UNIVERSITY OF BIRMINGHAM

A research project report submitted by

Kevin Lyons

as part of the requirement for the degree of

MRes in Antimicrobials and Antimicrobial Resistance (AAMR)

This research project was carried out under the supervision of Prof David Grainger and Dr Stephan Heeb

Supported by
wellcome trust

Submitted 30th July 2018

Word Count: 20,289

Table of contents

Acknowledgements

I would like to thank Prof David Grainger and Dr Stephan Heeb for their intellectual guidance and understanding; Ms Rachel Kettles and Dr Manuel Romero for their everyday supervision and practical assistance; and Dr Gemma Warren, Ms Abi Colclough and Ms Maria Papangeli for providing materials, advice and further practical assistance.

I would also like to thank Prof Paul Williams and Prof Ian Henderson for giving me the opportunity to conduct the research reported here; and *The Wellcome Trust* for providing the financial backing, without which none of this would have been possible.

Abstract

The emergence and spread of molecular mechanisms which make bacteria resistant to antibiotics together constitute a significant threat to human health which is global and growing. Two research projects with implications for the understanding of bacterial pathogenicity and antibiotic resistance are presented here. Both projects concern gene regulation in Gram-negative bacteria, and originated from data sets acquired via genome-wide investigations: a ChIP-seq data set and an RNA-seq data set respectively. Results obtained in the first project suggest a negative role for the transcription factor MarA in *Escherichia coli* biofilm formation, through transcriptional activation of the *ycgZ-ymgA-ariR-ymgC* operon; whereas results from the second project suggest that a newly-identified sRNA in *Pseudomonas aeruginosa* binds to the post-transcriptional regulators RsmA and RsmN, potentially promoting the chronic infection lifestyle which is associated with biofilm formation. The relevance of both projects to the problem of antibiotic resistance stems in part from their association with systems capable of regulating the formation of biofilms. Biofilms are formed when communities of microbes attach to surfaces via a self-produced matrix of hydrated extracellular polymeric substances such as polysaccharides, proteins, nucleic acids and lipids. They are typically associated with an inherent increase in resistance to antibiotics and other molecules, and are associated with ~65–80% of all bacterial infections in humans. Understanding these structures, and devising strategies to disrupt them and/or prevent their formation is highly desirable, as they can have implications not only for human health, but also for the efficiency and safety of industrial processes, waste management and food production.

Introduction

In recent years, such a vast number of high-quality literature reviews have been published on the topic of antibiotics – their classification, their history and discovery, their development, their modes of action, their clinical uses and non-clinical uses, their overuse and misuse, their alternatives, the origins and evolution of antibiotic resistance, and the consequences of the emergence and spread of antibiotic-resistant bacteria – that it seems somewhat futile to attempt anything other than a very brief overview here. *Less talk, more action* appears to have become the unofficial motto of researchers, healthcare workers and campaigners currently engaged in the fight against antibiotic resistance; and this report attempts to adhere to that motto, putting the emphasis very much on the *action* (i.e. the original research), rather than spending too much time regurgitating well-worn introductory *talk*. That being said, convention requires that a document such as this begin with a few short sections of preamble, and so it does. However, readers would do well to refer to recent reviews for a more comprehensive analysis (Aminov 2010; Davies and Davies 2010; Allen *et al.* 2014; Fair and Tor 2014; Blair *et al.* 2015; Teillant *et al.* 2015; Simpkin *et al.* 2017).

Antibiotics

The general term *antimicrobial drug* refers to a drug used in the treatment and prevention of microbial infections. More specifically, these drugs can be classified as antibacterial, antifungal, antiviral or antiparasitic drugs based on their activity against bacteria, fungi, viruses and eukaryotic parasites respectively. Over time, the terms *antibacterial drug* and *antibiotic* have become synonymous, both referring to a drug used in the treatment and prevention of bacterial infections. Some commonly used antibiotics are listed in Table 1, along with their targets and known resistance mechanisms.

ANTIBIOTIC CLASS	EXAMPLE(S)	TARGET	MODE(S) OF RESISTANCE
β-Lactams	penicillins (ampicillin), cephalosporins (cephamycin), carbapenems (meropenem), monobactams (aztreonam)	peptidoglycan biosynthesis	hydrolysis, efflux, altered target
Aminoglycosides	gentamicin, streptomycin, spectinomycin, neomycin, kanamycin	translation (30S ribosomal subunit)	phosphorylation, acetylation, nucleotidylation, efflux, altered target
Glycopeptides	vancomycin, teicoplanin	peptidoglycan biosynthesis	reprogramming peptidoglycan biosynthesis
Tetracyclines	tetracycline, minocycline, tigecycline, doxycycline	translation (30S ribosomal subunit)	monooxygenation, efflux, altered target
Macrolides	erythromycin, azithromycin, clarithromycin	translation (50S ribosomal subunit)	hydrolysis, glycosylation, phosphorylation, efflux, altered target
Lincosamides	clindamycin	translation (50S ribosomal subunit)	nucleotidylation, efflux, altered target
Streptogramins	pristinamycin, quinupristin/dalfopristin	translation (50S ribosomal subunit)	C-O lyase (type B streptogramins), acetylation (type A streptogramins), efflux, altered target
Oxazolidinones	linezolid	translation (50S ribosomal subunit)	efflux, altered target
Phenicols	chloramphenicol	translation (50S ribosomal subunit)	acetylation, efflux, altered target
Quinolones	ciprofloxacin, ofloxacin	DNA replication (DNA gyrase and topoisomerase IV)	acetylation, efflux, altered target
Pyrimidines	trimethoprim	folic acid metabolism (dihydrofolate reductase; DHFR)	efflux, altered target
Sulfonamides	sulfamethoxazole, sulfadiazine	folic acid metabolism (dihydropteroate synthase; DHPS)	efflux, altered target
Rifamycins	rifampicin, rifabutin, rifapentine	transcription (RNA polymerase)	ADP-ribosylation, efflux, altered target
Lipopeptides	daptomycin	cell membrane	altered target
Cationic peptides	colistin	cell membrane	altered target, efflux

Table 1 | Simplified summary of some commonly used antibiotics, their targets and known resistance mechanisms. Table adapted from Davies and Davies 2010.

Spectrum of activity

Different antibiotics are used to treat and prevent infections involving different types of bacteria (Lewis 2013; Blair *et al.* 2015). For example, vancomycin, linezolid, daptomycin, and the combination drug quinupristin/dalfopristin are used against Gram-positive bacteria; aminoglycosides are mainly used against Gram-negative bacteria; and chloramphenicol, β-lactams and quinolones are used against both. The spectrum of activity of an antibiotic is dependent on its chemical structure, biostability and mode of action.

Modes of action

Different antibiotics target different components of the bacterial cell (Lewis 2013; Blair *et al.* 2015). For example, β-lactams and glycopeptides target the bacterial cell wall by disrupting peptidoglycan biosynthesis; aminoglycosides, tetracyclines, macrolides, lincosamides, streptogramins, oxazolidinones and phenicols target the ribosome (either the 30S or 50S ribosomal subunit), thereby disrupting protein synthesis; quinolones target the topoisomerase enzymes DNA gyrase and topoisomerase IV, thereby disrupting DNA replication; trimethoprim and sulfamethoxazole target different enzymes involved in the tetrahydrofolate synthesis pathway, thereby inhibiting folic acid metabolism; rifamycins target the RNA polymerase, thereby disrupting transcription; and daptomycin and colistin target the cell membrane, thereby causing membrane depolarisation or lysis respectively. Antibiotics are often described as being either *bacteriostatic* or *bactericidal*, depending on whether their mode of action mediates, respectively, bacterial growth inhibition or bacterial cell death.

Resistance mechanisms

Some bacteria are inherently resistant to one or more antibiotics. For example, they might produce an outer membrane which inhibits penetration of antibiotics into the cell, or express chromosomally-encoded efflux pumps which quickly remove incoming antibiotics from the cell before they can do much damage. Alternatively, and perhaps more worryingly, bacteria can become resistant to antibiotics by mutation, or by the acquisition of a resistance mechanism via horizontal gene transfer. Many *intrinsic* and *acquired* mechanisms of resistance have been reported in the literature, and each can typically be described as adhering to one of a number of general patterns, including enzymatic inactivation or modification of the antibiotic, decreased penetration, increased efflux, target modification, target overproduction, and pathway bypass (Blair *et al.* 2015).

Antibiotic discovery and development

Most antibiotics in use today were isolated by screening soil-derived actinomycetes during the 1940s to 1960s. Unfortunately, the discovery of new antibiotics via this method has since slowed almost to a halt. Recent attempts to establish new platforms for antibiotic discovery – for example, via rational drug design or target-based screens of large compound libraries – have had mixed results (Payne *et al.* 2007; Lewis 2013), suggesting that the principal focus of future research should perhaps be turned from conventional antibiotic development to alternative methods of treatment and prevention (Allen *et al.* 2014).

The problem of antibiotic resistance

The ability to prevent, control and eradicate bacterial infections through the use of antibiotics has revolutionised modern medicine. Now, however, many antibiotics are failing, due to the rapid emergence and spread of antibiotic-resistant bacteria (Blair *et al.* 2015). Antibiotic resistance is a natural phenomenon. However, the production of large quantities of antibiotics by humans, followed by their use and release into the environment has greatly increased the exposure of bacteria to these molecules, thereby promoting the emergence and spread of resistance mechanisms (Davies and Davies 2010). The situation is dire. Around the world, previously-treatable infections are once again becoming deadly, and medical procedures which rely on antibiotics – including cancer chemotherapy, organ transplants, caesarean sections, and hip and knee replacements – cannot, in many cases, be performed without substantial risk of infection (Teillant *et al.* 2015). The problem of antibiotic resistance is compounded by the fact that pharmaceutical companies naturally favour the development and manufacture of highly-profitable drugs – such as those used to treat cancer and chronic illnesses like rheumatoid arthritis, diabetes and asthma – over the development of less-profitable drugs such as antibiotics (Overbye and Barrett 2005; Fair and Tor 2014). Despite the establishment of a myriad number of national and international antibiotic research and development initiatives, the antibiotic development pipeline has slowed to a trickle (Årdal *et al.* 2018). Worse yet, of the few drugs still in development, most will not pass clinical trials, do not target high priority pathogens, and are derived from molecules to which resistance has already emerged (Simpkin *et al.* 2017).

Alternatives to antibiotics

In the face of inevitable and relentless resistance, further conventional antibiotic development seems futile. However, all is not yet lost. A number of radical new therapeutic approaches and preventative measures have emerged as alternatives or ancillaries to antibiotics. These include the use of efflux pump inhibitors, phage therapy, host-directed therapies, predatory bacteria, antimicrobial peptides, bacteriocins, and anti-biofilm materials (Allen *et al.* 2014).

The future of antibiotic resistance

Many of the emerging alternatives to antibiotics appear promising. However, developing these new approaches takes time. In the short term, the use of complementary drug combinations, better hygiene and infection control management in hospitals (Weinstein 2001), and stricter regulations regarding antibiotic availability and disposal could help to minimise the emergence and spread of resistance (Chandy 2008). Also, the roles of education, outreach, and simple-language communication cannot be over-emphasised. Recent surveys conducted by *The Wellcome Trust* and the *World Health Organisation* showed that the concept of antibiotic resistance is widely misunderstood by the public, typically leading them to behave in ways which exacerbate the problem (Mendelson *et al.* 2017). There are no easy answers here, but a targeted multidisciplinary approach involving global monitoring initiatives and co-operation between industry, universities and hospitals, seems like our best chance of avoiding a return to the pre-antibiotic era.

Research projects

This MRes programme requires participants to undertake two 15-week research projects: one at the University of Birmingham, and one at the University of Nottingham.

My first research project was carried out under the supervision of Prof David Grainger at the University of Birmingham and focused on investigating the role of MarA and the *ycgZ-ymgA-ariR-ymgC* (YYAY) operon in *E. coli* biofilm formation. MarA is a transcriptional regulatory protein which has been implicated in resistance to multiple antibiotics via the control of multidrug efflux (McMurry, George and Levy 1994; Zhang, Rosner and Martin 2008; Ruiz and Levy 2010), porin production (Cohen, McMurry and Levy 1988; Chubiz and Rao 2011), DNA repair (Sharma *et al.* 2017) and lipid trafficking (Sharma *et al.* 2017). Before my arrival in the lab, a chromatin immunoprecipitation and sequencing (ChIP-seq) data set had been generated to identify high-affinity MarA-binding sites across the entire genome of the enterotoxigenic *Escherichia coli* (ETEC) strain H10407 (Sharma *et al.* 2017). One of these MarA targets was the P_{ycgZ} promoter, controlling expression of the YYAY operon (Lee *et al.* 2007). Results of crystal violet biofilm formation assays and β-galactosidase assays performed during this project indicate a negative role for MarA in *E. coli* biofilm formation via MarA-mediated transcriptional activation of the YYAY operon. This suggests a model whereby induction of the *marRAB* operon not only leads to increased levels of resistance to multiple antibiotics (George and Levy 1983), but also, perhaps, to reduced biofilm formation. Given that biofilms have been directly associated with bacterial pathogenesis and antibiotic resistance, this finding has potential implications for a better understanding of infections involving *E. coli* and other Gram-negatives.

My second research project was carried out under the supervision of Dr Stephan Heeb at the University of Nottingham and was primarily concerned with determining whether a newly-identified sRNA from *P. aeruginosa* binds to the post-transcriptional regulators RsmA and RsmN. These are RNA-binding proteins which regulate the expression of a large number of *P. aeruginosa* genes (Brencic and Lory 2009). In general terms, RsmA and RsmN act to repress chronic infection and promote acute infection (Brencic and Lory 2009); and regulatory sRNAs, such as RsmY and RsmZ, counteract these functions by binding to them (Brencic *et al.* 2009). Before my arrival in the lab, an RNA sequencing (RNA-seq) data set had been generated to identify RsmN-binding RNAs in *P. aeruginosa* strain PAO1-L (Romero *et al.* 2018). Results of electrophoretic mobility shift assays (EMSAs) performed during this project indicate that the sRNA does indeed bind RsmA and RsmN, indicating a potential role for this sRNA in counteracting their activities under certain conditions, with potential consequences for biofilm formation and antibiotic resistance.

Research Project #1

MarA reduces biofilm formation by activating expression of the *ycgZ-ymgA-ariR-ymgC* operon in *Escherichia coli*

ABSTRACT

The chromosomally-encoded multiple antibiotic resistance (*marRAB*) operon is widespread among enteric bacteria. In *Escherichia coli*, induction of the operon leads to increased expression of MarA, a transcriptional regulatory protein which has been implicated in resistance to multiple antibiotics via the control of multidrug efflux, porin production, DNA repair and lipid trafficking. MarA activates gene expression by using a dual helix-turn-helix motif to bind to an ~20 bp degenerate DNA sequence (marbox) upstream of various genes and operons. A recent chromatin immunoprecipitation and sequencing (ChIP-seq) study identified 28 high-affinity MarA binding sites present in the genomes of both the enterotoxigenic *E. coli* (ETEC) strain H10407 and the laboratory strain *E. coli* K-12. Binding of MarA to one of these targets, a marbox in the P_{ycgZ} promoter region, was further confirmed via electrophoretic mobility shift assays (EMSAs). This promoter controls the expression of the *ycgZ-ymgA-ariR-ymgC* (YYAY) operon, which encodes four small proteins (78–90 aa) with roles in motility, biofilm formation, porin production and acid resistance. However, with the exception of AriR, the exact structures and functions of these proteins remain poorly characterised. Here, crystal violet biofilm formation assays and β-galactosidase assays were used to determine the role of the YYAY operon in biofilm formation and whether or not it is transcriptionally activated by MarA. Three main findings are presented: (i) deletion of the entire YYAY operon leads to increased biofilm formation compared to the wild type; (ii) expression from the P_{ycgZ} promoter is reduced in the absence of the marbox; and (iii) reduced biofilm formation observed in a strain containing multiple plasmid-borne copies of the YYAY operon is dependent on the presence of the marbox. Taken together, these findings suggest a novel role for MarA in addition to its role in mediating resistance to multiple antibiotics via the control of multidrug efflux, porin production, DNA repair and lipid trafficking: namely, reducing biofilm formation via activation of the YYAY operon. This is notable given the clinical roles of biofilms in mediating the persistence of *E. coli* infections and conferring increased resistance to antibiotics.

INTRODUCTION

Biofilms are formed when communities of microbes attach to surfaces via a self-produced matrix of hydrated extracellular polymeric substances (EPS) such as polysaccharides, proteins, nucleic acids and lipids (Flemming and Wingender 2010; Flemming *et al.* 2016). Bacterial biofilms are typically associated with an inherent increase in resistance to disinfectants (Bridier *et al.* 2011), host defences (Roilides *et al.* 2015), and, importantly, antibiotics (Stewart and Costerton 2001; Stewart 2015). According to the USA's *Centers for Disease Control and Prevention* (CDC) and the *National Institutes of Health* (NIH), biofilms are associated with ~65–80% of all bacterial infections in humans (Joo and Otto 2012). Understanding these structures, and devising strategies to disrupt them and/or prevent their formation is highly desirable, as they can have implications not only for human health (Dongari-Bagtzoglou 2008), but also for the efficiency and safety of industrial processes, waste management and food production (Brooks and Flint 2008).

The Gram-negative facultative anaerobic bacterium *Escherichia coli* has been used to study the process of bacterial biofilm formation for decades, due to the fact that it is a generally well-characterized and easily manipulable model organism (Beloin, Roux and Ghigo 2008; Blount 2015). Most importantly, however, it is also a clinically-relevant species (Anderson 2003; Kaper, Nataro and Mobley 2004). For example, it is one of the microbes most frequently associated with urinary tract infections (UTIs) (Soto *et al.* 2006), and the persistence of these infections over time is largely considered to be the result of biofilm formation (Anderson 2003; Soto *et al.* 2006).

The first indication that the *ycgZ-ymgA-ariR-ymgC* (YYAY) operon (Fig. 1) might be involved in *E. coli* biofilm formation came via a transcriptome study, which showed that the genes of the operon are repressed in young (5h or 7h) biofilms, induced in developed (15h) biofilms, and repressed again in mature (24h) biofilms (Domka *et al.* 2007). Since then, three of the four small proteins encoded by the YYAY operon – YmgA, AriR and YmgC – have been more directly implicated in the process (Lee *et al.* 2007; Tschowri, Busse and Hengge 2009). However, given that the YYAY operon is regulated by the nearby divergently-transcribed *bluF* and *bluR* genes – in response to blue light, cold, and starvation signals – the ability of its proteins to mediate their effects appears to be temperature-dependent (Tschowri, Busse and Hengge 2009). Because of this, it is currently somewhat challenging to combine data from investigations carried out at different temperatures into a coherent regulatory narrative; not to mention the fact that, in general, rather little is known about these proteins.

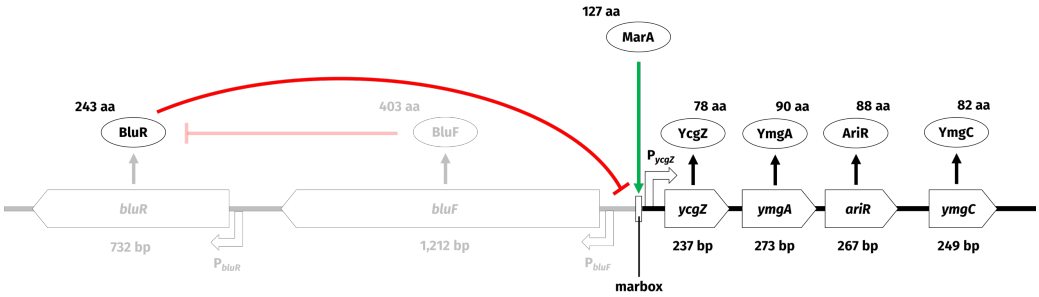

Fig. 1 | **The *E. coli ycgZ-ymgA-ariR-ymgC* (YYAY) operon.** The YYAY operon, the expression of which is controlled by the P_{ycgZ} promoter, encodes four small proteins (78–90 aa) with roles in motility, biofilm formation, porin production and acid resistance (Lee *et al.* 2007; Tschowri, Busse and Hengge 2009; Duval *et al.* 2017). However, with the exception of AriR, the exact structures and functions of these proteins remain poorly characterised. The operon is known to be regulated by the nearby divergently-transcribed *bluF* and *bluR* genes – in response to blue light, cold, and starvation signals (Tschowri, Busse and Hengge 2009); and a recent chromatin immunoprecipitation and sequencing (ChIP-seq) study suggests that MarA also plays a regulatory role in the expression of the operon by binding to a marbox in the P_{ycgZ} promoter region (Sharma *et al.* 2017).

To begin, however, it is worth noting that, at 37°C, knockout mutants of either *ymgA*, *ariR*, or *ymgC* have been previously shown to exhibit elevated levels of biofilm formation and increased motility (Lee *et al.* 2007); increased motility is known to positively influence biofilm formation in *E. coli* (Pratt and Kolter 1998). Hence, it appears that YmgA, AriR and YmgC typically play a negative role in biofilm formation at this temperature. Deletion of *ycgZ* had no effect.

AriR is the best characterised of these proteins (Lee *et al.* 2007; Tschowri, Busse and Hengge 2009). It forms a homodimer with geometry-dependent DNA-binding activity, has structural homology to the gene regulatory protein Hha, and is known to modulate the activity of the RcsC/RcsD/RcsB phosphorelay system at 28°C and 16°C, thereby activating acid resistance genes, activating production of the capsular polysaccharide colanic acid, and downregulating adhesive curli fimbriae (Tschowri, Busse and Hengge 2009). Colanic acid is involved in biofilm maturation (Danese, Pratt and Kolter 2000), whereas curli are involved in the initial surface attachment stage of biofilm formation as well as in biofilm maturation (Beloin, Roux and Ghigo 2008). YmgA is thought to play a similar role to AriR at 16°C, but not at 28°C (Tschowri, Busse and Hengge 2009). Little is known about YmgC.

YcgZ, despite being under the control of the same promoter as YmgA and AriR, appears to partially (or conditionally) counteract their effects (Tschowri, Busse and Hengge 2009). YcgZ is also known to conditionally decrease *ompF* expression (Duval *et al.* 2017), an effect which is typically associated with increased antibiotic resistance, due to the crucial role of OmpF in the passive diffusion of small hydrophilic antibiotics into the cell (Fernandez and Hancock 2012).

It might seem unusual that knockout mutants of *ymgA* and *ariR* exhibit elevated levels of biofilm formation given that they might be expected to produce less of the biofilm matrix component colanic acid, except for the fact that colanic acid production is not essential for the initial surface attachment stage (Danese, Pratt and Kolter 2000). During the early stages of *E. coli* biofilm formation, the initial attachment of cells to a surface is mediated by type I fimbriae, conjugative pili, and curli (Beloin, Roux and Ghigo 2008). As mentioned previously, knockout mutants of *ymgA* and *ariR* exhibit increased motility (Lee *et al.* 2007), and would presumably exhibit increased curli production (Tschowri, Busse and Hengge 2009). Both increased motility and increased curli production are known to have positive impacts on biofilm formation, which might explain the mutant phenotypes.

The principal aims of this work were (i) to establish whether or not the YYAY operon as a whole is involved in *E. coli* biofilm formation; (ii) whether or not the YYAY operon is transcriptionally regulated by the MarA protein; and (iii) if so, whether this MarA-mediated transcriptional regulation affects biofilm formation.

In the time since publication of the study showing that independent deletions of either *ymgA*, *ariR*, or *ymgC* increase levels of biofilm formation (Lee *et al.* 2007), it has been shown that these genes form an operon, with *ycgZ*, which is transcribed as a single ~1.4 kb polycistronic mRNA (Tschowri, Busse and Hengge 2009). Hence, it was resolved here to perform crystal violet biofilm formation assays to investigate the hypothesis that deletion of the entire operon would result in a similar increase in biofilm formation.

A recent chromatin immunoprecipitation and sequencing (ChIP-seq) study identified 28 high-affinity MarA binding sites present in the genomes of both the enterotoxigenic *E. coli* strain H10407 and the laboratory strain *E. coli* K-12 (Sharma *et al.* 2017). Binding of MarA to one of these targets, a marbox in the P_{ycgZ} promoter region, was further confirmed via electrophoretic mobility shift assays (EMSAs). MarA is a 127-amino-acid transcriptional regulator encoded by the chromosomal *marRAB* operon in *E. coli* (Fig. 2). This protein uses a dual helix-turn-helix motif to bind to an ~20 bp degenerate DNA sequence (AYNGCACNNWNNRYYAAAY), known as a marbox, located in the promoter regions of target genes and operons (Martin *et al.* 1999). MarA has been shown to be involved in regulating the expression of >60 chromosomal genes (increasing the expression of ~75%, while decreasing the expression of the other ~25%) (Barbosa and Levy 2000), and is involved in increasing cellular resistance to antibiotics, disinfectants, organic solvents and oxidative stress agents (Randall and Woodward 2002). This increased resistance – known as the *mar* phenotype – is more greatly observed at 30°C compared to 37°C (Maira-Litrán, Allison and Gilbert 2000), is induced by various different

structurally unrelated inducers (e.g. salicylate, chloramphenicol, tetracycline, acetaminophen, sodium benzoate, 2,4-dinitrophenol, cinnamate, carbonyl cyanide *m*-chlorophenylhydrazone, menadione, plumbagin) and is made possible by the decreased influx and increased efflux of specific molecules (Randall and Woodward 2002). For example, MarA downregulates expression of the OmpF porin by upregulating the *micF* antisense RNA, which disrupts translation of the *ompF* mRNA (Cohen *et al.* 1989). In addition, MarA upregulates the *acrAB* genes, which encode components of an efflux pump that promotes the efflux of antibiotics from the cell (Okusu, Ma and Nikaido 1996). Given that MarA is a known transcriptional regulator, and that it appears to bind to the P_{ycgZ} promoter, it was resolved here to investigate the hypothesis that this binding might transcriptionally activate or repress expression of the YYAY operon; and hence, that this MarA-mediated transcriptional regulation of the YYAY operon might impact levels of biofilm formation.

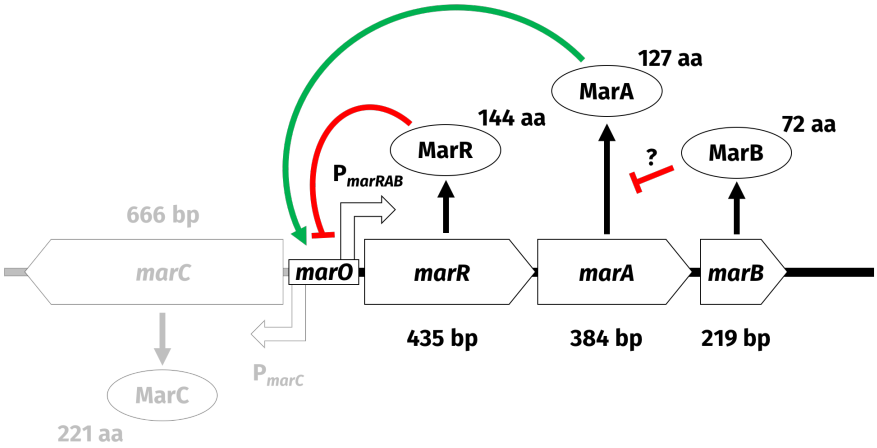

Fig. 2 | The *E. coli mar* locus. The *mar* locus of *E. coli* consists of two divergently-orientated transcriptional units, the *marC* gene and the *marRAB* operon, which share a common operator region, *marO*. The *marRAB* operon is more greatly expressed at 30°C compared to 37°C as measured by experiments with a *marO-lacZ* transcriptional fusion (Seoane and Levy 1995), and has previously been implicated in resistance to antibiotics, disinfectants, organic solvents and oxidative stress agents (Randall and Woodward 2002). The P_{marRAB} promoter is unique among all 2,102 *E. coli* promoters in that it regulates the transcription both of an autoactivator (MarA) and an autorepressor (MarR) (Prajapat, Jain and Saini 2015). This dual regulation enables the operon to exhibit a more diverse range of dynamics than genes controlled by a single regulator. Under non-stress conditions, MarR binds as a homodimer to two sites in P_{marRAB} leading to transcriptional repression of the operon (Martin and Rosner 1995). However, when inducing molecules are present, MarR binds, and is inhibited by, these inducers and the operon is derepressed due to MarA-mediated transcriptional activation (Martin *et al.* 1996). Constitutive expression of the *marRAB* operon has been reported in some clinical isolates, occurring as the result of mutations which prevent MarR-mediated autorepression (Ariza *et al.* 1994). MarB has been reported to reduce *marRAB* expression by reducing the rate of *marA* transcription; however, the exact mechanism is unknown (Vinué, McMurry and Levy 2013). The divergently-transcribed *marC* gene encodes a putative integral membrane protein with no known function in antibiotic resistance (Randall and Woodward 2002).

MATERIALS AND METHODS

Bacterial strains, plasmids and primers

E. coli strains and plasmids used in this work are listed in Table 2. All strains were grown at 37°C in LB broth with 200 rpm shaking, or on LB agar plates, unless otherwise indicated. When required, ampicillin, tetracycline or kanamycin was added to the broth and agar at concentrations of 100 µg/mL, 35 µg/mL and 50 µg/mL respectively, for the maintenance of plasmids.

Name	Genotype/Description	Source
Strains		
E. coli K-12 JCB387	Lac⁻ *nirB::lacZ*	(Page, Griffiths and Cole 1990)
E. coli K-12 JCB387 ΔYYAY	Δ*ycgZ-ymgA-ariR-ymgC* mutant	Rachel Kettles
E. coli K-12 M182 *hns::kan*	*hns::kan* mutant	(Haycocks *et al.* 2015)
E. coli K-12 JCB387 *hns::kan*	*hns::kan* mutant	this work
E. coli T7 Express	*fhuA2* Δλ Lon⁻ OmpT⁻ *lacZ::T7 gene1*	NEB
Plasmids		
pRW50	low-copy-number reporter plasmid with promoterless *lacZ* gene; Tet^R	(Lodge *et al.* 1992)
pRW50-marbox-P_{ycgZ}	pRW50 containing marbox and P_{ycgZ} promoter	Grainger Lab
pRW50-P_{ycgZ}	pRW50 containing P_{ycgZ} promoter	Grainger Lab
pRW50-mutmar-P_{ycgZ}	pRW50 containing mutated marbox, and P_{ycgZ} promoter	Grainger Lab
pBR322	low-copy-number plasmid; Tet^R and Amp^R	(Bolivar *et al.* 1977)
pBR322Δ*bla*	pBR322 derivative lacking most of the *bla* gene	this work
pBR322Δ*bla*-marbox-P_{ycgZ}-YYAY	pBR322Δ*bla* containing marbox, P_{ycgZ} promoter and YYAY operon	this work
pBR322Δ*bla*-P_{ycgZ}-YYAY	pBR322Δ*bla* containing P_{ycgZ} promoter and YYAY operon	this work
pBR322Δ*bla*-mutmar-P_{ycgZ}-YYAY	same as pBR322Δ*bla*-marbox-P_{ycgZ}-YYAY, but marbox is mutated	this work
pET28a	high-level expression plasmid with an IPTG-inducible promoter; Kan^R	Novagen
pET28a-MarA	pET28a with the *E. coli marA* gene cloned downstream of the IPTG-inducible promoter	Grainger Lab

Table 2 | List of bacterial strains and plasmids used in this work.

The enzymatic activity of the TEM-1 β-lactamase encoded by the *bla* gene of the pBR322 plasmid has previously been reported to interfere with peptidoglycan remodelling in pBR322-carrying *Pseudomonas aeruginosa* and *E. coli* strains (Gallant *et al.* 2005). This interference has been shown to disrupt the initial adhesion step of biofilm formation in these strains, leading to lower overall levels of biofilm formation. Disruption of the pBR322 *bla* gene was shown to eliminate this reduction, and hence in this work a variant of pBR322 lacking most of the *bla* gene (i.e. pBR322Δ*bla*) was generated for use in crystal violet biofilm formation assays, and transformants were selected using tetracycline instead of ampicillin.

Plasmid constructions

All novel plasmids used in this work, listed in Table 2, were constructed via conventional cloning methods according to the primers and descriptions in Table 3. For all cloning procedures, amplified PCR products and backbone vectors were digested with restriction enzymes, ligated using T4 DNA ligase, and transformed into *E. coli* cells via the standard heat-shock calcium-chloride method (Nakata, Tang and Yokoyama 1997). All new plasmids were verified before experimental use, via colony PCR and DNA sequencing.

Name	Sequence (5'→3')
To generate short piece of linker DNA for insertion into the AatII/VspI-cleaved pBR322 fragment, to make pBR322Δbla	
pBR322-linker-F	GCATCTATTAATCCAGCACTAACTACGATGCGCAGCGATAAGCAGGTAGGTCAACGTGCGATG CGTCATTCGGATTGCGATTTAGC
pBR322-linker-R	GCATGAGACGTCAAGTAACGATGCTCTGACTCGAAGATAGACTTGTGTTCTCTAAGCTAAATC GCAATCCGAATGACGCATCGCAC
To generate inserts to be cloned into pBR322Δbla, to make pBR322Δbla-marbox-YYAY, pBR322Δbla-YYAY and pBR322Δbla-mutmar-YYAY	
yyay-E-A-reverse	GCATTGGACGTCCTAAGAGAGCACGGATTC
yyay-E-A-comp1	CGCGACGAATTCATATGCATTAGCACTAATTGC
yyay-E-A-comp2	TGTGACGAATTCAATTTATCATTCTGTACACATATTTCGTAC
yyay-E-A-mutated	TGTGACGAATTCATATGCATTACGTCTAATTCGAAAAAATTAATTTATC
To confirm successful generation of pBR322Δbla-marbox-YYAY, pBR322Δbla-YYAY and pBR322Δbla-mutmar-YYAY via sequencing	
pBR322-rev	GCCCGCAGCATCCAGGGTGACGGTGCCGAG
pBR322-check	CTAACTACGATGCGCAGCGATAAGC
To confirm the ΔYYAY deletion via colony PCR	
ycg-for-1	GCTGCTGAATTCATATGCATTAGCACTAATTGCA
ymgC-down	GTGATACAGCTGATGTTTATTCTAAAACCTTACTCAAG
To confirm the hns::kan deletion via colony PCR	
hns-KO-up	AACTGAAAGCTTGTTGGCTGGAGTTTATCATAATTCG
hns-KO-down	AACTGACTCGAGGGTTGAACCGGTTGCTGATGTGACCGC

Table 3 | List of primers used in this work.

Crystal violet biofilm formation assays

A protocol similar to that described by Baugh *et al.* 2014, was used. 5 mL cultures of *E. coli* strains were grown in glass universal tubes at 37°C overnight in LB broth (Lennox) (Sigma-Aldrich, UK) with appropriate selection and 200 rpm shaking. Fresh LB broth was used to dilute these cultures to an optical density (OD) of 0.1 at 600 nm. For every diluted culture, 200 µL volumes were added to four wells in a Corning® Costar® 96-well flat-bottomed cell culture plate (Corning Incorporated, USA). Negative controls were provided by several wells filled with 200 µL volumes of LB broth. The plate was then incubated statically for 48 hr at 30°C or 37°C. All liquid was removed from wells with a multi-channel pipette, and wells were washed three times with water. Plate was left to dry upside-down for 30 minutes in an incubator. Biofilms were stained by adding 200 µL of 1% crystal violet (Sigma-Aldrich, UK) to each well

for 15 min at room temperature. Once again, all liquid was removed from wells with a multi-channel pipette, and wells were washed three times with water. Plate was left to dry upside-down for 30 min in an incubator. Stained biofilms were solubilised by adding 200 µL 70% ethanol to each well, and OD_{600} values were determined using a CLARIOstar® plate reader (BMG Labtech, Germany).

Congo red curli production assays

A protocol similar to that described by Baugh *et al.* 2014, was used. 5 mL cultures of *E. coli* strains were grown in glass universal tubes at 37°C overnight in LB broth with appropriate selection and 200 rpm shaking. These cultures were then diluted 1:10,000 in fresh LB broth without salt, made using tryptone (Sigma-Aldrich, UK), yeast extract (Sigma-Aldrich, UK) and water. 5 µL of each of these 1:10,000 dilutions were pipetted as a spot on LB agar plates without salt, made using tryptone, yeast extract and agar bacteriological (agar no. 1) (Sigma-Aldrich, UK), containing 40 µg/mL Congo red (Sigma-Aldrich, UK). These plates were incubated for 48 hr at 30°C or 37°C.

β-galactosidase assays

A protocol similar to that described by Miller 1972, was used. 5 mL cultures of *E. coli* strains were grown in glass universal tubes at 37°C overnight in LB broth with appropriate selection and 200 rpm shaking. An identical set of glass universal tubes containing media without bacteria was also prepared, and left at 4°C overnight. Next day, these tubes were warmed to 37°C and each inoculated with 200 µL of overnight culture. Bacteria were grown to OD_{650} = 0.3–0.5, via ~1.5 hr incubation at 37°C with 200 rpm shaking. 50 mg *ortho*-Nitrophenyl-β-galactoside (ONPG) and 338 µL β-mercaptoethanol were added to 125 mL Z-buffer (0.01 M KCl, 0.01 M $MgSO_4$, 0.06 M Na_2HPO_4, 0.03 M NaH_2PO_4). OD_{650} values of each culture were recorded, then each culture was lysed with two drops of toluene and two drops of sodium deoxycholate. Tubes were vortexed briefly and aerated at 37°C for 20 min with the lids off to allow toluene evaporation. 100 µL of each lysate were transferred to a separate glass test tube. An extra tube, containing 100 µL LB broth was prepared. 2.5 mL of the Z-buffer-ONPG-β-mercaptoethanol mixture was added to each tube. Tubes were vortexed immediately then incubated at 37°C for 10–20 min until a pale-yellow colour developed. Reactions were stopped by adding 1 mL sodium carbonate (1M) and vortexing again. The OD_{420} value of each reaction was determined, using the tube containing LB broth as a blank, and promoter activity was calculated in Miller units via the conventional formula.

P1 transduction

A protocol similar to that described by Thomason, Costantino and Court 2007, was used. To begin, 5 mL cultures of the donor strain (*E. coli* K-12 M182 *hns*::*kan*), and the two recipient strains (*E. coli* K-12 JCB387 and *E. coli* K-12 JCB387 ΔYYAY) were grown in glass universal tubes at 37°C overnight in LB broth with appropriate selection and 200 rpm shaking.

Preparation of lysate

50 µL of *E. coli* K-12 M182 *hns*::*kan* overnight was added to 5 mL fresh LB broth with 50 µL 20% glucose and 25 mL 1M $CaCl_2$ and incubated for 30–45 min at 37°C with 200 rpm shaking. 100 µL of P1*vir* stock was added and incubation continued for a further 3 hr until the culture lysed. A few drops of chloroform were added and incubation continued for ~5 min. The supernatant was transferred to multiple 1.5 mL microcentrifuge tubes. Tubes were centrifuged at 4°C for 10 min at ~9,200 x *g*. Finally, the supernatant (i.e. the P1 lysate) was transferred into a glass bijou and stored at 4°C.

P1 transduction

50 µL of *E. coli* K-12 JCB387 and *E. coli* K-12 JCB387 ΔYYAY overnights were added to separate 5 mL volumes of fresh LB broth and incubated at 37°C with 200 rpm shaking until the lysate was ready (~4 hr). 1.5 mL of each of these cultures were then centrifuged at top speed for 2 min. The supernatant was discarded and the cells re-suspended in 750 µL P1 salts (10 mM $CaCl_2$/5 mM $MgSO_4$). 100 µL of this mixture was added to varying amounts of lysate in glass bijous (i) 'no lysate' (negative control), (ii) 1 µL lysate, (iii) 10 µL lysate, (iv) 100 µL lysate. Phage was allowed to adsorb for 30 min at 37°C. 1 mL LB broth and 200 µL 1M sodium citrate were added, and incubation continued for 1 hr. Each mixture was centrifuged in microcentrifuge tubes at top speed for 2 min. The supernatant was discarded, the cells re-suspended in 100 µL LB broth and plated on LB agar plates containing kanamycin (50 µg/mL) and sodium citrate (5 mM). Successful introduction of the *hns*::*kan* deletion into *E. coli* K-12 JCB387 and *E. coli* K-12 JCB387 ΔYYAY was confirmed by colony PCR using MyTaq™ Red 2X Mix (Bioline, UK).

Statistics

Unpaired T tests for biofilm formation assays and β-galactosidase assays were performed using the QuickCalcs online tool (https://www.graphpad.com/quickcalcs/ttest1) (GraphPad Software). For the purposes of this work, *p*-values less than 0.05 were considered indicative of a statistically significant difference.

RESULTS

Deletion of the YYAY operon increases biofilm formation at 30°C

In the time since publication of the study showing that independent deletions of either *ymgA*, *ariR*, or *ymgC* increase levels of biofilm formation (Lee *et al.* 2007), it has been shown that these genes form an operon, with *ycgZ*, which is transcribed as a single ~1.4 kb polycistronic mRNA (Tschowri, Busse and Hengge 2009). Hence, it was resolved to determine here, for the first time, the overall impact of this operon on biofilm formation at 30°C by conducting crystal violet biofilm formation assays with the *E. coli* K-12 JCB387 wild type strain and a mutant lacking the entire YYAY operon. At 30°C, this ΔYYAY mutant exhibited significantly increased levels of biofilm formation compared to the wild type, indicating that the operon has a negative impact on biofilm formation at this temperature (Fig. 3).

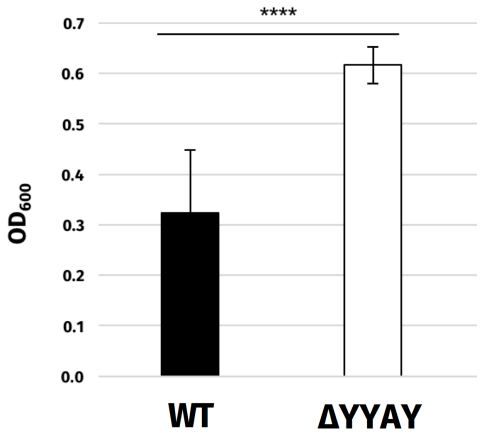

Fig. 3 | Deletion of the YYAY operon appears to increase levels of biofilm formation at 30°C, as measured by crystal violet biofilm formation assays. Data for the two *E. coli* K-12 JCB387 strains presented here are representative of three independent experiments, each involving two independent cell cultures of each strain, wherein four technical replicates of each independent cell culture were processed. This resulted in a data set of twenty-four independent readings for each strain. Columns indicate the mean of each data set, while error bars show the standard deviation. **** = P ≤ 0.0001.

Expression from the P_{ycgZ} promoter appears to be reduced in the absence of the marbox

A recent chromatin immunoprecipitation and sequencing (ChIP-seq) study identified 28 high-affinity MarA binding sites present in both the enterotoxigenic *E. coli* strain H10407 and the laboratory strain *E. coli* K-12 (Sharma *et al.* 2017). Binding of MarA to one of these targets, a marbox in the P_{ycgZ} promoter region, was further confirmed via electrophoretic mobility shift assays (EMSAs). Here, β-galactosidase assays were used to determine whether binding of MarA to the marbox in the P_{ycgZ} promoter region transcriptionally activates expression from the P_{ycgZ} promoter (Fig. 4). Expression from the P_{ycgZ} promoter appeared to be reduced in the absence of the marbox, suggesting that MarA does indeed activate transcription from the P_{ycgZ} promoter.

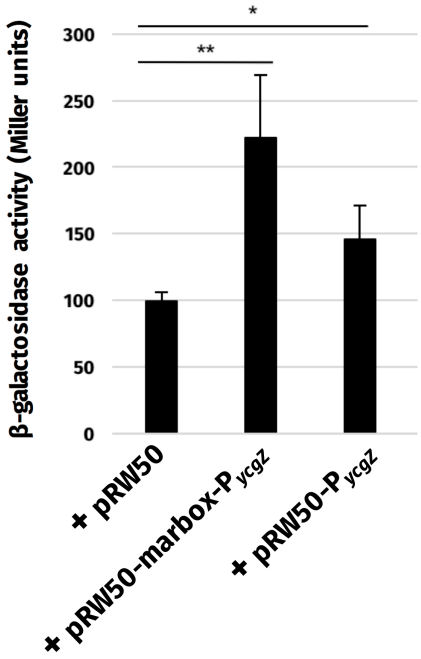

Fig. 4 | Expression from the P_{ycgZ} promoter, as measured by β-galactosidase assays, appears to be reduced in the absence of the marbox. When cloned upstream of the promoterless *lacZ* gene on the pRW50 plasmid, a short DNA fragment containing the upstream marbox and the P_{ycgZ} promoter significantly increases the expression of *lacZ* in *E. coli* T7 Express. Deletion of the marbox from this fragment appears to reduce this effect, suggesting that MarA transcriptionally activates the P_{ycgZ} promoter. The pRW50-marbox-P_{ycgZ} plasmid carries a short DNA fragment containing the P_{ycgZ} promoter and its upstream marbox, and the pRW50-P_{ycgZ} plasmid carries a shorter DNA fragment containing only the P_{ycgZ} promoter (no marbox). Data for each strain presented here represent the mean activities of three biological replicates. Columns indicate the mean of each data set, while error bars show the standard deviation. * = $P \leq 0.05$, ** = $P \leq 0.01$.

The negative role of the YYAY operon in biofilm formation at 30°C is dependent on the presence of the upstream marbox

Having established the negative role of the YYAY operon in biofilm formation at 30°C, and the possibility that MarA transcriptionally activates this operon, it was resolved to investigate whether MarA-mediated activation of the operon has any role in biofilm formation. The biofilm formation capabilities of three strains carrying derivatives of the pBR322Δ*bla* plasmid were tested, and compared to wild type levels and the levels exhibited by a strain carrying the pBR322Δ*bla* plasmid with no insert. The introduction of multiple plasmid-borne copies of the YYAY operon (along with its upstream marbox and P_{ycgZ} promoter) significantly reduced biofilm formation (Fig. 5). This effect was abolished in the absence of the marbox. These findings, combined with observations from the β-galactosidase assays, suggest that the negative role of the YYAY operon in biofilm formation at 30°C is dependent on binding of MarA to the marbox and subsequent MarA-mediated transcriptional activation of the YYAY operon.

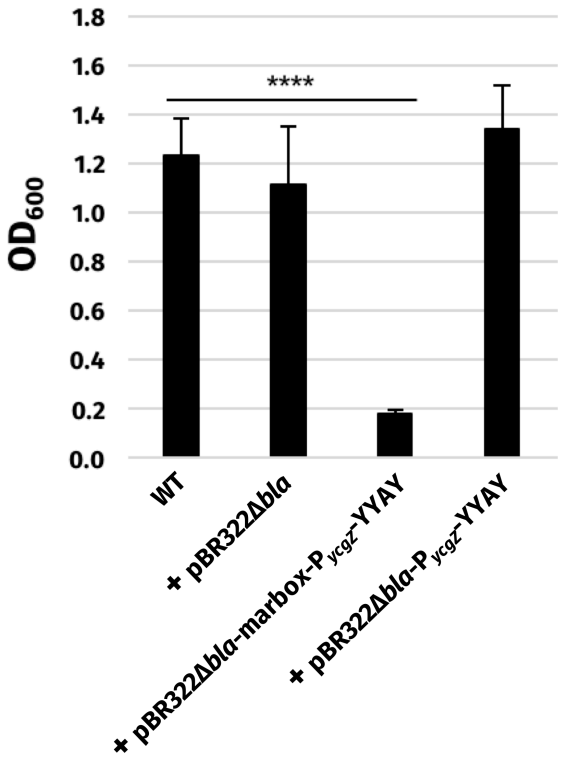

Fig. 5 | The negative role of the YYAY operon in biofilm formation at 30°C, as measured by crystal violet biofilm formation assays, is dependent on the presence of the upstream marbox. When cloned onto the pBR322Δ*bla* plasmid, a DNA fragment containing the upstream marbox, the P_{ycgZ} promoter and the YYAY operon significantly reduces biofilm formation in *E. coli* K-12 JCB387. Deletion of the marbox from this fragment restores biofilm formation to wild type levels, suggesting that the negative role of the YYAY operon in biofilm formation at 30°C is dependent on the presence of the upstream marbox. The pBR322Δ*bla*-marbox-P_{ycgZ}-YYAY plasmid carries a DNA fragment containing the upstream marbox, the P_{ycgZ} promoter and the YYAY operon, and the pBR322Δ*bla*-P_{ycgZ}-YYAY plasmid carries a shorter DNA fragment containing only the P_{ycgZ} promoter and the YYAY operon (no marbox). Data for each strain presented here are representative of three independent experiments, each involving two independent cell cultures of each strain, wherein four technical replicates of each independent cell culture were processed. This resulted in a data set of twenty-four independent readings for each strain. Columns indicate the mean of each data set, while error bars show the standard deviation. **** = $P \leq 0.0001$.

The role of MarA in regulating the YYAY operon was first investigated, in this work, through the use of crystal violet biofilm formation assays performed at 30°C with the *E. coli* K-12 JCB387 wild type strain and a ΔYYAY mutant (Fig. 3). These two strains were then transformed with the pBR322Δ*bla* plasmid and three pBR322Δ*bla* derivatives, and the biofilm formation capabilities of these transformants were determined and compared to those of the initial two strains (Fig. 6).

Deletion of the YYAY operon increased levels of biofilm formation at 30°C, and, as expected, transformation of wild type cells or the ΔYYAY mutant with the empty pBR322Δ*bla* plasmid had no significant effect on biofilm formation (P > 0.05). However, when a DNA fragment containing the upstream marbox, the P_{ycgZ} promoter and the YYAY operon was cloned onto this plasmid, a significant reduction in biofilm formation was observed in both the wild type and the ΔYYAY mutant. Deletion of the marbox from this fragment partially or completely restored biofilm formation in both strains, suggesting that the negative role of the YYAY operon in biofilm formation at 30°C is dependent on the presence of the upstream marbox. Finally, a fragment containing a supposedly inactive mutated version of the upstream marbox was associated with intermediate biofilm formation levels that were relatively lower for the wild type strain than for the ΔYYAY mutant; perhaps indicating that the mutated marbox is not entirely inactive.

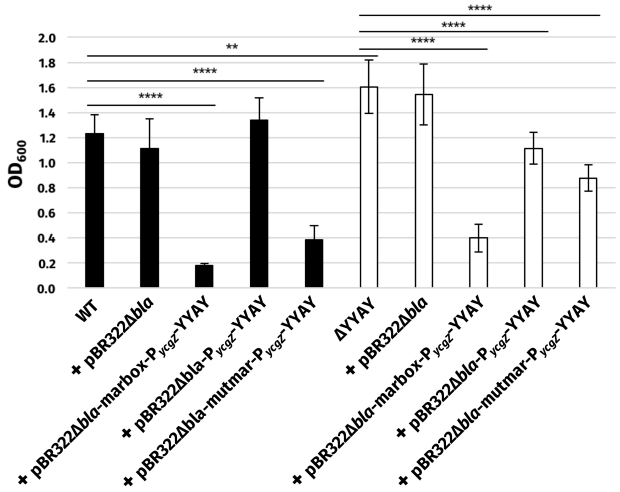

Fig. 6 | Biofilm formation levels of *E. coli* K-12 JCB387 and ΔYYAY strains at 30°C are reduced by the introduction of multiple plasmid-borne copies of the YYAY operon only when the upstream marbox is also included. *E. coli* K-12 JCB387 strains are shown in black, whereas *E. coli* K-12 JCB387 ΔYYAY strains are shown in white. The modified pBR322Δ*bla* plasmid was used here, instead of pBR322, due to reasons given in *Materials and Methods*. The pBR322Δ*bla*-marbox-P_{ycgZ}-YYAY plasmid carries a DNA fragment containing the upstream marbox, the P_{ycgZ} promoter and the YYAY operon; the pBR322Δ*bla*-P_{ycgZ}-YYAY plasmid carries a shorter DNA fragment containing only the P_{ycgZ} promoter and the YYAY operon (no marbox); and the pBR322Δ*bla*-mutmar-P_{ycgZ}-YYAY plasmid carries a DNA fragment containing a supposedly inactive mutated version of the upstream marbox, the P_{ycgZ} promoter and the YYAY operon. Transformation of wild type cells or the ΔYYAY mutant with the empty pBR322Δ*bla* plasmid had no significant effect on biofilm formation (P > 0.05). Data for each strain presented here are representative of three independent experiments, each involving two independent cell cultures of each strain, wherein four technical replicates of each independent cell culture were processed. This resulted in a data set of twenty-four independent readings for each strain. Columns indicate the mean of each data set, while error bars show the standard deviation. ** = P ≤ 0.01, **** = P ≤ 0.0001.

Lower overall levels of biofilm formation are observed for the *E. coli* K-12 JCB387 and ΔYYAY strains at 37°C compared to 30°C

Overall levels of biofilm formation were found to be lower for all strains at 37°C compared to 30°C (Fig. 7). Unfortunately, this made it difficult to observe significant differences in biofilm formation levels between the various plasmid-carrying strains.

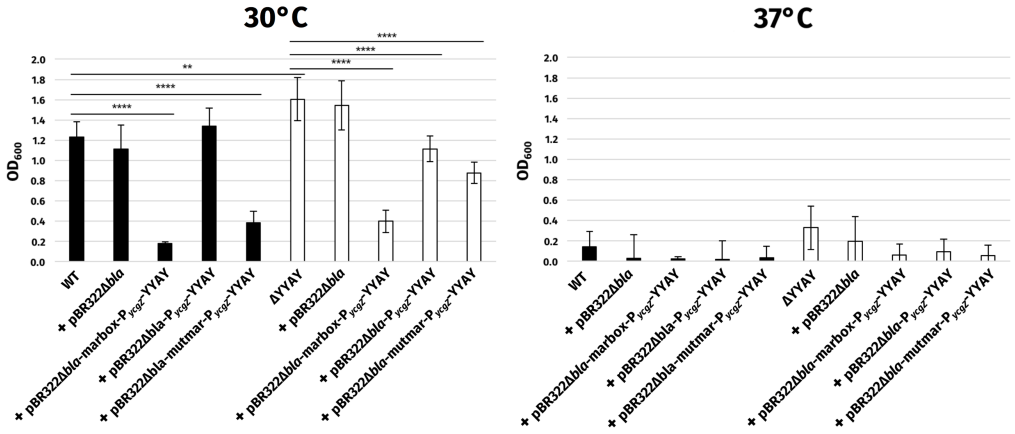

Fig. 7 | Lower overall levels of biofilm formation are observed for the *E. coli* K-12 JCB387 and ΔYYAY strains at 37°C compared to 30°C. These low levels of biofilm formation resulting from crystal violet biofilm formation assays performed at 37°C made it difficult to observe significant differences in biofilm formation levels between the various plasmid-carrying strains. *E. coli* K-12 JCB387 strains are shown in black, whereas *E. coli* K-12 JCB387 ΔYYAY strains are shown in white. Data for each strain presented here at 30°C are representative of three independent experiments, each involving two independent cell cultures of each strain, wherein four technical replicates of each independent cell culture were processed. This resulted in a data set of twenty-four independent readings for each strain. Data for each strain presented here at 37°C are representative of two independent experiments, each involving two independent cell cultures of each strain, wherein four technical replicates of each independent cell culture were processed. This resulted in a data set of sixteen independent readings for each strain. Columns indicate the mean of each data set, while error bars show the standard deviation. ** = P ≤ 0.01, **** = P ≤ 0.0001.

H-NS may have some role in regulating YYAY expression

Deletion of the YYAY operon significantly increased biofilm formation at 30°C, but had no significant effect at 37°C (Fig. 8). This suggests that the operon plays a more important role at the lower temperature, as expected. Deletion of *hns* from *E. coli* K-12 strains typically results in decreased biofilm formation (Belik, Tarasova and Khmel' 2008; Hong, Wang and Wood 2010). Here, this effect was observed at 37°C, but not at 30°C. Deletion of both the YYAY operon and the *hns* gene together, did not significantly affect biofilm formation levels at either temperature. The reason for this is unknown.

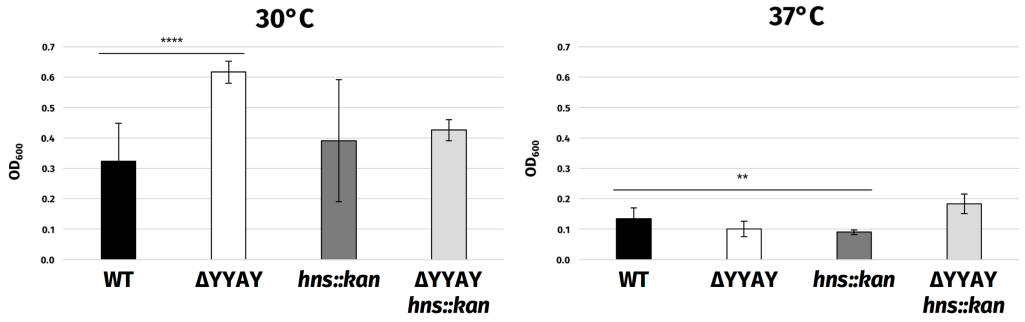

Fig. 8 | Results obtained from crystal violet biofilm formation assays performed with a *hns::kan* mutant and a ΔYYAY *hns::kan* double mutant were difficult to interpret. Whereas deletion of the YYAY operon appears to increase biofilm formation at 30°C, this effect is not observed at 37°C. Deletion of *hns* resulted in decreased biofilm formation at 37°C, but not at 30°C. Deletion of both the YYAY operon and the *hns* gene gave rise to biofilm formation levels which were similar to levels observed for wild type *E. coli* K-12 JCB387. Data for each strain presented here are representative of two independent experiments, each involving two independent cell cultures of each strain, wherein four technical replicates of each independent cell culture were processed. This resulted in a data set of sixteen independent readings for each strain. Columns indicate the mean of each data set, while error bars show the standard deviation. ** = $P \leq 0.01$, **** = $P \leq 0.0001$.

DISCUSSION

In this research project, crystal violet biofilm formation assays and β-galactosidase assays were used to investigate the role of MarA and the YYAY operon in *E. coli* biofilm formation. Three main findings are presented: (i) deletion of the entire YYAY operon leads to increased biofilm formation compared to the wild type; (ii) expression from the P_{ycgZ} promoter is reduced in the absence of the marbox; and (iii) reduced biofilm formation observed in a strain containing multiple plasmid-borne copies of the YYAY operon is dependent on the presence of the marbox. Overall, these findings suggest a novel role for MarA in addition to its role in mediating resistance to multiple antibiotics via the control of multidrug efflux, porin production, DNA repair and lipid trafficking (Cohen *et al.* 1989; Okusu, Ma and Nikaido 1996; Sharma *et al.* 2017): namely, reducing biofilm formation via activation of the YYAY operon.

During this project, crystal violet biofilm formation assays were routinely performed at 30°C. This was considered appropriate due to the fact that lower temperatures tend to increase the expression of biofilm-associated genes in *E. coli* K-12 (White-Ziegler *et al.* 2008), because MarA is thought to be expressed at greater levels at this temperature than at 37°C (Seoane and Levy 1995), and because BluR is thought to strongly repress the YYAY operon at 37°C (Duval *et al.* 2017).

The *E. coli* proteins SoxS and Rob have regulatory effects which overlap with those of MarA, and also bind to marbox sequences (Martin and Rosner 2002). In fact, most known marboxes preferentially bind SoxS or Rob (Martin *et al.* 2002). However, given that the MarA ChIP-seq study identified only those *in vivo* targets where binding of MarA is not occluded by other proteins, it seems unlikely that *soxS* and *rob* play any major role in regulating the activity of the P_{ycgZ} promoter. Nevertheless, perhaps it would be worth verifying this by repeating some of the experiments presented here with *soxS* and *rob* mutants, in order to determine with greater certainty that it is indeed MarA which plays the central role. In addition, it may also be worthwhile to repeat some of these experiments with strains in which MarA is overexpressed. If the MarA-mediated effects described here are accurate, overexpressing MarA would perhaps be expected to augment them, thereby providing further evidence of their validity. For example, if MarA truly does reduce biofilm formation via transcriptional activation of the YYAY operon, then overexpressing MarA might perhaps be expected to decrease biofilm formation in wild type cells, but not in a ΔYYAY mutant.

It has previously been shown that low temperatures increase expression of biofilm-, cold-shock- and RpoS-dependent genes in *E. coli* K-12 (White-Ziegler *et al.* 2008). However, something which becomes clear from a review of the relevant literature is that many studies conduct biofilm formation assays, gene expression analyses

and other investigations at different temperatures. This can make comparisons between studies difficult, since YYAY expression has been shown to be temperature-dependent (Tschowri, Busse and Hengge 2009). For example, YcgZ is upregulated 46-fold at 25°C compared to 37°C, and mutation of the YYAY-repressor gene *bluR* upregulates YYAY expression 100-fold at 37°C, but only 4-fold at 25°C, indicating strong BluR-mediated repression of YYAY at 37°C but not at 25°C (Duval *et al.* 2017). The global DNA-binding transcriptional regulator and nucleoid-associated protein H-NS has also been shown via ChIP-seq to bind and possibly repress the AT-rich YYAY operon, and particularly the *ariR* gene, at 37°C (Kahramanoglou *et al.* 2011). If the ultimate objective is to build a coherent regulatory narrative of the MarA and YYAY systems, then greater attention should be paid to these temperature differences in future work.

Porin downregulation is a trend typically associated with increased antibiotic resistance due the crucial role of porins such as OmpF in the passive diffusion of small hydrophilic antibiotics into the cell (Fernandez and Hancock 2012). MarA has previously been shown to downregulate expression of the OmpF porin in *E. coli* by upregulating the *micF* antisense RNA, which disrupts translation of the *ompF* mRNA (Cohen *et al.* 1989). Independent expression of YcgZ is also known to downregulate expression of the OmpF porin via an unknown mechanism (Duval *et al.* 2017). Given that the work presented here appears to suggest that MarA transcriptionally activates the YYAY operon, it is at least conceivable that under some conditions MarA might increase YcgZ levels in such a way as to downregulate OmpF expression, and hence antibiotic susceptibility, via this pathway. Considering this possibility, and the clinical roles of biofilms in mediating the persistence of *E. coli* infections and conferring increased resistance to antibiotics, future work in this area should perhaps test the antibiotic resistance profiles of the strains used here, and make use of known pathogenic *E. coli* strains and clinical isolates to determine whether the findings presented here can be replicated in clinically-relevant strains. Ultimately, investigations such as this should be carried out with a view to developing a better understanding of *E. coli* pathogenesis. Hence, if/when a role for MarA in biofilm formation has been established in pathogenic *E. coli* strains, steps should then of course be taken to understand the impact of such a finding in a biological context using relevant infection models.

Given the clinical relevance of biofilms, any system which contributes to or reduces their formation is of potential therapeutic interest, as such systems may constitute future targets for the treatment of biofilm infections through the use of biofilm-matrix-degrading enzymes, quorum-sensing signals, surfactants and small molecule inhibitors (Kaplan 2010). The MarA-mediated *mar* phenotype, involved in increasing cellular resistance to antibiotics, disinfectants, organic

solvents and oxidative stress agents, is induced by various different structurally unrelated inducers (e.g. salicylate, chloramphenicol, tetracycline, acetaminophen, sodium benzoate, 2,4-dinitrophenol, cinnamate, carbonyl cyanide *m*-chlorophenylhydrazone, menadione, plumbagin) (Randall and Woodward 2002). In broad terms, this phenotype is generally induced as a response to unfavourable conditions. Phenomena such as reduced biofilm formation and increased biofilm dispersal also occur in response to unfavourable conditions, such as nutrient limitation (Kaplan 2010). If the observations reported here are accurate, the role of MarA in reduced biofilm formation might be considered a new feature of the *mar* phenotype, such that upon induction of the *marRAB* operon not only do cells become increasingly resistant to antibiotics, disinfectants, organic solvents and oxidative stress agents, but they also become less likely to form biofilms, perhaps enabling them to disperse to different locations where conditions for biofilm formation are more favourable.

SUPPLEMENTARY RESULTS AND DISCUSSION

Does H-NS regulate YYAY expression?

The YYAY operon is not thought to be expressed at very high levels at 37°C (Duval *et al.* 2017). An RNA-seq study carried out at this temperature revealed only a very small number of cDNA reads mapping to *ycgZ* and *ymgA*, and none to either *ariR* or *ymgC* (Grainger *et al.*, unpublished). This is probably largely the result of BluR-mediated repression of the YYAY operon at this temperature (Duval *et al.* 2017). However, the global DNA-binding transcriptional regulator and nucleoid-associated protein H-NS has been shown via ChIP-seq to also bind and possibly repress the AT-rich YYAY operon, and particularly the *ariR* gene, at 37°C (Kahramanoglou *et al.* 2011) (Fig. 9). Hence, in an attempt to better understand the effect of H-NS on YYAY expression, it was resolved to repeat all crystal violet biofilm formation assays from this work with a *hns* deletion mutant. However, this approach is somewhat problematic given how remarkably pleiotropic the *hns* deletion is in *E. coli* (Yamada *et al.* 1991). In short, H-NS regulates such a vast array of genes, that the effects on biofilm formation of any change in YYAY expression occurring in the *hns* mutant would likely not be easily identifiable among the multitude of other changes. Hence, observations from the assays with *hns* mutants are likely of limited interest. Perhaps a more specific approach would have been better here: for example, investigating the role of H-NS in the expression of a P_{ycgZ} transcriptional fusion.

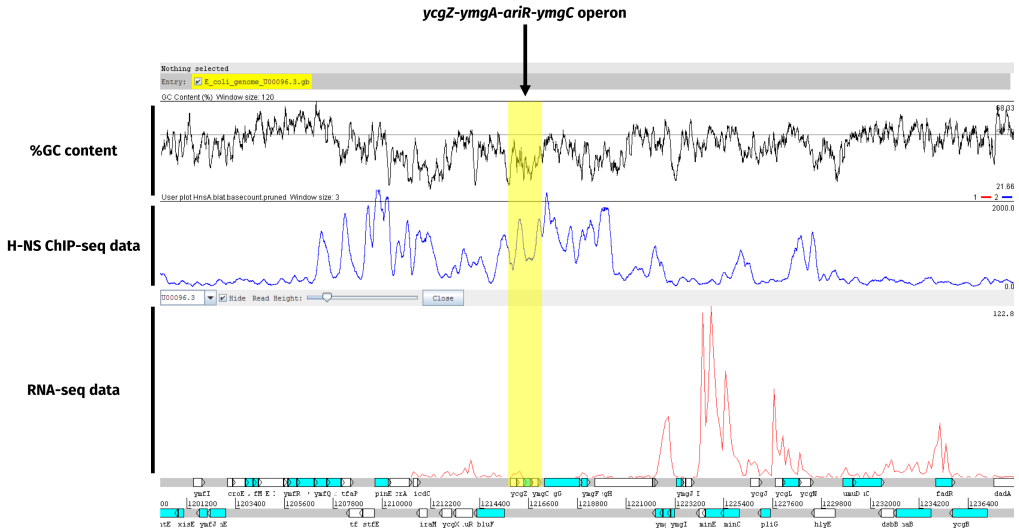

Fig. 9 | According to ChIP-seq data (blue) from Kahramanoglou *et al.* 2011, the AT-rich YYAY operon is bound by H-NS in *E. coli* K-12 cells grown at 37°C. H-NS-mediated repression of the YYAY operon at this temperature may in part explain why only a very small number of sequencing reads mapping to *ycgZ* and *ymgA* were observed in RNA-seq data from a separate investigation (red), and none to either *ariR* or *ymgC* (Grainger *et al.*, unpublished).

Crystal violet biofilm formation assays with strains carrying the pBR322Δ*bla*-derived plasmids appeared to suggest some role for H-NS in regulating expression of the YYAY operon, as differences in biofilm formation levels were detected between strains (Fig. 10). However, at this stage, no serious conclusions can be drawn.

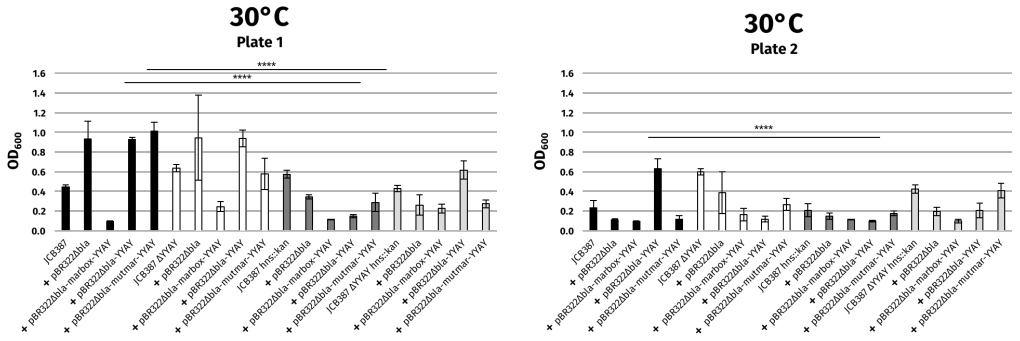

Fig. 10 | **Preliminary data suggest some role for H-NS in regulating the expression of the YYAY operon.** However, for unknown reasons, two identical preliminary crystal violet biofilm formation assays conducted with the same strains on two different 96-well plates yielded somewhat different results (as shown above). Hence, no serious conclusions can be drawn and the precise nature of this role remains unknown. Data for each strain presented here are representative of a two independent experiments, each involving two independent cell cultures of each strain, wherein four technical replicates of each independent cell culture were processed. This resulted in a data set of sixteen independent readings for each strain. Columns indicate the mean of each data set, while error bars show the standard deviation. **** = P ≤ 0.0001.

Congo red curli production assays

The AriR and YmgA proteins of the YYAY operon have previously been implicated in downregulating the production of curli fimbriae (Tschowri, Busse and Hengge 2009). These are extracellular structures required for biofilm maturation (Kikuchi *et al.* 2005). Here, Congo red curli production assays were used in an attempt to determine whether deletion of the entire YYAY operon and/or *hns* affects curli production at 30°C or 37°C (Fig. 11). Congo red curli production assays were also performed with the wild type and ΔYYAY mutant strains carrying the pBR322Δ*bla*-derived plasmids (not shown), but no obvious differences were observed between strains. In retrospect, this experimental approach was inherently flawed since curli are typically only produced at temperatures below 30°C in *E. coli* K-12 strains such as JCB387 (Kikuchi *et al.* 2005). Hence, all observations from these assays are likely of limited significance. An investigation of the relevant literature suggests that 28°C is the standard temperature for the study of curli expression (Tschowri, Busse and Hengge 2009), so repeating these assays at this temperature, or perhaps even a lower one may prove useful. In addition, the assay used in this work was purely qualitative and should perhaps be followed by a quantitative approach such as the spectrophotometry technique described by Baugh *et al.* 2014. A more quantitative approach may reveal subtle differences in curli production between strains that might be overlooked by the method used here.

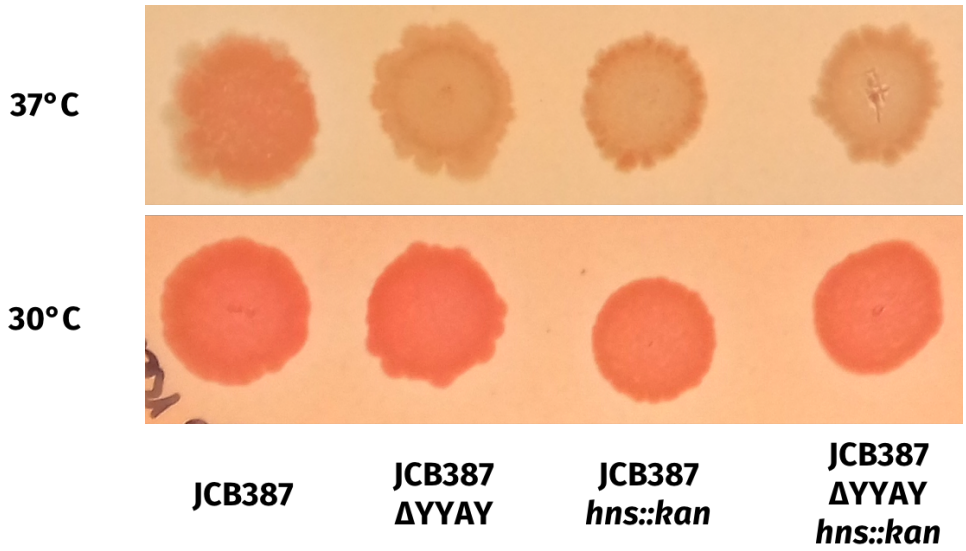

Fig. 11 | Example results from the qualitative Congo red curli production assays. The intensity of the red colour is indicative of the level of curli production. Wild type *E. coli* K-12 JCB387 appears slightly paler at 37°C than at 30°C. However, JCB387 ΔYYAY, JCB387 *hns::kan* and JCB387 ΔYYAY *hns::kan* all appear more noticeably paler at 37°C than at 30°C. Hence, deleting the YYAY operon, the *hns* gene, or both, from JCB387 appears to reduce curli production at 37°C, but not 30°C. Results shown here are representative of two biological replicates.

Can overexpressing MarA reduce biofilm formation by upregulating the YYAY operon?

High-level expression of MarA in the HC91255 clinical strain of uropathogenic *E. coli* (UPEC) is known to decrease the expression of the type I fimbriae genes *fimA* (which encodes the major structural subunit) and *fimB* (which encodes a site-specific recombinase) (Vila and Soto 2012). This effect is induced by disrupting the function of the negative regulator of the *marRAB* operon, MarR, either by the addition of salicylate, or by mutation of the *marR* gene. Decreased expression of the *fimA* and *fimB* genes leads to disruption of the initial surface attachment step of biofilm formation, and hence an overall reduction in biofilm formation.

In an attempt to determine the effect of high-level MarA expression on the formation of *E. coli* K-12 biofilms, further crystal violet biofilm formation assays were carried out here using strains transformed with the MarA-pET28a plasmid, where the *marA* gene is under the control of a T7 promoter (not shown). Unfortunately, the strain used for these assays, *E. coli* T7 Express (which expresses T7 RNA polymerase in an IPTG-inducible manner from a gene inserted into its chromosomal *lac* operon) does not form biofilms. Future work could use a different overexpression system to bypass this issue.

Further β-galactosidase assays were also performed, in an attempt to determine the effect of high-level MarA expression on the activity of the P_{ycgZ} promoter (Fig. 12). Background levels of β-galactosidase activity are observed in *E. coli* T7 Express cells carrying the pRW50 plasmid due to background expression of the promoterless *lacZ* gene on pRW50. Cloning a DNA fragment which contains the P_{ycgZ} promoter and its upstream marbox onto the pRW50 plasmid, upstream of the promoterless *lacZ* gene, led to increased levels of *lacZ* expression, suggesting that MarA binds to the marbox and transcriptionally activates the P_{ycgZ} promoter. Deletion of the marbox from this fragment appeared to reduce this effect. Attempts to overexpress MarA via the addition of IPTG produced unusual results. Namely, the activity of the P_{ycgZ} promoter appeared to be increased independently of a functional marbox. The cause of this is unknown, although perhaps the use of a different IPTG concentration or another overexpression system would be helpful.

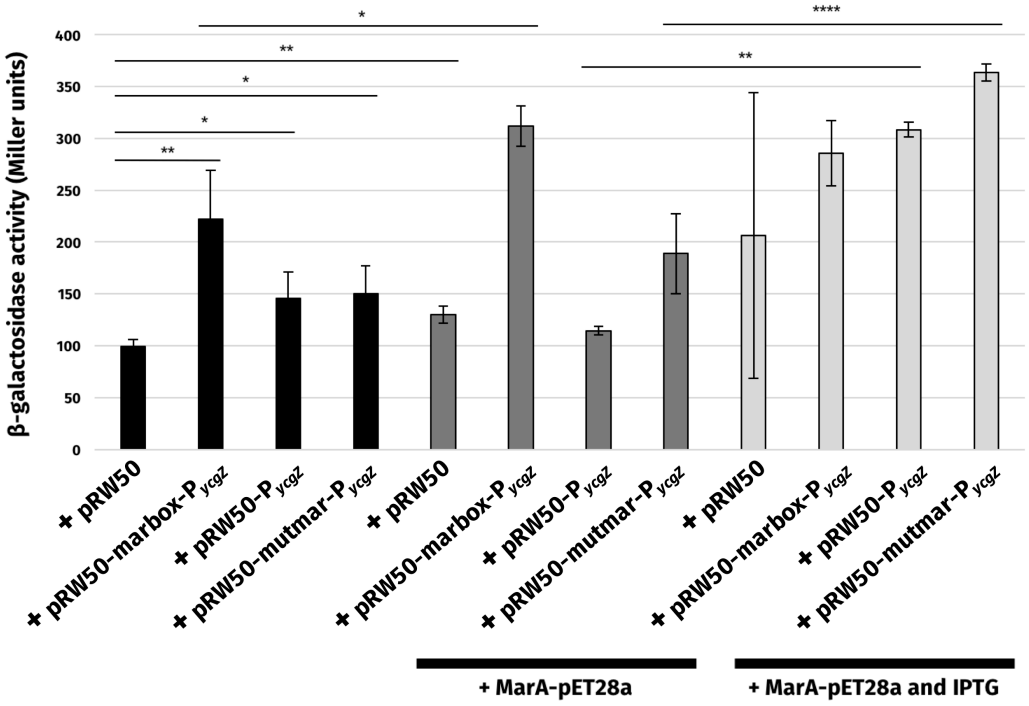

Fig. 12 | Expression from the P_{ycgZ} promoter, as measured by β-galactosidase assays, appears to be reduced in the absence of the marbox; although attempts to overexpress MarA via the addition of IPTG produced unusual results. Three different pRW50 derivatives were used in this assay, each containing a different fragment of DNA inserted upstream of the promoterless *lacZ* gene. The pRW50-marbox-P_{ycgZ} plasmid carries a DNA fragment containing the P_{ycgZ} promoter and its upstream marbox; the pRW50-P_{ycgZ} plasmid carries a DNA fragment containing only the P_{ycgZ} promoter; and the pRW50-mutmar-P_{ycgZ} plasmid carries a DNA fragment containing the P_{ycgZ} promoter and a mutated version of the upstream marbox. Data for each strain presented here represent the mean activities of three biological replicates. Activities of the strains shown in black were tested in the presence and absence of IPTG (not shown here), and the differences were not statistically significant ($P > 0.05$), indicating that IPTG has no statistically significant inhibitory effect on β-galactosidase production in these assays. Data for each strain presented here represent the mean activities of three biological replicates. Columns indicate the mean of each data set, while error bars show the standard deviation. * = $P \leq 0.05$, ** = $P \leq 0.01$, **** = $P \leq 0.0001$.

Research Project #2

A newly-identified sRNA from *Pseudomonas aeruginosa* binds to the post-transcriptional regulators RsmA and RsmN

ABSTRACT

The RNA-binding post-transcriptional regulators RsmA and RsmN of *Pseudomonas aeruginosa* are homologues of the *Escherichia coli* CsrA protein. Deletion of *rsmA* alters the expression of over 500 *P. aeruginosa* genes, including genes involved in iron acquisition, quorum sensing, multidrug efflux and motility. All known RsmN targets are shared with RsmA, indicating that the two proteins play similar roles, although Δ*rsmN* mutants typically exhibit less pronounced phenotypes than Δ*rsmA* mutants, probably due the fact that the proteins exhibit different binding affinities. In general terms, RsmA and RsmN act to repress the chronic infection lifestyle and promote acute infection. Regulatory sRNAs, such as RsmY and RsmZ, counteract these functions. A recent RNA-seq study identified 503 transcripts that bind to RsmN *in vivo*. *In silico* analysis of the RNA-seq data revealed an ~130 bp gene suspected to encode an RsmN-binding sRNA. The predicted secondary structure of the putative new sRNA was found to contain several GGA motifs in single-stranded regions (an attribute known to be important for binding to RsmA and RsmN); and a Rho-independent transcription terminator was predicted near the 3' end. Here, electrophoretic mobility shift assays (EMSAs) – using purified His-tagged RsmA and RsmN, and *in vitro* synthesised fluorescently-labelled sRNA – were used to prove binding of the putative new sRNA to RsmA and RsmN *in vitro*. Until recently, *P. aeruginosa* was thought to encode only two sRNAs capable of regulating RsmA and RsmN (i.e. RsmY and RsmZ). Now, it seems possible that the organism encodes at least five (RsmY, RsmZ, RsmW, RsmV, and the putative new sRNA reported here), which raises questions as to their distinct regulatory roles.

INTRODUCTION

Members of the CsrA/RsmA family of RNA-binding proteins post-transcriptionally regulate the expression of many genes involved in processes such as carbon metabolism, quorum sensing, virulence factor production, siderophore production, motility and stress responses in a wide variety of bacteria (Vakulskas *et al.* 2015). The first of these proteins to be discovered, CsrA, was identified in *Escherichia coli* (Romeo *et al.* 1993), and they still remain best characterised in Gram-negative bacteria, such as in pathogenic and non-pathogenic strains of *Escherichia coli*, as well as in *Shigella* spp., *Pseudomonas* spp., *Vibrio* spp., *Yersinia* spp., *Legionella pneumophila* and a number of important Gram-negative plant pathogens (Vakulskas *et al.* 2015); although a functional CsrA homologue has also been reported in *Bacillus subtilis* (Yakhnin *et al.* 2007). CsrA/RsmA family proteins are known to directly activate and/or repress gene expression by a variety of different molecular mechanisms (Vakulskas *et al.* 2015). For example, binding of the *Pseudomonas aeruginosa* RsmA protein to the 5' untranslated leader region of the *phz2* mRNA activates translation of the transcript by preventing formation of a stem-loop structure which would otherwise block ribosome binding (Ren *et al.* 2014); whereas binding of RsmA to a stem-loop structure in the 5' untranslated leader region of the *pslA* mRNA prevents the ribosome from accessing the Shine-Delgarno sequence, thereby repressing translation of the transcript (Irie *et al.* 2010).

Many pseudomonads produce two or more CsrA/RsmA family proteins: for example, RsmA and RsmN in *P. aeruginosa* (Pessi *et al.* 2001; Marden *et al.* 2013; Morris *et al.* 2013); RsmA and RsmE in *P. protegens* CHA0 (Reimmann *et al.* 2005); and RsmA, RsmE and RsmI in *P. putida* (Huertas-Rosales *et al.* 2017). *In vivo*, these proteins are thought to be negatively regulated under certain conditions by being sequestered away from their usual mRNA targets through binding to various sRNAs: for example, RsmY and RsmZ (and possibly RsmW and RsmV) in *P. aeruginosa* (Valverde *et al.* 2003; Heurlier *et al.* 2004; Brencic *et al.* 2009; Miller *et al.* 2016; Janssen *et al.* 2018); RsmX, RsmY and RsmZ in *P. protegens* CHA0; and RsmY and RsmZ in *P. putida* (Huertas-Rosales, Ramos-González and Espinosa-Urgel 2016).

The RsmA and RsmN proteins of *P. aeruginosa* form ~14 kDa homodimers in solution which have slightly different secondary structures, but which share a similarly positioned arginine residue critical for RNA-binding activity (Gutierrez *et al.* 2005; Rife *et al.* 2005; Heeb *et al.* 2006; Marden *et al.* 2013; Morris *et al.* 2013). RsmA exerts a positive regulatory influence on the expression of genes associated with an acute infection lifestyle – such as genes encoding lipase (*lipAH*) (Heurlier *et al.* 2004), rhamnolipids (*rhlAB*) (Heurlier *et al.* 2004), and those involved in swarming motility and type III secretion (Heurlier *et*

al. 2004; Mulcahy *et al.* 2006), and the production of type IV pili (Brencic and Lory 2009) – and negatively regulates products associated with a chronic infection lifestyle – such as genes involved in the production of hydrogen cyanide (HCN) and pyocyanin (Heurlier *et al.* 2004), genes involved in type VI secretion (Brencic and Lory 2009), and genes involved in the production of exopolysaccharides which promote biofilm formation (*pel* and *psl*) (Irie *et al.* 2010). Deletion of *rsmA* leads to increased biofilm formation, and the increased expression of HCN, pyocyanin and T6SS-related products (Pessi *et al.* 2001; Brencic and Lory 2009; Irie *et al.* 2010). While these effects are not observed in an *rsmN* single mutant, simultaneous deletion of both *rsmA* and *rsmN* leads to phenotypes which are more pronounced than those of the *rsmA* single mutant, suggesting that RsmN fulfils a similar (but seemingly less influential) role to RsmA. For example, elevated levels of biofilm formation are observed in the *rsmA rsmN* double mutant compared to the *rsmA* single mutant (Marden *et al.* 2013).

Several sRNAs are known to inhibit the RsmA and RsmN proteins in *P. aeruginosa*. The GacA/GacS two component regulatory system and two orphan sensor kinases LadS and RetS control the transcription of the sRNAs RsmY and RsmZ in response to unknown environmental signals (Valverde *et al.* 2003; Heurlier *et al.* 2004; Brencic *et al.* 2009); and two non-GacA/GacS-regulated sRNAs RsmW and RsmV have also been described (Miller *et al.* 2016; Janssen *et al.* 2018).

Binding between CsrA/RsmA proteins and RNA molecules (either mRNAs or sRNAs) is dependent on certain primary and secondary structural characteristics (Schulmeyer *et al.* 2016). A number of studies have helped to elucidate these requirements. For example, the *E. coli* CsrA-binding RNA consensus sequence (RUACARGGAUGU) has been previously determined, where R is either A or G, the ACA and GGA motifs are 100% conserved, and the GU motif is 98% conserved (Dubey *et al.* 2005); and a systematic evolution of ligands by exponential enrichment (SELEX) experiment to identify high affinity and specific targets of *P. aeruginosa* RsmA and RsmN revealed a common consensus sequence (CANGGAYG) (Schulmeyer *et al.* 2016). These and other studies have revealed a general binding pattern, suggesting that CsrA/RsmA proteins favour binding to GGA-motif-containing sequences exposed in single-stranded regions of RNA molecules, such as in the loop region of stem-loop structures (Dubey *et al.* 2005; Schubert *et al.* 2007; Schulmeyer *et al.* 2016).

A recent RNA-seq study identified 503 transcripts that bind to RsmN *in vivo* (Romero *et al.* 2018). *In silico* analysis of the RNA-seq data revealed an ~130 bp gene suspected to encode an RsmN-binding sRNA (Romero, unpublished). The RNA-seq study strongly suggested that this sRNA binds to RsmN *in vivo* due to the fact that it was enriched >3.5-fold in cDNA libraries from RsmN-bound RNA samples, compared to those from

total RNA samples. This initial observation was supported by further *in silico* analysis and prediction of the secondary structure of the newly-identified sRNA, which was found to contain several GGA motifs in single-stranded regions (an attribute known to be important for binding to RsmA and RsmN); and a Rho-independent transcription terminator was predicted near the 3' end.

The principal objective of this work was to investigate, *in vitro*, the hypothesis derived from the RNA-seq study – that the new sRNA binds to the RsmN protein. Given that all known RsmN targets are shared with RsmA, it was resolved to also determine whether the new sRNA binds to RsmA, with the hypothesis that this binding would also occur.

Motility in *P. aeruginosa* – traditionally characterised as either swarming, swimming or twitching (Henrichsen 1972) – is typically associated with the acute infection lifestyle promoted by RsmA and RsmN; and non-motility with the chronic infection lifestyle (Mahenthiralingam, Campbell and Speert 1994). Swimming and swarming are largely flagellum-dependent, whereas twitching depends on type IV pili (Toutain, Zegans and O'Toole 2005). RsmA promotes swarming motility in *P. aeruginosa*, whereas the sRNAs RsmY and RsmZ have the opposite effect (Heurlier *et al.* 2004; Kay *et al.* 2006). Hence, one might expect overexpression of the new sRNA to reduce motility via inhibition of RsmA and RsmN. This hypothesis was also investigated.

MATERIALS AND METHODS

Bacterial strains, plasmids and primers

Bacterial strains and plasmids used in this work are listed in Table 4. *E. coli* strain DH5α was used for the transformation, maintenance and storage of plasmids, whereas *E. coli* strain C41(DE3) was used for the overexpression and purification of His-tagged versions of the RsmA and RsmN proteins. *P. aeruginosa* strain PAO1-L (PAO1; Lausanne subline) was used for the preparation of genomic DNA for PCRs and subsequent construction of plasmids; as well as for the motility assays. All E. *coli* and *P. aeruginosa* strains were grown at 37°C in LB broth with 200 rpm shaking, or on LB agar plates, unless otherwise indicated. When required, carbenicillin or tetracycline was added to the broth and agar at concentrations of 100 µg/mL and 10 µg/mL respectively, for the maintenance of plasmids. Primers used in this work and their functions are listed in Table 5.

Name	Genotype/Description	Source
Strains		
E. coli DH5α	F^- *endA1 glnV44 thi-1 recA1* relA1 gyrA96 deoR nupG purB20 φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, hsdR17($r_K^-m_K^+$), λ^- laboratory strain for cloning and plasmid maintenance	Stratagene
E. coli C41(DE3)	F^- *ompT gal dcm* $hsdS_B(r_B^- m_B^-)$(DE3) DE3 = λ prophage with the T7 RNA polymerase gene and $lacI^q$ OverExpression C41(DE3), T7 gene expression strain	Lucigen
P. aeruginosa PAO1-L	PAO1 subline, University of Lausanne	Dieter Haas
Plasmids		
pminiRSETA	modified derivative of the expression vector pRSETA (Invitrogen) in which the enterokinase site has been replaced by a thrombin recognition site; Amp^R	(Heeb *et al.* 2006)
pminiRSETA::*rsmA*	pminiRSETA carrying the 186 bp *rsmA* gene from PAO1-L	this work
pminiRSETA::*rsmN*	pminiRSETA carrying the 216 bp *rsmN* gene from PAO1-L	unpublished
pUCP20	high-copy-number *Escherichia-Pseudomonas* shuttle vector derived from pUCP18 by digestion with NdeI and SfuI, followed by blunt-end ligation; Amp^R	(West *et al.* 1994)
pUCP20::sRNA	pUCP20 carrying a 1041 bp insert containing the putative sRNA gene and surrounding sequence	this work
pME6000	cloning vector derived from the broad-host-range Gram-negative cloning vector pBBR1MCS; Tet^R	(Maurhofer *et al.* 1998)
pME6000::sRNA	pME6000 carrying a 353 bp insert containing the putative sRNA gene region	this work

Table 4 | List of bacterial strains and plasmids used in this work.

Name	Sequence (5'→3')
To amplify approximate putative sRNA gene region and surrounding sequence from *P. aeruginosa* PAO1-L for insertion into pUCP20	
sRNA_pUCP20_F	TATGGATCCAATGAAAGCACCGGCAAG
sRNA_pUCP20_R	TATGAATTCCCTGTCGATCTTCATGCG
To screen putative pUCP20::sRNA transformants of *E. coli* and *P. aeruginosa* via colony PCR	
pUCP20_cPCR_F	CTGCAAGGCGATTAAGTTGG
pUCP20_cPCR_R	GGCCGATTCATTAATGCAGC
To confirm successful insertion of approximate sRNA gene region and surrounding sequence into pUCP20 via sequencing	
sRNA_pUCP20_F330	ATGTTGCGAGCGAAGATG
sRNA_pUCP20_R330	ATCTTCGCTCGCAACATC
sRNA_pUCP20_F790	AGTGTTCACGCCGAAATC
sRNA_pUCP20_R790	TTTCGGCGTGAACACTTC
To amplify approximate putative sRNA gene region for *in vitro* transcription to make ATTO700-labelled RNA for EMSAs	
sRNA_T7_EMSA_F	TTTTCTGCAGTAATACGACTCACTATAGGGCGCACACTGTACTGCACCC
sRNA_L_EMSA_R	AAAAAAAACCCCCCCCCAAGGCCTGATCGCAGACGGC
To amplify *rsmY* gene for *in vitro* transcription to make ATTO700-labelled RNA and unlabelled competitor RNA for EMSAs	
rsmY_T7_EMSA_F	TTTTCTGCAGTAATACGACTCACTATAGGGTCAGGACATTGCGCAGGAA
rsmY_L_EMSA_R	AAAAAAAACCCCCCCCCGCCTTTTGGGCGGGGTTTTG
rsmY_UL_EMSA_R	GCCTTTTGGGCGGGGTTTTG
To amplify region upstream of *pslA* gene for *in vitro* transcription to make unlabelled competitor RNA for EMSAs	
pslA_T7_EMSA_F	TAATACGACTCACTATAGGGACAAAGCCACTATCGACGAAT
pslA_UL_EMSA_R	CATGTTGTTTGCTCTGCCGA
To amplify the *rsmA* gene from *P. aeruginosa* PAO1-L for insertion into pminiRSETA, to enable purification of RsmA protein	
rsmA_pminiRSETA_F	TATGGATCCATGCTGATTCTGACTCGTCG
rsmA_pminiRSETA_R	TATGAATTCTTAATGGTTTGGCTCTTGATCTTTC
To amplify approximate sRNA gene region for insertion into pME6000	
sRNA_pME6000_F	TATGAATTCTGGCCTGGAAATCAGCGAAG
sRNA_pME6000_R	TATGGATCCATTTCGGCGTGAACACTTCC
To screen putative pME6000::sRNA transformants of *E. coli* and *P. aeruginosa* via colony PCR	
pME6000_cPCR_F	TCATTAATGCAGCTGGCACG
pME6000_cPCR_R	ACTTATTCAGGCGTAGCACC
To confirm pminiRSETA and pME6000 insertions via sequencing	
T7P	TAATACGACTCACTATAGGG
T7T	TATGCTAGTTATTGCTCAGCGG
T3	ATTAACCCTCACTAAAGGG

Table 5 | List of primers used in this work. GGATCC = BamHI restriction site; GAATTC = EcoRI restriction site

Plasmid constructions

All novel plasmids used in this work were constructed via conventional cloning methods. For all cloning procedures, amplified PCR products and backbone vectors were digested with both EcoRI and BamHI restriction enzymes, ligated using T4 DNA ligase, and transformed into *E. coli* DH5α cells via the standard heat-shock calcium-chloride method (Nakata, Tang and Yokoyama 1997). Transformations of purified plasmids into *P. aeruginosa* PAO1-L were performed via a similar method (Chuanchuen, Narasaki and Schweizer 2002). Primers sRNA_pUCP20_F and sRNA_pUCP20_R were used to amplify, via PCR, an 1041 bp region of *P. aeruginosa* PAO1-L genomic DNA, containing the

putative sRNA gene region and surrounding sequence for insertion into pUCP20. A strain of *P. aeruginosa* transformed with this plasmid was used in swarming, swimming and twitching assays. Primers rsmA_pminiRSETA_F and rsmA_pminiRSETA_R were used to amplify, via PCR, the 186 bp *rsmA* gene from *P. aeruginosa* PAO1-L genomic DNA, for insertion into pminiRSETA. This plasmid was used to overexpress and purify a His-tagged version of the RsmA protein. Primers sRNA_pME6000_F and sRNA_pME6000_R were used to amplify, via PCR, a 353 bp region of *P. aeruginosa* PAO1-L genomic DNA, for insertion into pME6000. This plasmid was intended for use in further motility assays. However, these were not performed due to time constraints. All other primers were used as outlined in Table 5; either to validate the three newly-constructed plasmids described above, by colony PCR or DNA sequencing, or to amplify various regions of *P. aeruginosa* PAO1-L genomic DNA for *in vitro* transcription to produce RNAs to be used in EMSAs.

Purification of 6xHis-Thr-RsmA and 6xHis-Thr-RsmN proteins for EMSAs

Two derivatives of the pminiRSETA expression plasmid – pminiRSETA::*rsmA* and pminiRSETA::*rsmN* – were used in two separate purification procedures, to express His-tagged versions of the RsmA and RsmN proteins from *P. aeruginosa* PAO1-L in *E. coli* C41(DE3) cells (Maurhofer *et al.* 1998). The ~8.5 kDa, 74-amino-acid 6xHis-Thr-RsmA protein and the ~9.3 kDa, 84-amino-acid 6xHis-Thr-RsmN protein were purified from the C41(DE3) cells using a method similar to that described by Romero *et al*., 2018. In each case, an overnight culture (10 mL) of the *E. coli* C41(DE3) strain harbouring the relevant pminiRSETA derivative was used to inoculate LB broth (800 mL) containing carbenicillin (100 µg/mL) in a 2 L glass flask. This newly-inoculated volume of LB broth was then incubated at 37°C with shaking at 200 rpm until reaching an OD_{600} of between 0.6–0.9 (log/exponential phase). This incubation step took ~7 h in the case of the C41(DE3) strain harbouring the pminiRSETA::*rsmA* plasmid, and ~4 h for pminiRSETA::*rsmN*. Expression of the His-tagged protein was then induced by the addition of IPTG (0.3 mM), and overnight incubation at 30°C with shaking at 200 rpm (~16–20 h). The OD_{600} values of the pminiRSETA::*rsmA* and pminiRSETA::*rsmN* cultures after overnight incubation were ~1.7 and ~1.3 respectively. The 800 mL culture was divided into four separate 200 mL volumes in four 250 mL centrifuge containers and centrifuged at 4°C for 15 min at 4,000 rpm. Each of the four resultant pellets was transferred to a 50 mL Falcon tube and stored at –80°C until required. In the case of His-tagged RsmN, one of these pellets was used to purify the protein by using the Ni-NTA Fast Start Kit (Qiagen) following the manufacturer's procedure. In the case of His-tagged RsmA, only approximately one quarter of a pellet was required to purify the protein by using the Ni-NTA Fast Start Kit (Qiagen), due to the large quantity of protein

produced. In each case, samples taken at various stages of the purification process were analysed via separation on a 10% SDS-PAGE gel for 1 h at 200 V, and in each case the final elution stage yielded 2 mL of ~0.8–1.2 mg/mL protein. SDS-PAGE gels were stained with 0.1% Coomassie Brilliant Blue R-250, 50% methanol and 10% glacial acetic acid for 1 h, followed by destaining with 40% methanol and 10% glacial acetic acid for 1 h (in the presence of a piece of tissue paper, to accelerate the process).

In vitro transcription, RNA purification and fluorescent labelling for EMSAs

Regions of *P. aeruginosa* PAO1-L genomic DNA corresponding to the newly-identified putative sRNA gene and the *rsmY* gene were amplified by PCR using primers designed to incorporate a T7 promoter at the 5' end and a 17 nt extension at the 3' end. The resultant dsDNA PCR products were then transcribed into RNAs for use in EMSAs, using the MAXIscript™ T7 Transcription Kit (Life Technologies). In order to enable their visualisation in EMSAs, an ATTO700-labelled primer was then hybridised to the 3' extension of each of these RNAs, using a method previously described by Ying, Fourmy and Yoshizawa 2007. In each case, the RNA and the ATTO700-labelled primer were mixed in Tris-EDTA buffer solution, pH 8.0 (Sigma-Aldrich) in a primer:RNA ratio of ~20:1. This mixture was incubated in a thermocycler for 5 min at 95°C, and then the temperature was decreased very slowly to 25°C over a period of ~2 h (at a rate of 0.01°C/sec) to allow hybridisation to occur. The RNA:DNA hybrid was stored at –80°C until required. Regions of genomic DNA corresponding to the *rsmY* gene and an ~150 bp region upstream of the *pslA* gene were also amplified by PCR and transcribed into RNA using the MAXIscript kit, to be used as unlabelled competitor RNAs. Hence in these cases, a conventional reverse primer was used with the T7-promoter-incorporating forward primer, as a 3' extension was not required.

Initial EMSAs to prove binding of newly-identified sRNA to RsmA and RsmN

PCR tubes containing a range of different concentrations of 6xHis-Thr-RsmA or 6xHis-Thr-RsmN (from 0 nM to 800 nM) were incubated in a thermocycler at 37°C for 30 min, with either the ATTO700-labelled RNA corresponding to the putative sRNA region, or ATTO700-labelled RsmY, in 1X binding buffer (10 mM Tris-HCl [pH 7.5], 10 mM $MgCl_2$, 100 mM KCl), with 1.1 µg/µL yeast tRNA (Thermo Fisher Scientific), 7.5% (v/v) glycerol, and 0.2 units SUPERase• In™ RNase Inhibitor (Thermo Fisher Scientific). The final concentration of the ATTO700-labelled RNA in each reaction was 10 nM. After this incubation step, Bromophenol Blue (0.01% w/v) was added to each sample, before separation of all samples at 4°C on

a 6% (w/v) native (i.e. non-denaturing) polyacrylamide TBE gel (47 mM Tris, 45 mM boric acid, 1 mM EDTA [pH 8.3]) at 150 V for 1 h, in 0.5X TBE buffer. Before sample loading, each gel was pre-run for 15 min at 200V in an attempt to eliminate RNases and other impurities. Gel imaging and image processing were performed using a 9201 Odyssey Imaging System (LI-COR Biosciences) and Image Studio V5.0 software respectively.

Competition EMSAs to prove specificity of binding

EMSAs to determine the specificity of the binding between ATTO700-labelled RNA and the two proteins were carried out by a similar method, using an excess of unlabelled competitor RNA (either RsmY or the unlabelled RNA corresponding to a ~150 bp region upstream of the *pslA* gene). In this case, a fixed protein concentration was used, and the final concentration of the ATTO700-labelled RNA in each reaction was reduced to 5 nM.

Motility assays

Motility assays were carried out to determine the swarming, swimming and twitching abilities of a *P. aeruginosa* PAO1-L strain carrying the high-copy-number plasmid pUCP20::sRNA, compared to the wild type and a strain carrying pUCP20 with no insert. Swarming agar (5 g/L Bacto™ Agar, 8 g/L Nutrient Broth No. 2), swimming agar (3 g/L Bacto™ Agar, 8 g/L Nutrient Broth No. 2) and twitching agar (20 g/L tryptone, 10 g/L yeast extract, 20 g/L NaCl, 10 g/L Agar Bacteriological (Agar No.1)) were prepared and autoclaved. All media was placed at 55°C for at least 30 min. Glucose was added to the swarming and swimming agar to a final concentration of 0.5 % (w/v). Media volumes of exactly 20 mL were transferred into Petri dishes and the agar allowed to solidify at room temperature with the lid closed for 40 min. The solidified plates were then allowed to dry with the lids off in a microbiological safety cabinet for 20 min. The centre of each plate was then inoculated with 5 µL of overnight culture and placed right-side-up at 37°C. Swarming and swimming plates were inoculated by gently pipetting culture onto the agar surface, and were incubated at 37°C unstacked on an even surface for ~7.5 h. Twitching plates were inoculated by piercing the agar surface and pipetting culture into the perforation, and were incubated at 37°C in an airtight container with a piece of moist tissue paper for ~48 h. All reagents for media preparation were from Thermo Fisher Scientific, except for the Bacto™ Agar (Becton Dickinson) and NaCl (Sigma-Aldrich).

RESULTS

Identification of putative new sRNA and *in silico* analysis

Manual inspection of RNA-seq data from Romero *et al.*, 2018, revealed a novel ~130 bp gene suspected to encode an RsmN-binding sRNA, which was enriched >3.5-fold in cDNA libraries from RsmN-bound RNA samples, compared to those from total RNA samples – indicating that the transcript binds to RsmN *in vivo* (Romero, unpublished) (Fig. 13). This finding warranted further investigation, as no sRNA gene has yet been annotated at this genomic location. Preliminary inspection of the primary sequence and approximate length of the sRNA gene, and the observation that the encoded sRNA binds to RsmN *in vivo*, suggested that the newly-identified sRNA may be similar in structure and function to the *P. aeruginosa* sRNAs RsmY, RsmZ, RsmW, and RsmV, which have previously been shown to bind both RsmA and RsmN *in vitro* (Kay *et al.* 2006; Brencic *et al.* 2009; Miller *et al.* 2016; Janssen *et al.* 2018).

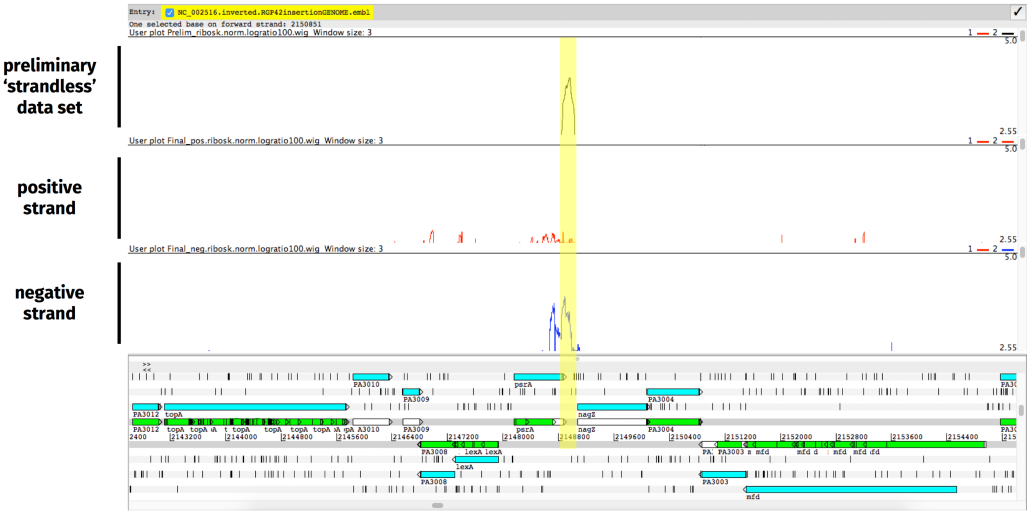

Fig. 13 | *In silico* identification of the approximate location of an sRNA gene encoding a putative unannotated RsmN-binding sRNA on the negative strand of the *P. aeruginosa* PAO1-L genome. Visualisation – using the genome browser and annotation tool, Artemis (Carver *et al.* 2012) – of RNA-seq data from Romero *et al.*, 2018, mapped to the *P. aeruginosa* PAO1-L genome in the region containing the *psrA* and *nagZ* genes, revealed a peak reaching a maximum enrichment index of approximately 4 in the preliminary 'strandless' RNA-seq dataset, and a peak reaching a maximum enrichment index of approximately 3.9 in the negative strand RNA-seq dataset (Romero, unpublished). These peaks indicate that the RNA-derived cDNA reads mapping to the negative strand in this genomic region were enriched >3.5-fold in cDNA libraries from RsmN-bound RNA samples, compared to those from total RNA samples, and suggest the existence of an as-of-yet unannotated ~130 bp sRNA gene on the opposite strand to both *psrA* and *nagZ*, complementary to a portion of the intergenic region between the two genes. The 702 bp *psrA* gene encodes the transcriptional regulator PsrA (*Pseudomonas* sigma regulator A) (Kojic and Venturi 2001), and the 999 bp *nagZ* gene encodes the β-N-acetyl-D-glucosaminidase NagZ (involved in resistance to β-lactams) (Asgarali *et al.* 2009; Zamorano *et al.* 2010).

Further *in silico* analysis of the sRNA gene and its surrounding sequence was carried out before commencing any *in vitro* or *in vivo* investigation. For starters, the sRNA gene and its putative upstream regulatory sequence was analysed using *Virtual Footprint* (Munch 2003; Munch *et al.* 2005), *TransTermHP* (Kingsford, Ayanbule and Salzberg 2007), *BPROM* and *BacPP*. Binding sites for known *E. coli* and *P. aeruginosa* transcription factors were predicted in the region upstream of the new sRNA gene, and two –10 elements were predicted within ~100 bp of the approximate 5' end of the gene, possibly indicating the presence of a promoter (Fig. 14). In addition, a Rho-independent transcriptional terminator was predicted near the approximate 3' end of the gene, which may indicate the transcriptional terminator of the sRNA gene.

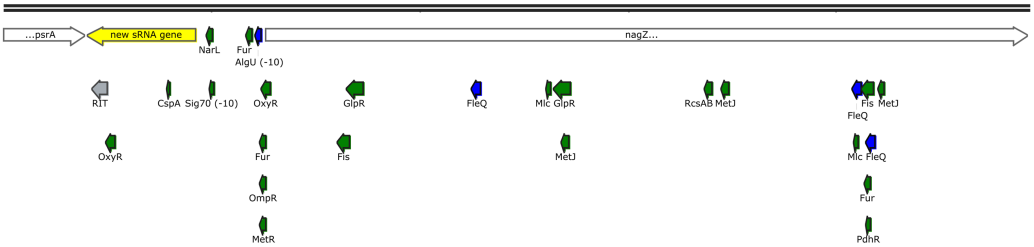

Fig. 14 | Predicted transcription factor binding sites and predicted Rho-independent terminator in the genomic region containing the putative new sRNA gene. Transcription factor binding sites were predicted using the online *Virtual Footprint* bacterial promoter analyser V3.0 (Munch 2003; Munch *et al.* 2005). Sites were predicted based on known binding sites of transcription factors in *E. coli* K-12 (green) and *P. aeruginosa* PAO1 (blue). The *TransTermHP* system predicts an 18 bp intrinsic (i.e. Rho-independent) transcriptional terminator (CGGCGCCCTGGGGCGCCG; grey, and labelled RIT above) near the 3' end of the new sRNA gene (Kingsford, Ayanbule and Salzberg 2007). This corresponds to the RIT predicted near the 3' end of the new sRNA by *mfold*. As of yet, the precise 5' and 3' boundaries of the new sRNA gene have yet to be defined on a single-nucleotide level. Web-based bacterial promoter prediction programmes *BPROM* (http://www.softberry.com) and *BacPP* (http://www.bacpp.bioinfoucs.com/home) failed to predict any sigma factor binding sites upstream of the new sRNA gene, although *Virtual Footprint* predicted two potential –10 elements. This figure was prepared using *SnapGene* V2.3.2.

Predicted structures for RsmY, RsmZ, RsmW and RsmV from PAO1 have been published previously (Heurlier *et al.* 2004, 2004; Miller *et al.* 2016; Janssen *et al.* 2018). Hence, it was resolved to predict the secondary structure of the putative new sRNA in an attempt to identify structural similarities with these other predicted structures. The secondary structure of the putative new sRNA was predicted using *mfold* and was found to contain several GGA motifs exposed in single-stranded regions, at least one of which appears to be presented in the loop region of a stem-loop structure and may enable binding to RsmA and RsmN (Fig. 15). The predicted structure of the new putative sRNA differs somewhat from the predicted structures of the other known sRNAs in *P. aeruginosa* PAO1, in that it appears to contain only a single, large stem-loop structure, rather than multiple smaller ones – although this may simply be an artefact of the prediction method. The Rho-independent transcriptional terminator predicted via analysis of the putative sRNA gene sequence using *TransTermHP* is also present in the predicted sRNA secondary structure, near the 3' end.

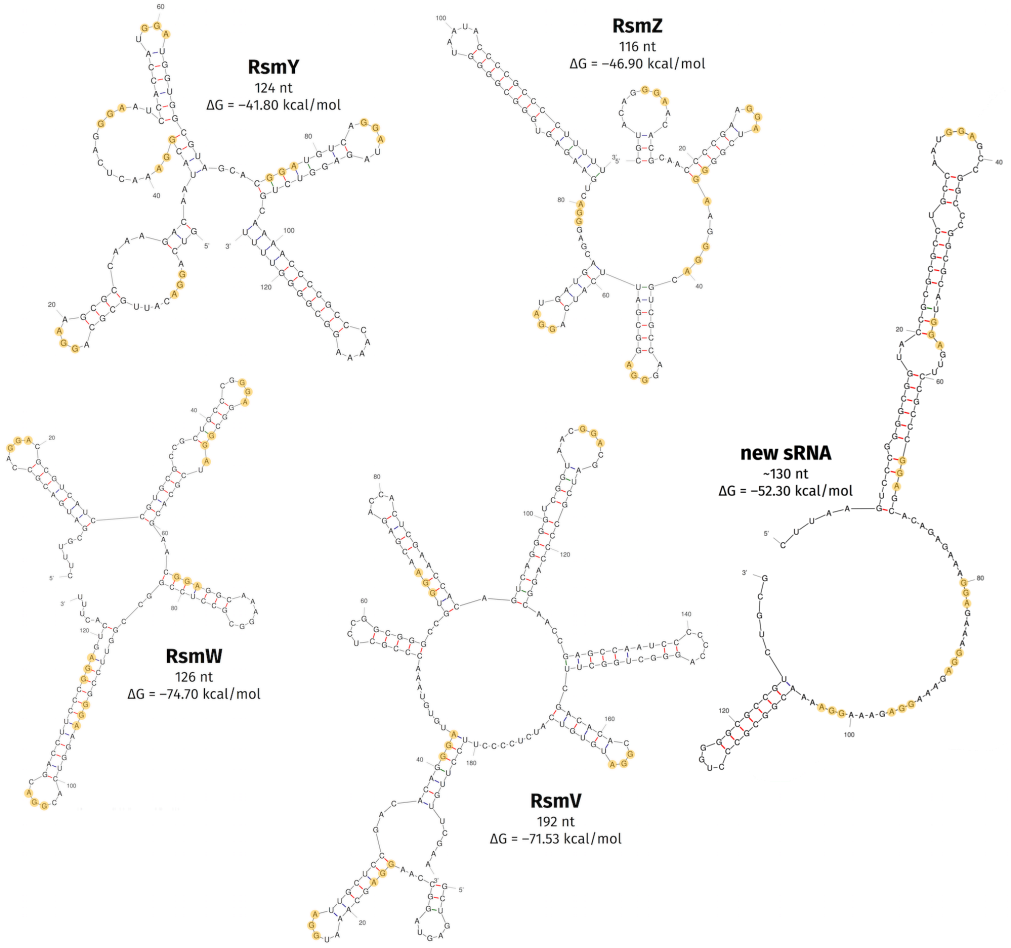

Fig. 15 | Predicted secondary structures of the Rsm family of sRNAs from *P. aeruginosa* PAO1. All secondary structures, including that of the newly-identified sRNA, were predicted using the default settings of the *mfold* web server for nucleic acid folding, at a folding temperature of 37°C (Zuker 2003). ΔG values represent the stability of the predicted structure at this temperature. In all cases, except RsmW, several different predicted structures were generated by *mfold*. The above structures were chosen on the basis of their relatively high stability in comparison to many of the other predicted structures, and their having relatively many exposed GGA motifs (yellow) in single-stranded regions. GGA motifs are indicated, as RsmA and RsmN have a preference for binding to GGA motifs situated in single-stranded regions of RNA stem-loop structures (Dubey *et al.* 2003; Gutierrez *et al.* 2005; Schubert *et al.* 2007; Schulmeyer *et al.* 2016). Predicted structures for RsmY, RsmZ, RsmW and RsmV from PAO1 have been published previously (Heurlier *et al.* 2004, 2004; Miller *et al.* 2016; Janssen *et al.* 2018).

Purification of 6xHis-Thr-RsmA and 6xHis-Thr-RsmN proteins for EMSAs

The preliminary *in silico* analysis described above (Fig. 13–15) appeared to suggest sufficient similarity between the new putative sRNA and known PAO1 Rsm family sRNAs to warrant further investigation. Given that these sRNAs are known to bind RsmA and RsmN *in vitro*, it was resolved to test the newly-identified sRNA for this activity. In preparation for EMSAs, His-tagged versions of the RsmA and RsmN proteins from *P. aeruginosa* PAO1-L (6xHis-Thr-RsmA and 6xHis-Thr-RsmN) were purified from *E. coli* (Fig. 16); and RNA targets were synthesised by *in vitro* transcription and fluorescently labelled as described in *Materials and Methods*.

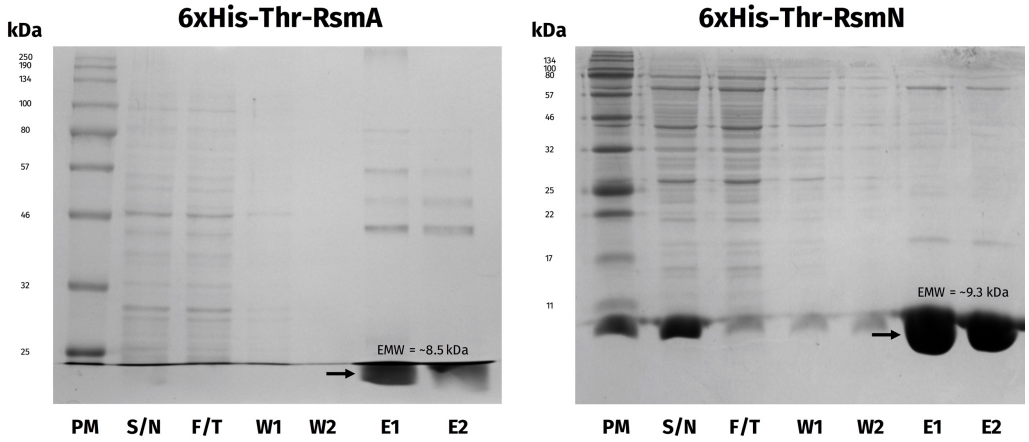

Fig. 16 | Purification of 6xHis-Thr-RsmA and 6xHis-Thr-RsmN proteins for EMSAs. Two derivatives of the pminiRSETA expression plasmid – pminiRSETA::*rsmA* and pminiRSETA::*rsmN* – were used in two separate purification procedures, to express His-tagged versions of the RsmA and RsmN proteins from *P. aeruginosa* PAO1-L in *E. coli* C41(DE3) cells. Left: 10% SDS-PAGE gel showing separation of samples taken at various stages of the purification process for the ~8.5 kDa, 74-amino-acid 6xHis-Thr-RsmA protein. Right: 10% SDS-PAGE gel showing separation of samples taken at various stages of the purification process for the ~9.3 kDa, 84-amino-acid 6xHis-Thr-RsmN protein. In each case, the leftmost lane shows the Color-coded Prestained Protein Marker, Broad Range (11-250 kDa) (Cell Signaling Technologies) (PM); followed by – from left to right – the initial cell lysate supernatant (S/N), the flow-through from the purification column (F/T), wash #1 (W1), wash #2 (W2), eluate #1 (E1), eluate #2 (E2). The final two eluate samples contain the desired His-tagged protein, indicated by the pronounced bands of low molecular weight. EMW = expected molecular weight.

The newly-identified sRNA binds to RsmA and RsmN *in vitro*

EMSAs showed that the new sRNA is capable of binding to RsmA and RsmN *in vitro* (Fig. 17). This was not entirely surprising, due to the presence of several GGA motifs in the sequence of the new sRNA and, in fact, binding to RsmN was positively expected based on the *in vivo* binding implied by the RNA-seq data from Romero *et al.*, 2018. Some Rsm family sRNAs are known to bind multiple CsrA/RsmA family protein molecules at a time (Jimenez *et al.* 2012). Unfortunately, the bands in the EMSA images appear too diffuse to make precise claims about exactly how many protein molecules are bound by the new sRNA in each assay.

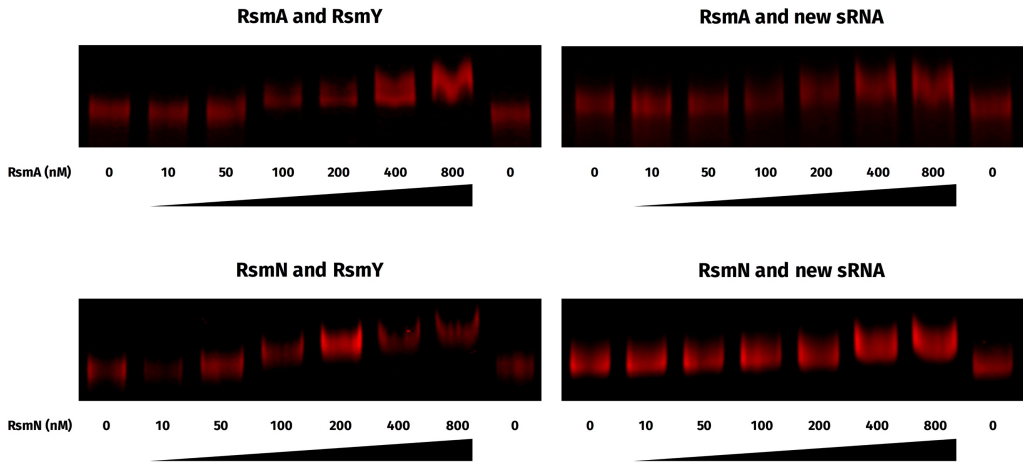

Fig. 17 | The newly-identified sRNA binds to RsmA and RsmN *in vitro*. Left: The RsmY sRNA is known to bind to both RsmA and RsmN proteins (Marden *et al.* 2013; Romero *et al.* 2018); and hence was used as a positive control. Band shifts observed upon co-incubation of ATTO700-labelled RsmY sRNA with increasing concentrations of either of the two purified proteins showed that this method is a viable means by which to test binding of the newly-identified sRNA to RsmA and RsmN. Right: Subsequent EMSAs, using the ATTO700-labelled newly-identified sRNA and the RsmA and RsmA proteins, also resulted in band shifts, indicating that, like the RsmY sRNA, this newly-identified sRNA has the ability to bind to both RsmA and RsmN.

Competition EMSAs to prove specificity of binding

Having shown that the new sRNA does indeed bind RsmA and RsmN *in vitro*, an attempt was made to prove that this binding is specific (Fig. 18). One way of doing this is to add increasing concentrations of either an unlabelled specific competitor RNA or an unlabelled non-specific competitor RNA into the EMSA mix; the rationale being that addition of an excess of specific competitor RNA will negatively impact binding of the fluorescently-labelled RNA to the protein, and addition of an excess of non-specific competitor RNA should have no effect. An unlabelled RNA corresponding to the RsmY sRNA was synthesised as a specific competitor, and an unlabelled RNA corresponding to the leader region of the *pslA* mRNA was synthesised as the non-specific competitor. As expected, binding of the fluorescently-labelled sRNA to RsmA was abolished upon addition of increasing concentrations of unlabelled specific competitors RsmY and *pslA*. Similarly, binding of the labelled sRNA to RsmN was abolished upon addition of increasing concentrations of the unlabelled specific competitor RsmY. However, given that unlabelled *pslA* is an appropriate non-specific competitor for RsmN, addition of an excess of this RNA was not expected to have an effect on binding of the labelled new sRNA to RsmN. It appears to reduce levels of binding here. Hence, a serious conclusion regarding the specificity of the binding between the new sRNA and the RsmA and RsmN proteins cannot be drawn from these results.

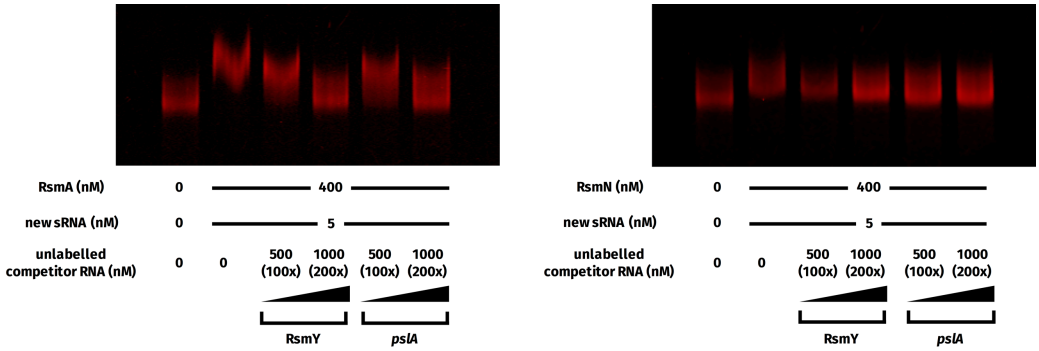

Fig. 18 | Competition EMSAs to prove specificity of binding of newly-identified sRNA to RsmA and RsmN. Increasing concentrations of competitor RNAs were added to the EMSA mixes as described in *Materials and Methods*. In each case, Lane 1 is representative of 'no binding' and Lane 2 is representative of 'binding'. Unfortunately, no serious conclusion regarding the specificity of the binding between the new sRNA and the RsmA and RsmN proteins can be drawn from these results.

Motility assays

Motility is associated with the acute infection lifestyle promoted by RsmA and RsmN. Having shown that the new sRNA binds to RsmA and RsmN *in vitro*, and knowing that RsmY and RsmZ inhibit the activity of RsmA *in vivo*, it seemed conceivable that introducing the new sRNA into *P. aeruginosa* PAO1-L on the high-copy-number plasmid pUCP20 might negatively impact bacterial motility by augmenting the natural role of the sRNA. Unfortunately, however, under the experimental conditions used, no significant motility differences were observed between wild type PAO1-L and a PAO1-L strain carrying the putative new sRNA region and surrounding sequence on the high-copy-number plasmid pUCP20 (Fig. 19).

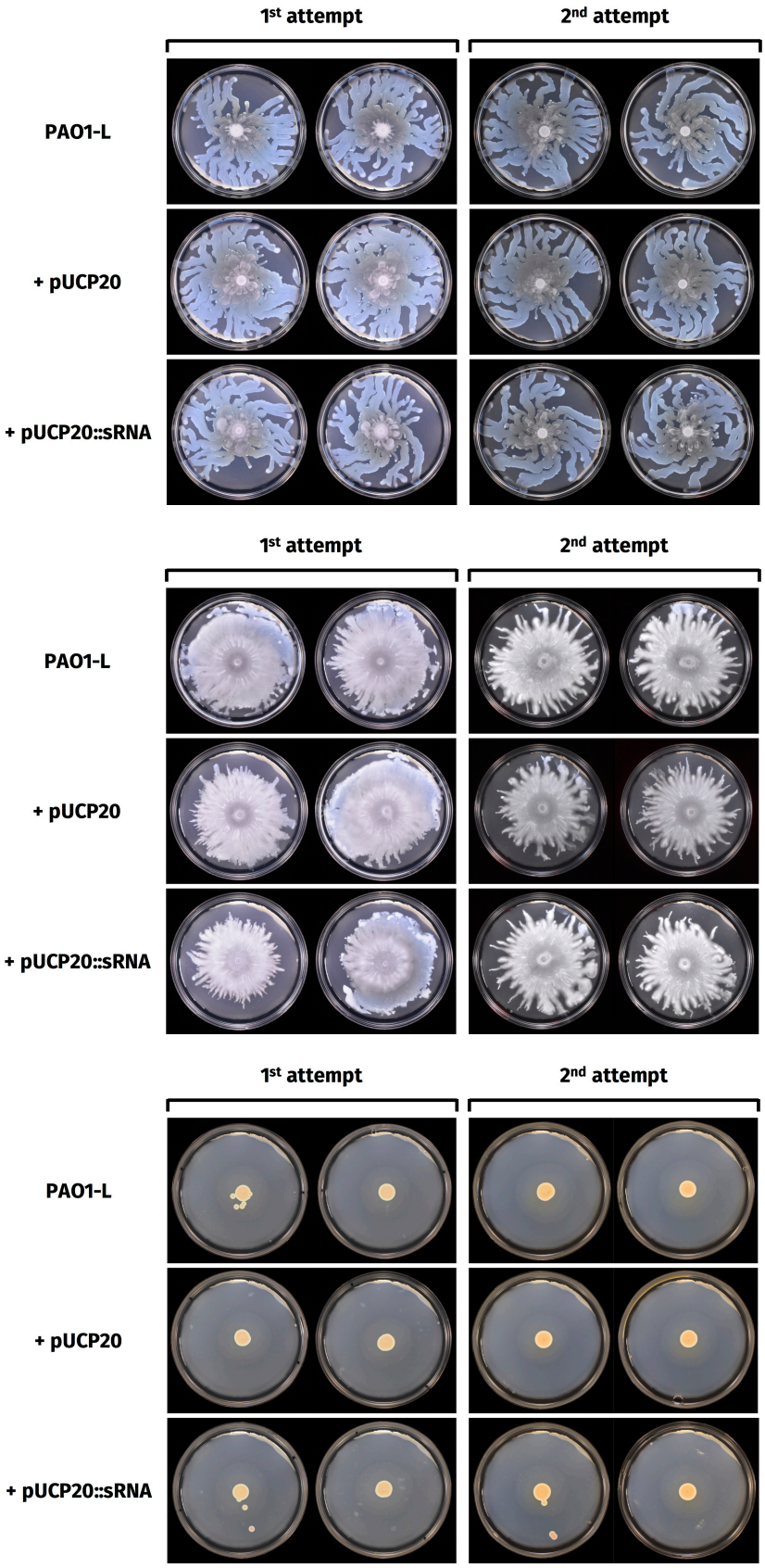

Fig. 19 | Swarming, swimming and twitching motility assays. Motility assays were carried out to determine the swarming (top), swimming (middle) and twitching (bottom) abilities of a *P. aeruginosa* PAO1-L strain carrying the high-copy-number plasmid pUCP20::sRNA, compared to the wild type and a strain carrying pUCP20 with no insert. No significant differences were detected between the three strains.

DISCUSSION

The principal finding of this work is that a newly-identified *P. aeruginosa* sRNA – encoded by an as-of-yet unannotated sRNA gene on the opposite strand to both *psrA* and *nagZ*, complementary to a portion of the intergenic region between the two genes – was shown to bind RsmA and RsmN *in vitro* via EMSAs (Fig. 5). Binding of the sRNA to RsmN was expected based on a previous RNA-seq study (Romero *et al.* 2018), and given that all known full length RsmN-binding targets also bind to RsmA, the fact that the newly-identified sRNA binds to RsmA was also expected. Unfortunately, however, no positive claims about the specificity of the binding interaction can be made from this work, since, in the case of RsmA, no appropriate non-specific inhibitor was used, and in the case of RsmN, addition of the non-specific competitor RNA (corresponding to the *pslA* leader region) appeared to reduce levels of binding. Both of the competition EMSAs ought to be repeated: in the case of RsmA, with an appropriate new unlabelled non-specific competitor RNA (instead of the *pslA* leader region); and in the case of RsmN, perhaps with different concentrations of unlabelled non-specific competitor RNA.

It has previously been suggested that RsmA can bind mRNA targets containing a single CANGGAYG sequence, but that RsmN requires at least two such sequences (Schulmeyer *et al.* 2016). These CANGGAYG sequences are typically predicted to be exposed in single-stranded pentaloop or hexaloop regions of RNA stem-loop structures, where they can be bound by the CsrA/RsmA protein via a 'molecular clamp' mechanism (Schubert *et al.* 2007). Despite the requirement for at least two GGA motifs for RsmN binding, the predicted secondary structure of the new sRNA appears to contain only a single GGA motif exposed in a stem-loop structure, and yet appears to bind RsmN *in vivo* (Romero *et al.* 2018), and in *vitro* (this work). Previously predicted secondary structures of Rsm family sRNAs (Fig. 3) suggest that if the new sRNA is indeed an RsmN-binding sRNA (as suggested by the RNA-seq study and the EMSAs presented here), its true secondary structure might perhaps be different from that predicted by the *mfold* programme.

P. aeruginosa exhibits three distinct types of motility: swarming, swimming and twitching (Henrichsen 1972). Swimming and swarming are predominantly flagellum-dependent, although swarming also requires the production of rhamnolipid surfactants which assist the cells in moving over surfaces (Toutain, Zegans and O'Toole 2005). In contrast, twitching motility involves the extension and retraction of type IV pili (Burrows 2012). Motility is typically associated with the acute infection lifestyle promoted by RsmA and RsmN, and non-motility with the chronic infection lifestyle (Mahenthiralingam, Campbell and Speert 1994). Hence, phenotypic differences in this regard are relevant in the study of CsrA/RsmA proteins and their regulatory sRNAs. For example,

overexpression of *rsmZ* in PAO1 has been shown to reduce swarming ability (presumably through binding and inhibiting RsmA) to levels comparable with that of an *rsmA* mutant (Heurlier *et al.* 2004), and RsmW has been shown to reduce swarming in an *rsmY rsmZ* double mutant (Miller *et al.* 2016). Motility assays conducted with a PAO1-L strain harbouring the high-copy-number *Escherichia-Pseudomonas* shuttle vector pUCP20 carrying the sRNA region and surrounding sequence did not reveal any significant differences compared to the wild type (Fig. 7). Given that the properties of the natural promoter controlling expression of this new sRNA are unknown, it is possible that the sRNA was simply not expressed under the experimental conditions used, or not expressed at very high levels. Alternatively, perhaps the promoter is active, but the sRNA does not have a major role to play in the regulation of motility. Repeating these motility assays with a plasmid in which the sRNA gene is under the control of a strong, constitutive promoter (e.g. pME6000) may prove worthwhile. The rationale of using a pME6000-derived plasmid to enable forced overexpression of the sRNA region in PAO1-L cells would be to potentially augment the natural effects of the sRNA in the hopes of observing an exaggerated phenotype. Such a plasmid was constructed during this project, but further motility assays could not be performed due to time constraints. Also, in the future, using a transcriptional fusion or other methods to investigate the natural promoter of the new sRNA gene, and to determine the conditions under which the sRNA is expressed would be helpful. A similar investigation, monitoring how expression levels vary according to growth phase using RT-qPCR, has already been performed for the other PAO1 Rsm family sRNAs (Janssen *et al.* 2018), and differences were observed, indicating that the functions of these sRNAs are not entirely redundant.

As recently as 2012, only the RsmA protein and the RsmY and RsmZ sRNAs were known. Since then, an additional protein (RsmN) (Marden *et al.* 2013; Morris *et al.* 2013; Romero *et al.* 2018), and two additional Rsm family sRNAs (RsmW and RsmV) have been reported (Miller *et al.* 2016; Janssen *et al.* 2018), and indeed, the work presented here might lead to the eventual reporting of a fifth Rsm family sRNA. RsmW has been shown to bind RsmA *in vitro* and to be able to replace the functions of RsmY and RsmZ in an *rsmY rsmZ* double mutant (Miller *et al.* 2016), and RsmV has been shown to bind both RsmA and RsmN *in vitro* and *in vivo*, and to bind RsmN with >10-fold higher affinity than RsmY and RsmZ (Janssen *et al.* 2018). The new putative sRNA investigated here also appears to bind RsmA and RsmN *in vitro*, and would therefore be a worthy subject for further phenotypic analyses. Perhaps investigating the ability of the new sRNA to replace the functions of RsmY and RsmZ in an *rsmY rsmZ* mutant would be worthwhile, with a focus on swarming, biofilm formation and other RsmA-related phenotypes; given that this approach proved useful in the case of RsmW.

Ultimately, all of these investigations should be carried out with a view to developing a better understanding of *P. aeruginosa* pathogenesis. Hence, if/when a biological role for the new sRNA is identified, perhaps in biofilm formation or motility, steps should then of course be taken to understand the impact of the finding(s) in a biological context using relevant infection models, such as *ex vivo* lung models, the wax moth (*Galleria mellonella*) or mice.

Overall, it seems our understanding of the complexity of the interactions between the CsrA/RsmA family proteins and their regulatory sRNAs has increased substantially in recent years. Known differences in temporal expression (Kay *et al.* 2006; Janssen *et al.* 2018), as well as differences in binding affinities suggest distinct roles for each of these sRNAs in regulating RsmA and RsmN activity in response to changing growth conditions in this highly adaptive organism. In addition, given that CsrA/RsmA family proteins are known to affect resistance to multiple antibiotics (Brencic and Lory 2009; Ricci *et al.* 2017), these proteins, their associated sRNAs and regulatory systems constitute promising targets for the future development of novel antibacterial strategies.

Future work

Suggestions as to how each of the two research projects might be developed into the future have already been given in the *Discussion* sections to the above reports. These suggestions and others are summarised here.

Research Project #1: MarA reduces biofilm formation by activating expression of the *ycgZ-ymgA-ariR-ymgC* operon in *Escherichia coli*

- the β-galactosidase assays reported here could be repeated at 30°C to confirm MarA-mediated regulation of the YYAY operon at this temperature (here assays were performed at 37°C)
- β-galactosidase assays and crystal violet biofilm formation assays could be performed with *soxS* and *rob* mutants to verify that it is indeed MarA, and not SoxS or Rob, which plays the central role in regulating P_{ycgZ} activity
- β-galactosidase assays and crystal violet biofilm formation assays could be performed with strains overexpressing MarA, to determine if this augments the effects presented in this work
- the antibiotic resistance profiles of strains reported here could be determined as a means of analysing the effect that changes mediated by MarA and the YYAY operon (on biofilm formation and perhaps porin production) have on antibiotic resistance
- the structures and functions of the proteins encoded by the YYAY operon could be examined in greater detail (with the exception of AriR, they are all poorly characterised)
- a new inactive mutated marbox should be generated for use as a negative control in β-galactosidase assays and crystal violet biofilm formation assays, given that the one used here appeared to be exhibiting residual activity in some cases
- the role of H-NS in controlling expression of the YYAY operon could be examined, perhaps via the use of a P_{ycgZ} transcriptional fusion or other more targeted methods
- the Congo red curli production assays reported here could be repeated at a more appropriate temperature (≤28°C) to determine whether or not MarA-mediated regulation of the YYAY operon has any effect on curli production
- known pathogenic *E. coli* strains and clinical isolates could perhaps be employed to determine whether the findings presented here can be replicated in clinically-relevant strains
- relevant infection models could perhaps be used to examine the effects presented in this work in a biological context
- as a general note, greater attention should be paid to the temperatures at which techniques used in previous studies have been performed, given that expression of the YYAY operon appears to be temperature-dependent

Research Project #2: A newly-identified sRNA from *Pseudomonas aeruginosa* binds to the post-transcriptional regulators RsmA and RsmN

- the specificity of the binding interaction between the new sRNA and the RsmA/RsmN proteins should be confirmed by repeating the competition EMSAs with appropriate unlabelled specific and non-specific competitor RNAs
- the motility assays reported here should be repeated with strains carrying a plasmid in which the sRNA gene is under the control of a strong, constitutive promoter (e.g. pME6000), to potentially augment the natural effects of the sRNA in the hopes of observing an exaggerated phenotype
- transcriptional fusions or other methods could be used to investigate the activity of the natural promoter controlling expression of the new sRNA gene, and to determine the conditions under which the sRNA is normally expressed
- primer extension or rapid amplification of cDNA ends (RACE) could be used to determine the transcription start site and the 5' and 3' boundaries of the newly-identified sRNA gene, given that these cannot be determined from current RNA-seq data
- perhaps investigating the ability of the new sRNA to replace the functions of RsmY and RsmZ in an *rsmY rsmZ* mutant would be worthwhile, with a focus on swarming, biofilm formation and other RsmA-related phenotypes; this approach proved useful in the case of RsmW
- if/when a biological role for the new sRNA is identified, perhaps in biofilm formation or motility, steps should be taken to understand the impact of the finding(s) in a biological context using relevant infection models, such as *ex vivo* lung models, the wax moth (*Galleria mellonella*) or mice
- given that CsrA/RsmA family proteins are known to affect resistance to multiple antibiotics the antibiotic resistance profile of a strain overexpressing the new sRNA could perhaps be examined and compared to that of the wild type strain
- known pathogenic strains and clinical isolates of *P. aeruginosa* could perhaps be employed to determine whether findings observed in the laboratory strain used here can be replicated in clinically-relevant strains
- given how many new Rsm family sRNAs have been described in recent years, perhaps there are still more of these sRNAs to be discovered

References

Allen HK, Trachsel J, Looft T *et al.* Finding alternatives to antibiotics. *Ann N Y Acad Sci* 2014;**1323**:91–100.

Aminov RI. A brief history of the antibiotic era: lessons learned and challenges for the future. *Front Microbiol* 2010;**1**:134.

Anderson GG. Intracellular bacterial biofilm-like pods in urinary tract infections. *Science* 2003;**301**:105–7.

Årdal C, Findlay D, Savic M *et al.* Revitalizing the antibiotic pipeline: stimulating innovation while driving sustainable use and global access. DRIVE-AB Report. 2018.

Ariza RR, Cohen SP, Bachhawat N *et al.* Repressor mutations in the *marRAB* operon that activate oxidative stress genes and multiple antibiotic resistance in *Escherichia coli*. *J Bacteriol* 1994;**176**:143–8.

Asgarali A, Stubbs KA, Oliver A *et al.* Inactivation of the glycoside hydrolase NagZ attenuates antipseudomonal β-lactam resistance in *Pseudomonas aeruginosa*. *Antimicrob Agents Chemother* 2009;**53**:2274–82.

Barbosa TM, Levy SB. Differential expression of over 60 chromosomal genes in *Escherichia coli* by constitutive expression of MarA. *J Bacteriol* 2000;**182**:3467–74.

Baugh S, Phillips CR, Ekanayaka AS *et al.* Inhibition of multidrug efflux as a strategy to prevent biofilm formation. *J Antimicrob Chemother* 2014;**69**:673–81.

Belik AS, Tarasova NN, Khmel' IA. Regulation of biofilm formation in *Escherichia coli* K12: effect of mutations in the genes HNS, STRA, LON, and RPON. *Mol Genet Microbiol Virol* 2008;**23**:159–62.

Beloin C, Roux A, Ghigo J-M. *Escherichia coli* biofilms. *Curr Top Microbiol Immunol* 2008;**322**:249–89.

Blair JMA, Webber MA, Baylay AJ *et al.* Molecular mechanisms of antibiotic resistance. *Nat Rev Microbiol* 2015;**13**:42–51.

Blount ZD. The unexhausted potential of *E. coli*. *Elife* 2015;**4**:e05826.

Bolivar F, Rodriguez RL, Greene PJ *et al.* Construction and characterization of new cloning vehicle: II. A multipurpose cloning system. *Gene* 1977;**2**:95–113.

Brencic A, Lory S. Determination of the regulon and identification of novel mRNA targets of *Pseudomonas aeruginosa* RsmA. *Mol Microbiol* 2009;**72**:612–32.

Brencic A, McFarland KA, McManus HR *et al.* The GacS/GacA signal transduction system of *Pseudomonas aeruginosa* acts exclusively through its control over the transcription of the RsmY and RsmZ regulatory small RNAs. *Mol Microbiol* 2009;**73**:434–45.

Bridier A, Briandet R, Thomas V *et al.* Resistance of bacterial biofilms to disinfectants: a review. *Biofouling* 2011;**27**:1017–32.

Brooks JD, Flint SH. Biofilms in the food industry: problems and potential solutions. *Int J Food Sci Technol* 2008;**43**:2163–76.

Burrows LL. *Pseudomonas aeruginosa* twitching motility: type IV pili in action. *Annu Rev Microbiol* 2012;**66**:493–520.

Carver T, Harris SR, Berriman M *et al.* Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* 2012;**28**:464–9.

Chandy SJ. Consequences of irrational use of antibiotics. *Indian J Med Ethics* 2008;**5**:174–5.

Chuanchuen R, Narasaki CT, Schweizer HP. Benchtop and microcentrifuge preparation of *Pseudomonas aeruginosa* competent cells. *BioTechniques* 2002;**33**:760–63.

Chubiz LM, Rao CV. Role of the *mar-sox-rob* regulon in regulating outer membrane porin expression. *J Bacteriol* 2011;**193**:2252–60.

Cohen SP, McMurry LM, Hooper DC *et al.* Cross-resistance to fluoroquinolones in multiple-antibiotic-resistant (Mar) *Escherichia coli* selected by tetracycline or chloramphenicol: decreased drug accumulation associated with membrane changes in addition to OmpF reduction. *Antimicrob Agents Chemother* 1989;**33**:1318–25.

Cohen SP, McMurry LM, Levy SB. *marA* locus causes decreased expression of OmpF porin in multiple-antibiotic-resistant (Mar) mutants of *Escherichia coli*. *J Bacteriol* 1988;**170**:5416–22.

Danese PN, Pratt LA, Kolter R. Exopolysaccharide production is required for development of *Escherichia coli* K-12 biofilm architecture. *J Bacteriol* 2000;**182**:3593–6.

Davies J, Davies D. Origins and evolution of antibiotic resistance. *Microbiol Mol Biol Rev* 2010;**74**:417–33.

Domka J, Lee J, Bansal T *et al.* Temporal gene-expression in *Escherichia coli* K-12 biofilms. *Environ Microbiol* 2007;**9**:332–46.

Dongari-Bagtzoglou A. Pathogenesis of mucosal biofilm infections: challenges and progress. *Expert Rev Anti Infect Ther* 2008;**6**:201–8.

Dubey AK, Baker CS, Romeo T *et al.* RNA sequence and secondary structure participate in high-affinity CsrA-RNA interaction. *RNA* 2005;**11**:1579–87.

Dubey AK, Baker CS, Suzuki K *et al.* CsrA regulates translation of the *Escherichia coli* carbon starvation gene, *cstA*, by blocking ribosome access to the *cstA* transcript. *J Bacteriol* 2003;**185**:4450–60.

Duval V, Foster K, Brewster J *et al.* A novel regulatory cascade involving BluR, YcgZ, and Lon controls the expression of *Escherichia coli* OmpF porin. *Front Microbiol* 2017;**8**:1148.

Fair RJ, Tor Y. Antibiotics and bacterial resistance in the 21st century. *Perspect Med Chem* 2014;**6**:25–64.

Fernandez L, Hancock REW. Adaptive and mutational resistance: role of porins and efflux pumps in drug resistance. *Clin Microbiol Rev* 2012;**25**:661–81.

Flemming H-C, Wingender J. The biofilm matrix. *Nat Rev Microbiol* 2010;**8**:623–33.

Flemming H-C, Wingender J, Szewzyk U *et al.* Biofilms: an emergent form of bacterial life. *Nat Rev Microbiol* 2016;**14**:563–75.

Gallant CV, Daniels C, Leung JM *et al.* Common β-lactamases inhibit bacterial biofilm formation. *Mol Microbiol* 2005;**58**:1012–24.

George AM, Levy SB. Amplifiable resistance to tetracycline, chloramphenicol, and other antibiotics in *Escherichia coli*: involvement of a non-plasmid-determined efflux of tetracycline. *J Bacteriol* 1983;**155**:531–540.

Gutierrez P, Li Y, Osborne MJ *et al.* Solution structure of the carbon storage regulator protein CsrA from *Escherichia coli*. *J Bacteriol* 2005;**187**:3496–501.

Haycocks JRJ, Sharma P, Stringer AM *et al.* The molecular basis for control of ETEC enterotoxin expression in response to environment and host. *PLoS Pathog* 2015;**11**:e1004605.

Heeb S, Kuehne SA, Bycroft M *et al.* Functional analysis of the post-transcriptional regulator RsmA reveals a novel RNA-binding site. *J Mol Biol* 2006;**355**:1026–36.

Henrichsen J. Bacterial surface translocation: a survey and a classification. *Microbiol Mol Biol Rev* 1972;**36**:478–503.

Heurlier K, Williams F, Heeb S *et al.* Positive control of swarming, rhamnolipid synthesis, and lipase production by the posttranscriptional RsmA/RsmZ system in *Pseudomonas aeruginosa* PAO1. *J Bacteriol* 2004;**186**:2936–45.

Hong SH, Wang X, Wood TK. Controlling biofilm formation, prophage excision and cell death by rewiring global regulator H-NS of *Escherichia coli*. *Microb Biotechnol* 2010;**3**:344–56.

Huertas-Rosales Ó, Ramos-González MI, Espinosa-Urgel M. Self-regulation and interplay of Rsm family proteins modulate the lifestyle of *Pseudomonas putida*. *Appl Environ Microbiol* 2016;**82**:5673–86.

Huertas-Rosales Ó, Romero M, Heeb S *et al.* The *Pseudomonas putida* CsrA/RsmA homologues negatively affect c-di-GMP pools and biofilm formation through the GGDEF/EAL response regulator CfcR. *Environ Microbiol* 2017;**19**:3551–66.

Irie Y, Starkey M, Edwards AN *et al. Pseudomonas aeruginosa* biofilm matrix polysaccharide Psl is regulated transcriptionally by RpoS and post-transcriptionally by RsmA. *Mol Microbiol* 2010:158–72.

Janssen KH, Diaz MR, Gode CJ *et al.* RsmV a small non-coding regulatory RNA in *Pseudomonas aeruginosa* that sequesters RsmA and RsmF from target mRNAs. *J Bacteriol* 2018:JB.00277-18. [Epub ahead of print]

Jimenez PN, Koch G, Thompson JA *et al.* The multiple signaling systems regulating virulence in *Pseudomonas aeruginosa*. *Microbiol Mol Biol Rev* 2012;**76**:46–65.

Joo H-S, Otto M. Molecular basis of *in vivo* biofilm formation by bacterial pathogens. *Chem Biol* 2012;**19**:1503–13.

Kahramanoglou C, Seshasayee ASN, Prieto AI *et al.* Direct and indirect effects of H-NS and Fis on global gene expression control in *Escherichia coli*. *Nucleic Acids Res* 2011;**39**:2073–91.

Kaper JB, Nataro JP, Mobley HLT. Pathogenic *Escherichia coli*. *Nat Rev Microbiol* 2004;**2**:123–40.

Kaplan JB. Biofilm dispersal: mechanisms, clinical implications, and potential therapeutic uses. *J Dent Res* 2010;**89**:205–18.

Kay E, Humair B, Denervaud V *et al.* Two GacA-dependent small RNAs modulate the quorum-sensing response in *Pseudomonas aeruginosa*. *J Bacteriol* 2006;**188**:6026–33.

Kikuchi T, Mizunoe Y, Takade A *et al.* Curli fibers are required for development of biofilm architecture in *Escherichia coli* K-12 and enhance bacterial adherence to human uroepithelial cells. *Microbiol Immunol* 2005;**49**:875–84.

Kingsford CL, Ayanbule K, Salzberg SL. Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. *Genome Biol* 2007;**8**:R22.

Kojic M, Venturi V. Regulation of *rpoS* gene expression in *Pseudomonas*: involvement of a TetR family regulator. *J Bacteriol* 2001;**183**:3712–20.

Lee J, Page R, García-Contreras R *et al.* Structure and function of the *Escherichia coli* protein YmgB: a protein critical for biofilm formation and acid-resistance. *J Mol Biol* 2007;**373**:11–26.

Lewis K. Platforms for antibiotic discovery. *Nat Rev Drug Discov* 2013;**12**:371–87.

Lodge J, Fear J, Busby S *et al.* Broad host range plasmids carrying the *Escherichia coli* lactose and galactose operons. *FEMS Microbiol Lett* 1992;**95**:271–6.

Mahenthiralingam E, Campbell ME, Speert DP. Nonmotility and phagocytic resistance of *Pseudomonas aeruginosa* isolates from chronically colonized patients with cystic fibrosis. *Infect Immun* 1994;**62**:10.

Maira-Litrán T, Allison DG, Gilbert P. Expression of the multiple antibiotic resistance operon (*mar*) during growth of *Escherichia coli* as a biofilm. *J Appl Microbiol* 2000;**88**:243–7.

Marden JN, Diaz MR, Walton WG *et al.* An unusual CsrA family member operates in series with RsmA to amplify posttranscriptional responses in *Pseudomonas aeruginosa*. *Proc Natl Acad Sci* 2013;**110**:15055–60.

Martin RG, Gillette WK, Martin NI *et al.* Complex formation between activator and RNA polymerase as the basis for transcriptional activation by MarA and SoxS in *Escherichia coli*. *Mol Microbiol* 2002;**43**:355–70.

Martin RG, Gillette WK, Rhee S *et al.* Structural requirements for marbox function in transcriptional activation of *mar*/*sox*/*rob* regulon promoters in *Escherichia coli*: sequence, orientation and spatial relationship to the core promoter. *Mol Microbiol* 1999;**34**:431–41.

Martin RG, Rosner JL. Binding of purified multiple antibiotic-resistance repressor protein (MarR) to *mar* operator sequences. *Proc Natl Acad Sci* 1995;**92**:5456–60.

Martin RG, Rosner JL. Genomics of the *marA*/*soxS*/*rob* regulon of *Escherichia coli*: identification of directly activated promoters by application of molecular genetics and informatics to microarray data. *Mol Microbiol* 2002;**44**:1611–24.

Maurhofer M, Reimmann C, Schmidli-Sacherer P *et al.* Salicylic acid biosynthetic genes expressed in *Pseudomonas fluorescens* strain P3 improve the induction of systemic resistance in tobacco against tobacco necrosis virus. *Phytopathology* 1998;**88**:678–84.

McMurry LM, George AM, Levy SB. Active efflux of chloramphenicol in susceptible *Escherichia coli* strains and in multiple-antibiotic-resistant (Mar) mutants. *Antimicrob Agents Chemother* 1994;**38**:542–6.

Mendelson M, Balasegaram M, Jinks T *et al.* Antibiotic resistance has a language problem. *Nature* 2017;**545**:23–5.

Miller CL, Romero M, Karna SLR *et al.* RsmW, *Pseudomonas aeruginosa* small non-coding RsmA-binding RNA upregulated in biofilm versus planktonic growth conditions. *BMC Microbiol* 2016;**16**:155.

Miller JH. *Experiments in Molecular Genetics.* Cold Spring Harbor Laboratory, 1972.

Morris ER, Hall G, Li C *et al.* Structural rearrangement in an RsmA/CsrA ortholog of *Pseudomonas aeruginosa* creates a dimeric RNA-binding protein, RsmN. *Structure* 2013;**21**:1659–71.

Mulcahy H, O'Callaghan J, O'Grady EP *et al.* The posttranscriptional regulator RsmA plays a role in the interaction between *Pseudomonas aeruginosa* and human airway epithelial cells by positively regulating the type III secretion system. *Infect Immun* 2006;**74**:3012–5.

Munch R. PRODORIC: prokaryotic database of gene regulation. *Nucleic Acids Res* 2003;**31**:266–9.

Munch R, Hiller K, Grote A *et al.* Virtual Footprint and PRODORIC: an integrative framework for regulon prediction in prokaryotes. *Bioinformatics* 2005;**21**:4187–9.

Nakata Y, Tang X, Yokoyama KK. Preparation of competent cells for high-efficiency plasmid transformation of *Escherichia coli*. *CDNA Libr Protoc* 1997:129–37.

Okusu H, Ma D, Nikaido H. AcrAB efflux pump plays a major role in the antibiotic resistance phenotype of *Escherichia coli* multiple-antibiotic-resistance (Mar) mutants. *J Bacteriol* 1996;**178**:306–8.

Overbye K, Barrett J. Antibiotics: where did we go wrong? *Drug Discov Today* 2005;**10**:45–52.

Page L, Griffiths L, Cole JA. Different physiological roles of two independent pathways for nitrite reduction to ammonia by enteric bacteria. *Arch Microbiol* 1990;**154**:349–54.

Payne DJ, Gwynn MN, Holmes DJ *et al.* Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat Rev Drug Discov* 2007;**6**:29–40.

Pessi G, Williams F, Hindle Z *et al.* The global posttranscriptional regulator RsmA modulates production of virulence determinants and N-acylhomoserine lactones in *Pseudomonas aeruginosa*. *J Bacteriol* 2001;**183**:6676–83.

Prajapat MK, Jain K, Saini S. Control of *marRAB* operon in *Escherichia coli* via autoactivation and autorepression. *Biophys J* 2015;**109**:1497–508.

Pratt LA, Kolter R. Genetic analysis of *Escherichia coli* biofilm formation: roles of flagella, motility, chemotaxis and type I pili. *Mol Microbiol* 1998;**30**:285–93.

Randall LP, Woodward MJ. The multiple antibiotic resistance (*mar*) locus and its significance. *Res Vet Sci* 2002;**72**:87–93.

Reimmann C, Valverde C, Kay E *et al.* Posttranscriptional repression of GacS/GacA-controlled genes by the RNA-binding protein RsmE acting together with RsmA in the biocontrol strain *Pseudomonas fluorescens* CHA0. *J Bacteriol* 2005;**187**:276–85.

Ren B, Shen H, Lu ZJ *et al.* The *phzA2-G2* transcript exhibits direct RsmA-mediated activation in *Pseudomonas aeruginosa* M18. *PLoS ONE* 2014;**9**:e89653.

Ricci V, Attah V, Overton T *et al.* CsrA maximizes expression of the AcrAB multidrug resistance transporter. *Nucleic Acids Res* 2017;**45**:12798–807.

Rife C, Schwarzenbacher R, McMullan D *et al.* Crystal structure of the global regulatory protein CsrA from *Pseudomonas putida* at 2.05Å resolution reveals a new fold. *Proteins Struct Funct Bioinforma* 2005;**61**:449–53.

Roilides E, Simitsopoulou M, Katragkou A *et al.* How biofilms evade host defenses. *Microbiol Spectr* 2015;**3**:MB-0012-2014.

Romeo T, Gong M, Liu MY *et al.* Identification and molecular characterization of *csrA*, a pleiotropic gene from *Escherichia coli* that affects glycogen biosynthesis, gluconeogenesis, cell size, and surface properties. *J Bacteriol* 1993;**175**:4744–55.

Romero M, Silistre H, Lovelock L *et al.* Genome-wide mapping of the RNA targets of the *Pseudomonas aeruginosa* riboregulatory protein RsmN. *Nucleic Acids Res* 2018. [Epub ahead of print]

Ruiz C, Levy SB. Many chromosomal genes modulate MarA-mediated multidrug resistance in *Escherichia coli*. *Antimicrob Agents Chemother* 2010;**54**:2125–34.

Schubert M, Lapouge K, Duss O *et al.* Molecular basis of messenger RNA recognition by the specific bacterial repressing clamp RsmA/CsrA. *Nat Struct Mol Biol* 2007;**14**:807–13.

Schulmeyer KH, Diaz MR, Bair TB *et al.* Primary and secondary sequence structure requirements for recognition and discrimination of target RNAs by *Pseudomonas aeruginosa* RsmA and RsmF. *J Bacteriol* 2016;**198**:2458–69.

Seoane AS, Levy SB. Characterization of MarR, the repressor of the multiple antibiotic resistance (*mar*) operon in *Escherichia coli*. *J Bacteriol* 1995;**177**:3414–19.

Sharma P, Haycocks JRJ, Middlemiss AD *et al.* The multiple antibiotic resistance operon of enteric bacteria controls DNA repair and outer membrane integrity. *Nat Commun* 2017;**8**:1444.

Simpkin VL, Renwick MJ, Kelly R *et al.* Incentivising innovation in antibiotic drug discovery and development: progress, challenges and next steps. *J Antibiot (Tokyo)* 2017;**70**:1087–96.

Soto SM, Smithson A, Horcajada JP *et al.* Implication of biofilm formation in the persistence of urinary tract infection caused by uropathogenic *Escherichia coli*. *Clin Microbiol Infect* 2006;**12**:1034–6.

Stewart PS. Antimicrobial tolerance in biofilms. *Microbiol Spectr* 2015;**3**:MB-0010-2014.

Stewart PS, Costerton JW. Antibiotic resistance of bacteria in biofilms. *The Lancet* 2001;**358**:135–8.

Teillant A, Gandra S, Barter D *et al.* Potential burden of antibiotic resistance on surgery and cancer chemotherapy antibiotic prophylaxis in the USA: a literature review and modelling study. *Lancet Infect Dis* 2015;**15**:1429–37.

Thomason LC, Costantino N, Court DL. *E. coli* genome manipulation by P1 transduction. In: Ausubel FM, Brent R, Kingston RE, et al. (eds.). *Current Protocols in Molecular Biology*. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2007, Chapter 1: Unit 1.17.

Toutain CM, Zegans ME, O'Toole GA. Evidence for two flagellar stators and their role in the motility of *Pseudomonas aeruginosa*. *J Bacteriol* 2005;**187**:771–7.

Tschowri N, Busse S, Hengge R. The BLUF-EAL protein YcgF acts as a direct anti-repressor in a blue-light response of *Escherichia coli*. *Genes Dev* 2009;**23**:522–34.

Vakulskas CA, Potts AH, Babitzke P *et al.* Regulation of bacterial virulence by Csr (Rsm) systems. *Microbiol Mol Biol Rev* 2015;**79**:193–224.

Valverde C, Heeb S, Keel C *et al.* RsmY, a small regulatory RNA, is required in concert with RsmZ for GacA-dependent expression of biocontrol traits in *Pseudomonas fluorescens* CHA0. *Mol Microbiol* 2003;**50**:1361–79.

Vila J, Soto SM. Salicylate increases the expression of *marA* and reduces *in vitro* biofilm formation in uropathogenic *Escherichia coli* by decreasing type 1 fimbriae expression. *Virulence* 2012;**3**:280–5.

Vinué L, McMurry LM, Levy SB. The 216-bp *marB* gene of the *marRAB* operon in *Escherichia coli* encodes a periplasmic protein which reduces the transcription rate of *marA*. *FEMS Microbiol Lett* 2013;**345**:49–55.

Weinstein R. Controlling antimicrobial resistance in hospitals: infection control and use of antibiotics. *Emerg Infect Dis* 2001;**7**:188–92.

West SEH, Schweizer HP, Dall C *et al.* Construction of improved *Escherichia-Pseudomonas* shuttle vectors derived from pUC18/19 and sequence of the region required for their replication in *Pseudomonas aeruginosa*. *Gene* 1994;**148**:81–86.

White-Ziegler CA, Um S, Perez NM *et al.* Low temperature (23°C) increases expression of biofilm-, cold-shock- and RpoS-dependent genes in *Escherichia coli* K-12. *Microbiology* 2008;**154**:148–66.

Yakhnin H, Pandit P, Petty TJ *et al.* CsrA of *Bacillus subtilis* regulates translation initiation of the gene encoding the flagellin protein (*hag*) by blocking ribosome binding. *Mol Microbiol* 2007;**64**:1605–20.

Yamada H, Yoshida T, Tanaka K *et al.* Molecular analysis of the *Escherichia coli hns* gene encoding a DNA-binding protein, which preferentially recognizes curved DNA sequences. *Mol Gen Genet* 1991;**230**:332–6.

Ying B-W, Fourmy D, Yoshizawa S. Substitution of the use of radioactivity by fluorescence for biochemical studies of RNA. *RNA* 2007;**13**:2042–50.

Zamorano L, Reeve TM, Deng L *et al.* NagZ inactivation prevents and reverts β-lactam resistance, driven by AmpD and PBP 4 mutations, in *Pseudomonas aeruginosa*. *Antimicrob Agents Chemother* 2010;**54**:3557–63.

Zhang A, Rosner JL, Martin RG. Transcriptional activation by MarA, SoxS and Rob of two *tolC* promoters using one binding site: a complex promoter configuration for *tolC* in *Escherichia coli*. *Mol Microbiol* 2008;**69**:1450–5.

Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 2003;**31**:3406–15.