THE UNIVERSITY
of EDINBURGH

THE UNIVERSITY
of EDINBURGH

Identification of non-coding RNA interactions that dictate *Staphylococcus aureus* virulence

Stuart William McKellar

September 2019

Thesis submitted for the degree of Doctor of Philosophy

Declaration

I declare that this thesis was composed by myself and that the work contained therein is my own, except where explicitly stated otherwise. The work has not been submitted for any other degree or professional qualification.

Stuart McKellar
September 2019

Abstract

Staphylococcus aureus is a bacterium which has gathered much attention over the past decade due to the emergence of both antibiotic-resistant and hyper-aggressive strains. These pose a significant threat to human health, particularly to individuals already weakened through other illnesses.

When *S. aureus* enters the human bloodstream, it must adapt in order to survive the challenging conditions faced. In particular, it must respond to the nutritional environment of the blood which is depleted in essential cofactors such as free iron. Additionally, *S. aureus* must survive attacks from the host immune system which will attempt to kill the invader through phagocytes and the production of antibodies.

The work carried out here aimed to understand how small, non-coding RNAs (sRNAs) regulate *S. aureus*' adaptation to the host bloodstream. These sRNAs are typically associated with regulating the translational efficiency and stability of mRNAs. Through use of a technique called "UV cross-linking, ligation and sequencing of hybrids" (CLASH), I identified novel targets of many sRNAs. In particular, I studied how RsaA, an sRNA involved in membrane homeostasis, regulates the translation of a transmembrane transporter involved in antiseptic and antibiotic resistance. Additionally, I identified interactions between RsaE, an sRNA involved in metabolism, and several toxin mRNAs from the phenol-soluble modulin class. This is a novel example of the direct link between cellular metabolism and virulence.

However, the most striking finding was that not only do sRNAs target mRNAs, but that they also target each other. I focused on two distinct sRNA – sRNA interactions; one between RsaA and RNAIII, and another between RsaE and RsaOG. The interaction between RsaE and RsaOG is an example of a so-called 'sponging interaction', where RsaOG is able to antagonise the activity of RsaE. This has the effect of freeing RsaE's targets from their regulation. I hypothesise that this ultimately induces the necessary metabolic changes required in order to survive the immediate nutritional stresses incurred after entering the bloodstream. Regarding RsaA and RNAIII, I hypothesise that this interaction is responsible for balancing virulent versus dormant behaviour. I suggest that RsaA is able to induce the destruction of RNAIII in order to steer the cell away from aggressive behaviour and into a more latent state.

Additionally, the interaction between these two sRNAs also appears to operate on exquisitely short timescales, demonstrating how capable bacteria are at adapting to stresses.

Ultimately, this work suggests that interactions between sRNAs are likely to be widespread and form a crucial aspect of stress responses in general. The experiments detailed herein have certainly not exhausted the produced data and I suspect that it will be utilised further in the future.

Lay abstract

Staphylococcus aureus is a bacterium which has gathered much attention over the past decade due to the emergence of antibiotic-resistant strains, namely the hospital 'superbug' MRSA. This poses a significant threat to human life, particularly to those individuals already weakened through other illnesses.

One of the most dangerous infections we can suffer from *S. aureus* is bacteraemia; when the bacterium enters our bloodstream. This allows *S. aureus* to disseminate around the body and induces panic in our own immune system that can cause sepsis if left unresolved.

After entering our blood, *S. aureus* is challenged with several stresses that it must survive. Firstly, it must find ways of obtaining the food and nutrients it needs. Next, the host individual will elevate their body temperature as part of the inflammatory response to infection. Such increased temperatures also cause stress for the invading bacteria, similarly to how we may feel hot, sensitive inflamed areas as painful. In addition, our defensive white blood cells will engage the bacterium and try to kill it. Finally, if antibiotics are administered to the patient then the bacterium will need to resist these as well. Thus, *S. aureus* must survive a wide variety of challenges in its adaptation to our bloodstream

In this thesis, I studied how *S. aureus* adapts using small, regulatory molecules to allow it to control its production of proteins, the machinery of the cell. It is known that these small, regulatory molecules are produced at different rates during stress, and I studied what they target, when they become active and what the final outcome for the cell is. This type of science is aimed at understanding the fundamental aspects of biology and laying the groundwork for the innovation of future therapies.

Acknowledgements

To my supervisor, Sander Granneman. I cannot stress enough how much your input has shaped me as a scientist. Principally, I thank you for your endless calm and patience (which I have no doubt pushed at times). You have ensured that I have never felt like a burden. You've celebrated my successes and made me feel like a genuine contributor, all while welcoming the inevitable problems and issues and treating these with tact and sensitivity. You have created a positive environment for learning, sharing, professional growth and, ultimately; performing science. Finally, I thank you for your inspiring passion and optimism which has filtered through into this previously most pessimistic Glaswegian.

To Ira. You are true friendship personified. Never before have I known somebody that I am completely unafraid to rely on and whose support is unwavering. How grey my PhD would have been without the laughs, the guilty-pleasure gossiping and the philosophical conversations that we have had together. Thank you. I truly believe that you will succeed in life.

Additionally, I thank the current and past members of the Granneman lab. Pedro, I thank you for our political and literary discussions and for your heroic genetic work. Iva, Tsui and Ross, I thank for their thoughtful discussions and good vibes. I also thank Amy Pickering for her all of her assistance and time.

Next, I am grateful to my thesis committee of David Tollervey, Ross Fitzgerald and David Gally who have all volunteered their time to meet and discuss my work. I thank my *viva* committee for their time. I also thank my collaborators Jai Tree and Julia Wong for their help with strains and experiments.

Finally, it is beyond my linguistic capabilities to put into accurate words how important my parents have been in supporting me throughout not just my education but also my entire life. I cannot express how the bravery and kindness of my mum has shaped me over the past year, and I will never forget how she has lead by example in how to cope with the trauma that can be life.

Contents

Abbreviations

Tab. 0.1: Abbreviations used in this thesis

Abbreviation	**Word**
AIP	Autoinducing peptide
AMP	Antimicrobial peptides
BHI	Brain heart infusion
CA-MRSA	Community-acquired MRSA
ClfB	Clumping factor B
DOR	Dense overlapping regulon
FFL	Feed forward loop
FnBP	Fibronectin binding protein
HA-MRSA	Hospital-acquired MRSA
Hld	Haemolysin δ
IL	Interleukin
Multilocus sequence type	MLST
MRSA	Methicillin-resistant *Staphylococcus aureus*
NAI	2-methylnicotinic acid imidazolide
NFL	Negative feedback loop
PrA	Protein A
PSM	Phenol-soluble modulin
PTR	Post-transcriptional regulation
RBP	RNA binding protein
RNP	Ribonucleoprotein complex
RBS	Ribosomal binding site
RNAP	RNA polymerase
SCV	Small colony variant
SD	Shine-Dalgarno
SIM	Single input module
TCA	Tricarboxylic acid
TF	Transcription factor
TLR	Toll-like receptor
TSA	Tryptic soy agar
TSB	Tryptic soy broth

1 Introduction

> *"My delight may be conceived when there were revealed to me beautiful tangles, tufts and chains of round organisms in great numbers, which stood out clear and distinct among the pus cells and debris...".*
>
> **— Sir Alexander Ogston**

1.1 The discovery of *Staphylococcus aureus*

Staphylococcus aureus was first detailed in the late 19th century by the Scottish surgeon Sir Alexander Ogston. This era in surgery was marked by an increasing appreciation of the importance of sterility and the often dire consequences of bacterial infection. Ogston was unconvinced by his contemporaries' opinion that air was the causative agent of wound suppuration and instead he examined a smear of pus he isolated from one of his patient's wounds under a microscope (Elek, 1959). His observations form the introductory quote to this chapter.

The image of "beautiful tangles, tufts and chains" led Ogston to name these '*Staphylococci*' from the Greek '*staphyle*', meaning a bunch of grapes. This nomenclature was furthered by the German surgeon Anton Rosenbach, who isolated two strains of *Staphylococci* and named them after their colony colour; *Staphylococcus aureus* (golden) and *Staphylococcus albus* (white) (Rosenbach, 1884). Ogston would additionally demonstrate the virulence of these microorganisms by injecting pus into guinea pigs and mice and detail the resulting abscesses. He also noted how *Staphylococci* would enter into the blood of the recipient animal and progress into septicaemia (Ogston, 1880).

1.2 *S. aureus* today

S. aureus is now realised as a ubiquitous bacterium that is commonly found on the skin and mucosal membranes of the body, in particular the nose and throat.

Indeed, it is estimated that 30% of the human population is colonised by *S. aureus* (van Belkum et al., 2009). Such infections are usually asymptomatic but can become a deadly threat to human health under particular circumstances. The events that lead to *S. aureus* acting pathogenically are not entirely understood, but immunocompromisation (including that caused by HIV/AIDS) and strain variation are thought to be the major determinants, with other factors such as diabetes also playing a role (Liu, 2009).

S. aureus frequently enters the body through an open wound or from an established niche such as the nasal passages (Liu, 2009). Additionally, nosocomial infections of *S. aureus* are associated with surgery and contaminated healthcare equipment such as catheters (Fowler et al., 2003). *S. aureus* may then overcome host defences and disseminate throughout the body, where it can cause a phenomenally broad range of diseases; endocarditis (infection of the heart), osteomyelitis (infection of the bone), pneumonia (infection of the lungs), soft skin infections, septicaemia (infection of the blood) and sepsis (Gordon and Lowy, 2008). The outcome of this threat is predictable; an estimate in 2012 attributed *S. aureus* with causing more deaths than AIDS, tuberculosis and viral hepatitis combined (Hal et al., 2012) and the Centres for Disease Control estimated that 23,000 deaths per year in the USA are associated with *S. aureus* (CDC, 2018). Additionally, there is an obvious economic burden, with one estimate in 2013 predicting that *S. aureus* costs the US economy 1.4-13 billion USD per year (Lee et al., 2013).

With regards to strain variation within *S. aureus*, antibiotic susceptibility has played a key role in its emergence as a global health concern. Methicillin, a β-lactam class antibiotic, was introduced as an effective agent against *S. aureus* in 1959. However, resistance was reported within a year (Gordon and Lowy, 2008). Methicillin's mode of action is through inhibiting cell wall synthesis by binding Penicillin-binding proteins (PBPs), a class of transpeptidases that are essential in the production of the peptidoglycan layer of the bacterial cell wall. Resistance to methicillin is conferred by the *mecA* gene, which encodes for an alternative PBP, PBP2a, that has decreased affinity for β-lactams. This *mecA* gene is encoded on a mobile genetic element, SCC*mec*, and is flanked by recombinase genes that facilitate intra- and interspecies horizontal transmission (Gordon and Lowy, 2008). Strains which acquire this resistance are referred to as 'Methicillin-resistant *Staphylococcus aureus*', or MRSA. These strains were originally associated with nosocomial infections and collectively identified as 'healthcare-associated MRSA' (HA-MRSA). Their antibiotic resistance makes them a great burden in such settings, where infected individuals typically have weakened immune systems and rely on antibiotic administration for protection. Indeed, Dantes et al., 2013, reported 65,000 cases of invasive HA-MRSA in the USA in 2011.

Following the emergence of MRSA strains in the 1960s, *S. aureus* infections were overall contained in the healthcare setting and in particular to the elderly and unwell. However, the 1990s saw the emergence of new, methicillin-resistant strains that were observed to infect healthy and young individuals with no recent healthcare episodes (Gordon and Lowy, 2008). Such strains are collectively identified as 'community-acquired MRSA' (CA-MRSA). In 1989 and 1991, an outbreak of CA-MRSA was observed in western Australia and the surrounding area (Udo et al., 1993; Gosbell et al., 2001). Between 1997 and 1999, four children died in North Dakota, USA, through a CA-MRSA strain that had a level of previously unwitnessed aggression. (Centers for Disease Control and Prevention, 1999)

However, CA-MRSA started to grip attention with the emergence of a strain now termed USA300. This strain was initially reported on from America and spread quickly through skin-to-skin contact. In particular, outbreaks were observed amongst prison inmates, soldiers and athletes (Gordon and Lowy, 2008). By 2011, USA300 was the most common isolate recovered from infections at all body sites (Diekema et al., 2014).

In 2005, a closely-related relative of USA300, termed 'USA300-Latin American variant', was isolated in Colombia and spread rapidly throughout the northern region of South America (Planet, 2017). USA300 has also been observed in France, Switzerland and England, although has not been able to establish itself to the same degree as observed in the Americas (Glaser et al., 2016; Toleman et al., 2016; Von Dach et al., 2016).

Strains of MRSA can be defined through a form of DNA fingerprinting. The bacterial chromosome is purified, digested with the rare cutting SmaI restriction enzyme and then ran on a gel and separated through pulse-field gel electrophoresis (Goering and Winters, 1992; McDougal et al., 2003). Polymorphisms in seven housekeeping genes (*arcC, aroE, glpF, gmk, pta, tpi and yqiL*) are used to define a unique allelic profile (Enright et al., 2000). This allelic profile is then matched to a bank of 53 known 'multilocus sequence types' (MLST). Regarding USA300, this strain is defined as being MLST8, positive for Panton-Valentine leukocidine, SCC*mec* type IV and containing the arginine catabolic mobile element (Tenover et al., 2006). Additionally, USA300 has been observed to express a number of additional virulence factors, such as phenol-soluble modulins and δ- and α-toxin (Wang et al., 2007). These are thought to contribute to the increased aggression and destructive capabilities of USA300.

1.3 Initial events of infection – from a *Staphylococcal* perspective

S. aureus infections most frequently originate from nasal carriages (Kluytmans et al., 1997). Individuals colonised here can either be persistent or only intermittent carriers (Williams, 1963). Interestingly, while persistent carriers are frequently colonised by only a single strain, intermittent carriers may harbour different ones over time (Eriksen et al., 1995; Vanden Bergh et al., 1999).

S. aureus is usually inoculated into the nose through direct contact with the hands (Wertheim et al., 2006). Once here, it must survive the defence systems of the host. Nasal secretions contain immunoglobulins, lysozyme and antimicrobial peptides in order to kill and immobilise invading bacteria (Chen and Fang, 2004; Kirkeby et al., 2000; Lee et al., 2002). *S. aureus* resistance to these is conferred by several mechanisms. For example, host secreted cationic antimicrobial peptides are thought to disrupt bacterial membranes to induce membrane depolarisation and therefore kill the cell (Bradshaw, 2003). However, *S. aureus* is able to resist these through reducing the net negative charge of its cell wall and thus decreasing the affinity of the peptide for its membrane (Peschel, 2002). One proposed mechanism for this involves MprF, which modifies phosphatidylglycerol with L-lysine (Peschel et al., 2001).

Secreted immunoglobulins facilitate phagocytosis and opsonisation once bound to surface antigens on their target. Once bound, the Fc region points outwards and is recognised by complement. However, *S. aureus* expresses protein A (prA) and Sbi on its cell surface, both of which bind the Fc region of IgG in a manner to leave it in an incorrect orientation for complement binding (Smith et al., 2012) . Indeed, prA-deficient bacteria are more susceptible to neutrophil-mediated phagocytosis and serum opsonins, and have reduced virulence in mouse models. (Palmqvist et al., 2002; Patel et al., 1987). Interestingly, prA also has an offensive role in virulence; it binds tumour necrosis factor receptor 1 on lung epithelial cells, which leads to a proinflammatory response mediated by expression of IL-8 to attract the immune system (Gómez et al., 2004). This increased inflammation damages the airway, loosens tight junctions between cells and furthers bacterial invasion. (Foster et al., 2014).

Following survival of the innate immune system, *S. aureus* must properly adhere to the nasal carriage. The epithelial wall of a nostril is fully keratinised and this offers an attachment site for *S. aureus* (Bibel et al., 1982). Clumping factor B (ClfB) and bacterial cell wall-associated teichoic acid have been demonstrated to be key

mediators in this (Weidenmaier et al., 2004; Wertheim et al., 2008). ClfB binds the C-terminal of keratin 10, a cell surface component of squamous cells (Walsh et al., 2004). Other proteins, including SdrC, SdrD and SasG, are also thought to contribute to host cell adhesion but are not fully understood (Foster et al., 2014).

S. aureus can also enter host cells with a variety of fibronectin binding proteins. FnBPA and FnBPB have been shown to be key mediators, both binding fibronectin arrays on the host cell surface. Once bound, this complex is recognised by integrins and internalisation of the bacterium ensues through endocytosis (Schwarz-Linek et al., 2004).

Infection of the nasal carriages with *S. aureus* been found to be a risk factor in the development of nosocomial infections after surgery (Kalmeijer et al., 2000; Kluytmans et al., 1995). Here, the nasal population may serve as a reservoir of *S. aureus* which can be spread to skin breakages during surgery. Additionally, nasal infection has been associated with systemic infections following liver cirrhosis and transplantation, in HIV-infected patients and for patients in intensive care units. (Chang et al., 1998; Nguyen et al., 1999; Pujol et al., 1996). Here, the weakened patient's immune system may be unable to keep *S. aureus* in the nasal carriages under control. Indeed, in at least 80% of cases, the nosocomial infection comes from an isolate of *S. aureus* that was already found in the patients mucosal membranes prior to their admittance (Wertheim et al., 2004).

Although the nasal carriages are thought to be the most common source of *S. aureus* for an infection, it is important to note that other sites in the body can be colonised independently as well, such as the skin, perineum and pharynx (Armstrong-Esther, 1976; Ridley, 1959). Additionally, contaminated hospital catheters have been found to be a prime source of *S. aureus* and these can deliver the bacterium directly into the blood. (Wertheim et al., 2004).

Once *S. aureus* delves deeper into its host tissue, it can no longer scavenge nutrients from the environment. Instead, these must be taken from the host. As such, considerable metabolic remodelling must occur to facilitate macromolecule catabolism. Additionally, *S. aureus* must produce virulence factors to induce host cell lysis and ultimately cause release of their nutrients. The exact changes in the metabolic profile of *S. aureus* in pathogenic conditions is not well understood, but some linkage between metabolism and virulence has been found. For example, starvation induces a decrease in intracellular GTP and branched-chain amino acids. These molecules act as allosteric activators of the CodY transcription factor and their loss results in the upregulation of CodY-repressed genes, including virulence factors such as haemolysins and RNAIII (Majerczyk et al., 2008; Stenz et al., 2011).

Iron is another cofactor whose loss is closely associated with virulence. This metal is essential to the growth and proliferation of *S. aureus* (Hammer and Skaar, 2011). Indeed, it is theorised that loss of free iron is used by *S. aureus* as a signal that it is in a vertebrate host through the iron-sensing Fur transcription factor (Skaar, 2010). Similarly to CodY and branched-chain amino acids, the absence of iron leads to Fur repression and then upregulation of several of its targets such including haemolysins, cytotoxins and iron-uptake transporters (Torres et al., 2010). Showing the importance of Fur-mediated adaption, deletion of this transcription factor results in *S. aureus* that are highly susceptible to neutrophils and also decreases the bacterial burden that they can cause (Torres et al., 2010).

Thus, *S. aureus*' transition from commensal to pathogenic bacterium is not a simple opportunistic act or molecular switch. From its normal niche within the nasal carriages, it may develop into a more systemic infection if the host immune system is weakened and therefore cannot keep *S. aureus* under control. Alternatively, surgeries or hospital equipment can bring *S. aureus* into direct contact with the interior of the body. Under these conditions, the limited nutrient availability forces *S. aureus* to induce virulent gene expression in order to extract the necessary materials for life.

1.4 Co-ordinating gene expression

The central dogma of biology whereby DNA is transcribed into RNA which is translated into protein describes the foundation of gene expression. Different environmental cues will induce transcription of different RNAs which will produce different proteins. Alternatively stated, the transcriptome of the cell will be modified by the act of transcription in response to stress and this change will be reflected in the expressed proteome, and therefore the behaviour, of the cell.

However, transcription is not the only means of altering the transcriptome. Crudely put, this act can be viewed as 'adding' to the transcriptome – but what of the RNAs already present within the cell that are no longer desired? For example, the shutting down of a particular metabolic pathway could be just as crucial for survival as activating another as it could drain the limited resources of the cell in non-essential programs. Here, the role of RNases and non-coding RNAs are critical.

This next section will summarise transcription in bacteria and detail transcriptional stress responses in *S. aureus*. Afterwards, it will describe the role of post-transcriptional regulation in shaping the transcriptome, and therefore proteome, of the cell and its demonstrated roles in stress adaptation and virulence.

Transcription in Gram-positive bacteria

Transcription of RNA from DNA is performed by the holoenzyme RNA polymerase (RNAP). Within eubacteria, the core of this enzyme is formed from four subunits; two α, one β and one β'; all of these are essential for viability. RNAP relies upon so-called 'sigma factors' (σ) for promoter recognition and the initiation of transcription (Helmann and Chamberlin, 1988). In Gram-positive bacteria, there also exists three smaller, accessory subunits; δ, ϵ and ω (Burgess, 1969). These are thought to aid RNAP folding, assembly, promoter specificity and recycling (Weiss and Shaw, 2015).

Each σ factor has its own regulon, usually a set of genes that are involved in similar cellular processes (e.g. 'housekeeping' processes, virulence, stress survival, etc). These factors direct RNAP to their promoter elements located -10 and -35 base-pairs upstream of the transcriptional initiation point. In *E. coli*, the RNAP core enzyme is expressed at mostly constant levels throughout growth and is the limiting factor in DNA transcription. Thus, competition occurs between σ factors during stress in order to dictate which genes are expressed (Maeda et al., 2000).

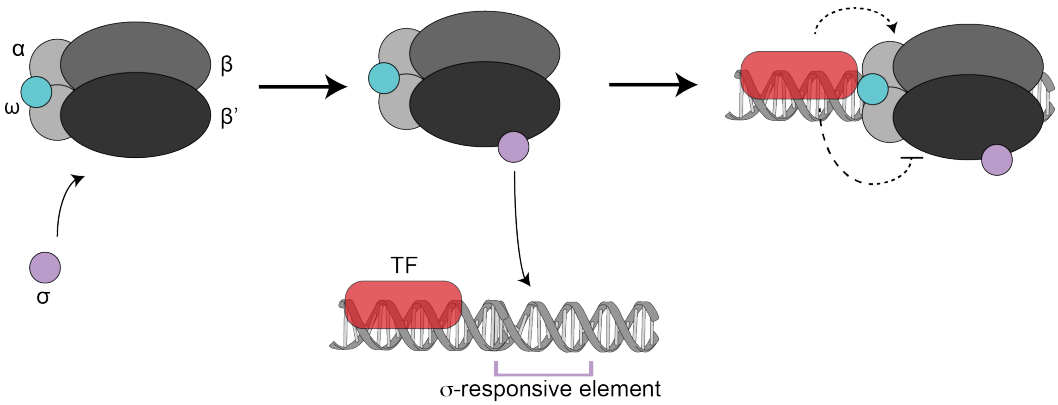

Fig. 1.1: Transcriptional initiation in Gram-positive bacteria. The σ factor directs RNA polymerase, composed of its core $\alpha_2\omega\beta\beta$' subunits, to its target promoter sites. Transcriptional factors (TFs) can also further modulate the activity of RNAP by interacting with the holoenzyme.

Within *S. aureus*, four σ factors have been identified. σ^A is considered to be the housekeeping factor that is responsible for the transcription of most genes in active, exponentially growing cells, acting homologously to σ^{70} in *E. coli* (Deora et al., 1996).

The principle sigma factor responsible for the expression of virulence genes is σ^B. A tiling array assay managed to identify 145 promoters as being dependent on this sigma factor (Mäder et al., 2016). A first role of σ^B is to direct the expression of several haemolysins and leukocidins, including *hla, hlg* and *lukDE* (Tuchscherr et al.,

2015). σ^B expression has also been associated with chronic, long term infections, reflecting its role in virulence promotion (Tuchscherr et al., 2015). However, σ^B is not only associated with virulence; it is also involved in regulating cell envelope composition, membrane transport and metabolism. Exemplifying this extensive network, σ^B is induced in response to thermal and alkaline stress, suggesting a role in stress adaptation and survival (Pané-Farré et al., 2006). However, even in the absence of stress, σ^B displays a basal level of expression in *S. aureus* which implies a role in general housekeeping (Mäder et al., 2016). Interestingly, this basal expression is not observed in *B. subtilis*, a non-pathogenic Gram-positive bacterium (Mäder et al, 2016). Thus, σ^B in *S. aureus* may have evolved to take on a more central role.

σ^S is an additional stress-responsive factor. It is induced upon exposure to compounds that induce DNA damage or cell wall stress, during nutrient starvation and also through growth in serum (Miller et al., 2012). Deletion of σ^S led *S. aureus* to be more prone to dying under these conditions (Shaw et al., 2008). Additionally, *S. aureus* σ^S mutants exhibited a decreased infectious burden in a murine model of septic arthritis, indicating a role in virulence (Shaw et al., 2008).

Although *S. aureus* is not genetically competent upon normal conditions, bioinformatic analysis found a sigma factor that was homologous to one responsible for competence in *B. subtilis*, σ^H (Morikawa et al., 2003). Under standard laboratory conditions, σ^H was not detectably expressed in *S. aureus*. However, artificial induction led to the expression of *comE* and *comG*, components of the DNA uptake machinery (Morikawa et al., 2003). Despite the initial failure to detect σ^H expression, further study managed to identify a small subpopulation of cells with expression under specific growth conditions. Importantly, this subpopulation was able uptake the SCC*mec*II genetic element to acquire methicillin resistance (Morikawa, 2012). As such, the authors hypothesised that a minor fraction of *S. aureus* exists in a competent state, conferred through σ^H, and this facilitates the uptake of e.g. antibiotic resistance cassettes.

In addition to sigma factors, the activity of RNAP can also be modulated by transcription factors (TFs). These bind to DNA, typically at sites spatially near RNAP, and can positively or negatively modulate its activity. Thus, TFs can be regarded as activators or repressors of gene transcription. Additionally, TFs contain a second domain that functions as a signal sensor through binding either small ligands or proteins that regulate its DNA-binding or RNAP-interacting activity. (Mascher et al., 2006). Thus, the architecture of TFs allows them to interpret environmental cues or cellular signals and relay the information by modulating the activity of RNAP in order to change transcription. Figure 1.2 summarises the core principle of transcriptional initiation in *S. aureus*.

Around 145 different TFs have been identified in *S. aureus* (Ibarra et al., 2013). This number was obtained through bioinformatic screening, and interestingly the authors noted that only around half of these had been characterised at the time of study. Using sequence conservation, hypothetical functionality was attributed to all but 9 of these unknown TFs. This resulted in the potential identification of novel two component regulatory systems and virulence and metabolism-related TFs (Ibarra et al., 2013).

Although there is clearly much more to understand about TF biology in *S. aureus*, some are well characterised. CodY and Fur were discussed in Section 1.4. AgrA is involved in the quorum-sensing network and will be discussed in Section 1.6; others of particular importance to this thesis are MgrA and Rot.

Strains lacking *mgrA* display attenuated virulence in an rabbit endocarditis model (Crosby et al., 2016). MgrA has been found to regulate around 350 genes in *S. aureus*, including important virulence-related transcripts such as nuclease, α-toxin, coagulase, protease, prA and autolysins. Additionally, MgrA also negatively regulates efflux pumps involved in vancomycin export and as such has a direct link to antibiotic resistance (Chen et al., 2006). Finally, MgrA has been found to be involved in regulating biofilm biology. It represses biofilm formation, at least in part, through negative regulation of *srtA*, or sortase, which functions to anchor cell wall proteins (Mazmanian et al., 2001, Trotonda et al., 2008). Additionally, MgrA also represses the expression of phenol-soluble modulins in cells within a biofilm, which effectively act as surfactants to induce cell release (Jiang et al., 2018). Thus, MgrA's extensive regulon affects many aspect of *S. aureus* biology.

Rot, or 'repressor of toxins', is a TF that is thought to dampen the expression of virulence-related genes such as lipase, haemolysins and secreted proteases (Saïd-Salim et al., 2003). It is also heavily involved in promoting biofilm formation and deletion of *rot* was found to greatly reduce the biofilm burden in a murine catheter model of infection (Mootz et al., 2015). This was confirmed as being a result of Rot repressing the expression of extracellular proteases and the authors were able to demonstrate direct Rot binding to their promoter regions, and their upregulation in the absence of Rot (Mootz et al., 2015). As such, Rot is another crucial transcription factor in *S. aureus* virulence.

Transcript elongation succeeds initiation. In this, RNAP extends the nascent RNA transcript as it moves along the DNA template. However, this process is also highly regulated. For example, RNAP contains a proofreading mechanism where, once a copying error is detected, it can backtrack along the gene, leaving the 3' end of the RNA exposed. During this state, RNAP pauses and erroneous transcript is trimmed back to the active RNAP core (Shaevitz et al., 2003). Pausing can also be a process

involved in premature termination, and thus be a component of gene expression regulation (Washburn and Gottesman, 2015).

Once RNAP reaches the end of a gene, the transcript must be terminated. This can be mediated in a gene-autonomous manner, whereby the structure of the produced RNA causes a stem-loop to form that induces the RNA to drop off of the polymerase. Alternatively, a factor called Rho can bind to the transcript and induce RNA release. Although *rho* in essential in *E. coli*, it can be deleted without a loss of viability or virulence in *S. aureus* (Washburn et al., 2001). Thus, it seems that Rho-independent termination is the dominant mode of termination in *S. aureus*.

Interestingly, non-coding RNAs (ncRNAs) have been shown to be able to attenuate Rho-mediated transcriptional termination in *E. coli*. Although Rho canonically binds to the 3' end of a transcript to terminate transcription of a finished RNA, it is also able to bind to the 5' UTR of a nascent RNA in order to induce premature termination. This occurs in a remarkable number of genes in *E. coli* and has been hypothesised to be a form of stress adaptation. ncRNAs play a role in the halting of this premature termination ('antitermination'), having been found to bind to the 5' UTR of targets previously targeted by Rho and preventing its binding (Sedlyarova et al., 2016). This mechanism therefore allows the cell to exist in a primed state where it can immediately activate stress-responsive genes (after halting premature termination) as they are already being transcribed.

Post-transcriptional regulation

Whilst transcription controls which RNAs are produced, the cell must also regulate the RNAs already existing within it. This revolves around controlling the localisation, stability and translational efficiency of RNAs, and is termed 'post-transcriptional regulation' (PTR, Figure 1.3A). Indeed, it has been shown in a wide variety of organisms that global mRNA levels do not correlate well with their respective protein levels, thus indicating the extensive regulation that mRNAs are subject to (Csárdi et al., 2015; Maier et al., 2009; Picard et al., 2009; Schwanhäusser et al., 2011).

PTR ultimately serves to direct and control the kinetics of a response. For example, in the case of a negative response, simply attenuating the transcription of a target gene would stop new mRNAs being produced but those already existing within the cell would continue to be actively translated. PTR is able to induce destruction of unwanted mRNAs on very short time scales and thus facilitate immediate rewiring of the transcriptome and proteome. This makes PTR a particularly important facet of gene regulation during stress responses (Gottesman et al., 2006). Additionally, PTR is able to suppress transcriptional noise of particular genes; this establishes

a dynamic where a minimum threshold number of mRNAs must be produced in order to induce a response. This can be useful for controlling the expression level of potentially hazardous genes or for avoiding target expression in response to a weak or transient signal (Nitzan et al., 2017). Additionally, PTR is able to shape gene expression profiles into patterns more complex than simply 'on' or 'off', and instead create pulses of varying amplitude and duration (discussed further in Section 1.10). Finally, whereas TFs typically regulate sets of functionally related genes, PTR is able to control the expression of a wide variety of mRNAs that may be regulated by different TFs or σ factors, thereby allowing cross-talk between various pathways.

In *S. aureus*, PTR has been found to be involved in the regulation of metabolism, cell death and virulence. Exemplifying this, RNAIII is already known to positively regulate the translation of *hla* and also *mgrA*, which itself drives transcription of *hla*. As such, PTR acts to drive the expression of Hla in a coherent feed-forward loop, and as a result induces expression at a maximal rate. Further examples of PTR will be given in Section 1.7 and how they integrate into regulatory networks in Section 1.9.

1.5 RNA-binding proteins and sRNAs; the major players of PTR

PTR relies on recognition of target transcripts to be repressed or activated, and proteins to carry these acts out. In bacteria, the act of recognition is frequently carried out by a class of ncRNAs often referred to as small RNAs (sRNAs). RNA-binding proteins (RBPs) also act in concert with these sRNAs to modulate target stability and translational efficiency. These players are described in more detail in Sections 1.7, 1.8 and 1.9.

The canonical mode of ncRNA-mediated post-transcriptional repression of an mRNA is represented in Figure 1.3B. In this example, an sRNA base-pairs with the Shine-Dalgarno (SD) sequence and blocks the 30S ribosome from initiating translation. Furthermore, this duplex can recruit RNases to degrade the mRNA. Conversely, sRNAs can also induce translation of their target mRNA, as shown in Figure 1.3C. In this example, an autoinhibitory loop in the mRNA prevents its translation. However, this loop can be melted by an sRNA to free the SD sequence, potentially aided by RNase-mediated cleavage.

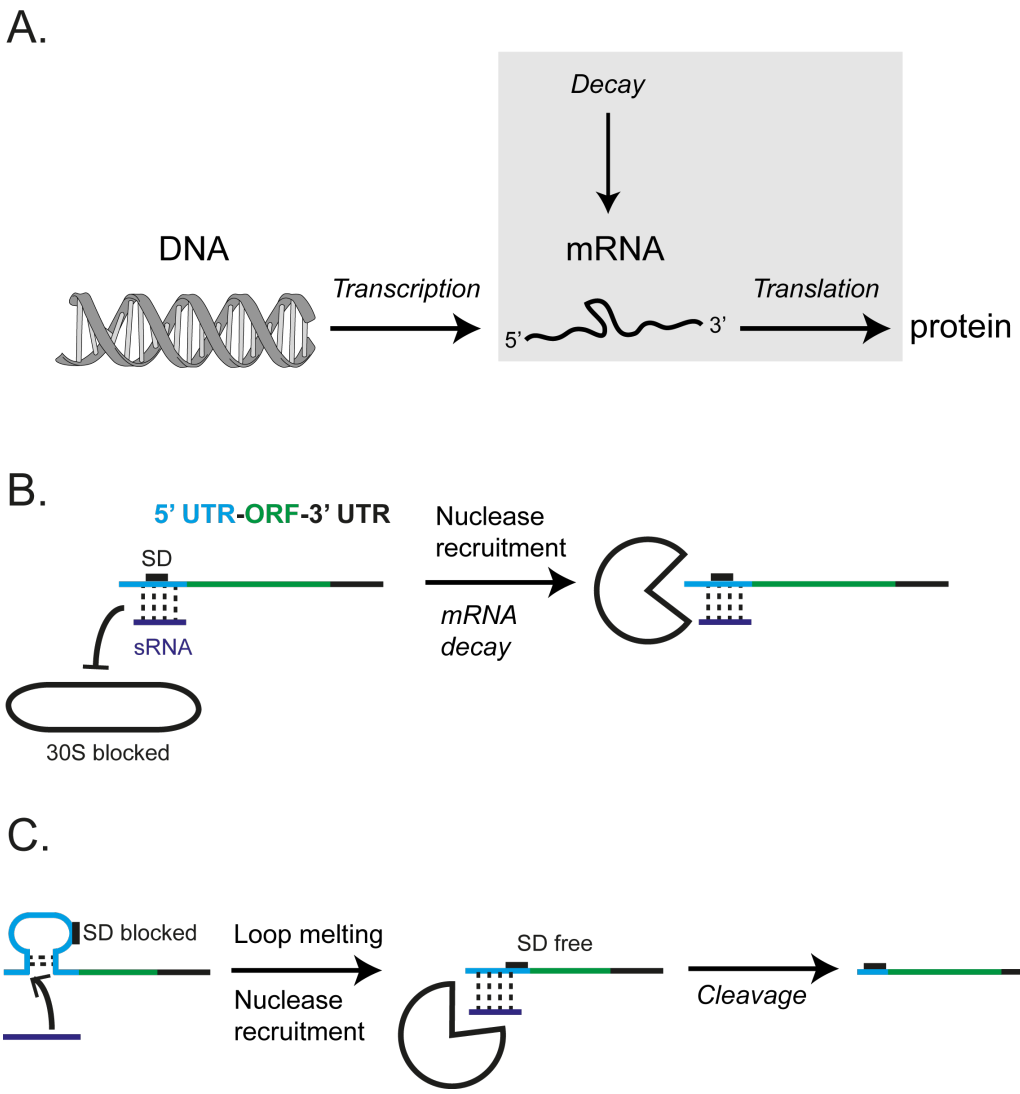

Fig. 1.2: The canonical mechanisms of post-transcriptional regulation. **A.**The central dogma of biology, with the region under which post-transcriptional regulation exerts its control highlighted. **B.** The typical mode of mRNA translational repression by an sRNA. The Shine-Dalgarno sequence (SD) is occluded through sRNA binding, preventing ribosome assembly. Furthermore, a nuclease can be recruited in order to degrade the sRNA-mRNA duplex. **C.** An example of target upregulation by an sRNA. An autoinhibitory loop on the mRNA which would otherwise occlude the SD sequence is melted by an sRNA and then a recruited nuclease cleaves the duplex in order to remove the inhibitory region.

Architecture of sRNAs

Small RNAs are heterogenous in their size, structure and function. Although they are typically 50-300 nucleotides in length, there are some of over 1,000 nucleotides. Additionally, although typically non-coding, some do encode for short peptides (Balaban and Novick, 1995).

Expression of sRNAs can either be *cis* to the target mRNA, i.e. encoded on the opposite strand at the same genomic locus, or *trans* to the target mRNA, being encoded at a separate locus (Figure 1.4).

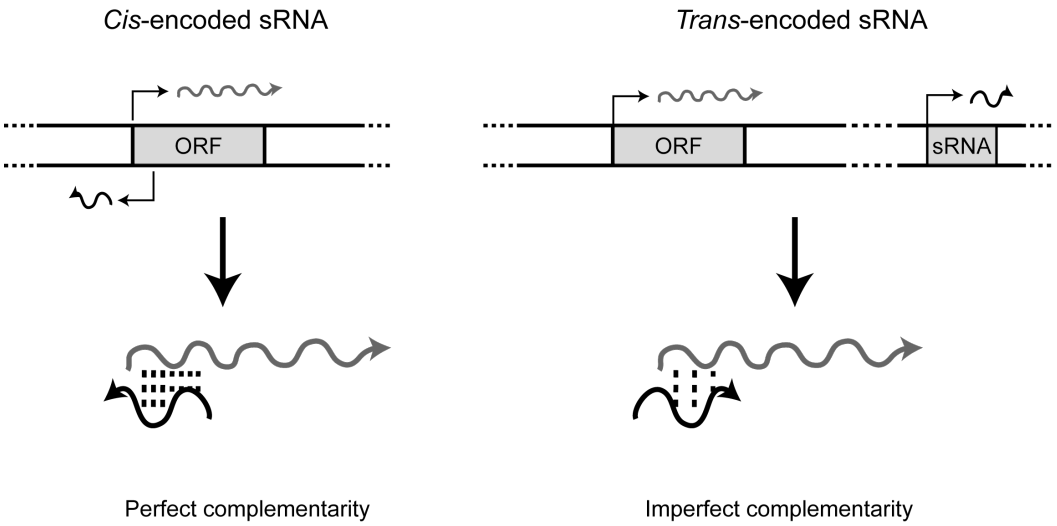

Fig. 1.3: The architecture of *cis*- and *trans*-encoded sRNAs. *Cis*-encoded sRNAs are encoded on the opposite strand of their target and exhibit perfect complementarity. *Trans*-encoded sRNAs are encoded at a distant genomic locus, and thus exhibit only partial complementarity to their targets. Figure adapted and redrawn from Prasse et al., 2013.

Currently there are over 500 sRNAs annotated in *S. aureus*, with the majority of these being identified through RNA sequencing experiments and *in silico* predictions (Liu et al., 2018; Sassi et al., 2015). However, there is disagreement about how many of these are *bona fide* sRNAs with a dedicated function in the cell, as opposed to transcriptional noise or unstable transcripts; one estimate suggests that there are only around 50 genuine sRNAs in *S. aureus*, discussed further in Section 1.8 (Liu et al., 2018).

Cis-encoded sRNAs exhibit perfect complementarity with their target. Despite this, the interaction between them is still thought to directed by an initial 'kissing interaction' starting between a few free nucleotides in accessible loops, before inducing conformational changes in the two RNAs to facilitate longer interactions (Guillet et al., 2013). Identified *cis* sRNAs to date revolve around replication machinery, transposases and toxin-antitoxin systems. The natural *Staphylococcal* plasmid pT181

regulates its own replication through production of RNAI, which is encoded *cis*- to the plasmid replication initiation protein, repC (Novick et al., 1989). Other examples include RsaOX, which exhibits perfect complementarity to the putative transposase SA0062 and RsaOW with the transposase IS1181 (Bohn et al., 2010).

An example of a *cis*-encoded sRNA involved in a toxin-antitoxin system in *S. aureus* is SprF1, which is part of the type I toxin-antitoxin SprG/SprF pair. SprG1 encodes for two cytolytic peptides from a single open reading frame and these are capable of inducing autolysis in *S. aureus*, but can also be secreted to lyse host erythrocytes. Interestingly, these peptides are also capable of lysing *E. coli* and *Pseudomonas aeruginosa*, two Gram-negative species of bacteria that frequently compete with *S. aureus*. Thus, the SprG1 toxin has roles in autolysis, erythrocyte killing (and therefore perhaps an indirect role in iron scavenging) and finally competition with other bacterial species. The SprF1 RNA is a *cis*-acting sRNA encoded on the opposite strand to SprG1 and is predicted to interfere with SprG1 translation and stability through canonical means (Pinel-Marie et al., 2014.). In agreement with their behaviour of a type I toxin-antitoxin system, the SprG1 sRNA has an incredibly long half-life of 120 minutes while SprF1 has a half-life of 10 minutes. As such, SprF1 must be transcribed at much higher rates than SprG1 (Pinel-Marie et al., 2014).

Interestingly, there is an example of a *cis*-encoded sRNA acting in *trans* with its target in *S. aureus*. The sRNA SprA1$_{AS}$ is part of type I toxin-antitoxin system. Here, the stable toxin mRNA, SprA1, encodes for a cytolytic peptide that inserts into the host cell membrane and disrupts its integrity. In order for the cell to survive, it must continuously produce the labile SprA1$_{AS}$ antitoxin RNA to block translation of SprA1. SprA1$_{AS}$ is transcribed from the complementary DNA strand of SprA1. However, despite exhibiting perfect complementarity at its 3' end with the 3' end of *sprA1* in a 35 nucleotide tract, SprA1$_{AS}$ acts in trans by base-pairing to the 5' end of *sprA1* to occlude the ribosomal binding site (RBS) (Sayed et al., 2012). There also exists a proposed SprA2/SprA2$_{AS}$ pair in *S. aureus*, based upon sequence identity, that may operate in a similar manner (Guillet et al., 2013).

Due to the fact that *trans*-acting sRNAs are encoded at distant genomic loci, they exhibit only partial, and at times interrupted, complementarity with their target. In *S. aureus*, these interactions are often extended in comparison to *E. coli*. This is potentially due to its AT-rich genome; as the interaction between A and T is weaker than that between G and C, a longer interaction is needed to compensate if it is primarily composed of A-T bonds (Guillet et al., 2013). Interestingly, many of the *bona fide* sRNAs rely on a conserved seed sequence to initiate binding to their target (Liu et al., 2018). Exemplifying this, RsaE contains two UCCC seed sequences located in conserved hairpins that are both capable of repressing target mRNAs (Rochat et al., 2018). RNAIII also uses UCCC sequences found in three of its loops to interact

with target SD sequences (Bronesky et al., 2016). Currently, most published sRNA – mRNA interactions in *S. aureus* involve the sRNA pairing to the 5' end of the mRNA to prevent translation, but exceptions are known involving RNAIII and SprX (discussed further in Sections 1.7 and 1.8).

1.6 RNAIII as a paradigm of sRNA-mediated regulation in *S. aureus*

RNAIII is a 516-nucleotide long sRNA that operates as the effector molecule of the quorum sensing, *agr* operon in *S. aureus*. This operon acts as a global regulatory system that responds to cell density in order to control the expression of myriad virulence factors. Indeed, given its crucial role in pathogenicity, there has been considerable interest in the *agr* pathway as a potential therapeutic target (Painter et al., 2014).

The quorum sensing machinery relies upon four proteins; AgrA, AgrB, AgrC and AgrD, whose expression is driven by the P2 promoter. The main effector molecule, RNAIII, is driven from the P3 promoter slightly downstream of the *agr* operon. The *agrD* locus produces a 46 amino acid-long propeptide that is then processed by the transmembrane protein AgrB. The product, autoinducing peptide (AIP), is then secreted and binds to a transmembrane receptor kinase, AgrC, and induces its autophosphorylation. This phosphorylated AgrC is then able to phosphorylate and activate the cytosolic transcription factor AgrA, which dimerises and then binds to the intergenic region between the P2 and P3 promoters in order to further drive expression of the *agr* operon and also RNAIII (Painter et al., 2014). The *agr* operon is pictorialised in Figure 1.5.

Interestingly, four allelic variants of AgrD, and therefore AIP, exist. Each variant can only induce the autophosphorylation of their respective AgrC variant and they inhibit the activity of others. When binding to their complementary AgrC, they induce a conformational change in the C-terminal cytoplasmic helix that subsequently brings the sensory and kinase domains together, enabling autophosphorylation. When a heterologous AIP binds to AgrC, the cytoplasmic helix twists in the opposite direction and so the sensory and kinase domains remain distant (Wang et al., 2014).

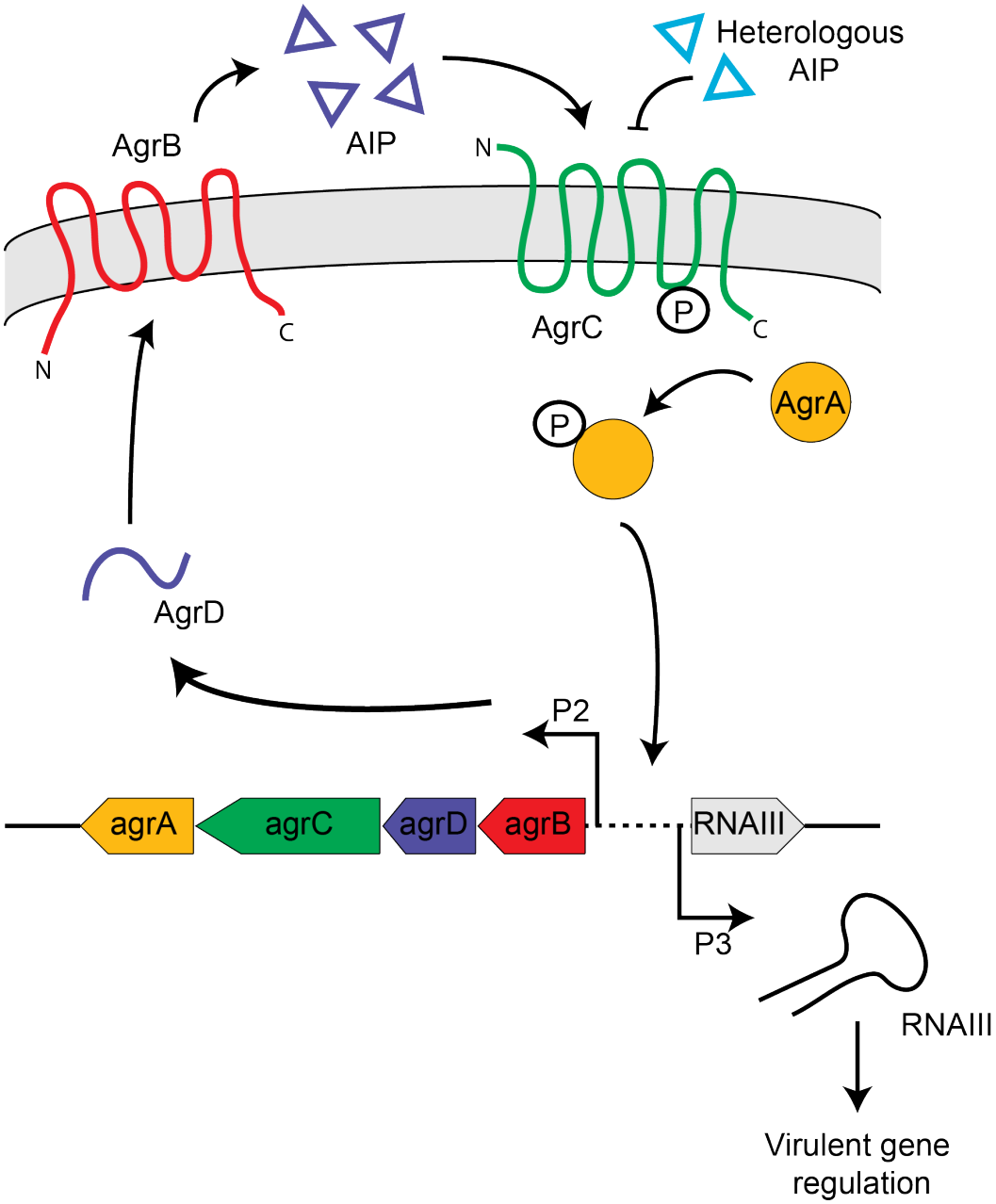

Fig. 1.4: The structure and mechanics of the quorum sensing system in *S. aureus*. AgrD is processed into autoinducing peptide (AIP) through the action of the transmembrane protein AgrB. AIP is then secreted and binds the transmembrane kinase receptor AgrC, which then autophosphorylates and activates AgrA through phosphorylation. AgrA then drives expression from the P2 promoter to upregulate the *agr* operon, and also the P3 promoter to drive RNAIII expression. RNAIII then drives virulent gene expression. Adapted and redrawn from Painter et al, 2014.

The *agr* pathway is also integrated into nutrient sensing networks. CodY, a transcriptional repressor, acts to sense intracellular carbon and nitrogen sources. During exponential growth these nutrient sources are high in concentration and CodY acts to indirectly repress the *agr* operon (Roux et al., 2014). Thus, *agr*-regulated virulence genes are coupled to nutrient status within the cell.

Once RNAIII expression is activated at high cell density, it acts to mediate the switch from the expression of cell-surface associated proteins to pathogenic behaviour, including the production and secretion of toxins. RNAIII itself encodes for a small, 26 amino acid-long cytolytic peptide, δ-haemolysin (Hld), that can insert into membranes and cause cell lysis (Verdon et al., 2009). However, the principle mode of action of RNAIII is through post-transcriptional regulation of its target mRNAs.

RNAIII is involved in the only documented cases of an sRNA upregulating the translation of an mRNA in *S. aureus*. First, the 5' UTR of the *hld* region binds to the 5' UTR of α-haemolysin and acts to free the SD sequence and recruit the ribosome (Novick et al., 1993). Secondly, RNAIII is known to stabilise the mRNA of a global pro-virulence transcription factor, MgrA (Gupta et al., 2015). Here, both the 5' and 3' end of RNAIII interact with the long 5' UTR of the *mgrA* mRNA at two distinct regions and prevent its destruction by RNases. MgrA acts to repress biofilm formation, promote metabolic remodelling , and positively regulates the expression of several toxins including *hlg, lukDE* and *lukMF* (Luong et al., 2006). Thus, RNAIII can indirectly cause major shifts in cell behaviour through regulating MgrA. Finally, RNAIII has been shown to stabilise the *map* which encodes for a surface adhesion protein that can prevent leukocyte adhesion and induce T-cell death (Ivain et al., 2017; Liu et al., 2011).

However, RNAIII also acts as a negative regulator to repress targets and induce RNase-mediated degradation (Boisset et al., 2007; Chevalier et al., 2010; Tomasini et al., 2014) and it is not thought that RNAIII needs a protein chaperone to aid its activity (Bronesky et al., 2016). Through direct base-pairing interactions, RNAIII downregulates the expression of *sbi, prA* and *fnBP* (Boisset et al., 2007; Chabelskaya et al., 2014; Huntzinger et al., 2005). Interestingly, the interaction between RNAIII and *sbi* occurs at two distinct sites on the *sbi* mRNA; one at the 5' UTR involving the 3' end of RNAIII and another within the coding sequence that utilised the 5' end of RNAIII (Chabelskaya et al., 2014). This mechanism of action where RNAIII binds both the translational initiation region at the 5' UTR and within the coding sequence is also observed for another of its targets, *coa* (Chevalier et al., 2010).

RNAIII also represses the translation of *rot* ('Repressor of toxins'), a TF that directly binds to the promoter regions of many toxin transcripts and blocks their transcription (Saïd-Salim et al., 2003). Rot is also known to stimulate the expression of PrA, and

thus RNAIII represses the translation of this mRNA both directly and indirectly (Oscarsson et al., 2006). RNAIII also represses the translation of *lytM*, a hydrolytic enzyme involved in regulating cell wall homeostasis (Chunhua et al., 2012). The activity of RNAIII is summarised in Figure 1.6.

In conclusion, RNAIII is a multifaceted RNA that acts to switch the cell from a more dormant, 'defensive' state into pathogenic behaviour upon the reaching of a critical cell density. This act is carried out on multiple levels. Firstly, RNAIII downregulates the 'front line' proteins on the cellular membrane that are involved in host immune system avoidance. Secondly, RNAIII stimulates the production of toxins which are then secreted in order to attack the host. However, RNAIII also modulates key transcription factors in the cell by downregulating *rot* and stabilising *mgrA*, thus also changing the core transcriptional program of the cell. As such, RNAIII can be regarded as a post-transcriptional molecular switch.

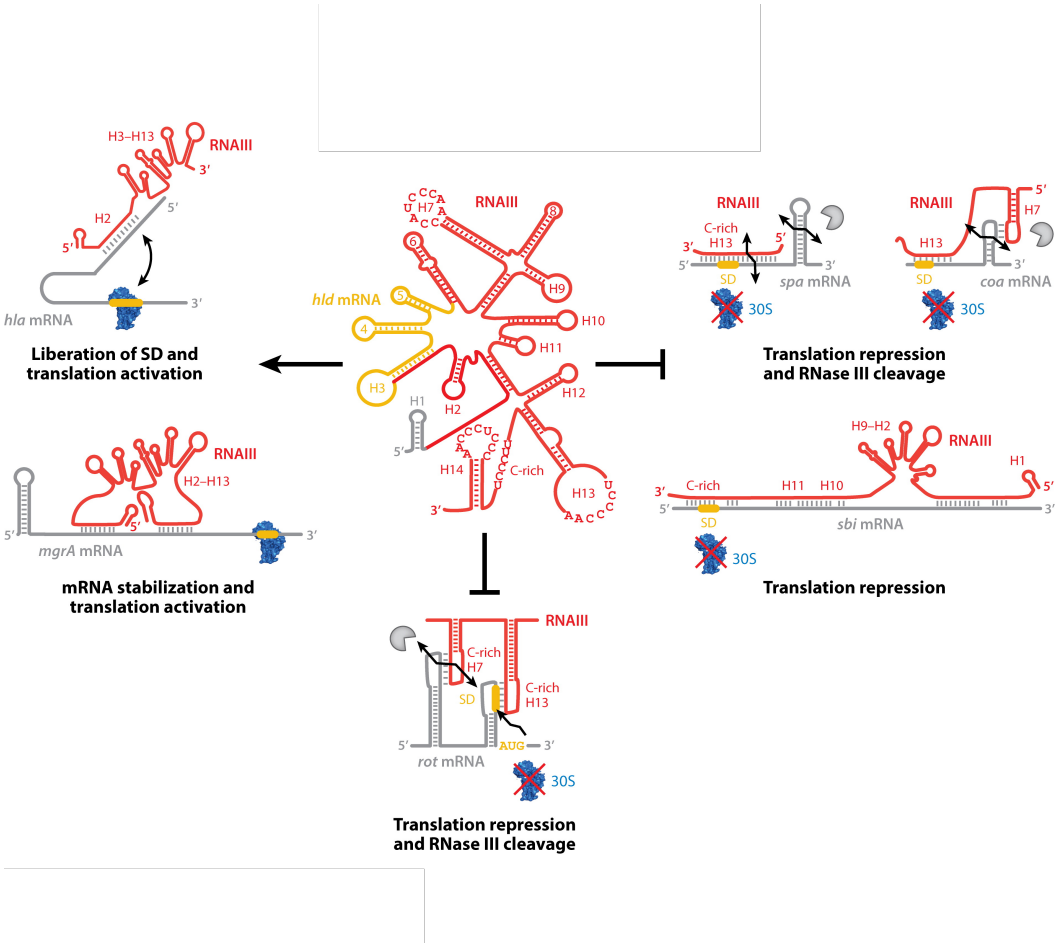

Fig. 1.5: The targets and regulatory mechanisms of RNAIII. RNAIII negatively regulates *spa, coa, sbi* and *rot* through classical RBS occlusion. In the cases of *spa, coa* and *rot*, this is accompanied by RNase III recruitment and transcript degradation. Conversely, RNAIII is able to stimulate translation of *hla* and *mgrA* through transcript stabilisation and liberation of the SD from autoinhibitory loops. This image was taken from Bronesky et al, 2016.

1.7 Other sRNAs in *S. aureus*

The Spr and Rsa family of sRNAs are a heterogenous collection that have been identified through *in silico* predictions and subsequent validation. Their shared names do not represent collective function but rather their discovery.

The Spr family

The Spr ('Small pathogenicity island RNA') family of RNAs was identified in 2005, at a time when only RNAI and RNAIII were known (Pichon and Felden, 2005). Multiple approaches were used in order to identify the original Spr sRNAs; total RNA was resolved on an agarose gel and stained in order to find highly expressed ncRNAs, and intergenic regions were screened computationally and through microarrays. This identified 7 sRNAs expressed from pathogenicity islands, named SprA to SprG, and these vary in length from 90 to 400 nucleotides. Since their discovery, other sRNAs have been added to this family (such as SprX), despite being encoded in the core genome.

The original paper managed to obtain *in vitro* evidence that SprA regulated the translation of an ABC transporter, *SA2216*. The role of other Spr sRNAs has since been expanded. SprX is now known to directly bind the RBS of *spoVG*, repressing its translation, through a C-rich loop at its 3' end. SpoVG is a DNA-binding protein that is involved in methicillin and vancomycin resistance. It regulates the expression of extracellular protease, nuclease and lipase and has a role in capsule formation. Indeed, plasmid-borne overexpression of SprX reduces *S. aureus* resistance to glycopeptide antibiotics (Eyraud et al., 2014). SprX is additionally known to repress the translation of *ecb*, a protein involved in inactivating host blood complement (Ivain et al., 2017) and upregulate the translation of *clfB, coa* and *hld*, and has subsequent influence on the formation of biofilms and on virulence (Kathirvel et al., 2016). SprD is known to repress the immune-evasive protein *sbi* in a similar manner and deleting this sRNA results in a serious attenuation of virulence (Chabelskaya et al., 2014, 2010). SprC is also known to be a negative regulator of virulence and control *Staphylococcal* uptake by monocytes and macrophages; null-mutants for SprC displayed a higher level of virulence and increased uptake by phagocytes. This was found to be at least partially a result of SprC repressing the translation of *atl*, an adhesin/invasin that interacts with host cells. As such, SprC was hypothesised to be involved in attenuating virulence and therefore play a role in host-*S. aureus* commensialism (Le Pabic et al., 2015).

The Rsa family

The Rsa ('RNA from *S. aureus*') family of sRNAs was originally identified through *in silico* predictions based on intergenic regions conserved in numerous Gram-positive bacteria and also throughout *S. aureus* strains (Geissmann et al., 2009). Putative sRNAs were then confirmed through microarray analysis and Northern blotting. This identified 11 previously unknown sRNAs, named RsaA to RsaK. All but RsaK are predominantly expressed in stationary phase in rich medium, but their expression profiles change significantly in stringent medium, indicating that their expression patterns are tightly controlled and respond to external conditions (Geissmann et al., 2009). Adding to this idea, RsaA, RsaE and RsaI were induced under acidic pH and oxidative stress (while RsaF and RsaG were not), and the expression of RsaC and RsaD were significantly upregulated by cold shock (Geissmann et al., 2009).

The function of several of these sRNAs has now been elucidated. RsaE, a key metabolic regulator, has been found to be conserved across firmicutes and in particular *Bacillus subtilis* (Durand et al., 2015). In *S. aureus*, RsaE regulates the tricarboxylic acid (TCA) cycle by downregulating the translation of several component enzymes, namely *sucC, sucD, citB, citC*, and *citZ* (Bohn et al., 2010). RsaE also has a role in regulating the amino acid composition of the cell; its expression induces transcription of valine, leucine and isoleucine-responsive operons (Guillet et al., 2013). Additionally, it downregulates the arginine degradation pathway (Rochat et al., 2018), and it has also been shown to downregulate the *opp-3* operon that encodes for oligopeptide transporters (Bohn et al., 2010). RsaE also downregulates a component of the purine biosynthesis pathway, *fhs* (Bohn et al., 2010). RsaE's regulation of the TCA cycle and amino acid metabolism has led to suggestions that it facilitates the transition between energy metabolisms from late-exponential phase to stationary phase, where acetate and amino acids are used as alternative carbon sources (Guillet et al., 2013). Finally, as RsaE expression is dependent upon the *agr* pathway and its expression is σ^B-dependent, it has also been suggested that it may play a role in virulence (Geissmann et al., 2009). In *B. subtilis*, RsaE (known as RoxS) has been found to respond to oxidative stress, be involved in the regulation of redox-related transcripts and also been predicted to regulate additional genes involved in electron transfer (Durand et al., 2015, Durand et al., 2017).

The σ^B-dependent sRNA RsaA is involved in attenuation of virulence (Romilly et al., 2014; Tomasini et al., 2017). An initial study by Romilly et al, 2014, deleted the coding sequence for RsaA and analysed the cytosolic protein content. This identified *spoVG* as a potential RsaA target, but also a number of proteins that are part of the MgrA regulon. Further study also found that expressing RsaA from a plasmid decreased the amount of MgrA protein produced. RsaA, through its UCCC seed sequences, was subsequently found to bind to two distinct regions on *mgrA*; one sequestering the RBS (and therefore blocking translation), and another which acts to recruit RNase III to induce degradation. Deletion of RsaA also changed the

protein levels of several cell surface proteins such as PrA, Ebh_1 and Ebh_3, ClfB and others, although this effect is likely to be indirect and through regulation of MgrA (Tomasini et al., 2017). The mutation also reduces biofilm formation and inhibits capsule production, in agreement with its role as a negative regulator of MgrA. Additionally, deletion was also found to induce a higher bacterial load in the blood and spleen of mouse models (Romilly et al., 2014). Thus, RsaA is hypothesised to attenuate invasiveness and aggressive pathogenicity, and therefore may be favoured in commensal interactions with the host.

Several RsaA targets were identified through MAPS, where the target RNA is tagged with an MS2 affinity tag, purified, and the RNAs bound to it identified through sequencing (discussed further in Section 1.12). This revealed that RsaA has a role in regulating components of the cell surface (Tomasini et al., 2017). This confirmed that *mgrA* is a major target of RsaA, but also identified members of the *Staphylococcal secretory antigen A class* (*SsaA*), which are involved in peptidoglycan metabolism, and *flr* (encoding for FLIPr), a secreted immunomodulatory protein that inhibits complement and IgG opsonisation of *S. aureus*. The distinctive UCCC motif in RsaA was found to be essential for interaction with these transcripts and it binds to the RBS to inhibit mRNA translation.

The Teg and Sau families

The Teg ('Transcript from experimental method from Geneva') family are a loosely connected group of putative sRNAs originally identified through RNA sequencing by the Francois group (Beaume et al., 2010). Originally, 150 were identified from intergenic regions in the N315 core genome and 9 from a plasmid. Thirty-one were hypothesised to be *cis*-acting riboswitches and 57 to be *trans*-acting, independent transcripts. Additionally, 23 were the product of anti-sense transcription at a protein coding gene and exhibited perfect complementarity with their respective mRNA; several of these were antisense for transposes and also present in multiple copies. In order to validate a subsection of these, RT-qPCR was performed on 26 chosen putative sRNAs. Most were found to be highly expressed at stationary phase in rich medium, and that several also responded to thermal, oxidative, or pH stress (Beaume et al., 2010).

The Francois group then collaborated with the Stinear group in Australia to perform similar work in two Australian *S. aureus* strains derived from ST239 (Howden et al., 2013). One strain, JKD6008, displays intermediate vancomycin resistance while the other, JKD6009, is vancomycin-sensitive. RNAseq was performed on each of these strains after 2 and 6 hours of exposure to 4 different antibiotics and a no antibiotic control. In total, they identified 409 putative sRNA sequences, with 79 of these being encoded in the accessory genome. Of these 409, 171 were found to be products of

antisense transcription at mRNAs and 94 contained a 'UCCC' motif that commonly blocks the RBS, as seen for RsaA, RsaE and RNAIII. Additionally, they identified 70% of the aforementioned Tegs in their analyses. In particular, 47 putative sRNAs were identified that displayed significant differential regulation in response to antibiotic exposure.

Although no validation of these transcripts has been performed, Howden et al, 2013, noted that many of the sRNAs responsive to an antibiotic are anti-sense transcripts for mRNAs that operate in protein synthesis and ribosomal function. For example, sRNA242 was downregulated in all antibiotic conditions and this sequence is antisense to the 30S ribosomal protein S1. Thus, there may be a subset of antisense sRNAs that act to regulate the expression of ribosomal and protein production genes.

Similar RNAseq studies have been performed in a so-called 'small colony variant' (SCV) of *S. aureus* (Abu-Qatouseh et al., 2010). These SCVs display slower growth, reduced virulence, decreased sensitivity to antibiotics and typically produce chronic infections. RNAseq and computational analyses identified 142 potential sRNAs, each termed 'Sau-n' ('*S. au*reus ncRNA'). Seventy-eight of these were found to be products of anti-sense transcription of known mRNAs, 16 partially overlapped with known ORFs (and so may represent UTRs), and 48 derived from intergenic regions. Eighty Sau transcripts were chosen for validation through Northern blots and signal was detected for 18 of these. Comparing expression patterns of SCV and WT strains led to the discovery of several Sau transcripts which exhibit differential regulation in SCVs, such Sau-27 (an intergenic transcript) being lost in SCVs or Sau-66 (an antisense product) being upregulated in SCVs (Abu-Qatouseh et al., 2010).

Finally, a tiling array study managed to identify 7 transcriptional units which were theorised to be potential *trans*-acting sRNAs, simply named e.g. "S35" (Mäder et al, 2016). Two of these were observed to highly expressed in a large number of tested conditions, two were upregulated in conditions that mimic bloodstream infection and one was highly expressed when grown in human plasma. As such, some of these candidates sRNAs may be involved in virulence.

sRNA curation in S. aureus

Given that multiple studies have been published each with their own naming convention, an attempt to curate sRNAs in *S. aureus* has been attempted (Sassi et al., 2015). For example, the transcript encoding for RsaA has been discovered in all the studies mentioned above. As such, it has also named Sau-64, Teg88, sRNA132 and finally given the systematic name srn_1510 by Sassi et al; the same applies for RsaE (Sau20, Teg92, sRNA183, srn_2130). Integrating all of these identifications together

has led to the curation of 539 sRNA sequences in NCTC8325, 530 in USA300 and 607 in JKD6008.

However, it is clear that only a handful of sRNAs have any meaningful functionality attributed to them. It is therefore difficult to distinguish between *bona fide* sRNAs from transcriptional noise and poorly annotated UTR sequences. Additionally, given that *trans*-acting sRNAs are of greater interest than *cis*-actors due to their ability to regulate a large number of transcripts and integrate several pathways and cellular responses, it is important to be able to correctly identify these for future study. As such, greater stringency must be applied when defining transcripts as putative *trans*-acting sRNAs.

A study in 2018 attempted to perform such filtering of the sRNA database in HG003, a NCTC8325 derivative (Liu et al., 2018). They defined a *trans*-acting sRNAs as one with its own promoter and terminator; does not overlap with any gene from the opposite strand; is not a processed UTR; is not a product of premature termination; and contains a transcription factor/σ binding site. They identified putative promoters in their strain using RegPrecise, which relies on comparative genomics to identify TF binding sites (Novichkov et al., 2013); in order for an sRNA to be defined as having its own promoter, it required one of these binding sites between -100 and +50 nucleotides of its 5' end. In order to identify terminators, the authors assumed that *trans*-acting sRNAs utilise Rho-independent terminators and so searched for these using TransTermHP (Kingsford et al., 2007). This tool searches for charactersitic markers of Rho-independent terminators, such as an A-rich tail in the 5' end of the gene, followed by a short, low-energy hairpin and finally ended with a T-rich tail.

From their pool of 527 putative sRNAs, they concluded that only 46 were genuine, *bona fide, trans*-acting sRNAs, with the rest mostly being asRNAs and UTR sequences (Liu et al, 2018). However, it should be noted that processed UTRs have been shown to produce functional sRNAs in Gram-negative bacteria and these will be missed by Liu et al's strict definition (reviewed by Chao et al., 2012 and Ren et al., 2017, Iosub et al., 2019). Overall, although the pool of sRNAs appear to be large in *S. aureus*, scepticism is required in their interpretation.

1.8 RNA-binding proteins and sRNA-mediated regulation

RNAs, including sRNAs, frequently require the activity of an RNA-binding protein (RBP) in order to exert their function, and it is the resultant ribonucleoprotein complex (RNP) that exerts the functional outcome. As such, the roles of RBPs have

been a significant focus of research into post-transcriptional regulation, particularly in Gram-negative bacteria. This section will detail the current knowledge of RBPs in *S. aureus* and, where appropriate, compare this to current knowledge in Gram-negatives such as *E. coli* and *Salmonella typhimurium*, and also the Gram-positive *B. subtilis*.

Hfq

Hfq is an RNA chaperone that plays a central role in the functionality of *trans*-encoded sRNA in several bacterial species. Overall, Hfq acts to stabilise sRNAs by protecting them from ribonuclease cleavage and facilitates sRNA-target interactions (Storz et al., 2011; Vogel and Luisi, 2011). It is present in a number of bacterial phyla, including several firmicutes, actinobacteria and α-, β- and γ-proteobacteria.

Hfq operates as a homohexamer, forming a donut-like shape with two RNA-binding faces on either side. Each face has a specific RNA target motif, with the proximal face binding AU-rich sequences and the distal binding poly(A) sequences (Link et al., 2009; Schumacher et al., 2002). This dual-activity of Hfq facilitates its binding to an sRNA on one face and an mRNA on another. This increases the local concentration of sRNA and target and may also induce conformational changes in the RNAs to encourage base pairing. In *E. coli*, Hfq is also able to recruit RNase E in order to degrade sRNA-mRNA pairs (Bandyra et al., 2012; Waters et al., 2017).

Deletion of Hfq in many pathogenic bacteria, such as uropathogenic *E. coli* and *S. typhimurium*, leads to growth defects, increased susceptibility to stress and reduced virulence (Chao and Vogel, 2010). In *Salmonella*, Hfq is thought to regulate around 20% of all genes, either directly or indirectly, and this includes those involved in host cell invasion, motility, central metabolism, cell wall regulation and two-component pathways (Chao and Vogel, 2010).

Although Hfq is found in many *S. aureus* strains, its role and importance in post-transcriptional regulation is unclear. An initial study found that *S. aureus* Hfq was able to bind RNAIII and *spa* RNAs *in vitro* (Huntzinger et al., 2005). However, deleting the Hfq coding sequence in three *S. aureus* strains (COL, Newman and RN6390) did not show significant changes in RNAIII expression or Protein A production (Bohn et al., 2007). Additionally, this study examined the RN6390 Δ*hfq* strain in more detail and performed around 2000 phenotypical assays (such as nutrient source usage, sensitivity to chemical stresses, etc) but could detect no significant or robust phenotypes. As such, this study concluded that either RNA-RNA interactions in *S. aureus* do not require the aid of an RNA chaperone, or that currently unidentified one(s) performs this role. In agreement with the former idea, several

trans-acting sRNAs have been found to act completely independently of Hfq (Guillet et al., 2013).

Western blotting analyses in wild-type RN6390 and COL revealed that Hfq is undetectable in these strains (Liu et al., 2010). This is potentially due to the loss of the gene's major promoter (Geisinger et al., 2006). However, RN6390 and COL are laboratory-adapted strains and may not represent accurately an *S. aureus* strain in the outside environment. In agreement with this, deleting *hfq* in *S. aureus* 8325-4 led to changed expression profiles of 116 genes as detected by microarray analysis (fold change ≥ 1.5), and immunoprecipitating Hfq identified 49 of these mRNAs. Included in these 49 were *sbi, sucD* and *rot*; all mRNAs with known sRNA regulators. Additionally, deleting Hfq resulted in a decrease of virulence in a murine peritonitis infection model and an increase in surface carotenoid pigment (Liu et al., 2010). This effect on carotenoid pigment has also been confirmed in the N315 strain (Castro et al., 2011). As such, Hfq may not be an indispensable tool across all *S. aureus* strains and may exhibit strain specificity. Given that *S. aureus* does not express RNase E, it may act to recruit other cellular nucleases, as described below.

Finally, deletion of Hfq does not cause a strong phenotype in *B. subtilis* (Rochat et al., 2015). This has led to the suggestion the need for a general chaperone may not be necessary in species with more stable sRNA-target interactions. This fits with the idea of extended interactions due to the AT-rich genome, as described in Section 1.5. Alternatively, it may be that species such as *B. subtilis* and *S. aureus* rely on more condition-specific RNA chaperones, as opposed to a single, global one (Mars et al., 2016). However, at the time of writing, dedicated RNA chaperones have not yet been validated in *S. aureus*.

RNase III

RNase III is a magnesium-dependent endoribonuclease that contains a C-terminal RNA-binding domain and an N-terminal ribonuclease domain. It targets double-stranded RNA, such as sRNA-mRNA duplexes, loop-loop interactions and stacked helices (Chevalier et al., 2008). It operates as a homodimer and cleavage results in a 5'-phosphate and a 3' dinucleotide overhang. In *B. subtilis* and *E. coli*, RNase III has been shown to be involved in 30S rRNA processing in order to generate the 16S and 23S subunits (Srivastava and Schlessinger, 1990) and also regulates the expression of phages, plasmids and cellular genes (Wagner et al., 2002). As RNase III only recognises double-stranded RNA, it can be recruited through *trans*-acting sRNAs.

Deletion of RNase III in *S. aureus* is viable and mutants appear without growth defects (Huntzinger et al., 2005). In constrast to this, RNase III is essential in *B.*

subtilis, and deletion in *E. coli* results in slow growth (Herskovitz and Bechhofer, 2000; Studier, 1975). The lethality in *B. subtilis* upon RNase III removal is a result of its essential role in suppressing the translation of two toxin mRNAs, *txpA* and *yonT* which are part of toxin-antitoxin systems of two prophages (Durand et al, 2012). These are targetted for degradation through the action of asRNAs which act to recruit RNase III (Durand et al, 2012).

However, RNase III mutants in *S. aureus* display impaired processing of several virulence-related transcripts. For example, Δ*rnc* strains cannot degrade *spa* mRNA encoding for Protein A. As described in Section 1.7, the 3' end of RNAIII is able to duplex with *spa* to block the RBS, and this dsRNA structure is recognised by RNase III and cleaved (Chevalier et al., 2008). This mechanism of action is not restricted to *spa* and has also been demonstrated for *fnBP, coa, sa1000, ssaA*, and *rot* (Boisset et al., 2007) and confirmed *in vitro* (Chevalier et al., 2008). Corroborating the role of RNase III in regulating pathogenicity, a Δ*rnc* mutant was shown to secrete fewer immunogenic proteins and toxins and had diminished virulence in a murine acute peritonitis model.

RNase III has been the subject of RIP-seq experiments in order to identify its cellular targets. As baits, both a wild-type RNase III and a catalytically-dead (E135A) mutant were used (Lioliou et al., 2012). Corroborating its known links to rRNA processing, rRNA transcripts were identified as major targets of RNase III . Additionally, catalytically-dead RNase III was found to bind antisense transcripts of 1175 mRNAs (44% of the genome). As such, RNase III likely plays a role in degrading *cis*-encoded sRNAs in duplex with their partner mRNA, but also potentially products of pervasive transcription. These pervasive transcription products are the result of e.g. failed termination events and continued RNA polymerase elongation, promoter sequences present in intergenic regions and bidirectional promoters (Wade et al, 2014). These products are undesirable and non-coding transcripts that are then targetted for degradation. RNase III was also found to bind key virulence factors such as *rot* and *spa*, and interestingly also bound 58 different ncRNAs including sRNAs. These include numerous sRNAs from the Spr and Rsa families, and RNAIII (Lioliou et al., 2012). Thus, RNase III is also predicted to be a key player in sRNA-mediated decay of mRNAs.

In conclusion, RNase III is an important RNase within the cell with clear links to the regulation of pathogenesis. It is involved in degrading mRNAs that have been targeted for destruction by sRNAs and this applies to key virulence factors.

RNase Y

RNase Y has been described as the functional homologue of RNase E in *S. aureus*, which operates as the central degradosome scaffold in *E. coli*. It is a 5' monophosphate-dependent endonuclease and has been found to be responsible for initiating bulk mRNA decay in *B. subtilis* (Durand et al., 2012). Indeed, depletion of RNase Y led to the dysregulation of 1,261 transcripts, representing just over a quarter of *B. subtilis*' gene expression. For many transripts after RNase Y cleavage, RNase J1 then degrades the rest of the transcript (Durand et al, 2012). Additionally, depletion of RNase Y led to the stabilisation of both known sRNAs and asRNAs, indicating a role of RNase Y in degrading several of these transcripts. (Durand et al, 2012).

In *S. aureus*, RNase Y is essential for virulence (Marincola Gabriella et al., 2012). RNase Y was found to be key in processing the *saePQRS* operon, of which *saeR* and *saeS* act as a two-component system to regulate the expression of numerous virulence genes such as *spa, nuc, coa,* and several haemolysin mRNAs. Deletion of *rny* resulted in the differential regulation of 569 genes as detected by microarray and included in these were the sRNAs RsaA and Sau63. The most common gene ontology terms associated with the deregulated ORFs were pathogenesis, proteolysis, transport and metabolic processes. In agreement with RNase Y's role in pathogenesis, a Δ*rny* mutant was found to have attenuated virulence in a murine bacteraemia model (Marincola Gabriella et al., 2012). Interestingly, RNase Y was found not to regulate RNAIII or any component of the *agr* pathway, thus this regulation of virulence was proposed to be independent of the quorum sensing system.

RNase Y has also been hypothesised to be the backbone scaffold of the RNA degradosome in both *S. aureus* and *B. subtilis*. After RNase Y cleavage, RNase J1 (a 5'-3' exonuclease) and PNPase (a 3'-5' exonuclease) subsequently degrade targeted transcripts transcript (Durand et al., 2012). Deletion of the N-terminal tail of RNase Y induced a growth defect (as opposed to deletion of the entire locus, which does not) and thus cytosolic RNase Y may scavenge degradosome components and mislocalise them (Khemici et al., 2015). The degradosome is discussed further in the section below.

Khemici et al, 2015, also identified which transcripts are directly controlled by RNase Y by mapping transcript ends through RNAseq. They found that deletion of RNase Y significantly affected 248 transcripts; significantly less than that published by Marincola et al., 2012. However, *sarA* and *sarR* were found by Khemici et al., corroborating Marincola et al's description of RNase Y regulating the *sae* operon. However, Khemici et al. found *agrA* and *agrC* upregulated upon RNase Y deletion, which was not found in Marincola's data. Khemici et al. also found six sRNAs stabilised in the absence of RNase Y, including RsaA, SprX and RNAIII. Using 5' end mapping, Khemici et al. identified 99 RNase Y cleavage sites, of which 50 were

in ORFs and the remainder in UTRs, ncRNAs and intergenic regions. The authors concluded from these results that RNase Y is the prime regulator of only around 100 transcripts and therefore its activity is under considerable regulation and control (Khemici et al., 2015). In conclusion, RNase Y is likely not an RNase involved in bulk mRNA degradation in *S. aureus*, but rather involved in the processing and degradation of virulence-related transcripts.

The degradosome – E. coli, B. subtilis and S. aureus

In *E. coli*, bulk mRNA degradation is mediated by a holoenzyme complex called the degradosome. It is composed of the endonuclease RNase E, the 3'-exonuclease polynucleotide phosphorylase (PNPase), RNA helicase B (RhlB) and the metabolic enzyme Enolase (Carpousis, 2002; Py et al., 1994). RNase E is thought to be the central player of this complex as it is the scaffold upon which the complex is built, and it is also the initiating factor in RNA decay (Carpousis, 2002). The *E. coli* degradosome is pictorialised in Figure 1.7A.

RNase E contains an N-terminal catalytic domain and a long, unstructured C-terminal domain that contains RNA and protein-binding sites. It is on this C-terminal domain that the degradosome is assembled (Carpousis, 2007). PNPase acts to degrade mRNAs at the 3' end through phosphorolysis, while RhlB melts short stretches of duplexed RNA in order for it to be fed into RNase E/PNPase (Carpousis, 2007).

The functionality of Enolase in the degradosome is unclear. Enolase is a glycolytic enzyme that catalyses the dehydration of phosphoglycerate to phosphoenolpyruvate. It is an abundant cellular protein, but only a small fraction of it is bound to the degradosome. It is thought that Enolase acts to somehow link information about the energetic state in the cell to mRNA degradation, but the molecular basis or consequences of this is unknown (Carpousis, 2007). Corroborating this idea in *E. coli*, Enolase has been found to mediate the rapid degradation of a glucose transporter, *ptsG*, in response to phosphosugar stress (Morita et al., 2004). Additionally, Enolase has been found to be able to directly bind RNA in yeast and mammalian cells (Entelis et al., 2006; Hernández-Pérez et al., 2011).

In comparison to *E. coli*, much less is known about the degradosome in Gram-positive species. Most do not contain a direct RNase E orthologue, including *B. subtilis* and *S. aureus*. In *B. subtilis*, a proposed degradosome is composed of the nucleases RNase J1, RNase J2, PNPase and RNase Y, alongside cofactors CshA (an RNA helicase), Enolase and phosphofructokinase (Pfk) (Commichau et al., 2009; Lehnik-Habrink et al., 2010). RNase J1 and J2 are both bifunctional nucleases with endo- and 5'-3' exonuclease activity and may act to initiate mRNA degradation, while CshA is hypothesised to melt structured RNA regions in a manner similar to RhlB (Roux

et al., 2011). RNase Y is known to be membrane-bound in both *B. subtilis* and *S. aureus* and is thought to be the central anchor to the complex (Cascante-Estepa et al., 2016, Khemici et al., 2015).

Only limited analyses have been performed on the *S. aureus* degradosome in the form of bacterial two-hybrid assays and *in vitro* binding experiments. This led to the conclusion that RNase J1 and J2 may act as a heteromultimer, and that J1 may recruit PNPase and CshA. Additionally, CshA may bind RNase Y, Enolase, Pfk, and an additional cellular RNase predicted to be involved in bulk mRNA degradation called RnpA (Olson et al., 2011; Roux et al., 2011). Thus, the proposed degradosome is composed mainly of RNase J1 and J2 acting to initiate bulk mRNA degradation (potentially alongside RnpA) while RNase Y may target a specific subset of RNAs, such as those with specific structural features (Roux et al., 2011).

Interestingly, CshA has been found to protect essential RNAs from degradation by the toxin RNase MazF in *S. aureus* (Kim et al., 2016). This RNase is part of a toxin-antitoxin system, where its ribonuclease activity is inhibited by the labile MazE antitoxin. Whether this toxin RNase has any physiological benefit to the cell is still under debate, but one hypothesis suggests that it acts to clear unnecessary RNAs within the cell at times of extreme stress, and may also induce drug-tolerance and biofilm formation (Schuster and Bertram, 2016). Interestingly, overexpression of MazF induced cellular stasis but not cell death, and therefore it was proposed that MazF activity inhibits growth and could therefore promote survival (Fu et al., 2009). Transcriptomic analysis following MazF induction showed that mRNAs encoding for virulence proteins such as *hla* and *spa* were degraded, while essential 'housekeeping' genes such as *recA, sarA* and *gyrB* were protected (Fu et al., 2009). Using Northwestern blots, it was found that these proteins were likely bound to a protective RBP and this was identified as being CshA in subsequent studies (Fu et al., 2009; Kim et al., 2016). Deletion of *cshA* results in a loss of viability upon MazF induction and a decrease in the expression of 22 sRNAs (alongside many housekeeping mRNAs) (Kim et al., 2016). This result also suggests that certain sRNAs are essential in cellular survival in times of duress. Thus, CshA may have multiple roles within the cell; it may act as part of the degradosome to unwind structured RNA regions but may also exist as a cytosolic pool independent of the degradosome where it can protect key RNAs from degradation.

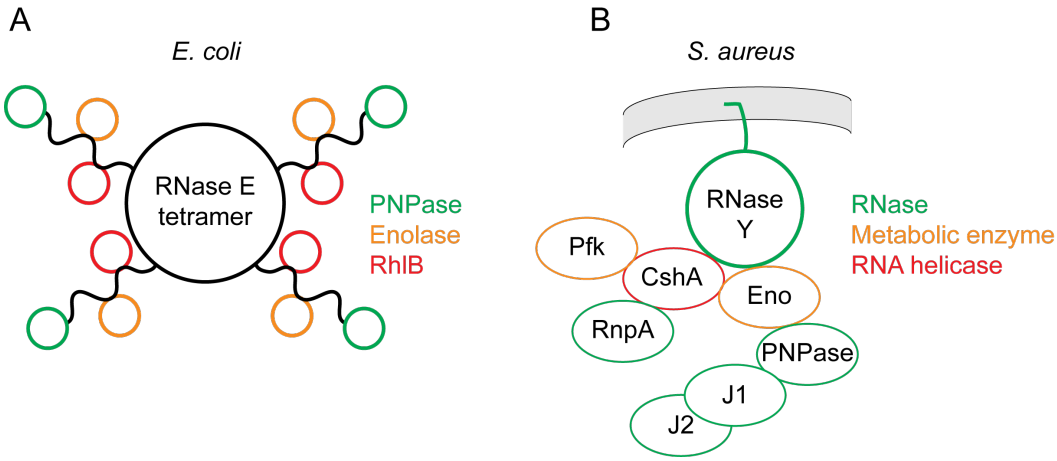

Fig. 1.6: The degradosome of *E. coli* and proposed degradosome of *S. aureus*. **A.** *E. coli* degradosome is built upon a tetramer RNase E, which acts as a central scaffold for PNPase, Enolase and RhlB via its unstructured C-terminal domains. In addition to this, RNase E initiates bulk mRNA decay. **B.** The *S. aureus* degradosome is not as well understood as that in *E. coli*. It is potentially based built upon the membrane-bound RNase Y, which can then interact with Enolase and CshA. Enolase can then recruit PNPase, which in turn recruits the RNase J1 and J2 heteromer. CshA can interact with both Pfk and RnpA. Data based off of (Khemici et al., 2015; Roux et al., 2011).

1.9 sRNAs and their integration into regulatory networks

Given that sRNAs are capable of regulating the stability and translational efficiency of mRNAs, they are well placed to act in concert with transcription factors to drive and shape signalling responses.

Typically, a protein transcription factor or two-component system is used to sense a change in the environment or detect a stimulus. These will then stimulate or repress the expression of target genes, including sRNAs. This regulation exerted on the sRNAs then drives secondary responses in their targets. Importantly, some sRNAs can then feedback, positively or negatively, on the original transcription factor in order to shape the timings and dynamics of the response. Overall, these integrated transcription factor-sRNA networks can range from simple single input module (SIM) and negative feedback loops (NFLs) motifs to more complex dense overlapping regulons (DORs) and feedforward loops (FFLs) (Beisel and Storz, 2010).

Simple architectures; SIMs and NFLs

The SIM motif is the simplest form of a regulatory circuit that contains an sRNA. Here, a transcription factor drives the expression of an sRNA, and this sRNA in turn exerts regulatory effects on mRNA targets (Figure 1.8A). As such, the transcription factor has an indirect effect on the targets. As an example from *S. aureus*, the transcription factor AgrA drives expression of RNAIII, which then regulates numerous targets; exemplified here with its positive regulation of *hla* and negative regulation of *rot*. Computational predictions also suggest that a SIM can create a hierarchical response, where the mRNAs that have extensive base-pairing with the sRNA are regulated first and then those with less pairings afterwards. This would then add a temporal aspect to the response and also a dose-dependent effect where only a strong stimulus would induce enough sRNA for the lesser targets to be regulated (Beisel and Storz, 2010).

When the sRNA can repress its own inducing transcription factor, a negative feedback loop is created. This has the benefit of exerting tighter control over the sRNA and prevents over-accumulation which may have hazardous effects on the cell. Additionally, cell-to-cell variation is reduced as the inducing TF is inhibited at increasing rates as it accumulates, while low levels of TF are able to drive the expression of the sRNA relatively unopposed. Finally, such loops also produce a timed response that encode for their own shutdown; this is useful when inducing a stress response and then shutting it down when the stress is resolved. This is

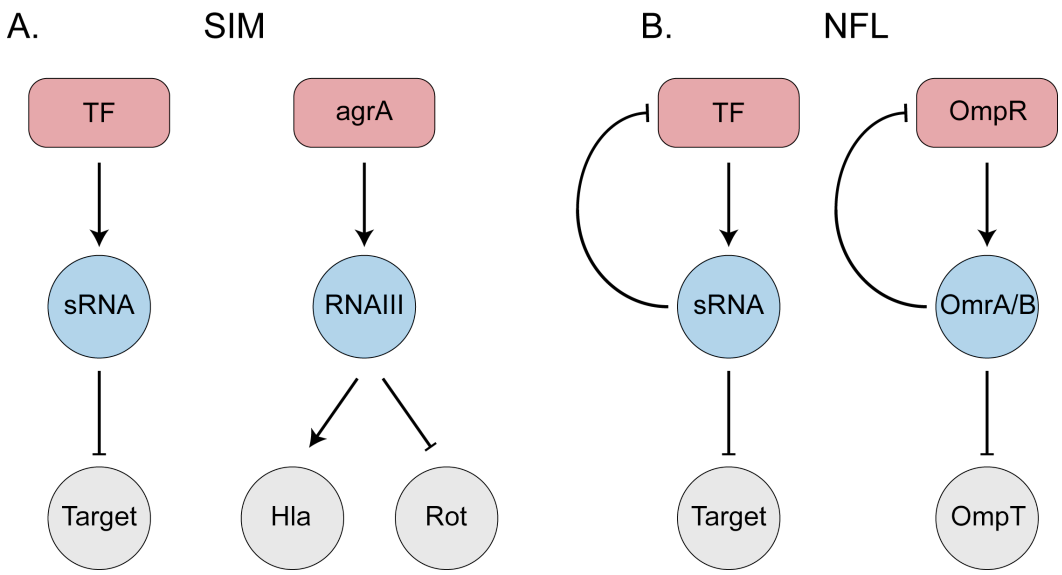

Fig. 1.7: The architecture of single input modules (SIMs) and negative feedback loops (NFLs) utilising sRNAs. **A.** SIMs consist of a single transcription factor that drives expression of an sRNA, which then regulates its targets directly. **B.** In an NFL, the sRNA then represses the translation of its own inducing transcription factor. Modified from Beisel and Storz, 2010.

represented in Figure 1.8B, where the transcription factor OmpR is induced by membrane stress in *E. coli*. OmpR then induces the expression of two sRNAs, OmrA and OmrB, which in turn negatively regulate several outer membrane proteins, including *ompT* (Guillier and Gottesman, 2006).

Complex architectures: DORs and FFLs

DORs represent a collection of multiple, overlapping SIMs. These are used to incorporate multiple environmental stimuli and buffer them against each other. As such, DORs may be used when controlling the expression rates of genes involved in multiple biological processes (Beisel and Storz, 2010). Figure 1.9A shows an example of a DOR involved in controlling the composition of the outer membrane, in particular the OMPs, in *E. coli* and *S. typhimurium*. These OMPs are critical determinants of cell viability as they determine what molecules can enter and exit the cell, have a profound effect on membrane integrity and are recognised by phages (Vogel, 2009). As such, their regulation is controlled through multiple stress responsive pathways; σ^E (extracytoplasmic stress response), osmotic shock (OmpR) and host cell invasion (HilD). Each of these pathways mediates its effect through an sRNA, which in return negatively regulates a particular set of OMPs. As such, this DOR mediates a coordinated response through a common set of sRNAs in order to respond to a variety of membrane stresses, ultimately tuning the outer membrane porin composition.

FFLs are composed of two regulators, either sRNA or TF, controlling the expression of a common gene. Every interaction in the circuit can be either positive or negative, and this yields eight different network configurations. These eight can be divided into two outcomes; coherent and incoherent. The coherent loops are formed when both regulators jointly regulate their target in the same manner. Incoherent loops are formed when the arms of regulation are in opposition to each other.

In Figure 1.9B, two examples of coherent FFLs are given. In *S. aureus*, RNAIII stabilises the transcription factor *mgrA* and promotes the translation of a toxin, *hla* (Gupta et al., 2015; Morfeldt et al., 1995). MgrA also acts in concert to drive expression of *hla*. Thus, the concerted act of both RNAIII and MgrA drives Hla expression faster than if they acted alone, and therefore toxin production can be achieved at the maximal possible velocity. Another example of an FFL is represented in Figure 1.9B from *E. coli*. It has been proposed that when in the gut, *E. coli* must populate its outer membrane with OmpC, which has a relatively small pore size (in comparison to porins such as OmpF) in order to exclude toxic bile salts. To facilitate a swift transition, OmpR acts to induce production of *ompC* mRNA, but also represses the expression of a repressor of OmpC, the sRNA MicA (Yoshida et al., 2006). Thus, OmpR activates OmpC and inactivates its repressor, supporting accumulation of OmpC. As such, coherent FFLs drive sustained responses at the maximum possible rate.

On the other hand, incoherent FFLs can act to drive pulses of expression, where a response is induced but ultimately timed to inactivate itself. They can also act to ensure that a particular response is only committed when absolutely appropriate, effectively acting to buffer transcriptional noise or low stimuli. The latter has been demonstrated in *E. coli*, where σ^E drives expression of both RybB and *htrG* in times of stress (Asakura and Kobayashi, 2009; Gogol et al., 2011). RybB is an sRNA that represses the translation of several mRNAs, including *htrG* which induces cell autolysis, likely through disruption of the cell membrane (Asakura and Kobayashi, 2009). In low-stress conditions, there is sufficient RybB to effectively repress *htrG* translation. However, when the stress is strong enough (such as sustained nutrient starvation), sufficient *htrG* is produced in order to overcome the RybB repression, and this induces cell death. As such, the population of *E. coli* benefits from reduced competition and may survive.

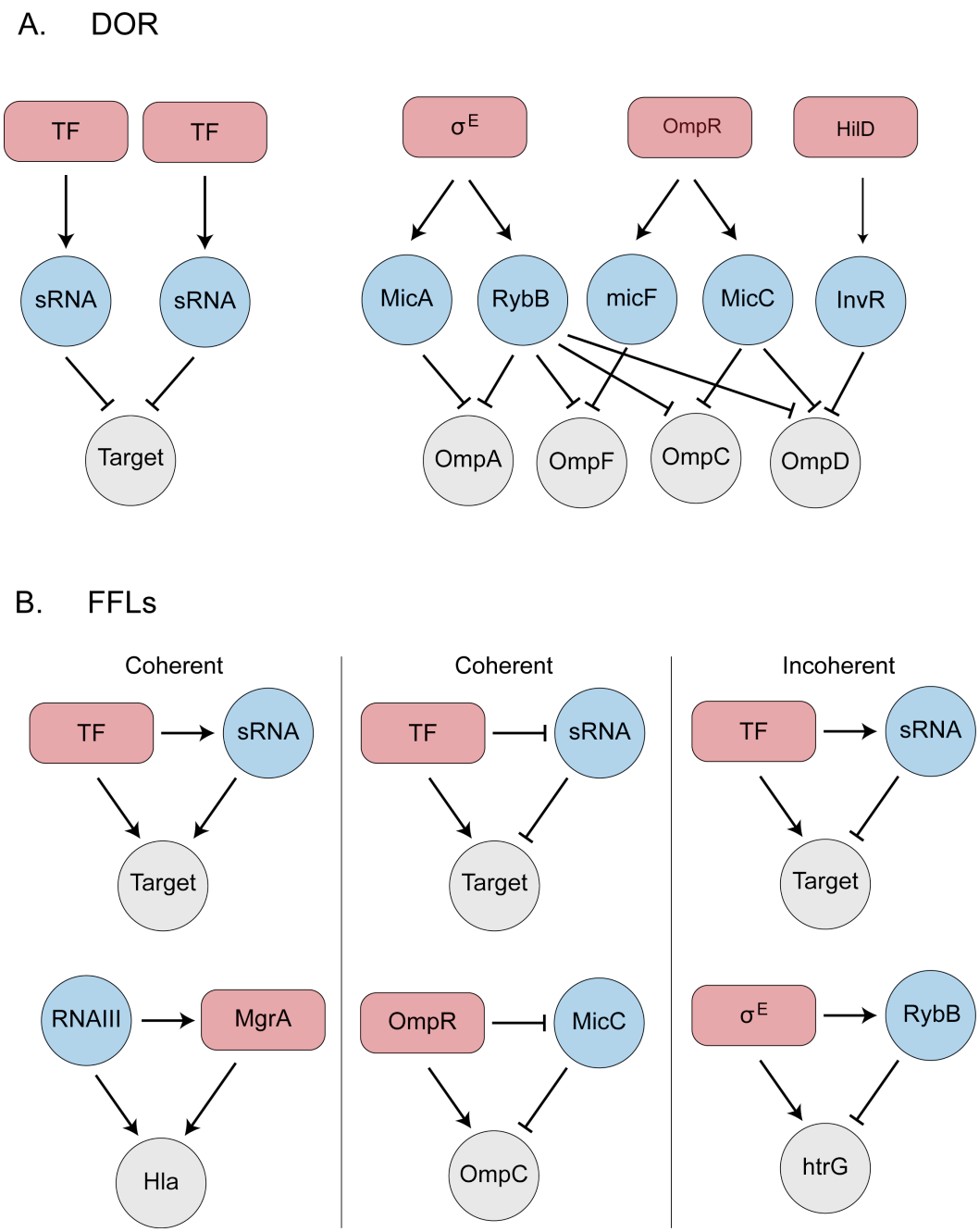

Fig. 1.8: The architecture of dense overlapping regulons (DORs) and feedforward loops (FFLs). **A.** DORs consist of several individual SIMs which converge on common targets. This allows different signalling pathways to act in concert or to buffer them against each other. **B.** FFLs are separated into coherent and incoherent architectures. Coherent loops drive repression or stimulation from both a transcriptional level and translational level, while incoherent loops contain properties of both repression and stimulation. Modified from Beisel and Storz, 2010.

1.10 sRNA sponges

RNA sponges, or decoys, are species which bind to and antagonise the activity of regulatory RNAs (Ebert et al., 2007). As such, sponges can act as indirect regulators of sRNA targets. However, unlike most sRNA-mRNA interactions, sRNA sponging interactions do not always lead to degradation. The first description of an RNA sponge was in plants, where phosphate starvation acts to induce expression of the ncRNA IPS1 and this in turn directly binds miR-399 to inhibit its activity. Indeed, overexpression of IPS1 leads to accumulation of miR-399 targets (Franco-Zorrilla et al., 2007). Interestingly, both miR-399 and IPS1 are induced by phosphate starvation and thus it is their relative transcription and degradation rates that act to dictate target expression.

The first example of sRNA sponging in prokaryotes was found in *S. typhimurium*. The constitutively expressed sRNA ChiX acts to repress the translation of *chiP* in the absence of chitooligosaccharides. *chiP* encodes for a chitoporin required for the uptake of these sugars, and upon their detection, the *chb* operon is transcribed. This operon is involved in chitooligosaccharide metabolism and is subsequently processed and one of the produced intercistronic fragments then acts as a sponge against ChiX. This subsequently relieves *chiP* of its repression (Figueroa-Bossi et al., 2009). Similarly, in *E. coli*, the 3' UTR of *pspG* acts as a sponge against the sRNA Spf (Melamed et al., 2016).

Sponges are also capable of affecting the activity of a whole regulon, and therefore numerous pathways concurrently. In *S. typhimurium* and *E. coli*, the sRNA GcvB is thought to regulate up to 1% of all mRNAs, including major ABC transporters, amino acid biosynthesis proteins and major transcription factors (Vanderpool, 2011). However, the processing of one of GcvB's target polycistrons, *gltIJKL*, releases an sRNA sponge, SroC, that binds to GcvB and induces its degradation by RNase E. Thus, GcvB-mediated destruction of the *gltIJKL* polycistron acts to eventually downregulate GcvB, which firstly facilitates *gltIJKL* translation. This also has wider implications for the cell as other GcvB targets are then relieved of their regulation (Miyakoshi et al., 2015).

sRNA sponges as buffers against noise

As sRNAs can act to cause widespread changes within the cell, their expression and activity must be tightly regulated. An example of such behaviour comes from *E. coli*.

External transcribed spacer (ETS) sequences in tRNAs are excised and usually degraded immediately. However, an ETS (3'ETSleuZ) produced from RNase E-mediated processing of the glyW-cysT-leuZ pre-tRNA base pairs with two sRNAs, RyhB and RybB (Lalaouna et al., 2015). Critically, the production of 3'ETSleuZ does not respond to the intracellular levels of RyhB or RybB and instead remains constant; as such, it is predicted to act as a buffer to antagonise transcriptional noise of these sRNAs. Thus, in order for RyhB or RybB to have functionality within the cell, their expression must pass a critical threshold which can only be induced by a continuous and strong stimulus. Overall, the action of such sRNA sponges decreases extrinsic and intrinsic noise of signalling responses, ensuring they are only activated when absolutely appropriate. This behaviour is represented graphically in Figure 1.10A.

Sponges as regulators of sRNA activity

Sponges can also act to shape the dynamics of sRNA responses. As discussed previously, several sRNAs in *S. aureus* have been shown to be downregulated by stress and therefore their regulation is likely to be important for adaptation. As shown in Figure 1.10B, the induction of an antagonistic sRNA sponge could help to downregulate the activity of an sRNA quicker than simply turning off its transcription (and could also induce its degradation through recruited RNases).

Sponges can also be used to generate 'pulses' of sRNA activity, where a stimulus induces both the sRNA and also its antagonistic sponge, perhaps utilising a slight delay. As shown in Figure 1.10C, the sRNA would be able to exert its activity while the sponge lags behind in expression/activity. However, when the sponge manages to reach a level that can effectively compete with the sRNA pool, the sRNA response is downregulated. Such a network could be important for controlling sRNAs with widespread effects across multiple pathways to ensure that they are effectively controlled and that they are active only across a limited timeframe.

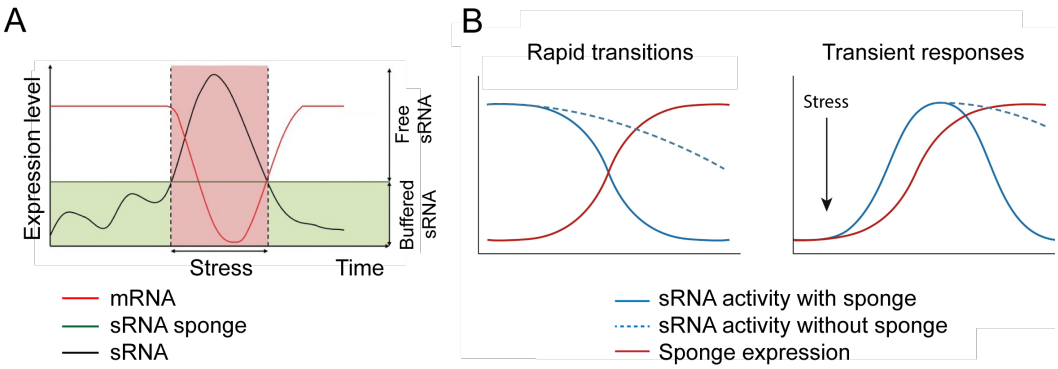

Fig. 1.9: The role of sRNA sponges in shaping gene expression profiles. **A.** sRNA sponges can act to buffer transcriptional noise of an sRNA, preventing unwanted repercussions and ensuring that responses only occur upon a strong stimulus. **B.** sRNA sponges can act to repress sRNA activity, leading to either rapid inactivation (left hand side) or a timed response (right hand side). Figure A adapted from Lalaouna et al, 2015, and B and C from Ebert and Sharp, 2010.

Currently, only a few examples of sRNA sponging have been demonstrated in bacteria, with the most well understood described above. However, it is becoming increasingly clear that there remains much to be discovered. In enterohemorrhagic *E. coli*, a study identified around 150 unique sRNA – sRNA interactions, suggesting that sponging and induced degradation of sRNAs by sponges could be a common phenomenon (Waters et al., 2017). Additionally, in *S. aureus*, RsaOG has been found to bind RsaE, RsaD and RsaG, although biological significance has yet to be found (Bronesky et al., 2018; Rochat et al., 2018).

1.11 Identifying sRNA – target interactions *in vivo*

Depending on one's definition of what constitutes an sRNA, there are between 50 and 500 in *S. aureus*. However, the targets of only a few sRNAs have been identified and validated, and even less is known about the dynamics of these interactions. Under what conditions does the interaction take place? How does the interaction respond to cellular stresses and external stimuli? How does the interaction confer a survival or growth advantage to the cell?

As described in Section 1.8, *in silico* predictions have been critical in identifying both sRNAs and their targets. However, predicting targets of sRNAs is difficult as the interaction region is typically small. Additionally, predictions remain just that until properly validated *in vivo*. *In silico* methods are also susceptible to false positives and negatives, where they incorrectly predict interactions where they don't exist and miss out those that do. Finally, a computational prediction offers no concrete information about the nature of the regulation, be it positive or negative. To gain genuine insight

into the global functions of regulatory RNAs, one must use techniques to capture their interactomes.

CLASH

Several techniques exist to capture sRNA interactomes on a global scale, including CLASH, MARIO, PARIS, and RILseq (Kudla et al., 2011; Lu et al., 2018; Melamed et al., 2016; Nguyen et al., 2016). Of interest to this thesis is CLASH (Crosslinking, ligation and sequencing of hybrids), a technique which uses an RBP as bait for capturing bound RNAs. Using a proximity-based ligation of basepaired RNAs, CLASH can identify RNAs in complex with each other. CLASH has already been utilised in yeast, mammalian cells and *E. coli* (Helwak and Tollervey, 2014; Iosub et al., 2018; Kudla et al., 2011; Waters et al., 2017). As such, it is an appropriate technique to discover novel sRNA – target interactions in *S. aureus*.

Briefly, CLASH involves expressing an RBP of interest tagged with a tandem affinity tag, namely an HTP (His6-TEV cleavage site-Protein A) or HTF (His6-TEV cleavage site-FLAG) tag (Figure 1.11). The recombinant cells are then exposed to 264 nm UV radiation in order to induce covalent bonds between RBPs and their cognate RNAs. The protein is subsequently purified using either IgG-agarose beads or anti-FLAG magnetic beads (dependent upon if HTP or HTF was used) for the primary capture, eluted using TEV protease, and then captured onto nickel-agarose beads under highly denaturing conditions. This latter step is crucial as the stringency enacted by the denaturing conditions ensures that covalently cross-linked RNAs are enriched and reduces noise by removing background binding to the beads. A fraction of the RBPs bound to the beads will have bound RNA-RNA duplexes and these duplexes can be ligated together to form a single molecule through RNA ligation. Sequencing adaptors are then ligated onto the ends of the RNAs, and then they are eluted from the protein, reverse transcribed, amplified by PCR and sequenced (Helwak and Tollervey, 2014; Kudla et al., 2011). The produced sequencing data will then contain so-called 'chimeric' reads which represent two distinct RNAs that were previously duplexed together. These chimeric reads can be detected using the Hyb pipeline (Travis et al., 2014). The read sequences can then be folded *in silico* in order to gain insight into how the RNA fragments form a duplex.

It is important to note that CLASH is a slightly modified version of an older protocol termed CRAC (Crosslinking and analysis of cDNAs; Figure 1.11) (Granneman et al., 2009). CRAC was developed as an improved CLIP protocol, utilising the HTP/HTF tag to reduce background noise as described above. The main difference between CRAC and CLASH is that CLASH utilises a dedicated intermolecular ligation step in order to induce RNA-RNA ligation, while CRAC foregoes this step. The enzymatic steps in CLASH are usually also done at lower temperatures to ensure that short RNA-

RNA interactions are preserved during the whole procedure (Helwak and Tollervey, 2014). As a result, the sequencing data from CRAC is a list of single reads which represent single RNAs bound to the RBP. However, as not all captured RBPs are bound to RNA duplexes and RNA-RNA ligation is very inefficient, CLASH will also give 'CRAC data', i.e. single reads. Representing this fact, the data obtained from performing CLASH on RNase E in *E. coli* contained chimeric reads at a rate of just 1% and around 99% single reads (Waters et al., 2017). Thus, performing CLASH allows one to detect RNA-RNA interactions but also uncover the RNA targets of a chosen RBP.

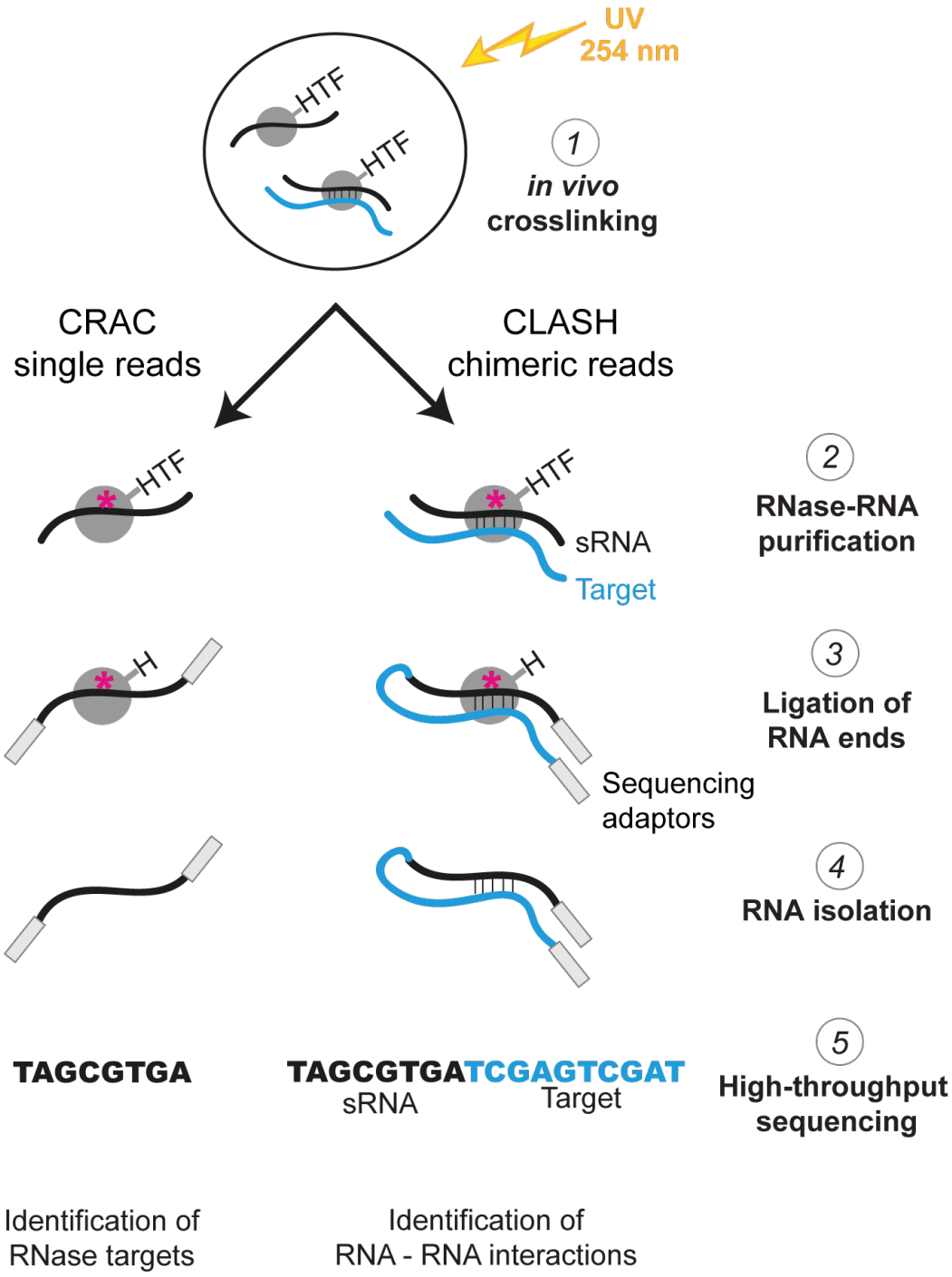

Fig. 1.10: The CRAC and CLASH methodologies. Both utilise an HTF-tagged RBP as bait, which is cross-linked to its bound RNAs through UV irradiation. The RBP is then captured under highly denaturing conditions following lysis. In CRAC, sequencing adaptors are ligated onto the bound RNAs, which are then isolated, reverse transcribed, and sequenced. This produces 'single read' data, where the RNAs bound to the RBP are identified. CLASH utilises the same methodologies, but includes a ligation step where RNA duplexes, such as sRNA-target interactions, are joined to create a single entity. Thus, CLASH produces chimeric reads, which can identify RNA interacting partners.

MAPS

Once an sRNA has been properly identified, one can tag the genetic sequence with a variety of RNA-based tags. This allows the RNA and its interacting partners to be pulled down in a very similar manner to that of tagged proteins (Carrier et al., 2016). One technique that has gained traction in the RNA field is termed MAPS (MS2-affinity purification coupled with RNA sequencing') (Lalaouna et al., 2015). Here, the sRNA of interest is genetically tagged with the MS2 stem-loop aptamer RNA. This tag is derived from the bacteriophage genome and its usefulness stems from the fact that it will bind tightly the MS2 coat protein (this protein itself tagged, traditionally with maltose binding protein). Thus, one can generate a cell lysate from a tagged strain and apply this to an amylose resin on which MBP-MS2 has been immobilised. The column can then be washed, the MS2::sRNA-target complexes eluted with maltose, and then reverse transcribed for RNA sequencing. Additionally, LC-MS/MS can be performed on the eluates in order to identify protein binding partners of the sRNA, which can help to identify e.g. chaperones required or RNases which could imply a degradation event (Lalaouna et al., 2015).

One potential bottleneck of the MAPS technology is that it requires prior knowledge of an RNA to be an effective bait. Additionally, it operates on a single RNA scale; although this can give valuable insight into the functionality of a chosen RNA, it does not give a truly global snapshot of the sRNA networks operating within a cell *in vivo*. As discussed in Section 1.8, MAPS has already been applied to RsaA, RsaE and RsaC in *S. aureus* to good effect, demonstrating its viability in this organism (Bronesky et al., 2019; Lalaouna et al.; Tomasini et al., 2017).

1.12 Investigating the role of sRNAs in mediating adaptation to stress

Several Rsa, Spr and Teg transcripts have already been demonstrated to exhibit changed expression in response to antibiotic, oxidative, pH and thermal stresses (Section 1.8). Although these research fields have not been exhausted, it is even less clear how sRNAs shape the transcriptome in virulent conditions. As detailed in Section 1.7 and 1.8, RNAIII and SprD have been shown to be crucial players in regulating virulence-related transcripts. However, the studies that elucidated these interactions were carried out using static models that do not accurately recapitulate the environment when *S. aureus* becomes virulent; they relied on techniques such as genetic knockouts followed by northern/western blots in rich culture mediums, or *in vitro* methods such as ribosomal toeprinting. It still remains unclear how sRNAs

immediately shape the transcriptome of *S. aureus* on a global level in response to virulent conditions.

CLASH is well suited for answering such questions. The Granneman lab has collaborated with a private company, UVO-3, to create a system that facilitates the crosslinking of cells on very short timescales followed by rapid harvesting through vacuum filtration. This system has already demonstrated its utility in studying how the yeast *S. cerevisiae* adapts to glucose deprivation on minute timescales (Nues et al., 2017). This system was used here in combination with CLASH to study how sRNAs shape *S. aureus*' adaptation to conditions that mimic the stresses incurred when entering the human bloodstream.

A study by Mäder et al, 2016, used tiling arrays to examine the transcriptome architecture of cells grown under various conditions. They found that a commercially available medium, RPMI-1640, induces a very similar response in *S. aureus* as human blood plasma, with particular overlap in iron-responsive genes (Mäder et al., 2016). As such, CLASH was performed on cells exposed to this medium to examine how sRNAs shape the response to the stresses induced when *S. aureus* enters the human bloodstream. By identifying the targets of sRNAs, CLASH acted as a first step in elucidating the functionality of sRNAs that are active in these conditions, allowing further experimentation and validation.

1.13 Bibliography

Abraham NM, Jefferson KK. 2012. *Staphylococcus aureus* clumping factor B mediates biofilm formation in the absence of calcium. Microbiology (Reading, Engl) 158:1504–1512. doi:10.1099/mic.0.057018-0

Abu-Qatouseh LF, Chinni SV, Seggewiß J, Proctor RA, Brosius J, Rozhdestvensky TS, Peters G, von Eiff C, Becker K. 2010. Identification of differentially expressed small non-protein-coding RNAs in *Staphylococcus aureus* displaying both the normal and the small-colony variant phenotype. Journal of Molecular Medicine 88:565–575. doi:10.1007/s00109-010-0597-2

Asakura Y, Kobayashi I. 2009. From damaged genome to cell surface: transcriptome changes during bacterial cell death triggered by loss of a restriction–modification gene complex. Nucleic Acids Research 37:3021–3031. doi:10.1093/nar/gkp148

Balaban N, Novick RP. 1995. Translation of RNAIII, the *Staphylococcus aureus* agr regulatory RNA molecule, can be activated by a 3'-end deletion. FEMS Microbiol Lett 133:155–161. doi:10.1111/j.1574-6968.1995.tb07877.x

Bandyra KJ, Said N, Pfeiffer V, Górna MW, Vogel J, Luisi BF. 2012. The seed region of a small RNA drives the controlled destruction of the target mRNA by the endoribonuclease RNase E. Mol Cell 47:943–953. doi:10.1016/j.molcel.2012.07.015

Beaume M, Hernandez D, Farinelli L, Deluen C, Linder P, Gaspin C, Romby P, Schrenzel J, Francois P. 2010. Cartography of Methicillin-Resistant *S. aureus* Transcripts: Detection, Orientation and Temporal Expression during Growth Phase and Stress Conditions. PLOS ONE 5:e10725. doi:10.1371/journal.pone.0010725

Beisel CL, Storz G. 2010. Base pairing small RNAs and their roles in global regulatory networks. FEMS Microbiol Rev 34:866–882. doi:10.1111/j.1574-6976.2010.00241.x

Bekeredjian-Ding I, Stein C, Uebele J. 2017. The Innate Immune Response Against *Staphylococcus aureus* Microbiology, Pathology, Immunology, Therapy and Prophylaxis, Current Topics in Microbiology and Immunology. Cham: Springer International Publishing. pp. 385–418. doi:10.1007/82_2015_5004

Bohn C, Rigoulay C, Bouloc P. 2007. No detectable effect of RNA-binding protein Hfq absence in *Staphylococcus aureus*. BMC Microbiol 7:10. doi:10.1186/1471-2180-7-10

Bohn C, Rigoulay C, Chabelskaya S, Sharma CM, Marchais A, Skorski P, Borezée-Durant E, Barbet R, Jacquet E, Jacq A, Gautheret D, Felden B, Vogel J, Bouloc P. 2010. Experimental discovery of small RNAs in *Staphylococcus aureus* reveals a riboregulator of central metabolism. Nucleic Acids Res 38:6620–6636. doi:10.1093/nar/gkq462

Boisset S, Geissmann T, Huntzinger E, Fechter P, Bendridi N, Possedko M, Chevalier C, Helfer AC, Benito Y, Jacquier A, Gaspin C, Vandenesch F, Romby P. 2007. *Staphylococcus aureus* RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. Genes Dev 21:1353–1366. doi:10.1101/gad.423507

Bronesky D, Desgranges E, Corvaglia A, François P, Caballero CJ, Prado L, Toledo-Arana A, Lasa I, Moreau K, Vandenesch F, Marzi S, Romby P, Caldelari I. 2019. A multifaceted small RNA modulates gene expression upon glucose limitation in *Staphylococcus aureus*. The EMBO Journal 38:e99363. doi:10.15252/embj.201899363

Bronesky D, Desgranges E, Corvaglia A, François P, Caballero CJ, Prado L, Toledo-Arana A, Lasa I, Moreau K, Vandenesch F, Marzi S, Romby P, Caldelari I. 2018. A dual sRNA in *Staphylococcus aureus* induces a metabolic switch responding to glucose consumption. bioRxiv 278127. doi:10.1101/278127

Bronesky D, Wu Z, Marzi S, Walter P, Geissmann T, Moreau K, Vandenesch F, Caldelari I, Romby P. 2016. *Staphylococcus aureus* RNAIII and Its Regulon Link Quorum Sensing, Stress Responses, Metabolic Adaptation, and Regulation of Virulence Gene Expression. Annual Review of Microbiology 70:299–316. doi:10.1146/annurev-micro-102215-095708

Burgess RR. 1969. Separation and characterization of the subunits of ribonucleic acid polymerase. J Biol Chem 244:6168–6176.

Burian M, Rautenberg M, Kohler T, Fritz M, Krismer B, Unger C, Hoffmann WH, Peschel A, Wolz C, Goerke C. 2010. Temporal expression of adhesion factors and activity of global regulators during establishment of *Staphylococcus aureus* nasal colonization. J Infect Dis 201:1414–1421. doi:10.1086/651619

Carpousis AJ. 2007. The RNA Degradosome of *Escherichia coli*: An mRNA-Degrading Machine Assembled on RNase E. Annu Rev Microbiol 61:71–87. doi:10.1146/annurev.micro.61.080706.093440

Carpousis AJ. 2002. The *Escherichia coli* RNA degradosome: structure, function and relationship in other ribonucleolytic multienzyme complexes. Biochem Soc Trans 30:150–155.

Carrier M-C, Lalaouna D, Massé E. 2016. A game of tag: MAPS catches up on RNA interactomes. RNA Biol 13:473–476.
doi:10.1080/15476286.2016.1156830

Castro SL, Nelman-Gonzalez M, Nickerson CA, Ott CM. 2011. Induction of attachment-independent biofilm formation and repression of Hfq expression by low-fluid-shear culture of *Staphylococcus aureus*. Appl Environ Microbiol 77:6368–6378. doi:10.1128/AEM.00175-11

CDC. 2018. The biggest antibiotic-resistant threats in the U.S. Centers for Disease Control and Prevention. https://www.cdc.gov/drugresistance/biggest_threats.html Centers for Disease Control and Prevention (CDC). 1999. Four pediatric deaths from community-acquired methicillin-resistant *Staphylococcus aureus* — Minnesota and North Dakota, 1997-1999. MMWR Morb Mortal Wkly Rep 48:707–710.

Chabelskaya S, Bordeau V, Felden B. 2014. Dual RNA regulatory control of a *Staphylococcus aureus* virulence factor. Nucleic Acids Res 42:4847–4858. doi:10.1093/nar/gku119

Chabelskaya S, Gaillot O, Felden B. 2010. A *Staphylococcus aureus* Small RNA Is Required for Bacterial Virulence and Regulates the Expression of an Immune-Evasion Molecule. PLOS Pathogens 6:e1000927. doi:10.1371/journal.ppat.1000927

Chao Y, Papenfort K, Reinhardt R, Sharma CM, Vogel J. 2012. An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs. EMBO J 31:4005–4019. doi:10.1038/emboj.2012.229

Chao Y, Vogel J. 2010. The role of Hfq in bacterial pathogens. Current Opinion in Microbiology 13:24–33. doi:10.1016/j.mib.2010.01.001

Chen PR, Bae T, Williams WA, Duguid EM, Rice PA, Schneewind O, He C. 2006. An oxidation-sensing mechanism is used by the global regulator MgrA in Staphylococcus aureus. Nat Chem Biol 2:591–595. doi:10.1038/nchembio820

Chevalier C, Boisset S, Romilly C, Masquida B, Fechter P, Geissmann T, Vandenesch F, Romby P. 2010. *Staphylococcus aureus* RNAIII Binds to Two Distant Regions of coa mRNA to Arrest Translation and Promote mRNA Degradation. PLOS Pathogens 6:e1000809. doi:10.1371/journal.ppat.1000809

Chevalier C, Huntzinger E, Fechter P, Boisset S, Vandenesch F, Romby P, Geissmann T. 2008. *Staphylococcus aureus* endoribonuclease III purification and properties. Meth Enzymol 447:309–327. doi:10.1016/S0076-6879(08)02216-7

Chunhua M, Yu L, Yaping G, Jie D, Qiang L, Xiaorong T, Guang Y. 2012. The expression of LytM is down-regulated by RNAIII in *Staphylococcus aureus*. Journal of Basic Microbiology 52:636–641. doi:10.1002/jobm.201100426

Commichau FM, Rothe FM, Herzberg C, Wagner E, Hellwig D, Lehnik-Habrink M, Hammer E, Völker U, Stülke J. 2009. Novel activities of glycolytic enzymes in *Bacillus subtilis*: interactions with essential proteins involved in mRNA processing. Mol Cell Proteomics 8:1350–1360. doi:10.1074/mcp.M800546-MCP200

Crosby HA, Schlievert PM, Merriman JA, King JM, Salgado-Pabón W, Horswill AR. 2016. The Staphylococcus aureus Global Regulator MgrA Modulates Clumping and Virulence by Controlling Surface Protein Expression. PLOS Pathog 12:e1005604. doi:10.1371/journal.ppat.1005604

Csárdi G, Franks A, Choi DS, Airoldi EM, Drummond DA. 2015. Accounting for Experimental Noise Reveals That mRNA Levels, Amplified by Post-Transcriptional Processes, Largely Determine Steady-State Protein Levels in Yeast. PLOS Genetics 11:e1005206. doi:10.1371/journal.pgen.1005206

Cucarella C, Solano C, Valle J, Amorena B, Lasa I, Penadés JR. 2001. Bap, a *Staphylococcus aureus* surface protein involved in biofilm formation. J Bacteriol 183:2888–2896. doi:10.1128/JB.183.9.2888-2896.2001

Dantes R, Mu Y, Belflower R, Aragon D, Dumyati G, Harrison LH, Lessa FC, Lynfield R, Nadle J, Petit S, Ray SM, Schaffner W, Townes J, Fridkin S. 2013. National Burden of Invasive Methicillin-Resistant *Staphylococcus aureus* Infections, United States, 2011. JAMA Intern Med 173:1970–1978. doi:10.1001/jamainternmed.2013.10423

Deora R, Misra TK. 1996. Characterization of the primary sigma factor of Staphylococcus aureus. J Biol Chem 271:21828–21834. doi:10.1074/jbc.271.36.21828

Diekema DJ, Richter SS, Heilmann KP, Dohrn CL, Riahi F, Tendolkar S, McDanel JS, Doern GV. 2014. Continued emergence of USA300 methicillin-resistant *Staphylococcus aureus* in the United States: results from a nationwide surveillance study. Infect Control Hosp Epidemiol 35:285–292. doi:10.1086/675283

Durand S, Braun F, Helfer A-C, Romby P, Condon C. 2017. sRNA-mediated activation of gene expression by inhibition of 5'-3' exonucleolytic mRNA degradation. eLife 6:e23602. doi:10.7554/eLife.23602

Durand S, Braun F, Lioliou E, Romilly C, Helfer A-C, Kuhn L, Quittot N, Nicolas P, Romby P, Condon C. 2015. A Nitric Oxide Regulated Small RNA Controls Expres-

sion of Genes Involved in Redox Homeostasis in *Bacillus subtilis*. PLOS Genetics 11:e1004957. doi:10.1371/journal.pgen.1004957

Durand S, Gilet L, Bessières P, Nicolas P, Condon C. 2012. Three Essential Ribonucleases—RNase Y, J1, and III—Control the Abundance of a Majority of *Bacillus subtilis* mRNAs. PLOS Genetics 8:e1002520. doi:10.1371/journal.pgen.1002520

Ebert MS, Neilson JR, Sharp PA. 2007. MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. Nat Methods 4. doi:10.1038/nmeth1079

Elek S. 1959. Staphylococcus pyogenes and its relation to disease. Livingstone: Edinburgh: E&S.

Enright MC, Day NPJ, Davies CE, Peacock SJ, Spratt BG. 2000. Multilocus Sequence Typing for Characterization of Methicillin-Resistant and Methicillin-Susceptible Clones of *Staphylococcus aureus*. J Clin Microbiol 38:1008–1015. doi:10.1128/JCM.38.3.1008-1015.2000

Entelis N, Brandina I, Kamenski P, Krasheninnikov IA, Martin RP, Tarassov I. 2006. A glycolytic enzyme, Enolase, is recruited as a cofactor of tRNA targeting toward mitochondria in *Saccharomyces cerevisiae*. Genes Dev 20:1609–1620. doi:10.1101/gad.385706

Eyraud A, Tattevin P, Chabelskaya S, Felden B. 2014. A small RNA controls a protein regulator involved in antibiotic resistance in *Staphylococcus aureus*. Nucleic Acids Res 42:4892–4905. doi:10.1093/nar/gku149

Figueroa-Bossi N, Valentini M, Malleret L, Fiorini F, Bossi L. 2009. Caught at its own game: regulatory small RNA inactivated by an inducible transcript mimicking its target. Genes Dev 23:2004–2015. doi:10.1101/gad.541609

Finlay BB, Cossart P. 1997. Exploitation of Mammalian Host Cell Functions by Bacterial Pathogens. Science 276:718–725. doi:10.1126/science.276.5313.718

Flannagan Ronald S., Heit Bryan, Heinrichs David E. 2015. Intracellular replication of *Staphylococcus aureus* in mature phagolysosomes in macrophages precedes host cell death, and bacterial escape and dissemination. Cellular Microbiology 18:514–535. doi:10.1111/cmi.12527

Foster TJ, Geoghegan JA, Ganesh VK, Höök M. 2014. Adhesion, invasion and evasion: the many functions of the surface proteins of *Staphylococcus aureus*. Nature Reviews Microbiology 12:49–62. doi:10.1038/nrmicro3161

Fowler VG, Olsen MK, Corey GR, Woods CW, Cabell CH, Reller LB, Cheng AC, Dudley T, Oddone EZ. 2003. Clinical Identifiers of Complicated *Staphylococcus aureus* Bacteremia. Arch Intern Med 163:2066–2072. doi:10.1001/archinte.163.17.2066

Franco-Zorrilla JM, Valli A, Todesco M, Mateos I, Puga MI, Rubio-Somoza I, Leyva A, Weigel D, García JA, Paz-Ares J. 2007. Target mimicry provides a new mechanism for regulation of microRNA activity. Nat Genet 39:1033–1037. doi:10.1038/ng2079

Fu Z, Tamber S, Memmi G, Donegan NP, Cheung AL. 2009. Overexpression of MazFSa in *Staphylococcus aureus* Induces Bacteriostasis by Selectively Targeting mRNAs for Cleavage. Journal of Bacteriology 191:2051–2059. doi:10.1128/JB.00907-08

Geisinger E, Adhikari RP, Jin R, Ross HF, Novick RP. 2006. Inhibition of rot translation by RNAIII, a key feature of agr function. Mol Microbiol 61:1038–1048. doi:10.1111/j.1365-2958.2006.05292.x

Geissmann T, Chevalier C, Cros M-J, Boisset S, Fechter P, Noirot C, Schrenzel J, François P, Vandenesch F, Gaspin C, Romby P. 2009. A search for small noncoding RNAs in *Staphylococcus aureus* reveals a conserved sequence motif for regulation. Nucleic Acids Research 37:7239–7257. doi:10.1093/nar/gkp668

Glaser P, Martins-Simões P, Villain A, Barbier M, Tristan A, Bouchier C, Ma L, Bes M, Laurent F, Guillemot D, Wirth T, Vandenesch F. 2016. Demography and Intercontinental Spread of the USA300 Community-Acquired Methicillin-Resistant *Staphylococcus aureus* Lineage. mBio 7:e02183-15. doi:10.1128/mBio.02183-15

Goering RV, Winters MA. 1992. Rapid method for epidemiological evaluation of Gram-positive *cocci* by field inversion gel electrophoresis. J Clin Microbiol 30:577–580.

Gogol EB, Rhodius VA, Papenfort K, Vogel J, Gross CA. 2011. Small RNAs endow a transcriptional activator with essential repressor functions for single-tier control of a global stress regulon. PNAS 108:12875–12880. doi:10.1073/pnas.1109379108

Gómez MI, Lee A, Reddy B, Muir A, Soong G, Pitt A, Cheung A, Prince A. 2004. *Staphylococcus aureus* protein A induces airway epithelial inflammatory responses by activating TNFR1. Nat Med 10:842–848. doi:10.1038/nm1079

González-Zorn B, Senna JPM, Fiette L, Shorte S, Testard A, Chignard M, Courvalin P, Grillot-Courvalin C. 2005. Bacterial and host factors implicated in nasal carriage

of methicillin-resistant *Staphylococcus aureus* in mice. Infect Immun 73:1847–1851. doi:10.1128/IAI.73.3.1847-1851.2005

Gordon RJ, Lowy FD. 2008. Pathogenesis of Methicillin-Resistant *Staphylococcus aureus* Infection. Clinical Infectious Diseases 46:S350–S359. doi:10.1086/533591

Gosbell IB, Mercer JL, Neville SA, Crone SA, Chant KG, Jalaludin BB, Munro R. 2001. Non-multiresistant and multiresistant methicillin-resistant *Staphylococcus aureus* in community-acquired infections. Med J Aust 174:627–630.

Gottesman S, McCullen CA, Guillier M, Vanderpool CK, Majdalani N, Benhammou J, Thompson KM, FitzGerald PC, Sowa NA, FitzGerald DJ. 2006. Small RNA regulators and the bacterial response to stress. Cold Spring Harb Symp Quant Biol 71:1–11. doi:10.1101/sqb.2006.71.016

Granneman S, Kudla G, Petfalski E, Tollervey D. 2009. Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. PNAS 106:9613–9618. doi:10.1073/pnas.0901997106

Guillet J, Hallier M, Felden B. 2013. Emerging Functions for the *Staphylococcus aureus* RNome. PLoS Pathog 9. doi:10.1371/journal.ppat.1003767

Guillier M, Gottesman S. 2006. Remodelling of the *Escherichia coli* outer membrane by two small regulatory RNAs. Mol Microbiol 59:231–247. doi:10.1111/j.1365-2958.2005.04929.x

Gupta RK, Luong TT, Lee CY. 2015. RNAIII of the *Staphylococcus aureus* agr system activates global regulator MgrA by stabilizing mRNA. PNAS 112:14036–14041. doi:10.1073/pnas.1509251112

Hal SJ van, Jensen SO, Vaska VL, Espedido BA, Paterson DL, Gosbell IB. 2012. Predictors of Mortality in *Staphylococcus aureus* Bacteremia. Clinical Microbiology Reviews 25:362–386. doi:10.1128/CMR.05022-11

Hall-Stoodley L, Costerton JW, Stoodley P. 2004. Bacterial biofilms: from the natural environment to infectious diseases. Nat Rev Microbiol 2:95–108. doi:10.1038/nrmicro821

Helmann JD, Chamberlin MJ. 1988. Structure and function of bacterial sigma factors. Annu Rev Biochem 57:839–872. doi:10.1146/annurev.bi.57.070188.004203

Helwak A, Tollervey D. 2014. Mapping the miRNA interactome by cross-linking ligation and sequencing of hybrids (CLASH). Nature Protocols 9:711–728. doi:10.1038/nprot.2014.043

Hernández-Pérez L, Depardón F, Fernández- Ramírez F, Sánchez-Trujillo A, Bermúdez-Crúz RM, Dangott L, Montañez C. 2011. α-Enolase binds to RNA. Biochimie 93:1520–1528. doi:10.1016/j.biochi.2011.05.007

Herskovitz MA, Bechhofer DH. 2000. Endoribonuclease RNase III is essential in *Bacillus subtilis*. Mol Microbiol 38:1027–1033.

Howden BP, Beaume M, Harrison PF, Hernandez D, Schrenzel J, Seemann T, Francois P, Stinear TP. 2013. Analysis of the Small RNA Transcriptional Response in Multidrug-Resistant *Staphylococcus aureus* after Antimicrobial Exposure. Antimicrob Agents Chemother 57:3864–3874. doi:10.1128/AAC.00263-13

Huntzinger E, Boisset S, Saveanu C, Benito Y, Geissmann T, Namane A, Lina G, Etienne J, Ehresmann B, Ehresmann C, Jacquier A, Vandenesch F, Romby P. 2005. *Staphylococcus aureus* RNAIII and the endoribonuclease III coordinately regulate spa gene expression. EMBO J 24:824–835. doi:10.1038/sj.emboj.7600572

Ibarra JA, Pérez-Rueda E, Carroll RK, Shaw LN. 2013. Global analysis of transcriptional regulators in *Staphylococcus aureus*. BMC Genomics 14:126. doi:10.1186/1471-2164-14-126

Iosub IA, Marchioretto M, Sy B, McKellar S, Nieken KJ, Nues RW van, Tree JJ, Viero G, Granneman S. 2018. Hfq CLASH uncovers sRNA-target interaction networks enhancing adaptation to nutrient availability. bioRxiv 481986. doi:10.1101/481986

Ivain L, Bordeau V, Eyraud A, Hallier M, Dreano S, Tattevin P, Felden B, Chabelskaya S. 2017. An in vivo reporter assay for sRNA-directed gene control in Gram-positive bacteria: identifying a novel sRNA target in *Staphylococcus aureus*. Nucleic Acids Research 45:4994–5007. doi:10.1093/nar/gkx190

Jiang Q, Jin Z, Sun B. 2018. MgrA negatively Regulates Biofilm Formation and Detachment by Repressing the Expression of psm Operons in Staphylococcus aureus. Appl Environ Microbiol AEM.01008-18. doi:10.1128/AEM.01008-18

Jin MS, Kim SE, Heo JY, Lee ME, Kim HM, Paik S-G, Lee H, Lee J-O. 2007. Crystal structure of the TLR1-TLR2 heterodimer induced by binding of a tri-acylated lipopeptide. Cell 130:1071–1082. doi:10.1016/j.cell.2007.09.008

Kang JY, Nan X, Jin MS, Youn S-J, Ryu YH, Mah S, Han SH, Lee H, Paik S-G, Lee J-O. 2009. Recognition of lipopeptide patterns by Toll-like receptor 2-Toll-like receptor 6 heterodimer. Immunity 31:873–884. doi:10.1016/j.immuni.2009.09.018

Kang S-S, Noh SY, Park O-J, Yun C-H, Han SH. 2015. *Staphylococcus aureus* induces IL-8 expression through its lipoproteins in the human intestinal epithelial cell, Caco-2. Cytokine, IL 4/IL 13 family 75:174–180. doi:10.1016/j.cyto.2015.04.017

Kathirvel M, Buchad H, Nair M. 2016. Enhancement of the pathogenicity of *Staphylococcus aureus* strain Newman by a small noncoding RNA SprX1. Medical Microbiology and Immunology 205:563–574. doi:10.1007/s00430-016-0467-9

Khemici V, Prados J, Linder P, Redder P. 2015. Decay-Initiating Endoribonucleolytic Cleavage by RNase Y Is Kept under Tight Control via Sequence Preference and Subcellular Localisation. PLOS Genetics 11:e1005577. doi:10.1371/journal.pgen.1005577

Kim S, Corvaglia A-R, Léo S, Cheung A, Francois P. 2016. Characterization of RNA Helicase CshA and Its Role in Protecting mRNAs and Small RNAs of *Staphylococcus aureus* Strain Newman. Infect Immun 84:833–844. doi:10.1128/IAI.01042-15

Kudla G, Granneman S, Hahn D, Beggs JD, Tollervey D. 2011. Cross-linking, ligation, and sequencing of hybrids reveals RNA–RNA interactions in yeast. PNAS 108:10010–10015. doi:10.1073/pnas.1017386108

Lalaouna D, Baude J, Wu Z, Tomasini A, Chicher J, Marzi S, Vandenesch F, Romby P, Caldelari I, Moreau K. RsaC sRNA modulates the oxidative stress response of *Staphylococcus aureus* during manganese starvation. Nucleic Acids Res. doi:10.1093/nar/gkz728

Lalaouna D, Carrier M-C, Semsey S, Brouard J-S, Wang J, Wade JT, Massé E. 2015. A 3' External Transcribed Spacer in a tRNA Transcript Acts as a Sponge for Small RNAs to Prevent Transcriptional Noise. Molecular Cell 58:393–405. doi:10.1016/j.molcel.2015.03.013

Lee BY, Singh A, David MZ, Bartsch SM, Slayton RB, Huang SS, Zimmer SM, Potter MA, Macal CM, Lauderdale DS, Miller LG, Daum RS. 2013. The economic burden of community-associated methicillin-resistant *Staphylococcus aureus* (CA-MRSA). Clinical Microbiology and Infection 19:528–536. doi:10.1111/j.1469-0691.2012.03914.x

Lehnik-Habrink M, Pförtner H, Rempeters L, Pietack N, Herzberg C, Stülke J. 2010. The RNA degradosome in *Bacillus subtilis*: identification of CshA as the major RNA he-

licase in the multiprotein complex. Mol Microbiol 77:958–971. doi:10.1111/j.1365-2958.2010.07264.x

Le Pabic H, Germain-Amiot N, Bordeau V, Felden B. 2015. A bacterial regulatory RNA attenuates virulence, spread and human host cell phagocytosis. Nucleic Acids Res 43:9232–9248. doi:10.1093/nar/gkv783

Lina G, Boutite F, Tristan A, Bes M, Etienne J, Vandenesch F. 2003. Bacterial competition for human nasal cavity colonization: role of *Staphylococcal* agr alleles. Appl Environ Microbiol 69:18–23.

Link TM, Valentin-Hansen P, Brennan RG. 2009. Structure of *Escherichia coli* Hfq bound to polyriboadenylate RNA. Proc Natl Acad Sci USA 106:19292–19297. doi:10.1073/pnas.0908744106

Lioliou E, Sharma CM, Caldelari I, Helfer A-C, Fechter P, Vandenesch F, Vogel J, Romby P. 2012. Global Regulatory Functions of the *Staphylococcus aureus* Endoribonuclease III in Gene Expression. PLOS Genetics 8:e1002782. doi:10.1371/journal.pgen.1002782

Liu GY. 2009. Molecular Pathogenesis of *Staphylococcus aureus* Infection. Pediatr Res 65:71R-77R. doi:10.1203/PDR.0b013e31819dc44d

Liu W, Rochat T, Toffano-Nioche C, Le Lam TN, Bouloc P, Morvan C. 2018. Assessment of Bona Fide sRNAs in *Staphylococcus aureus*. Front Microbiol 9. doi:10.3389/fmicb.2018.00228

Liu Y, Mu C, Ying X, Li W, Wu N, Dong J, Gao Y, Shao N, Fan M, Yang G. 2011. RNAIII activates map expression by forming an RNA–RNA complex in *Staphylococcus aureus*. FEBS Letters 585:899–905. doi:10.1016/j.febslet.2011.02.021

Liu Y, Wu N, Dong J, Gao Y, Zhang X, Mu C, Shao N, Yang G. 2010. Hfq is a global regulator that controls the pathogenicity of *Staphylococcus aureus*. PLoS ONE 5. doi:10.1371/journal.pone.0013069

Lu Z, Gong J, Zhang QC. 2018. PARIS: Psoralen Analysis of RNA Interactions and Structures with High Throughput and Resolution. Methods Mol Biol 1649:59–84. doi:10.1007/978-1-4939-7213-5_4

Luong TT, Dunman PM, Murphy E, Projan SJ, Lee CY. 2006. Transcription Profiling of the mgrA Regulon in *Staphylococcus aureus*. J Bacteriol 188:1899–1910. doi:10.1128/JB.188.5.1899-1910.2006

Mäder U, Nicolas P, Depke M, Pané-Farré J, Debarbouille M, Kooi-Pol MM van der, Guérin C, Dérozier S, Hiron A, Jarmer H, Leduc A, Michalik S, Reilman E, Schaffer M, Schmidt F, Bessières P, Noirot P, Hecker M, Msadek T, Völker U, Dijl JM van. 2016. *Staphylococcus aureus* Transcriptome Architecture: From Laboratory to Infection-Mimicking Conditions. PLOS Genetics 12:e1005962. doi:10.1371/journal.pgen.1005962

Maeda H, Fujita N, Ishihama A. 2000. Competition among seven Escherichia coli sigma subunits: relative binding affinities to the core RNA polymerase. Nucleic Acids Res 28:3497–3503.

Maier T, Güell M, Serrano L. 2009. Correlation of mRNA and protein in complex biological samples. FEBS Lett 583:3966–3973. doi:10.1016/j.febslet.2009.10.036

Marincola Gabriella, Schäfer Tina, Behler Juliane, Bernhardt Jörg, Ohlsen Knut, Goerke Christiane, Wolz Christiane. 2012. RNase Y of *Staphylococcus aureus* and its role in the activation of virulence genes. Molecular Microbiology 85:817–832. doi:10.1111/j.1365-2958.2012.08144.x

Mars RAT, Nicolas P, Denham EL, van Dijl JM. 2016. Regulatory RNAs in *Bacillus subtilis*: a Gram-Positive Perspective on Bacterial RNA-Mediated Regulation of Gene Expression. Microbiology and Molecular Biology Reviews 80:1029–1057. doi:10.1128/MMBR.00026-16

Mascher T, Helmann JD, Unden G. 2006. Stimulus Perception in Bacterial Signal-Transducing Histidine Kinases. Microbiol Mol Biol Rev 70:910–938. doi:10.1128/MMBR.00020-06

Mazmanian SK, Ton-That H, Schneewind O. 2001. Sortase-catalysed anchoring of surface proteins to the cell wall of Staphylococcus aureus. Mol Microbiol 40:1049–1057. doi:10.1046/j.1365-2958.2001.02411.x

McDougal LK, Steward CD, Killgore GE, Chaitram JM, McAllister SK, Tenover FC. 2003. Pulsed-Field Gel Electrophoresis Typing of Oxacillin-Resistant *Staphylococcus aureus* Isolates from the United States: Establishing a National Database. J Clin Microbiol 41:5113–5120. doi:10.1128/JCM.41.11.5113-5120.2003

Melamed S, Peer A, Faigenbaum-Romm R, Gatt YE, Reiss N, Bar A, Altuvia Y, Argaman L, Margalit H. 2016. Global Mapping of Small RNA-Target Interactions in Bacteria. Molecular Cell 63:884–897. doi:10.1016/j.molcel.2016.07.026

Merino N, Toledo-Arana A, Vergara-Irigaray M, Valle J, Solano C, Calvo E, Lopez JA, Foster TJ, Penadés JR, Lasa I. 2009. Protein A-Mediated Multicellular Behavior in

Staphylococcus aureus. Journal of Bacteriology 191:832–843. doi:10.1128/JB.01222-08

Miller HK, Carroll RK, Burda WN, Krute CN, Davenport JE, Shaw LN. 2012. The Extracytoplasmic Function Sigma Factor S Protects against both Intracellular and Extracytoplasmic Stresses in Staphylococcus aureus. J Bacteriol 194:4342–4354. doi:10.1128/JB.00484-12

Miyakoshi M, Chao Y, Vogel J. 2015. Cross talk between ABC transporter mRNAs via a target mRNA-derived sponge of the GcvB small RNA. EMBO J 34:1478–1492. doi:10.15252/embj.201490546

Mootz JM, Benson MA, Heim CE, Crosby HA, Kavanaugh JS, Dunman PM, Kielian T, Torres VJ, Horswill AR. 2015. Rot is a key regulator of Staphylococcus aureus biofilm formation. Mol Microbiol 96:388–404. doi:10.1111/mmi.12943

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

Morikawa K, Inose Y, Okamura H, Maruyama A, Hayashi H, Takeyasu K, Ohta T. 2003. A new staphylococcal sigma factor in the conserved gene cassette: functional significance and implication for the evolutionary processes. Genes to Cells 8:699–712. doi:10.1046/j.1365-2443.2003.00668.x

Morikawa K, Takemura AJ, Inose Y, Tsai M, Thi LTN, Ohta T, Msadek T. 2012. Expression of a Cryptic Secondary Sigma Factor Gene Unveils Natural Competence for DNA Transformation in Staphylococcus aureus. PLOS Pathogens 8:e1003003. doi:10.1371/journal.ppat.1003003

Morita T, Kawamoto H, Mizota T, Inada T, Aiba H. 2004. Enolase in the RNA degradosome plays a crucial role in the rapid decay of glucose transporter mRNA in the response to phosphosugar stress in *Escherichia coli*. Molecular Microbiology 54:1063–1075. doi:10.1111/j.1365-2958.2004.04329.x

Nguyen TC, Cao X, Yu P, Xiao S, Lu J, Biase FH, Sridhar B, Huang N, Zhang K, Zhong S. 2016. Mapping RNA–RNA interactome and RNA structure in vivo by MARIO. Nature Communications 7:12023. doi:10.1038/ncomms12023

Niebuhr M, Baumert K, Werfel T. 2010. TLR-2-mediated cytokine and chemokine secretion in human keratinocytes. Exp Dermatol 19:873–877. doi:10.1111/j.1600-0625.2010.01140.x

Nitzan M, Rehani R, Margalit H. 2017. Integration of Bacterial Small RNAs in Regulatory Networks. Annual Review of Biophysics 46:131–148. doi:10.1146/annurev-biophys-070816-034058

Novichkov PS, Kazakov AE, Ravcheev DA, Leyn SA, Kovaleva GY, Sutormin RA, Kazanov MD, Riehl W, Arkin AP, Dubchak I, Rodionov DA. 2013. RegPrecise 3.0–a resource for genome-scale exploration of transcriptional regulation in bacteria. BMC Genomics 14:745. doi:10.1186/1471-2164-14-745

Novick RP, Iordanescu S, Projan SJ, Kornblum J, Edelman I. 1989. pT181 plasmid replication is regulated by a countertranscript-driven transcriptional attenuator. Cell 59:395–404.

Novick RP, Ross HF, Projan SJ, Kornblum J, Kreiswirth B, Moghazeh S. 1993. Synthesis of *Staphylococcal* virulence factors is controlled by a regulatory RNA molecule. EMBO J 12:3967–3975.

Nues R van, Schweikert G, Leau E de, Selega A, Langford A, Franklin R, Iosub I, Wadsworth P, Sanguinetti G, Granneman S. 2017. Kinetic CRAC uncovers a role for Nab3 in determining gene expression profiles during stress. Nature Communications 8:12. doi:10.1038/s41467-017-00025-5

Ogston A. 1880. Ueber Abscesse. Archiv für Klinische Chirugie 25:588–600.

Olson PD, Kuechenmeister LJ, Anderson KL, Daily S, Beenken KE, Roux CM, Reniere ML, Lewis TL, Weiss WJ, Pulse M, Nguyen P, Simecka JW, Morrison JM, Sayood K, Asojo OA, Smeltzer MS, Skaar EP, Dunman PM. 2011. Small Molecule Inhibitors of *Staphylococcus aureus* RnpA Alter Cellular mRNA Turnover, Exhibit Antimicrobial Activity, and Attenuate Pathogenesis. PLOS Pathogens 7:e1001287. doi:10.1371/journal.ppat.1001287

O'Neill E, Pozzi C, Houston P, Humphreys H, Robinson DA, Loughman A, Foster TJ, O'Gara JP. 2008. A Novel *Staphylococcus aureus* Biofilm Phenotype Mediated by the Fibronectin-Binding Proteins, FnBPA and FnBPB. Journal of Bacteriology 190:3835–3850. doi:10.1128/JB.00167-08

Oscarsson J, Tegmark-Wisell K, Arvidson S. 2006. Coordinated and differential control of aureolysin (aur) and serine protease (sspA) transcription in *Staphylococcus aureus* by sarA, rot and agr (RNAIII). International Journal of Medical Microbiology 296:365–380. doi:10.1016/j.ijmm.2006.02.019

Painter KL, Krishna A, Wigneshweraraj S, Edwards AM. 2014. What role does the quorum-sensing accessory gene regulator system play during *Staphylococcus aureus* bacteremia? Trends in Microbiology 22:676–685. doi:10.1016/j.tim.2014.09.002

Palmqvist N, Foster T, Tarkowski A, Josefsson E. 2002. Protein A is a virulence factor in *Staphylococcus aureus* arthritis and septic death. Microb Pathog 33:239–249.

Pané-Farré J, Jonas B, Förstner K, Engelmann S, Hecker M. 2006. The B regulon in Staphylococcus aureus and its regulation. International Journal of Medical Microbiology 296:237–258. doi:10.1016/j.ijmm.2005.11.011

Patel AH, Nowlan P, Weavers ED, Foster T. 1987. Virulence of protein A-deficient and alpha-toxin-deficient mutants of *Staphylococcus aureus* isolated by allele replacement. Infect Immun 55:3103–3110.

Picard F, Dressaire C, Girbal L, Cocaign-Bousquet M. 2009. Examination of post-transcriptional regulations in prokaryotes by integrative biology. C R Biol 332:958–973. doi:10.1016/j.crvi.2009.09.005

Pichon C, Felden B. 2005. Small RNA genes expressed from *Staphylococcus aureus* genomic and pathogenicity islands with specific expression among pathogenic strains. PNAS 102:14249–14254. doi:10.1073/pnas.0503838102

Pinel-Marie M-L, Brielle R, Felden B. 2014. Dual Toxic-Peptide-Coding *Staphylococcus aureus* RNA under Antisense Regulation Targets Host Cells and Bacterial Rivals Unequally. Cell Reports 7:424–435. doi:10.1016/j.celrep.2014.03.012

Planet PJ. 2017. Life After USA300: The Rise and Fall of a Superbug. J Infect Dis 215:S71–S77. doi:10.1093/infdis/jiw444

Prasse D, Ehlers C, Backofen R, Schmitz RA. 2013. Regulatory RNAs in archaea: first target identification in *Methanoarchaea*. Biochemical Society Transactions 41:344–349. doi:10.1042/BST20120280

Py B, Causton H, Mudd EA, Higgins CF. 1994. A protein complex mediating mRNA degradation in *Escherichia coli*. Mol Microbiol 14:717–729.

Regev-Yochay G, Trzciński K, Thompson CM, Malley R, Lipsitch M. 2006. Interference between *Streptococcus pneumoniae* and *Staphylococcus aureus*: *In Vitro* Hydrogen Peroxide-Mediated Killing by *Streptococcus pneumoniae*. Journal of Bacteriology 188:4996–5001. doi:10.1128/JB.00317-06

Ren G-X, Guo X-P, Sun Y-C. 2017. Regulatory 3' Untranslated Regions of Bacterial mRNAs. Front Microbiol 8. doi:10.3389/fmicb.2017.01276

Rochat T, Bohn C, Morvan C, Le Lam TN, Razvi F, Pain A, Toffano-Nioche C, Ponien P, Jacq A, Jacquet E, Fey PD, Gautheret D, Bouloc P. 2018. The conserved regulatory RNA RsaE down-regulates the arginine degradation pathway in *Staphylococcus aureus*. Nucleic Acids Research 46:8803–8816. doi:10.1093/nar/gky584

Rochat T, Delumeau O, Figueroa-Bossi N, Noirot P, Bossi L, Dervyn E, Bouloc P. 2015. Tracking the Elusive Function of *Bacillus subtilis* Hfq. PLOS ONE 10:e0124977. doi:10.1371/journal.pone.0124977

Romilly C, Lays C, Tomasini A, Caldelari I, Benito Y, Hammann P, Geissmann T, Boisset S, Romby P, Vandenesch F. 2014. A Non-Coding RNA Promotes Bacterial Persistence and Decreases Virulence by Regulating a Regulator in *Staphylococcus aureus*. PLOS Pathogens 10:e1003979. doi:10.1371/journal.ppat.1003979

Rosenbach A. 1884. Mikro-Organismen bei den Wund-Infections-Krankheiten des Menschen. Wiesbaden: Bergmann.

Roux A, Todd DA, Velazquez JV, Cech NB, Sonenshein AL. 2014. CodY-Mediated Regulation of the *Staphylococcus aureus* Agr System Integrates Nutritional and Population Density Signals. Journal of Bacteriology 196:1184–1196. doi:10.1128/JB.00128-13

Roux CM, DeMuth JP, Dunman PM. 2011. Characterization of Components of the *Staphylococcus aureus* mRNA Degradosome Holoenzyme-Like Complex. J Bacteriol 193:5520–5526. doi:10.1128/JB.05485-11

Saïd-Salim B, Dunman PM, McAleese FM, Macapagal D, Murphy E, McNamara PJ, Arvidson S, Foster TJ, Projan SJ, Kreiswirth BN. 2003. Global Regulation of *Staphylococcus aureus* Genes by Rot. Journal of Bacteriology 185:610–619. doi:10.1128/JB.185.2.610-619.2003

Sassi M, Augagneur Y, Mauro T, Ivain L, Chabelskaya S, Hallier M, Sallou O, Felden B. 2015. SRD: a *Staphylococcus* regulatory RNA database. RNA 21:1005–1017. doi:10.1261/rna.049346.114

Sayed N, Jousselin A, Felden B. 2012. A cis-antisense RNA acts in trans in *Staphylococcus aureus* to control translation of a human cytolytic peptide. Nature Structural & Molecular Biology 19:105–112. doi:10.1038/nsmb.2193

Schroeder K, Jularic M, Horsburgh SM, Hirschhausen N, Neumann C, Bertling A, Schulte A, Foster S, Kehrel BE, Peters G, Heilmann C. 2009. Molecular Characterization of a Novel *Staphylococcus aureus* Surface Protein (SasC) Involved in Cell Aggregation and Biofilm Accumulation. PLOS ONE 4:e7567. doi:10.1371/journal.pone.0007567

Schumacher MA, Pearson RF, Møller T, Valentin-Hansen P, Brennan RG. 2002. Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: a bacterial Sm-like protein. EMBO J 21:3546–3556. doi:10.1093/emboj/cdf322

Schuster CF, Bertram R. 2016. Toxin-Antitoxin Systems of *Staphylococcus aureus*. Toxins 8:140. doi:10.3390/toxins8050140

Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M. 2011. Global quantification of mammalian gene expression control. Nature 473:337–342. doi:10.1038/nature10098

Schwarz-Linek U, Höök M, Potts JR. 2004. The molecular basis of fibronectin-mediated bacterial adherence to host cells. Mol Microbiol 52:631–641. doi:10.1111/j.1365-2958.2004.04027.x

Sedlyarova N, Shamovsky I, Bharati BK, Epshtein V, Chen J, Gottesman S, Schroeder R, Nudler E. 2016. sRNA-Mediated Control of Transcription Termination in *E. coli*. Cell 167:111-121.e13. doi:10.1016/j.cell.2016.09.004

Sendi P, Proctor RA. 2009. *Staphylococcus aureus* as an intracellular pathogen: the role of small colony variants. Trends in Microbiology 17:54–58. doi:10.1016/j.tim.2008.11.004

Shaevitz JW, Abbondanzieri EA, Landick R, Block SM. 2003. Backtracking by single RNA polymerase molecules observed at near-base-pair resolution. Nature 426:684–687. doi:10.1038/nature02191

Shaw LN, Aish J, Davenport JE, Brown MC, Lithgow JK, Simmonite K, Crossley H, Travis J, Potempa J, Foster SJ. 2006. Investigations into sigmaB-modulated regulatory pathways governing extracellular virulence determinant production in *Staphylococcus aureus*. J Bacteriol 188:6070–6080. doi:10.1128/JB.00551-06

Shaw LN, Lindholm C, Prajsnar TK, Miller HK, Brown MC, Golonka E, Stewart GC, Tarkowski A, Potempa J. 2008. Identification and Characterization of S, a Novel Component of the Staphylococcus aureus Stress and Virulence Responses. PLOS ONE 3:e3844. doi:10.1371/journal.pone.0003844

Smith EJ, Corrigan RM, van der Sluis T, Gründling A, Speziale P, Geoghegan JA, Foster TJ. 2012. The immune evasion protein Sbi of *Staphylococcus aureus* occurs both extracellularly and anchored to the cell envelope by binding lipoteichoic acid. Mol Microbiol 83:789–804. doi:10.1111/j.1365-2958.2011.07966.x

Srivastava AK, Schlessinger D. 1990. Mechanism and regulation of bacterial ribosomal RNA processing. Annu Rev Microbiol 44:105–129. doi:10.1146/annurev.mi.44.100190.000541

Storz G, Vogel J, Wassarman KM. 2011. Regulation by Small RNAs in Bacteria: Expanding Frontiers. Mol Cell 43:880–891. doi:10.1016/j.molcel.2011.08.022

Studier FW. 1975. Genetic mapping of a mutation that causes ribonucleases III deficiency in *Escherichia coli*. J Bacteriol 124:307–316.

Takai T, Chen X, Xie Y, Vu AT, Le TA, Kinoshita H, Kawasaki J, Kamijo S, Hara M, Ushio H, Baba T, Hiramatsu K, Ikeda S, Ogawa H, Okumura K. 2014. TSLP Expression Induced via Toll-Like Receptor Pathways in Human Keratinocytes In: Conn PM, editor. Methods in Enzymology, Endosome Signaling Part B. Academic Press. pp. 371–387. doi:10.1016/B978-0-12-397925-4.00021-3

Takeuchi O, Hoshino K, Akira S. 2000. Cutting edge: TLR2-deficient and MyD88-deficient mice are highly susceptible to *Staphylococcus aureus* infection. J Immunol 165:5392–5396.

Tenover FC, McDougal LK, Goering RV, Killgore G, Projan SJ, Patel JB, Dunman PM. 2006. Characterization of a strain of community-associated methicillin-resistant *Staphylococcus aureus* widely disseminated in the United States. J Clin Microbiol 44:108–118. doi:10.1128/JCM.44.1.108-118.2006

Toleman MS, Reuter S, Coll F, Harrison EM, Blane B, Brown NM, Török ME, Parkhill J, Peacock SJ. 2016. Systematic Surveillance Detects Multiple Silent Introductions and Household Transmission of Methicillin-Resistant *Staphylococcus aureus* USA300 in the East of England. J Infect Dis 214:447–453. doi:10.1093/infdis/jiw166

Tomasini A, François P, Howden BP, Fechter P, Romby P, Caldelari I. 2014. The importance of regulatory RNAs in *Staphylococcus aureus*. Infection, Genetics and Evolution 21:616–626. doi:10.1016/j.meegid.2013.11.016

Tomasini A, Moreau K, Chicher J, Geissmann T, Vandenesch F, Romby P, Marzi S, Caldelari I. 2017. The RNA targetome of *Staphylococcus aureus* non-coding RNA

RsaA: impact on cell surface properties and defense mechanisms. Nucleic Acids Res 45:6746–6760. doi:10.1093/nar/gkx219

Travis AJ, Moody J, Helwak A, Tollervey D, Kudla G. 2014. Hyb: A bioinformatics pipeline for the analysis of CLASH (crosslinking, ligation and sequencing of hybrids) data. Methods 65:263–273. doi:10.1016/j.ymeth.2013.10.015

Tuchscherr L, Bischoff M, Lattar SM, Noto Llana M, Pförtner H, Niemann S, Geraci J, Van de Vyver H, Fraunholz MJ, Cheung AL, Herrmann M, Völker U, Sordelli DO, Peters G, Löffler B. 2015. Sigma Factor SigB Is Crucial to Mediate Staphylococcus aureus Adaptation during Chronic Infections. PLoS Pathog 11. doi:10.1371/journal.ppat.1004870

Udo EE, Pearman JW, Grubb WB. 1993. Genetic analysis of community isolates of methicillin-resistant *Staphylococcus aureus* in Western Australia. Journal of Hospital Infection 25:97–108. doi:10.1016/0195-6701(93)90100-E

van Belkum A, Melles DC, Nouwen J, van Leeuwen WB, van Wamel W, Vos MC, Wertheim HFL, Verbrugh HA. 2009. Co-evolutionary aspects of human colonisation and infection by *Staphylococcus aureus*. Infection, Genetics and Evolution 9:32–47. doi:10.1016/j.meegid.2008.09.012

Vanderpool CK. 2011. Combined experimental and computational strategies define an expansive regulon for GcvB small RNA. Mol Microbiol 81:1129–1132. doi:10.1111/j.1365-2958.2011.07780.x

Verdon J, Girardin N, Lacombe C, Berjeaud J-M, Héchard Y. 2009. delta-hemolysin, an update on a membrane-interacting peptide. Peptides 30:817–823. doi:10.1016/j.peptides.2008.12.017

Vogel J. 2009. A rough guide to the non-coding RNA world of Salmonella. Molecular Microbiology 71:1–11. doi:10.1111/j.1365-2958.2008.06505.x

Vogel J, Luisi BF. 2011. Hfq and its constellation of RNA. Nat Rev Microbiol 9:578–589. doi:10.1038/nrmicro2615

Von Dach E, Diene SM, Fankhauser C, Schrenzel J, Harbarth S, François P. 2016. Comparative Genomics of Community-Associated Methicillin-Resistant *Staphylococcus aureus* Shows the Emergence of Clone ST8-USA300 in Geneva, Switzerland. J Infect Dis 213:1370–1379. doi:10.1093/infdis/jiv489

Wade JT, Grainger DC. 2014. Pervasive transcription: illuminating the dark matter of bacterial transcriptomes. Nat Rev Microbiol 12:647–653. doi:10.1038/nrmicro3316

Wagner EGH, Altuvia S, Romby P. 2002. 12 - Antisense RNAs in Bacteria and Their Genetic Elements In: Dunlap JC, Wu C -ting, editors. Advances in Genetics, Homology Effects. Academic Press. pp. 361–398. doi:10.1016/S0065-2660(02)46013-0

Walsh EJ, O'Brien LM, Liang X, Hook M, Foster TJ. 2004. Clumping factor B, a fibrinogen-binding MSCRAMM (microbial surface components recognizing adhesive matrix molecules) adhesin of *Staphylococcus aureus*, also binds to the tail region of type I cytokeratin 10. J Biol Chem 279:50691–50699. doi:10.1074/jbc.M408713200

Wang B, Zhao A, Novick RP, Muir TW. 2014. Activation and Inhibition of the Receptor Histidine Kinase AgrC Occurs through Opposite Helical Transduction Motions. Molecular Cell 53:929–940. doi:10.1016/j.molcel.2014.02.029

Wang R, Braughton KR, Kretschmer D, Bach T-HL, Queck SY, Li M, Kennedy AD, Dorward DW, Klebanoff SJ, Peschel A, DeLeo FR, Otto M. 2007. Identification of novel cytolytic peptides as key virulence determinants for community-associated MRSA. Nature Medicine 13:1510–1514. doi:10.1038/nm1656

Washburn RS, Gottesman ME. 2015. Regulation of Transcription Elongation and Termination. Biomolecules 5:1063–1078. doi:10.3390/biom5021063

Washburn RS, Marra A, Bryant AP, Rosenberg M, Gentry DR. 2001. rho Is Not Essential for Viability or Virulence in *Staphylococcus aureus*. Antimicrobial Agents and Chemotherapy 45:1099–1103. doi:10.1128/AAC.45.4.1099-1103.2001

Waters SA, McAteer SP, Kudla G, Pang I, Deshpande NP, Amos TG, Leong KW, Wilkins MR, Strugnell R, Gally DL, Tollervey D, Tree JJ. 2017. Small RNA interactome of pathogenic *E. coli* revealed through crosslinking of RNase E. The EMBO Journal 36:374–387. doi:10.15252/embj.201694639

Weidenmaier C, Kokai-Kun JF, Kristian SA, Chanturiya T, Kalbacher H, Gross M, Nicholson G, Neumeister B, Mond JJ, Peschel A. 2004. Role of teichoic acids in *Staphylococcus aureus* nasal colonization, a major risk factor in nosocomial infections. Nat Med 10:243–245. doi:10.1038/nm991

Weiss A, Shaw LN. 2015. Small things considered: the small accessory subunits of RNA polymerase in Gram-positive bacteria. FEMS Microbiol Rev 39:541–554. doi:10.1093/femsre/fuv005

Wertheim HFL, Walsh E, Choudhurry R, Melles DC, Boelens HAM, Miajlovic H, Verbrugh HA, Foster T, van Belkum A. 2008. Key role for clumping factor B in

Staphylococcus aureus nasal colonization of humans. PLoS Med 5:e17. doi:10.1371/journal.pmed.0050017

Yoshida T, Qin L, Egger LA, Inouye M. 2006. Transcription Regulation of ompF and ompC by a Single Transcription Factor, OmpR. J Biol Chem 281:17114–17123. doi:10.1074/jbc.M602112200

Materials and methods

2

2.1 Bacterial strains and culture conditions

An overview of all *E. coli* and *S. aureus* strains used in this study is provided in Table 2.1. The *E. coli* DH5α strain was used for general plasmid propagation, while the *S. aureus* RN4220 strain was used as an intermediate for transforming plasmids into *S. aureus* USA300. RN450 was used to produce and harvest phage for transduction of USA300.

CLASH was performed in the *S. aureus* strains JKD6008 RNase III::HTF, JKD6008 RNase Y::HTF and USA300 RNase III::HTF. CRAC was performed on USA300 RpoC::HTF. In these strains, the genes encoding for RNase III (*rnc*), RNase Y (*rny*) or RpoC were chromosomally tagged by Pedro Arede Rei with the dual affinity HTF tag (His6-TEV-3xFLAG). These strains generated through allelic exchange using the pIMAY plasmid (Monk et al., 2012) with the help of Pedro Arede Rei.

Both JKD6008 strains were provided by Jai Tree (UNSW, Sydney), USA300 and RN450 were provided by Ross Fitzgerald (Roslin Institute, Edinburgh), and RN4220 by Baolin Sun (University of Science and Technology of China, China).

For plasmid propagation, E. coli DH5α were grown in lysogeny broth (LB) at 37°C with shaking at 200 rpm. The medium was supplemented with either ampicillin ($100\,\mu$g/mL) or chloramphenicol ($25\,\mu$g/mL) when required. Overnight cultures were grown in 5 mL volumes and purified with a miniprep kit according to manufacturer's instructions (ThermoFisher).

Tab. 2.1: Strains used in this study

Strain	Genotype	Reference
E. coli DH5α	*fhuA2 lac*(del) U169 *phoA glnV44 80'ωlacZ*(del)*M15 gyrA96 recA1 relA1 endA1 thi-1hsdR17*	Taylor et al, 1993
S. aureus RN4220	Restriction-deficient derivative of reference strain NCTC8325	R Novick
S. aureus RN450 Ω80α	Encodes for 80α phage under the control of a mitomycin C-responsive promoter	R Novick
S. aureus USA300 LAC	Prototype MRSA strain, ST8, β-lactamase positive	McDougal et al, 2003; Diep et al, 2006
S. aureus JKD6008 RNase III::HTF	*rnc*::HTF	Jai Tree, UNSW
S. aureus JKD6008 RNase Y::HTF	*rny*::HTF	Jai Tree, UNSW
S. aureus USA300 RNase III::HTF	*rnc*::HTF	This study, Pedro Arede Rei
S. aureus USA300 LAC rpoC::HTF	*rpoC*::HTF	This study, Pedro Arede Rei

2.2 Shift experiments and UV crosslinking

Cells were firstly plated out onto TSA from a glycerol stock and a single colony was inoculated into 5 mL of TSB and grown overnight at 3737°C with shaking. This overnight culture was then diluted 1:500 into 500 mL of TSB and grown until OD_{600}=3 with the same conditions.

For CLASH, a 65 mL aliquot of the culture was taken at OD_{600}=3 and crosslinked in a Vari-X-Linker (UVO3) with 250 mJ of UV. The cells were then quickly harvested with a vaccuum filtration device (UVO3) and flash-frozen in liquid nitrogen. Afterwards, another 65 mL aliquot was was harvested through vacuum filtration and then were resuspended in 65 mL of RPMI medium (Gibco). After 15 minutes, the cells were crosslinked and harvested as described above.

For the RNAseq and qPCR experiments, the experimental set up was similar. 50 mL of culture was shifted to RPMI and samples were then taken after 5, 10, 15 and 30 minutes, plus a time zero (before shift) reference sample. 5 mL of culture was taken and harvested through centrifugation at 13,000 g for 1 minute at room temperature in 5 mL tubes.

As a control in the qPCR and RNAseq experiments, an independent experiment where the cells were resuspended back into their own TSB supernatant was performed in order to study any mechanical stresses as a result of the shift.

The composition of TSB as follows; tryptone 17 g/L, soytone 3.0 g/L, glucose 2.5 g/L, sodium chloride 5 g/L, dipotassium phosphate 2.5 g/L. The composition of RPMI can be found on the ThermoFisher website, reference R7513.

2.3 Construction of pIMAY plasmids for HTF tagging

Table 2.2 lists the plasmids used and constructed, Table 2.3 lists the primers, and Table 2.4 all gene fragments used in in this study. This includes those for the construction of pIMAY plasmids required for tagging genes with the HTF tag.

Tab. 2.2: Plasmids used in this study

Plasmid	Description	Characteristics	Reference
pBS1539-HTF	Cloning vector containing HTF tag sequence	Amp^r	Nues et al., 2017
pIMAY	Vector for allelic exchange in S. aureus	Cm^r, pWV01ts replicon	Monk et al., 2012
pUC19	Cloning vector and template for in vitro transcription	Amp^r, pMB1 replicon	Norrander et al., 1983
pCN33-GFP	Shuttle vector containing sfGFP under a pTufA constitutive promoter	Em^r, pT181 replicon	Ivain et al., 2017
pICS3	Shuttle vector for sRNA expression	Cm^r, pC194 replicon	Ivain et al., 2017
pRMC2	Shuttle vector for inducible sRNA expression under pTetO control	Cm^r, pC194 replicon	Corrigan and Foster, 2009
pJET1.2	Cloning vector for blunt-end ligation of PCR products and gBlocks	Amp^r, pBR322 replicon	ThermoFisher

For tagging genes with the HTF tag, Gibson assembly was used to create the required plasmid. Firstly, a forward primer was designed to anneal to the sequence 500 basepairs upstream of the natural stop codon of the target gene, and a sequence designed to overlap with the pIMAY plasmid, upstream of the KpnI site, was added to the 5' end (Table 2.3, pIMAY fw). A reverse primer was then generated to amplify this sequence, finishing on the nucleotide immediately before the stop codon. To this reverse primer, a sequence designed to overlap with the HTF tag was added to the 5' end (Table 2.3, HTF overlap 1).

Secondly, primers were designed to amplify the HTF tag from the pBS1539::HTF plasmid. The forward primer was designed with a 20 nucleotide overlap of sequence immediately before the stop codon of the gene to be tagged. The reverse primer was then designed to contain the natural stop codon and then 17 nucleotides upstream of it. To these sequences, HTF fw and HTF rev (Table 2.3) were added respectively.

Thirdly, a reverse primer was designed to anneal to the sequence 500 basepairs downstream of the natural stop codon of the target gene, and a sequence designed to overlap with pIMAY downstream of the KpnI site was added (Table 2.3,pIMAY rev 1). A forward primer, starting at the natural stop codon, was designed to amplify this sequence and contained a 20 nucleotide overlap with the HTF tag (Table 2.3, HTF overlap 2).

The PCRs were performed using Q5 polymerase (NEB) with the manufacturer's recommended conditions. The annealing temperature was calculated through use of a Tm calculator (NEB, https://tmcalculator.neb.com/). PCR products were then resolved on an agarose gel and the band of interest gel purified using the MinElute gel extraction kit (Qiagen). In order to assemble the fragments into pIMAY, $2\mu g$ of the plasmid backbone was first digested with 2 U of KpnI-HF (NEB). The fragments were then assembled into pIMAY using the Gibson Assembly Mastermix (NEB) according to manufacturer's instructions. The Gibson Assembly products were then transformed into *E. coli* DH5α, single colonies were then grown, the plasmids purified and finally confirmed via sequencing.

Tab. 2.3: Primers

Oligonucleotide name	**Sequence (5' to 3')**	**Purpose**
pIMAY fw 1	TCACTAAAGGGAACAAAAGC TGG	Overlaps pIMAY upstream to the digested KpnI site.
HTF overlap 1	ATAATCATGGTGATGGTGAT GGTGCTCCATGGATCC	Overlaps with the start of the HTF tag

Tab. 2.3: Primers

Oligonucleotide name	Sequence (5' to 3')	Purpose
HTF fw	GGATCCATGGAGCACCATCA	Amplifies HTF tag
HTF rev	ACGCGGCCGCAGAATTCTCA	Amplifies HTF tag
HTF overlap 2	GATGACGACGATAAAGACTA CAAAGATGACGACGATAAAT	Overlaps with the end of the HTF tag
pIMAY rev 1	GTCGACCTCGAGGGGGGGC C	Overlaps pIMAY upstream to the digested KpnI site
RNase III 1-500 fw	TCACTAAAGGGAACAAAAGC TGGAACATCCCAACTTGCCA GAA	Generates first fragment for RNase III::HTF
RNase III 1-500 rev	ATAATCATGGTGATGGTGAT GGTGCTCCATGGATCCTTTA ATTTGTTTTAATTGCTTAT	Generates first fragment for RNase III::HTF
RNase III 501-656 fw	TATAAGCAATTAAAACAAAT TAAAGGATCCATGGAGCACC ATCA	Generates HTF tag with RNase III overlap
RNase III 501-656 rev	GTCGTATCATATAAATTTCT AACGCGGCCGCAGAATTCTC A	Generates HTF tag with RNase III overlap
RNase III 657-1156 fw	GATGACGACGATAAAGACTA CAAAGATGACGACGATAAAT AGAAATTTATATGATACGAC	Generates last fragment for RNase III::HTF
RNase III 657-1156 rev	GTCGACCTCGAGGGGGGGC CATTATTTATGTAGTACTCA	Generates last fragment for RNaseIII::HTF
RNase Y 1-500 fw	CACTAAAGGGAACAAAAGCT GGTAGGCGAAGATGAGACA TT	Generates first fragment for RNase Y::HTF
RNase Y 1-500 rev	TGATGGTGCTCCATGGATCC TTTCGCATATTCTACTGCT	Generates first fragment for RNaseY::HTF
RNase Y 501-656 fw	TAGAGCAGTAGAATATGCGA AAGGATCCATGGAGCACCAT CA	Generates HTF tag with RNase Y overlap

Tab. 2.3: Primers

Oligonucleotide name	Sequence (5' to 3')	Purpose
RNase Y 501-656 rev	CTAATTTGTGAGGGAGACAAAAATTAACGCGGCCGCAGAATTCTCA	Generates HTF tag with RNase Y overlap
RNase Y 657-1156 fw	GACTACAAAGATGACGACGATAAATAATTTTTGTCTCCCTCACAAATTAGTGAGGGAGC	Generates last fragment for RNase Y::HTF
RNase Y 657-1156 rev	GACTACAAAGATGACGACGATAAATAATTTTTGTCTCCCTCACAAATT	Generates last fragment for RNaseY:::HTF
RpoC 1-500 fw	CTAAAGGGAACAAAAGCTGGTTGCTGGATTAAACGCGAC	Generates first fragment for RpoC::HTF
RpoC 1-500 rev	TGATGGTGCTCCATGGATCCTTCCGTTACTTCAGTTTGA	Generates first fragment for RpoC::HTF
RpoC 501-656 fw	AATCTCAAACTGAAGTAACGGAAGGATCCATGGAGCACCATCA	Generates HTF tag with RpoC overlap
RpoC 501-656 rev	ATTAGCCTCTGTTATATACTTGTTACTCACGCGGCCGCAGAATT	Generates HTF tag with RpoC overlap
RpoC 676-1156 fw	GAGAATTCTGCGGCCGCGTGAGTAACAAGTATATAACAGAGGCTAA	Generates last fragment for RpoC::HTF
RpoC 657-1156 rev	TCGACCTCGAGGGGGGGCCCGTAGTTGGCATGTGATATGT	Generates last fragment of RpoC::HTF
RNAIII rev	TATGAATTCCAACATTAGACTTATTCATATATTTTAACGGCG	Amplification of RNAIII for cloning into pRMC2 and pICS3
RNAIII fw	ATACTGCAGAAAATTTGTTTGATTTTTAATGGATAATGTGATATAATGGTACCTAACTAGATCACAGAGATGTGATGGAAAATAG	Amplification of RNAIII for cloning into pRMC2 and pICS3

Tab. 2.3: Primers

Oligonucleotide name	Sequence (5' to 3')	Purpose
RsaA rev	TATGAATTCAAATCTGTATG TTTAAACTTATATATATGTG CTAATGTATTATC	Amplification of RsaA for cloning into pRMC2 and pICS3
RsaA fw	ATACTGCAGAAAATTTGTTT GATTTTTAATGGATAATGTG ATATAATGGTACCAACCATT ACAAAAATTGTATAGAGTAG CGAC	Amplification of RsaA for cloning into pRMC2 and pICS3
pIMAY MCS fw	TACATGTCAAGAATAAACTG CCAAAGC	Screening for presence of pIMAY
pIMAY MCS rev	AATACCTGTGACGGAAGATC ACTTCG	Screening for presence of pIMAY

Tab. 2.4: Gene fragments used in this study

Gene fragment	Sequence (5' to 3')
qacAB pCN33	GCGCAGATCTGTTGTAATATGTAAAAAAATAGATTATA ATCTTATAGACCGATCGCACGGTCTATAAGGATTGGAG GGAACTTAAATGATTTCATTTTTACAAAAACTACTGATG ATATCCGAG
esxA pCN33	GCGAGATCTTAAAAAGTTTTGATAAACTTAAAATATTCA GGAGGTTTCTAGTTATGGCAATGATTAAGATGAGTCCA GAGGAAATCAGAGCAAAATCGCAATCTTACGGGCAAGG TTCAGATATCGCG
RsaA	ATAGGATCCTAATACGACTCACTATAGGGGTTAACCAT TACAAAAATTGTATAGAGTAGCGACTGTATAATTTCTAT TGAGGTTAACGTTTATATGTAGTGATAGTAGTTAAAGT TCTCCCAAGGAAGACTACTCGGGTACACTTTGCTATGA GCAAAGTGTACTTTGTTATTGATACTCGAGGTCGACAT A

Tab. 2.4: Gene fragments used in this study

Gene fragment	Sequence (5' to 3')
RNAIII interacting	ATAGGATCCTAATACGACTCACTATAGGGATAGTACTA AAAGTATGAGTTATTAAGCCATCCCAACTTAATAACCAT GTAAAATTAGCAAGTGATAACATTTGCTAGTAGAGTTA GTTTCCTTGGACTCAGTGCTATGTATTTTTCTTAATTAT CATTACAGATAATTATTTCTAGCATGTAAGCTACTCGAG GTCGACATA
RNAIII non-interacting	ATAGGATCCTAATACGACTCACTATAGGGAGTGATTTC AATGGCACAAGATATCATTTCAACAATCGGTGACTTAG TAAAATGGATTATCGACACAGTGAACAAATTCACTAAA AAATAAGATGAATAATTAATTACTTTCATTGTAAATTTG TTATCTCTCGAGGTCGACATA
RsaE	ATAGGATCCTAATACGACTCACTATAGGGTGAAATTAA TCACATAACAAACATACCCCTTTGTTTGAAGTGAAAAAT TTCTCCCATCCCCTTTGTTTAGCGTCGTGTATTCAGCTC GAGGTCGACATA
RsaE single mutant	ATAGGATCCTAATACGACTCACTATAGGGTGAAATTAA TCACATAACAAACATACCCCTTTGTTTGAAGTGAAAAAT TTCTGGGATGGGGTTTGTTTAGCGTCGTGTATTCAGCT CGAGGTCGACATA
RsaE double mutant	ATAGGATCCTAATACGACTCACTATAGGGTGAAATTAA TCACATAACAAACATAGGGGTTTGTTTGAAGTGAAAAA TTTCTGGGATGGGGTTTGTTTAGCGTCGTGTATTCAGC TCGAGGTCGACATA
RsaOG	ATAGGATCCTAATACGACTCACTATAGGGTAACAGGGG GAGCGATTAAACAAAGGGGTAGAGCTATATAACAGATA GCTTTTAATCGTTCAAGTTACATCACAACATCTTCAATA TTTATTACTTACTTTCCTTTCTATTTGTCGGCTAGCACG TGACTAGCCGACTCGAGGTCGACATA
09902	ATAGGATCCTAATACGACTCACTATAGGGATTTCATAA CAAACAAAGGAGGTCTTTCACATGGGTATCATTGCAGG AATCATTAAATTCATTAAAGGATTAATTGAGAAATTCAC TGGTAAGTACTCGAGGTCGACATA

2.4 Construction of inducible sRNA expression plasmids

For inducible expression of RsaA and RNAIII, the pRMC2 expression vector was used (Corrigan and Foster, 2009). This utilises a pTetO promoter which induces gene expression upon addition of anhydrotetracycline.

For cloning RsaA and RNAIII into pRMC2, PCR was used to amplify the sequence from isolated genomic DNA using the respective primers (Table 2.3). The PCR was carried out as previously described in Section 2.2, and the products were resolved on an agarose gel and the band of interest purified. The produced RsaA and RNAIII fragments were then blunt-end cloned into pJET1.2, transformed into DH5α and confirmed via sequencing following purification.

The inserts were then cut out of pJET1.2 using 3 U of EcoRI-HF and 3 U of KpnI-HF (NEB) in reactions containing $3\mu g$ of recombinant plasmid. Digests were carried out at 37°C for 1 hour. Concurrently, $3\mu g$ of pRMC2 was digested in the same way, and then dephosphorylated for 1 hour at 37°C using 1 U of rSAP (NEB). Following digestion, all reactions were resolved on an agarose gel, and digested pRMC2 and isolated RsaA and RNAIII fragments were purified using the MinElute gel extraction kit.

Following extraction, RsaA and RNAIII were ligated into pRMC2. Ligations were set up using 50 ng of digested pRMC2, and inserts were added in 3:1 ratios, with 400 U of T4 DNA ligase. Reactions were left overnight at 16°C, and subsequently transformed into DH5α. Plasmids were purified from single colonies and screened through restriction digest using KpnI-HF and EcoRI-HF, as described above.

2.5 Construction of mRNA-GFP fusion and partner sRNA plasmids

For examining the role of selected sRNAs in regulating the translation of chosen targets, sRNAs were cloned with the pTufA promoter into pICS3, while a portion of the 5' UTR and coding sequence of the target was cloned into pCN33-GFP.

Generation of pICS3::RsaA and pICS3::RNAIII was performed as described in Sections 2.2 and 2.3. Importantly, the produced products included a constitutive promoter, pTufA, when utilising the PstI and EcoRI sites.

For producing the GFP fusions in pCN33, gene fragments were designed to include around 80 nucleotides of the 5' UTR, the start codon, and 30 nucleotides of the coding sequence. These gene fragments were cloned into pJET1.2, transformed into DH5α and confirmed by sequencing. The inserts were then excised using EcoRV and PstI, and ligated into pCN33 as described in Section 2.3.

2.6 Construction of *in vitro* transcription plasmids

For producing in vitro transcribed RNA, chosen sRNAs were cloned into the pUC19 plasmid. This plasmid was used due to its lack of T7 promoter and high copy number.

Chosen sRNA constructs were designed to include for a BamHI site at the 5' end, followed by a T7 promoter, the sRNA sequence, a XhoI site, and finally ending in a SalI site. The BamHI and SalI sites are used for cloning into pUC19, while the XhoI site is used for linearising the plasmids for transcription.

The constructs for RsaA, RNAIII interacting, RNAIII non-interacting, RNAIII, RsaE, RsaOG, *09902, 09903* and 5S are shown in Table 2.4, and were ordered from IDT as gBlocks. Two mutant versions of RsaE in which one or both of the CCC motifs were mutated to GGG (RsaE single mutant and RsaE double mutant) were also produced in order to study their roles in mediating the interaction between PSMs and RsaE. These gene fragments were cloned into pUC19 as described above and in section 2.3.

2.7 Genomic DNA extraction

Extraction of genomic DNA from *S. aureus* was accomplished through the Bacterial Genomic DNA Extraction Kit (EdgeBio) according to manufacturer's instructions, with one modification. From an overnight culture, 1 mL of cells were harvested through centrifugation at 13.4K for 1 minute. The supernatant was decanted, and the pellet resuspended in $100\,\mu$l of the kit's spheroplast buffer. To this, $5\,\mu$l of 10 mg/mL lysostaphin (Prospect Bio) solution was added and incubated for 1 hour at 37°C in order to degrade the outer cell wall. The manufacturer's instructions were subsequently followed, with the final DNA pellet resuspended in $50\,\mu$l of water.

2.8 *E. coli* chemically competent cell generation and heat shock transformation

Chemically-competent *E. coli* cells were produced using the Inoue method. Briefly, a single colony was inoculated into 25 mL of LB and grown overnight at 37°C, shaking at 200 rpm. This overnight culture was then used to inoculate 250 mL of LB. The culture was then grown overnight at 18°C, shaking at 200 rpm.

Culture growth was continued until a final OD_{600} value of 0.5, and afterwards the cells were transferred to an ice water bath for 10 minutes. The cells were then spun for 10 minutes at 4,600 g, 4°C, and the supernatant was decanted and residual medium removed with a pipette. The cells were then resuspended in 80 mL of ice-cold Inoue transformation buffer (55 mM $MnCl_2.4H_2O$, 15 mM $CaCl_2.2H_2O$, 250 mM KCl, 10 mM PIPES). The cells were then collected again through centrifugation and subsequently resuspended into 20 mL of Inoue transformation buffer, and 1.5 mL of DMSO was added and the cells left on ice for 10 minutes. The cells were then aliquoted and flash-frozen in liquid nitrogen before being stored at -80°C.

For transformation, $100\,\mu$l of cells were thawed on ice for 10 minutes. The plasmid to be transformed was then added ($1\,\mu$l for whole plasmids, or $10\,\mu$l for ligation products) and the samples left on ice for 30 minutes. The cells were then heat shocked at 42°C for 60 seconds, left on ice for 2 minutes, and then recovered through addition of $800\,\mu$l of SOC medium and incubation at 37°C for 1 hour with shaking. Cells were subsequently pelleted through centrifugation at 3,000 g for 3 minutes, before being resuspended in around $200\,\mu$l of residual medium and plated onto LB agar with the appropriate antibiotic. Cells were then grown overnight at 37°C.

2.9 *S. aureus* electrocompetent cell generation and electroporation

S. aureus RN4220 was plated out onto TSA (Oxoid) and a single colony was inoculated into 5 mL of TSB (Oxoid) and grown overnight at 37°C. The next day, this overnight culture was diluted 1:500 into 500 mL of TSB and grown to OD_{600} = 0.5.

The cells were then transferred to an ice-water bath for 10 minutes and collected afterwards through centrifugation at 4600 g for 10 minutes at 4°C. The cell pellet

was resuspended in 50 mL of ice cold 0.5 M sucrose, and then the cells were collected again through centrifugation. The supernatant was subsequently discarded and the cells resuspended in 25 mL of sucrose. The cells were then left on ice for 20 minutes before being collected again. The supernatant was discarded and the cells resuspended in $500\,\mu$l of sucrose. Aliquots of $50\,\mu$l were then dispensed and flash-frozen in liquid nitrogen, and stored at -80°C until use.

Competent cells were thawed on ice when required for 10 minutes. For routine plasmids, 250-500 ng of plasmid was added to the cells in a volume no more than $5\,\mu$l. The cells were then transferred to an ice-cold 0.1 cm cuvette (BioRad) and pulsed in a BioRad cell-pulser with the following conditions: 200 ω, 25 μF, 2.1 kV. Immediately after electroporation, the cells were recovered in 1 mL of prewarmed TSB and grown for 1 hour at 37°C, before being pelleted and plated out onto TSA with the appropriate antibiotic as previously described.

For pIMAY constructs, plasmids were concentrated to around $5\,\mu$l/μg and $10\,\mu$l of plasmid was used for electroporation. The plasmids were electroporated as described above, but the cells recovered in TSB for 2 hours at 30°C. Following plating onto TSA-chloramphenicol $5\,\mu$g/mL, the cells were grown for 48-72 hours at 30°C.

2.10 Phage transduction of USA300

Phages were generated using the RN450 strain. First, cells were plated onto TSA and grown overnight at 37°C. Afterwards, a single colony was inoculated into 5 mL of TSB and grown overnight at 37°C with shaking at 200 rpm. The following day, $250\,\mu$l of overnight culture was inoculated into 25 mL of TSB and grown until OD_{600} = 0.3. Mitomycin C was added to a final concentration of $2\,\mu$g/mL in order to induce phage production and the cells shaken slowly with 80 rpms at 32°C until completely lysed. The phages were then filtered through a 0.4 μm filter and stored at 4°C until use.

The plasmid to be transduced into USA300 was firstly electroporated in RN4220 as described in Section 2.8. The produced RN4220::pDNA strain was then grown overnight at 37°C on a BHI slant (Oxoid). The cells were then resuspended in 1 mL of TSB, and $CaCl_2$ was added to a final concentration of 5 mM. To six Falcon tubes, $10\,\mu$l of these cells were added.

The phages were then serially diluted in phage buffer (1 mM $MgSO_4$, 4 mM $CaCl_2$, 50 mM Tris pH7.8, 100 mM NaCl, 0.1% gelatin) from 10^{-3} to 10^{-8}, and $10\,\mu$l of each dilution was added to a 15 mL Falcon tube containing the cells. Afterwards, 3 mL

of liquid phage top agar (0.3% casamino acids, 0.3% yeast extract, 100 mM NaCl, 0.5% agar) was added and the cells poured and spread onto 20 mL plates of phage bottom agar (0.3% casamino acids, 0.3% yeast extract, 100 mM NaCl, 1.5% agar). The plates were then incubated overnight at 30°C, and USA300 was also inoculated onto a BHI slant to grow overnight at 37°C.

The transduction plate that gave near-confluent lysis was then selected and 2 mL of phage buffer was added. The plate was left at 4°C for 1 hour, and then the top layer of agar was scraped off and put into a Falcon tube, along with the phage buffer. This was then centrifuged at 4600 g for 30 minutes at 4°C, and then the supernatant collected and filtered twice through 0.4 μm filters. These plasmid-containing phages were then stored at 4°C until use.

The USA300 cells grown on the BHI slant were recovered in 1 mL of TSB and $CaCl_2$ was added to a final concentration of 5 mM. Afterwards, $50\,\mu$l of these cells were added to a Falcon tube containing $150\,\mu$l of phage buffer, and then $50\,\mu$l of plasmid-containing phages were added. The Falcon tubes were then moved to a 37°C incubator for 20 minutes, shaking at 200 rpm. Afterwards, 3 mL of liquid 0.3 GL top agar (0.3% casamino acids, 0.3% yeast extract, 100 mM NaCl, 0.15% sodium lactate, 0.1% glycerol, 1.5 mM trisodium citrate, 0.5% agar, pH 8) was added and then the cells poured over 20 mL plates of 0.3 GL bottom agar (0.3 GL top agar but with 1.5% agar) containing the appropriate antibiotic. The plates were then incubated overnight at 37°C and single colonies chosen for validation.

2.11 Colony PCR

Colony PCR was used on S. aureus in order to confirm the presence of pIMAY. A portion of the colony of interest was resuspended in $100\,\mu$l of lysis buffer (20 mM Tris pH 8, 3 mM MgCl2, 0.5% Tween 20, 0.5% NP-40, $60\,\mu g/\mu$L proteinase K) and then incubated at 55°C for 1 hour, followed by an incubation at 95°C for 10 minutes. The contents were then centrifuged at 20000 g for 10 minutes in order to pellet cell debris, and then PCR was performed using $5\,\mu$l of the supernatant and Taq polymerase as per the manufacturer's instructions (NEB).

2.12 FACS sorting

pCN33-GFP and its derivatives were transformed into RN4220 by electroporation as previously described. Electrocompetent cells were prepared from the resulting transformants and pICS3:: RsaA or pICS3::RNAIII were transformed by electropora-

tion. Single isolated colonies were inoculated into 5 mL of BHI and grown at 37°C overnight, shaking at 200 rpm.

Cultures were diluted the day afterwards 1:40 into 2 mL of phosphate buffered saline. Translation of GFP fusions was monitored on the LSRFortessa Special Order Research Product (BD) from a $500\,\mu$l aliquot of PBS-diluted samples on a 530/30 nm bandpass filter. Sample acquisition of 100,000 events was performed on the built-in Diva (LSRFortessa SORP) software. Median fluorescence intensity (MFI) was determined using the FlowJo software where FSC and SSC were used to gate any fluorescence attributed to cellular background. The average median fluorescence intensity and standard deviations were calculated and plotted. These experiments were performed by Julia Wong in Jai Tree's group.

2.13 *S. aureus* genetic manipulation using pIMAY

Following pIMAY electroporation into RN4220, single colonies were selected and grown in TSB at 30°C overnight. Plasmid-containing phages were then generated and transduced into USA300 as described in Section 2.9, with the modification that overnight growth was performed at 30°C. Colonies were then screened through colony PCR using pIMAY MCS fw and pIMAY MCS rev to confirm the presence of replicating pIMAY in the cytosol.

Integration of the plasmid into the chromosome was then forced through growth in TSB-Cm at 37°C for 48 hours, and then re-inoculating into fresh TSB-Cm and growing again at 37°C for 24 hours. The culture was then serially diluted from 10^{-1} to 10^{-4}, and $50\,\mu$l of each dilution plated out onto TSA-Cm. Plates were grown overnight at 37°C, and then colonies were screened for integration of pIMAY into the chromosome through colony PCR. Potential integrants were screened for integration at the 5' end of the gene of interest through a forward primer designed to anneal 550 basepairs upstream of the gene, amplified using the MCS rev primer. Integration at the 3' end of the gene of interest was tested through a reverse primer designed to anneal 550 basepairs upstream of the gene, amplified using the MCS fw primer.

Once integration into the chromosome was confirmed, excision of the pIMAY backbone was forced through growth at 30°C for 4 days in TSB. Following this, the culture was serially diluted from 10^{-1} to 10^{-4} and $50\,\mu$l of each dilution was replica-plated out onto TSA and TSA-Cm and grown overnight at 37°C. Colonies which grew on TSA but not on TSA-Cm were then screened for loss of the pIMAY plasmid through colony PCR with MCS fw and MCS rev primers. Colonies confirmed to be

negative for pIMAY were then tested for successful HTF tag integration through Western blotting or qPCR in the case of sRNA deletion.

2.14 Northern blotting

Cells were harvested in 5 mL Eppendorf tubes and collected via centrifugation at 20000 g for 1 minute. Cells were resuspended in $550\,\mu$l of GTC-phenol and lysed via vortexing with $100\,\mu$l of zirconia beads (Biospek Products). RNA was then extracted through acid guanidinium thiocyanate-phenol-chloroform extraction as described in Chomczynski and Sacchi, 1987.

RNA was resolved on an 8% polyacrylamide TBE-urea gel and then transferred to a nitrocellulose membrane through electroblotting for 4 hours at 50 V. RNA was then crosslinked to the membrane through exposure to 1200 mJ of 254 nm UV radiation.

For hybridisation, membranes were firstly prehybridised in 7 mL of UltraHyb (Ambion) for 1 hour at 42°C. A DNA primer was concurrently radiolabelled using 3μL of 32P-γATP (Perkin Elmer) and 30 U of T4 PNK (NEB), afterwards being cleaned up through a Quick Spin Oligo Column (Roche). Following prehybridisation, the radiolabelled probe was added to the UltraHyb and left to hybridise overnight at 42°C. Membranes were then washed twice in 2xSSC with 0.5% SDS for 10 minutes, before rinsing briefly in 2xSSC. Membranes were then imaged using a phosphorimager screen and FujiFilm FLA-5100 scanner using the IP-S filter. For 5S rRNA, imaging was also performed through radiosensitive film (GE Healthcare).

Probes for northern blotting used in this study are detailed in Table 2.5.

Tab. 2.5: Oligonucleotide used for Northern blots

Probe	**Sequence (5' to 3')**
RsaA	GAGTAGTCTTCCTTGGGAGAAC
RNAIII	GCACTGAGTCCAAGGAAACTAACT
5S	TTAACTTCTGTGTTCGGCATGGGA

2.15 *In vitro* RNA transcription and RNA radiolabelling

$10\,\mu g$ plasmid was first linearised with 10 U of XhoI for 2 hours at 37°C and then isolated after being ran on an agarose gel. *In vitro* transcription was then carried out using a MEGAscript T7 transcription kit (ThermoFisher) as per the manufacturer's instructions. Briefly, the reagents were thawed on ice and then transcription reactions set up using $2\,\mu g$ of plasmid as a template in a $20\,\mu l$ volume. Reactions were incubated at 37°C for 2 hours. Afterwards, $3\,\mu l$ of 10X TURBO DNase buffer and $25\,\mu l$ of water were added, followed by 4 U of TURBO DNase in order to degrade the plasmid. The RNA was then purified through addition of RNAClean XP beads and processed as per the manufacturer's instructions (Beckman Coulter). The beads were finally resuspended in $11\,\mu l$ of structure buffer (25 mM Tris pH 7.5, 150 mM KCl, 1 mM $MgCl_2$). In order to check the purity of produced RNA, 1 μL was resolved on a 6% polyacrylamide TBE-urea gel and stained with SYBR Safe.

To radiolabel RNA samples, $1\,\mu g$ of RNA was firstly dephosphorylated using 1 U of TSAP (Promega). The RNA was then purified using RNAClean XP beads and eluted in $10\,\mu l$ of water. The dephosphorylated RNA was then radiolabelled using 10 U of T4 PNK and $3\,\mu l$ of ^{32}P-γATP (Perkin Elmer) for 1 hour at 37°C. The RNA was again purified using RNAClean XP beads, and resuspended in $10\,\mu l$ of 1X FA dye (NEB). The RNA was then heat denatured at 85°C for 3 minutes and snap-chilled on ice for 5 minutes before being resolved on a 6% polyacrylamide TBE-urea gel. The RNA was imaged through radiography, and the radioactive band excised and extracted by shaking in extraction buffer (20 mM Tris pH 7.5, 100 mM NaOAc, 10 mM EDTA, 0.1% SDS). Finally, the RNA was purified through phenol:chloroform extraction as described previously and resuspended in structure buffer. The concentration was obtained through use of a Qubit 4 and the RNA broad spectrum dye (Invitrogen).

2.16 Electrophoretic mobility shift assays

The RNAs of interest, generated through *in vitro* transcription as described above, were first refolded in a thermal cycler by heating to 95°C for 1 minute, then cooled slowly to 25°C and finally incubated at 25°C for 20 minutes. Binding reactions between the radiolabelled mRNA and cold sRNA were then set up in 1:0, 1:5, 1:10, 1:20, 1:80 and 1:320 molar ratios of mRNA:sRNA, and incubated at 25°C for 20 minutes. Native loading buffer was added to a 1X concentration (10% sucrose, 0.1X TBE, 0.04% bromophenol blue) and then RNA complexes were resolved on an 6%

acrylamide TBE gel, running at 400 V for 2 hours. The gel was then dried under vacuum for 1 hour at 80°C and imaged as described in section Northern blotting.

2.17 Structure probing

In vitro transcription was used to generate $1\,\mu g$ of RNAIII interacting and RsaA as previously described. The RNA was then refolded through incubation at 95°C for 1 minute, and then cooled slowly to 25°C over 2 minutes, and incubated at 25°C for 15 minutes. The SHAPE reagent 2-methylnicotinic acid imidazolide (NAI) was then added to a final concentration of 50 mM, and the reaction left to proceed for 20 minutes at room temperature.

This modified RNA was then used as an input for reverse transcription using SuperScript III (Invitrogen) using a ^{32}P-radiolabelled primer. Primers were annealed to the RNA through heating to 85°C for 3 minutes, followed by snap-chilling on ice. The reverse transcription was then performed for 1 hour at 45°C and afterwards treated with 25 U of RNase If (NEB) for 30 minutes at 37°C. The reaction was stopped through addition of 2X RNA loading dye (NEB). Sequencing ladders were generated using the Sequenase v2.0 kit (Affymetrix) according to the manufacturer's instructions, utilising plasmid used for the original *in vitro* transcription. Samples were then resolved on a 6% polyacrylamide TBE-urea gel, dried under vacuum for 1 hour at 80°C and then imaged using a phosphorimager screen and a FujiFilm FLA-5100 scanner

2.18 CLASH

Described here is the final CLASH protocol following optimisation.

S. aureus strains expressing chromosomally tagged RNase Y and RNase III were grown overnight in TSB at 37°C with shaking at 200 rpm. This starter culture was then diluted 1:500 in fresh TSB and regrown to OD_{600}=3. As a control, an untagged parental strain was used to examine background protein binding to the antibody and nickel beads.

Cells were then crosslinked and harvested as described in Section 2.2. For CLASH libraries, cells were shifted to RPMI for 15 minutes and crosslinked with 250 mJ of 254 nm UV radiation. Following harvesting, the filters were stored in Falcon tubes at -80°C until required.

Cells were harvested from the filters through two washes with 10 mL of PBS, and then pelleted through centrifugation at 4600 g for 10 minutes, 4°C. The supernatant was discarded and the cell pellets weighed. The pellets were then resuspended in 2 volumes of TN150-lysostaphin (50 mM Tris pH 7.8, 150 mM NaCl, $100\,\mu$gμL lysostaphin, 0.1% NP-40, 0.5% Triton X-100), and 60 U of DNase RQ1 (Promega) and 200 U of SUPERasin (Invitrogen) were added. The cells were then incubated for 1 hour at 20°C in order for the lysostaphin to degrade the outer cell wall. The cells were then transferred to 15 mL Falcon tubes and lysed through bead beading with 0.1 mm zirconia beads (Biospek Products) for 5 minutes, and 2 volumes of TN150 antipeptidase was then added (50 mM Tris pH 7.8, 150 mM NaCl, 0.1% NP-40, 0.5% Triton X-100, 10 mM EDTA, 1 mini cOmplete protease inhibitor per 10 mL (Roche)) The beads were then separated from the lysate by centrifugation at 4600 g for 20 minutes at 4°C and then the lysate transferred to Eppendorfs. The insoluble and soluble fractions of the lysate were then separated through centrifugation at 20000 g for 20 minutes at 4°C.

Magnetic anti-FLAG M2 beads (Sigma Aldrich) were washed three times in TN150 (50 mM Tris pH 7.8, 150 mM NaCl, 0.1% NP-40, 0.5% Triton X-100), with $75\,\mu$l taken for each sample. The beads were resuspended in a noted volume, dependent upon the number of samples, and distributed equally between the cleared lysates. The beads were then incubated with the lysate for 2 hours at 4°C with rotation.

Following capture, the beads were washed three times in TN1000 (50 mM Tris pH 7.8, 1 M NaCl, 0.1% NP-40, 0.5% Triton X-100) for 10 minutes at 4°C with rotation. The beads were then rinsed three times in TN150, and then resuspended in a final volume of $250\,\mu$l TN150. In order to cleave the RNases from the FLAG beads, $10\,\mu$l of homemade TEV protease was added, and the samples incubated for 2 hours at room temperature with rotation.

Following cleavage, an extra $350\,\mu$l of TN150 was added to the samples and the eluate collected following separation from the beads using a magnetic rack. The eluates were then RNase digested through incubation for 7 minutes at 20°C with $1\,\mu$l of a 1:100 dilution of RNace-It (Agilent). The RNase digestion was stopped through the addition of 0.4 g of guanidium hydrochloride (GuHCl, Sigma Aldrich). Following digestion, $100\,\mu$l of nickel-NTA agarose beads (Qiagen) was added, prewashed in wash buffer 1 (50 mM Tris pH 7.8, 0.1% NP-40, 5 mM β-mercaptoethanol, 0.5% Triton X-100, 300 mM NaCl, 10 mM imidazole, 6M GuHCl) and proteins were captured overnight.

The capture solutions were then transferred to Pierce spin columns and the beads washed three times with wash buffer 1 and three times with NP-PNK (50 mM Tris pH 7.8, 0.1% NP-40, 5 mM β-mercaptoethanol, 0.5% Triton X-100, 10 mM

$MgCl_2$). Afterwards, the RNAs were dephosphorylated on-column using 4 U of TSAP (Promega) in the presence of 80 U of rRNasin (Promega) in 1X PNK buffer (50 mM Tris pH 7.8, 10 mM MgCl2, 10 mM β-mercaptoethanol, 0.1% Triton X-100) for 1 hour at 20°C. The beads were then washed once with wash buffer 1 in order to inactivate the enzyme and then three times with NP-PNK to remove residual guanidium. The RNAs were then radiolabelled at the 5' end using 30 U of T4 PNK and 3μl of ^{32}P-ATP in 1X PNK buffer for 100 minutes at 20°C, after which cold ATP to a final concentration of 1 mM was added and the reaction left to proceed for another 40 minutes in order to ensure complete phosphorylation of the RNAs at the 5' end. Following phosphorylation, the beads were washed three times with wash buffer 1 and three times with NP-PNK.

In order to increase the number of chimeric reads attained, a dedicated intermolecular ligation step was performed. In this, 15 U of T4 RNA ligase 1 was used to ligate the 5' end of one RNA to the 3' end of one in complex with it. This reaction was carried out in the presence of 40 U of rRNasin and 1 mM ATP, and left to proceed for 2 hours. Afterwards, the beads were washed once with wash buffer 1 and three times with NP-PNK.

Sequencing adaptors were then ligated to the ends of the RNAs. First, an L5 adaptor was ligated to the 5' end using 200 μmoles of adaptor and 40 U of T4 RNA ligase, in the presence of 80 U of rRNasin and 1 mM ATP in 1X PNK buffer, for 16 hours at 16°C. The beads were subsequently washed once with wash buffer I, and three times with NP-PNK. Afterwards, 60 μmoles of App_PE adaptor was ligated onto the 3' end using 600 U of T4 RNA Ligase 2 truncated K227Q (NEB). This reaction was carried out in 1X PNK buffer with 10% PEG-8000 and 30 U of rRNasin for 6 hours at 25°C. Afterwards, the beads were washed once in wash buffer I and three times in wash buffer two (50 mM Tris pH 7.8, 10 mM β-mercaptoethanol, 0.1% NP-40, 0.5% Triton X-100, 50 mM NaCl, 10 mM imidazole).

The protein-RNA complexes were then eluted from the beads through addition of 200μl of elution buffer (wash buffer two with 250 mM imidazole), repeated for a total of two times. The proteins were then pooled and precipitated through addition of trichloroacetic acid (Sigma Aldrich) to a final concentration of 20%, and left to precipitate on ice for 20 minutes. The samples were then centrifuged at 20000 g for 20 minutes at 4°C. The pellets were then washed with 800μl of acetone and dried for two minutes in a fume hood, before being dissolved in 20μl of 1X NuPAGE buffer (Novex) and resolved on a 4-12% Bis-Tris gel. The protein-RNA complexes were visualised through autoradiography as described previously and the gel piece containing these excised. The RNAs were then extracted through incubation in 4 mL of extraction buffer (50 mM Tris pH 7.8, 0.1% NP-40, 5 mM β-mercaptoethanol, 1% SDS, 5 mM EDTA, 50 mM NaCl, $60\mu g/\mu$L proteinase K) at 55°C for 2 hours.

Following this, the RNAs were purified through phenol:chloroform extraction as previously described and the RNA resuspended in $20\,\mu$l of DEPC-water.

The RNAs were afterwards reverse transcribed using the PE_reverse primer and SuperScript IV according to the manufacturer's instructions. The template RNA was subsequently degraded through addition of 10 U of RNase H (NEB). Afterwards, the cDNA was purified through RNAClean XP beads and resuspended in a final volume of $11\,\mu$l. Half of this cDNA was then used a template for PCR using Pfu polymerase, using a BC and the P5 primers. The cycling conditions were as follows: 95°C for 2 min; 24 cycles of 95°C for 20s, 52°C for 30s and 72°C for 1 min; final extension of 72°C for 5 min. The PCR product was treated with 40 U of Exonuclease I (NEB) in order to degrade free primer and the DNA library purified using RNAClean XP beads. The library was then resolved on a 2% MetaPhor (Lonza) gel and 175-300 bp fragments were excised and gel extracted through a MinElute column. The library was quantified using a 2100 Bioanalyzer and a DNA HS assay (Agilent).

Individual libraries were then pooled together to produce an equimolar solution and sequenced through paired-end sequencing on an Illumina HiSeq 4000 platform (Edinburgh Genomics).

All L5 and BC adaptors are described previously (Iosub et al, 2019).

2.19 Western blotting

Western blotting was used to analyse the efficiency of protein capture when optimising CLASH. Samples for western blotting were taken after cell lysis, lysate clearing, and TEV elution. Samples were resolved on a homemade 8% polyacrylamide gel and transferred to a nitrocellulose membrane for 2 hours at 100V. Membranes were blocked for 1 hour in blocking solution (5% non-fat milk, 0.1% Tween-20 in 1X phosphate-buffered saline). Primary antibody probing was performed overnight at 4°C using a rabbit anti-TAP antibody (1/5000, ThermoFisher). The membrane was then washed in PBST (PBS with 0.1% Tween-20) for 10 minutes, three times. In order to detect the primary antibody, an HRP-linked goat anti-rabbit antibody (1/500, Abcam) in blocking solution was incubated with the membrane for 1 hour at room temperature. The membrane was washed again three times with PBST, rinsed in PBS and imaged using Pearce enhanced chemiluminescence solutions (ThermoFisher) according to the manufacturer's instructions.

While optimising CLASH, the protein-RNA complexes were blotted onto nitrocellulose using an iBlot (Invitrogen), and the proteins visualised by Western blotting after autoradiography using the method described above.

2.20 CRAC

CRAC was carried out on S. aureus USA300 *rpoC*::HTF. This was performed in a nearly identical way as the CLASH protocol, with the only differences being that the lysostaphin treatment, TSAP and T4 PNK labelling steps were carried out at 37°C. The dedicated intermolecular ligation step of CLASH was also left out.

2.21 RNAtag-Seq

RNA sequencing libraries of cells grown in TSB and shifted to RPMI medium were generated utilising the RNAtag Seq protocol (Shishkin et al., 2015).

RNA was first extracted from cells as previously described and then quantified using a 2100 Bioanalyzer and the RNA nano assay. Ten samples were processed simultaneously with $1\,\mu$g taken for each. 20 U of SUPERasin was then added and the volume increased to $16\,\mu$l. Afterwards, $4\,\mu$l of 10X FastAP buffer (ThermoFisher) was added and then the RNA was fragmented in a thermal cycler through incubation at 92°C for 2 minutes. The samples were then snap-chilled on ice. The RNA was then dephosphorylated and DNase-treated. 10 U of FastAP and 8 U of TURBO DNase were added, alongside 40 U of rRNasin. The reaction was left to proceed for 30 minutes at 37°C. Afterwards, the RNA was purified using RNAClean XP beads and eluted in $11\,\mu$l. A unique sequencing adaptor was then ligated onto the 3' end of the RNA. $10\,\mu$l of RNA was taken and 200 μmoles of adaptor was ligated onto it through addition of 36 U of T4 RNA ligase 1 in the manufacturer's T4 RNA ligase buffer, supplemented with 9% DMSO, 1 mM ATP, 20% PEG-8000 and 16 U of rRNasin. The reaction was left to proceed for 90 minutes at 22°C. After the ligation step, $120\,\mu$l of RLT buffer (Qiagen) was added to inhibit the enzyme. The samples were then pooled and extracted through phenol:chloroform precipitation, with the final RNA pellet resuspended in $26\,\mu$l of DEPC-H2O. The concentration of the RNA sample was determined using a Qubit 4 and the RNA broad range assay and 5 μg of sample was then taken, made up to $26\,\mu$l and used as input for rRNA depletion via a Ribo-Zero kit (Illumina), carried out according to the manufacturer's instructions. The rRNA-depleted RNA was then reverse transcribed into cDNA using the AR2 primer and SuperScript IV, according to the manufacturer's instructions. Afterwards, the RNA was degraded through addition of NaOH to a final concentration of 100 mM. The samples were incubated

at 70°C for 12 minutes, at then HCl was added to a final concentration of 100 mM in order to neutralise the NaOH. The cDNA was then purified using RNAClean XP beads and eluted in $11\,\mu$l of water. The 3Tr3 adaptor was then ligated onto the 5' end of the cDNA. $1\,\mu$l of 80 μM adaptor was added to the DNA using CircLigase according to the manufacturer's instructions. The reaction was then cleaned up using RNAClean XP beads, and then this clean up repeated a second time in order to ensure that as much primer as possible had been removed. The cDNA was then amplified using AccuPrime polymerase according to the manufacturer's instructions, utilising the 2P_univP5 and 2P_rev_2 primers. The cycling conditions were as follows: 98°C for 2 min; 20 cycles of 98°C for 30s, 55°C for 30s and 65°C for 1 min; final extension of 65°C for 10 min. The PCR product was treated with 40 U of Exonuclease I (NEB) in order to degrade free primers and the DNA library was purified using RNAClean XP beads. The library was then resolved on a 2% MetaPhor (Lonza) gel and 175-300 bp fragments were excised and gel extracted through a MinElute column. The library was quantified using a 2100 Bioanalyzer and a DNA HS assay (Agilent).

The RNA adaptors and PCR primers used for are detailed by Shishkin et al, 2015.

2.22 Reverse-transcription quantitative PCR

qRT-PCR was performed on RNA samples extracted from cells shifted from TSB to RPMI and from the TSB to TSB control.

First, a DNase treatment step was performed to ensure that no contaminating DNA was present in the sample; $8\,\mu g$ of RNA was taken and digested with 1.8 U of TURBO DNase according to manufacturer's instructions for 1 hour at 37°C in the presence of 2 U of SUPERasin. The RNA was purified using RNAClean XP beads and then eluted in a final volume of $20\,\mu$l.

The RNA concentration was then quantified using a Qubit 4 and the RNA broad range assay and diluted to a concentration of $5\,\mu g/\mu$L. Afterwards, the RNA was aliquoted and stored at -80°C until required.

The PCR was then performed using Luna Universal One-Step RT-qPCR kit (NEB), according to the manufacturer's instructions, and 5 ng of RNA. The PCR was run on a LightCycler 480 (Roche) with the cycling conditions as recommended by NEB: 55°C for 10 min for reverse transcription; 95°C for 1 min for initial denaturation; 40 cycles with single acquisition mode of 95°C for 10s and then 60°C for 30s; melt curve at 65°C (0.11 ramp rate with 5 acquisitions per °C, continuous).

The IDEAS2.0 software was used to calculate Ct values using the absolute quantification/fit points method with default parameters, and the fidelity of the PCR was examined through melt curve genotyping analyses. To calculate the relative fold-change of genes, the $2^{d}dCt$ method was employed using 5S as a control. Each qPCR experiment was performed in technical triplicate. For final data analyses, the mean and standard error of the mean of three biological triplicates was calculated and plotted.

2.23 Bioinformatic analyses

Pre-processing of raw sequencing data

Raw sequencing data were first processed using the pyCRAC package (Webb et al., 2014), available from https://bitbucket.org/sgrann/. In particular, the CRAC_Pipeline_PE.py pipeline was used with the following command:

```
CRAC_Pipeline_PE.py -f forward_reads.fastq -r reverse_reads.fastq --novoindex
reference_genome.novoindex --gtf USA300/JKD6008.gtf -c chromosome_length_file.txt
-b L5_adaptors.txt -a adaper_file.fasta
```

This pipeline firstly demultiplexed the data based on the in-read barcode sequences found in the L5 adaptors using pyBarcodeFilter.py. Flexbar then removed the 3' adaptor sequences and any flanking nucleotides with a Phred score below 23. The reads were then collapsed to remove PCR duplicates and then aligned to either the *S. aureus* JKD6008 or USA300 genome using NovoAlign (www.novoalign.com) and a Gene Transfer Format (GTF) file. The GTF file was kindly provided by Jai Tree. Note that in the analyses carried out here, UTR coordinates were simply defined as 100 nucleotides upstream and downstream of the start and end codons respectively. The pyReadCounters.py script then used the output from NovoAlign to quantify the number of reads for each transcriptional unit.

Hyb analyses

The hyb pipeline was used to detect and annotate chimeric reads (Travis et al., 2014). As hyb does not accept paired-end data, FLASh was used to combine the forward and reverse reads of the demultiplexed .fastq files into a single contig with the following parameters (FLASh is available from https://ccb.jhu.edu/software/FLASH/):

```
flash forward_reads.fastq reverse_reads.fastq -m 5 -M 65 -O -r 65
```

FLASh generates an output file containing the uncombined forward and reverse reads and the uncombined forward reads were combined with the FLASh output:

```
cat combined_forward_reverse.fastq uncombined_forward > final.fastq
```

A hyb database was then created and exported using a .fasta file of the genome:

```
make_hyb_db genome.fasta
export HYB_DB=/hyb_db
```

Hyb was then used to detect chimeras:

```
hyb preprocess qc=flexbar link=NAGATCGGAAGAGCACACG check detect
align=bowtie2 in=final.fastq db=hyb_db anti=1
```

In order to calculate the folding energy and to map the genomic coordinates to gene names, custom scripts written by Sander Granneman were used. When visualising hybrids using a genome browser, the .ua_hyb output file was converted to a GTF file using a custom script, written by Sander Granneman. GTF files could then be converted to .sgr files if desired, using the pyCRAC pyGTF2sgr.py script.

RNAtag-Seq analysis

Following pre-processing of the data using the pyCRAC_Pipeline_PE as previously described, the output from pyReadCounters.py for each experimental condition was merged together. In order to normalise the data and to account for variations in sequencing depth, the data was converted into 'Transcripts per million reads' (TPM). The number of raw read counts for each gene was first divided by the length of the gene in order to create a normalised transcript expression value. The sum of all normalised transcript values was then divided by 1,000,000 in order to create a scaling factor, and then each normalised transcript expression value was divided by this scaling factor to yield the TPM value. All data could then be normalised to the t0 sample when desired in order to examine relative fold change.

In order to cluster genes into common expression patterns, the 'Short Time-series Expression Miner' (STEM) was used (Ernst and Bar-Joseph, 2006). Data was $\log_2$ normalised and clustered using the STEM clustering method with 50 model profiles. In order to examine which genes showed changed expression following shift to RPMI as compared to the control, Differential expression analysis 2 (DESeq2 (Love et al., 2014)) was used. Only differentially expressed genes with an adjusted p-value of 0.05 or lower were considered for future analysis.

Conservation analyses and sequence alignments

When comparing sRNA sequences between *S. aureus* strains, the *Staphylococcal* Regulatory RNA Database (SRD, Sassi et al., 2015) was used to retrieve the homologous sequences. When comparing sequences between Staphylococcal species, BLAST was used with the *S. aureus* USA300 sequence used as a reference. JalView was then used to align the sequences using the MAFFT algorithm and visualise the conservation (Waterhouse et al., 2009).

2.24 Bibliography

Chomczynski P, Sacchi N. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Analytical Biochemistry 162:156–159. doi:10.1016/0003-2697(87)90021-2

Corrigan RM, Foster TJ. 2009. An improved tetracycline-inducible expression vector for *Staphylococcus aureus*. Plasmid 61:126–129. doi:10.1016/j.plasmid.2008.10.001

Diep BA, Gill SR, Chang RF, Phan TH, Chen JH, Davidson MG, Lin F, Lin J, Carleton HA, Mongodin EF, Sensabaugh GF, Perdreau-Remington F. 2006. Complete genome sequence of USA300, an epidemic clone of community-acquired meticillin-resistant *Staphylococcus aureus* 367:9.

Ernst J, Bar-Joseph Z. 2006. STEM: a tool for the analysis of short time series gene expression data. BMC Bioinformatics 7:191. doi:10.1186/1471-2105-7-191

Ivain L, Bordeau V, Eyraud A, Hallier M, Dreano S, Tattevin P, Felden B, Chabelskaya S. 2017. An *in vivo* reporter assay for sRNA-directed gene control in Gram-positive bacteria: identifying a novel sRNA target in *Staphylococcus aureus*. Nucleic Acids Research 45:4994–5007. doi:10.1093/nar/gkx190

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15. doi:10.1186/s13059-014-0550-8

McDougal LK, Steward CD, Killgore GE, Chaitram JM, McAllister SK, Tenover FC. 2003. Pulsed-Field Gel Electrophoresis Typing of Oxacillin-Resistant *Staphylococcus aureus* Isolates from the United States: Establishing a National Database. J Clin Microbiol 41:5113–5120. doi:10.1128/JCM.41.11.5113-5120.2003

Monk IR, Shah IM, Xu M, Tan M-W, Foster TJ. 2012. Transforming the Untransformable: Application of Direct Transformation To Manipulate Genetically *Staphylococcus aureus* and *Staphylococcus epidermidis*. mBio 3:e00277-11. doi:10.1128/mBio.00277-11

Nair D, Memmi G, Hernandez D, Bard J, Beaume M, Gill S, Francois P, Cheung AL. 2011. Whole-Genome Sequencing of *Staphylococcus aureus Strain RN4220*, a Key Laboratory Strain Used in Virulence Research, Identifies Mutations That Affect Not Only Virulence Factors but Also the Fitness of the Strain. Journal of Bacteriology 193:2332–2335. doi:10.1128/JB.00027-11

Norrander J, Kempe T, Messing J. 1983. Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene 26:101–106. doi:10.1016/0378-1119(83)90040-9

Nues R van, Schweikert G, Leau E de, Selega A, Langford A, Franklin R, Iosub I, Wadsworth P, Sanguinetti G, Granneman S. 2017. Kinetic CRAC uncovers a role for Nab3 in determining gene expression profiles during stress. Nature Communications 8:12. doi:10.1038/s41467-017-00025-5

Sassi M, Augagneur Y, Mauro T, Ivain L, Chabelskaya S, Hallier M, Sallou O, Felden B. 2015. SRD: a *Staphylococcus* regulatory RNA database. RNA 21:1005–1017. doi:10.1261/rna.049346.114

Shishkin AA, Giannoukos G, Kucukural A, Ciulla D, Busby M, Surka C, Chen J, Bhattacharyya RP, Rudy RF, Patel MM, Novod N, Hung DT, Gnirke A, Garber M, Guttman M, Livny J. 2015. Simultaneous generation of many RNA-seq libraries in a single reaction. Nature Methods 12:323–325. doi:10.1038/nmeth.3313

Taylor RG, Walker DC, Mclnnes RR. 1993. *E. coli* host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. Nucleic Acids Research 21:1677–1678. doi:10.1093/nar/21.7.1677

Travis AJ, Moody J, Helwak A, Tollervey D, Kudla G. 2014. Hyb: A bioinformatics pipeline for the analysis of CLASH (crosslinking, ligation and sequencing of hybrids) data. Methods 65:263–273. doi:10.1016/j.ymeth.2013.10.015

Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. 2009. Jalview Version 2–a multiple sequence alignment editor and analysis workbench. Bioinformatics 25:1189–1191. doi:10.1093/bioinformatics/btp033

Waters SA, McAteer SP, Kudla G, Pang I, Deshpande NP, Amos TG, Leong KW, Wilkins MR, Strugnell R, Gally DL, Tollervey D, Tree JJ. 2017. Small RNA interactome of pathogenic *E. coli* revealed through crosslinking of RNase E. The EMBO Journal 36:374–387. doi:10.15252/embj.201694639

Webb S, Hector RD, Kudla G, Granneman S. 2014. PAR-CLIP data indicate that Nrd1-Nab3-dependent transcription termination regulates expression of hundreds of protein coding genes in yeast. Genome Biology 15:R8. doi:10.1186/gb-2014-15-1-r8

CLASH in *S. aureus*

3

3.1 Introduction

RNAs often rely on RNA-binding proteins (RBPs) to exert their function and this has led to great interest in understanding the dynamics between RBPs and RNAs. A first breakthrough came with the emergence of RIP-chip, where an RBP is immunoprecipitated and its bound RNAs identified through microarrays or sequencing (Keene et al., 2006). Building on RIP-chip led to the development of improved protocols such as CLIP, HITS-CLIP and CRAC (Granneman et al., 2009; Licatalosi et al., 2008; Ule et al., 2003).

In order to study RNA-RNA interactions *in vivo*, a modified CRAC technique, termed CLASH, was developed (Kudla et al., 2011). This technique relies on tagging an RBP with a tandem affinity purification tag, followed by UV-crosslinking of cells expressing this recombinant RBP in order to covalently link it to its bound RNAs. The RBP is then purified through two affinity capture steps, the second of which is carried out under denaturing conditions to greatly reduce background signal. Crucially, if two RNAs bound to an isolated RBP are in close proximity to one another (such as during an sRNA-mRNA interaction), enzymatic treatment can induce an intermolecular ligation between them to make a single hybrid RNA. Adaptor sequences are then ligated onto bound RNAs, the RNA is isolated and finally a cDNA library of the RNA is made and sequenced (Figure 1.10).

One particular technical challenge in CRAC and CLASH is retaining RNA integrity until the RNase digestion step. The aim of this step is to trim the cellular RNAs to sizes of around 50-100 bases to facilitate sequencing. After trimming, the RNAs are then transferred to a highly denaturing buffer, thus protecting them from further degradation. However, RNA degradation can occur before this digestion step, particularly following cell lysis due to the release of cellular RNases. This degradation results in RNAs which are too short to sequence effectively. As such, steps must be taken to minimise RNA degradation.

Another challenge is optimisation of the crosslinking step. Cells are exposed to 254 nm UV radiation in order to induce covalent bonds between RBPs and their bound RNAs. However, over-crosslinking has been reported to induce RNA and protein

degradation (personal communications with Ross Cordiner). Additionally, as over-crosslinking involves increasing the UV irradiation time, the exposed cells activate DNA damage response pathways which can introduce noise into the produced experimental data (Nues et al., 2017). As such, optimisations to find a UV exposure time that produces sufficient RBP-RNA crosslinking but minimises RNA degradation must be performed.

CRAC and CLASH have been successfully utilised in human cells, yeast and *E. coli*, (Granneman et al., 2009; Helwak and Tollervey, 2014; Iosub et al., 2018; Waters et al., 2017). However, they have never been performed in Gram-positive bacteria. In this study, an initial experiment was performed using a published CLASH protocol optimised for use in cultured human cells (Helwak and Tollervey, 2014), but this method resulted in both very poor RNA and protein recovery. Thus, it was hypothesised that *S. aureus* itself presents unique challenges to the CLASH protocol. As such, the first aim of this project was to create an optimised CLASH method for use in *S. aureus*.

In this chapter I detail the numerous optimisation steps that were undertaken which aimed to improve cell lysis and RBP recovery, alongside minimising RNA degradation. Additionally, in order to maximise the potential of the data after sequencing, numerous additions were made to the *in silico* data analysis pipeline. In sum, the optimisations performed resulted in the creation of a CLASH protocol that is applicable for use in *S. aureus*.

3.2 Optimisation of CLASH for *S. aureus*

RNase III and RNase Y were chosen as candidate bait proteins for the discovery of sRNA – target interactions as a result of their already proven role in mediating sRNA activity (Section 1.10).

The outcome of the initial CLASH test using the published methodology was very poor, with the autoradiogram requiring a 56-hour exposure in order to visualise the radiolabelled RNA (Figure 3.1A, upper panel). The requirement for such a long exposure represents the fact that not enough RNA was being obtained after the purification. This could be due to inefficient RBP purification (and thus a poor harvest of crosslinked RNA-RBP complexes) and RNA degradation occurring during the purification itself. The western blot showed that very little protein was actually purified (Figure 3.1A, lower panel), and since these proteins are themselves RNases, both purification efficiency and minimising RNA degradation were taken as avenues of optimisation. Overall, my first aim was to reduce this 56-hour autoradiogram

exposure to around a 3-hour exposure, which would represent purifying enough RNA to obtain high quality CLASH data. We used RNase III for the optimisations as this was considered a higher priority than RNase Y due to its known dsRNA binding activity. RNase III is also a soluble protein, as opposed to the membrane-bound RNase Y, thus was thought to have a higher chance of successful purification. The optimisations were performed in the *S. aureus* strain JKD6008 due to the necessary tagged strains already being available (kind gift from Jai Tree, Sydney).

Due to expense, modifying several purification variables was tested concurrently. Firstly, I investigated if treating *S. auerus* with lysostaphin, an enzyme which degrades the outer peptidoglycan layer of the cell wall, prior to mechanical disruption would improve cell lysis. In combination with this, EDTA was added to the buffer used in the primary anti-FLAG capture. Since RNase Y and RNase III are dependent on divalent ions for catalysis, I hypothesised that the addition of EDTA would inactivate the activity of these RNases, as well as that of other endogenous RNases. However, as lysostaphin is zinc-dependent, EDTA was only added after this initial enzymatic treatment. Finally, we also tested if longer UV exposure times would aid crosslinking efficiency.

As seen in Figure 3.1.1B, changing the UV exposure time had no major impact on crosslinking efficiency. UV exposure times of 13, 30, 60 and 90 seconds were tested using the UVO3 Vari-X-linker (https://www.vari-x-link.com). These values relate to roughly 55, 125, 250 and 375 mJ of UV. The 13 second exposure gave both the best autoradiogram and western blot signal, with a continuous decrease in intensities being observed as the cells were crosslinked for longer periods. Similar results in human cells have been observed in the Granneman lab (Ross Cordiner, unpublished data) and it is hypothesised that excessive UV exposure results in both RNA and protein degradation. Thus, a UV exposure time of 13 seconds was chosen for future crosslinkings.

The combined addition of a lysostaphin treatment and EDTA in the primary capture buffer had a positive effect on both protein and RNA recovery. In the experiment represented by Figure 3.1B, sufficient radiolabelled RNA was present to facilitate an overnight exposure for the autoradiogram, as opposed to the previous 56 hours. The lysostaphin treatment is likely to have improved significantly the cell lysis and therefore RNase III recovery, as concluded from the strong western blot signal. It is impossible to deduce absolutely whether EDTA helped to reduce RNA degradation due to the concurrent testing of multiple experimental variables, but it was maintained in future experiments regardless.

CLASH has been performed in *E. coli* on RNase E. (Waters et al, 2017). Here, the authors suggested that the first anti-FLAG step should be done overnight to

get optimal recovery of the protein. However, I hypothesized that this lengthy incubation step would result in unwanted degradation of the RNA. RNAs only become protected from degradation during the secondary capture as this is carried out under denaturing conditions. Thus, I next tested if reducing the initial primary FLAG capture to two hours would reduce RNA degradation. Although an overnight capture will presumably capture more protein than that of a two hour, this leaves the RNAs crosslinked to the RBP in contact with endogenous RNases for an extended period of time. Thus, getting to the secondary capture stage as quickly as possible was hypothesised to maximise RNA recovery, even if at the cost of reduced protein capture. In order to compensate for this reduced capture time, the amount of anti-FLAG resin was increased 8-fold. Additionally, it was noted that autoradiograms from previous experiments gave too defined a signal, indicating over-digestion of the crosslinked RNAs. Thus, different concentrations of RNase-IT, the RNase cocktail used for RNA digestion, were tested.

As seen in Figure 3.1C, the two-hour FLAG capture greatly increased the amount of intact RNAs present, facilitating a three-hour autoradiogram exposure. In the sample untreated with RNase-It, (Figure 3.1C, lane 1), it is also observed that a wide range of RNAs are being pulled down with the protein, with signal extending from 32 kDa up to 58 kDa. Agreeably, it is also seen that decreasing the amount of RNase-IT results in an a more diffuse autoradiogram signal, and the 1/100 concentration was chosen for future experiments.

As RNase Y is known to be membrane bound (Khemici et al., 2015), Triton X- 100 (TX-100) was added to all buffers in an attempt to improve RNase Y purification. This was hypothesised to aid RNase Y solubility but it was unknown if the presence of TX-100 would negatively impact the affinity capture or enzymatic steps. Thus, included in this test was RNase III as a positive control. Additionally, a parental, untagged strain was used in order to visualise non-specific binding to the beads.

Figure 3.1D shows that TX-100 had no negative impacts on the purification of RNase III nor on the phosphorylation of its bound RNAs. As such, I concluded that TX-100 has no noticeable adverse effects on the experimental system. With regards to RNase Y, the protein purification is very inefficient in the absence of TX-100, with a poor autoradiogram signal observed after 3 hours and only a small amount of protein in the western blot. However, the addition of TX-100 greatly improved protein capture ($\sim$4X by densitometry), and this facilitated an observable autoradiograph signal after three hours of exposure. Finally, the parental strain exhibits no observable signal, indicating that the signal observed is specific to the chosen proteins.

Using the described optimisations, CLASH was performed on RNase Y and RNase III in cells grown in TSB, and also on cells transferred from TSB to RPMI medium.

As seen in Figure 3.1E, upper panel, the signal from the autoradiogram of RNases purified from cells exposed to RPMI is significantly stronger than that of cells grown only in TSB, indicating that these enzymes are highly active during the stress response and are binding more RNAs. Additionally, this experiment required just a 1.5-hour autoradiogram exposure, representing the cumulative effects of including the optimisations. The cDNA libraries from these CLASH experiments are shown in Figure 3.1E, lower panel, and these were successfully sequenced.

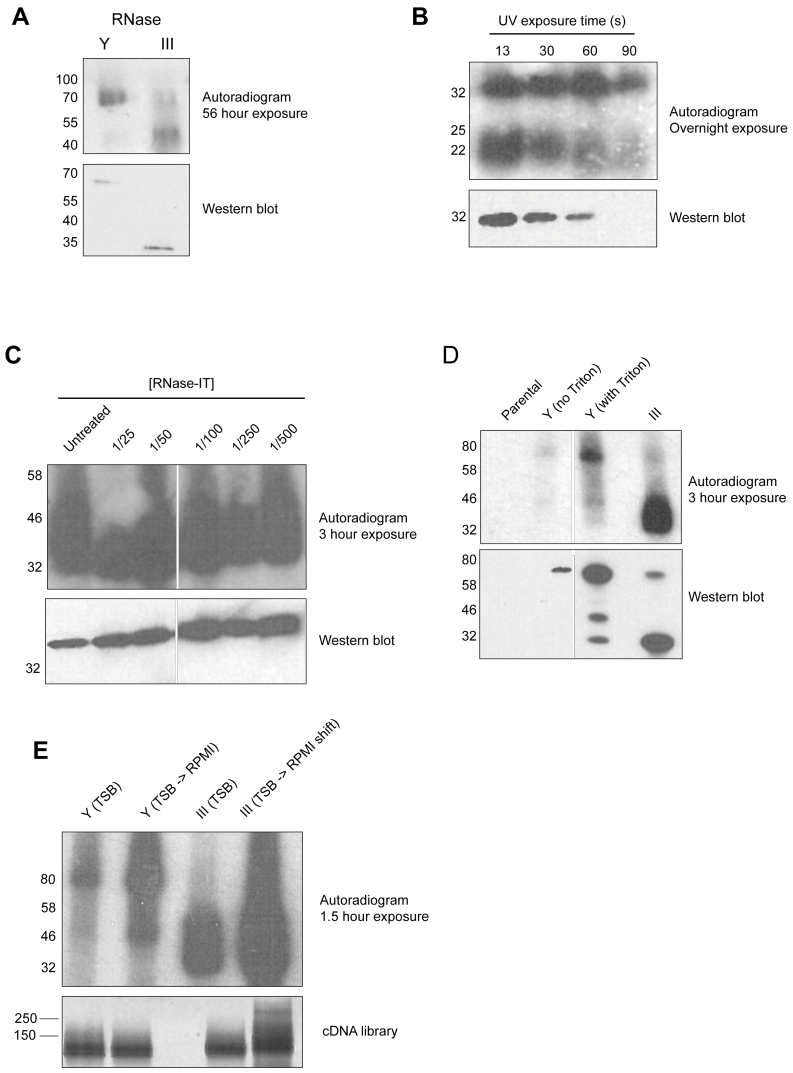

Fig. 3.1: Optimisation of the CLASH protocol. **A.** Initial CLASH experiment using published protocols. The very long autoradiogram exposure represents poor efficiency of RNA recovery. **B.** Testing the effect of different crosslinking times (13, 30, 60 and 90 seconds) and also the addition of a lysostaphin treatment and EDTA to the primary capture buffer on recovery of crosslinked RNAs. RNase III was used as the bait. **C.** Testing the effect of different RNase-IT concentrations (untreated, 1/25, 1/50, 1/100, 1/250, 1/500), and also a two hour FLAG capture, on recovery of crosslinked RNAs. RNase III was used as the bait. **D.** Testing the effect of Triton X-100 on recovery of crosslinked RNAs with RNase Y and RNase III. **E.** Autoradiogram and cDNA library of a successful CLASH experiment on RNase Y and RNase III, obtained from cells grown in TSB and also from cells shifted from TSB to RPMI for 15 minutes.

3.3 Discussion

CLASH has proven to be a useful tool for identifying RNA-RNA interactions in human cells, yeast and *E. coli*, but use of an already published protocol failed when applied to *S. aureus*. As such, optimisations were performed and these yielded conditions to be used for subsequent CLASH experiments.

Firstly, this optimised protocol uses a lysostaphin treatment to aid cell lysis. This reflects the inherent difficulty in lysing Gram-positive bacteria due to their tough outer cell wall composed of peptidoglycan. Previous CLASH protocols utilised only bead beating for cell lysis, but it is clear that this was not sufficient for *S. aureus*. In *S. aureus*, the peptidoglycan wall is typically 20-30 nm thick and is crosslinked through pentaglycine bridges. Lysostaphin is able to hydrolyse these crosslinks and thus greatly reduce the strength of the cell wall, and this is likely to facilitate subsequent lysis through bead beating and detergent usage.

Additionally, the protocol described here has taken several steps to minimise RNA degradation during the initial capture stages. EDTA was added to the primary capture buffer and the FLAG capture was reduced to 2 hours. It is hypothesised that EDTA is able to chelate the metal ions that are essential for RNase activity, thus inactivating them, while reducing the FLAG capture time decreases the contact between the RNAs and the RNases.

Finally, the protocol described here addresses other logistical issues necessary for successful CLASH. After testing several UV exposure times, 125 mJ of UV has been chosen for RNA-RBP crosslinking which represents around 30 seconds in the Vari-X-Linker. Additionally, a 1/100 RNase-IT concentration is used for trimming bound RNAs, and it was confirmed that Triton X-100 could be used in order to aid solubility of membrane-bound proteins without any adverse effects on enzymatic steps. Overall, these optimisations will be able to be collated into a small, publishable paper to allow other groups to perform these experiments.

3.4 Bibliography

Granneman S, Kudla G, Petfalski E, Tollervey D. 2009. Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. PNAS 106:9613–9618. doi:10.1073/pnas.0901997106

Helwak A, Tollervey D. 2014. Mapping the miRNA interactome by cross-linking ligation and sequencing of hybrids (CLASH). Nature Protocols 9:711–728. doi:10.1038/nprot.2014.043

Iosub IA, Marchioretto M, Sy B, McKellar S, Nieken KJ, Nues RW van, Tree JJ, Viero G, Granneman S. 2018. Hfq CLASH uncovers sRNA-target interaction networks enhancing adaptation to nutrient availability. bioRxiv 481986. doi:10.1101/481986

Keene JD, Komisarow JM, Friedersdorf MB. 2006. RIP-Chip: the isolation and identification of mRNAs, microRNAs and protein components of ribonucleoprotein complexes from cell extracts. Nat Protoc 1:302–307. doi:10.1038/nprot.2006.47

Khemici V, Prados J, Linder P, Redder P. 2015. Decay-Initiating Endoribonucleolytic Cleavage by RNase Y Is Kept under Tight Control via Sequence Preference and Subcellular Localisation. PLOS Genetics 11:e1005577. doi:10.1371/journal.pgen.1005577

Kudla G, Granneman S, Hahn D, Beggs JD, Tollervey D. 2011. Cross-linking, ligation, and sequencing of hybrids reveals RNA–RNA interactions in yeast. PNAS 108:10010–10015. doi:10.1073/pnas.1017386108

Licatalosi DD, Mele A, Fak JJ, Ule J, Kayikci M, Chi SW, Clark TA, Schweitzer AC, Blume JE, Wang X, Darnell JC, Darnell RB. 2008. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature 456:464–469. doi:10.1038/nature07488

Nues R van, Schweikert G, Leau E de, Selega A, Langford A, Franklin R, Iosub I, Wadsworth P, Sanguinetti G, Granneman S. 2017. Kinetic CRAC uncovers a role for Nab3 in determining gene expression profiles during stress. Nature Communications 8:12. doi:10.1038/s41467-017-00025-5

Ule J, Jensen KB, Ruggiu M, Mele A, Ule A, Darnell RB. 2003. CLIP identifies Nova-regulated RNA networks in the brain. Science 302:1212–1215. doi:10.1126/science.1090095

Waters SA, McAteer SP, Kudla G, Pang I, Deshpande NP, Amos TG, Leong KW, Wilkins MR, Strugnell R, Gally DL, Tollervey D, Tree JJ. 2017. Small RNA interactome of pathogenic *E. coli* revealed through crosslinking of RNase E. The EMBO Journal 36:374–387. doi:10.15252/embj.201694639

4 Global analyses of RNA expression and sRNA activity

4.1 Introduction

As described in Chapter 3, RNase III and RNase Y were chosen as bait proteins for CLASH due to their already known roles in mediating RNA degradation and sRNA-mediated decay. These RBPs have also been implicated in the regulation of virulence (Introduction, Section 1.9), and thus were hypothesised to be active players in *S. aureus*' adaptation to the human bloodstream.

Using the optimised CLASH protocol as described in Chapter 3, I aimed to identify sRNAs that are involved in adaptation to the stresses *S. aureus* encounters when entering the human bloodstream. Additionally, as CLASH also gives single-read data ('CRAC data'), I also examined how these RNases themselves shape adaptation through their degradation of target RNAs.

CLASH was performed on *S. aureus* in a rich medium, TSB, and on cells which had been shifted to a medium called RPMI 1640. RPMI was originally designed to culture human leukocytes but has been shown in a study by Mäder et al, 2016, to induce a very similar response in *S. aureus* as human blood plasma, with particular overlap in iron-responsive genes. In order to move cells from one medium to another, a vacuum-based filtration system that facilitates shifting on minute timescales was utilised (Nues et al, 2017). As such, a model was created to mimic as well as possible the stresses induced when *S. aureus* moves from the environment into the human bloodstream (Figure 4.1).

In order to complement the CLASH data, I also performed RNAseq on cells in order to deduce RNA steady-state levels. Importantly, we also performed RNAseq on cell shifted back to their original medium, which allowed us to determine whether the filtration step itself induces significant stress. Additionally, CRAC on the polymerase (via the RpoC subunit) was performed because the data can be used as a proxy for transcription. Finally, with the help of a postdoctoral researcher in the lab (Liangcui Chu), I am also preparing protein samples to quantify changes in protein levels by mass-spectrometry. This would give me a readout what proteins are being

synthesized and degraded during the stress. Collectively, these analyses provided a detailed systems-level view of how RNA transcription and RNA degradation shape RNA stability and protein abundance.

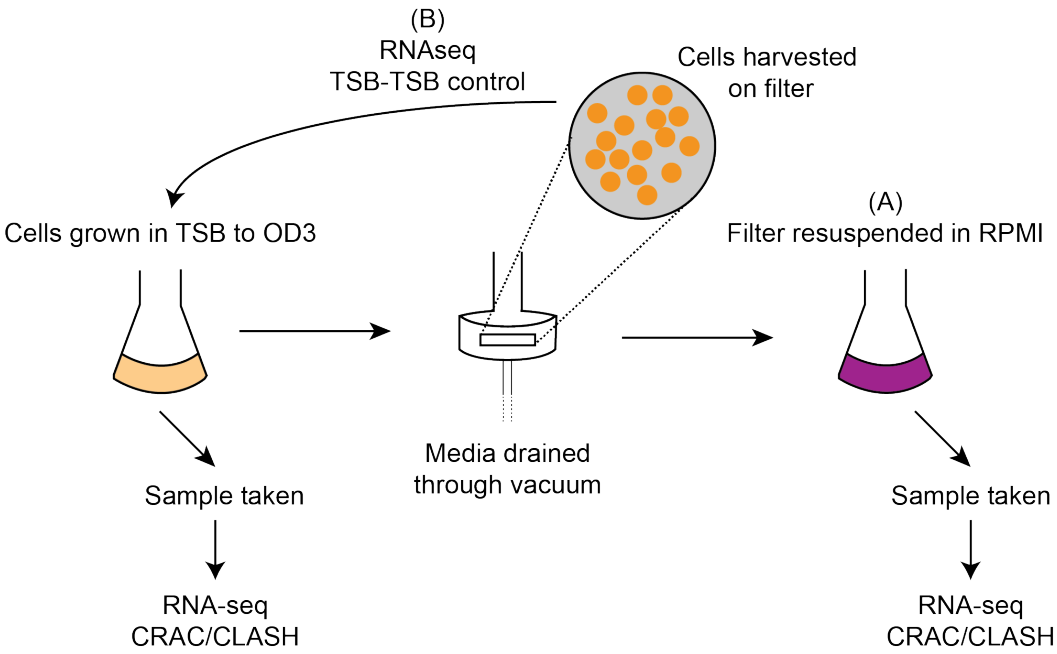

Fig. 4.1: Experimental set up for global analyses. Cells were grown in TSB until OD_{600}=3, and then a reference sample was taken. The cells were then harvested through vacuum filtration and resuspended in RPMI medium **A.** For CLASH, a sample was taken after 15 minutes and subsequently crosslinked. For RNAseq, samples were taken after 5, 10, 15 and 30 minutes. As a control for the RNAseq data, cells were shifted back into the original medium (**B.**) and samples were taken after 5, 10, 15 and 30 minutes.

The results presented in this chapter show firstly that there is a large degree of change in the expression of sRNAs after the imposition of stress induced by RPMI medium, implying that sRNAs are active players in the adaptation process. Through CLASH, targets for many of these sRNAs can be identified and several of these change in response to stress, implying that they are specific to the adaptation process. Finally, I identified a large number of sRNA-sRNA interactions, which suggests that sRNA activity is also regulated through other sRNAs. This chapter will overall seek to show that CLASH is a suitable method for identifying sRNA interactions in *S. aureus* and suggest novel targets for several *trans*-acting sRNAs.

4.2 sRNA expression undergoes significant remodelling in response to RPMI stress

RNAseq was performed on cells that had been shifted to RPMI and on cells shifted back to the original TSB medium in order to examine the changes in RNA steady

state levels in response to these stresses (Figure 4.1). In order to visualise the immediate responses, cell samples were taken after 5, 10, 15 and 30 minutes.

Firstly, I examined whether the shift to RPMI incurred cell death or caused a slowing of growth. I shifted an aliquot of cells to RPMI and then took OD_{600} measurements after 5, 10, 15 and 30 minutes, comparing to an unshifted control (Figure 4.2). This revealed that the shifted cells grew at the a comparable rate to the unshifted control. Thus, I concluded that shifting *S. aureus* from TSB to RPMI does not prevent the cells from growing normally.

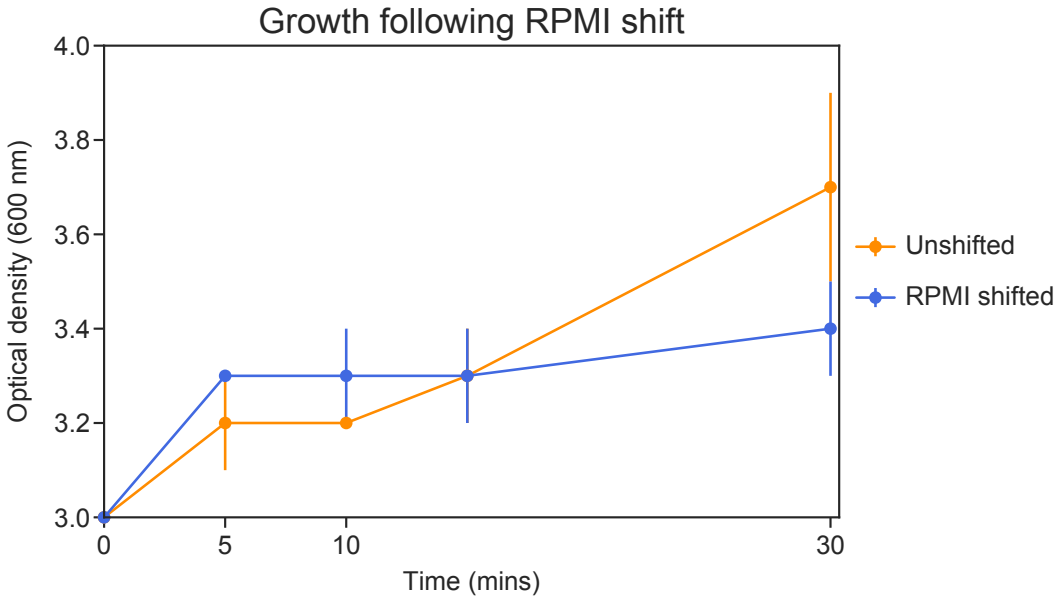

Fig. 4.2: Optical density measurements following the shift to RPMI as compared to an unshifted control.

Next, the expression of sRNAs after the shift was examined. A list of sRNAs was curated from the *Staphylococcal* Regulatory RNA Database (Sassi et al., 2015) and the expression of sRNAs that had statistically significant change after 30 minutes was plotted (DESeq2, Benjamini-Hochberg adjusted p-values < 0.05; Figure 4.3A, left panel). Figure 4.3A, right panel, shows the expression of these sRNAs in the control TSB-TSB shift. The sRNAs were also grouped into clusters of similar expression patterns (Figure 4.3B).

After 30 minutes of exposure, 55 sRNAs were observed to have statistically significantly changed expression after the shift to RPMI. Significance was calculated as compared to the 30 minute timepoint of the control TSB-TSB shift. Of these 55, 7 are known to be *bona fide, trans*-acting sRNAs, as defined by Liu et al, 2018; RNAIII, RsaD, RsaOG, SprA3, srn_3810, srn_1520, srn_2975 and srn_4520. Additionally, RsaOG and RsaH (which also meet the criteria for *trans*-acting sRNAs) were observed to have significantly changed expression after 15 minutes of exposure, but then

return to basal levels by 30 minutes. The specific expression profiles of these sRNAs is shown in the form of line plots in Figure 4.4.

The expression profiles show that sRNA responses are not simplistic up- or downregulation, but rather are complex, dynamic and change throughout the timecourse. Figure 4.3B shows the expression of the sRNAs in chosen clusters and their mean expression profile.

In cluster 1, it can be observed that two sRNAs exhibit strong and continuous downregulation for 15 minutes before starting to recover. Thus, these sRNAs may be unwanted during the immediate response to stress and so are repressed, but are then relieved of this repression in later adaptation stages and start to return to basal levels. Both of the sRNAs in this clusters are also slowly decreasing in expression in the TSB-TSB shift, but not to the same degree. Interestingly, the pro-virulence factor RNAIII is included in this cluster, which is significantly downregulated after the RPMI shift, as shown in Figure 4.4.

Cluster 2 shows sRNAs which are initially repressed during the first 5 minutes but are then actively upregulated to levels above their initial expression. Thus, these sRNAs may be unwanted in the immediate face of stress (and so are destroyed) but may play a role in the later stages of adaptation and so actively upregulated afterwards. RsaD is included in this cluster, which is already known to respond to oxidative stress and contains the distinctive UCCC motif of the Rsa class (Geissmann et al., 2009). The expression of RsaD is shown in Figure 4.4, where it is observed that is strongly upregulated almost 10-fold after 30 minutes in RPMI, as compared to its flat expression in the TSB control.

Clusters 3 and 6 show sRNAs which are upregulated during the first 5 minutes of adaptation but then downregulated afterwards until 15 minutes when they appear to be relieved of their repression. Cluster 4 exhibits a similar overall pattern but acts slower. Thus, these three clusters of sRNAs may be important for the initial response to stress but then must be removed afterwards. The sRNAs srn_3810 and srn_1520 belong to cluster 3, but nothing is known about their functionality. SprA3, RsaOG and RsaH belong to cluster 6 and these sRNAs are better understood. SprA3 is encoded in a pathogenicity island and is a target of RNase III (Lioliou et al., 2012; Pichon and Felden, 2005) and also exhibits strong sequence similarity to the anti-toxin sRNA SprA. RsaOG, also known as RsaI, is involved in an intricate network of multiple sRNAs which regulate sugar metabolism and uptake, and has also been predicted to play a role in the metabolic balancing of virulence and dormancy (Bronesky et al., 2019). RsaH has been reported to accumulate in either pre-stationary phase or stationary phase, but its function has not been found (Bohn et al., 2010; Geissmann et al., 2009). Both RsaOG and RsaH have a continuous decrease in expression in

the TSB-TSB shift, thus their strong upregulation during the first 15 minutes of RPMI exposure appears to be specific to the adaptation process. The individual profiles of these sRNAs, as shown in Figure 4.4, contains a common 'triangular' shape where each sRNA is upregulated for a short period of time before being repressed afterwards.

Finally, cluster 5 contains sRNAs which are strongly upregulated after stress, and this continues throughout the whole time course. Thus, these sRNAs may play an active role in shaping adaptation. None of the sRNAs in this cluster meet the criteria for *trans*-acting sRNAs as defined by Liu et al., 2018, and nothing is known of their function.

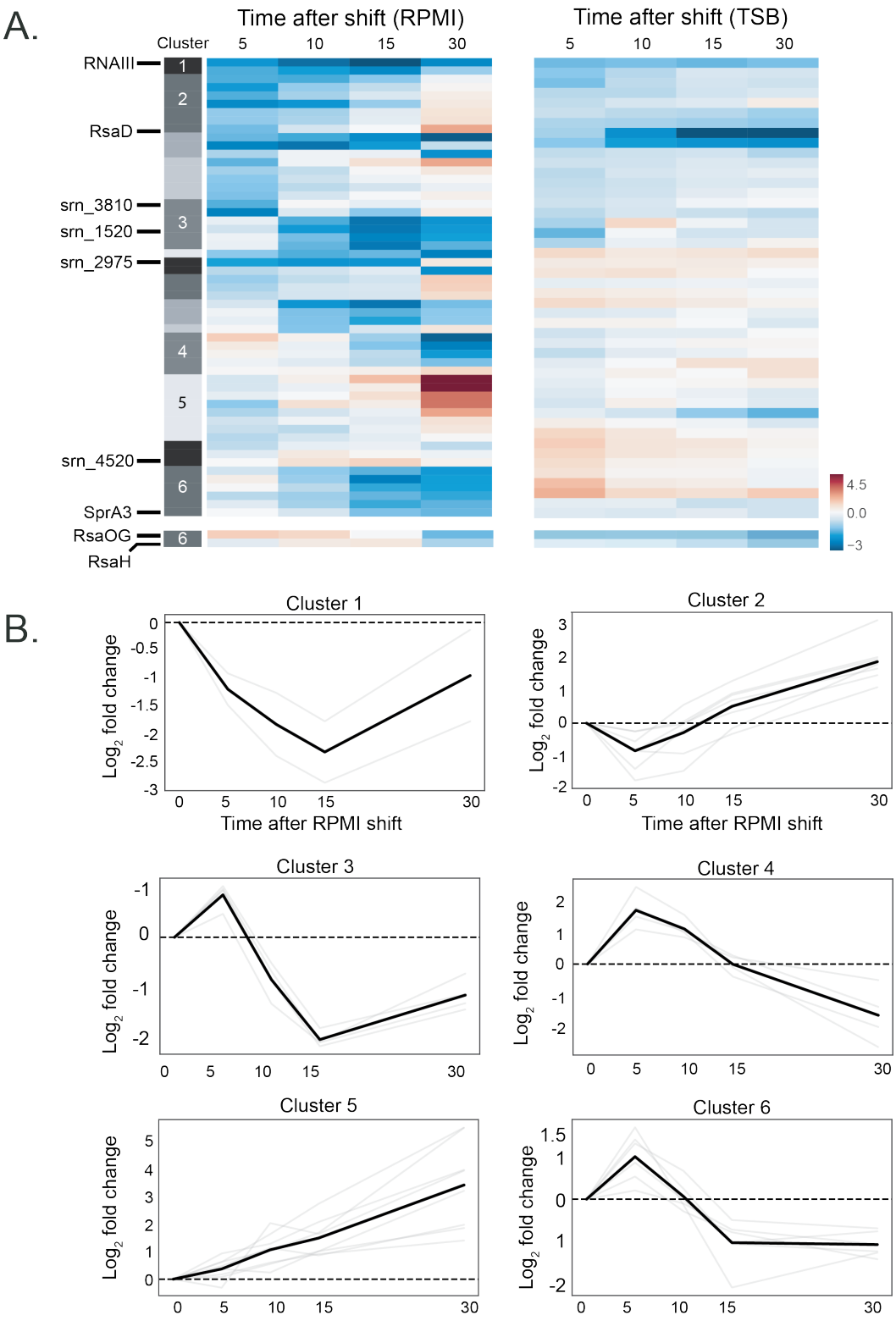

Fig. 4.3: sRNA expression in response to RPMI stress. **A.** Heatmap showing sRNAs with statistically significant changed expression after 30 minutes of RPMI exposure, as compared to the control. sRNAs have been clustered into patterns of similar behaviour, as calculated by STEM (Ernst and Bar-Joseph, 2006). *Bone fide, trans*-acting sRNAs are labelled (Liu et al., 2018). The expression of RsaOG and RsaH has been also added, both of which exhibit statistically significantly changed expression after 15 minutes but then recover by 30 minutes. **B.** Line graphs showing chosen expression clusters.

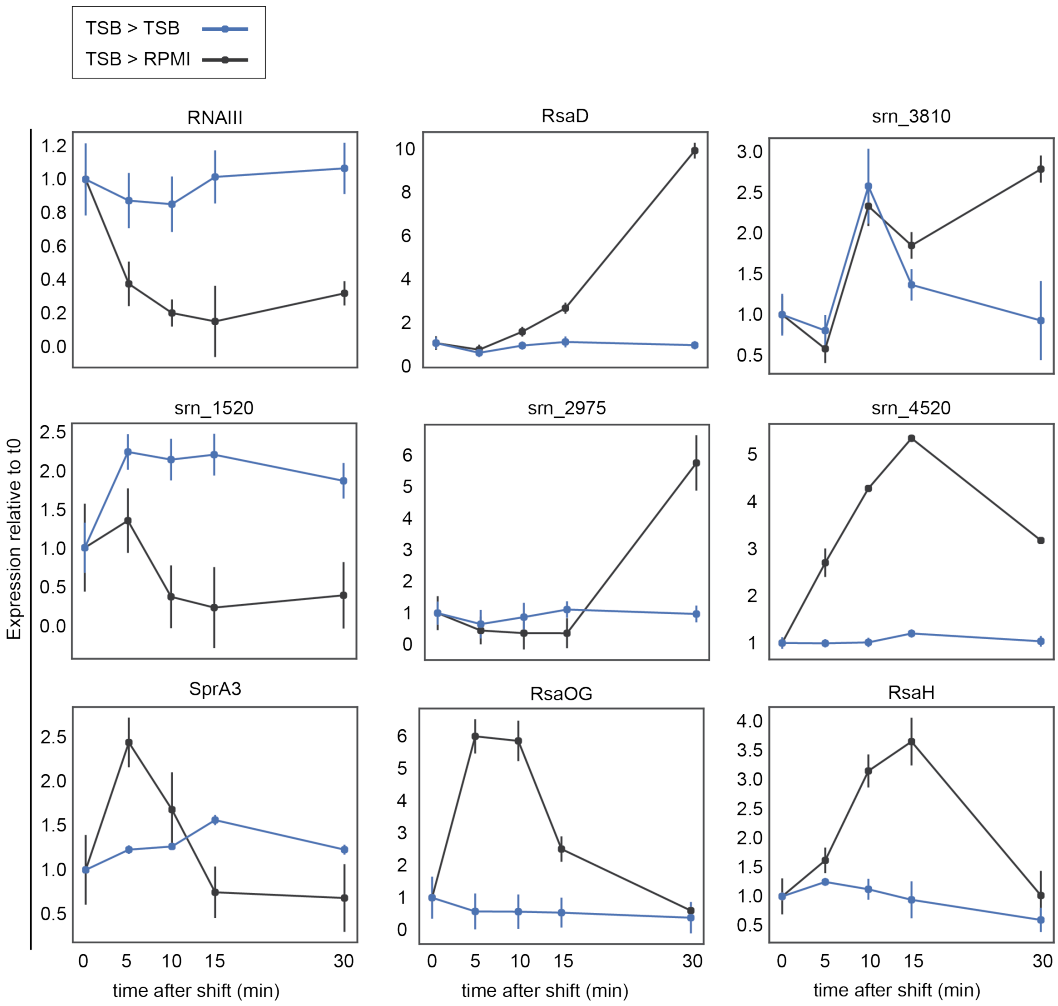

Fig. 4.4: Expression of *bona fide, trans*-acting sRNAs that exhibit significantly changed expression after either 15 or 30 minutes of RPMI exposure.

After examining the expression of sRNAs in response to RPMI, the expression of the entire transcriptome was examined. As seen in Figure 4.5A, the transcriptome undergoes significant remodelling in response to RPMI medium. Overall, 735 genes exhibit significantly changed expression after 30 minutes of exposure. In a similar manner to the sRNA responses, many of the expression profiles are dynamic and changing. This is in opposition to the expression patterns seen in the TSB-TSB control shift, where genes are simply either up- or down-regulated continuously. This leads to the idea that the physical act of shifting the cells from one medium to another incurs minimal non-specific stresses and rather the expression changes are simply due to the effect of time. In support of this, genes which are known to decrease in expression during stationary phase versus exponential, such as *spa, fruB, ald, ilvA*, and *adhE*, also decrease over the time course. Conversely, genes which are known to be upregulated in stationary phase, such as *ssaA* and *SAUSA300_0229*, also increase over the TSB shift time course (Weiss et al., 2016).

Examining the different clusters of expression in RPMI shows similar patterns to that observed for the previously discussed sRNAs (Figure 4.5B). Cluster 1 contains genes which are initially downregulated for the first five minutes but are then upregulated afterwards. STRING analyses reveals that 'metal ion binding' is in an enriched term for this cluster, and several genes involved in iron transport can be identified such as *SAUSA300_1028, SAUSA300_1029, SAUSA30_2134, SAUSA300_1978*, and *SAUSA300_1005*. Additionally, lipoteichoic acid synthesis, a central component of the cell wall, was also an enriched STRING term.

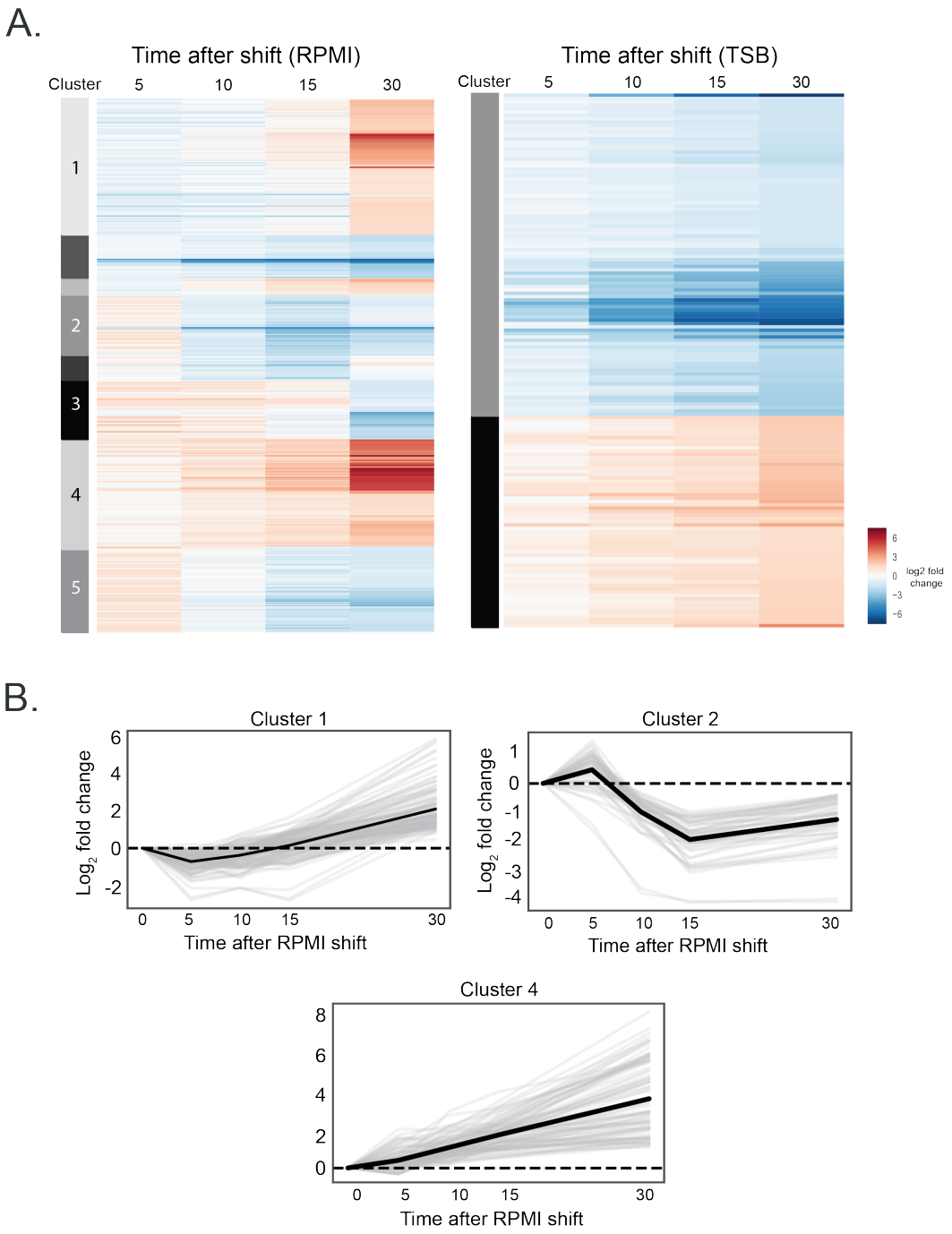

Fig. 4.5: Transcriptome-wide changes in response to RPMI. **A.** Left: heatmap showing genes with statistically significant changed expression after 30 minutes of RPMI exposure, as compared to the control. Genes have been clustered into patterns of similar behaviour as calculated by STEM. Right: heatmap showing genes with significantly changed expression after 30 minute of TSB shift **B.** Line graphs showing chosen expression clusters.

Clusters 2, 3 and 5 display a common pattern of behaviour, with varying amplitudes and timing; for brevity, only cluster 2 is shown in Figure 4.5B. Overall, these clusters increase initially, with clusters 3 and 6 reaching their maximum expression by five minutes and cluster 4 reaching this at 15 minutes. After reaching the maximum, their expression is decreased and drops to levels below the basal. STRING analysis reveals that this group is enriched for genes involved in metabolic processes, and indeed 'glycolysis' is a statistically significant enriched Uniprot term due to genes such as *eno, gpmL* and *pgk*. Additionally, 'virulence' is also a statistically significant enriched annotation term for these clusters due to genes such as *clfA, sarT, sspA, eno* and *crtN*.

The fifth cluster from the RPMI data includes genes which are significantly upregulated throughout the timecourse. This differs from the TSB control upregulated cluster as the RPMI cluster reaches a far higher maximum expression. In the TSB data, only 7 genes are upregulated at least 4-fold, with the maximum being a ~9-fold change. Regarding the RPMI data, 173 genes are upregulated at least 4-fold, and over 40 genes have changes of at least 32-fold. This cluster 5 also has enrichment for metabolic genes, and in particular (branched-chain) amino acid biosynthesis, carboxylic acid biosynthesis and diaminopimelate biosynthesis.

Manual examination of the transcripts with the highest degree of up- and down-regulation gives insight into the types of metabolic changes the cell must undergo in order to survive. Figure 4.6 shows the results of a DESeq analysis comparing the expression of genes in TSB versus those in RPMI after 30 minutes of exposure (Love et al., 2014). This reveals that there is a large cluster of iron-uptake related transcripts with increased expression. Several operons involved in amino acid uptake and biosynthesis are also increased, such as the *opp, leu*, and *ilv* operons. There is also remodelling of sugar metabolism, with the fructose transporters *fruA* and *fruB* being strongly downregulated. Additionally, *SAUSA300_0755*, a glycolytic operon regulator; *malR*, a maltose uptake repressor; and *SAUSA300_0194*, a sucrose transporter, are also downregulated.

4.3 CLASH identifies *bona fide* sRNA-target interactions in *S. aureus*

Given that there is significant change in the expression of sRNAs, I hypothesised that some of these could play a role in the adaptation process. CLASH was used to identify the targets of many of these sRNAs in both TSB and after 15 minutes of RPMI shift, and thus allowing one to begin to uncover the role of selected sRNAs in shaping adaptation.

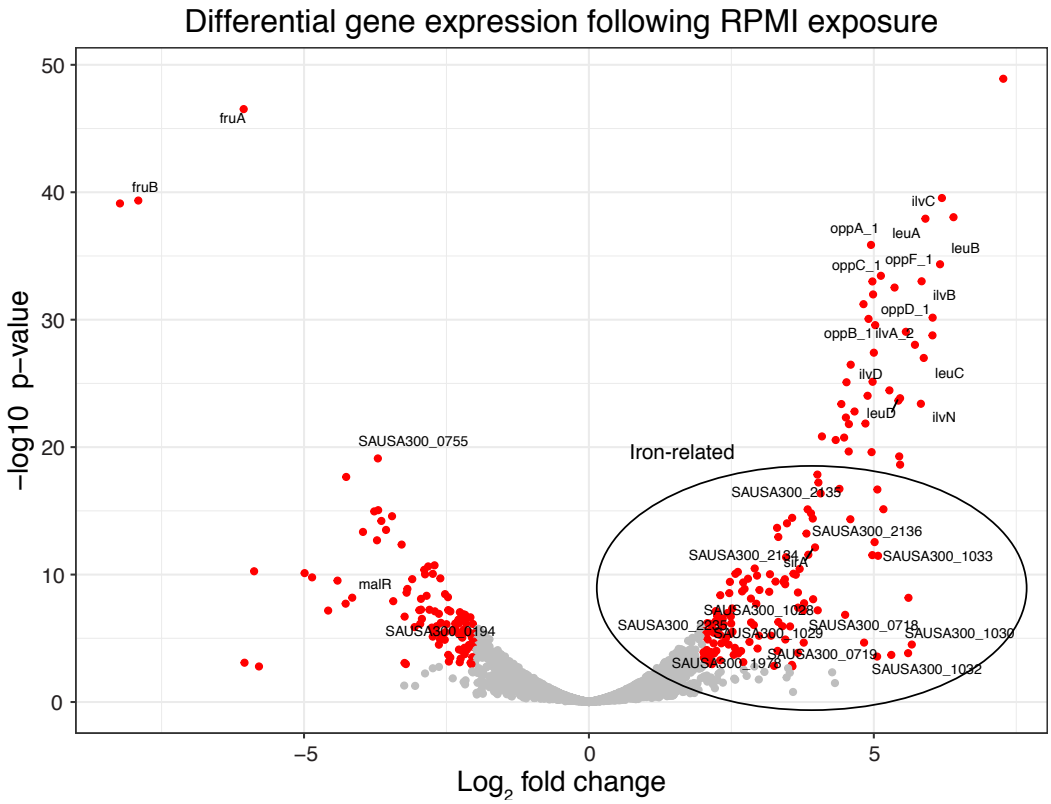

Fig. 4.6: Differential gene expression following RPMI shift. Expression 30 minutes after RPMI shift was compared to the reference, unshifted sample.

As *S. aureus* is known to show considerable genetic variability between strains, CLASH was performed on an Australian VISA strain (vancomycin-intermediate sensitive; JKD6008) and an American VSSA strain with enhanced virulence (vancomycin-sensitive; USA300 LAC). Overall, JKD6008 is viewed as an *S. aureus* strain with a wide range of antibiotic restistances, but not only that is particularly virulent. It is ST239 and contains a type 3 SCC*mec* plasmid. Additionally, JKD6008 contains a ϕSPβ-like prophage that has been hypothesised to potentially confer aminoglycoside resistance (Holden et al., 2010).

Conversely, USA300 is a strain that is considered to be extremely aggressive and a major driver of community-acquired MRSA. For example, in some parts of the United States, USA300 was isolated from almost every soft skin and tissue infection (Moran et al., 2006, Talan et al., 2011). USA300 is ST8 and contains a type 4 SCC*mec* cassette. Additionally, USA300 contains the *msr(A)* erythromycin resistance genes and a prophage-encoded Panton-Valentine leukocidin (McDougal et al., 2003). In an attempt to explain the increased pathogenicity of USA300, its genome was compared to 10 other *S. aureus* strain. This revealed the presence of a novel pathogenicity island containing two enterotoxins and the arginine catabolic mobile genetic element (Diep et al., 2006). Additionally, USA300 has been observed to have an increased capacity to express toxins encoded on the core genome, such as Hla and phenol-

soluble modulins, however the genetic basis of this has not been precisely explained (Li et al., 2009).

Interestingly, it has been proposed that ST239 strains, such as JKD6008, is a hybrid between ST8 and ST30, created by a large chromosomal replacement (Robinson and Enright, 2004).

For JKD6008, CLASH was performed on both RNase III and RNase Y, while only RNase III was used in USA300 as this proved to be a much better bait in enriching for sRNA-target interactions.

An average of 1% of the total reads from RNase III CLASH represented hybrids, although significant variability was observed between samples (Figure 4.7A). This number is in line with previous studies (Iosub et al., 2018; Waters et al., 2017). However, RNase Y did not yield as many hybrids, averaging only ~0.3%. This was surprisingly lower than the parental control, which yielded ~0.4%. However, it is important to note that CLASH performed on the parental strains produced a very small number of reads (around 500,000) thus any non-specific hybrids account for a relatively large percentage. Conversely, RNase III and RNase Y produced a very large number of total reads; around 96 million and 18 million respectively. Thus, the percentages represent a large number of interactions. Note that in these analyses, tRNA – rRNA interactions were removed as these are highly abundant RNAs that are the most likely to form non-specific interactions. Moreover, predicted rRNA – rRNA and tRNA-tRNA interactions were also not considered as their rRNA and tRNA genes are present in many copies and the sequence identity is very similar, making it difficult to determine if these were indeed intermolecular interactions.

Examining the composition of the hybrids reveals RNase III CLASH in particular was able to enrich for sRNA – mRNA hybrids. CLASH on the untagged parental strains allowed me to calculate the background binding of cellular RNAs to the binding resin. This revealed 23 unique hybrids involving sRNAs in TSB and 12 in RPMI, representing ~5.5% and ~4% of the total interactions found (Figure 4.7B). However, CLASH on RNase III in both JKD6008 and USA300 increased this number dramatically. In USA300, 1800 hybrids containing sRNAs were identified in TSB and 26627 in RPMI. For JKD6008, 7866 were identified in TSB and 6390 in RPMI (Figure 4.7B).

CLASH on RNase Y was only performed in JKD6008 during the initial stages of this thesis. This is because it was far less effective at identifying RNA interactions in general and also those involving sRNAs in particular. I identified just under 600 sRNA-containing hybrids with RNase Y in TSB and RPMI (Figure 4.7B and C). Thus, RNase III was prioritised in future experiments.

However, it is important to note that a single interaction can be represented by many hybrids. Collapsing the hybrid data into single interactions revealed the number of unique interactions obtained (Figure 4.7D). RNase III CLASH yielded 208 unique interactions between sRNAs and mRNAs in TSB, with 78 occurring within the protein coding sequence (CDS) of the mRNA and 130 within the 5' UTR. For RPMI, 229 interactions between sRNAs and mRNAs were found, with 136 within the CDS and 93 within the 5' UTR. I also identified 51 interactions between two sRNAs in TSB and 180 in RPMI. Interestingly, there is a huge number of interactions between intergenic regions and sRNAs, and between intergenic regions and 5' UTRs in the USA300 RPMI dataset (Figure 4.7D)

Examining the RNase Y CLASH data in JKD6008 reveals that it uncovers a similar number of unique interactions as RNase III CLASH (Figure 4.7D). Although these have been filtered for statistical significance, the vast majority of these interactions are supported by only a single hybrid. Exemplifying this, only 17 interactions between sRNAs and mRNAs were supported by more than 1 hybrid in the RNase Y data. This reemphasizes that RNase Y is not a great bait for identifying sRNA-target interactions and suggests it may not play a major role in this process.

The statistical analysis pipeline used for this analysis takes into account the number of single reads for an RNA and compares this to the number of hybrids obtained (Waters et al., 2017). Such a methodology is weak at filtering non-specific hybrids which involve RNAs with a low number of reads. Although the lab has found that hybrids represented by only a single read can represent genuine interactions (and this thesis has identified known interactions in *S. aureus* represented by only single hybrid), I treat these with greater suspicion. This is because proximity-dependent ligation experiments are subject to false-positives as a result of spurious ligation events, mapping artefacts and errors incurred during reverse transcription and PCR (Ramani et al., 2015). As such, RNase III was focused on for future experiments due to it being able to extract a greater number of hybrids, and as such increase the confidence in the data.

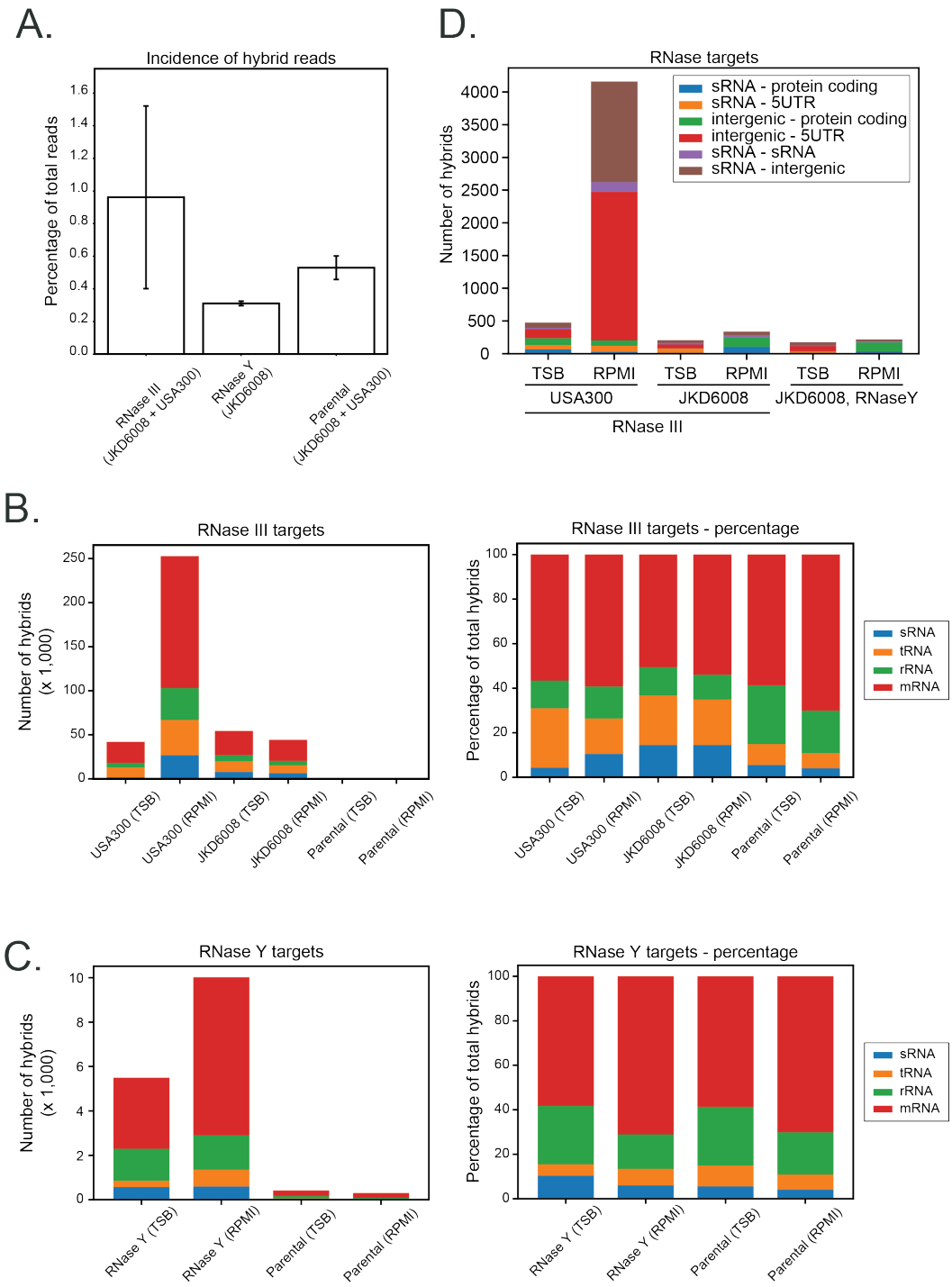

Fig. 4.7: CLASH hybrid incidence and composition. **A.** The incidence of hybrid reads in CLASH for RNase III, RNase Y and the parental strains. Both JKD6008 and USA300 data have been merged for the RNase III and the parental strain data, while RNase Y CLASH was only carried out in JKD6008. **B.** and **C.** Hybrid composition for RNase III and RNase Y in USA300 and JKD6008, alongside the respective parental strain. Left hand panel shows the total number of hybrids involving each RNA class and this data is presented as a percentage in the right hand panel. Interactions between rRNAs and tRNAs were removed from the data and the parental strain data (USA300 and JKD6008) were merged. Note that for the purpose of this analysis, reads mapping to 5' UTR, CDS and 3' UTR were merged into the umbrella term of 'mRNA'. **D.** Total number of each interaction type for RNase III and RNase Y in TSB and RPMI. .

To further these analyses, I examined between which RNA classes these interactions were occurring (Figure 4.8). Regarding RNase III in USA300 and JKD6008, interactions between sRNAs and both the 5' UTR and protein coding regions of mRNAs represent significant classes, representing from ~6% to 11% of the total hybrid reads. Interactions between sRNAs and other sRNAs also consistently represented around 2% of the hybrid reads. Interestingly, interactions between intergenic regions and 5' UTRs, coding sequences and sRNAs represent significant interaction classes; these intergenic regions may represent a pool of potentially undiscovered sRNAs. Overall, the hybrid incidence is similar when comparing the data from TSB to RPMI in USA300, although I note that interactions between 5' UTRs are only found in RPMI and not significantly in TSB. Some more pronounced differences are observed in JKD6008 where interactions between protein coding sequences and other protein coding sequences are the most abundant transcript class in RPMI but are not present significantly in TSB. This is difficult to explain and these interactions could represent poorly annotated genes. Additionally, interactions between 5' UTRs and other 5' UTRs are only found in TSB and not in RPMI.

The data for RNase Y is less rich. In TSB, "other combinations" make up the most abundant interaction type; these represent lowly abundant interactions, but also interactions in which one composing RNA maps to multiple locations e.g. a 5' UTR and then into the coding sequence. Of note, interactions between sRNAs and either 5 UTRs or protein coding sequences represent roughly 6% of the interactome. However, interactions between intergenic regions represent significant proportions of the interactome in both TSB and RPMI, as are interactions between protein coding regions in RPMI.

4.4 CLASH identifies canonical sRNA – mRNA interactions

In order to gain confidence in the CLASH data, I examined which regions of mRNAs were being targeted by sRNAs. As sRNAs canonically bind to the 5' UTR of mRNAs, it was hypothesised that reads for mRNAs found in sRNA – mRNA hybrids should be enriched in this region. Two methods were employed to examine this.

Firstly, the read profile around the starter codon was examined. Here, reads for mRNAs found in sRNA – mRNA hybrids were mapped across the AUG (Figure 4.9A). To generate this, the data was collapsed so that identical hybrids are only counted once in order to prevent a single, highly abundant interaction dominating the data. This revealed that for both USA300 and JKD6008, mRNA reads mapped preferentially to the 5' UTR and start codon and then immediately decreased downstream.

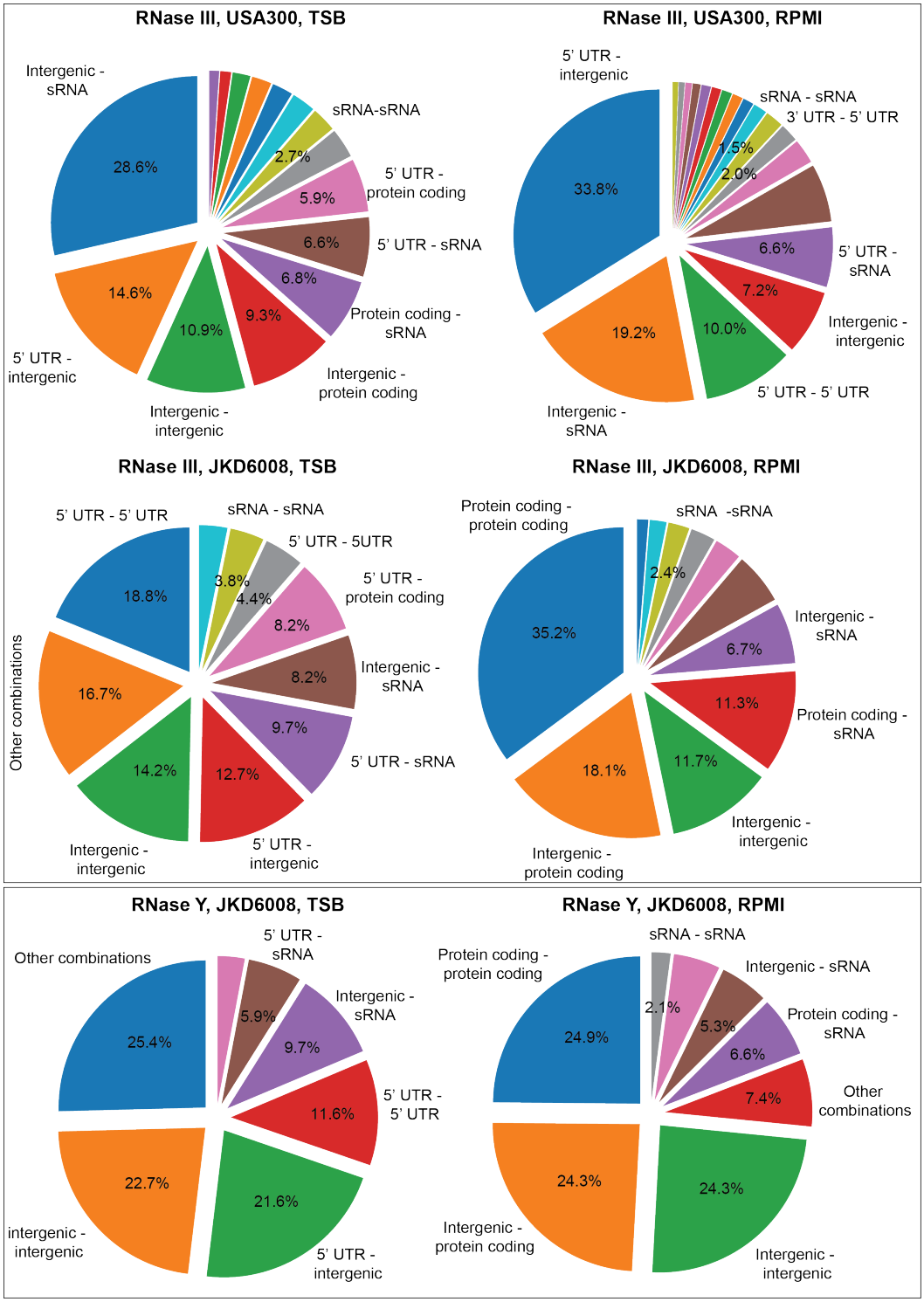

Fig. 4.8: Intermolecular RNA transcript combinations as detected by RNase III and RNase Y CLASH. Interactions were first filtered for statistical significance and then interactions containing rRNAs and tRNAs were removed. Interaction types of interest are annotated, and 'Other combinations' represents merged lowly abundant combinations. Each RNA interaction was then plotted as a percentage of the total. Note, a single interaction can be represented by many individual hybrids.

Secondly, mRNAs involved in sRNA – mRNA hybrids were separated into 100 equally-sized fragments ('bins') and then the reads were mapped relative to these (Figure 4.9B). This data was again processed such that each hybrid was only counted once. This revealed that for both USA300 and JKD6008, reads were again concentrated at the 5' end of the mRNAs. However, for both for JKD6008 and USA300, there is a sharp spike in the middle of the coding sequence, particularly true for the USA300 data.

The mRNAs which mapped to this 'spike' region were examined and this revealed that many of these were components of operons. Of the 18 genes within this region, 9 belonged to operons. A tentative suggestion is that sRNAs could be targeting a central portion of the gene in order to recruit RNase III to mediate processing of the operon. Additionally, *rot* was found within this spike region and this is known to be targeted by RNAIII within its coding sequence (Boisset et al., 2007). Thus, these interactions may also represent genuine targeting in the middle of the transcript.

Next I examined if CLASH recovered interactions containing validated sRNAs. As discussed in Section 1.8, a previous study by the Bouloc group bioinformatically filtered the current list of sRNAs to try ascertain which of these are *trans*-acting, *bona fide* sRNAs. For this, they only considered sRNAs which contain their own promoter and terminator and do not overlap any other gene or antisense transcription (Liu et al., 2018). Although this strict definition is imperfect as it excludes UTR-derived sRNAs, it serves as a useful tool for examining if our data contains confirmed sRNAs. As shown in Figure 4.9C, CLASH recovered 16 of the 39 '*bona fide*' sRNAs. This is a positive result as I did not expect to recover all the sRNAs in *S. aureus* as many are only expressed at specific growth stages and under specific conditions.

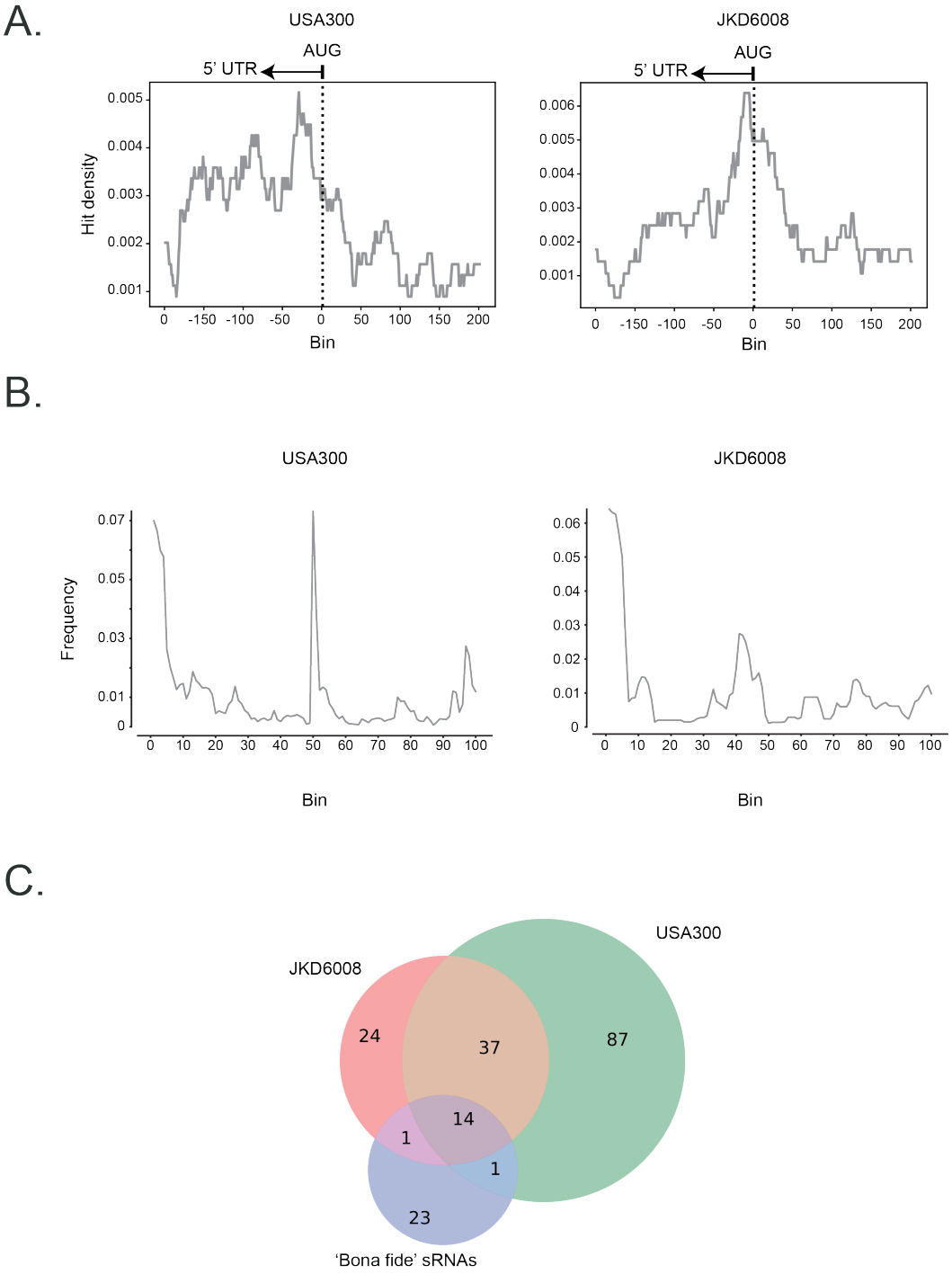

Fig. 4.9: RNase III CLASH identifies canonical sRNA – mRNA interactions. **A.** The reads mapping to mRNAs involved in sRNA – mRNA interactions were plotted relative to the starter codon. **B.** The same reads were plotted across the mRNA genomic sequence, split into 100 equally sized bins. **C.** All of the sRNAs involved in sRNA-mRNA interactions in both USA300 and JKD6008 were extracted and the overlap between them was plotted as a Venn diagram. Additionally, in order to see how many of these sRNAs are highly verified, the curated list of '*bona fide*' sRNAs as described by Liu et al., 2018, was integrated into this Venn diagram.

4.5 CLASH identifies known sRNA interactions

I next examined if the CLASH data contained known sRNA interactions. As discussed in the Introduction, several sRNA – mRNA and sRNA – sRNA interactions have been validated in *S. aureus* and uncovering some of these would give credence to any novel interactions proposed.

The most abundant sRNA – mRNA interaction detected was between *sprA1* and SprA1$_{AS}$ with over 4,500 unique hybrids found across the pooled datasets. This interaction was previously found by the Felden group (Sayed et al., 2012). Although the structure of the sprA1 - SprA1$_{AS}$ duplex has not been experimentally validated, it was computationally predicted to involve the SD sequence and start codon. Figure 4.10A shows the interaction uncovered through CLASH, visualised *in silico* through RNAcofold (Lorenz et al., 2011). The region highlighted in blue on Figure 4.10A denotes the regions predicted to be involved in base pairing by Sayed et al, 2012 and excellent overlap is seen with the CLASH data.

The sRNA SprX is known to regulate the translation of *spoVG*, a protein involved in antibiotic resistance, the production of virulence factors and capsule regulation (Eyraud et al., 2014). The structure of this interaction has been found through *in vitro* structure probing, and again the precise interaction sites were identified through CLASH with over 200 unique hybrids identified (Figure 4.10B). Although the start codon of *spoVG*, which is known to be targeted by SprX, is not shown in duplex, this is due to inaccurate prediction by RNAcofold as opposed to inaccuracies in CLASH.

Another known interaction identified through CLASH is between SprF and *sprG* (Pinel-Marie et al., 2014). No structure probing or *in silico* structural predictions were made in the original paper, but it was found that SprF represses the translation of *sprG*. The CLASH data uncovered both the start codon of *sprG* and a portion of 5' UTR which fits the canonical mode of translational repression (Figure 4.10C).

Many interactions between RsaA and RsaE with already known targets were also found. These include RsaA's interaction with *mgrA, SsaA2_3* and *JKD6008_01954*, and RsaE's interaction with RsaOG, *purK, opp3-B, sucD* and *purH* (Figure 4.11A). As RsaA, RsaE and RNAIII will be discussed in dedicated chapters, these interactions are detailed later. Some of these interactions were only uncovered with a small number of respective hybrids. Thus, this demonstrates that interactions found with CLASH are not only valid when in abundance, but that even small numbers of hybrids can represent true, *bona fide* interactions.

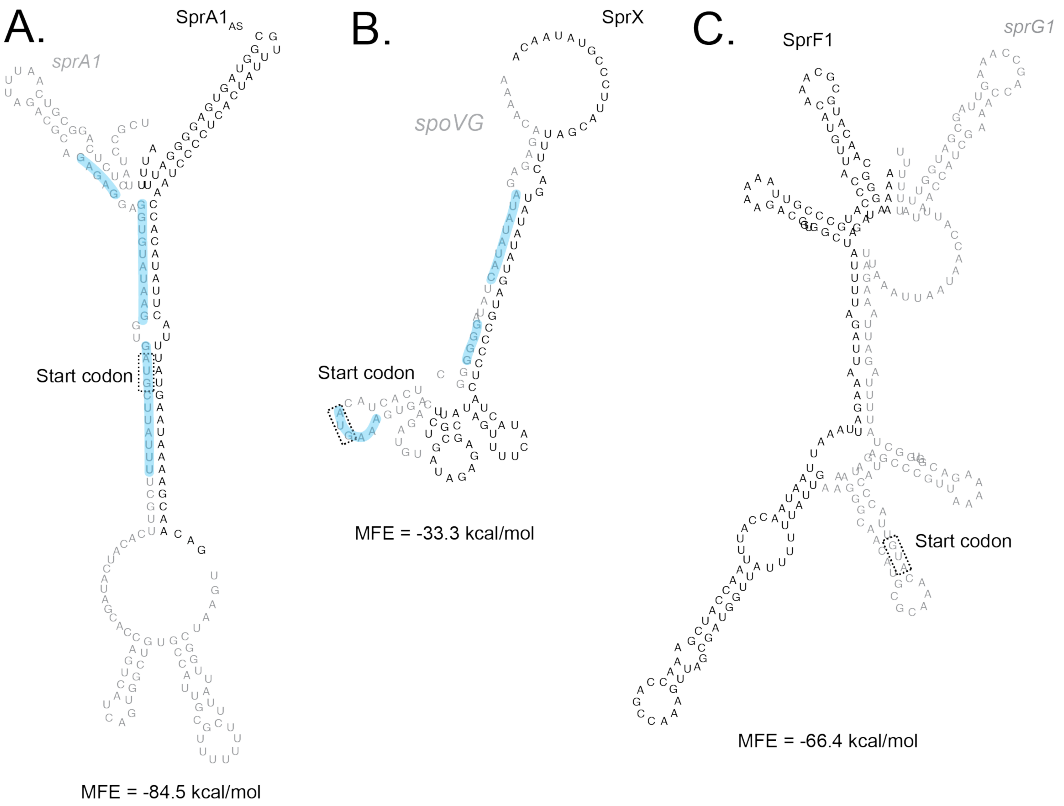

Fig. 4.10: Examples of known targets identified through CLASH; SprX-*spoVG* (**A.**), SprF1-*sprG1* (**B.**) and *sprA1*-SprA1$_{AS}$ (**C.**). Structures were generated by extracting the CLASH read sequences as input for RNAcofold, which was used to predict the interaction structure. For SprA1$_{AS}$ - *sprA1*, computational predictions of the interaction were made by the authors and are highlighted in blue (Sayed et al., 2012). For SprX-*spoVG*, the structure of this interaction has been validated previously through *in vitro* structure probing, and the regions of *spoVG* known to be targeted by SprX are highlighted in blue (Eyraud et al., 2014).

As will be detailed in subsequent chapters, CLASH also identified a large number of novel interactions. Figure 4.11A shows the individual interactomes of RsaA, RsaE and RNAIII as uncovered from CLASH experiments performed in this study. The interactions composing these interactomes have been filtered for statistical significance and any interactions with rRNAs and tRNAs have also been removed. This latter point is due to the fact that I am unsure if interactions between rRNAs and tRNAs are genuine or artefacts due to these transcripts' abundance. However, it can be seen that CLASH is able to identify potentially novel mRNA and sRNA targets of all of these well-known sRNAs. Future studies to be discussed in this thesis will focus on RsaA's interaction with RNAIII and *qacAB*, and RsaE's interaction with RsaOG, *SAA6008_09901* and *SAA6008_09902*. Finally, Figure 4.11B shows an example of a whole cell interactome as obtained from CLASH; in this case, from RNase III after RPMI exposure in USA300. This Figure illustrates the power of CLASH to identify complex networks of interactions, alongside smaller connections.

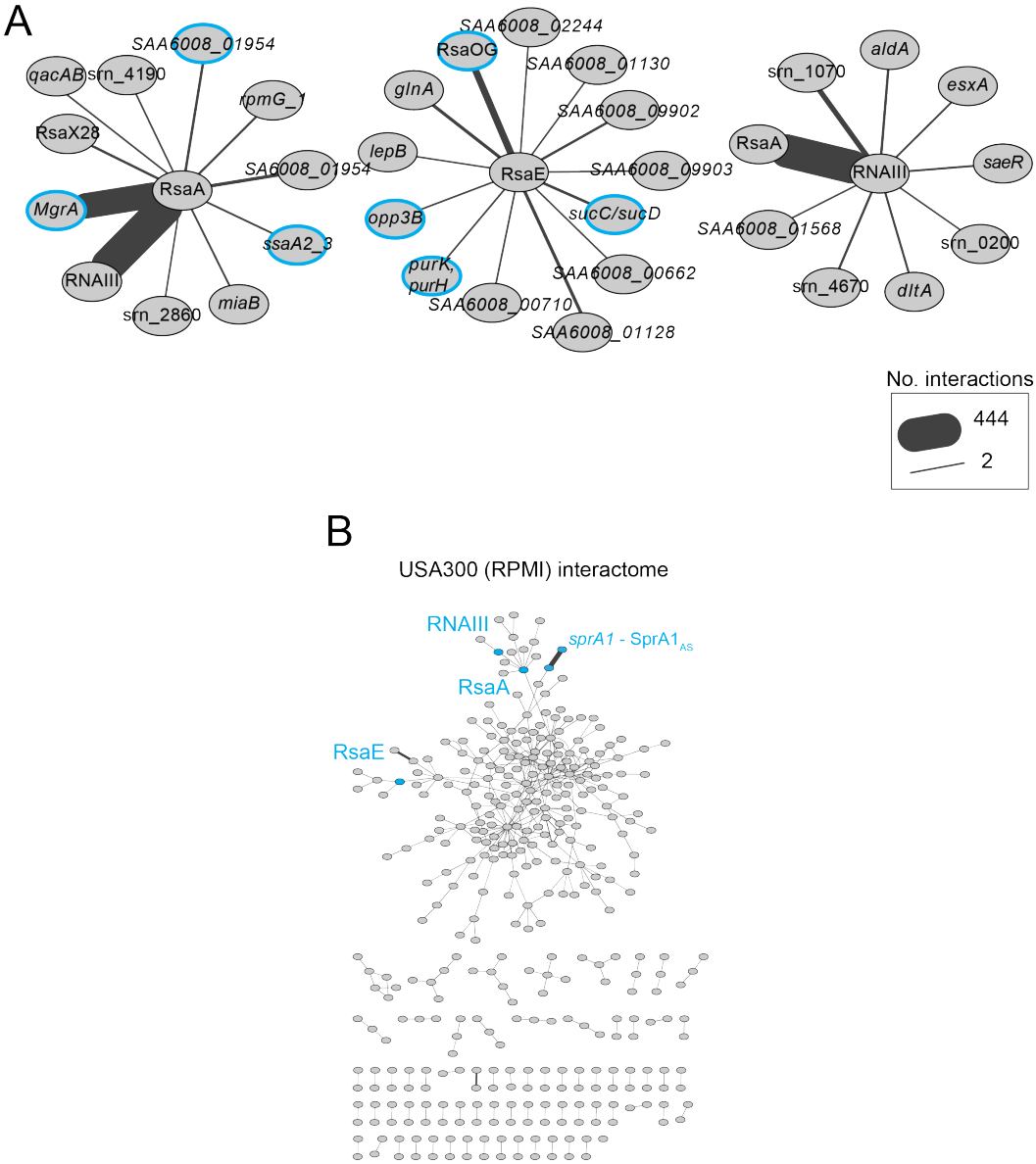

Fig. 4.11: Examples of interactomes obtained from CLASH. **A.** Individual interactomes of RsaA, RsaE and RNAIII from merging all the RNase III and RNase Y CLASH datasets. Interactions that have already been reported are highlighted in blue. These data have been filtered for statistical significance and any interactions with rRNAs and tRNAs have been removed. The edge weight indicates the number of unique interactions obtained. **B.** Overall interactome obtained from CLASH on RNase III in USA300. The most abundant interaction, that between *sprA1* and SprA1$_{AS}$ is highlighted. Additionally, the locations of RsaA, RNAIII and RsaE are shown. These data have been filtered for statistical significance and any interactions with rRNAs and tRNAs have been removed. The edge weight indicates the number of unique interactions obtained.

From the data presented here and in Section 4.5, I determine that CLASH is able to recover sRNA – mRNA interactions that agree with the current paradigm of sRNA-mediated regulation. Additionally, CLASH is able to identify interactions involving validated sRNAs and many known interactions are recovered - these

positive validations give credence to any novel interactions proposed by CLASH. Thus, I conclude that CLASH is a suitable method for identifying both validated and novel sRNA – target interactions.

4.6 CRAC analysis reveals target classes of RNase III and RNase Y

Although I am primarily interested in RNA – RNA interactions, the CLASH methodology also provides data on the single RNAs bound to the precipitated protein. This is relevant to this study as CLASH was performed on two RNases and as such the target RNAs may be transcripts targeted for degradation. Thus, examination of these single reads may reveal target classes that are marked for destruction after the shift to RPMI.

In both JKD6008 and USA300, I observed that RNase III increasingly targets protein coding sequences after the shift to RPMI (Figure 4.12A, upper panel) – in USA300, from ∼4% in TSB to ∼11% in RPMI and in JKD6008, from ∼10% to ∼13%. Additionally, 5' and 3' UTRs are also increasingly targeted at a similar rate, and small decreases are observed for rRNAs, tRNAs and intergenic regions. The shift does not seem to impact the targeting sRNAs which consistently represent ∼6% of the total targets. Thus, RNase III may increasingly target mRNAs for degradation after the shift to RPMI, as represented by increasing targeting of UTRs and protein coding sequences.

I next examined if any common sequence motifs could be found within RNase III's target genes (Table 4.1). This was performed by filtering the CRAC data for false-positives and then generating peaks of RNase III binding. Afterwards, these peaks were analysed for common motifs between 4 and 8 nucleotides long. These motifs are scored by 'Z-score', which represents how many standard deviations the motif exhibits from the mean binding. Looking at the top 5 motifs found in both TSB and RPMI, it is clear that RNase III has a preference for purine-rich sequence such as GAAG, AGA and GGA.

Tab. 4.1: RNase III target motifs

Condition	Motif	Z-score
TSB	AGAG	6.81
	GGAA	5.90
	GAAG	5.63
	GAGC	5.58
	GAAGA	5.50
RPMI	GAAG	7.59
	AGAG	7.57
	GAAGA	7.39
	TGAAG	7.32
	GAGA	6.77

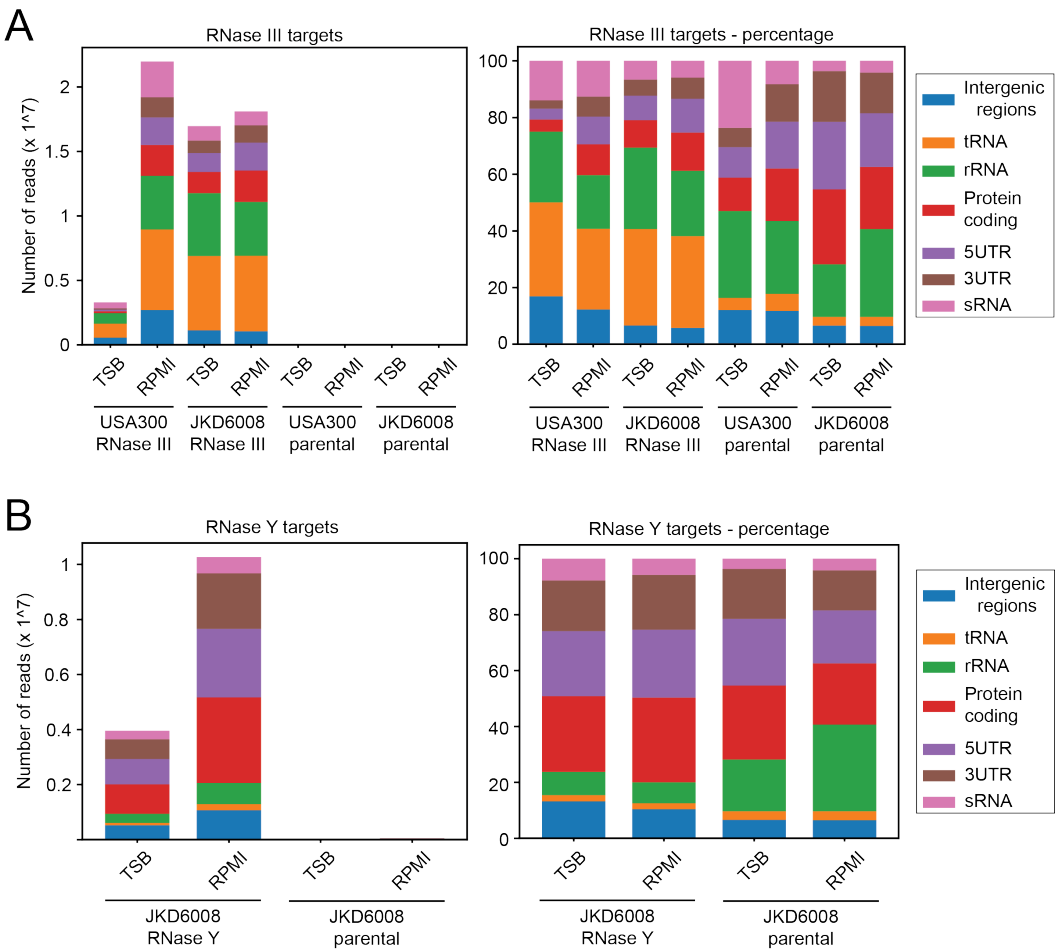

Fig. 4.12: RNase III and RNase Y target composition. Each panel shows the target classes of RNase III (**A.**) and RNase Y (**B.**) as raw counts (left) and as a percentage of the total (right).

RNA classes targeted by RNase Y do not significantly respond to the shift to RPMI. A small decrease is observed in intergenic regions and sRNAs (~3% and ~2% respectively) and a small increase for protein coding (~3%) after the shift to RPMI, but other transcript classes remain virtually stable (Figure 4.12B, upper panel). Interestingly, although RNase Y has not been associated with targeting rRNA transcripts, I identify the 23S, 16S and 5S transcripts as being targets. Looking at the target motifs again revealed a preference for G-A rich sequences (Table 4.2).

Tab. 4.2: RNase Y target motifs

Condition	Motif	Z-score
TSB	GGAG	12.20
	GAGA	11.26
	AGAG	10.40
	GAAG	10.02
	GAGG	9.60
RPMI	GAAG	12.32
	GAGA	11.14
	TGAAG	10.74
	GAAGA	10.54
	AAGA	10.48

4.7 Discussion

The first aim of this chapter was to explore the transcriptome of *S. aureus* in response to RPMI stress and confirm that sRNAs are likely to form a major component of the stress response. Next, this chapter aimed to prove that CLASH was a suitable method for detecting sRNA-target interactions. This latter aim was achieved through examining the data from a global perspective to see if found interactions fit the canonical sRNA mode of activity, and then at a finer level to examine if known sRNA interactions were identified.

sRNAs exhibit complex responses in response to stress

One of the most interesting aspects of the data presented in this chapter is the dynamism exhibited by sRNAs in response to RPMI. Many sRNAs were observed to have complex expression patterns that exhibited sharp changes on short timescales. Additionally, many sRNAs switched from being downregulated to upregulated, or *vice versa*, at different points in the timecourse. Firstly, this data shows the power of short-term timecourse experiments to resolve complex cell responses. Whilst the *Staphylococcal* field has traditionally examined sRNA expression profiles after time periods ranging from half an hour to 6 hours after stress (Bronesky et al., 2019; Geissmann et al., 2009; Howden et al., 2013), I demonstrate here that sRNAs are capable of acting on timescales as short as 5 minutes. Additionally, the entire response profile of some sRNAs to stress takes less than 30 minutes, such as RsaOG and RsaH in response to RPMI. These responses may be crucial for the initial stages of adaptation but would be missed in by traditional approaches to study them. Finally, the fact that some sRNAs can exhibit both up- and down-regulation within a 30-minute window suggests that within a specific stress, a single sRNA can be only desired at a particular time period, e.g. during the initial moment of adaptation but becomes unwanted and destroyed thereafter. As such, the role of sRNAs in stress adaptation may not be a simple case of 'turned off' or 'turned on'.

RPMI medium causes S. aureus to undergo significant transcriptional remodelling

After exposing the cells to RPMI, a huge number of transcripts exhibited significant differential expression. Some study into the behaviour of *S. aureus* in RPMI medium has already been carried out by the field. Dörries and Lalk, 2013, grew the *S. aureus* strains COL and HG001 in RPMI and found that it uptakes amino acids in the exponential phase of growth. Here, cells were shifted at late-exponential phase into RPMI and several operons involved in amino acid uptake and biosynthesis increased significantly in expression. In my data, I observed an upregulation in amino acid transporters, such as the *opp* operon, and also in the *ilv* operon which is involved

in the biosynthesis of branched chain amino acids. Given that CodY is known to bind branched chain amino acids such as isoleucine, CodY could be mediating this response (Majerczyk et al., 2008). Mäder et al, 2016, found that growing cells in RPMI led to increased expression in many iron-responsive genes, and here a large cluster of iron-related transcripts were observed to increase in expression after the shift. This could be mediated by inactivation of the Fur repressive transcription factor as a result of low intracellular iron after the shift to RPMI.

However, the RNAseq data carried out here also contributes novel findings. It was found that immediately after the shift to RPMI, several genes involved in glycolysis were upregulated during the first 5 minutes but then downregulated thereafter to levels below basal. As such, it may be that *S. aureus* tries to adapt to the initial phases of stress by upregulating glycolysis as a means to immediately generate the energy it requires to survive, before shutting it off – perhaps to utilise a different metabolic pathway afterwards.

Additionally, several virulence-related genes exhibited the same expression pattern as these glycolytic genes. The link between metabolism and virulence is becoming increasingly appreciated in bacteriology and is discussed at length in Chapter 5. However, I briefly mention here that this small correlation also adds to this idea.

Thus, the RNAseq data here serves to show that the experimental system recapitulates what is already known in the field, giving weight to the idea that this is a useful tool for examining sRNA-mediated regulation. Secondly, it gives further insight into the nature of sRNA expression, demonstrating their changing behavioural patterns. Finally, this data will be further discussed in Chapters 5 and 6 to examine the expression of novel targets of sRNAs to gain insight into sRNA-mediated destruction and sRNA sponging.

CLASH uncover sRNA – target interactions in S. aureus

Using RNase III as a bait protein for CLASH discovered many unique sRNA – target interactions. The first aim of my CLASH analysis was to demonstrate that this dataset was of high quality and could be used to uncover novel interactions. Satisfying this aim, I was able to identify many known interactions and also recover a significant number of interactions containing validated sRNAs. Secondly, examining the data from a global perspective revealed that interactions found between sRNAs and mRNAs frequently obeyed the canonical thought of sRNA-mediated regulation, with binding enriched in the 5' UTR and start codon region of the mRNA. As such, I conclude that the CLASH data presented here is a useful tool for uncovering novel sRNA targets.

One surprising aspect of the CLASH data is the large number of interactions between intergenic regions and mRNAs (including their 5' UTRs). This potentially suggests that there is a pool of unidentified sRNAs in *S. aureus* currently annotated as intergenic regions. Indeed, such regions are already a topic of focus in the bacteriology field (Argaman et al., 2001; Fuli et al., 2017; Rath et al., 2017; Tsai et al., 2015). One of the main advantages in using CLASH to discover these potentially novel sRNAs is that it also uncovers their potential targets and as such gives immediate insight into their possible function and avenues of future research.

Additionally, RNase III CLASH consistently found that 2-4% of the total interactome was represented by sRNA-sRNA interactions. The interaction between sRNAs themselves is becoming increasingly appreciated, and CLASH on RNase E in *E. coli* identified 150 sRNA-sRNA interactions, representing just under 2% of the total captured interactome (Waters et al., 2017). Additionally, Hfq CLASH in *E. coli* identified 90 sRNA-sRNA interactions, again representing just under 2% of the total (Iosub et al., 2018). Thus, interactions between sRNAs may be a consistent and sizeable proportion of the interactome across different bacterial species.

However, it is important to note that as sRNAs are able to bind multiple mRNA targets, they integrate many different signals into a single molecule. Thus, the single act of one sRNA binding to and sponging another away from its targets has the capacity to induce changes in the stability or translational efficiency of many target mRNAs. As such, while interactions between sRNAs may represent only ~2% of the interactome, the impact of these interactions may be much larger.

The RNase III CLASH data in RPMI identified a large number of interactions between sRNAs and intergenic regions. Similarly to that discussed above relating to mRNAs, these may represent novel sRNA – sRNA interactions. If this was the case, then this would greatly increase the incidence of this interaction type and would suggest that sRNA – sRNA interactions are more abundant than previously thought. Additionally, as RNase III was used to identify these, this would place RNase III as a major regulator of these interactions.

RNase III CLASH was also able to identify potentially novel mRNA targets of relatively well-understood sRNAs, namely RsaA, RsaE and RNAIII. It is noteworthy that for RsaA and RsaE, already known mRNA targets were identified. However, this was not the case for RNAIII. This is initially surprising as the targetome of RNAIII is expansive; it may be that at the growth stage CLASH was performed, RNAIII has already induced the degradation of its negatively regulated targets such as *sbi* and *rot*. Alternatively, it is possible that some interactions between RNAIII and mRNA targets are not directly recognized by RNase III and therefore not efficiently captured.

Examining the novel targets of RsaA, RsaE and RNAIII reveals encouraging avenues of future research. RsaA was found to interact with an antiseptic drug transporter, *qacAB*. This transporter belongs to the qac class and is involved in the export of quaternary ammonium compounds, intercalating dyes and β-lactam antibiotics (Jaglic and Cervinkova, 2018). Currently, RsaA is understood to indirectly regulate cell membrane homeostasis through its regulation of *mgrA*, a transcription factor that in turn regulates many membrane proteins (Tomasini et al., 2017). Additionally, MgrA is known to regulate capsule formation which is linked to antibiotic resistance (Campos et al., 2004; Geisinger and Isberg, 2015) and deletion of *mgrA*, in combination with *sarA*, causes increased susceptibility to cell wall-active antibiotics (Trotonda et al., 2009). Thus, the finding that RsaA may directly regulate a membrane protein involved in antibiotic and antiseptic resistance fits its already known role in membrane homeostasis, but with the novelty of being a direct interaction. This interaction is further investigated in Chapter 5.

RsaE was also found to interact with three other sRNAs; RNAIII, srn_2860 and srn_4190. Both srn_2860 and srn_4190 were originally identified by Howden et al, 2013, and are thought to be products of antisense transcription. As such, they were not considered for future study. However, the interaction with RNAIII was recovered with almost 450 unique hybrids. These sRNAs are already known to have a common target between them, *mgrA*, and also have opposing roles; RsaA promotes dormancy, while RNAIII promotes virulence (Gupta et al., 2015; Romilly et al., 2014). Thus, due to the already recognised importance of these sRNAs, this interaction was chosen for future study and is described further in Chapter 5.

Regarding RsaE, the most interesting candidate targets were *SAA6008_09902* and *SAA6008_09903*. These are α phenol-soluble modulins (PSMs), a class of virulence factors involved in host cell lysis, biofilm remodelling and intracellular escape (Cheung et al., 2014; Grosz Magdalena et al., 2014; Xu et al., 2017). Currently, RsaE is known to be involved in metabolic regulation, and so this interaction suggests a direct link between virulence and metabolism. Due to the importance of these behaviours, the interaction between RsaE and the αPSMs is further discussed and studied in Chapter 6.

Regarding RNAIII, this was observed to interact with four sRNAs. One, RsaA, has been discussed above. However, it was also found to interact with srn_1070 (otherwise known as Sau-41). This transcript is highly expressed in exponential and stationary phase, and is not present in small colony variants – a very similar pattern of expression as seen for RNAIII (Abu-Qatouseh et al., 2010). However, it is unclear if this is a genuine sRNA and it has been noted that it could be a 5' UTR of SAS014, a hypothetical protein (Felden et al., 2011). Thus, it could be an sRNA that acts in concert with RNAIII or potentially a novel mRNA target that, judging

from its expression pattern, may be positively regulated by RNAIII. The second 'sRNA' to interact with RNAIII, srn_4670 (otherwise known as RsaOT), has also been hypothesised to be a product of 3' end processing of its neighbouring mRNA, as it starts a mere 6 nucleotides after the stop codon (Bohn et al., 2010). Teg4 was identified by Howden et al, 2013, but has no meaningful information attributed to it. As such, I subsequently focused on RNAIII's interaction with RsaA.

RNAIII was also found to potentially interact with two virulence factors; *esxA* and *saeR*. EsxA is a secreted peptide that regulates host cell apoptosis (Korea et al., 2014; Sundaramoorthy et al., 2008) and it is interesting that RNAIII, a virulence-related RNA, is found directly interacting with it. This interaction is further discussed and validated in Chapter 5.

SaeR is the transcription factor response regulator of the *sae* two-component regulatory system (Giraudo et al., 1999). SaeR is activated when the sensor, SaeS, autophosphorylates in response to certain environmental signals such as human neutrophil peptides and calprotectin (Cho et al., 2015; Geiger et al., 2008). In turn, SaeR activates the expression of many virulence related genes such as *coa, nuc, hla, lukE, sbi, tst* and many other exotoxins. Interestingly, the *agr* operon has been found to positively regulate transcription of *sae* P1 promoter which induces transcription of the entire operon, including *saeR* (Geiger et al., 2008). Additionally, RNAIII is required for *sae* transcription (Novick and Jiang, 2003). However, it has been noted that the *sae* operon does not contain any AgrA binding sequences and that therefore the effect has been hypothesised to be indirect (Liu et al., 2016). However, the finding that RNAIII potentially interacts directly with *saeR* at the post-transcriptional level may partially explain these findings; RNAIII may stabilise the *saeR* transcript, allowing its translation. Although this interaction was not chosen for further study, it would be interesting to continue this through further validation.

Thus, overall, the CLASH methodology has successfully completed its aims. Firstly, it has proven that it can recover verified interactions. Secondly, it has successfully recovered novel sRNA-target interactions that fit the behavioural profile of selected sRNAs (such as RsaA's interaction with *qacAB*) and also suggested novel interactions that open up new avenues of research (e.g. linking metabolism virulence through RsaE). These interactions have been taken forward and further studied, as described in Chapters 5 and 6.

CRAC identifies RNA targets of RNase III and RNase Y

Another advantage to the CLASH protocol is its ability to give details on the single RNA targets of the protein of interest. In this chapter, I have detailed these analyses

from a global perspective. However, the CLASH data will be described further in Chapters 5 and 6 in order to model degradation of selected target genes.

A first observation from examining the single read data obtained from RNase III CLASH is that all classes of RNAs are recovered. This fits the known profile of RNase III, which has already been shown to be involved in rRNA and tRNA processing, turnover of sRNAs and mRNAs and processing of sRNA-mRNA duplexes in *S. aureus* (Lioliou et al., 2012). Additionally, RNase III has been found to target untranslated regions in *Streptococcus pyogenes* (Le Rhun et al., 2017) and here I also observed binding to 5' and 3' UTRs of mRNAs. Regarding specific transcripts, RNase III has also been found to bind many different sRNAs including RsaA, RsaE and RNAIII (Lioliou et al., 2012). This is encouraging as it gives credence to the interactions I uncovered between these sRNAs and their targets using CLASH.

RNase III has also been found to be a major regulator of antisense transcripts (Lasa et al., 2011; Lioliou et al., 2012). Such transcripts can be generated from transcription of the opposite strand of protein coding sequences or from overlapping long 5' and 3' UTRs of genes from opposite strands. This produces a noncoding transcript with perfect complementarity to the coding sequence and can cover the 5' or 3' end, coding sequence or even the entire gene (Lasa et al., 2011). RNase III has been found to target duplexes created by the interaction between sense and antisense transcripts (Lasa et al., 2011). Such targeting of antisense transcripts was beyond the scope of the work presented in this thesis, but recently analyses performed by Sander Granneman have demonstrated that these interactions are present in the data. As such, further interrogation of the data could further add to the list of antisense RNAs targeted by RNase III.

Examination of the target motif of RNase III revealed a preference for GA-rich sequences. However, this was also recovered in the RNase Y data. Although these sequences are highly enriched, it is potentially suspicious that two distinct nucleases recovered the exact same target motifs. However, *in vitro* studies have shown a preference in RNase III for G-C pairings in *E. coli* (Altuvia et al., 2018). It may also be that RNase III recognises the structure of its targets more so than any specific nucleotide tract. This fits with the known data of RNase III as it specifically targets double-stranded RNA and in particular stem structures.

Additionally, Altuvia et al, 2018, found that RNase III is capable of cleaving intramolecular stem structures formed within an individual RNA. This avenue of research could be greatly furthered by the CLASH data, as CLASH is capable of not only detecting intermolecular interactions (i.e. RNA1 – RNA2) but also intramolecular interactions (i.e. RNA1 – RNA1). Although not presented in this thesis, the CLASH data contains huge numbers of intramolecular interactions which could be

the subject of further analysis to examine the structures of RNase III targets. Indeed, a collaboration has been set up with Dr Grzegorz Kudla to further examine the structures of targeted RNAs and preliminary results indicate that RNase III targets are generally more folded compared to other regions of the genome, contain tetraloop structures and are depleted in long regions of unfolded nucleotides. This research will be continued and included in any upcoming publication.

RNase Y has also been subject of some study in the *Staphylococcal* field. Deletion mutants and mapping of the 5' ends of its targets identified around 250 targets of RNase Y, consisting of mRNAs and sRNAs (Khemici et al., 2015). Some of these sRNAs were recovered in this study such as RsaA, SprX and RNAIII, as were members the *sae* operon that have been shown to be targeted by RNase Y (Marincola Gabriella et al., 2012). However, the authors noted that they did not identify any convincing rRNA transcripts subject to RNase Y-mediated cleavage which disagrees with the data presented here. The authors used a silica column-based method for their RNA purification which is noted to select for RNAs above 200 nucleotides; as such, they may have missed the short 5S transcript that we identify as an RNase Y target. However, this does not explain the why I identify 23S and 16S transcripts as RNase Y targets and they did not.

Khemici et al, 2015, also found that RNase Y does not have a strong target motif, noting only that they consistently found a guanine nucleotide prior to any cleavage site. They proposed that secondary structure recognition may be important for RNase Y targeting. As such, my data could also be further analysed for secondary structure in a similar manner to that being carried out for RNase III. However, it is noted that in *B. subtilis*, RNase Y is thought to target AU-rich regions (Shahbabian et al., 2009).

Furthermore, RNase Y is a membrane-bound RNase that is incorporated into a large degradosome-like complex (Marincola Gabriella et al., 2012; Roux et al., 2011). Removal of the membrane anchor results in a severe growth defect (Marincola Gabriella et al., 2012), and so it may be that RNase Y target acquisition is mediated through the delivery of selected RNAs to its subcellular location through degradosome partners. If this were the case, then this may remove the need for a consensus RNase Y target motif and instead place this burden on the RBP partners.

Future directions

Already described above is how the single read data will be mined to examine structural motifs targeted by RNase III and potentially RNase Y. However, the main aim of this chapter and its constituent analyses was to identify sRNA-target interactions for future study. This has been achieved and the interactions between

RsaA, RNAIII and *qacAB*; RsaE, *SAA6008_09902* and *SAA6008_09902*; and RNAIII and *esxA* have been selected for future study. These are discussed in the upcoming Chapter 5 and Chapter 6. However, the CLASH data as a whole will be of great interest to the *Staphylococcal* field and I suspect that, when published, it will be mined by other groups for potentially novel sRNA-target interactions. These will undoubtedly involve known sRNAs but also novel, unidentified ones currently annotated as intergenic regions. As such, the CLASH data presented here has great potential in expanding the field considerably.

4.8 Bibliography

Abu-Qatouseh LF, Chinni SV, Seggewiß J, Proctor RA, Brosius J, Rozhdestvensky TS, Peters G, von Eiff C, Becker K. 2010. Identification of differentially expressed small non-protein-coding RNAs in *Staphylococcus aureus* displaying both the normal and the small-colony variant phenotype. J Mol Med 88:565–575. doi:10.1007/s00109-010-0597-2

Altuvia Y, Bar A, Reiss N, Karavani E, Argaman L, Margalit H. 2018. *In vivo* cleavage rules and target repertoire of RNase III in *Escherichia coli*. Nucleic Acids Res 46:10380–10394. doi:10.1093/nar/gky684

Argaman L, Hershberg R, Vogel J, Bejerano G, Wagner EGH, Margalit H, Altuvia S. 2001. Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. Curr Biol 11:941–950. doi:10.1016/S0960-9822(01)00270-6

Bohn C, Rigoulay C, Chabelskaya S, Sharma CM, Marchais A, Skorski P, Borezée-Durant E, Barbet R, Jacquet E, Jacq A, Gautheret D, Felden B, Vogel J, Bouloc P. 2010. Experimental discovery of small RNAs in *Staphylococcus aureus* reveals a riboregulator of central metabolism. Nucleic Acids Res 38:6620–6636. doi:10.1093/nar/gkq462

Boisset S, Geissmann T, Huntzinger E, Fechter P, Bendridi N, Possedko M, Chevalier C, Helfer AC, Benito Y, Jacquier A, Gaspin C, Vandenesch F, Romby P. 2007. *Staphylococcus aureus* RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. Genes Dev 21:1353–1366. doi:10.1101/gad.423507

Bronesky D, Desgranges E, Corvaglia A, François P, Caballero CJ, Prado L, Toledo-Arana A, Lasa I, Moreau K, Vandenesch F, Marzi S, Romby P, Caldelari I. 2019. A multifaceted small RNA modulates gene expression upon glucose limitation in *Staphylococcus aureus*. EMBO J 38:e99363. doi:10.15252/embj.201899363

Campos MA, Vargas MA, Regueiro V, Llompart CM, Albertí S, Bengoechea JA. 2004. Capsule Polysaccharide Mediates Bacterial Resistance to Antimicrobial Peptides. Infect Immun 72:7107–7114. doi:10.1128/IAI.72.12.7107-7114.2004

Cheung GYC, Joo H-S, Chatterjee SS, Otto M. 2014. Phenol-soluble modulins – critical determinants of *Staphylococcal* virulence. FEMS Microbiol Rev 38:698–719. doi:10.1111/1574-6976.12057

Cho H, Jeong D-W, Liu Q, Yeo W-S, Vogl T, Skaar EP, Chazin WJ, Bae T. 2015. Calprotectin Increases the Activity of the SaeRS Two Component System and Murine Mortality during *Staphylococcus aureus* Infections. PLoS Pathog 11. doi:10.1371/journal.ppat.1005026

Diep BA, Gill SR, Chang RF, Phan TH, Chen JH, Davidson MG, Lin F, Lin J, Carleton HA, Mongodin EF, Sensabaugh GF, Perdreau-Remington F. 2006. Complete genome sequence of USA300, an epidemic clone of community-acquired meticillin-resistant Staphylococcus aureus 367:9.

Ernst J, Bar-Joseph Z. 2006. STEM: a tool for the analysis of short time series gene expression data. BMC Bioinformatics 7:191. doi:10.1186/1471-2105-7-191

Eyraud A, Tattevin P, Chabelskaya S, Felden B. 2014. A small RNA controls a protein regulator involved in antibiotic resistance in *Staphylococcus aureus*. Nucleic Acids Res 42:4892–4905. doi:10.1093/nar/gku149

Felden B, Vandenesch F, Bouloc P, Romby P. 2011. The *Staphylococcus aureus* RNome and Its Commitment to Virulence. PLOS Pathog 7:e1002006. doi:10.1371/journal.ppat.1002006

Fuli X, Wenlong Z, Xiao W, Jing Z, Baohai H, Zhengzheng Z, Bin-Guang M, Youguo L. 2017. A Genome-Wide Prediction and Identification of Intergenic Small RNAs by Comparative Analysis in *Mesorhizobium huakuii* 7653R. Front Microbiol 8. doi:10.3389/fmicb.2017.01730

Geiger T, Goerke C, Mainiero M, Kraus D, Wolz C. 2008. The Virulence Regulator Sae of *Staphylococcus aureus*: Promoter Activities and Response to Phagocytosis-Related Signals. J Bacteriol 190:3419–3428. doi:10.1128/JB.01927-07

Geisinger E, Isberg RR. 2015. Antibiotic Modulation of Capsular Exopolysaccharide and Virulence in Acinetobacter baumannii. PLOS Pathog 11:e1004691. doi:10.1371/journal.ppat.1004691

Geissmann T, Chevalier C, Cros M-J, Boisset S, Fechter P, Noirot C, Schrenzel J, François P, Vandenesch F, Gaspin C, Romby P. 2009. A search for small noncoding RNAs in *Staphylococcus aureus* reveals a conserved sequence motif for regulation. Nucleic Acids Res 37:7239–7257. doi:10.1093/nar/gkp668

Giraudo AT, Calzolari A, Cataldi AA, Bogni C, Nagel R. 1999. The sae locus of *Staphylococcus aureus* encodes a two-component regulatory system. FEMS Microbiol Lett 177:15–22. doi:10.1111/j.1574-6968.1999.tb13707.x

Grosz Magdalena, Kolter Julia, Paprotka Kerstin, Winkler Ann-Cathrin, Schäfer Daniel, Chatterjee Som Subra, Geiger Tobias, Wolz Christiane, Ohlsen Knut, Otto Michael, Rudel Thomas, Sinha Bhanu, Fraunholz Martin. 2014. Cytoplasmic replication of *Staphylococcus aureus* upon phagosomal escape triggered by phenol-soluble modulin α. Cell Microbiol 16:451–465. doi:10.1111/cmi.12233

Gupta RK, Luong TT, Lee CY. 2015. RNAIII of the *Staphylococcus aureus* agr system activates global regulator MgrA by stabilizing mRNA. Proc Natl Acad Sci 112:14036–14041. doi:10.1073/pnas.1509251112

Holden MTG, Lindsay JA, Corton C, Quail MA, Cockfield JD, Pathak S, Batra R, Parkhill J, Bentley SD, Edgeworth JD. 2010. Genome Sequence of a Recently Emerged, Highly Transmissible, Multi-Antibiotic- and Antiseptic-Resistant Variant of Methicillin-Resistant Staphylococcus aureus, Sequence Type 239 (TW). J Bacteriol 192:888–892. doi:10.1128/JB.01255-09

Howden BP, Beaume M, Harrison PF, Hernandez D, Schrenzel J, Seemann T, Francois P, Stinear TP. 2013. Analysis of the Small RNA Transcriptional Response in Multidrug-Resistant *Staphylococcus aureus* after Antimicrobial Exposure. Antimicrob Agents Chemother 57:3864–3874. doi:10.1128/AAC.00263-13

Iosub IA, Marchioretto M, Sy B, McKellar S, Nieken KJ, Nues RW van, Tree JJ, Viero G, Granneman S. 2018. Hfq CLASH uncovers sRNA-target interaction networks enhancing adaptation to nutrient availability. bioRxiv 481986. doi:10.1101/481986

Jaglic Z, Cervinkova D. 2018. Genetic basis of resistance to quaternary ammonium compounds; the qac genes and their role: a review. Veterinární Medicína 57:275–281. doi:10.17221/6013-VETMED

Khemici V, Prados J, Linder P, Redder P. 2015. Decay-Initiating Endoribonucleolytic Cleavage by RNase Y Is Kept under Tight Control via Sequence Preference and Subcellular Localisation. PLOS Genet 11:e1005577. doi:10.1371/journal.pgen.1005577

Korea CG, Balsamo G, Pezzicoli A, Merakou C, Tavarini S, Bagnoli F, Serruto D, Unnikrishnan M. 2014. Staphylococcal Esx proteins modulate apoptosis and release of intracellular *Staphylococcus aureus* during infection in epithelial cells. Infect Immun 82:4144–4153. doi:10.1128/IAI.01576-14

Lasa I, Toledo-Arana A, Dobin A, Villanueva M, de los Mozos IR, Vergara-Irigaray M, Segura V, Fagegaltier D, Penadés JR, Valle J, Solano C, Gingeras TR. 2011. Genome-wide antisense transcription drives mRNA processing in bacteria. Proc Natl Acad Sci U S A 108:20172–20177. doi:10.1073/pnas.1113521108

Le Rhun A, Lécrivain A-L, Reimegård J, Proux-Wéra E, Broglia L, Della Beffa C, Charpentier E. 2017. Identification of endoribonuclease specific cleavage positions reveals novel targets of RNase III in *Streptococcus pyogenes*. Nucleic Acids Res 45:2329–2340. doi:10.1093/nar/gkw1316

Li M, Diep BA, Villaruz AE, Braughton KR, Jiang X, DeLeo FR, Chambers HF, Lu Y, Otto M. 2009. Evolution of virulence in epidemic community-associated methicillin-resistant Staphylococcus aureus. Proc Natl Acad Sci 106:5883–5888. doi:10.1073/pnas.0900743106

Lioliou E, Sharma CM, Caldelari I, Helfer A-C, Fechter P, Vandenesch F, Vogel J, Romby P. 2012. Global Regulatory Functions of the *Staphylococcus aureus* Endoribonuclease III in Gene Expression. PLOS Genet 8:e1002782. doi:10.1371/journal.pgen.1002782

Liu Q, Yeo W-S, Bae T. 2016. The SaeRS Two-Component System of *Staphylococcus aureus*. Genes 7. doi:10.3390/genes7100081

Liu W, Rochat T, Toffano-Nioche C, Le Lam *TN, Bouloc P, Morvan* C. 2018. Assessment of Bona Fide sRNAs in Staphylococcus aureus. Front Microbiol 9. doi:10.3389/fmicb.2018.00228

Lorenz R, Bernhart SH, Höner zu Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL. 2011. ViennaRNA Package 2.0. Algorithms Mol Biol 6. doi:10.1186/1748-7188-6-26

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15. doi:10.1186/s13059-014-0550-8

Mäder U, Nicolas P, Depke M, Pané-Farré J, Debarbouille M, Kooi-Pol MM van der, Guérin C, Dérozier S, Hiron A, Jarmer H, Leduc A, Michalik S, Reilman E, Schaffer M, Schmidt F, Bessières P, Noirot P, Hecker M, Msadek T, Völker U, Dijl JM van. 2016. *Staphylococcus aureus* Transcriptome Architecture: From Laboratory to Infection-Mimicking Conditions. PLOS Genet 12:e1005962. doi:10.1371/journal.pgen.1005962

Marincola Gabriella, Schäfer Tina, Behler Juliane, Bernhardt Jörg, Ohlsen Knut, Goerke Christiane, Wolz Christiane. 2012. RNase Y of *Staphylococcus aureus* and its role in the activation of virulence genes. Mol Microbiol 85:817–832. doi:10.1111/j.1365-2958.2012.08144.x

McDougal LK, Steward CD, Killgore GE, Chaitram JM, McAllister SK, Tenover FC. 2003. Pulsed-Field Gel Electrophoresis Typing of Oxacillin-Resistant Staphylococcus aureus Isolates from the United States: Establishing a National Database. J Clin Microbiol 41:5113–5120. doi:10.1128/JCM.41.11.5113-5120.2003

Moran GJ, Gorwitz RJ, McDougal LK. 2006. Methicillin-Resistant S. aureus Infections among Patients in the Emergency Department. N Engl J Med 9.

Novick RP, Jiang D. 2003. The *Staphylococcal* saeRS system coordinates environmental signals with agr quorum sensing. Microbiol Read Engl 149:2709–2717. doi:10.1099/mic.0.26575-0

Pichon C, Felden B. 2005. Small RNA genes expressed from *Staphylococcus aureus* genomic and pathogenicity islands with specific expression among pathogenic strains. Proc Natl Acad Sci 102:14249–14254. doi:10.1073/pnas.0503838102

Pinel-Marie M-L, Brielle R, Felden B. 2014. Dual Toxic-Peptide-Coding *Staphylococcus aureus* RNA under Antisense Regulation Targets Host Cells and Bacterial Rivals Unequally. Cell Rep 7:424–435. doi:10.1016/j.celrep.2014.03.012

Ramani V, Qiu R, Shendure J. 2015. High-throughput determination of RNA structure by proximity ligation. Nat Biotechnol 33:980–984. doi:10.1038/nbt.3289

Rath EC, Pitman S, Cho KH, Bai Y. 2017. Identification of *Streptococcal* small RNAs that are putative targets of RNase III through bioinformatics analysis of RNA sequencing data. BMC Bioinformatics 18:540. doi:10.1186/s12859-017-1897-0

Robinson DA, Enright MC. 2004. Evolution of Staphylococcus aureus by Large Chromosomal Replacements. J Bacteriol 186:1060–1064. doi:10.1128/JB.186.4.1060-1064.2004

Romilly C, Lays C, Tomasini A, Caldelari I, Benito Y, Hammann P, Geissmann T, Boisset S, Romby P, Vandenesch F. 2014. A Non-Coding RNA Promotes Bacterial Persistence and Decreases Virulence by Regulating a Regulator in *Staphylococcus aureus*. PLOS Pathog 10:e1003979. doi:10.1371/journal.ppat.1003979

Roux CM, DeMuth JP, Dunman PM. 2011. Characterization of Components of the *Staphylococcus aureus* mRNA Degradosome Holoenzyme-Like Complex. J Bacteriol 193:5520–5526. doi:10.1128/JB.05485-11

Sassi M, Augagneur Y, Mauro T, Ivain L, Chabelskaya S, Hallier M, Sallou O, Felden B. 2015. SRD: a *Staphylococcus* regulatory RNA database. RNA 21:1005–1017. doi:10.1261/rna.049346.114

Sayed N, Jousselin A, Felden B. 2012. A *cis*-antisense RNA acts in trans in *Staphylococcus aureus* to control translation of a human cytolytic peptide. Nat Struct Mol Biol 19:105–112. doi:10.1038/nsmb.2193

Shahbabian K, Jamalli A, Zig L, Putzer H. 2009. RNase Y, a novel endoribonuclease, initiates riboswitch turnover in Bacillus subtilis. EMBO J 28:3523–3533. doi:10.1038/emboj.2009.283

Sundaramoorthy R, Fyfe PK, Hunter WN. 2008. Structure of *Staphylococcus aureus* EsxA suggests a contribution to virulence by action as a transport chaperone and/or adaptor protein. J Mol Biol 383:603–614. doi:10.1016/j.jmb.2008.08.047

Talan DA, Krishnadasan A, Gorwitz RJ, Fosheim GE, Limbago B, Albrecht V, Moran GJ. 2011. Comparison of Staphylococcus aureus From Skin and Soft-Tissue Infections in US Emergency Department Patients, 2004 and 2008. Clin Infect Dis 53:144–149. doi:10.1093/cid/cir308

Tomasini A, Moreau K, Chicher J, Geissmann T, Vandenesch F, Romby P, Marzi S, Caldelari I. 2017. The RNA targetome of *Staphylococcus aureus* non-coding RNA RsaA: impact on cell surface properties and defense mechanisms. Nucleic Acids Res 45:6746–6760. doi:10.1093/nar/gkx219

Trotonda MP, Xiong YQ, Memmi G, Bayer AS, Cheung AL. 2009. Role of mgrA and sarA in methicillin-resistant *Staphylococcus aureus* autolysis and resistance to cell wall-active antibiotics. J Infect Dis 199:209–218. doi:10.1086/595740

Tsai C-H, Liao R, Chou B, Palumbo M, Contreras LM. 2015. Genome-Wide Analyses in Bacteria Show Small-RNA Enrichment for Long and Conserved Intergenic Regions. J Bacteriol 197:40–50. doi:10.1128/JB.02359-14

Waters SA, McAteer SP, Kudla G, Pang I, Deshpande NP, Amos TG, Leong KW, Wilkins MR, Strugnell R, Gally DL, Tollervey D, Tree JJ. 2017. Small RNA interactome of pathogenic *E. coli* revealed through crosslinking of RNase E. EMBO J 36:374–387. doi:10.15252/embj.201694639

Weiss A, Broach WH, Shaw LN. 2016. Characterizing the transcriptional adaptation of *Staphylococcus aureus* to stationary phase growth. Pathog Dis 74:ftw046. doi:10.1093/femspd/ftw046

Xu T, Wang X-Y, Cui P, Zhang Y-M, Zhang W-H, Zhang Y. 2017. The Agr Quorum Sensing System Represses Persister Formation through Regulation of Phenol Soluble Modulins in *Staphylococcus aureus*. Front Microbiol 8. doi:10.3389/fmicb.2017.02189

5 RsaA and RNAIII interact *in vivo* and are targeted by RNase III

5.1 Introduction

As discussed in Chapter 1, sRNAs have already been shown to play key roles in crucial signalling pathways including virulence and antibiotic resistance. RNAIII is the most well characterised sRNA in *S. aureus*. This is a pro-virulence factor that acts as the main effector from the quorum-sensing *agr* pathway (Section 1.7). It acts to switch the cell from a state of dormancy into one of aggression. This is accomplished through its regulation of several target mRNAs; it stimulates translation of hla and mgrA, whilst inducing RNase III-mediated destruction of *rot, spa, coa*, and *sbi*.

As discussed in Section 1.8, RsaA is thought to be involved in the attenuation of virulence. RsaA is known to bind to *mgrA* at two distinct regions; one interaction covers the ribosomal binding site and prevents translation, while the second recruits RNase III and induces degradation. Additionally, RsaA has been found to be involved in regulating cell surface proteins such as ClfB, Ebh and Protein A.

The CLASH experiments revealed a large number of interactions between RsaA and RNAIII in TSB medium. When the cells were shifted to RPMI, this interaction was then found at a greatly reduced incidence. Additionally, the RNA-seq data showed that RNAIII is immediately downregulated after the shift. Thus, I hypothesised that RNAIII and RsaA are capable of interacting *in vivo*, that this interaction responds to the stresses incurred, and that it may regulate the stability of RNAIII. This chapter will detail validations of this interaction and hypothesise it helps to shape *S. aureus*' adaptation.

5.2 CLASH identifies known targets of RsaA

In order to gain greater insight into the reliability of CLASH, the data was examined to see if already known targets of RsaA could be identified. One of most well studied targets of RsaA is the pro-virulence transcription factor, *mgrA*. RsaA is known to bind to the 5' end of *mgrA* to prevent the 30S ribosomal subunit from binding, and also

to the 3' end (Romilly et al., 2014). For JKD6008, the CLASH data obtained around 300 unique hybrids between RsaA and *mgrA* in RPMI and 28 in TSB; in USA300, 61 unique hybrids were identified in RPMI. All of these mapped to the 5' end of *mgrA*. As seen in Figure 5.1, the hybrids obtained through CLASH recover the entire 3'-end interaction, including the bases covering the AUG of the start codon.

SsaA belongs to the Staphylococcal secretory antigen A family which is involved in peptidoglycan metabolism. *SsaA2_3* and *SsaA_2* were identified as targets of RsaA through MAPS, and then validated through *in vitro* EMSAs and toeprinting assays (Tomasini et al., 2017). This found that RsaA binds to the RBS and prevents translational initiation. The interaction between RsaA and *Ssa2_3* was recovered with CLASH, but only a single hybrid containing a short fragment of *SsaA2_3* was identified. The reads relating to *SsaA2_3* mapped to the 5' UTR exclusively and did not extend into the SD sequence. Thus, the CLASH data validates the *SsaA2_3* interaction with RsaA but with a slightly different binding region.

MAPS also uncovered an SH3 domain-containing protein interacting with RsaA, named *HG001_01977* in the HG001 strain (Tomasini et al., 2017). This is equivalent to *JKD6008_01954* locus in the JKD6008 strain, and this interaction was also uncovered with CLASH. However, Tomasini et al., 2017, did not manage to identify precisely the region of interaction. One of the advantages of CLASH over MAPS is that CLASH gives insight into which regions of two RNAs interact together, and data described here finds that RsaA binds to the 5' UTR of *JKD6008_01954*, before the SD sequence.

In conclusion, CLASH is able to confirm already known interactions between RsaA and its targets. This helps to confirm the CLASH approach and gives credence to novel interactions for RsaA.

5.3 CLASH identifies RNAIII as a target for RsaA

The most abundant interaction uncovered for RsaA was with RNAIII. This interaction was found in both the JKD6008 and USA300 data. When summing the total number of interactions in TSB and in RPMI, it was found that ~5X more interactions were identified in TSB versus RPMI; 282 versus 57 respectively (Figure 5.2A). Although CLASH is not strictly quantitative, such a large difference postulated the idea that this interaction could respond to specific mediums. When hybrids were mapped to the genomic locus of each gene, it was found that is the 5' end of RsaA which interacts with the 3' end of RNAIII (Figure 5.2B). For RNAIII, hybrid reads involved with base-pairing to RsaA mapped to the 8th, 9th, and 10th hairpin structures, using

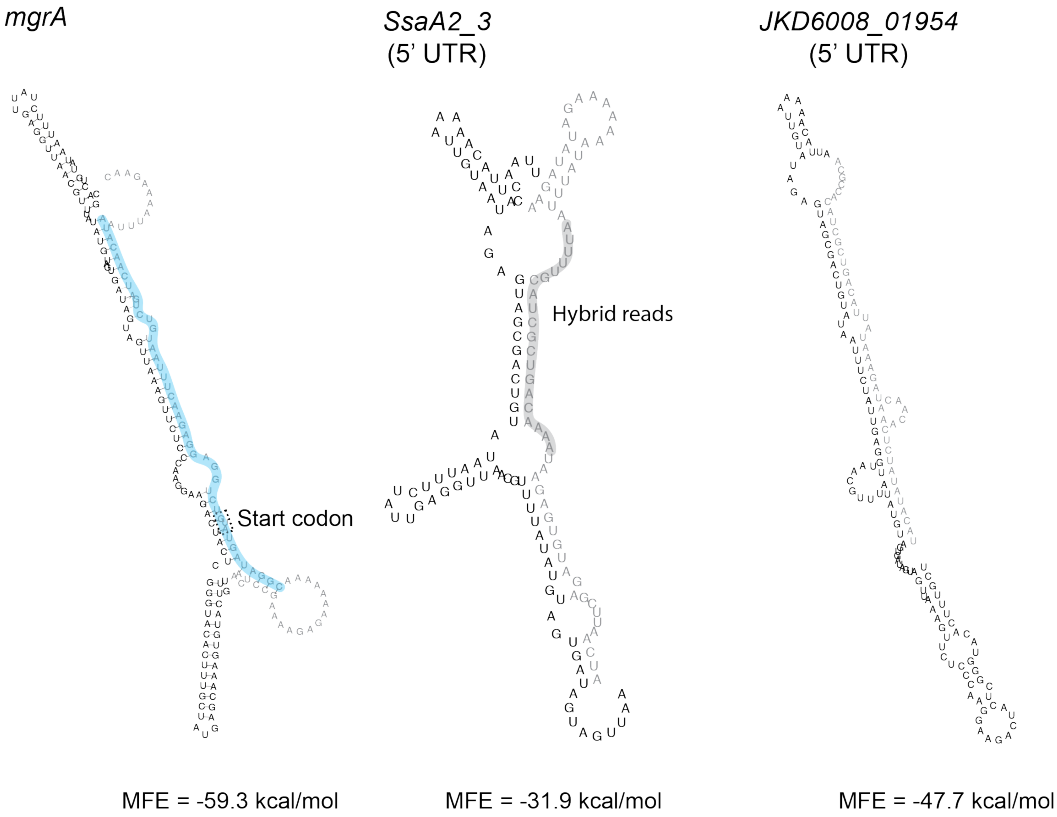

Fig. 5.1: Predicted interactions between RsaA and its known targets. RsaA is shown in grey and the targets in black. For *mgrA*, the 5' site of interaction found by Romilly et al is highlighted in blue. Structures were generated by extracting the CLASH read sequences as input for RNAcofold, which was used to predict the interaction structure. For the interaction with *SsaA2_3*, such a short read was uncovered (highlighted grey) that 20 nucleotides from either side of the hybrid were added to the 5' and 3' end.

the nomenclature from Bronesky et al., 2016. Interestingly, this region in RNAIII is far from the small ORF encoding for δ-haemolysin and thus I hypothesised that this interaction regulates the activity or stability of these sRNAs, as opposed to regulating translation of *hld*. Indeed, when the nucleotides involved in this interaction are extracted and folded *in silico*, an extended region of interaction is observed with a very favourable minimum folding energy of -93.30 kcal/mol.

As sequence conservation across bacterial species is indicative of a preserved functionality, conservation analysis was performed on RsaA and the interacting region of RNAIII (Figure 5.3B). RsaA does not display particularly strong conservation overall, although there exists a 'ACGTATATA' motif that is somewhat preserved throughout the examined species. Additionally, the characteristic 'UCCC' motif ('TCCC' in figure Figure 5.2C) of the Rsa sRNAs is conserved between *S. aureus*, *S. argenteus* and *S. simiae*. The interacting region of RNAIII exhibits stronger conservation, with significant variation only observed in *S. argenteus*. However, this conservation could also be related to the fact that the hairpin structures involved in interacting with

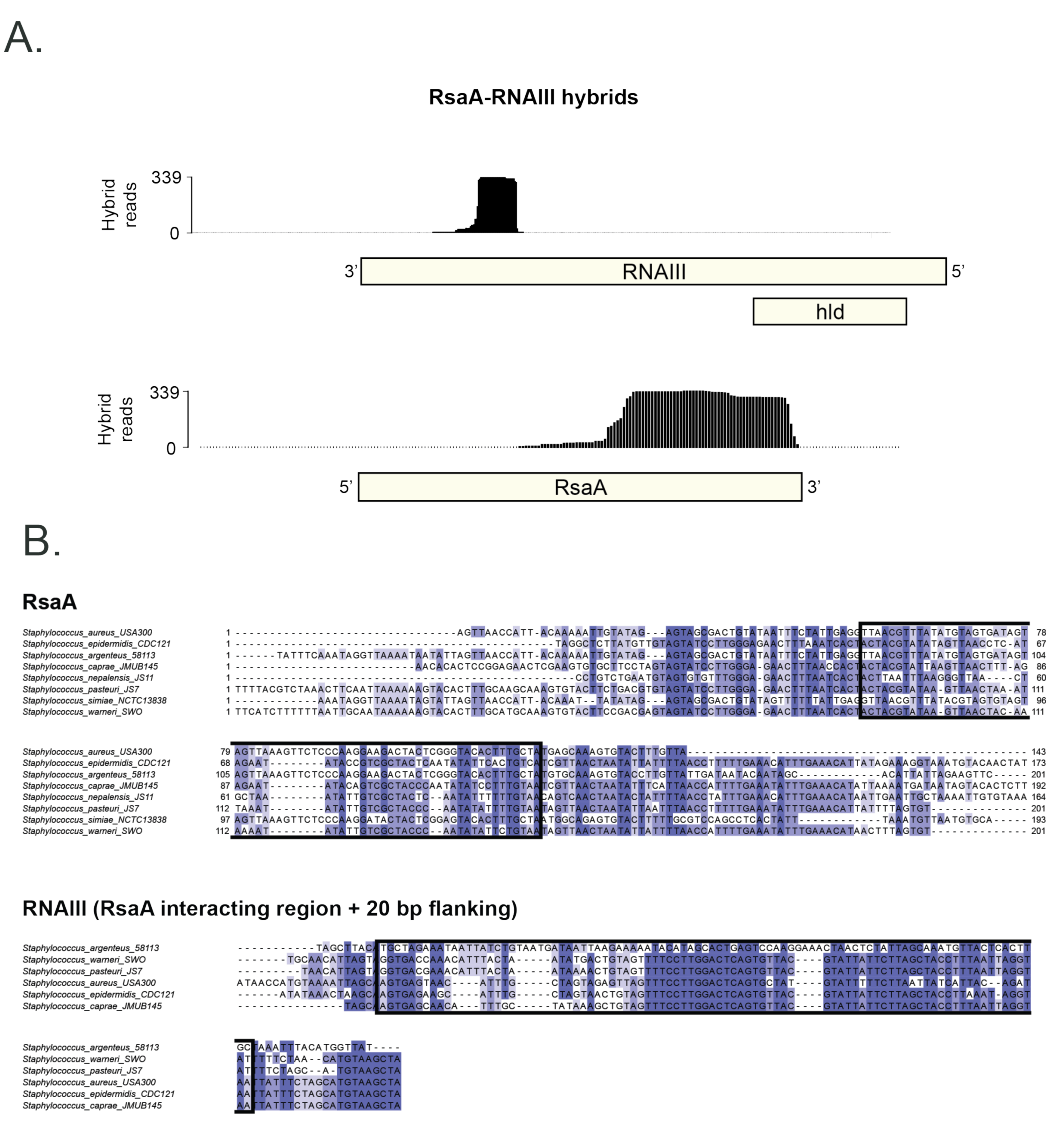

Fig. 5.2: Proposed interaction between RsaA and RNAIII. **A.** Mapping the hybrids to the genomic loci of their respective gene. **B.** Conservation analysis of RsaA and the interacting region of RNAIII (with 20 flanking nucleotides on both ends). The boxed regions indicate those which are involved in the interaction.

RsaA could also be used for interacting with other transcripts; hairpin 10 is already known to interact with *coa*.

Figure 5.3 shows the regions of interaction onto the structures of the sRNAs themselves. This shows clearly how interacting region of RNAIII (Figure 5.3A) map to the 8th, 9th and 10th hairpins, while for RsaA this region is found at the 3' end and covers the functional UCCC motif (Figure 5.3B). For *in silico* analyses, a portion of RNAIII from the 7th to 11th hairpin was taken and folded with the full length RsaA. This revealed an extended region of interaction which melts the 9th helix of RNAIII and the UCCC-containing helix of RsaA (Figure 5.3C). The interaction also has a favourable folding energy of -93.90 kcal/mol. Thus, overall, it is predicted that it

is the activity of hairpin 9 of RNAIII interacting with the UCCC-containing loop of RsaA that mediates the interaction.

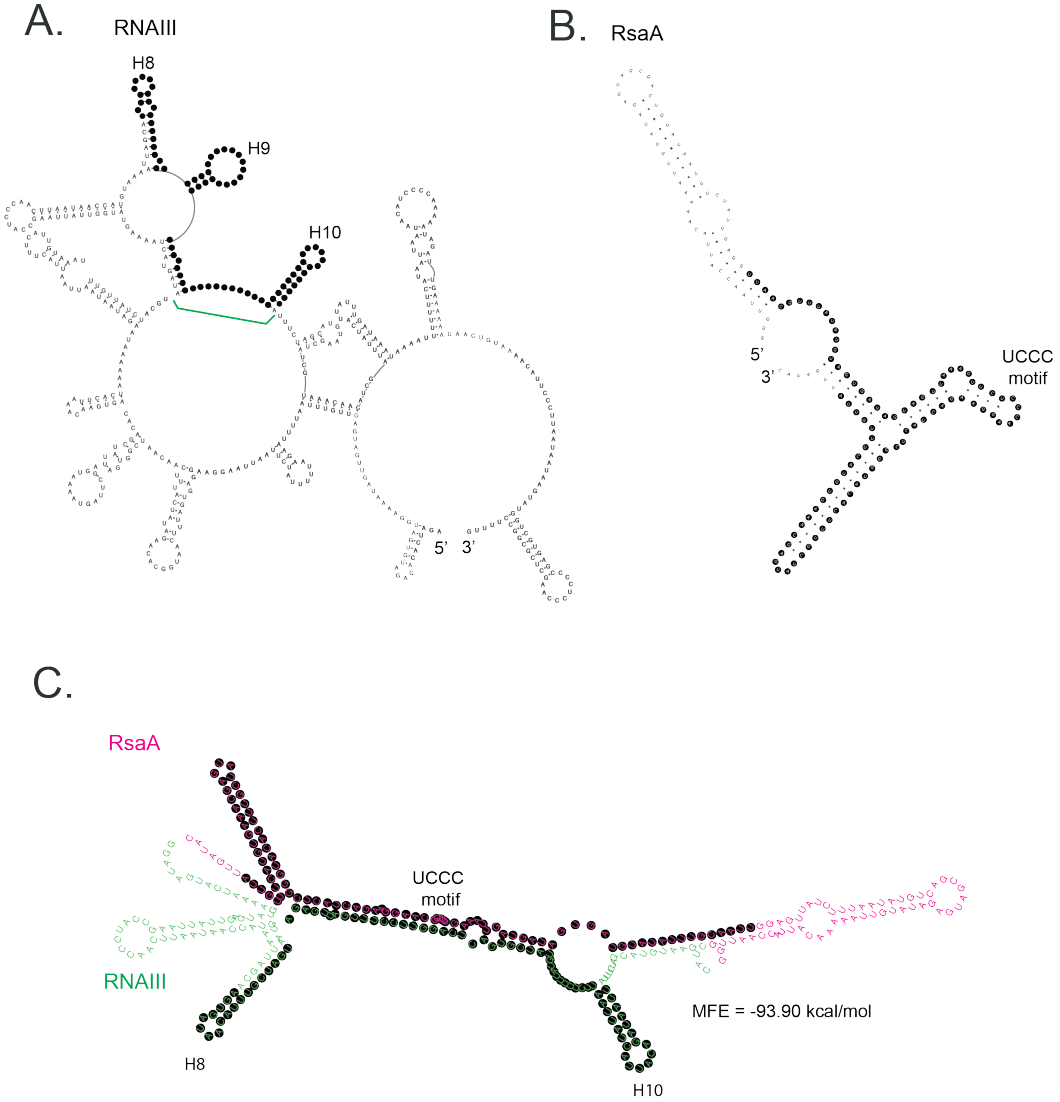

Fig. 5.3: Regions of interaction for RNAIII and RsaA. **A.** RNAIII structure with the hybrid reads identified in CLASH interacting with RsaA outlined in black. The nucleotides in between the green line were taken for *in silico* folding analysis as shown in C. **B.** RsaA structure with the hybrid reads interacting with RNAIII in black. The UCCC motif is labelled. **C.** *In silico* predicted structure of RNAIII (7th – 11th helices) interacting with RsaA. The nucleotides surrounded by the black border were identified in CLASH.

5.4 RsaA and RNAIII are able to interact specifically *in vitro*

In order to validate the idea that these two RNAs interact specifically through the 9th helix of RNAIII, *in vitro* electrophoretic mobility shift assays (EMSAs) were per-

formed. For this, two synthetic fragments of RNAIII were generated *in vitro*; RNAIII interacting, consisting of helices 7 to 11; and RNAIII non-interacting, consisting of helices 2 to 6. This RNAIII non-interacting construct was to act as a negative control, being specifically chosen as no reads from this region of RNAIII were found interacting with RsaA in the CLASH data. However, folding this region with RsaA *in silico* gives a similar MFE energy to RNAIII interacting domain; -72.80 kcal/mol versus -93.90 kcal/mol respectively. Thus, I predicted that this region of RNAIII would serve as a fair negative control to examine non-specific binding of RNAIII to RsaA *in vitro*.

RNAIII interacting and RsaA were first confirmed to fold as expected through SHAPE-based structure probing. This method uses a chemical probe, 2-methylnicotinic acid imidazolide (NAI), that forms adducts with flexible nucleotides, which are predominantly found in single-stranded regions such as bulges or internal loops. These adducts can terminate reverse transcription (Deigan et al., 2009). When using low concentrations of NAI, the location of these modification sites can be identified by primer-extension analysis, which enabled me to detect flexible regions in the *in vitro* folded RNAs.

Structure probing experiments for both RNAIII interacting and RsaA indicated that the *in vitro* produced RNAs folded largely as expected as most of the NAI modification sites mapped to regions that were predicted to be single-stranded (Kolb et al., 2000). For RNAIII interacting, flexible nucleotides were found just before helix 7 and in the loop region. This pattern was also observed for helices 8, 9 and 10 (Figure 5.4B). Regarding RsaA, free nucleotides were identified in bulges that formed from imperfect base-pairing sequences, and at the first loop region, although no free nucleotides were detected at the UCCC motif. Due to time constrains, the structure of RNAIII non-interacting was not probed. However, overall, it was deduced that RNAIII interacting and RsaA fold as expected. I also plan to probe the RNAIII-RsaA duplex to see if I can validate the structure predicted by the CLASH data.

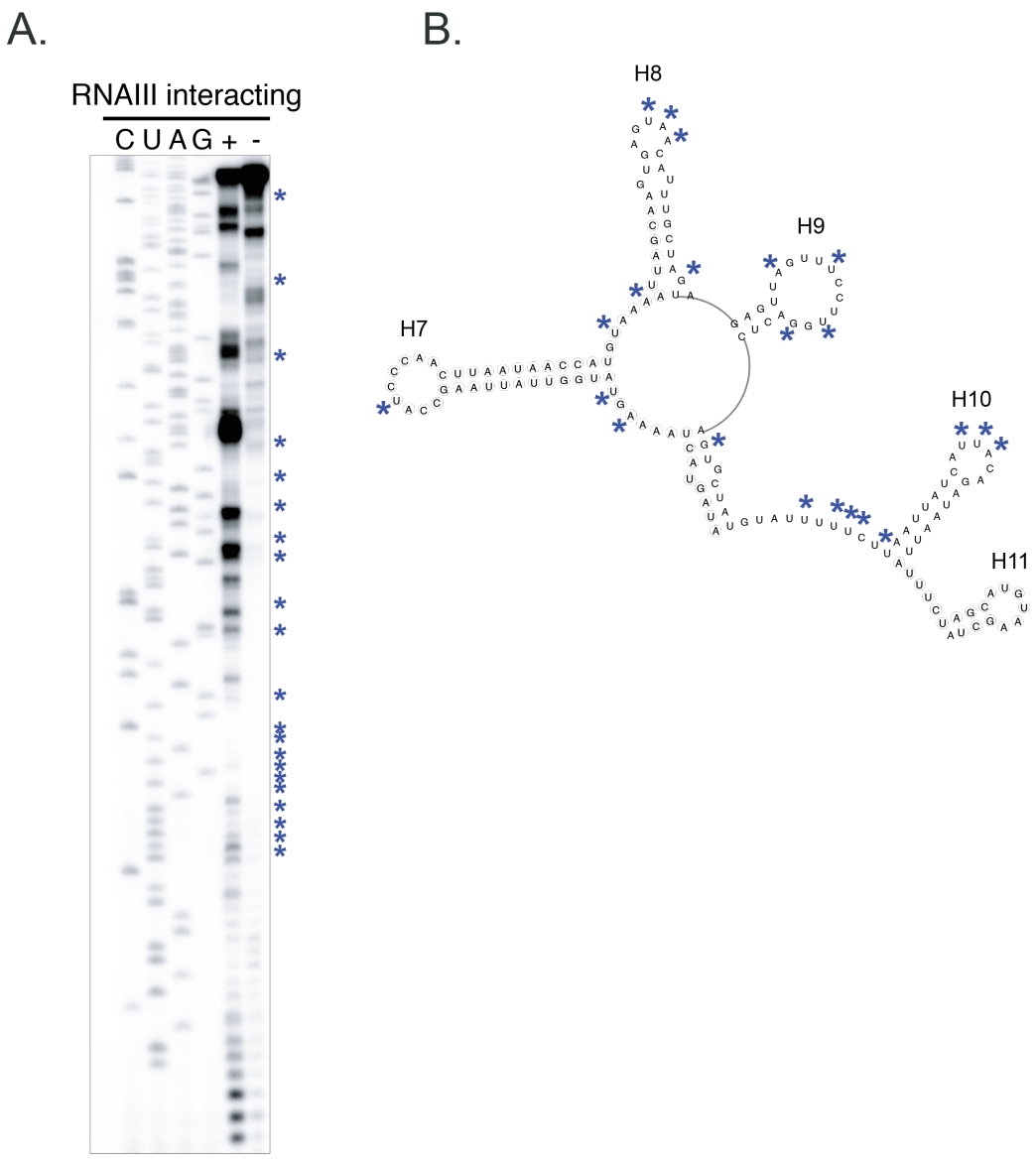

Fig. 5.4: Structure probing of RNAIII interacting. **A.** Primer extension gel after NAI treatment. Asterisks indicate sites of modification. **B.** The structure of RNAIII interacting, taken from the full length structure of RNAIII as produced by (Kolb et al., 2000), with sites of modifications marked with asterisks.

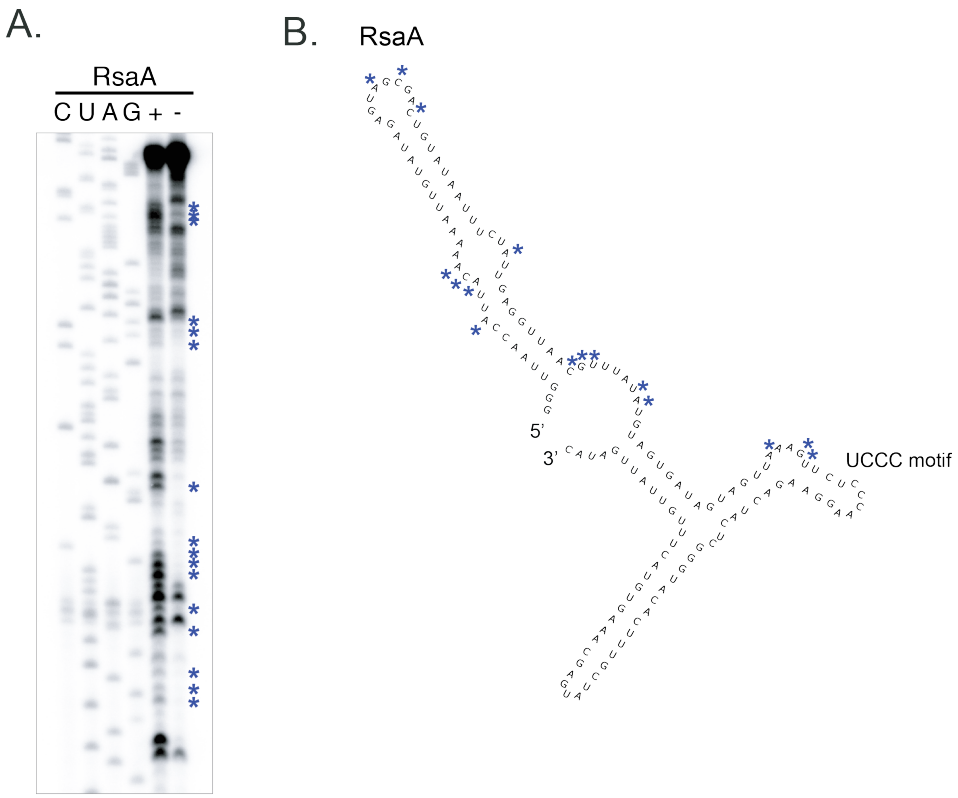

Fig. 5.5: Structure probing of RsaA interacting. **A.** Primer extension gel after NAI treatment. Asterisks indicate sites of modification. **B.** The structure of RsaA, as predicted by RNAcofold, with sites of modifications marked with asterisks.

After confirming that *in vitro* transcribed 'RNAIII interacting' fragment and RsaA fold correctly, EMSAs were performed to examine their proposed interaction. As shown in Figure 5.6, only RNAIII interacting is able to form a complex with RsaA and adding increasing amounts of RNAIII interacting increases the amount of complex produced. Quantifying this (Figure 5.6B) shows that the dynamics of this interaction fit a classical saturation binding curve, with RNAIII able to occupy almost the entire pool of RsaA when in 320-fold molar excess of RsaA, with around a 60-molar excess able to bind half of the pool. Additionally, multiple bands can be observed to form between RNAIII interacting and RsaA, and thus the duplex may exist in multiple, slightly different conformations *in vitro*.

Conversely, RNAIII non-interacting is not able to form any specific interactions with RsaA. Observed in Figure 5.6A (and quantified in Figure 5.6B) is RNAIII non-interacting forming a duplex at a rate comparable to RsaA by itself, i.e. the background oligomerisation of RsaA. Increasing the molar ratio of RNAIII non-interacting does not substantially increase the formation of duplex, and thus it is concluded that the signal observed for RNAIII non-interacting mostly relates to negligible RsaA self-interaction. Thus, the interaction between RsaA and RNAIII interacting is specific.

It is overall concluded that RsaA and RNAIII interact specifically through a region in RNAIII found between helices 8 and 10, and a region containing the UCCC-motif in RsaA. *In silico* predictions lead helix 9 to be the main region responsible for base pairing with RsaA, but future experiments would be necessary to absolutely confirm this.

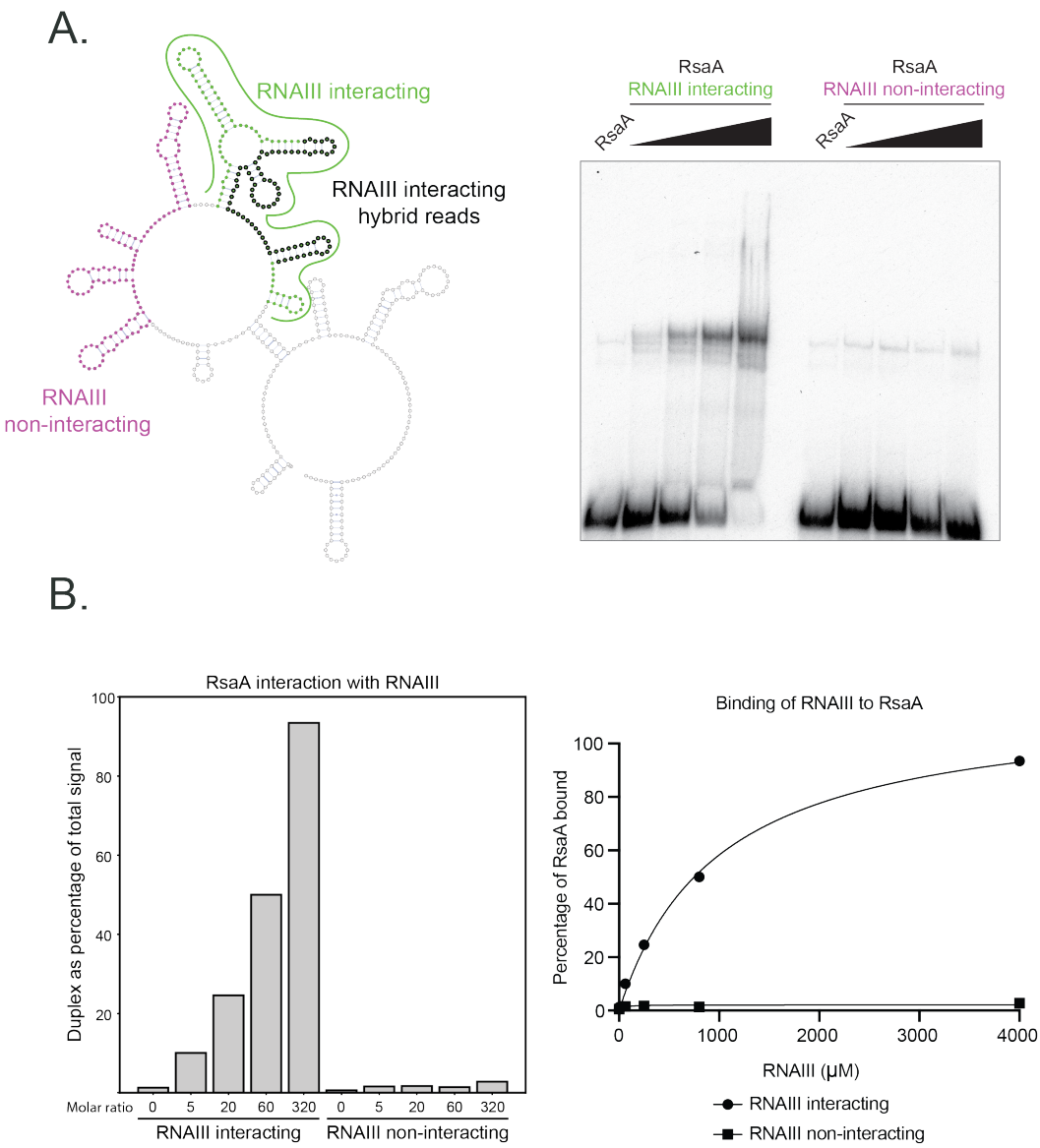

Fig. 5.6: *In vitro* interaction of RsaA and RNAIII. **A.** Left: Structure of RNAIII with the 'RNAIII interacting' and 'RNAIII non-interacting' constructs indicated. Right: EMSA between radiolabelled RsaA and increasing molar concentrations of RNAIII interacting and non-interacting **B.** Quantification of RsaA – RNAIII interaction. Left: histogram showing the formation of the duplex as a percentage of the total RNA. Right: binding curves fit to non-linear regression analysis.

5.5 The interaction between RsaA and RNAIII responds to stress

When I examined the CLASH data, I found that the number of interactions between RsaA and RNAIII was much higher in the unshifted, TSB sample versus the RPMI; 282 versus 57 respectively (Figure 5.7A). Although CLASH is not a strictly quantitative technique, this alluded to the idea that the interaction between RsaA and RNAIII could respond to stress.

In order to explore why the number of interactions was decreasing, the expression of RsaA and RNAIII was measured by qPCR following shift to RPMI medium. Although this had already been obtained in the form of RNAseq (Section 4.2), it was noted that the data for these two sRNAs was noisy and so qPCRs were predicted to give a more reliable insight into their dynamics.

Figure 5.7B shows that when the cells are shifted back into TSB as a negative control, their expression slowly increases over the entire time-course, although noise is observed for the final 30-minute sample. However, when the cells are shifted to RPMI, the expression of RNAIII immediately decreases; after just 5 minutes, its expression is decreased by ~75%. This downregulation continues to 10 minutes where it is downregulated ~85% before starting to slowly recover. Comparing the expression of RNAIII in RPMI to TSB reveals that the downregulation at 5 and 10 minutes is statistically significant, although the final, 30-minute timepoint is not due to the large degree of error for the control TSB shift.

Conversely, the expression of RsaA increases over time. After 5 minutes, RsaA is upregulated ~1.5-fold as compared to the reference sample, and by 10 minutes this has increased to ~2-fold. For comparison, the TSB shifted control only increases to ~1.1-fold, although the difference between the two is not quite statistically significant ($p = 0.09$) due to the degree of error.

Overall, the mirror image expression pattern of RsaA and RNAIII helps to explain why fewer interactions are found in RPMI versus TSB. Based on these data I hypothesize that after the shift to RPMI, RNAIII is immediately downregulated, presumably in part through base-pairing interactions with RsaA, and only a small fraction of it remains after just 5 minutes. Considering that CLASH was performed on cells after 15 minutes of RPMI incubation, it is predicted that only around ~20% of the original pool of RNAIII remained, and thus there is less RNAIII for RsaA to interact with.

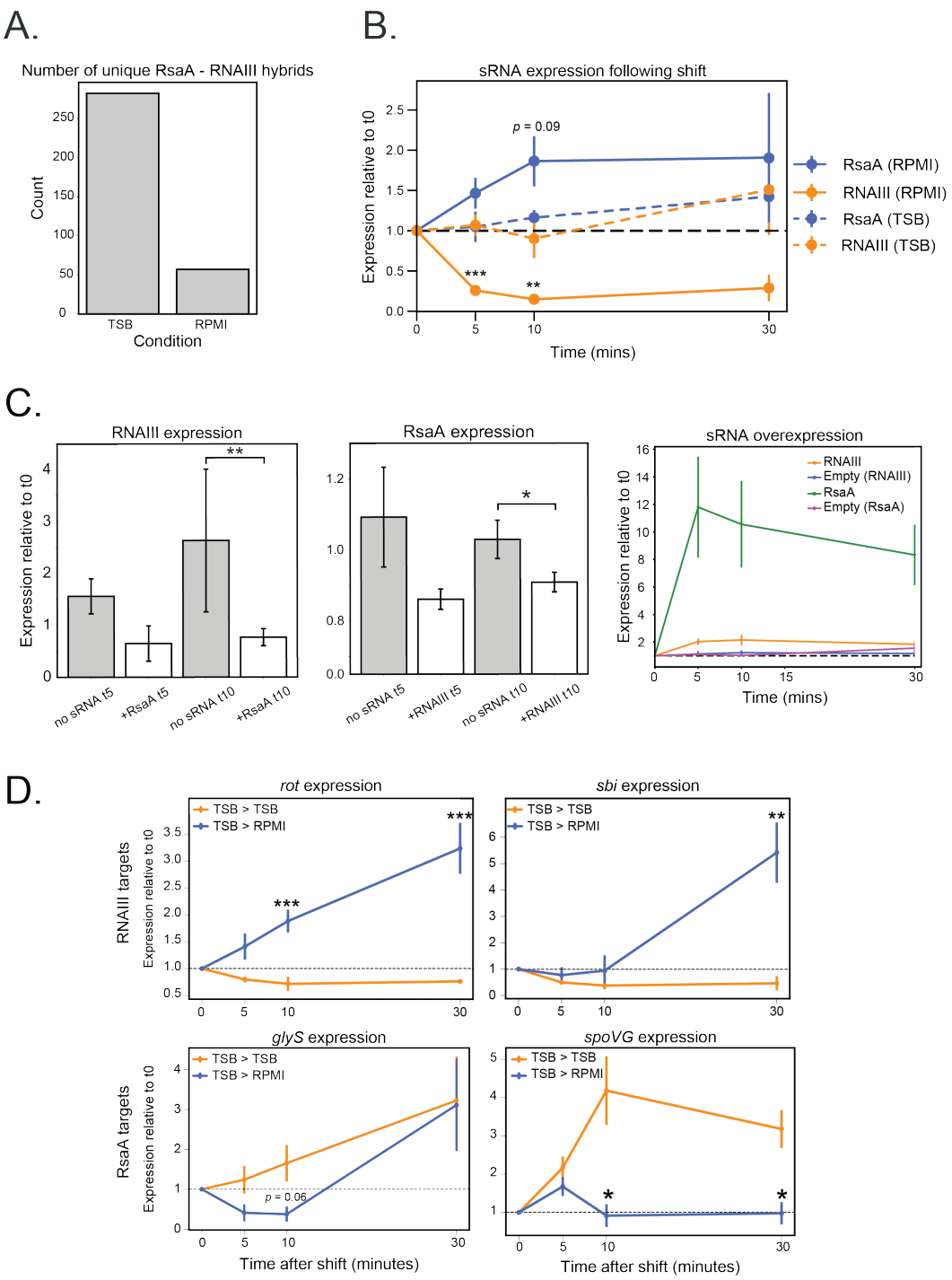

Fig. 5.7: *In vivo* dynamics of RsaA and RNAIII **A.** Number of hybrids between RsaA and RNAIII in TSB and RPMI. **B.** Expression of RsaA (blue) and RNAIII (orange) after shift to TSB (dashed line) and RPMI (solid line) as measured by qPCR. The statistics are generated from comparing the expression in RPMI at that particular timepoint to TSB. **C.** Expression of RsaA and RNAIII after overexpression of the other, as measured by qPCR. Far right panel shows the efficiency of overexpression **D.** Expression of the targets of RsaA and RNA after TSB (orange) and RPMI (blue) shift. In all panels, * = 0.05; ** = 0.01; *** = < 0.01 in Student's unpaired t test, and error bars relate to standard error of the mean.

5.6 Overexpression of RsaA downregulates RNAIII and *vice versa*

The fact that this interaction was discovered using RNase III as bait suggests that RsaA binds to RNAIII and recruits this double-strand-targeting RNase in order to degrade the duplex. Therefore, an obvious question is whether over-expression of RNAIII results in down-regulation of RsaA and *vice versa*. To test this, I overexpressed both sRNAs from a multi-copy plasmid, pRMC2 (Corrigan and Foster, 2009). The effect of over-expression was assessed after 5 and 10 minutes after inducing sRNA expression as I wanted to avoid any secondary effects. Moreover, I observed that USA300 shuts down the expression of these sRNAs within 30 minutes of over-expression (data not shown), suggesting that the expression of these sRNAs is under tight control. While RsaA could be highly over-expressed from a plasmid (~12-fold after just 5 minutes; Figure 5.7C far right panel), I was only able to increase RNAIII cellular levels by ~2-fold. This is likely explained by the fact that these experiments were carried out at OD 3, when the *agr* pathway is active and so RNAIII is already a major, highly expressed transcript within the cell. Thus, the plasmid-mediated induction does not have such a strong effect on the relative amount of RNAIII within the cell. A possible solution to this problem would be to repeat the over-expression experiment in an RNAIII deletion strain, which is currently being made.

When examining the expression of RNAIII after induction of RsaA, I observed a modest, but reproducible and statistically significant, decrease in the levels of RNAIII. In the empty control plasmid, the expression of RNAIII increases with time. Already after 5 minutes of induction, RNAIII has increased ~1.5-fold in the empty plasmid control. Although a large degree of error is observed, by 10 minutes this has increased to ~2.5-fold (Figure 5.7C). Such a strong increase in RNAIII expression after just 10 minutes once reaching OD_{600} 3 is not observed in the TSB to TSB shifts (Figure 5.7B) and thus it may be that the *agr* pathway responds to anhydrotetracycline which is used to induce expression from the vector.

However, inducing RsaA expression leads to a small decrease in the amount of RNAIII within the cell and prevents its future induction. After inducing RsaA, around 35% of the original pool of RNAIII is destroyed after 5 minutes. RNAIII starts to recover after 10 minutes, but only to ~80% of the original, and this is statistically significant in comparison to the empty plasmid control (Figure 5.7C). Thus, overexpression of RsaA is able to induce destruction of RNAIII and prevent its accumulation in post-exponential phase.

The RNA steady state levels of RsaA was then examined after induction of RNAIII. RsaA exhibited no significant changes in expression after inducing the empty plasmid control (~1.3-fold and ~1.1-fold for 5 and 10 minutes respectively). However, when RNAIII was induced, a ~15% and 10% reduction in RsaA levels were observed respectively after 5 and 10 minutes, and this was statistically significant for the 10-minute time point. Thus, although I was unable to highly over-express RNAIII in this strain, the ~2-fold increase in RNAIII levels were sufficient to cause a small but significant reduction in RsaA levels.

In conclusion, overexpression of RsaA and RNAIII is able to cause a statistically significant decrease in RNA levels of the other. This data is consistent with what I observed in the RPMI shift, when RsaA is induced and then RNAIII is subsequently downregulated. Thus, this supports my hypothesis that RsaA is able to induce destruction of RNAIII through recruitment of RNase III.

5.7 The induction of RsaA and destruction of RNAIII following RPMI shift has effects on downstream mRNAs

Once a relationship and understanding of the dynamic between RsaA and RNAIII had been established, I next wanted to determine what are the functional consequences of such an interaction. My assumption was that if RsaA levels increased in RPMI, the levels of the mRNA targets that are negatively regulated by RsaA should decrease. *Vice versa*, I predicted that the levels of mRNAs negatively regulated by RNAIII would increase due to RNAIII's destruction. To test this hypothesis, qPCR on the targets of RNAIII and RsaA following the shift to RPMI and the TSB control samples was performed.

RNAIII is known to regulate many mRNA targets within the cell, as described in Section 1.7. Of particular importance is the transcription factor *rot* and the immune evasive membrane protein *sbi*. RNAIII is known to repress the translation and induce destruction of both of these mRNAs and so it was predicted that their levels would increase upon RNAIII downregulation. As seen in Figure 5.7D, the expression of both *sbi* and *rot* is upregulated following the shift to RPMI medium (and subsequent destruction of RNAIII). However, the dynamics are different, as *rot* is immediately and continuously upregulated throughout the whole time-course while *sbi* is only induced at the latest 30-minute timepoint. This behaviour is in opposition to the TSB control, where both transcripts both decrease continuously over time.

RsaA is has been found to regulate the translational efficiency of several targets *in vitro*, but *in vivo* data is lacking. Deletion of RsaA results in increased levels of *glyS* which encodes for a tRNA synthetase, and *spoVG*, which encodes a sporulation factor involved in antibiotic resistance (Romilly et al., 2014). As deletion of RsaA induces their upregulation, their interaction with RsaA is predicted to be a negative one. Consistent with this idea, *glyS* is downregulated immediately after the shift to RPMI (Figure 5.7D), and continues to be so during the first 10 minutes ($p = 0.06$), and thereafter returns to a level very similar to the control. As such, I predict that RsaA is able to induce *glyS* destruction and this occurs when RsaA is upregulated in RPMI. With regards to *spoVG*, this transcript is increasing over time in the TSB control but is unable to be upregulated in the RPMI shift and instead remains at basal levels. As such, it may be that RsaA is able to prevent *spoVG* induction.

I cannot rule out the possibility that the observed effect was purely due to transcriptional changes but the fact that there is a negative correlation between the sRNA levels and their respective targets suggests that sRNA-mediated degradation at least contributes to the changes in gene expression. In fact, my current working hypothesis is that a combination of both changed transcription of mRNAs combined with a changing stability of the sRNAs that target these messengers has a profound effect on the mRNA induction kinetics and results in much faster responses (also see Discussion). To test this hypothesis, the lab has recently made a RsaA-null strain and I will repeat the TSB-RPMI shift experiment with this mutant and quantify the changes in RNAIII levels and its targets. The hope is that this experiment will provide additional evidence to support my model.

In conclusion, I propose that RsaA-mediated destruction of RNAIII is able to cause downstream effects on the targets of RNAIII. These effects are observed already after 5 minutes for rot and 30 minutes for sbi, and thus may form part of the early-stage adaptation process that occurs when *S. aureus* is shifted into RPMI medium, and ergo the host blood stream. Additionally, upregulated of RsaA appears to induce immediate destruction of a tested tRNA synthetase, glyS, and may also play a role in shaping the expression of spoVG. Thus, induced RsaA may downregulate its own mRNA targets, RNAIII, and also act as an indirect activator of RNAIII targets.

5.8 Transcription and degradation help to shape the dynamics of RNAIII, RsaA and their target mRNAs

As mentioned above, I can not rule out the possibility that expression of RNAIII during the RPMI shift is also, at least in part, shaped by changes in RNA polymerase transcription. In theory, RNAIII may have a very short half-life and the polymerase could immediately stop transcribing it in response to RPMI, resulting in its degradation. This model would also explain its expression profile in RPMI. Thus, to test this I would need measure transcription rates in *S. aureus*.In order to measure nascent transcription, I performed CRAC on RNA polymerase before and at several timepoints after the shift to RPMI. The assumption was that the bulk of the RNA recovered from RNA polymerase CRAC experiments would be RNA that is actively being transcribed. Thus, this allowed a direct read out of what RNAs in the cell are actively producing or repressed during the stress response. To facilitate the CRAC experiments, Pedro Arede Rei produced a USA300 strain in which the RpoC subunit was fused to an HTF tag. Pedro then discovered that RpoC could indeed be efficiently cross-linked to RNA and this prompted me to perform these analyses. It also reasoned that most of the RNA recovered from RNase III CLASH experiment was RNA that was in the process of being degraded. Hence, by comparing the RNA polymerase CRAC data with the RNase III CLASH data, I was able to identify transcripts that were not only repressed in RPMI (transcriptional response) but also likely degraded by RNase III (post-transcriptional response). It is important to note that the data for RNase III only includes a single timepoint, 15 minutes after the RPMI shift.

Following the shift to RPMI, the polymerase binding to RNAIII decreased rapidly. Before the shift, just under 4,000 transcripts per million (TPM) encode for RNAIII and this falls to just 125 after 5 minutes of RPMI incubation, a reduction of ~97%. This repression is maintained throughout the whole time-course (Figure 5.6). Considering that the expression of RNAIII decreases by ~80% at this point, there is strong agreement between the RNA steady-state levels and the changes in the transcription rate. When examining the RNase III binding to RNAIII, I observed a ~50% decrease in binding 15 minutes after the shift to RPMI. This is interesting as although the expression of RNAIII has decreased by around 80% after 15 minutes of RPMI incubation, RNase III binding to RNAIII is still surprisingly high. This implies that down-regulation of RNAIII is controlled both at the transcriptional and post-transcriptional level.

To my surprise, the RpoC CRAC data suggested that transcription of RsaA actually decreased after the shift; from ~6,000 TPM before the shift to ~3,000 after 15

minutes (Figure 5.8). In contrast, the binding of RNase III to RsaA increased ~3-fold after 15 minutes in RPMI. Both of these behaviours are in opposition to the upregulation of RsaA observed after the shift to RPMI (Section 5.6). A plausible explanation for this observation is that although RsaA transcription decreases, the half-life of RsaA levels actually increases during the stress response. Increased RsaA steady state levels is then expected to recruit more RNase III, which may explain the increase in RsaA signal in the RNase III RPMI CLASH data. Alternative explanations are provided in the Discussion section.

In Section 5.6 I showed that several RNAIII targets were increased after the shift to RPMI. Most notably, *rot* increased immediately and continuously, while *sbi* increased after a short delay. Examining the polymerase binding to *rot* reveals that its transcription is actually decreased after the shift. The RNase binding data is noisy to the point where no meaningful conclusion can be made, but it is possible that there is little change in its binding after the shift. As RNAIII is downregulated after the shift, *rot* may now be freed from its RNAIII-mediated degradation and so accumulates despite it being transcribed at a lessened rate.

Regarding *sbi*, its transcription increases continuously throughout the timecourse, from 33 TPM to 95 by 30 minutes; a ~3-fold increase. Although the stoichiometry is not identical, this matches the expression pattern, where a ~5-fold increase is observed. Thus, the expression pattern of *sbi* may be shaped by both transcription as well as relief from RNAIII-mediated repression. RNase III binding to *sbi* also increases, which could be expected since more *sbi* is being produced.

Less detail can be obtained regarding the targets of RsaA. The RpoC data for *glyS* is too noisy for meaningful conclusions to be made regarding the first 15 minutes after the shift. However, overall it is likely that there is no change during this time, and so the initial decrease in *glyS* may be due to RsaA initiating its destruction. There appears to be a decrease in transcription moving from 15 to 30 minutes, which does not match the RNA steady-state levels (increasing strongly across these timepoints). Although the RNase binding data is noisy, a decrease is observed following the shift to RPMI. Thus, the initial decrease in *glyS* following the shift could be due to RsaA-mediated destruction, which then ends after around 15 minutes, allowing *glyS* to recover. Regarding *spoVG*, the transcription appears to be overall stable for the first 10 minutes of the shift and is then downregulated after 15 minutes and maintained. This fits the expression profile of spoVG, and no changes in RNase binding are observed. Thus, the dynamics of *spoVG* may be shaped entirely by transcription. Clearly some of these experiments need to be repeated, however, the results are encouraging; as RpoC also transcribes rRNA and tRNAs, which represent ~90% of the total cellular RNA, I was unsure whether we would actually be able to detect clear changes in mRNA transcription using RpoC CRAC. However, the results

do indicate that for mRNAs and sRNAs that are reasonably abundant, reliable data can be generated. We are also investigating whether different data normalization steps can improve the results.

In conclusion, it is clear that transcription plays a major role in shaping the expression of RNAIII. As such, its interaction with RsaA may be only one facet of several which control its expression in response to stress. The dynamics of RsaA appear more complicated, which displays decreasing transcription, increased RNase III binding, and yet upregulated steady-state levels. Some conclusions can be made for the targets of RNAIII; the expression of rot may be shaped primarily by the downregulation of RNAIII, while *sbi* undergoes both increased transcription and degradation. Regarding RsaA's targets, *glyS* appears to be shaped through RsaA-mediated recruitment of RNase III, while *spoVG*'s dynamics may be mostly transcriptional. Thus, the interaction of RsaA and RNAIII likely contributes to regulating the steady-state level of RNAIII, and by extension the targets of RNAIII, but this behaviour is also coupled to transcriptional effects.

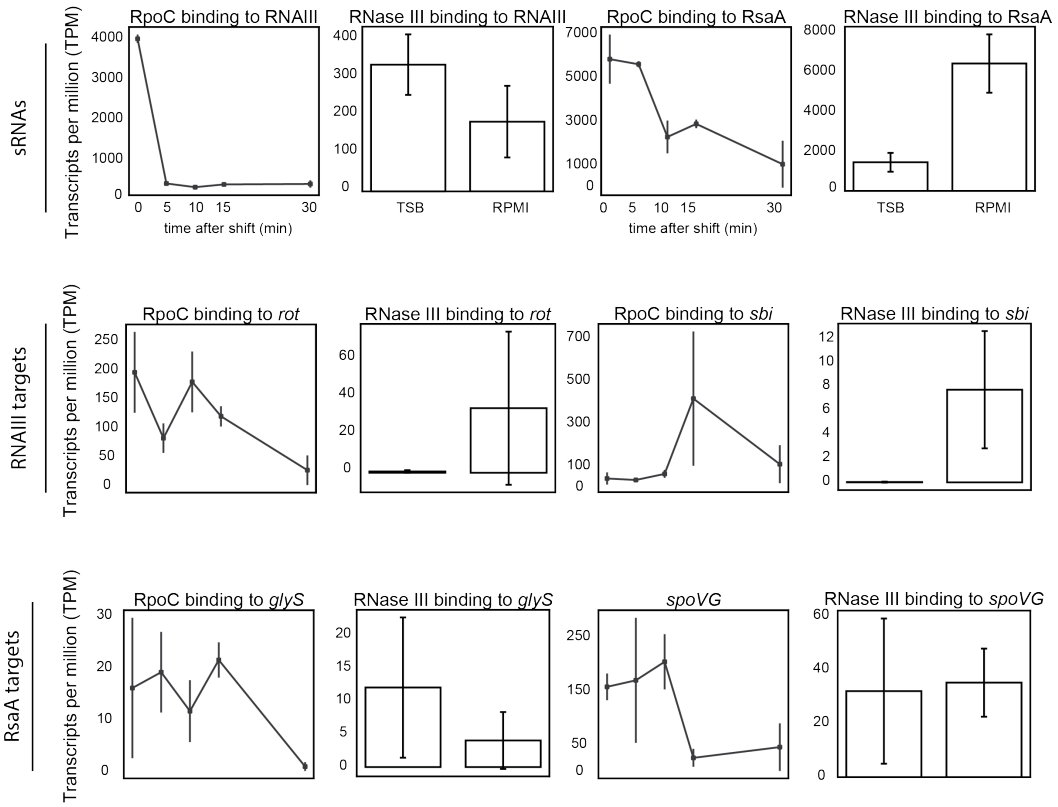

Fig. 5.8: Nascent transcription of RNAIII, RsaA and their targets following the shift to RPMI in USA300. Data obtained through performing CRAC on the polymerase. The polymerase data represents two independent biological replicates, while the RNase III data represents four.

5.9 Identification and validation of novel targets of RsaA and RNAIII

The CLASH data contains many potentially novel targets of RsaA and RNAIII, as discussed in Chapter 4, Section 4.6 and Section 5. Further examination of this reveals several potential targets of RsaA that agree with its already established role in regulating cell membrane composition. As discussed in Section 5.2, RsaA is already known to regulate several members of the Ssa family, which are involved in peptidoglycan metabolism. Additionally, RsaA binds *HG001_01977*, which is hypothesised to be an endopeptidase involved in breaking glycl-glycyl bridges, such as those found in the cell wall of *S. aureus*. The CLASH data for USA300 uncovered 42 unique interactions between RsaA and *SAUSA300_1921*, a peptidoglycan hydrolase of unknown function. Additionally, RsaA was found to bind *qacAB*, a peripheral membrane protein that acts as a wide-spectrum antiseptic transporter involved in exporting quaternary ammonium salts (e.g. cetrimide), intercalating dyes and biguanidines (e.g. chlorhexidine) (Jaglic and Cervinkova, 2018). Considering RsaA's established role in regulating the composition of the cell membrane, novel interactions between RsaA and peptidoglycan hydrolyses and transporters are worthy of further study.

Concerning RNAIII, I was particularly interested in its interaction with *esxA*. EsxA is known be involved in promoting host cell survival. Given RNAIII's known role in virulence, and it even encoding its own haemolysin, it was interesting to see an interaction between two behaviours which are seemingly in opposition with each other (Balaban and Novick, 1995; Sundaramoorthy et al., 2008). Thus, this interaction was selected for further study.

Of additional interest is the interaction between RNAIII and *saeR*, as discussed in Chapter 4, Section 5, but this interaction as only uncovered in a CLASH experiment performed near the conclusion of this work. Thus, time did not permit further study.

Thus, overall I chose to further elucidate the interactions between RsaA and *qacAB*, and RNAIII and *esxA*. Although these interactions were supported by a relatively low number of chimeras, folding of the chimeras suggested that the sRNA base-paired at or near the SD sequence, suggesting that these could be *bona fide* interactions. The interaction between RsaA and the 5' UTR of *qacAB* is of high interest as it may help us to improving the understanding antibiotic and antimicrobial resistance.

More detailed examination of the chimeric reads revealed that the start codons were recovered for each of the mRNAs. The *qacAB* fragment that was fused to RsaA contained the 'AU' region that was predicted to interact with RsaA when folded *in silico* with RNAcofold. Additionally, the interaction appears to start at the UCCC motif of RsaA (Figure 5.9A). As the first two nucleotides of the start codon were identified in the hybrid, this interaction was predicted to be a canonical example of sRNA-mediated translational repression. Folding of the RNAIII-*esxA* chimeras implied that helix 9 of RNAIII base-pairs with the *esxA* coding region. The start codon, however, seems to be sequestered in an intra-molecular loop formed in the *esxA* mRNA. Thus, how RNAIII would regulate *esxA* expression was not immediately obvious (Figure 5.9A).

In order to deduce the role of RsaA and RNAIII on the translational efficiency of *qacAB* and *esxA* respectively, I decided to use an *in vivo* GFP-reporter assay that was recently published by the Felden group (Ivain et al., 2017). Although it is an artificial system, it can provide a clear indication on if the interaction identified by CLASH is also functional *in vivo*. The reasons for choosing this system was because of the enormous challenges we faced when integrating DNA fragments in the *S. aureus* genome. Moreover, similar reporter assays are routinely used in the Gram-negative field (Corcoran et al., 2012). Briefly, the regions identified in CLASH were fused in-frame to GFP and the produced fluorescence examined through FACS in the absence or presence of sRNA overexpression (Figure 5.9B). Here, it was found that overexpression of RsaA strongly reduces the fluorescence produced by the *qacAB-GFP* reporter around 10-fold. Thus, I conclude that RsaA represses the translation of *qacAB in vivo*. Interestingly, upon over-expression of RNAIII, the fluorescence of the EsxA-GFP reporter increased ~2-fold. Thus, RNAIII is predicted to stimulate the translation of *esxA in vivo*. The next goal is to use various biochemical approaches, including RNA structure probing, to dissect the mechanism of how RNAIII enhances EsxA translation.

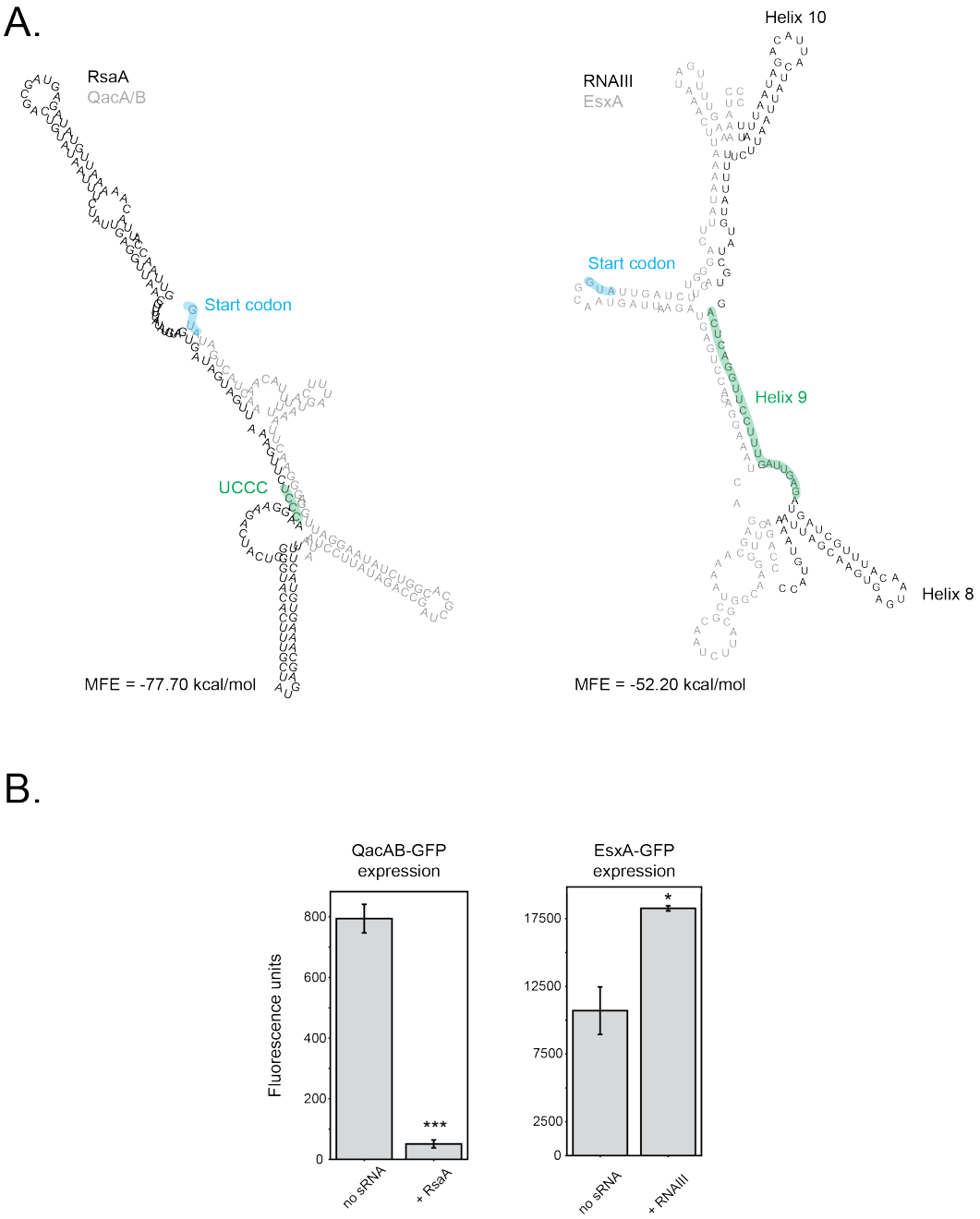

Fig. 5.9: Validation of novel targets of RsaA and RNAIII. **A.** Predicted secondary structures of RsaA and *qacAB*, and RNAIII and *esxA*. The reads mapping to each RNA in the CLASH data were extracted and folded with each other *in silico*. The sRNA is shown in black and the mRNA in grey. The start codon is highlighted in blue. **B.** FACS sorting showing the effect that overexpression of RsaA and RNAIII has on *qacAB* and *esxA* respectively. * = 0.05; *** = < 0.01 in Student's unpaired t test, error bars relate to standard deviation.

Other predicted RsaA-sRNA interactions

In general, the USA300 CLASH datasets contained many more interactions. In these datasets we uncovered several sRNA fragments fused to RsaA besides RNAIII. RsaA was found interacting with srn_2860 and RsaX28, and both of these interactions were calculated as being statistically significant, even though only 2 and 14 chimeras supported these interactions, respectively.

Extracting the CLASH reads and folding them in silico reveals that neither interaction is predicted to cover the UCCC motif of RsaA (Figure 5.10). Regarding srn_2860, a relatively long region of interaction is observed but this frequently contains unpaired nucleotides. With regards to RsaX28, although 14 unique hybrids were identified, only short reads were recovered. Thus, 50 nucleotides flanking the genomic region of the read were added. Here, RsaX28 forms a perfect, 6-nucleotide pairing at the 3' end of RsaA and another 10-nucleotide pairing with a single nucleotide bulge in the middle of RsaA.

When examining the expression of these sRNAs after 15 minutes of RPMI shift, neither are observed to increase significantly. Considering that RsaA is upregulated after the shift to RPMI, these interactions may represent true sponging interactions, as opposed to induced destruction. However, from the CLASH data it is impossible to predict which sRNA regulates which.

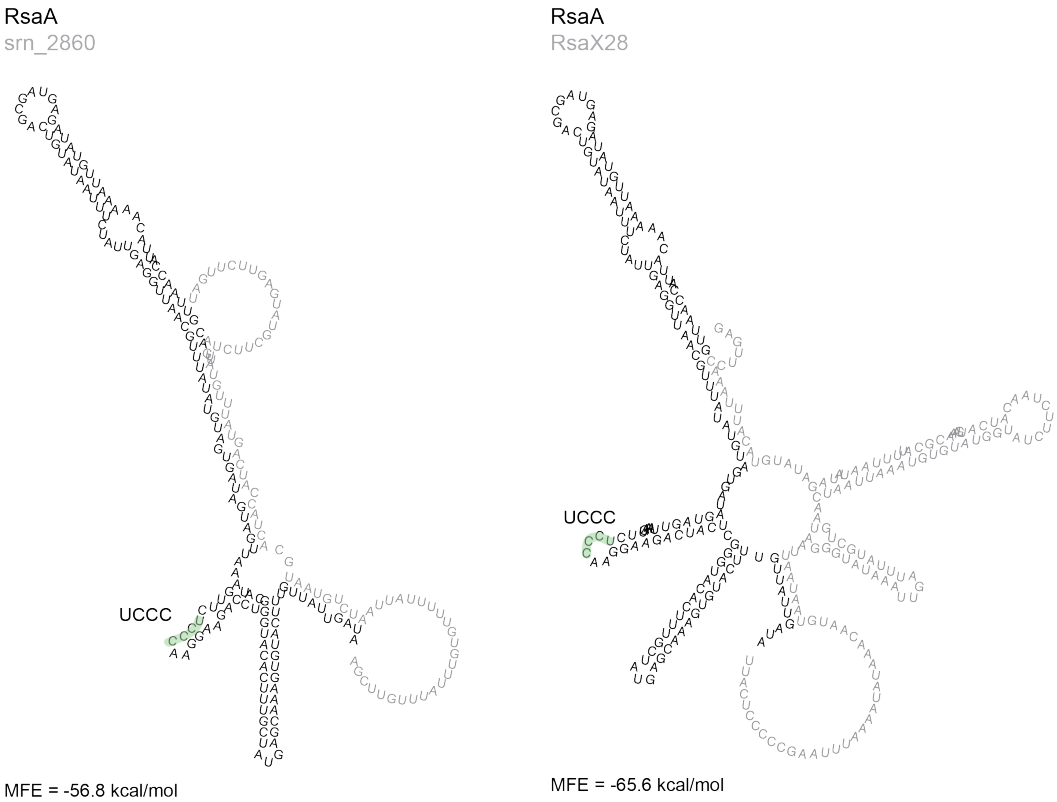

Fig. 5.10: Validation of novel targets of RsaA and RNAIII. Interaction between RsaA, and srn_2860 and RsaX28. RsaA is shown in black, and the interacting partner in grey. The UCCC seed motif of RsaA is highlighted in green. For RsaX28, 50 nucleotides from either side the interacting read was added.

5.10 Discussion

The work described in this chapter firstly illustrate CLASH's ability to recover known interactions of a well-understood sRNA, RsaA. This gives credence to the novel targets that CLASH uncovers, of which several have been chosen for further validation. The major, novel discovery of the work described here is that RsaA is able to bind RNAIII, an sRNA intimately involved in regulating *S. aureus*' transition to virulent behaviour and that this interaction may have downstream consequences for the mRNA targets of these sRNAs.

Other CLASH validates known targets of RsaA

Currently, detailed analyses describe how RsaA regulates the translation of *mgrA*, a pro-virulence transcription factor (Romilly et al., 2014); *flr*, an immunomodulatory protein involved in protection against complement and antibodies (Stemerding et al., 2013); *HG001_01977*, an SH3 domain-containing protein hypothesised to act as an endopeptidase involved in membrane modelling (Tomasini et al., 2017); and members of the *Ssa* family, which are involved in peptidoglycan synthesis (Dubrac et al., 2008; Tomasini et al., 2017). All of these interactions, except that with *flr*, were recovered in CLASH, giving credence to the CLASH methodology.

RsaA is known to bind the SD sequence of *mgrA* in order to block ribosome binding and repress translation (Romilly et al., 2014). This interaction was perfectly and abundantly recovered with CLASH. However, RsaA is also thought to bind the 3' end of *mgrA* and the authors were "tempted to speculate" that this interaction site may recruit RNase III. This idea is not supported in the data presented here; I did not detect any RsaA chimeras that had fragments of 3' end of *mgrA* in the RNase III CLASH data. However, a negative result in CLASH does not mean it does not exist and it is certainly possible that other dsRBPs mediate or recognize this interaction. Additionally, as different regions of RsaA bind to the 3' and 5' end of *mgrA* respectively, it is possible that a single RsaA molecule binds both ends of *mgrA*, forming a loop structure, and the double-stranded region at the 5' end recruits RNase III. Despite this, good agreement is found between the CLASH data and the *in vitro* validation performed by Romilly et al, 2014.

There are a number of potential reasons as to why the interaction between RsaA and *flr* was not identified. In their paper, Tomasini et al, 2017, expressed RsaA from a plasmid under the *agr* P3 promoter, such that RsaA was expressed abundantly only at stationary phase. This led them to performing their experiments at an OD_{600} of 5, while the CLASH experiments performed here were at an OD_{600} of 3. It may be that RsaA only regulates *flr* at certain ODs or that *flr* may only be expressed in the

stationary phase. Additionally, the expression of RsaA is known to increase with OD, and so the cell will contain more RsaA at OD 5 versus OD 3. Concomitantly, Tomasini et al found that their RsaA expressed from a plasmid was 1.7-fold more abundant than the wild type. Thus, the amount of RsaA in the cell in the MAPS study will be far higher than that for the experiments carried out here. *flr* may only be a weak target of RsaA and so may have only been uncovered due to the potential excess of RsaA.

A mutant RsaA-null strain was found to have increased levels of SpoVG, a protein involved in capsule production and antibiotic resistance and also glycyl-tRNA synthetase, *glyS*. However, no hybrids were identified between RsaA and these targets. Similarly to *flr*, these interactions may not take place in the conditions here.

In conclusion, the CLASH methodology has proven itself reliable for identifying targets of RsaA due to its recovery of several already known interacting partners.

RsaA interacts with RNAIII and potentially regulates its stability

The CLASH data also suggests new interacting partners of RsaA. The most interesting of these is with RNAIII, an sRNA known to be involved in driving the cell into a virulent state.

The primary evidence supporting an interaction between RsaA and RNAIII is a considerable number of unique interactions found in CLASH and the fact that this interaction was found both in JKD6008 and USA300. Reads mapping to helices 8, 9 and 10 of RNAIII were uncovered in CLASH, with helix 9 being predicted to directly interact with helices 2 and 3 of RsaA, covering the UCCC motif. The fact that this seed region of RsaA is recovered in the interaction is also good evidence of the interaction being genuine.

Additionally, the interaction was recapitulated *in vitro*, where *in vitro* transcribed RNAs were shown to interact with each other. Importantly, this interaction was proven to be specific, as RsaA was only able to interact with a selected portion of RNAIII, chosen due to its presence in the CLASH data. Conversely, RsaA was not able to interact with a second, control region of RNAIII. Thus, the interaction between RsaA and RNAIII is specific and occurs only with a particular region of RNAIII.

Interestingly, there were 5-fold more interactions between RsaA and RNAIII in TSB medium versus RPMI. This leads me to propose that this interaction responds to nutrient stress. Delving further into the dynamics of these two sRNAs revealed that after the shift to RPMI medium, RNAIII is immediately downregulated and the pool is destroyed. On the basis that RPMI has been shown to mimic human

blood plasma (Mäder et al., 2016), this finding matches already published data where transcriptomic analysis of *S. aureus* that had been incubated in human blood revealed that the *agr* locus was completely shut down, even at high cell densities (James et al., 2013; Malachowa et al., 2011). Part of the reason for this is thought to be due to host apolipoprotein B, which is able to sequester the agr signalling molecule, AIP (Peterson et al., 2008). However, as there is no apolipoprotein B in RPMI medium, this suggests that there is another reason for the decrease in RNAIII expression. One explanation is that the *agr* pathway is turned off in the face of stress and this connection has already been made with antibiotic stress, both for vancomycin and methicillin resistance (Howden et al., 2008; Moise et al., 2010; Pozzi et al., 2012; Sakoulas et al., 2002). Thus, the stress encountered in RPMI medium, such as iron depletion, may induce *agr* suppression.

With regards to the data discussed here, the key questions are therefore; What role does RsaA have in this downregulation of RNAIII? What is the functional outcome of this interaction?

Overexpression of RsaA was found to cause a modest, but reproducible, decrease in the amount of RNAIII within the cell. This over-expression also prevented RNAIII from being induced as was seen in the control. This finding fits well with the expression patterns of these two sRNAs after shifting to RPMI, when RsaA is upregulated and RNAIII is downregulated. Overall, this leads to the idea that RsaA is able to induce the degradation of RNAIII and prevent it from accumulating within the cell.

However, it is important to note the stoichiometry. Overexpression of RsaA leads to a 10-fold increase in the amount of RsaA and a 23% decrease in the amount of RNAIII after 10 minutes. On the other hand, the shift to RPMI induces just under a 2-fold increase in RsaA, yet an 85% decrease in RNAIII levels. Thus, it is likely that the behaviour of RNAIII can not only be explained through the action of RsaA.

It is clear that transcription also plays a role in shaping the expression of RNAIII. Following the shift to RPMI, the polymerase immediately stops transcribing RNAIII and this remains throughout the time course, which matches the expression profile. Thus, it may be that the bulk degradation of RNAIII immediately after the shift to RPMI is due to RsaA-mediated destruction, but the long-term dynamics are shaped by transcription. Despite the significant decrease in transcription and drop RNAIII steady-state levels, a relatively large number of interactions between RsaA and RNAIII are still able to be captured using RNase III after 15 minutes of RPMI exposure. It may be that any residual RNAIII or newly transcribed RNAIII is still preferentially destroyed by the cell under RPMI stress. This could be a result of its interaction with RsaA.

For most mRNAs that I analysed in the RpoC CRAC data, I detected what appeared to be a decrease in polymerase transcription, whereas few mRNAs showed increased transcription. This is a suspicious and so the lab is now investigating the possibility that the observed decrease in many RNA levels could be explained by issues with the normalization of the CRAC sequencing data. One possibility is that during stress the levels of most RNAs strongly decreases and standard normalization methods for library depth (such as TPM normalization) do not perform well on datasets where there are substantial changes in RNA levels. To solve this problem, we are currently testing other normalization methods. From experience with yeast and human cells we learned that rRNAs that often contaminate CRAC datasets can be reliably used for data normalization. Moreover, we are in the process of designing a modified CRAC protocol that allows us to use spike-ins.

The fact that RsaA is known to be involved in attenuating virulence puts it in direct opposition of RNAIII. Deletion of RsaA in *S. aureus* was found to induce a hypervirulent phenotype in a mouse sepsis model, with increased invasiveness observed. RsaA is known to promote dormancy, biofilm formation, and potentially favours chronic infections in vivo (Romilly et al., 2014). In addition, Romilly et al, 2014, postulated that RsaA may act as a suppressor of virulence, noting that RsaA has been found to be actively expressed in commensal *S. aureus* infections, such as those in nasal passages (Song et al., 2012). Their opinion was based on the idea that as *S. aureus* exists primarily as a commensal bacterium, i.e. one that does not cause damage to the host, it must have a means of halting its virulent behaviour in order not to cause excessive damage to the host. Indeed, non-coding RNAs that act to repress virulence have already been identified in *S. typhimurium*, where an antisense sRNA negatively regulates the *mgtC* virulence factor and deletion of the asRNA induces a hypervirulent phenotype (Lee and Groisman, 2010) . However, the data described here suggests that RsaA may be used as a very general means to rapidly shut down virulent behaviour whenever necessary, such as in the stressful conditions encountered in RPMI medium. Here, it may be of benefit to *S. aureus* to stop committing resources to virulent behaviours and instead focus on making the necessary metabolic remodelling in order to survive. Thus, RsaA may play a leading role in this state change through accelerating the downregulation of RNAIII and therefore attenuating virulence until the stress is survived.

However, this model still needs to explain why a large number of RsaA – RNAIII interactions were found in TSB, before the imposition of stress. It may be that RsaA is able to interact with RNAIII under these conditions, but RNase III does not induce any degradation. Alternatively, RsaA may continuously induce RNAIII degradation, but this is offset by very high rates of RNAIII transcription. Both of these models would suggest that the cell is permanently 'primed' to induce degradation of RNAIII in order to shut down virulence in times of stress, allowing rapid adaptation.

A potentially interesting alternative interpretation of the interaction between RsaA and RNAIII is that the interaction is bi-directional. That is, RNAIII would also able to regulate the activity of RsaA. A small piece of evidence currently supports this model. Overexpression of RNAIII was found to induce a small but reproducible and statistically-significant downregulation of RsaA. One reason why this downregulation was so small could be that RNAIII is such an abundant transcript within the cell at the optical density the experiment was performed at, that the degree of overexpression was not strong enough to cause significant effects within the cell. Indeed, plasmid-induced expression of RNAIII only caused a 2.5-fold upregulation of RNAIII by 10 minutes. It would be important to repeat this experiment in an RNAIII knock-out strain.

In the CLASH experiments performed here, cells were grown to late-exponential phase. Here, the *agr* pathway in *S. aureus* has become active, RNAIII is being increasingly transcribed and the cell is committing to a more virulent state. It may be that at OD_{600} 3 in TSB medium, RNAIII is actually binding to RsaA to mediate its inactivation or destruction. Both RsaA and RNAIII regulate the pro-virulence transcription factor *mgrA* in opposite fashions, with RsaA repressing and RNAIII stimulating its translation. Thus, RNAIII could increasingly bind RsaA as RNAIII expression increases with time, in order to halt RsaA's promotion of dormancy and repression of virulence. Thus, when CLASH was performed in TSB, interactions representing RNAIII repressing RsaA were captured. Upon the shift to RPMI, RNAIII is downregulated through the combined action of its decreased transcription and also through RsaA now acting as the regulator and inducing RNAIII destruction. This model also fits the observation that only a small number of RsaA-RNAIII interactions were found in RPMI.

This model fits the idea of a feed forward loop. RNAIII's potential mediation in the destruction of RsaA, an inhibitor of *mgrA* translation, would mean that it encourages the translation of *mgrA* both directly and indirectly (through destruction of an *mgrA* inhibitor). Additionally, if the cell intends to promote dormancy following the shift to RPMI through upregulation of RsaA, then RsaA's promotion of RNAIII destruction would also represent a feed forward loop as RsaA would be concomitantly shutting down an opposing virulence pathway.

The idea that RsaA is connected to the *agr* pathway, and therefore RNAIII, has been suggested before. RsaA is positively regulated by σ^B, and strains that have a fully functional σ^B express lower levels of RNAIII (Bischoff et al., 2004). *mgrA* is also positively controlled by σ^B, and thus RsaA acts in an incoherent feed-forward loop with MgrA. Additionally, MgrA is connected to the *agr* locus; mutations in MgrA reduce the levels of RNAIII, and MgrA also has several targets in common with RNAIII such as *hla* and *sbi* (Ingavale et al., 2005; Luong et al., 2006; Romilly et al.,

2014). Additionally, an entirely *in silico* analysis predicted that RsaA interacts with *agrC* (Subramanian et al., 2018).

RsaA may thus have several links to the *agr* system. However, the data described here is the first instance of a direct, *in vivo* connection between the two.

One crucial experiment in order to further validate this interaction would be to shift an RsaA deletion strain into RPMI medium and monitor the expression of RNAIII. If this model proposed here is correct, then kinetics of RNAIII reduction should change as it will largely be dependent on the sRNAs own half-life and its decrease in transcription.

Now that the USA300 RsaA knock-out strain available, I will also be able to perform cell survival studies to demonstrate the importance of RsaA shutting down RNAIII. For example, wild type and the RsaA deletion mutant could be shifted to RPMI medium and then have their growth measured in a plate reader in order to monitor their response to the stress. One hypothesis would be that only wild type cells are able to adapt and then survive in order to grow, while the RsaA mutant would be unable to turn off virulent gene expression and then fail to adapt to the nutrient stress. This experiment could be combined with staining in order to visualise cell death after the shift.

In order to explore the idea of a bi-directional interaction, the expression of RsaA in an *agr* knockout strain could be compared to wild type. Additionally, RNAIII could be overexpressed in the same strain through a plasmid, or at a low optical density before the *agr* pathway is activated. Both of these would allow a huge induction of RNAIII relative to that within the cell at the time of experimentation, and this may be enough to induce strong effects on RsaA.

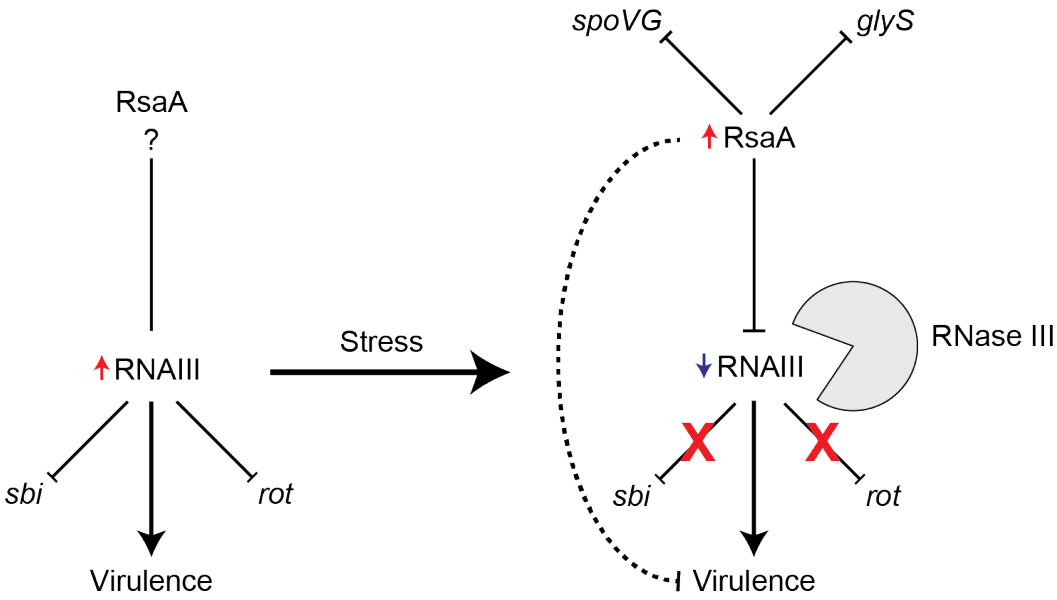

Fig. 5.11: The interaction between RsaA and RNAIII responds to stress. During unstressed conditions, e.g. in TSB, RNAIII and RsaA interact but the consequences of this are unknown. As cells grow and increase in density, the level of RNAIII increases (red arrow) and this drives virulent behaviour. This is in part due to the repression of *sbi* and *rot*. After the imposition of stress, it is theorised that RsaA inhibits RNAIII activity and recruits RNase III, which induces RNAIII destruction. This, in combination with decreased RNAIII transcription, causes RNAIII levels to fall (blue arrow). Additionally, RsaA is upregulated, theorised due to an increase in stability (red arrow). The destruction of RNAIII leads to its target transcripts being relieved of their repression, while targets of RsaA are now subject to regulation.

Novel mRNA targets validated for RsaA and RNAIII

RsaA is already known to be involved in regulating the composition of the cell wall and inhibiting capsule formation (Romilly et al., 2014; Tomasini et al., 2017). I also found RsaA binding to the 5'UTR of *qacAB*, which encodes for a wide-spectrum antiseptic transporter. This was chosen for further validation due to the current interest in understanding antiseptic and antibiotic resistance. *In vivo* GFP reporter assays suggest that RsaA is able to significantly downregulate the translation of *qacAB*. This interaction fits with the proposed, although not validated, interaction between RsaA and *spoVG*. RsaA was proposed to negatively regulate the translation of SpoVG due to its upregulation in an RsaA-null strain (Romilly et al., 2014). SpoVG is known to mediate methicillin and glycopeptide resistance (Eyraud et al., 2014), thus RsaA's inhibition of *qacAB* would fit an idea of it negatively repressing proteins involved in antibiotic resistance.

In order to further validate this interaction, it must be further probed. One means to do would be to generate an RsaA 'seed mutant', where the UCCC motif is mutated and thus the binding should be abrogated, and the regulation of *qacAB-GFP* removed. This could then be further strengthened by creating a compensatory mutation in the *qacAB-GFP* construct such that binding and regulation is restored. Alternatively, *in vitro* methods such as EMSAs, toeprinting assays and structure probing could delve into the binding dynamics.

I found that RNAIII positively regulates the translation of a virulence factor, EsxA. This is a surprising finding since the interaction was recovered with the nuclease RNase III. It is possible that RNase III cleavage of the duplex does not lead to degradation of the *esxA* mRNA and instead this processing event that alters the structure of the 5' UTR. This may then lead to freeing of the start codon from the stem structure. It was previously shown that RNase III processing of σ^B in *E. coli* can stimulate the translation of the mRNA (Freire et al., 2006).

Besides EsxA, RNAIII also enhances the translation of *hla*, a haemolysin implicated in inducing host cell lysis (Menzies and Kourteva, 2000). Furthermore, RNAIII itself also encodes for a small, cytolytic peptide, δ-haemolysin that is also involved in intracellular escape (Balaban and Novick, 1995; Giese et al., 2011). The positive regulation of *esxA*, a protein involved in host cell survival, and *hla*, a protein involved in host cell lysis, are seemingly at odds. It may be that temporal changes, such as in the timing of *esxA* and *hla* expression, dictate which behaviour RNAIII drives. Overall, this potentially puts RNAIII in the unique position of regulating both host cell death and host cell survival (Figure 5.12A).

An alternate hypothesis is that *esxA* and *hla* may be produced and secreted together. Indeed, it has been observed that *Helicobacter pylori* injects an anti-apoptotic protein, CagA, to offset the cytotoxic activity of a main virulence factor, VacA (Oldani et al., 2009). Thus, *hla* may be translated and secreted to e.g. allow escape from a phagolysosome, but *esxA* is also produced concurrently in order to ensure the cell as a whole is kept alive. This would help *S. aureus* to escape a perilous local environment within the cell but maintain itself within safety overall (Figure 5.12B).

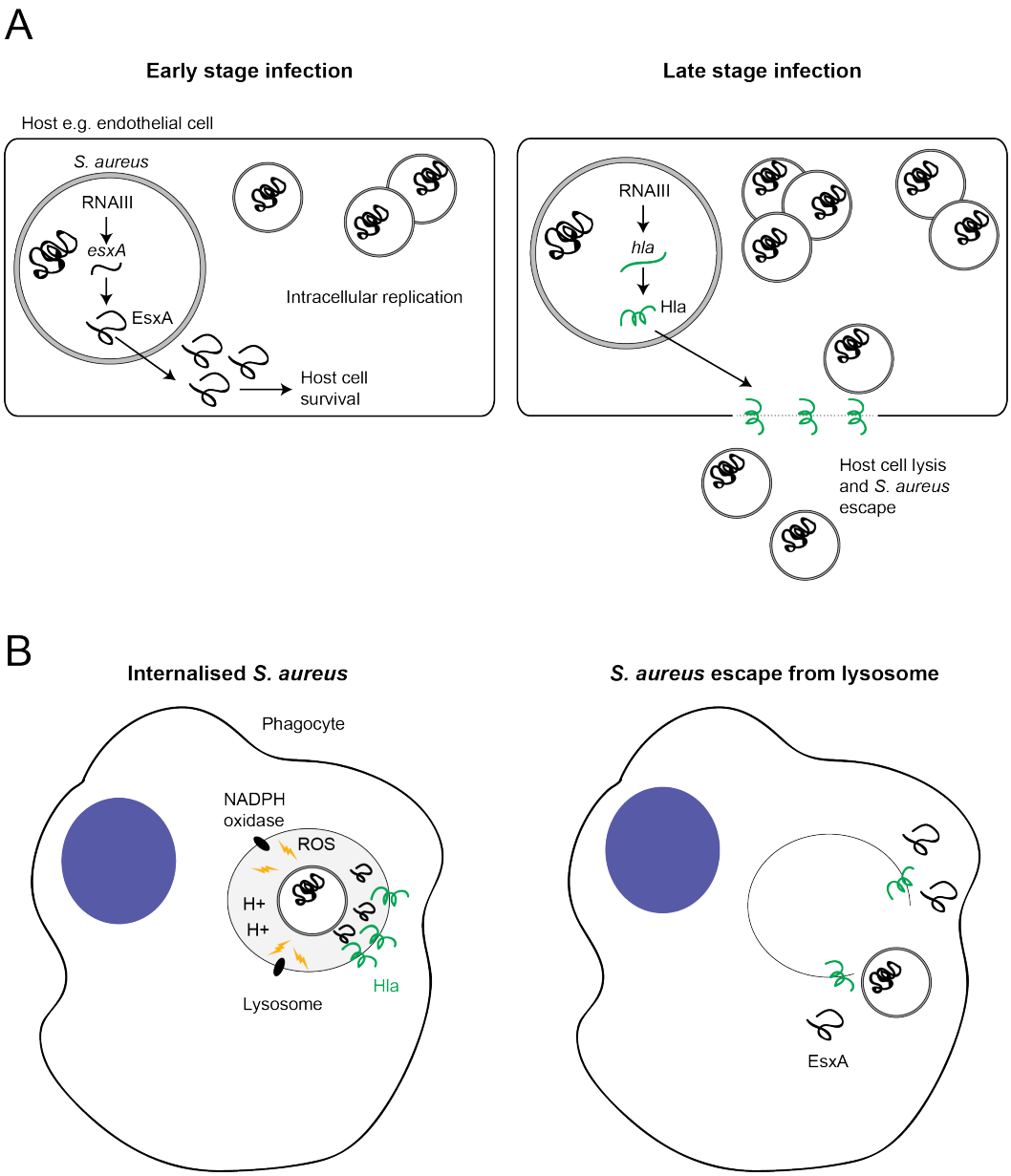

Fig. 5.12: Hypothesised dual role of RNAIII in regulating host cell survival and lysis. **A.** During the early infection stage of a host non-immune cell (e.g. an endothelial cell), RNAIII promotes the translation of *esxA*, which is then secreted into the host cell. EsxA then promotes host cell survival, facilitating *S. aureus* replication in a safe, nutrient-rich environment. After division, RNAIII then switches to promoting *hla* translation, which then integrates into the host cell membrane and induces lysis, facilitating *S. aureus* escape. **B.** Following internalisation by a host phagocyte, *S. aureus* may use *hla* in order to escape from the lysosome. At the same time, EsxA is secreted in order to offset any cytotoxic effects of bursting the lysosome due to its promotion of host cell survival. This potentially allows *S. aureus* to stay within a favourable environment.

RsaA may have other sRNA targets

In addition to RNAIII, RsaA was also found interacting with two other sRNAs; srn_2680 and RsaX28. Both of these interactions were found in USA300 after the RPMI shift.

All that is known about srn_2680 is that it is encoded antisense to MsrR, a membrane-associated transcriptional attenuator (Howden et al., 2013; Rossi et al., 2003). Interestingly, mutants for MsrR express higher levels of RNAIII, and thus MsrR could be a negative regulator of RNAIII (Rossi et al., 2003). Although highly speculative, srn_2680 may act as an antisense, *cis*-encoded sRNA to repress the transcription or translation of MsrR. Upon the shift to RPMI, RsaA may then sponge srn_2680 activity to allow MsrR to be transcribed and translated, and this may then have a role in shutting down RNAIII production.

RsaX28 is a *trans* acting, *bona fide* sRNA that has its own promoter, terminator, and does not overlap any other coding regions (Liu et al., 2018). It has been recorded as displaying different lengths in different *S. aureus* strains but appears to be around 1200 nucleotides long on average. Interestingly, it is known to upregulated after exposure to sub-inhibitory concentrations of antibiotics and is involved in virulence (Horn et al., 2018, p. 42). *S. aureus* RsaX28 mutant strains have a decreased ability to lyse red blood cells be due to an inability to transcribe *hla* (Horn et al., 2018). Thus, RsaX28 acts as a positive regulator of *hla*, but the mechanism behind this remains unexplained.

This chapter has proposed that RsaA plays a role in downregulating virulence by aiding the destruction of RNAIII. Incidentally, RNAIII also stimulates *hla* expression by encouraging its translation. Thus, *hla* is subject to regulation by both RNAIII and RsaX28. Each of these is potentially regulated by RsaA in times of nutrient stress, and so these interactions may represent RsaA attenuating *hla* expression, and virulence as a general concept, from several different angles. This idea could be further explored by confirming the interaction between RsaX28 and RsaA *in vitro*, and examining the expression of *hla* in an RsaA mutant strain after the shift to RPMI.

5.11 Bibliography

Balaban N, Novick RP. 1995. Translation of RNAIII, the *Staphylococcus aureus* agr regulatory RNA molecule, can be activated by a 3-end deletion. FEMS Microbiol Lett 133:155–161. doi:10.1111/j.1574-6968.1995.tb07877.x

Barrass JD, Reid JEA, Huang Y, Hector RD, Sanguinetti G, Beggs JD, Granneman S. 2015. Transcriptome-wide RNA processing kinetics revealed using extremely short 4tU labeling. Genome Biol 16:282. doi:10.1186/s13059-015-0848-1

Bischoff M, Dunman P, Kormanec J, Macapagal D, Murphy E, Mounts W, Berger-Bächi B, Projan S. 2004. Microarray-Based Analysis of the *Staphylococcus aureus* σB Regulon. J Bacteriol 186:4085–4099. doi:10.1128/JB.186.13.4085-4099.2004

Bronesky D, Wu Z, Marzi S, Walter P, Geissmann T, Moreau K, Vandenesch F, Caldelari I, Romby P. 2016. *Staphylococcus aureus* RNAIII and Its Regulon Link Quorum Sensing, Stress Responses, Metabolic Adaptation, and Regulation of Virulence Gene Expression. Annu Rev Microbiol 70:299–316. doi:10.1146/annurev-micro-102215-095708

Corcoran CP, Podkaminski D, Papenfort K, Urban JH, Hinton JCD, Vogel J. 2012. Superfolder GFP reporters validate diverse new mRNA targets of the classic porin regulator, MicF RNA. Mol Microbiol. doi:10.1111/j.1365-2958.2012.08031.x

Corrigan RM, Foster TJ. 2009. An improved tetracycline-inducible expression vector for *Staphylococcus aureus*. Plasmid 61:126–129. doi:10.1016/j.plasmid.2008.10.001

Deigan KE, Li TW, Mathews DH, Weeks KM. 2009. Accurate SHAPE-directed RNA structure determination. Proc Natl Acad Sci U S A 106:97–102. doi:10.1073/pnas.0806929106

Dubrac S, Bisicchia P, Devine KM, Msadek T. 2008. A matter of life and death: cell wall homeostasis and the WalKR (YycGF) essential signal transduction pathway. Mol Microbiol 70:1307–1322. doi:10.1111/j.1365-2958.2008.06483.x

Eyraud A, Tattevin P, Chabelskaya S, Felden B. 2014. A small RNA controls a protein regulator involved in antibiotic resistance in *Staphylococcus aureus*. Nucleic Acids Res 42:4892–4905. doi:10.1093/nar/gku149

Freire P, Amaral JD, Santos JM, Arraiano CM. 2006. Adaptation to carbon starvation: RNase III ensures normal expression levels of bolA1p mRNA and sigma(S). Biochimie 88:341–346. doi:10.1016/j.biochi.2005.09.004

Giese B, Glowinski F, Paprotka K, Dittmann S, Steiner T, Sinha B, Fraunholz MJ. 2011. Expression of -toxin by *Staphylococcus aureus* mediates escape from phago-endosomes of human epithelial and endothelial cells in the presence of β-toxin. Cell Microbiol 13:316–329. doi:10.1111/j.1462-5822.2010.01538.x

Horn J, Klepsch M, Manger M, Wolz C, Rudel T, Fraunholz M. 2018. Long Noncoding RNA SSR42 Controls *Staphylococcus aureus* Alpha-Toxin Transcription in Response to Environmental Stimuli. J Bacteriol 200. doi:10.1128/JB.00252-18

Howden BP, Beaume M, Harrison PF, Hernandez D, Schrenzel J, Seemann T, Francois P, Stinear TP. 2013. Analysis of the Small RNA Transcriptional Response in Multidrug-Resistant *Staphylococcus aureus* after Antimicrobial Exposure. Antimicrob Agents Chemother 57:3864–3874. doi:10.1128/AAC.00263-13

Howden BP, Smith DJ, Mansell A, Johnson PD, Ward PB, Stinear TP, Davies JK. 2008. Different bacterial gene expression patterns and attenuated host immune responses are associated with the evolution of low-level vancomycin resistance during persistent methicillin-resistant *Staphylococcus aureus* bacteraemia. BMC Microbiol 8:39. doi:10.1186/1471-2180-8-39

Ingavale S, Wamel W van, Luong TT, Lee CY, Cheung AL. 2005. Rat/MgrA, a Regulator of Autolysis, Is a Regulator of Virulence Genes in *Staphylococcus aureus*. Infect Immun 73:1423–1431. doi:10.1128/IAI.73.3.1423-1431.2005

Ivain L, Bordeau V, Eyraud A, Hallier M, Dreano S, Tattevin P, Felden B, Chabelskaya S. 2017. An *in vivo* reporter assay for sRNA-directed gene control in Gram-positive bacteria: identifying a novel sRNA target in *Staphylococcus aureus*. Nucleic Acids Res 45:4994–5007. doi:10.1093/nar/gkx190

Jaglic Z, Cervinkova D. 2018. Genetic basis of resistance to quaternary ammonium compounds; the qac genes and their role: a review. Veterinární Medicína 57:275–281. doi:10.17221/6013-VETMED

James EH, Edwards AM, Wigneshweraraj S. 2013. Transcriptional downregulation of agr expression in *Staphylococcus aureus* during growth in human serum can be overcome by constitutively active mutant forms of the sensor kinase AgrC. FEMS Microbiol Lett 349:153–162. doi:10.1111/1574-6968.12309

Kolb FA, Romby P, Lina G, Etienne J, Vandenesch F. 2000. Probing the structure of RNAIII, the *Staphylococcus aureus* agr regulatory RNA, and identification of the RNA domain involved in repression of protein A expression. RNA 6:668–679.

Korea CG, Balsamo G, Pezzicoli A, Merakou C, Tavarini S, Bagnoli F, Serruto D, Unnikrishnan M. 2014. *Staphylococcal* Esx proteins modulate apoptosis and release of intracellular Staphylococcus aureus during infection in epithelial cells. Infect Immun 82:4144–4153. doi:10.1128/IAI.01576-14

Lee E-J, Groisman EA. 2010. An antisense RNA that governs the expression kinetics of a multifunctional virulence gene. Mol Microbiol 76:1020–1033. doi:10.1111/j.1365-2958.2010.07161.x

Liu W, Rochat T, Toffano-Nioche C, Le Lam TN, Bouloc P, Morvan C. 2018. Assessment of Bona Fide sRNAs in *Staphylococcus aureus*. Front Microbiol 9. doi:10.3389/fmicb.2018.00228

Luong TT, Dunman PM, Murphy E, Projan SJ, Lee CY. 2006. Transcription Profiling of the mgrA Regulon in *Staphylococcus aureus*. J Bacteriol 188:1899–1910. doi:10.1128/JB.188.5.1899-1910.2006

Mäder U, Nicolas P, Depke M, Pané-Farré J, Debarbouille M, Kooi-Pol MM van der, Guérin C, Dérozier S, Hiron A, Jarmer H, Leduc A, Michalik S, Reilman E, Schaffer M, Schmidt F, Bessières P, Noirot P, Hecker M, Msadek T, Völker U, Dijl JM van. 2016. *Staphylococcus aureus* Transcriptome Architecture: From Laboratory to Infection-Mimicking Conditions. PLOS Genet 12:e1005962. doi:10.1371/journal.pgen.1005962

Malachowa N, Whitney AR, Kobayashi SD, Sturdevant DE, Kennedy AD, Braughton KR, Shabb DW, Diep BA, Chambers HF, Otto M, DeLeo FR. 2011. Global Changes in *Staphylococcus aureus* Gene Expression in Human Blood. PLOS ONE 6:e18617. doi:10.1371/journal.pone.0018617

Menzies BE, Kourteva I. 2000. Staphylococcus aureus α-toxin induces apoptosis in endothelial cells. FEMS Immunol Med Microbiol 29:39–45. doi:10.1111/j.1574-695X.2000.tb01503.x

Moise PA, Forrest A, Bayer AS, Xiong YQ, Yeaman MR, Sakoulas G. 2010. Factors Influencing Time to Vancomycin-Induced Clearance of Nonendocarditis Methicillin-Resistant *Staphylococcus aureus* Bacteremia: Role of Platelet Microbicidal Protein Killing and agr Genotypes. J Infect Dis 201:233–240. doi:10.1086/649429

Oldani A, Cormont M, Hofman V, Chiozzi V, Oregioni O, Canonici A, Sciullo A, Sommi P, Fabbri A, Ricci V, Boquet P. 2009. *Helicobacter pylori* counteracts the apoptotic action of its VacA toxin by injecting the CagA protein into gastric epithelial cells. PLoS Pathog 5:e1000603. doi:10.1371/journal.ppat.1000603

Peterson MM, Mack JL, Hall PR, Alsup AA, Alexander SM, Sully EK, Sawires YS, Cheung AL, Otto M, Gresham HD. 2008. Apolipoprotein B Is an Innate Barrier against Invasive *Staphylococcus aureus* Infection. Cell Host Microbe 4:555–566. doi:10.1016/j.chom.2008.10.001

Pozzi C, Waters EM, Rudkin JK, Schaeffer CR, Lohan AJ, Tong P, Loftus BJ, Pier GB, Fey PD, Massey RC, O'Gara JP. 2012. Methicillin Resistance Alters the Biofilm Phenotype and Attenuates Virulence in *Staphylococcus aureus* Device-Associated Infections. PLOS Pathog 8:e1002626. doi:10.1371/journal.ppat.1002626

Romilly C, Lays C, Tomasini A, Caldelari I, Benito Y, Hammann P, Geissmann T, Boisset S, Romby P, Vandenesch F. 2014. A Non-Coding RNA Promotes Bacterial Persistence and Decreases Virulence by Regulating a Regulator in *Staphylococcus aureus*. PLOS Pathog 10:e1003979. doi:10.1371/journal.ppat.1003979

Rossi J, Bischoff M, Wada A, Berger-Bächi B. 2003. MsrR, a putative cell envelope-associated element involved in *Staphylococcus aureus* sarA attenuation. Antimicrob Agents Chemother 47:2558–2564. doi:10.1128/aac.47.8.2558-2564.2003

Sakoulas G, Eliopoulos GM, Moellering RC, Wennersten C, Venkataraman L, Novick RP, Gold HS. 2002. Accessory Gene Regulator (agr) Locus in Geographically Diverse *Staphylococcus aureus* Isolates with Reduced Susceptibility to Vancomycin. Antimicrob Agents Chemother 46:1492–1502. doi:10.1128/AAC.46.5.1492-1502.2002

Song J, Lays C, Vandenesch F, Benito Y, Bes M, Chu Y, Lina G, Romby P, Geissmann T, Boisset S. 2012. The Expression of Small Regulatory RNAs in Clinical Samples Reflects the Different Life Styles of *Staphylococcus aureus* in Colonization vs. Infection. PLOS ONE 7:e37294. doi:10.1371/journal.pone.0037294

Stemerding AM, Köhl J, Pandey MK, Kuipers A, Leusen JH, Boross P, Nederend M, Vidarsson G, Weersink AYL, Winkel JGJ van de, Kessel KPM van, Strijp JAG van. 2013. *Staphylococcus aureus* Formyl Peptide Receptor–like 1 Inhibitor (FLIPr) and Its Homologue FLIPr-like Are Potent FcγR Antagonists That Inhibit IgG-Mediated Effector Functions. J Immunol 191:353–362. doi:10.4049/jimmunol.1203243

Subramanian D, Bhasuran B, Natarajan J. 2018. Genomic analysis of RNA-Seq and sRNA-Seq data identifies potential regulatory sRNAs and their functional roles in *Staphylococcus aureus*. Genomics. doi:10.1016/j.ygeno.2018.09.016

Sundaramoorthy R, Fyfe PK, Hunter WN. 2008. Structure of *Staphylococcus aureus* EsxA suggests a contribution to virulence by action as a transport chaperone and/or adaptor protein. J Mol Biol 383:603–614. doi:10.1016/j.jmb.2008.08.047

Tomasini A, Moreau K, Chicher J, Geissmann T, Vandenesch F, Romby P, Marzi S, Caldelari I. 2017. The RNA targetome of *Staphylococcus aureus* non-coding RNA RsaA: impact on cell surface properties and defense mechanisms. Nucleic Acids Res 45:6746–6760. doi:10.1093/nar/gkx219

6 RsaOG sponges RsaE in times of nutrient stress

6.1 Introduction

Virulence is a costly exercise and is thus intertwined with metabolism. The availability of energy and carbon resources dictate the expression of virulence genes. In support of this idea, it has been found that RNAIII is produced only in the absence of glucose and this repression was found to lead to elevated levels of *spa* (Seidl et al., 2006). Additionally, deletion of *ccpA*, a transcriptional regulator involved in metabolic regulation, also lowers RNAIII levels (Seidl et al., 2008), and was found to control the expression of toxic shock syndrome toxin 1, a major *Staphylococcal* virulence factor (Seidl et al., 2008). Finally, inactivation of the TCA cycle reduces the production of secreted virulence factors in *S. aureus* (Somerville et al., 2003, 2002).

Studies on CodY further demonstrating the link between virulence and metabolism. CodY is a carbon and nitrogen source sensor through its recognition of branched chain amino acids and is known to repress genes involved in amino acid transport, catabolism and biosynthesis (Sonenshein, 2005). CodY has also been implicated in regulating virulence factor synthesis in *Clostridium difficile, Listeria monocytogenes, Streptococcus pneumoniae and S. aureus* (Somerville and Proctor, 2009). Additionally, deletion of the *codY* locus in clinical isolates has been found to result in elevated *agr* activity and *hla*, and this in turn lead to increased haemolytic activity towards erythrocytes and increased biofilm formation (Majerczyk et al., 2008).

Additionally, MgrA, a pro-virulence transcription factor, responds to the intracellular redox conditions. It can be oxidised by TCA cycle-generated superoxide anions, which decreases its DNA binding potential and increases the expression of autolysin genes and the rate of cellular autolysis (Luong et al., 2006). When the TCA cycle in inactive, this autolysis is prevented.

It is important to note that the above descriptions are not exhaustive in describing the established link between *S. aureus* metabolism and virulence. For example, redox-responsive regulators, oxygen sensors and metal ion-responsive regulators

have also been linked to virulence factor production (reviewed by Somerville and Proctor, 2009).

Transcription factors are not alone in regulating the metabolic demands of the cell. It is becoming clear that post-transcriptional regulation is also important in regulating gene expression and coordinating adaptive responses. In *S. aureus*, the sRNA RsaE is believed to play a central role in regulating oligonucleotide transport, folate metabolism, arginine degradation and the TCA cycle. RsaE appears to be primarily expressed in the late-exponential stage, and is then downregulated in stationary (Bohn et al., 2010; Geissmann et al., 2009). However, it should be noted that the degree of change in expression between the exponential and the stationary phase varies between different *S. aureus* strains. Overall, it has been proposed that RsaE helps to coordinate the required metabolic remodelling for transitioning between these growth states.

RsaE is strongly conserved throughout the *Bacillales* order and has also been found to contribute to cellular NAD+/NADH balance in *B. subtilis* (Durand et al., 2015; Rochat et al., 2018). In *S. aureus*, RsaE has been shown to block the ribosomal binding site of its targets via its UCCC motif, triggering RNase III-mediated cleavage (Felden et al., 2011). In particular, RsaE is thought to downregulate the TCA cycle and purine biosynthesis upon the entry to stationary phase, when carbon sources become scarce.

The regulation of the TCA cycle by RsaE is exerted through its base-pairing with mRNA encoding succinyl-CoA synthetases (*sucC* and *sucD*), aconitase (*citB*) citrate synthase (*citZ*) and isocitrate dehydrogenase (*citC*). Additionally, it is also able to regulate *fhs* mRNA stability, which is involved in purine biosynthesis (Bohn et al., 2010; Geissmann et al., 2009). RsaE is also intimately involved in cellular amino acid availability through regulation of the oligopeptide transporter operon *opp-3*, and operons involved in valine, leucine, and isoleucine metabolism (Bohn et al., 2010).

Other targets of RsaE were identified through a technique called Hybrid-trap-seq (Rochat et al., 2018). This technique is similar to the aforementioned MAPS, however, it is entirely based on *in vitro* assembled interactions. In short, *in vitro* transcribed RsaE was produced, biotinylated, and then fixed onto magnetic streptavidin beads. These beads were then incubated in total cellular RNA extracts, washed, and then the RsaE binding partners were eluted and sequenced. In order to increase the confidence of found interactions, the expression of putative RsaE targets were investigated in a Δ*rsaE* strain, compared to both a wild type and a ΔrsaE strain where RsaE expression was induced through a plasmid. These analyses predicted direct interactions between RsaE and *rocF, rocD, ndh2* and *icaR*. These interactions

are thought to not only regulate translational efficiency but also mRNA stability as well. RocF and RocD are involved in the conversion of arginine to ornithine and then to glutamate semialdehyde respectively, and the authors also showed that a Δ*rsaE* strain grew faster in Complete Defined Medium where the only carbon source is amino acids (Rochat et al., 2018). Finally, the authors reported that RsaE contains two distinct UCCC motifs that function independently of each other in their regulation of *rocD* and *rocF* (Rochat et al., 2018). Regarding RsaE's regulation of *icaR*, this interaction was proposed to explain RsaE's promotion of biofilm production in *S. epidermidis* (Schoenfelder et al., 2019).

As RsaE is known to be involved in regulating cellular metabolism, I hypothesised that it could also be involved in regulating the virulent state of the cell. The CLASH data described here included interactions between RsaE and several α phenol soluble modulin (PSM) mRNAs. PSMs are small, secreted peptides that cause disruption of membranes through insertion via amphipathic helices, inducing pore formation (Cheung et al., 2014). PSMs make up around 60% of the total secreted peptides of *S. aureus*, indicating that these are critical factors in *S. aureus* biology (Wang et al., 2007).

All PSMs within *S. aureus* are produced from two operons, the α and the β. The αPSMs range from 20-25 amino acids in length, while the β are between 43 and 45 amino acids long (Cheung et al., 2014). The PSM operons are regulated directly by AgrA binding to their promoters and thus they are increasingly expressed as cellular density increases (Xu et al., 2017). Interestingly, PSMs have already been shown to be subject to regulation by sRNAs, with Teg41 positively regulating the production of αPSMs, and deletion of the crucial regulatory seed sequence of Teg41 attenuated virulence in a murine model of infection (Zapf et al., 2019).

Within the local *S. aureus* community, PSMs can be secreted to induce biofilm remodelling in order to enable cellular dissemination (García-Betancur et al., 2017). Additionally, they are able to act as critical virulence factors through their cytotoxic activity, and have been shown to induce lysis of osteoblasts, endothelial cells, epithelial cells, monocytes, erythrocytes and neutrophils (Cheung et al., 2014). The αPSMs also overall displaying increased lytic activity as compared to the βPSM class (Cheung et al., 2014; Wang et al., 2007). Finally, αPSMs have been shown to be crucial in *S. aureus* escape from phagosomes in both non-professional phagocytes (such as epithelial cells), and also from monocytes (Grosz Magdalena et al., 2014). Demonstrating the overall crucial nature of αPSMs in virulence, their deletion results in decreased skin legions and morbidity in a murine model of infection, and also decreased cortical bone destruction in a murine model of *S. aureus*-mediated osteomyelitis (Cassat et al., 2013; Wang et al., 2007).

Interestingly, Hybrid-trap-seq and MAPS also identified another sRNA, RsaOG, as a target of RsaE (Bronesky et al., 2019; Rochat et al., 2018). However, Bronesky et al were unable to recapitulate this interaction *in vitro*, and no further validations were attempted by Rochat et al. RsaOG (also known as RsaI) is highly conserved in *Staphylococcacea* and is involved in a signalling pathway which responds to glucose availability (Bronesky et al., 2019; Geissmann et al., 2009). In rich medium, RsaOG is known to be repressed by CcpA when glucose is present, but becomes induced in its absence (Bronesky et al., 2019). Using MAPS, the authors managed to identify several targets of RsaOG and proposed that RsaOG functions to switch the cell away from glucose utilisation. Of these targets, two were validated as causing translational repression; *glcU_2*, a glucose permease, and fructosamine-3-kinase. Additional targets involved in glucose catabolism were identified through MAPS, such as *treB, fba, gnd* and *tkt*. Interestingly, *icaR* was also identified as a target of RsaOG, and thus *icaR* may be subject to regulation from both RsaE and RsaOG.

Alongside these mRNAs and RsaE, RsaOG was also found to bind RsaG and RsaD through both MAPS and *in vitro* EMSAs (Bronesky et al., 2019). The functional consequences of these sRNA – sRNA interactions are unknown but were speculated to be involved in RsaOG's promotion of nitric oxide (NO) detoxification. From an infection perspective, RsaOG's promotion of NO resistance and regulation of metabolic rewiring was hypothesised to place it as a potential player in *S. aureus* survival within host cell phagocytes.

The interaction between RsaOG and RsaE, RsaG and RsaD links to the idea of sRNA sponging. As discussed in Section 1.11, sRNA sponging has been observed in the bacterial species *S. typhimurium* and *E. coli*, but has not been demonstrated in *S. aureus*. Although no biological significance has been attributed to the interaction between RsaOG and RsaE, RsaG or RsaD, it is very likely that sRNA sponging will be an important regulatory mechanism in *S. aureus* biology. This is because sRNA sponging allows responses to occur on rapid timescales and allows co-regulation of different adaptive responses (as described in Section 1.11).

As discussed in Chapter 3, CLASH identified a large number of RsaE and RsaOG interactions in RPMI medium. This was hypothesised to be a sponging interaction, as the expression of RsaOG was immediately upregulated in RPMI, but the stability of RsaE remained constant. This chapter will discuss the validation of this interaction and the findings that attribute biological functionality to it. Finally, *in vitro* validations of RsaE and phenol soluble modulins will be described.

6.2 CLASH identifies known mRNA targets of RsaE

I first asked how many of the known RsaE interactions were recovered from CLASH in order to increase the confidence in my data, similarly to what was described in Section 5.2. Through microarray analysis, *purK* was identified as being upregulated following RsaE deletion (Geissmann et al., 2009). Although this interaction was not further validated, the CLASH experiments performed here identified 4 unique interactions between RsaE and *purK*, occurring in the 5' UTR of *purK* at a region that could be a SD sequence ('AGGG'), pairing with the distinctive UCCC motif of RsaE (Figure 6.1). As described in Section 6.1, the *opp-3* operon is another target of RsaE, identified via microarray analysis and validated through EMSAs and toeprinting. Here, hybrids between RsaE and *opp-3B* were identified that covered the start codon and SD sequence, again through the UCCC motif of RsaE. Finally, a similar finding was observed for *sucD*, although the *in silico* structural prediction did not pair the start codon with RsaE.

The interaction between *purH* and RsaE was previously found in Hybrid-trap-seq (Rochat et al., 2018) and RsaE has already been implicated in the regulation of another member of the *pur* operon, *purK* (Geissmann et al., 2009). Here 14 interactions were found between *purH* and RsaE, however, this was found to occur within the coding region of *purH*. Additionally, when *in silico* structural predictions were made, the UCCC motif of RsaE was only very tenuously base-paired with *purH*, although an extended basepairing interaction is seen downstream. Given that this interaction does not mimic in the canonical mode of RsaE regulation, i.e. it does not occur at the 5' UTR and SD sequence/start codon region and does not utilising RsaE's UCCC motifs, this interaction requires further investigation before significant meaning is attributed to it.

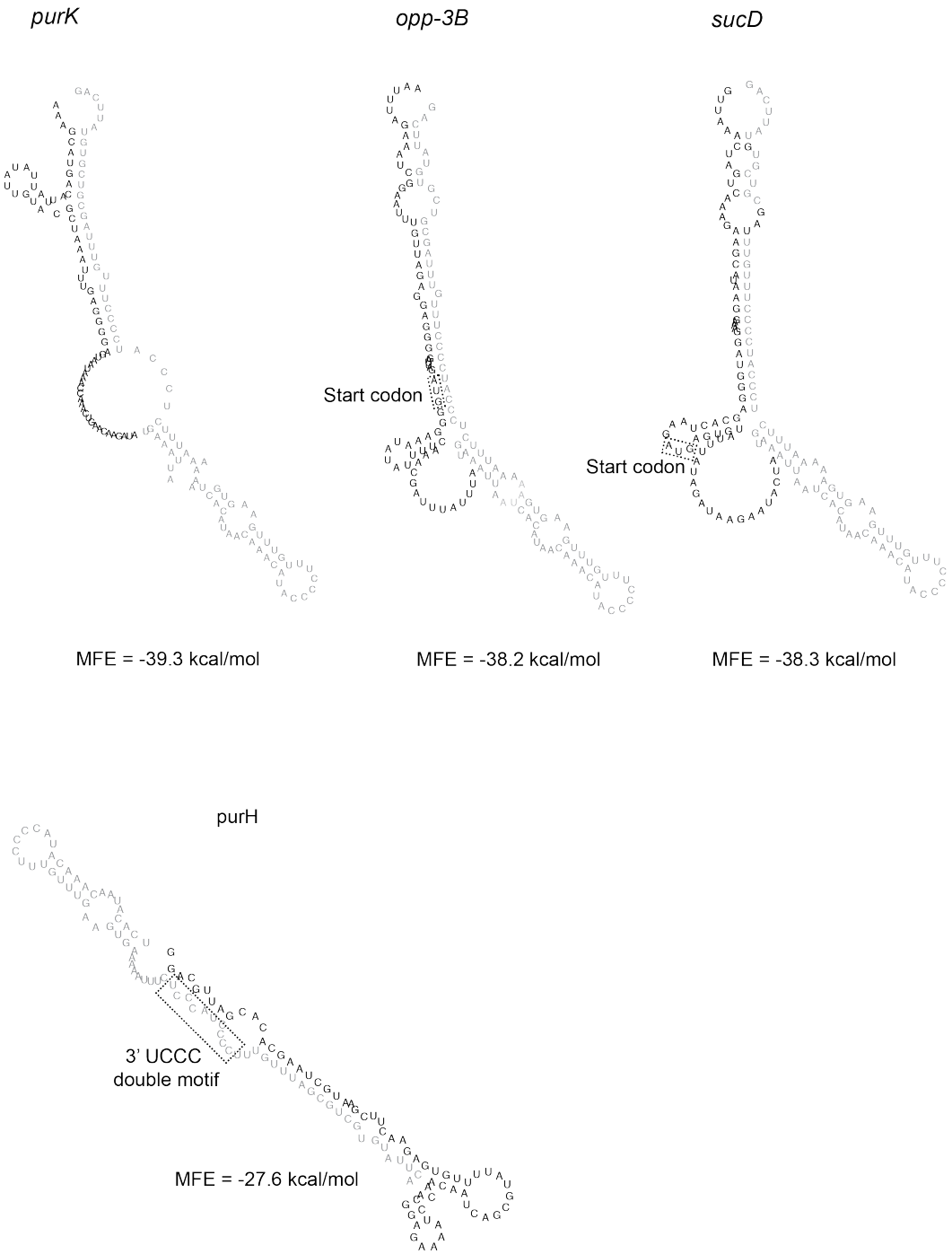

Fig. 6.1: Predicted interactions between RsaE and its known targets. RsaE is shown in grey and the mRNA targets in black. Structures were generated by extracting the CLASH read sequences as input for RNAcofold, which was used to predict the interaction structure. Due to the fact that short read fragments were recovered for *purK, opp3-B* and *sucD*, 20 nucleotides from either side of the hybrid were added to the 5' and 3' end.

6.3 The expression of RsaOG responds to RPMI stress, while RsaE does not

Reviewing the RNAseq data reveals that upon the shift to RPMI medium, RsaOG is immediately upregulated 6-fold by 5 minutes (Figure 6.2A). This is maintained until after 10 minutes, when steady state levels decrease rapidly, and RNA levels return to basal levels by 30 minutes. This is in contrast to the control TSB shift, where the expression of RsaOG remains virtually constant, perhaps decreasing slightly. Examining the RNA polymerase and RNase III binding to RsaOG explains how this trend is shaped. The shape of the RNA polymerase binding profile matches that of the RNA steady state levels where a sharp increase is observed after 5 minutes, which then decreases afterwards (Figure 6.2, middle row). However, it should be noted that the stoichiometry is not the same, with only a 2-fold increase in RNA polymerase RsaOG binding, as compared to a 6-fold increase in the expression levels. As such, the increase in RsaOG levels may reflect an increase in transcription and an increase in RNA stability. The RNase III data is also in agreement, as it can be seen that there in increased RNase III binding to RsaOG after 15 minutes of RPMI shift, which is when the RNA steady state levels start to decrease. Thus, it can be concluded that the expression of RsaOG responds positively to the shift to RPMI, and that the dynamics of this are shaped by RNase III and the RNA polymerase.

In contrast to this, the expression of RsaE remains basal through both the TSB and RPMI shifts, indicating that RsaE does not respond to the stresses imposed (Fig 6.2A). The RNA polymerase binding to RsaE actually decreases throughout the time course, however the RNase III data is noisy to the degree where no conclusions can be made. It may be that although the RNA polymerase decreases the rate of RsaE transcription, the cellular pool already existing is not targeted by RNase III to a degree that would change the steady state levels.

In order to explain the dramatic change in the RNA levels of RsaOG, it is important to consider the growth conditions at the time of the shift. TSB contains 0.2% glucose, the same as RPMI. At the late-exponential phase in TSB, I predict that the cells have depleted most, and if not then all, of the available glucose. Thus, the shift to RPMI actually induces a glucose replenishment. As RsaOG is known to be positively regulated by CcpA, its upregulation indicates that CcpA is activated after the RPMI shift (Mäder et al., 2016, Bronesky et al., 2019). This makes sense as CcpA is known to promote glycolysis and shut down the TCA cycle, which would be expected under conditions in which glucose is present (Seidl et al, 2009). Thus, I predict that the upregulation of RsaOG is mediated by the response of CcpA to glucose.

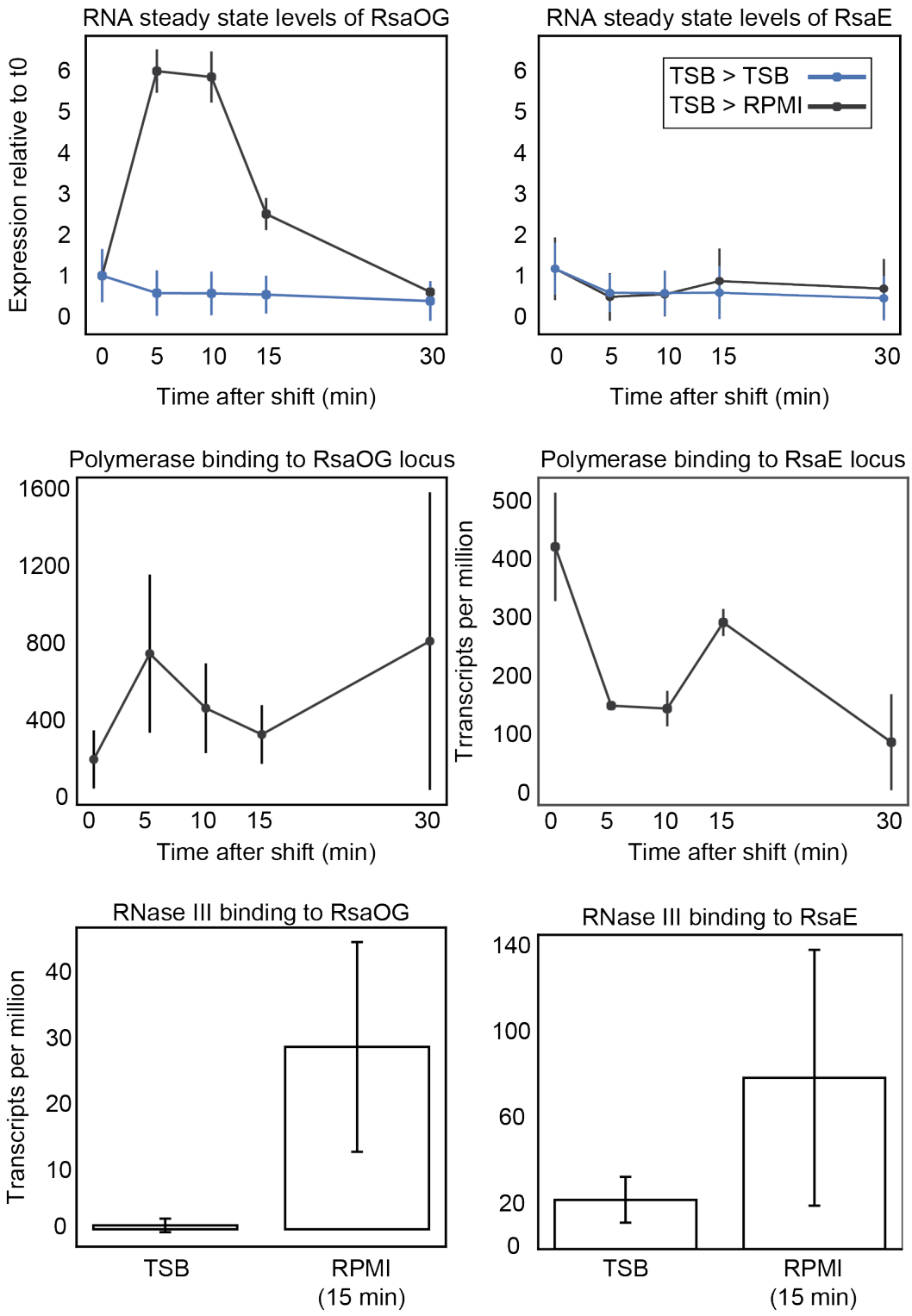

Fig. 6.2: Shaping the expression profile of RsaOG and RsaE. Top panels: RNA steady state levels of RsaOG and RsaE after the shift to RPMI, and a control TSB shift as measured by RNAseq. Middle: RNA polymerase binding to RsaOG and RsaE throughout the shift. Bottom: RNase III binding to RsaOG and RsaE in TSB and after 15 minutes of RPMI shift

6.4 RsaOG and RsaE interact *in vitro* through both of RsaE's UCCC motifs

Examination of the CLASH data revealed reasonably long interacting reads for both RsaE and RsaOG, and it was noted that both of the UCCC motifs of RsaE were included. Through *in silico* folding, it was predicted that only the 3' UCCC motif interacted with a GGG trimer in RsaOG (Figure 6.3A).

In order to examine this hypothesis, two mutant versions of RsaE were produced *in vitro*; one in which the 5' UCCC mutant was mutated to AGGG, and another mutant in which both the 5' and the 3' UCCC double motifs were mutated. EMSAs were then performed with radiolabelled RsaOG, as compared to wild type RsaE. As shown in Figure 6.3B, a strong band shift was observed when incubating RsaOG with RsaE. In order to quantify this, the amount of dimer (indicated with the chevron) was compared to the total amount of signal relating to the monomer plus the dimer. Although another band is observed above the proposed dimer band, this band is also observed when only RsaOG is present without RsaE. Thus, it may be that RsaOG is able to oligomerise, potentially into a trimer. In an attempt to examine formation of just an RsaOG – RsaE interaction, only the dimer band was examined as this is not observed when RsaOG is alone. Interestingly, the maximum occupancy was observed when RsaOG was incubated with just a 5-fold molar excess of RsaE, giving a binding of 60% of the total examined pool. Increasing the amount of RsaOG, even up to 320-fold excess, did not increase this binding.

Why RsaE binding does not increase in the presence of increasing amounts of RsaOG is not clear, however, the following experiments suggest that the interaction is specific. Binding to RsaOG is significantly disrupted upon mutation of one of RsaE's UCCC motifs. The maximum occupancy is reduced to around 25%, and this also only occurs at a higher molar excess of 20. Interestingly, increasing the molar excess does not change this, despite the fact that more UCCC motifs are being added to the reaction that could, in theory, compensate for the mutation. Once both of RsaE's UCCC motifs are removed, binding is completely abrogated, and dimer bands are no longer visibly observed. Thus, it is overall concluded that RsaE is able to interact with RsaOG using both of its UCCC motifs, and that these are able to bind RsaOG independently of each other, but that maximum RsaOG binding relies on their cooperative action. I am currently making compensatory mutations in RsaOG to determine if I can restore the interaction with the RsaE UCCC double mutant.

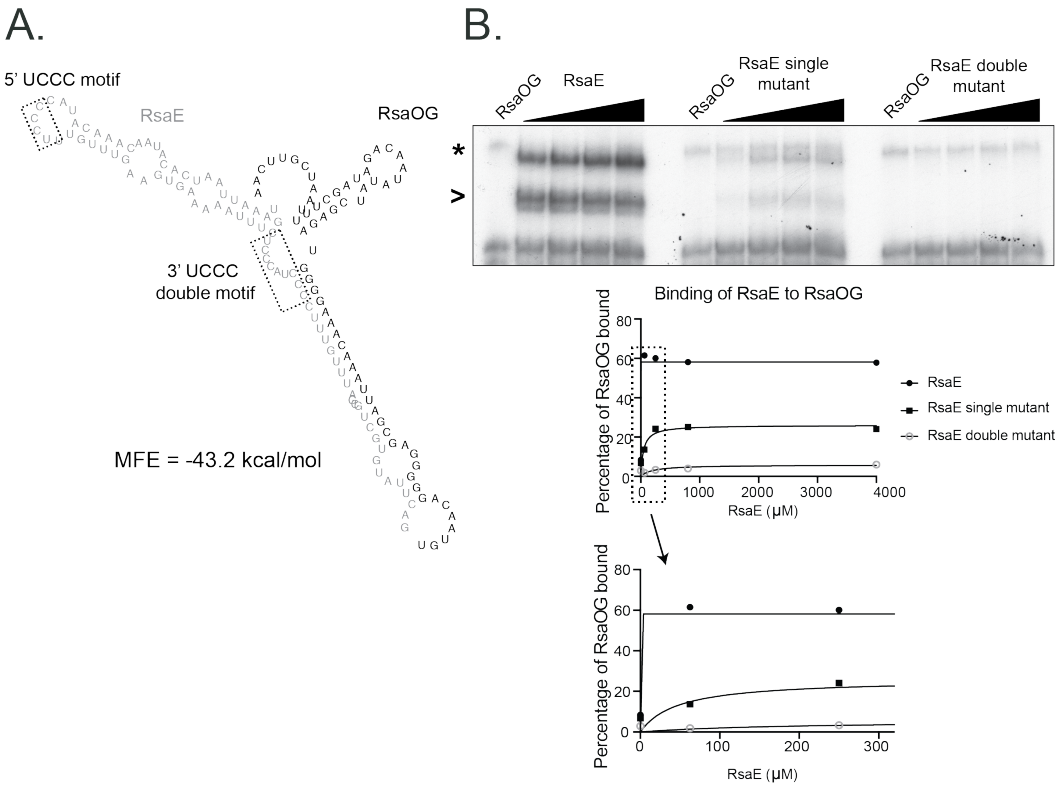

Fig. 6.3: *In vitro* interaction between RsaOG and RsaE. **A.** *In silico* predicted structure of the RsaE – RsaOG interaction. RsaE is shown in grey and RsaOG in black. Structures were generated by extracting the CLASH read sequences as input for RNAcofold, which was used to predict the interaction structure. The two UCCC motifs are highlighted. **B.** EMSA of radiolabelled RsaOG with RsaE WT, the single RsaE UCCC mutant, and the double mutant. Asterisk (*) indicates the proposed trimer, while the chevron (>) indicates the proposed dimer. Quantifications are shown below the gel, in which the amount of dimer was compared to the total monomer plus dimer signal.

6.5 RsaOG may bind RsaE in a sponging interaction that responds to RPMI

Examining the number of interactions found by CLASH between RsaOG and RsaE reveals a sharp increase upon the shift to RPMI medium. In TSB, 12 interactions were found while 63 were observed in RPMI (Figure 6.4A). Although CLASH is not considered a quantitative approach, the observed increase in RsaOG expression suggests that the newly transcribed RsaOG molecules are able to increasingly target RsaE. Given that RsaE steady state levels did not significantly change during the timecourse, this begs the question whether RsaOG is able to bind to RsaE but not induce its degradation, even though this interaction is recognized by RNase III. I hypothesize that RsaOG is a *bona fide* sRNA sponge that, by base-pairing with RsaE, sequesters RsaE away from its targets. To test this hypothesis, I examined the expression of several RsaE targets in the TSB-RPMI timecourse by qPCR. I reasoned that if RsaOG indeed sponges RsaE, then the targets that are normally negatively regulated by RsaE should now be upregulated or at least be stabilized. I chose *rocA, fumC* and *citB* for these analyses as they were identified as potential targets of RsaE through transcriptome analysis of an RsaE deletion strain (Rochat et al., 2018).

Firstly, consistent with my RNA-seq data, the qPCR data showed that RsaOG was significantly upregulated shortly after the shift to RPMI (Figure 6.4B, upper left panel). RsaOG was 6-fold upregulated after 5 minutes and these high levels were maintained until 10 minutes, after which they went back down to basal levels after 30 minutes. A similar expression pattern was observed for RsaE targets (Figure 6.4B). All of *rocA, fumC* and *citB* showed an immediate upregulation in response to the RPMI shift, albeit to varying degrees. *rocA* underwent the most significant increase, with around a 10-fold expression increase after just 5 minutes and decreased down to basal levels by 30 minutes in an expression profile shape identical to that of RsaOG. Regarding *fumC*, its expression is only increased to a maximum of around 1.75-fold by 10 minutes, before it starts to decrease afterwards. The *citB* and *rocA* mRNAs showed a similar profile as RsaOG, although the data for the latter contained noise. Overall, the targets of RsaE are immediately upregulated after the shift to RPMI which provides further evidence for a sponging interaction between RsaOG and RsaE.

It is important to note that *citB* and *rocA* have been shown to be repressed by CcpA (Seidl et al, 2009). However, their upregulation is in opposition to the hypothesis already discussed in which CcpA is activated after the shift to RPMI in order to shut down the TCA cycle and amino acid metabolism following glucose replenishment. Considering that CitB and RocA are members of these respective

metabolic processes, this overall implies that there are other transcription factors potentially competing with CcpA. As such, the expression of these genes likely involves multiple transcription factors.

In order to differentiate between post-transcriptional and transcriptional effects, the RNA polymerase binding to *rocA, fumC*, and *citB* was examined (Figure 6.4C). *citB* offers the clearest data, with the distinctive immediate upregulation observed, which is maintained until 10 minutes and starts to decrease afterwards. Although the data for *fumC* is noisier, it can be said that *fumC* also displays a sharp increase after 5 minutes, which returns to basal by the end of the time course.

Unfortunately, the t0 sample for *rocA* displays significant error which makes conclusions about the overall nature of the profile difficult. However, again it is observed that expression remains constant between 5 and 10 minutes and decreases afterwards.

Thus, I conclude that the RNA polymerase helps to shape the expression of these *rocA, fumC* and *citB*, as the profile of their RNA steady state levels matches the RNA polymerase binding. As such, the data strongly suggest that both transcriptional and post-transcriptional regulatory mechanisms act cooperatively to shape the gene expression pattern.

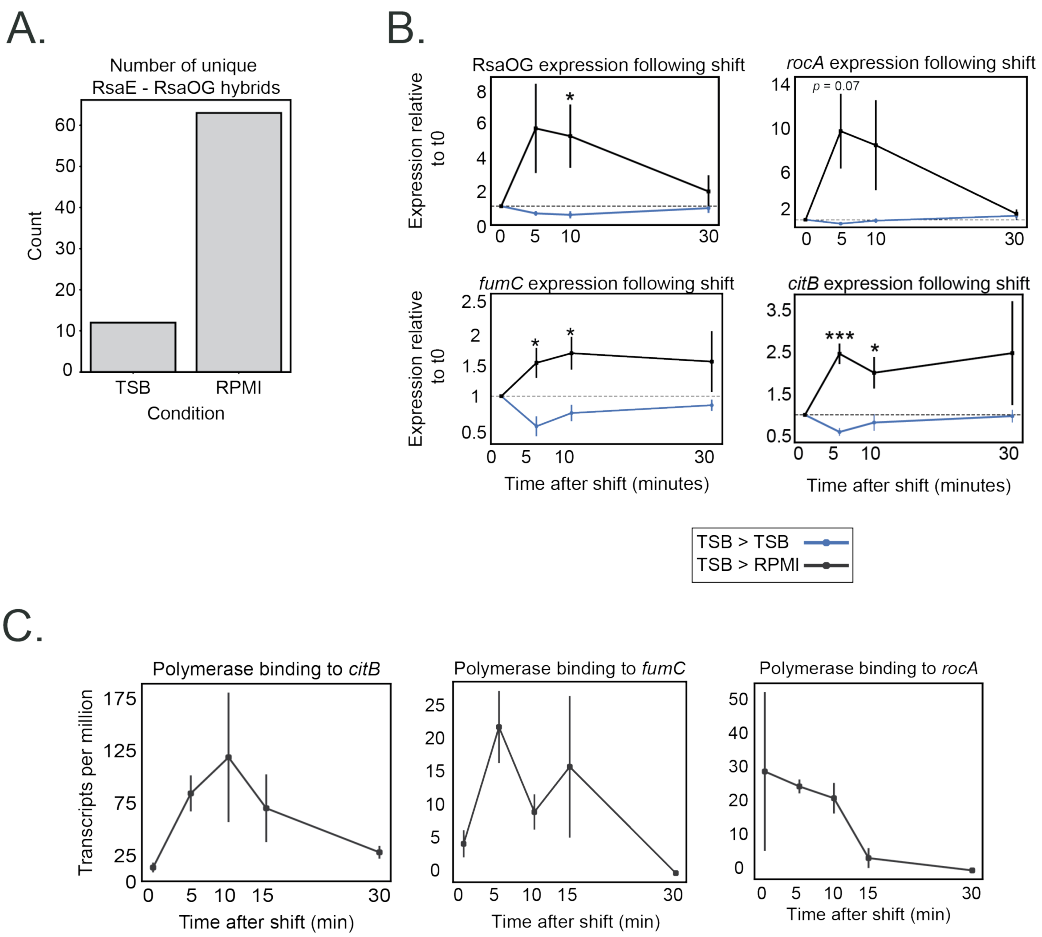

Fig. 6.4: The interaction between RsaE and RsaOG responds to RPMI stress and may dictate the fate of RsaE's targets. **A.** Number of interactions found between RsaE and RsaOG in TSB and after 15 minutes in RPMI. **B.** Expression of RsaOG, *rocA, fumC* and *citB* after the shift to RPMI, as measured by qPCR. Statistical significance was calculated through comparing the expression value in RPMI versus TSB. *p* values: * = 0.05, ** = 0.01, *** = <0.01 **C.** RNA polymerase binding to *citB, fumC,* and *rocA* following the shift.

Thus, the expression of RsaOG responds to the stresses induced by RPMI, which results in its rapid but transient upregulation. In these stressful conditions, an increased number of interactions were observed between RsaOG and RsaE, and this interaction was recapitulated *in vitro* and proven to be specific. Finally, through examination of the expression patterns of the targets of RsaE, it is hypothesised that this sponging of RsaE results in its targets being relieved of their previous repression, allowing them to be upregulated. However, transcriptional effects are likely to play a role in this dynamic as well.

6.6 RsaE binds phenol-soluble modulins *in vivo* and this can be recapitulated *in vitro*

Interestingly, my CLASH data revealed interactions between RsaE and PSMα2 plus PSMα3. All of these were identified in the TSB data, with 15 unique interactions between PSMα2 and 10 between PSMα3 found. *In silico* structure predictions of the CLASH reads show that it is the 3' UCCC motif of RsaE that is predicted to interact with a SD-like sequence of both of these PSMs (Figure 6.5 A and B, top panel).

In order to confirm these interactions, *in vivo* GFP reporter assays, like those described in Section 5.8 for RsaA with *qacAB* and RNAIII with *esxA*, were attempted by Jai Tree and Julia Wong in Sydney. Unfortunately, the PSM-GFP fusions did not exhibit sufficient fluorescence for meaningful interpretation to be made. It is conceivable that the PSM-GFP fusions were unstable or toxic to the cell. As an alternative approach to demonstrate the interaction between these RNA molecules, I performed *in vitro* EMSAs. As observed in Figure 6.5A and B, bottom panel, a specific band relating to duplex between RsaE and each PSM was observed (indicated by the chevron). When even a single UCCC motif was mutated, this interaction was completely abolished. It is also clear that the interaction between RsaE and PSMα3 is stronger than that between PSMα2, as specific signal is observed at lower molar ratios of RsaE. Thus, I conclude that RsaE interacts with these PSM transcripts *in vitro*, although I point out that a vast excess of RsaE is required (320 molar excess) to see a significant shift. Nevertheless, the data suggest that this interaction is specific and that only a single UCCC motif in RsaE is sufficient to support this interaction *in vitro*, in opposition to that observed for RsaOG.

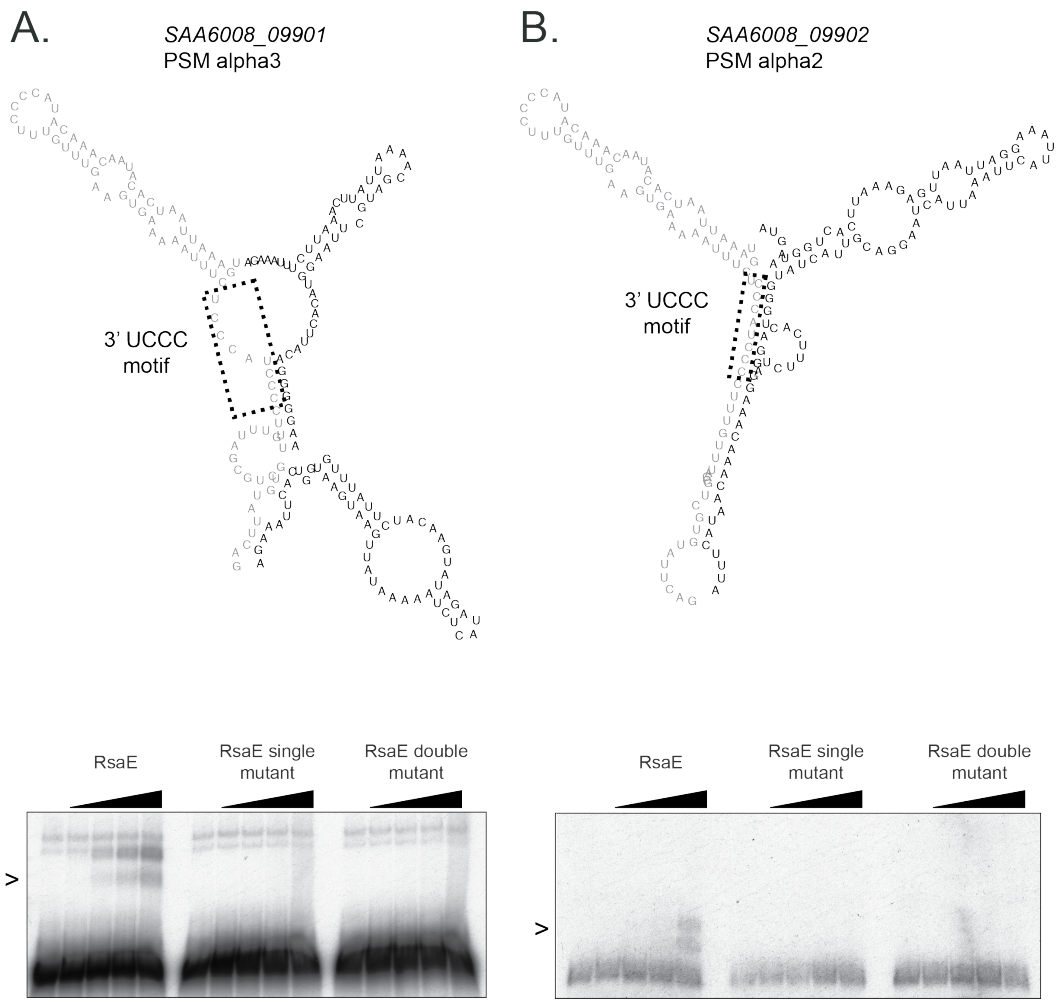

Fig. 6.5: Interaction between RsaE and PSMα3 (**A.**) and PSMα2 (**B.**). Top: predicted interactions between RsaE and the respective PSMs. RsaE is shown in grey and the PSMs in black. Structures were generated by extracting the CLASH read sequences as input for RNAcofold, which was used to predict the interaction structure. The two UCCC motifs are highlighted. Bottom: EMSA between the respective radiolabelled PSMs with RsaE WT, the single RsaE UCCC mutant, and the double mutant. The chevron (>) indicates the proposed dimer.

6.7 CLASH reveals novel RsaE targets

The CLASH data also uncovered other novel targets of RsaE which were manually inspected after filtering for statistical significance. I asked whether or not the interaction was found in multiple datasets, whether the interaction was discovered in both orientations (RNA1-RNA2, RNA2-RNA1), and whether or not they fitted the already known role for RsaE in regulating metabolic processes, and by extension virulence. This final proposed functionality of RsaE has been extended through the previously discussed interaction between RsaE and the αPSMs. This filtering left two proposed candidates; glutamine synthase, *glnA*, and an integrase of unknown function encoded in a pathogenicity island, *SAA6008_00346*. The interaction with *glnA* covers two nucleotides of the start codon and extends downstream of it and so is likely to represent a canonical repressive action. Glutamine synthase is part of the nitrogen metabolism pathway, and it catalyses the production of glutamine using glutamate and ammonia as precursors (Joo et al., 2018). This proposed interaction between RsaE and *glnA* fits with RsaE's proposed regulation of *rocD* and *rocF*, as described in Section 6.1, which are also involved in glutamine synthesis (Rochat et al., 2018). Thus, RsaE may have a more intimate role in nitrogen metabolism, and therefore amino acid anabolism and purine/pyrimidine synthesis, than previously thought.

The interaction between RsaE and *SAA6008_00346* was found in every CLASH dataset obtained. Examining the CLASH reads show that the start codon of *SAA6008_00346* was identified, an *in silico* folding predicts that the 3' UCCC double motif exhibits perfect base pairing with a downstream SD sequence. However, very little is known about *SAA6008_00346* and what is available comes from bioinformatic predicitons. *SAA6008_00346* is encoded in a pathogenicity island and is a proposed integrase with similarity to the ICEB1 class. If true, this would make *SAA6008_00346* a recombinase that is part of a mobile genetic element (Johnson and Grossman, 2015). Such mobile genetic elements are known to be activated in times of DNA damage, and although it is possible that the shift to RPMI medium induces such stress (either directly or indirectly), it may that the UV cross-linking used in CLASH activated this element's expression. Although previous studies have shown that the crosslinker used here greatly reduces contaminating signal from UV-mediated DNA damage, this still inevitably occurs (Nues et al., 2017). Regardless, it is interesting that RsaE is able to interact with this mRNA *in vivo*, and if this interaction were true then it may open up a new role for RsaE in regulating mobile genetic elements.

Thus, two mRNAs have been proposed as novel RsaE targets worthy of further exploration, *glnA* and *SAA6008_00346*. If these interactions were to pass future

validation, then this would further associate RsaE with nitrogen metabolism but also add a role in regulating genome integrity.

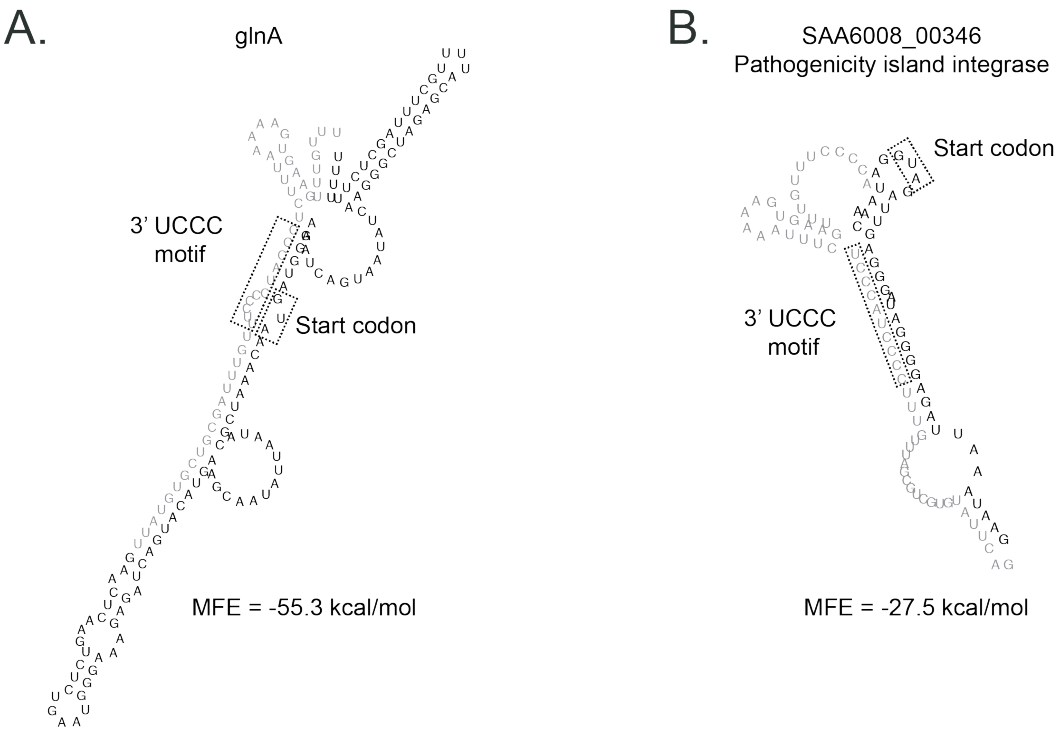

Fig. 6.6: Predicted interactions between RsaE and novel targets as predicted by CLASH. RsaE is shown in grey and the mRNA targets in black. Structures were generated by extracting the CLASH read sequences as input for RNAcofold, which was used to predict the interaction structure.

6.8 Discussion

The work described in this chapter further demonstrates that CLASH is able to recover known interactions of well-studied sRNAs, in this case RsaE. One of these targets is the sRNA RsaOG. Although this has already been described, its validation and the understanding of this interaction has remained elusive. This chapter has managed to demonstrate the specificity of this interaction and has started to attribute biological significance. Furthermore, CLASH was able to identify novel targets of RsaE from the phenol soluble modulin class, and these were further validated *in vitro*.

CLASH uncovers known RsaE target interactions

Through overexpression and microarray plus RT-qPCR analysis, RsaE was predicted to downregulate *gcvP* and *gcvT*; *fhs*; *sucC* and *sucD*; *citB*, *citC* and *citZ*; *opp3A* and *opp3B*; and *opp4A* and *opp4B* (Bohn et al., 2010). However, it should be noted that

in vitro EMSAs aimed at validating several of these interactions failed, with the sole exception of *opp3A* (Bohn et al., 2010). Additionally, RsaE deletion and microarray analysis found upregulation of *purK* (Geissmann et al., 2009), and Hybrid-trap-seq principally identified *purH, fumC, RsaOG, rocA, rocD* and *rocF* (Rochat et al., 2018).

From the known bank of interactions, CLASH managed to recover the interactions with *opp3-B, sucD, purK* and RsaOG (Figure 6.7). It should be noted that CLASH is not necessarily an exhaustive technique due to its reliance on the stochastic RNA ligation step. As such, I do not expect to recover every known interaction. Additionally, the previous *in vivo* studies that overexpressed RsaE were performed at OD_{600} values of 0.6 and 0.5 respectively, while the CLASH experiments were performed at an OD_{600} of 3. Thus, these interactions may not occur at the tested density.

The RsaE – RsaOG interaction responds to stress

RsaOG has already been hypothesised to be a binding partner of RsaE through Hybrid-trap-seq, but attempts to validate this interaction through *in vitro* EMSAs failed (Rochat et al., 2018). Here, 75 unique interactions between RsaOG and RsaE were identified and validation through EMSAs was performed. These EMSAs proved that the interaction is specific and occurs through both of RsaE's UCCC motifs.

It is not easy to explain why the interaction between RsaE and RsaOG was able to be recapitulated *in vitro* in this study and not by Rochat et al, 2018. Examination of the materials and methods reveal similar buffer compositions and molar ratios. However, there may have been some difference in the sample preparation and RNA refolding (personal communications).

The vast majority of RsaE – RsaOG interactions were found after the shift to RPMI. This was accompanied by a significant upregulation in the RNA steady state levels of RsaOG during the first 10 minutes of the shift, after which the levels return to basal by 30 minutes. This pattern of expression is also recapitulated into the RNA polymerase binding of its RNA, indicating that transcription is shaping the expression profile. However, although the duplex between RsaOG and RsaE is recovered using RNase III, the stability of RsaE remains constant throughout the time course. Thus, it may be that RNase III binds but does not degrade the duplex, or that any degradation is compensated by transcription. Such binding with a lack of degradation has been shown to occur in *E. coli*, where RNase III binds to but does not degrade the cIII gene of bacteriophage lambda (Altuvia et al., 1987). Additionally, RNase III is able to cleave just one strand of a double-stranded target (e.g. an sRNA-sRNA duplex), and so perhaps only RsaOG from the interaction is cleaved (Court et al., 2013).

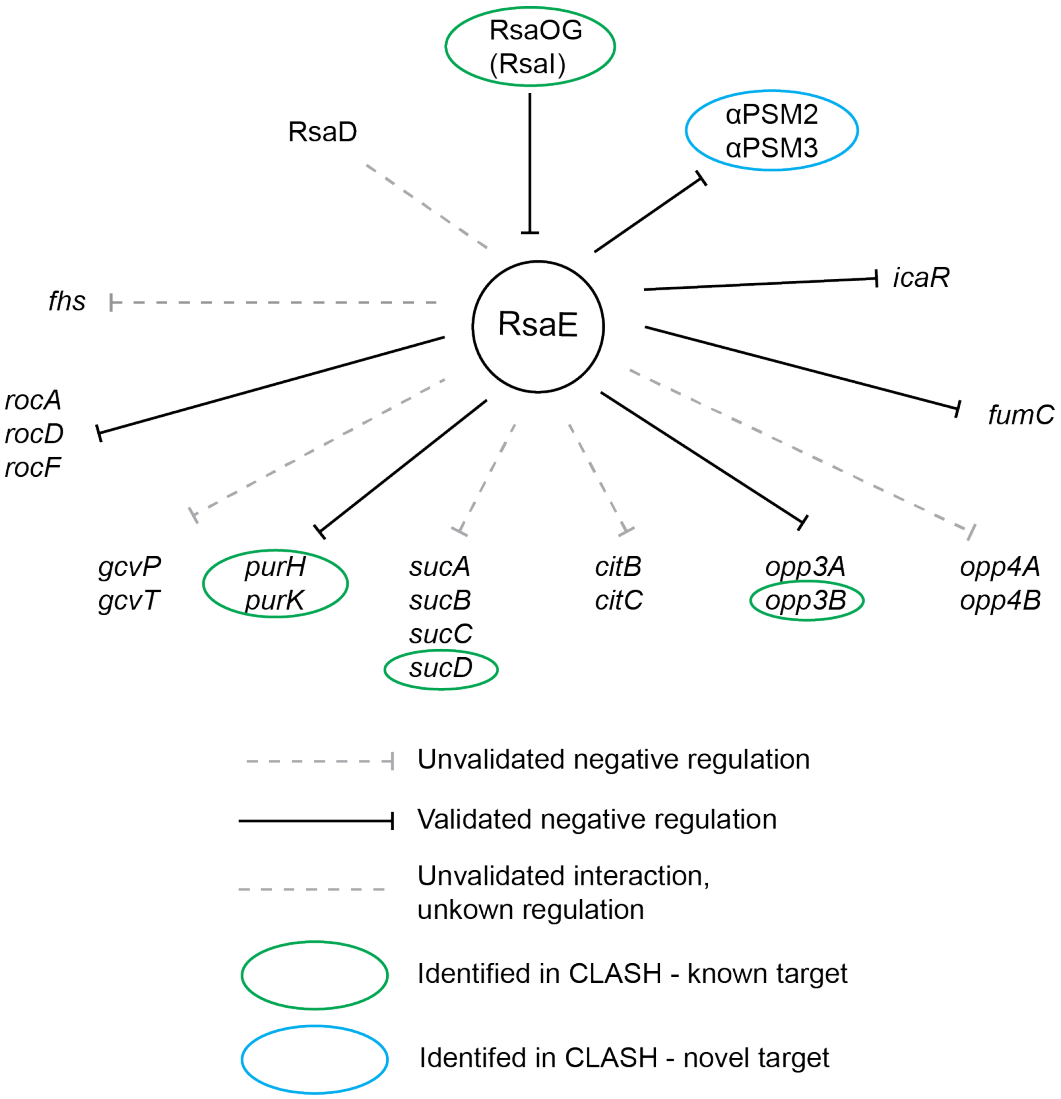

Fig. 6.7: Targetome of RsaE. Targets found in previous studies (Geissman et al, 2009; Rochat et al, 2018 and Bronesky et al, 2019) are shown alongside the newly proposed interaction with the αPSMs (blue circle). Known interactions which were recovered in this study are circled in green.

This leads to the model that upon the shift to RPMI, RsaOG is upregulated through increased transcription. The newly transcribed RsaOG molecules are then able to bind to RsaE, drawing it away from its targets, but do not induce significant downregulation (a so-called sponging interaction). The mRNAs that RsaE was previously targeting for degradation, such as *rocA, citB* and *fumC*, are now freed of their repression. As such, this model predicts that they would be upregulated after the shift for a period (Figure 6.8). Once RsaOG levels decrease as a result of decreased transcription and potentially degradation, RsaE is able to return to its original activity and repress its targets once more.

Examining the expression of the RsaE targets, such as *rocA, fumC* and *citB* gives credence to this model. After the shift to RPMI, these mRNAs were seen to immediately

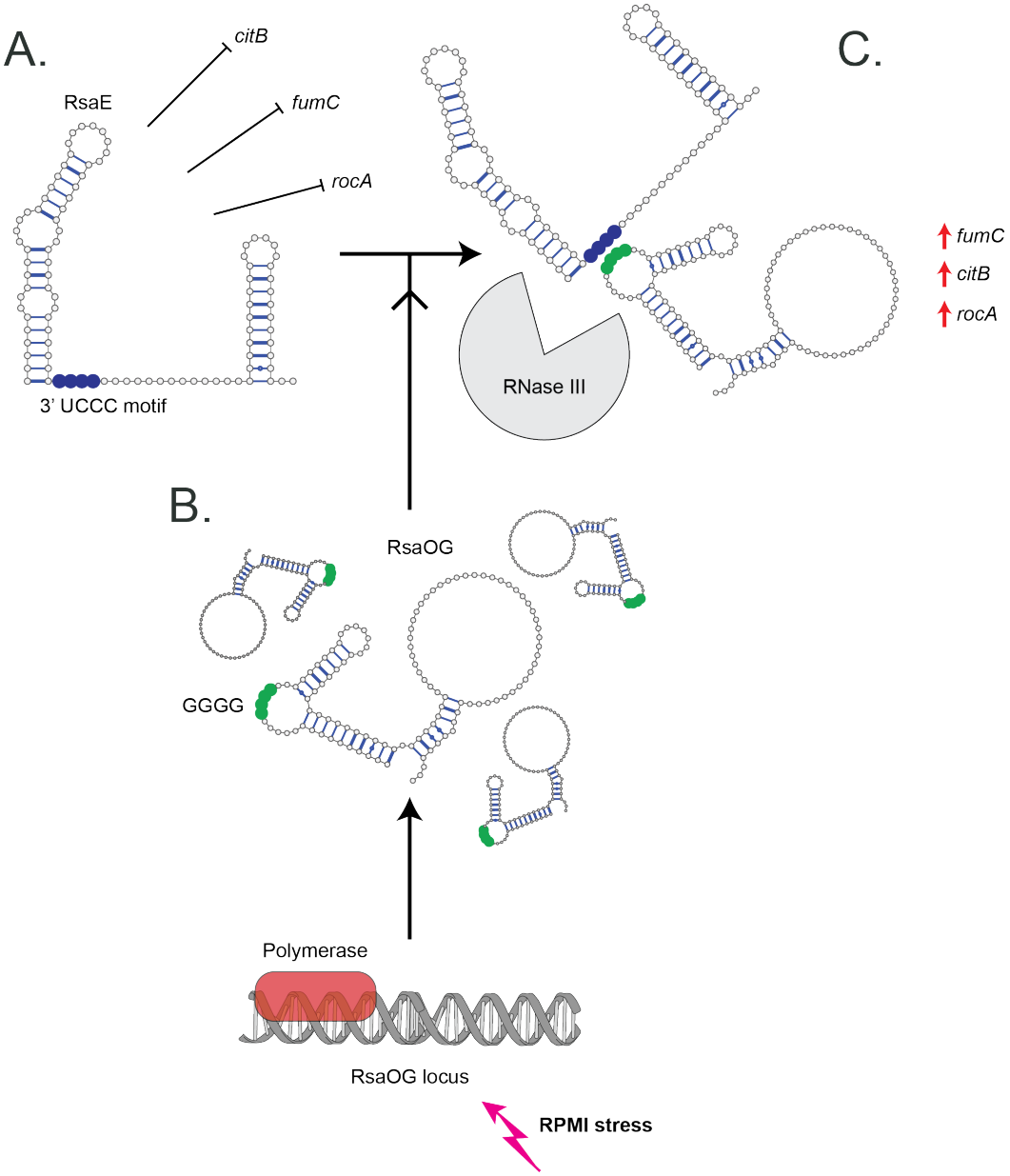

Fig. 6.8: RsaOG sponges RsaE in times of RPMI stress. **A.** Under favourable conditions, RsaE targets transcripts such as *citB, rocA* and *fumC* for degradation. **B.** Upon the imposition of stresses, such as those imposed upon the shift to RPMI medium, RsaOG is upregulated through increased transcription. **C.** The newly transcribed RsaOG molecules bind to RsaE. For simplicity, only the first 3' UCCC motif (blue) is shown to bind a G-rich sequence (green) in RsaOG, as this was uncovered in CLASH. This duplex is then targeted by RNase III, although no significant degradation occurs. The primary outcome of this interaction is that the mRNAs previously being repressed by RsaE are now relieved, and able to be upregulated and predicted to be translated afterwards (indicated by red arrows).

increase in stability before decreasing back down to basal levels by the end of the time course. Although *citB* and *fumC* also exhibit increased rates of transcription, this is not unexpected; transcription and post-transcriptional regulation act in tandem to

shape a response, and commitment to a survival pathway is likely to involve both as avenues of adaptation.

My current working hypothesis is that RsaOG is part of a coherent feed forward loop that enables rapid accumulation of RsaE mRNA targets by inhibiting RsaE activity. Thus, my prediction is that RsaOG expression should have a significant positive impact on the gene expression induction kinetics of RsaE targets. In order to test this model, two key experiments are currently underway. Firstly, a Δ*rsaOG* mutant must be shifted to RPMI medium and then the expression of RsaE targets determined by RT-qPCR. If RsaOG does indeed sponge RsaE away, the lack of RsaOG in this mutant should mean that the targets of RsaE should remain constant or upregulated after the shift to a much lesser degree. The required mutant strain has now been made and this experiment will be performed imminently.

Secondly, in order to examine if this relieving of RsaE-mediated regulation on its targets actually affects protein levels, mass spectroscopy on total cell lysates will be performed after the shift to RPMI. Investigating if proteins such as CitB are upregulated would help to demonstrate that this interaction does indeed dictates changes in the proteome, and this experiment could also be performed on a ΔrsaOG mutant to further prove that the sponging interaction is crucial for this.

RsaE regulates αPSM transcripts

S. aureus utilises the glycolytic, TCA cycle and pentose phosphate pathways in order to derive their required macronutrients (Somerville and Proctor, 2009). In rich medium conditions, the TCA cycle is typically repressed as the cells enter logarithmic growth and glycolysis is utilised as the primary mode of ATP generation (Collins and Lascelles, 1962; Strasters and Winkler, 1963). However, once *S. aureus* has depleted the pool of preferential carbon sources, such as glucose, then the TCA cycle becomes active and upregulated (Somerville et al., 2003).

RsaE is so far the only sRNA known in *S. aureus* to have an intimate relationship with metabolism. As described in Section 6.1, it regulates the TCA cycle purine biosynthesis and amino acid transport and metabolism. This chapter has proposed a novel class of targets for RsaE, the αPSMs. These are critical virulence factors which make up the majority of the secretome of *S. aureus*, and are involved in host cell escape and lysis, and biofilm restructuring. As such, these interactions place RsaE as a core link between metabolism and virulence.

The interaction between RsaE and the two αPSMs was found to require both of RsaE's UCCC motifs. Deletion of just the 5' UCCC motif completely abolished the interaction, despite the fact that all interactions found through CLASH occurred

within the 3' UCCC motif. It may be that the 5' motif is crucial for the initial binding or inducing structural rearrangements within RsaE. This behaviour is in opposition to that observed for the interaction with RsaOG, where deletion of the 5' UCCC motif reduced the binding capacity of RsaE but did not abolish it entirely. This shows that RsaE has different interaction kinetics for its different targets.

It is also observed that the interactions between RsaE and the αPSMs do not appear to be particularly strong *in vitro*, with significant shifts only observed upon addition of a 320-fold excess of RsaE. However, these experiments were performed only with the RNAs. It may be that these interactions require a chaperone *in vivo* to efficiently base pair. In Gram-negative species, the protein Hfq is the major RNA chaperone and is thought to impact the expression of more than 20% of the genome in *E. coli* and *S. enterica* (Chao et al., 2012; Tree et al., 2014). However, the role of Hfq in *S. aureus* is unclear, as discussed in Section 1.9. Another important RNA chaperone in Gram-negative species is ProQ, although a functional counterpart of ProQ does not appear to exist in *S. aureus* (Smirnov et al., 2017). However, this does not eliminate the idea that Hfq could be important in regulating an interaction between RsaE and αPSMs, or that another, unidentified chaperone could play a role.

All of the interactions between RsaE and the αPSMs were discovered in TSB. Thus, it may be that RsaE represses the translation of αPSMs in TSB, a nutrient-rich medium which contains all of its required cofactors. The fact that no interactions were observed in RPMI may hint that the αPSMs are relieved of their repression in this medium, resulting in their translation and secretion. This may be a means of *S. aureus* acquiring nutrients through lysis of surrounding cells.

RPMI and TSB have a similar glucose composition (2 g/L and 2.5 g/L respectively), and the shift was performed at the post-exponential phase when *S. aureus* will have depleted the preferential carbon sources such as glucose. Thus, it is not hypothesised that carbon is the limiting factor in RPMI medium. It may be alternative stresses imposed by the RPMI medium, such as a lack of iron, that induce *S. aureus* to secrete αPSMs in an attempt to lyse nearby cells and scavenge essential metal cofactors. However, this observation does potentially link the translation of virulence factors to the metabolic state of the cell via an sRNA.

In order to further validate this interaction, a collaboration has been set up with Ronan Carroll, where he will test the protein levels of the αPSMs in a ΔrsaE mutant and after overexpression of RsaE. This experiment is also being combined with haemolysis assays to monitor the ability of these strains to induce red blood cell lysis, presumably mediated through PSM activity.

However, other experiments could be performed to further verify the interaction *in vivo*, such as a MAPS-based approach. Although MAPS has already been performed on RsaE, this was carried out at an OD_{600} of 0.5, as opposed to the value of 3 used here. This may explain why the interaction with the αPSMs was not found. Additionally, MAPS could be used to enrich for proteins that bind to RsaE, and mass spectroscopy could identify a potential chaperone.

Another *in vivo* method that could be used is ribosome profiling. This method captures translating ribosomes in vivo, and then examines which RNAs are bound. As such, it provides a direct read out of translation *in vivo*. Ribosomal profiling could be performed after overexpression or deletion of RsaE, and thus investigate if RsaE is able to control translation of the αPSMs as expected. Finally, overexpression of RsaE could be combined with RNAseq in order to investigate if this induces αPSM downregulation, which would imply that RsaE is able to recruit RNase III in order to destroy these mRNAs (RT-qPCR is challenging due to the short length of the PSMs). Overall, these proposed experiments would further solidify the link between RsaE and the αPSMs and investigate the impact at the level of both translation and stability.

6.9 Bibliography

Altuvia S, Locker-Giladi H, Koby S, Ben-Nun O, Oppenheim AB. 1987. RNase III stimulates the translation of the cIII gene of bacteriophage lambda. Proc Natl Acad Sci U S A 84:6511–6515. doi:10.1073/pnas.84.18.6511

Bohn C, Rigoulay C, Chabelskaya S, Sharma CM, Marchais A, Skorski P, Borezée-Durant E, Barbet R, Jacquet E, Jacq A, Gautheret D, Felden B, Vogel J, Bouloc P. 2010. Experimental discovery of small RNAs in *Staphylococcus aureus* reveals a riboregulator of central metabolism. Nucleic Acids Res 38:6620–6636. doi:10.1093/nar/gkq462

Bronesky D, Desgranges E, Corvaglia A, François P, Caballero CJ, Prado L, Toledo-Arana A, Lasa I, Moreau K, Vandenesch F, Marzi S, Romby P, Caldelari I. 2019. A multifaceted small RNA modulates gene expression upon glucose limitation in *Staphylococcus aureus*. EMBO J 38:e99363. doi:10.15252/embj.201899363

Cassat JE, Hammer ND, Campbell JP, Benson MA, Perrien DS, Mrak LN, Smeltzer MS, Torres VJ, Skaar EP. 2013. A Secreted Bacterial Protease Tailors the *Staphylococcus aureus* Virulence Repertoire to Modulate Bone Remodeling during Osteomyelitis. Cell Host Microbe 13:759–772. doi:10.1016/j.chom.2013.05.003

Chao Y, Papenfort K, Reinhardt R, Sharma CM, Vogel J. 2012. An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs. EMBO J 31:4005–4019. doi:10.1038/emboj.2012.229

Cheung GYC, Joo H-S, Chatterjee SS, Otto M. 2014. Phenol-soluble modulins – critical determinants of *Staphylococcal virulence*. FEMS Microbiol Rev 38:698–719. doi:10.1111/1574-6976.12057

Collins FM, Lascelles J. 1962. The effect of growth conditions on oxidative and dehydrogenase activity in *Staphylococcus aureus*. J Gen Microbiol 29:531–535. doi:10.1099/00221287-29-3-531

Court DL, Gan J, Liang Y-H, Shaw GX, Tropea JE, Costantino N, Waugh DS, Ji X. 2013. RNase III: Genetics and function; structure and mechanism. Annu Rev Genet 47:405–431. doi:10.1146/annurev-genet-110711-155618

Durand S, Braun F, Lioliou E, Romilly C, Helfer A-C, Kuhn L, Quittot N, Nicolas P, Romby P, Condon C. 2015. A Nitric Oxide Regulated Small RNA Controls Expression of Genes Involved in Redox Homeostasis in *Bacillus subtilis*. PLOS Genet 11:e1004957. doi:10.1371/journal.pgen.1004957

Felden B, Vandenesch F, Bouloc P, Romby P. 2011. The *Staphylococcus aureus* RNome and Its Commitment to Virulence. PLOS Pathog 7:e1002006. doi:10.1371/journal.ppat.1002006

García-Betancur et al. 2017. Cell differentation defines acute and chronic infection cell types in *Staphylococcus aureus*. eLIFE.

Geissmann T, Chevalier C, Cros M-J, Boisset S, Fechter P, Noirot C, Schrenzel J, François P, Vandenesch F, Gaspin C, Romby P. 2009. A search for small noncoding RNAs in *Staphylococcus aureus* reveals a conserved sequence motif for regulation. Nucleic Acids Res 37:7239–7257. doi:10.1093/nar/gkp668

Grosz Magdalena, Kolter Julia, Paprotka Kerstin, Winkler Ann-Cathrin, Schäfer Daniel, Chatterjee Som Subra, Geiger Tobias, Wolz Christiane, Ohlsen Knut, Otto Michael, Rudel Thomas, Sinha Bhanu, Fraunholz Martin. 2014. Cytoplasmic replication of *Staphylococcus aureus* upon phagosomal escape triggered by phenol-soluble modulin α. Cell Microbiol 16:451–465. doi:10.1111/cmi.12233

Johnson CM, Grossman AD. 2015. Integrative and Conjugative Elements (ICEs): What They Do and How They Work. Annu Rev Genet 49:577–601. doi:10.1146/annurev-genet-112414-055018

Joo HK, Park YW, Jang YY, Lee JY. 2018. Structural Analysis of Glutamine Synthetase from *Helicobacter pylori*. Sci Rep 8:1–8. doi:10.1038/s41598-018-30191-5

Luong TT, Dunman PM, Murphy E, Projan SJ, Lee CY. 2006. Transcription Profiling of the mgrA Regulon in *Staphylococcus aureus*. J Bacteriol 188:1899–1910. doi:10.1128/JB.188.5.1899-1910.2006

Majerczyk CD, Sadykov MR, Luong TT, Lee C, Somerville GA, Sonenshein AL. 2008. *Staphylococcus aureus* CodY Negatively Regulates Virulence Gene Expression. J Bacteriol 190:2257–2265. doi:10.1128/JB.01545-07

Nues R van, Schweikert G, Leau E de, Selega A, Langford A, Franklin R, Iosub I, Wadsworth P, Sanguinetti G, Granneman S. 2017. Kinetic CRAC uncovers a role for Nab3 in determining gene expression profiles during stress. Nat Commun 8:12. doi:10.1038/s41467-017-00025-5

Rochat T, Bohn C, Morvan C, Le Lam TN, Razvi F, Pain A, Toffano-Nioche C, Ponien P, Jacq A, Jacquet E, Fey PD, Gautheret D, Bouloc P. 2018. The conserved regulatory RNA RsaE down-regulates the arginine degradation pathway in *Staphylococcus aureus*. Nucleic Acids Res 46:8803–8816. doi:10.1093/nar/gky584

Schoenfelder SMK, Lange C, Prakash SA, Marincola G, Lerch MF, Wencker FDR, Förstner KU, Sharma CM, Ziebuhr W. 2019. The small non-coding RNA RsaE influences extracellular matrix composition in *Staphylococcus epidermidis* biofilm communities. PLOS Pathog 15:e1007618. doi:10.1371/journal.ppat.1007618

Seidl K, Goerke C, Wolz C, Mack D, Berger-Bächi B, Bischoff M. 2008. *Staphylococcus aureus* CcpA Affects Biofilm Formation. Infect Immun 76:2044–2050. doi:10.1128/IAI.00035-08

Seidl K, Stucki M, Ruegg M, Goerke C, Wolz C, Harris L, Berger-Bächi B, Bischoff M. 2006. *Staphylococcus aureus* CcpA affects virulence determinant production and antibiotic resistance. Antimicrob Agents Chemother 50:1183–1194. doi:10.1128/AAC.50.4.1183-1194.2006

Smirnov A, Wang C, Drewry LL, Vogel J. 2017. Molecular mechanism of mRNA repression in trans by a ProQ-dependent small RNA. EMBO J 36:1029–1045. doi:10.15252/embj.201696127

Somerville GA, Chaussee MS, Morgan CI, Fitzgerald JR, Dorward DW, Reitzer LJ, Musser JM. 2002. *Staphylococcus aureus* aconitase inactivation unexpectedly inhibits post-exponential-phase growth and enhances stationary-phase survival. Infect Immun 70:6373–6382. doi:10.1128/iai.70.11.6373-6382.2002

Somerville GA, Proctor RA. 2009. At the Crossroads of Bacterial Metabolism and Virulence Factor Synthesis in *Staphylococci*. Microbiol Mol Biol Rev 73:233–248. doi:10.1128/MMBR.00005-09

Somerville GA, Saïd-Salim B, Wickman JM, Raffel SJ, Kreiswirth BN, Musser JM. 2003. Correlation of acetate catabolism and growth yield in *Staphylococcus aureus*: implications for host-pathogen interactions. Infect Immun 71:4724–4732. doi:10.1128/iai.71.8.4724-4732.2003

Sonenshein AL. 2005. CodY, a global regulator of stationary phase and virulence in Gram-positive bacteria. Curr Opin Microbiol 8:203–207. doi:10.1016/j.mib.2005.01.001

Strasters KC, Winkler KC. 1963. Carbohydrate metabolism of *Staphylococcus aureus*. J Gen Microbiol 33:213–229. doi:10.1099/00221287-33-2-213

Tree JJ, Granneman S, McAteer SP, Tollervey D, Gally DL. 2014. Identification of bacteriophage-encoded anti-sRNAs in pathogenic *Escherichia coli*. Mol Cell 55:199–213. doi:10.1016/j.molcel.2014.05.006

Wang R, Braughton KR, Kretschmer D, Bach T-HL, Queck SY, Li M, Kennedy AD, Dorward DW, Klebanoff SJ, Peschel A, DeLeo FR, Otto M. 2007. Identification of novel cytolytic peptides as key virulence determinants for community-associated MRSA. Nat Med 13:1510–1514. doi:10.1038/nm1656

Xu T, Wang X-Y, Cui P, Zhang Y-M, Zhang W-H, Zhang Y. 2017. The Agr Quorum Sensing System Represses Persister Formation through Regulation of Phenol Soluble Modulins in *Staphylococcus aureus*. Front Microbiol 8. doi:10.3389/fmicb.2017.02189

Zapf RL, Wiemels RE, Keogh RA, Holzschu DL, Howell KM, Trzeciak E, Caillet AR, King KA, Selhorst SA, Naldrett MJ, Bose JL, Carroll RK. 2019. The Small RNA Teg41 Regulates Expression of the Alpha Phenol-Soluble Modulins and Is Required for Virulence in *Staphylococcus aureus*. mBio 10. doi:10.1128/mBio.02484-18

7 Concluding remarks

The advent of next generation sequencing has revolutionised the field of post-transcriptional regulation. Originally, studies were carried out on a gene-by-gene basis and often required prior knowledge about the biological system of interest. For example, the elucidation of RNAIII as a post-transcriptional regulator in *S. aureus* relied on knowledge of the quorum sensing pathway, detailed structure of the *agr* locus and of its promoters (Novick et al., 1993; Peng et al., 1988; Recsei et al., 1986).

However, immunoprecipitation-based methods now allow us to capture interactomes with relative ease. Proximity-based ligation methods, such as CLASH and related methods (e.g. RIL-seq) are now being increasingly applied to a wider and wider variety of organisms. Such methods were originally applied to human and yeast cells (Helwak et al., 2013, Kudla et al., 2011) but have now been applied to Gram-negative bacteria such as *E. coli* (Iosub et al., 2018; Melamed et al., 2016; Waters et al., 2017). This thesis details the application and optimisation of CLASH to the Gram-positive bacterium *S. aureus*.

It is also noteworthy that interactome-capture methods which forego immunoprecipitation have now been developed and may help to circumvent the main issue of previous techniques; identifying a suitable RBP. These include methods such as SPLASH, PARIS and MARIO (Aw et al., 2016; Lu et al., 2018; Nguyen et al., 2016). Overall, these methods crosslink RNA duplexes together and then capture these on a global scale through bead-based purification or 2-dimensional gel electrophoresis. As such, they eliminate the need for a known RBP. However, one of the main reasons I chose CLASH over these methods is that utilising an RBP allows one to enrich for a specific class of RNA interaction, such as the sRNA-target interactions that were my focus here. Even so, these methods have still proven themselves to be powerful tools in further elucidating RNA interactions on a global scale and offer a different approach in future research.

I suspect that the produced interactomes from this study will have wide-reaching implications for the *Staphylococcal* field. Firstly, they confirm a number of already known interactions and therefore increase our confidence in their veracity. Secondly, many novel interactions have been found and a selection of these validated. Fi-

nally, the data may give leads into identifying novel sRNAs from regions previously annotated as intergenic or UTRs.

In this thesis I have identified and verified several novel interactions between sRNAs and mRNAs. This has increased the depth of our understanding of how particular sRNAs regulates a certain aspect of *S. aureus* biology, such as RsaA and membrane homeostasis. However, it has also directly linked distinct behaviours such as virulence and metabolism, such as via RsaE.

However, the most unique element of the work described here relates to the experiments detailing interactions between sRNAs. These have added to the knowledge of how sRNAs are able to coordinate a multitude of responses through a single binding interaction, inducing myriad downstream repercussions. Additionally, although requiring further verification, this work has hypothesised how these interactions can lead to the destruction of an sRNA population on exquisitely short timescales, such as that seen for RNAIII with RsaA. These experiments are ongoing and will be included in an upcoming publication.

Summary of the main findings and broader implications

The first aim of this work was to optimise CLASH for use in *S. aureus*, which was by no means facile (Chapter 3). Ultimately, most progress was found in reducing the contact time between the immunoprecipitated RNP complex and the cell lysate, thus minimising RNA degradation from cellular RNases. Additionally, when using RNases as bait proteins, this reduction in contact time is likely to have the extra benefit of bringing the experiment into the denaturing conditions of the secondary purification as fast as possible. This subsequently inactivates the RNase and presumably minimises degradation. The produced protocol allowed CLASH to be performed on two RNases; RNase III and RNase Y. Analyses detailed in Chapter 4 showed that this resulted in successful libraries which confirmed many sRNA-target interactions but also discovered novel ones, particularly observed for RNase III. As such, I hypothesise that this optimised CLASH methodology will of great interest to the field and indeed be applicable to any RBP – research in Sander Granneman's lab has already started to apply this protocol to other RNases, metabolic enzymes and RNA modifying enzymes.

I next focused on several validated sRNAs; RsaA, RsaE and RNAIII. This allowed me the luxury of being able to work in relatively well-understood research area which was of great benefit, and I thank the scientists whose previous hard work has greatly aided my own research.

Regarding RsaA and RNAIII, my work has suggesting a balance between these two sRNAs. I hypothesise that these sRNAs interact in the late-exponential phase of growth, where RNAIII may downregulate RsaA's activity as part of its virulent post-transcriptional program. However, this interaction responds to stress and I hypothesise that RsaA is then able to rapidly induce RNAIII destruction. Such a finding is novel for several reasons. First, it may suggest that the cell maintains a pool of RsaA just in case it needs to shut-down RNAIII in response to stress. Secondly, it may show how sRNAs can be the foundation of complex cell behaviours and how their regulation has wide-reaching implications. Finally, it is another example of how bacteria are able to rapidly change their gene expression profiles in response to stress.

My work on RsaA and RNAIII also elucidated novel mRNA targets for both of these sRNAs. Thus, despite their study of many years, it is likely that study of these sRNAs has not been exhausted. Additionally, the discovery that RNAIII (typically associated with inducing host cell lysis and tissue destruction) positives regulates *esxA*, a protein involved in host cell survival, shows how a single RNA can encode for seemingly contradictory behaviours. As such, a gap in our understanding clearly exists and more knowledge on the temporal control of sRNA-mediated regulation may aid in bridging this.

I next detailed the interaction between RsaE and RsaOG. This work adds to the knowledge of sRNA sponging and differs from the interaction between RsaA and RNAIII as it does not involve degradation. As such, this may represent an sRNA-sRNA interaction that the cell expects to be temporary, and so does not induce destruction of any partner.

I also uncovered an interaction between RsaE and several mRNAs encoding for toxins. This is of particular import as it adds to the growing connection between virulence and metabolism and is a direct representation of the link between these two cell pathways. It is unlikely that RsaE will be the only sRNA within *S. aureus* to link these two together and so I expect further study in the field to further this hypothesis.

This study also shows the power of merging different global analysis methods. Here, I used CRAC on RNA polymerase to model nascent transcription, RNA-seq to calculate RNA steady-state levels and CLASH on RNase III to model RNA degradation and detect RNA-RNA interactions. I also intend to augment these analyses with mass spectroscopy at a later date in order to calculate protein expression. As such, combining all of these methods allows one to examine a response to stress on several levels; RNA production, RNA interactions, RNA decay and the cumulative effect of all on the proteome. Such holistic analyses allow us to gain intimate

insight into how cells adapt to their changing environments. However, even this is not exhaustive; techniques such as ribosome profiling are also integrable into these datasets to examine the translatome and CHIP-seq could be used to examine how transcription factors shape the transcriptional response. As such, all the tools required for understanding how cells adapt to change now seem to be available. Our main limiting factors therefore appear to be predictable ones; time and money.

7.1 Bibliography

Aw JGA, Shen Y, Wilm A, Sun M, Lim XN, Boon K-L, Tapsin S, Chan Y-S, Tan C-P, Sim AYL, Zhang T, Susanto TT, Fu Z, Nagarajan N, Wan Y. 2016. In Vivo Mapping of Eukaryotic RNA Interactomes Reveals Principles of Higher-Order Organization and Regulation. Mol Cell 62:603–617. doi:10.1016/j.molcel.2016.04.028

Helwak A, Kudla G, Dudnakova T, Tollervey D. 2013. Mapping the Human miRNA Interactome by CLASH Reveals Frequent Noncanonical Binding. Cell 153:654–665. doi:10.1016/j.cell.2013.03.043

Iosub IA, Marchioretto M, Sy B, McKellar S, Nieken KJ, Nues RW van, Tree JJ, Viero G, Granneman S. 2018. Hfq CLASH uncovers sRNA-target interaction networks enhancing adaptation to nutrient availability. bioRxiv 481986. doi:10.1101/481986

Kudla G, Granneman S, Hahn D, Beggs JD, Tollervey D. 2011. Cross-linking, ligation, and sequencing of hybrids reveals RNA–RNA interactions in yeast. Proc Natl Acad Sci 108:10010–10015. doi:10.1073/pnas.1017386108

Lu Z, Gong J, Zhang QC. 2018. PARIS: Psoralen Analysis of RNA Interactions and Structures with High Throughput and Resolution. Methods Mol Biol Clifton NJ 1649:59–84. doi:10.1007/978-1-4939-7213-5_4

Melamed S, Peer A, Faigenbaum-Romm R, Gatt YE, Reiss N, Bar A, Altuvia Y, Argaman L, Margalit H. 2016. Global Mapping of Small RNA-Target Interactions in Bacteria. Mol Cell 63:884–897. doi:10.1016/j.molcel.2016.07.026

Nguyen TC, Cao X, Yu P, Xiao S, Lu J, Biase FH, Sridhar B, Huang N, Zhang K, Zhong S. 2016. Mapping RNA–RNA interactome and RNA structure in vivo by MARIO. Nat Commun 7:12023. doi:10.1038/ncomms12023

Novick RP, Ross HF, Projan SJ, Kornblum J, Kreiswirth B, Moghazeh S. 1993. Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. EMBO J 12:3967–3975.

Peng HL, Novick RP, Kreiswirth B, Kornblum J, Schlievert P. 1988. Cloning, characterization, and sequencing of an accessory gene regulator (agr) in Staphylococcus aureus. J Bacteriol 170:4365–4372.

Recsei P, Kreiswirth B, O'Reilly M, Schlievert P, Gruss A, Novick RP. 1986. Regulation of exoprotein gene expression in Staphylococcus aureus by agar. Mol Gen Genet MGG 202:58–61. doi:10.1007/bf00330517

Waters SA, McAteer SP, Kudla G, Pang I, Deshpande NP, Amos TG, Leong KW, Wilkins MR, Strugnell R, Gally DL, Tollervey D, Tree JJ. 2017. Small RNA interactome of pathogenic E. coli revealed through crosslinking of RNase E. EMBO J 36:374–387. doi:10.15252/embj.201694639

List of Figures

List of Tables

Colophon

This thesis was typeset with LaTeX 2_ε. It uses the *Clean Thesis* style developed by Ricardo Langner. The design of the *Clean Thesis* style is inspired by user guide documents from Apple Inc.

Download the *Clean Thesis* style at `http://cleanthesis.der-ric.de/`.