTGTTACACCTTGGCCA
ACAGCCTACTCTGAAATTTATGATTATCAGTGGAATGAAAAGTTGATA
CCATATTGGCAAGATTTAGCTGAGTTTGCAAAAGAGCAAGATGTAAAA
ATTGCCATAGAGTTGCATGCAGGATTTTTAGTGCATACACCATATACA
ATGTTGAAGTTACGTGAGGCTACAAATGAATATATCGGTGCTAACTTA
GATCCTAGTCATCTATGGTGGCAAGGTATTGACC

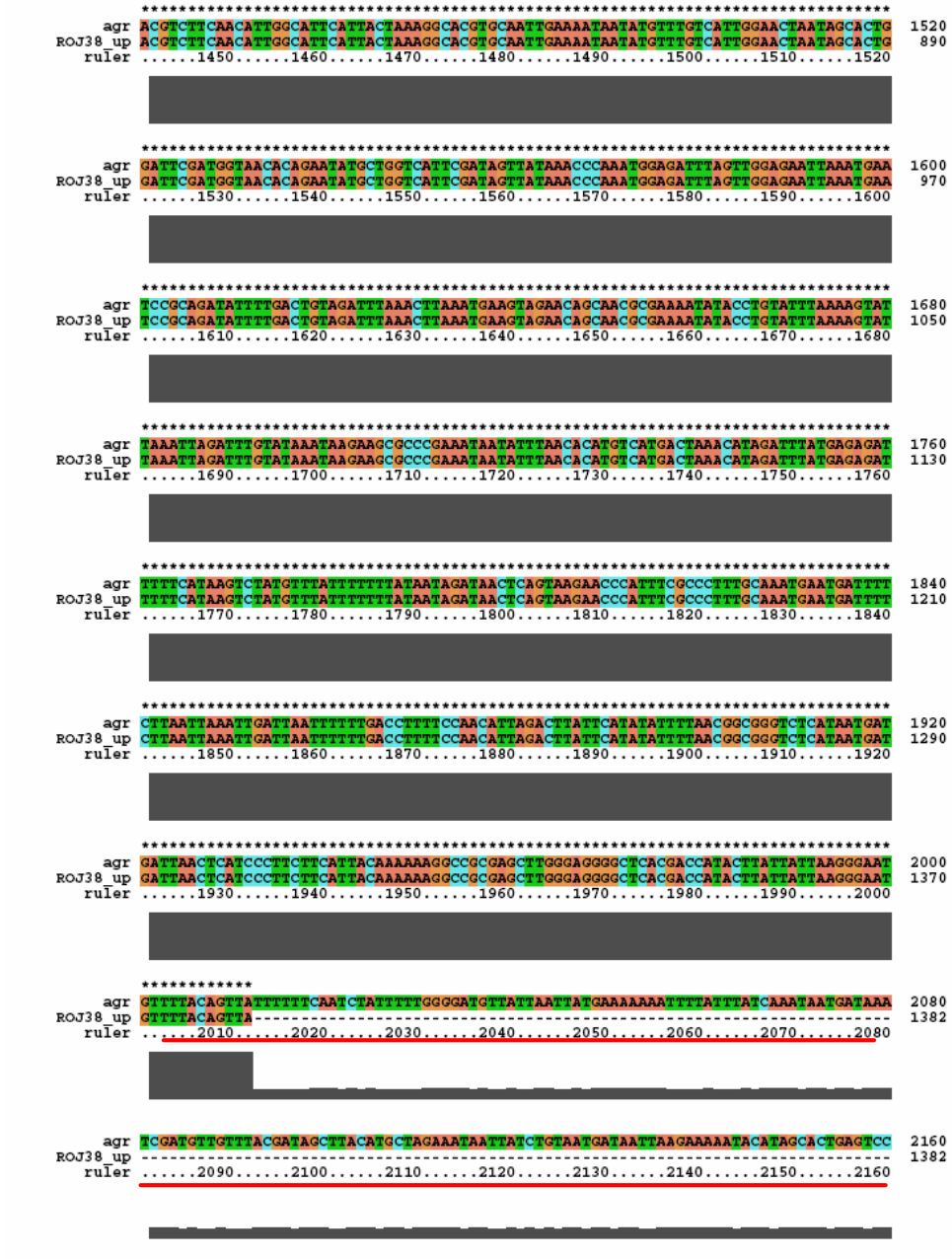
Appendix C

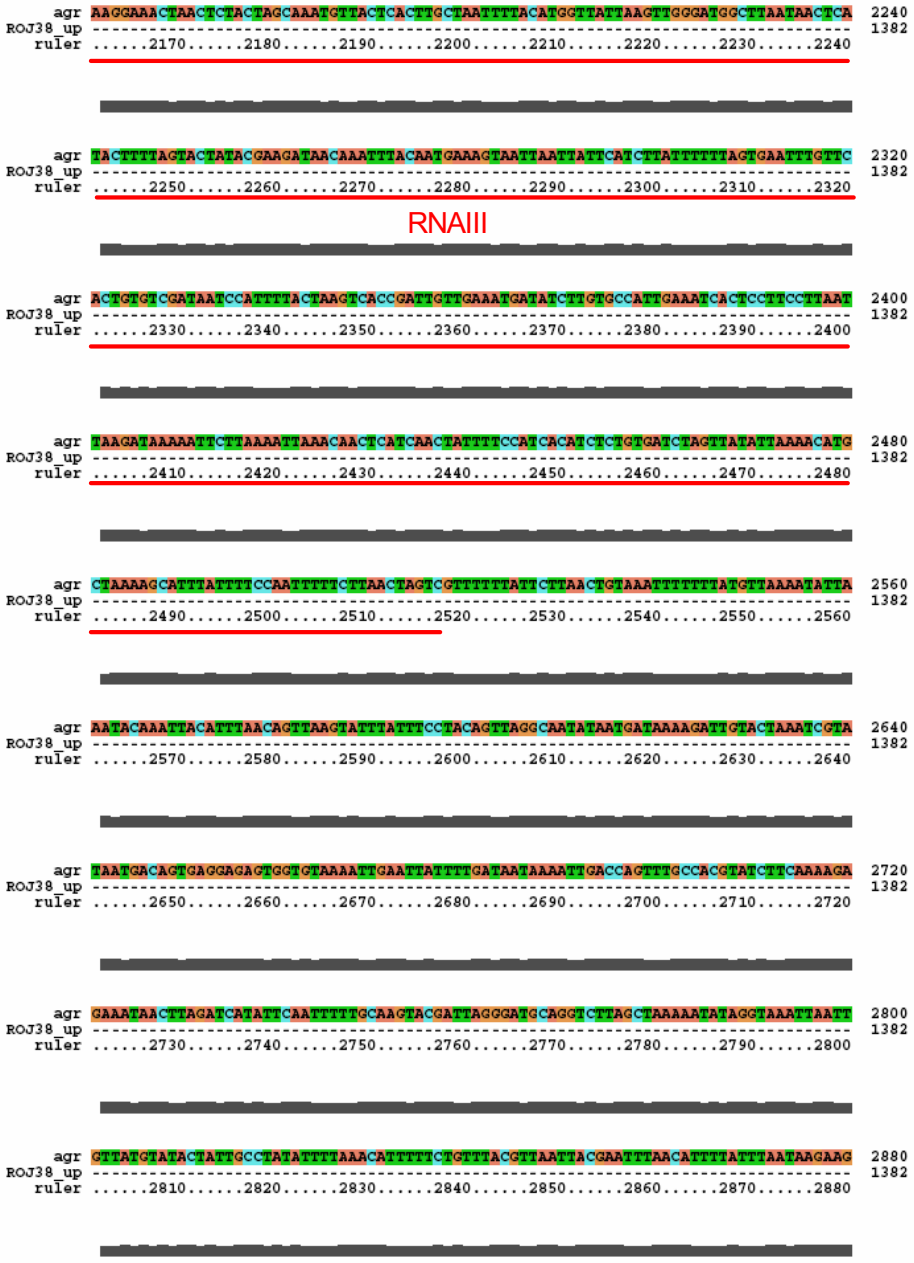
CLUSTAL X (1.81) MULTIPLE SEQUENCE ALIGNMENT

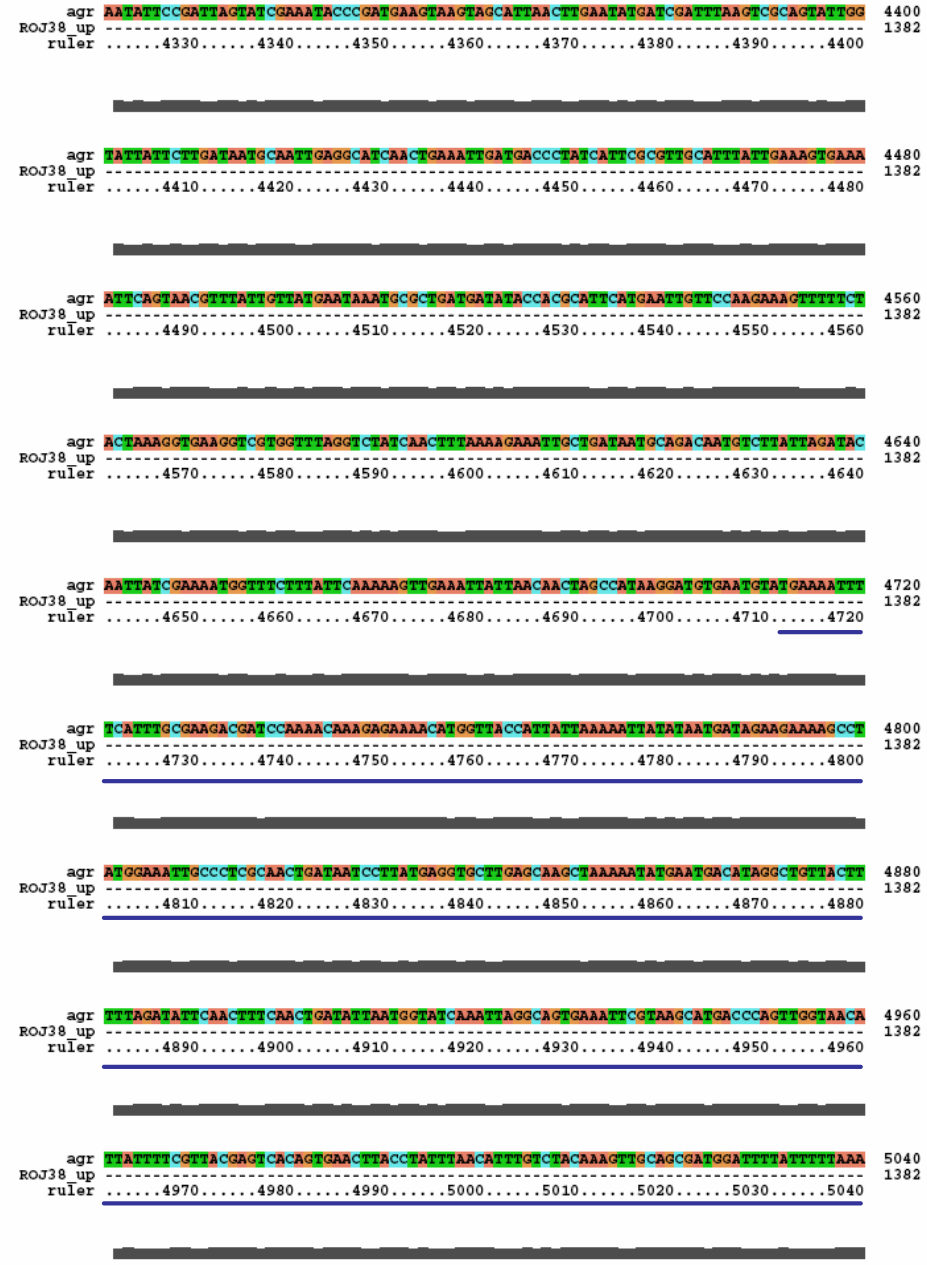
File: C:Documents and SettingsRasmus JensenDate: Sat Sep 27 11:14:13 2008quencessequencingl
Page 1 of 10

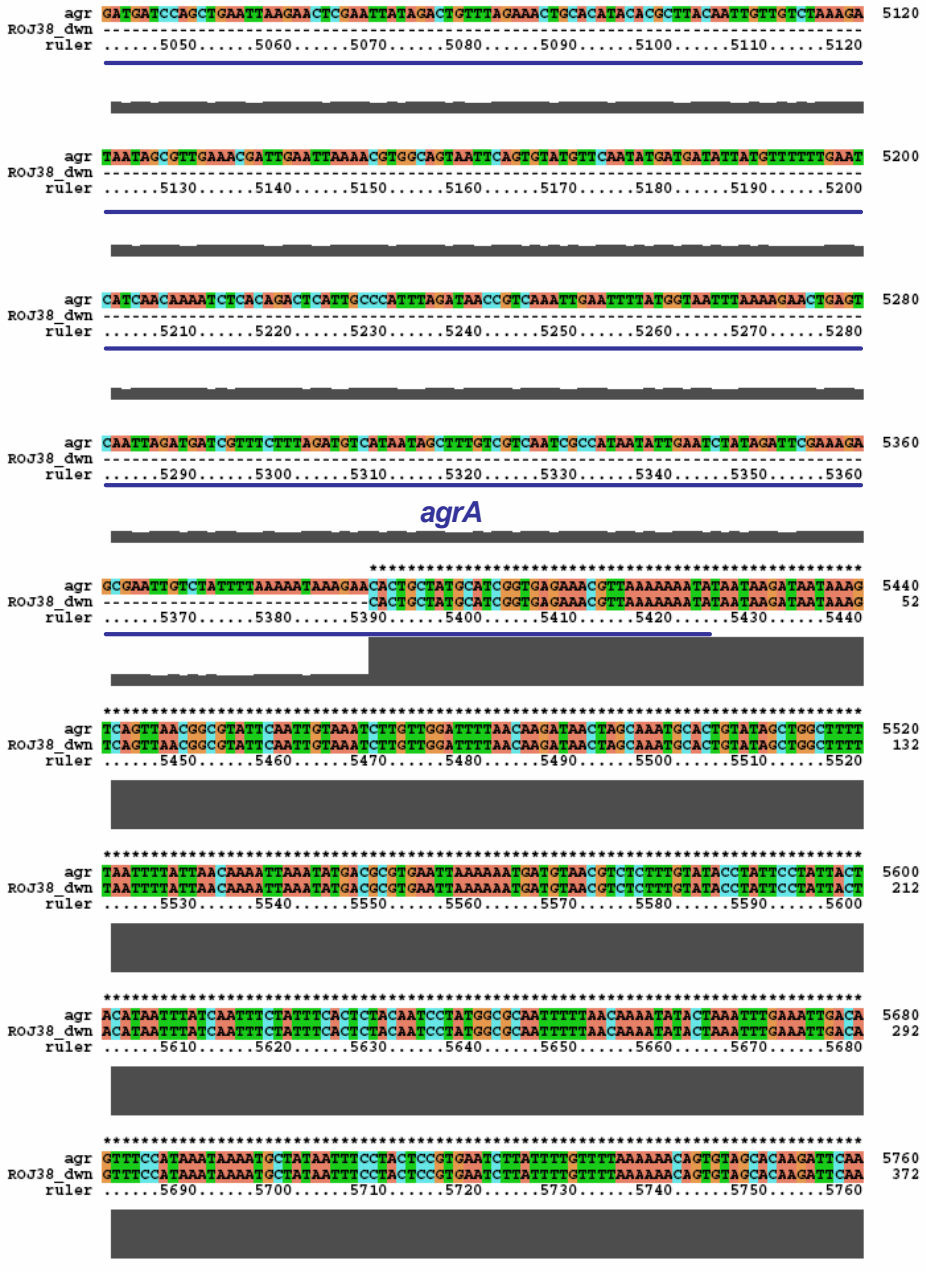
```
      agr TAATTTGAAAGATTCGATATCAATTAGCGATGACGAATTAGAATCAATGATTGCACATGCAGTTAAACATGTTCCTTCAT    80
 ROJ38_up --------------------------------------------------------------------------------
    ruler 1.......10........20........30........40........50........60........70........80

      agr CATTTAATTCTCAATCAACTAGAATTGTTCTCTTGTTAAATGACAACAATAATAAATTTTGGGATAACACAAAAGCAATT   160
 ROJ38_up --------------------------------------------------------------------------------
    ruler ........90.......100.......110.......120.......130.......140.......150.......160

      agr TTAAAAGAAGTCATGGGAGAGAATCGTGATTTCGAGCCAACTGAACAAAAAATTGATAACTTCAAACATTCATACGGTAC   240
 ROJ38_up --------------------------------------------------------------------------------
    ruler .......170.......180.......190.......200.......210.......220.......230.......240

      agr AATTTTGTTTTATGAAGATCAAGATGTTGTAAGTGGTTTACAAGAACAAATGCCAAACTACTATGATAATTTTGCAATAT   320
 ROJ38_up --------------------------------------------------------------------------------
    ruler .......250.......260.......270.......280.......290.......300.......310.......320

      agr GGTCAACTCAAACAAATGCCATGCATCAATTTGCAATTTGGACTGCATTAGCTACTAAAGGCATAGGTGCTTCATTACAA   400
 ROJ38_up --------------------------------------------------------------------------------
    ruler .......330.......340.......350.......360.......370.......380.......390.......400

      agr CACTATAATCCACTTGTTGATGAGATGACTTCAAATGAATTTAATATTCCTAAAAGTTGGAAACTTATTGCTCAAATGCC   480
 ROJ38_up --------------------------------------------------------------------------------
    ruler .......410.......420.......430.......440.......450.......460.......470.......480

      agr ATTTGGTGATATTCGTGAAGCTGCTGGCGAAAAAACATTTAATCCAGTGGAAGATCGTTTCGTAATAAAAAAATAATCAA   560
 ROJ38_up --------------------------------------------------------------------------------
    ruler .......490.......500.......510.......520.......530.......540.......550.......560

                                                                            **********
      agr TATATTGTTAGGAATGGGACAGAAATGATAAAGAATCATTAGTTATTCATTATGTAGTAGTTCTTACACATTAGCCGCTG   640
 ROJ38_up ----------------------------------------------------------------------TTAGCCGCTG    10
    ruler .......570.......580.......590.......600.......610.......620.......630.......640

          ********************************************************************************
      agr CTAATGTGCACTTAAGAATAGGAATACATGAGTAAAACTCATGTATAAGAAATACTAATTTCTAAAGAAAAAGTATTTCT   720
 ROJ38_up CTAATGTGCACTTAAGAATAGGAATACATGAGTAAAACTCATGTATAAGAAATACTAATTTCTAAAGAAAAAGTATTTCT    90
    ruler .......650.......660.......670.......680.......690.......700.......710.......720
```

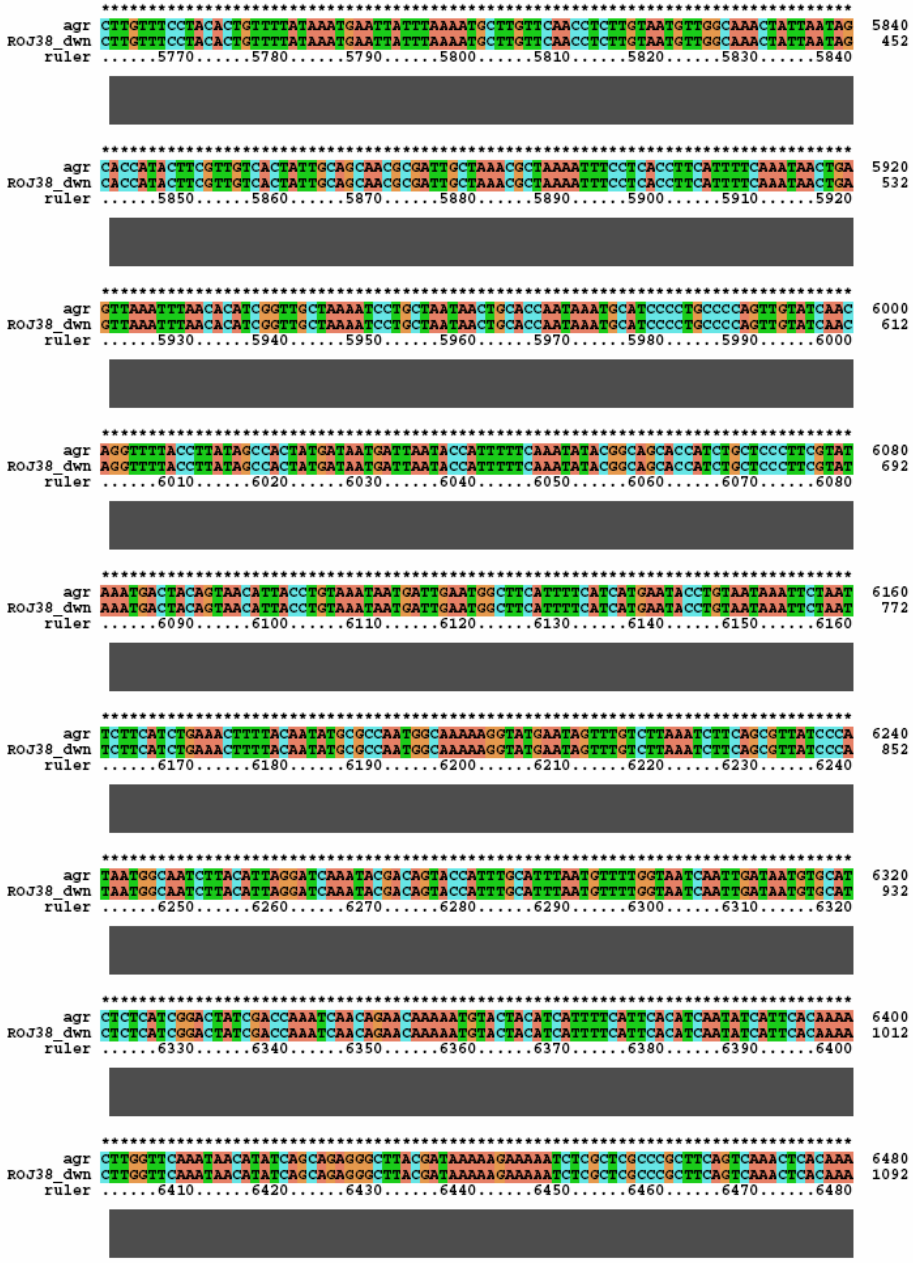
```
                                                                               
agr         TTATGTTAGAGCCCACCCCAACTTGCATTGTCTGTAGAAATTGGTGTTCCAATTTCTCTATGTTGGGGCCCCATCAACTA   800
ROJ38_up    TTATGTTAGAGCCCACCCCAACTTGCATTGTCTGTAGAAATTGGTGTTCCAATTTCTCTATGTTGGGGCCCCATCAACTA   170
ruler       .......730.......740.......750.......760.......770.......780.......790.......800

                                                                               
agr         CTGTCAATATAACTTTGTAGAGCCTAGGACTTTGATTTTATGTCTTGTGCTCGTTTTTTGCGTTTCAACATAACTTTTAG   880
ROJ38_up    CTGTCAATATAACTTTGTAGAGCCTAGGACTTTGATTTTATGTCTTGTGCTCGTTTTTTGCGTTTCAACATAACTTTTAG   250
ruler       .......810.......820.......830.......840.......850.......860.......870.......880

                                                                               
agr         TAAGTTATATATAAGATCTATCAAGGATGTGATGTTATGAAAGTCCAAATTTATCAATTACCGATTGTTTTTGGAGATAG   960
ROJ38_up    TAAGTTATATATAAGATCTATCAAGGATGTGATGTTATGAAAGTCCAAATTTATCAATTACCGATTGTTTTTGGAGATAG   330
ruler       .......890.......900.......910.......920.......930.......940.......950.......960

                                                                               
agr         TTCTAAAAATGAAACTCAAATAACGCAATGGTTCGAAAAAAATATGAATGCTGAAGTAGATGTAGTCGTTTTACCTGAAA  1040
ROJ38_up    TTCTAAAAATGAAACTCAAATAACGCAATGGTTCGAAAAAAATATGAATGCTGAAGTAGATGTAGTCGTTTTACCTGAAA   410
ruler       .......970.......980.......990......1000......1010......1020......1030......1040

                                                                               
agr         TGTGGAATAATGGCTATGATTTAGAGCATTTAAATGAAAAAGCAGATAATAATCTAGGACAAAGTTTTTCATTTATTAAA  1120
ROJ38_up    TGTGGAATAATGGCTATGATTTAGAGCATTTAAATGAAAAAGCAGATAATAATCTAGGACAAAGTTTTTCATTTATTAAA   490
ruler       ......1050......1060......1070......1080......1090......1100......1110......1120

                                                                               
agr         CATTTAGCTGAAAAATATAAAGTCGATATCGTAGCTGGGTCAGTTTCTAACATTAGAAACAATCAAATATTTAATACAGC  1200
ROJ38_up    CATTTAGCTGAAAAATATAAAGTCGATATCGTAGCTGGGTCAGTTTCTAACATTAGAAACAATCAAATATTTAATACAGC   570
ruler       ......1130......1140......1150......1160......1170......1180......1190......1200

                                                                               
agr         TTTTAGTGTGAATAAAAGTGGCCAGCTGATTAATGAATACGACAAAGTTCATTTAGTTCCTATGTTAAGAGAGCATGAAT  1280
ROJ38_up    TTTTAGTGTGAATAAAAGTGGCCAGCTGATTAATGAATACGACAAAGTTCATTTAGTTCCTATGTTAAGAGAGCATGAAT   650
ruler       ......1210......1220......1230......1240......1250......1260......1270......1280

                                                                               
agr         TTTTAACCGCTGGCGAATACGTAGCTGAACCCTTTCAATTGTCTGACGGTACTTATGTGACCCAACTTATCTGTTATGAT  1360
ROJ38_up    TTTTAACCGCTGGCGAATACGTAGCTGAACCCTTTCAATTGTCTGACGGTACTTATGTGACCCAACTTATCTGTTATGAT   730
ruler       ......1290......1300......1310......1320......1330......1340......1350......1360

                                                                               
agr         TTACGTTTTCCAGAATTATTAAGATATCCTGCTCGTAGTGGTGCTAAAATTGCATTTTATGTGGCGCAATGGCCTATGTC  1440
ROJ38_up    TTACGTTTTCCAGAATTATTAAGATATCCTGCTCGTAGTGGTGCTAAAATTGCATTTTATGTGGCGCAATGGCCTATGTC   810
ruler       ......1370......1380......1390......1400......1410......1420......1430......1440
```


agr ACGTCTTCAACATTGGCATTCATTACTAAAGGCACGTGCAATTGAAAATAATATGTTTGTCATTGGAACTAATAGCACTG 1520
ROJ38_up ACGTCTTCAACATTGGCATTCATTACTAAAGGCACGTGCAATTGAAAATAATATGTTTGTCATTGGAACTAATAGCACTG 890
ruler .....1450.....1460.....1470.....1480.....1490.....1500.....1510.....1520
agr GATTCGATGGTAACACAGAATATGCTGGTCATTCGATAGTTATAAACCCAAATGGAGATTTAGTTGGAGAATTAAATGAA 1600
ROJ38_up GATTCGATGGTAACACAGAATATGCTGGTCATTCGATAGTTATAAACCCAAATGGAGATTTAGTTGGAGAATTAAATGAA 970
ruler .....1530.....1540.....1550.....1560.....1570.....1580.....1590.....1600
agr TCCGCAGATATTTTGACTGTAGATTTAAACTTAAATGAAGTAGAACAGCAACGCGAAAATATACCTGTATTTAAAAGTAT 1680
ROJ38_up TCCGCAGATATTTTGACTGTAGATTTAAACTTAAATGAAGTAGAACAGCAACGCGAAAATATACCTGTATTTAAAAGTAT 1050
ruler .....1610.....1620.....1630.....1640.....1650.....1660.....1670.....1680
agr TAAATTAGATTTGTATAAATAAGAAGCGCCCGAAATAATATTTAACACATGTCATGACTAAACATAGATTTATGAGAGAT 1760
ROJ38_up TAAATTAGATTTGTATAAATAAGAAGCGCCCGAAATAATATTTAACACATGTCATGACTAAACATAGATTTATGAGAGAT 1130
ruler .....1690.....1700.....1710.....1720.....1730.....1740.....1750.....1760
agr TTTTCATAAGTCTATGTTTATTTTTTTATAATAGATAACTCAGTAAGAACCCATTTCGCCCTTTGCAAATGAATGATTTT 1840
ROJ38_up TTTTCATAAGTCTATGTTTATTTTTTTATAATAGATAACTCAGTAAGAACCCATTTCGCCCTTTGCAAATGAATGATTTT 1210
ruler .....1770.....1780.....1790.....1800.....1810.....1820.....1830.....1840
agr CTTAATTAAATTGATTAATTTTTTGACCTTTTCCAACATTAGACTTATTCATATATTTTAACGGCGGGTCTCATAATGAT 1920
ROJ38_up CTTAATTAAATTGATTAATTTTTTGACCTTTTCCAACATTAGACTTATTCATATATTTTAACGGCGGGTCTCATAATGAT 1290
ruler .....1850.....1860.....1870.....1880.....1890.....1900.....1910.....1920
agr GATTAACTCATCCCTTCTTCATTACAAAAAAGGCCGCGAGCTTGGGAGGGGCTCACGACCATACTTATTATTAAGGGAAT 2000
ROJ38_up GATTAACTCATCCCTTCTTCATTACAAAAAAGGCCGCGAGCTTGGGAGGGGCTCACGACCATACTTATTATTAAGGGAAT 1370
ruler .....1930.....1940.....1950.....1960.....1970.....1980.....1990.....2000
agr GTTTTACAGTTATTTTTTCAATCTATTTTTGGGGATGTTATTAATTATGAAAAAAATTTTATTTATCAAATAATGATAAA 2080
ROJ38_up GTTTTACAGTTA---------- 1382
ruler ......2010.....2020.....2030.....2040.....2050.....2060.....2070.....2080
agr TCGATGTTGTTTACGATAGCTTACATGCTAGAAATAATTATCTGTAATGATAATTAAGAAAAATACATAGCACTGAGTCC 2160
ruler .....2090.....2100.....2110.....2120.....2130.....2140.....2150.....2160


agr AAGGAAACTAACTCTACTAGCAAATGTTACTCACTTGCTAATTTTACATGGTTATTAAGTTGGGATGGCTTAATAACTCA 2240
ROJ38_up ---------- 1382
ruler .....2170.....2180.....2190.....2200.....2210.....2220.....2230.....2240
agr TACTTTTAGTACTATACGAAGATAACAAATTTACAATGAAAGTAATTAATTATTCATCTTATTTTTTAGTGAATTTGTTC 2320
ROJ38_up ---------- 1382
ruler .....2250.....2260.....2270.....2280.....2290.....2300.....2310.....2320
RNAIII
agr ACTGTGTCGATAATCCATTTTACTAAGTCACCGATTGTTGAAATGATATCTTGTGCCATTGAAATCACTCCTTCCTTAAT 2400
ROJ38_up ---------- 1382
ruler .....2330.....2340.....2350.....2360.....2370.....2380.....2390.....2400
agr TAAGATAAAAATTCTTAAAATTAAACAACTCATCAACTATTTTCCATCACATCTCTGTGATCTAGTTATATTAAAACATG 2480
ROJ38_up ---------- 1382
ruler .....2410.....2420.....2430.....2440.....2450.....2460.....2470.....2480
agr CTAAAAGCATTTATTTTCCAATTTTTCTTAACTAGTCGTTTTTTATTCTTAACTGTAAATTTTTTTATGTTAAAATTATTA 2560
ROJ38_up ---------- 1382
ruler .....2490.....2500.....2510.....2520.....2530.....2540.....2550.....2560
agr AATACAAATTACATTTAACAGTTAAGTATTTATTTCCTACAGTTAGGCAATATAATGATAAAAGATTGTACTAAATCGTA 2640
ROJ38_up ---------- 1382
ruler .....2570.....2580.....2590.....2600.....2610.....2620.....2630.....2640
agr TAATGACAGTGAGGAGAGTGGTGTAAAATTGAATTATTTTGATAATAAAATTGACCAGTTTGCCACGTATCTTCAAAAGA 2720
ROJ38_up ---------- 1382
ruler .....2650.....2660.....2670.....2680.....2690.....2700.....2710.....2720
agr GAAATAACTTAGATCATATTCAATTTTTGCAAGTACGATTAGGGATGCAGGTCTTAGCTAAAAATATAGGTAAATTAATT 2800
ROJ38_up ---------- 1382
ruler .....2730.....2740.....2750.....2760.....2770.....2780.....2790.....2800
agr GTTATGTATACTATTGCCTATATTTTAAACATTTTTCTGTTTACGTTAATTACGAATTTAACATTTTATTTAATAAGAAG 2880
ROJ38_up ---------- 1382
ruler .....2810.....2820.....2830.....2840.....2850.....2860.....2870.....2880


agr AATATTCCGATTAGTATCGAAATACCCGATGAAGTAAGTAGCATTAACTTGAATATGATCGATTTAAGTCGCAGTATTGG 4400
ROJ38_up ---------- 1382
ruler .....4330.....4340.....4350.....4360.....4370.....4380.....4390.....4400
agr TATTATTCTTGATAATGCAATTGAGGCATCAACTGAAATTGATGACCCTATCATTCGCGTTGCATTTATTGAAAGTGAAA 4480
ROJ38_up ---------- 1382
ruler .....4410.....4420.....4430.....4440.....4450.....4460.....4470.....4480
agr ATTCAGTAACGTTTATTGTTATGAATAAATGCGCTGATGATATACCACGCATTCATGAATTGTTCCAAGAAAGTTTTTCT 4560
ROJ38_up ---------- 1382
ruler .....4490.....4500.....4510.....4520.....4530.....4540.....4550.....4560
agr ACTAAAGGTGAAGGTCGTGGTTTAGGTCTATCAACTTTAAAAGAAATTGCTGATAATGCAGACAATGTCTTATTAGATAC 4640
ROJ38_up ---------- 1382
ruler .....4570.....4580.....4590.....4600.....4610.....4620.....4630.....4640
agr AATTATCGAAAATGGTTTCTTTATTCAAAAAGTTGAAATTATTAACAACTAGCCATAAGGATGTGAATGTATGAAAATTT 4720
ROJ38_up ---------- 1382
ruler .....4650.....4660.....4670.....4680.....4690.....4700.....4710.....4720
agr TCATTTGCGAAGACGATCCAAAACAAAGAGAAAACATGGTTACCATTATTAAAAATTATATAATGATAGAAGAAAAGCCT 4800
ROJ38_up ---------- 1382
ruler .....4730.....4740.....4750.....4760.....4770.....4780.....4790.....4800
agr ATGGAAATTGCCCTCGCAACTGATAATCCTTATGAGGTGCTTGAGCAAGCTAAAAATATGAATGACATAGGCTGTTACTT 4880
ROJ38_up ---------- 1382
ruler .....4810.....4820.....4830.....4840.....4850.....4860.....4870.....4880
agr TTTAGATATTCAACTTTCAACTGATATTAATGGTATCAAATTAGGCAGTGAAATTCGTAAGCATGACCCAGTTGGTAACA 4960
ROJ38_up ---------- 1382
ruler .....4890.....4900.....4910.....4920.....4930.....4940.....4950.....4960
agr TTATTTTCGTTACGAGTCACAGTGAACTTACCTATTTAACATTTGTCTACAAAGTTGCAGCGATGGATTTTATTTTTAAA 5040
ROJ38_up ---------- 1382
ruler .....4970.....4980.....4990.....5000.....5010.....5020.....5030.....5040


agrA


agr CTTGTTTCCTACACTGTTTTATAAATGAATTATTTAAAATGCTTGTTCAACCTCTTGTAATGTTGGCAAACTATTAATAG 5840
ROJ38_dwn CTTGTTTCCTACACTGTTTTATAAATGAATTATTTAAAATGCTTGTTCAACCTCTTGTAATGTTGGCAAACTATTAATAG 452
ruler .....5770.....5780.....5790.....5800.....5810.....5820.....5830.....5840
agr CACCATACTTCGTTGTCACTATTGCAGCAACGCGATTGCTAAACGCTAAAATTTCCTCACCTTCATTTTCAAATAACTGA 5920
ROJ38_dwn CACCATACTTCGTTGTCACTATTGCAGCAACGCGATTGCTAAACGCTAAAATTTCCTCACCTTCATTTTCAAATAACTGA 532
ruler .....5850.....5860.....5870.....5880.....5890.....5900.....5910.....5920
agr GTTAAATTTAACACATCGGTTGCTAAAATCCTGCTAATAACTGCACCAATAAATGCATCCCCTGCCCCAGTTGTATCAAC 6000
ROJ38_dwn GTTAAATTTAACACATCGGTTGCTAAAATCCTGCTAATAACTGCACCAATAAATGCATCCCCTGCCCCAGTTGTATCAAC 612
ruler .....5930.....5940.....5950.....5960.....5970.....5980.....5990.....6000
agr AGGTTTTACCTTATAGCCACTATGATAATGATTAATACCATTTTTCAAATATACGGCAGCACCATCTGCTCCCTTCGTAT 6080
ROJ38_dwn AGGTTTTACCTTATAGCCACTATGATAATGATTAATACCATTTTTCAAATATACGGCAGCACCATCTGCTCCCTTCGTAT 692
ruler .....6010.....6020.....6030.....6040.....6050.....6060.....6070.....6080
agr AAATGACTACAGTAACATTACCTGTAAATAATGATTGAATGGCTTCATTTTCATCATGAATACCTGTAATAAATTCTAAT 6160
ROJ38_dwn AAATGACTACAGTAACATTACCTGTAAATAATGATTGAATGGCTTCATTTTCATCATGAATACCTGTAATAAATTCTAAT 772
ruler .....6090.....6100.....6110.....6120.....6130.....6140.....6150.....6160
agr TCTTCATCTGAAACTTTTACAATATGCGCCAATGGCAAAAAGGTATGAATAGTTTGTCTTAAATCTTCAGCGTTATCCCA 6240
ROJ38_dwn TCTTCATCTGAAACTTTTACAATATGCGCCAATGGCAAAAAGGTATGAATAGTTTGTCTTAAATCTTCAGCGTTATCCCA 852
ruler .....6170.....6180.....6190.....6200.....6210.....6220.....6230.....6240
agr TAATGGCAATCTTACATTAGGATCAAATACGACAGTACCATTTGCATTTAATGTTTTGGTAATCAATTGATAATGTGCAT 6320
ROJ38_dwn TAATGGCAATCTTACATTAGGATCAAATACGACAGTACCATTTGCATTTAATGTTTTGGTAATCAATTGATAATGTGCAT 932
ruler .....6250.....6260.....6270.....6280.....6290.....6300.....6310.....6320
agr CTCTCATCGGACTATCGACCAAATCAACAGAACAAAAATGTACTACATCATTTTCATTCACATCAATATCATTCACAAAA 6400
ROJ38_dwn CTCTCATCGGACTATCGACCAAATCAACAGAACAAAAATGTACTACATCATTTTCATTCACATCAATATCATTCACAAAA 1012
ruler .....6330.....6340.....6350.....6360.....6370.....6380.....6390.....6400
agr CTTGGTTCAAATAACATATCAGCAGAGGGCTTACGATAAAAAGAAAAATCTCGCTCGCCCGCTTCAGTCAAACTCACAAA 6480
ROJ38_dwn CTTGGTTCAAATAACATATCAGCAGAGGGCTTACGATAAAAAGAAAAATCTCGCTCGCCCGCTTCAGTCAAACTCACAAA 1092
ruler .....6410.....6420.....6430.....6440.....6450.....6460.....6470.....6480

CLUSTAL X (1.81) MULTIPLE SEQUENCE ALIGNMENT

File: C:Documents and SettingsRasmus JensenMy DocumentsWork PhDSequencessequencingI Date: Sat Sep 27 11:13:31 2008
Page 10 of 10

```
            ****************
       agr  AGCTAACGCAGTGTTTGCTTCATTCGTGCGATACACTTTTGATACATCTACACCAATACTTGATATCGTTTCAATAATGC  6560
ROJ38_dwn  AGCTAACGCAGTGTTT----------------------------------------------------------------  1108
     ruler  ......6490......6500......6510......6520......6530......6540......6550......6560

       agr  TATCTCCAAATGCATCATTACCTAATTGTGTAATCATATACGCTTGTTGTCCTAACTTTTGAACTGTACAAGCTACATTG  6640
ROJ38_dwn  --------------------------------------------------------------------------------  1108
     ruler  ......6570......6580......6590......6600......6610......6620......6630......6640

       agr  CATGGCGCGCCACCAATTTGCTTCGTAAATGTTTGAACATCTTTTAAATTCGCATTTGTAACATTTGGAATAAAATCAAT  6720
ROJ38_dwn  --------------------------------------------------------------------------------  1108
     ruler  ......6650......6660......6670......6680......6690......6700......6710......6720

       agr  TAACGCTTCTCCTATTGAAAATAGACGTCTCATTTATGTTCATCCTTTAAATCATATTTAGTAAATTGTAAATATACTTG  6800
ROJ38_dwn  --------------------------------------------------------------------------------  1108
     ruler  ......6730......6740......6750......6760......6770......6780......6790......6800

       agr  CCCTGATTCTGTTGATGTTTTAATACCTAGTGCGTCTTCTGTCGGGAAAATTCGAGATGTCAAAACACGTTCTCCATCAT  6880
ROJ38_dwn  --------------------------------------------------------------------------------  1108
     ruler  ......6810......6820......6830......6840......6850......6860......6870......6880

       agr  TACAGAAAATTTCTATACTAGATGTATCAACAAAAATTTGTAATTGCTTTAATGGCGTGTC  6941
ROJ38_dwn  -------------------------------------------------------------  1108
     ruler  ......6890......6900......6910......6920......6930......6940.
```

Publications